

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."*

THE RADIO AND ELECTRONIC ENGINEER

The Journal of the Institution of Electronic and Radio Engineers

VOLUME 38 No. 4

OCTOBER 1969

Exporting Expertise

IN electronic engineering projects today, the determination of viability, or to put it bluntly, the answer to the question 'will it pay?' can call for highly-developed techniques. To mention only two of these techniques, computer methods will often be employed for statistical analysis of design data or for simulation of operational conditions. Such aids and many others are becoming more and more important in the way they can complement the wealth of practical experience that has to be brought to bear at all stages, first by the development engineer, then the design engineer, and finally the production engineer. Right from the initial stages the electronic engineer works in close collaboration with engineers and scientists of other disciplines, particularly in major projects in such fields as space technology and ocean technology. Very considerable expertise is called for in this kind of work as well as in more purely electronic schemes such as communications networks and air traffic control.

It is thus apparent that in a project of any size the advantages are all on the side of the larger organization which has the financial resources to employ sufficient men and materials: the advantages are also on the side of the country which has a pool of skilled manpower on which to draw. Inevitably, the resources of skilled manpower in sufficient numbers to carry out complicated projects will be easier to find in the highly industrialized countries than in those countries in Asia, Africa and South America which historically have been primarily agricultural. The desire to move into the second half of the twentieth century on the part of these developing countries is well understood but if they are to achieve industrial expansion they will need to seek help, for some years to come, from the countries in which technical experience and 'know how' already exist.

What is the nature of the assistance required by the developing countries? Firstly, there is a need for communications in the broadest sense—transportation, on land and in the air, internal and external telecommunication, broadcasting, both sound and television. Then come industrial installations for processing raw materials of mineral or agricultural origin, and for fabricating consumer goods. The general modernization of commerce and education is also an area where new techniques are necessary. In all these fields the role to be played by electronics has become very apparent through recent developments in the industrial countries of the West, where the growing emphasis on 'systems' has led to the increasing prominence of the 'systems engineer'. The broad background knowledge, theoretical as well as practical, for this branch of engineering must rely on the opportunity for training and experience only available in an industrialized country.

Since the Industrial Revolution in the eighteenth and nineteenth centuries British engineering industry has traditionally been outward looking in the way it has helped developing countries, initially through the construction of complete transportation networks, irrigation systems and the like. Today much of the capability for such civil engineering work is becoming available within those countries: it is the specialized techniques of the 'science-based industries'—aircraft, chemicals, electronics, nuclear power, for instance—the techniques calling for large numbers of highly proficient and experienced engineers, that are called for in greatest measures. Engineering consultancy, the provision of services such as data processing, information processing and computing, and the operation of installed systems as well as the design and production of the systems, are techniques which are needed.

It is surely a challenge to the British electronics industry in particular to look outward in selling its skill and experience in the engineering of systems of all kinds, by the 'turn-key contract' or 'package deal', as well as in services. Exports of this kind, both in visible 'hardware' and in invisible expertise, are vital contributions to the benefit of mankind.

F. W. S.

INSTITUTION NOTICES

Closer Liaison between I.E.E. and I.E.R.E.

The following is the text of a statement issued jointly by the I.E.R.E. and I.E.E. on 10th October 1969. (A similar announcement is appearing in I.E.E. publications):

'A committee has been formed by the Institution of Electrical Engineers and the Institution of Electronic and Radio Engineers to discuss the activities of the two Institutions and generally to effect a closer liaison between them.

'The terms of reference of the Committee, which is known as the I.E.E.-I.E.R.E. Joint Liaison Committee, are:

- (a) To examine, and to report on, the advantages, disadvantages and problems of possible methods of combining the activities of the two Institutions in a manner which would be in the best interests of the members of both Institutions.
- (b) To make recommendations for the closer working together of the two Institutions.
- (c) To report regularly to the Councils of both Institutions.

'The I.E.E. and the I.E.R.E. have worked closely together in the past as members of the Standing Committee of Kindred Societies and in other ways. Many joint meetings have been held, and the recent issue of a combined meetings programme is one example of the way in which the two Institutions are collaborating.

'The work of the two Institutions has been the subject of informal discussions in the past months, and the setting-up of this Joint Liaison Committee is an outcome of the discussions.'

The formal statement above does not of course do more than hint at the work of the Joint Liaison Committee. Aspects of co-operation at present under discussion are: the reconciliation of any anomalies which exist between the requirements for corresponding grades of membership in the two Institutions; formulation of a common viewpoint on the qualification, registration and learned-society activities of technician engineers; and consideration of policies on overseas activities, e.g. local offices and recruitment.

Reprints of *Journal Papers*

Reprints are prepared of all papers published in the *Journal* and copies may be obtained from the Institution, price 5s. 0d. each (post free). Requests for reprints may be made using the form which is included in the end pages of most issues of the *Journal*. It is particularly asked that remittances be sent with orders to avoid book-keeping entries and thus reduce handling costs.

Annual General Meeting

Members resident outside Great Britain who do not receive the *Proceedings of the I.E.R.E.* may obtain a copy of the September-October 1969 issue containing the Annual Report of the Council for 1968-9 on application to the Publications Department. The Report was presented at the Annual General Meeting held in London on 22nd October. It was followed by the Inaugural Address of Mr. Harvey F. Schwarz, the 21st President of the Institution. Mr. Schwarz's Address 'The Engineer in State and Private Enterprise' will be printed in the January 1970 issue of *The Radio and Electronic Engineer*.

C.E.I. Graham Clark Lecture 1969

The 15th Graham Clark Lecture will be given on 18th November 1969, by Sir Frederick Brundrett, K.C.B., K.B.E. at the Institution of Electrical Engineers, Savoy Place, London, W.C.2; it will start at 5.45 p.m.

The title of Sir Frederick's Lecture will be 'The Sea of Opportunity—What the Oceans have to offer the Engineer'. This is a broad appraisal of the importance of the oceans and the need to exploit them for the benefit of the world as a whole and of this country in particular. He will examine this country's record of achievement and will demonstrate the necessity of incorporating the engineer and his outlook in the planning and development teams to a much greater extent than hitherto.

The Graham Clark Lecture is an annual event organized by C.E.I., the theme of the lectures being 'The place of Engineering in relation to Society as a whole'. Their principal object is to make better known the essential functions of the professional engineer in modern society, in particular his role in the evolution of National Policy. The Lectures are named after a one-time secretary of the Institution of Civil Engineers who was particularly concerned with this aspect of an engineer's function.

Correspondence with the Institution

Members are reminded that it considerably helps the Institution's secretariat to deal speedily with correspondence if membership numbers (and grades) are stated on *all* letters etc. sent to Bedford Square.

The addition of postcodes to computer-stored addresses is being effected when entries such as subscription payments, etc., are made to a member's record. Postcode information should therefore always be given when communicating with the Institution so that it can be incorporated at a convenient time.

Radar Polarization Comparisons in Sea-clutter Suppression by Decorrelation and Constant False Alarm Rate Receivers

By

J. CRONEY,
B.Sc., Ph.D.†

AND

A. WORONCOW,
Dipl.Ing., C.Eng., M.I.E.R.E.

Summary: An X-band radar with back-to-back horizontally and vertically polarized antennas of equal gain and beamwidths was built. The antennas were enclosed in a symmetrical aerodynamic radome permitting speeds of rotation from 20 rev/min to 600 rev/min. The radar power was switched every half revolution from one antenna to the other, so that a sea-clutter picture presented on the top 180° of the p.p.i. display was identically repeated on the lower 180°, but with the orthogonal polarization. Some of the different characteristics of sea-clutter on the two polarizations were strikingly illustrated by this virtually simultaneous presentation. The sea-clutter smoothing effects obtained by decorrelation and integration are demonstrated. The r.m.s. sea-clutter amplitude was brought down to that of the noise by a true i.f. (amplifying logarithmically at the input frequency) type of constant false alarm rate (c.f.a.r.) receiver developed during this work. Recent work on logarithmic receivers has shown that the output noise fluctuations do not always become constant and independent of input r.m.s. noise amplitude, even though a near perfect logarithmic input-output law may be obtained on pulse or continuous wave (c.w.) inputs. The performance of the true i.f. logarithmic amplifier has been investigated using a wide-band noise source (the output of which may be varied through 50 dB) to feed the amplifier. The conditions which must be observed if the output noise fluctuations are to remain constant while the input noise varies over this range have been established. When these conditions are met there is no significant impairment of detectability of weak signals against noise when using a logarithmic receiver, compared with a linear receiver.

1. Introduction

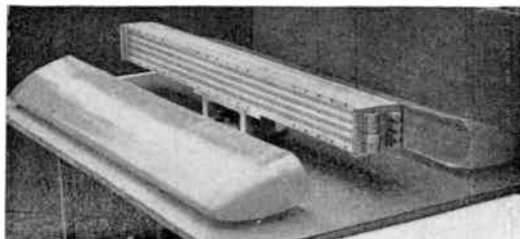
In an earlier paper¹ the first author has demonstrated the improvements which can be realized in the detection of small surface targets in sea clutter by the use of high-speed antenna scanning to decorrelate the clutter echoes. This earlier work used a vertically-polarized transmission from an antenna designed² to rotate at speeds up to 1400 rev/min. A number of workers^{3, 4} have noted a pronounced 'spikiness' or coarseness of the sea-clutter returns from horizontal polarization as compared with vertical polarization from some sea states. Macdonald³ gives some A-scan illustrations which clearly show this increased 'spikiness' especially at angles near grazing incidence. In the present work the authors use a matched pair of antennas horizontally and vertically polarized, and mounted back to back, which may be rotated at speeds up to 600 rev/min. The radar is switched from one to the other every half revolution. The object is to see if there is any advantage to one polarization or the other, in the ability to smooth the sea clutter by decorrelation and constant false alarm rate (c.f.a.r.) techniques, using a visual p.p.i. display.

Recent work in A.S.W.E.⁵ has shown wide variations in the quality of the c.f.a.r. performance against noise of various types of logarithmic receiver even though they may show very good logarithmic laws for c.w. or pulse inputs. For the present work a true i.f. type of logarithmic amplifier developed by the authors was used.⁶ The factors necessary to ensure a constant false alarm rate on a varying r.m.s. amplitude of noise input are investigated, leading to possible reasons why a loss of detectability on weak signals is sometimes experienced when using logarithmic receivers.

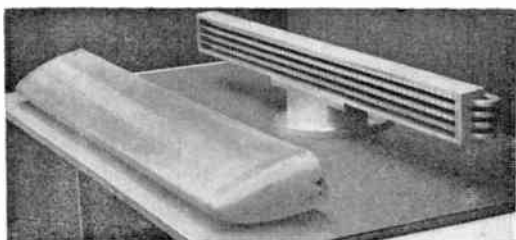
2. Radar System

The radar system used was a conventional X-band magnetron (15 kW peak) transmitter, pulsed by a 0.1 μs modulator in which the repetition rate could be varied from 1 to 10 kHz. The back-to-back antenna assembly in its aerodynamic radome could be rotated at speeds of 20 rev/min to 600 rev/min. Each antenna consisted of a stack of four linear arrays and Fig. 1 makes the arrangement clear. The important parameters of the two antennas, measured at the operating

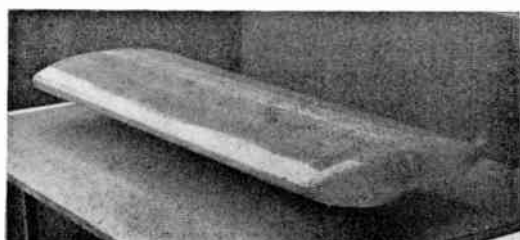
† Admiralty Surface Weapons Establishment Extension, Hambrook, Chichester, Sussex.



(a) Back-to-back antenna, vertically polarized stack of arrays.



(b) Back-to-back antenna, horizontally polarized stack of arrays.



(c) Antenna in radome, dimensions: length 1.83 m (6 ft) 14 cm (5.5 in).

Fig. 1.

frequency, were as shown in the following Table.

| | Vertically-polarized | Horizontally-polarized |
|----------------------------|----------------------|------------------------|
| Horizontal beamwidth | 1.55° | 1.66° |
| Vertical beamwidth | 16° | 14° |
| Gain | 29.95 ± 0.2 dB | 30.1 ± 0.2 dB |
| Worst horizontal side-lobe | -28.5 dB | -27 dB |
| Overall match | 1.2 | 1.2 |
| Power in end load | -17 dB | -15 dB |

The antennas were fed through a twin-channel rotating joint. Below this joint a hybrid waveguide assembly with a 180° sector shorting vane, working in an air-gap, switched the radar power for one half revolution to one channel of the rotating joint and for the next half revolution to the other, the switching time being 3 ms. The two channels of the rotating joint had losses which varied slightly throughout each revolution of the antenna. These losses were measured

and added to the measured differences between the two antennas themselves. This gave figures for probable differences between the two effective antenna gains ($G_H - G_V$) which could be 0.47 dB at the minimum and 0.95 dB at the maximum. This degree of balance was felt to be close enough for the sea clutter comparisons. The antenna was sited on a bank overlooking the sea at a height of 15 m (50 ft) above sea level. As the antenna was 30 m (100 ft) away from the transmit/receive equipment the power was fed to it through a low-loss circular waveguide operating in the H_{01} mode which effectively eliminated some 5 dB of radar loss. This was developed by the Marconi Company⁷ and over a period of two years use in all weathers proved remarkably stable.

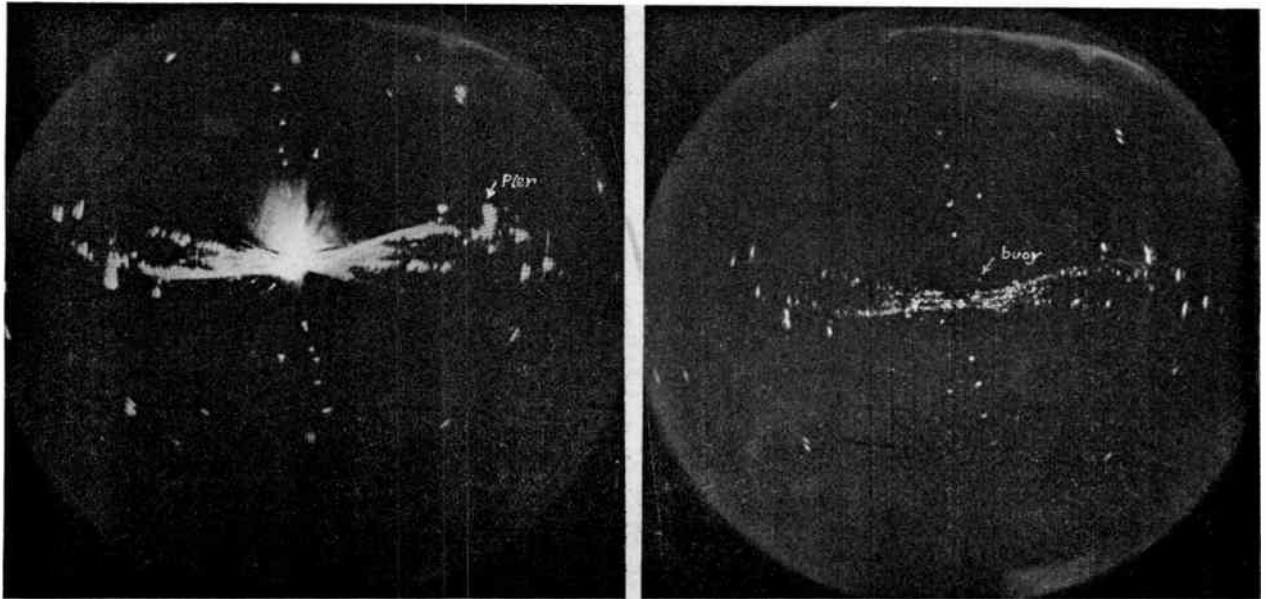
The logarithmic receiver used to provide suppression of the sea-clutter down to the noise level was of the true i.f. type.⁶ The properties of this receiver are dealt with later in the paper under the section dealing with c.f.a.r. characteristics. The technique used for the observations was to keep the repetition rate set to 4500 pulse/s. With the antenna running at 600 rev/min, and assuming a mean horizontal beamwidth of 1.6° for the antennas, this gave two pulses per beamwidth repeated every 0.1 seconds. When the antenna was switched to 20 rev/min the system gave 60 pulses per beamwidth (at the same repetition rate) repeated every 3 seconds, so that the mean radar power was the same at the two antenna speeds. In a previous paper⁴ the first author has shown that the decorrelation time of sea-clutter to be in the region of 8–14 ms, and has measured some 6 dB of improvement in detectability of small signals in clutter by switching between the heavily correlated 20 rev/min system above, and the largely decorrelated 600 rev/min system. It was found essential to retain two pulses per beamwidth (strongly correlated) in the 600 rev/min system to avoid antenna scanning losses. The purpose of the two antenna speeds in the present work therefore was not to establish any measured value for correlation improvements but to see by direct comparison if there was any noticeable difference between decorrelation behaviour on the two polarizations.

3. Clutter Comparisons

The first comparison was made on a fairly smooth sea, waves 15 cm (6 in) to 22 cm (9 in), and is shown in Fig. 2. The total range displayed in all the p.p.i. photographs is 2.24 km (1.4 miles). The coastline runs approximately from left to right. There is a pier at 1.6 km (1 mile) which is clearly marked in Fig. 2(a). The upper 180° of the picture gives the sea aspect with vertical polarization and the lower 180° the same sea aspect with horizontal polarization. In the Fig. 2(a) the fast time-constant after the c.f.a.r. receiver has been switched out in order to show clearly the extent

of the sea clutter. We see that in this sea-state the clutter on vertical polarization is considerably worse than on horizontal polarization. In fact the clutter with vertical polarization extends to a little over three times the range of that with horizontal polarization.

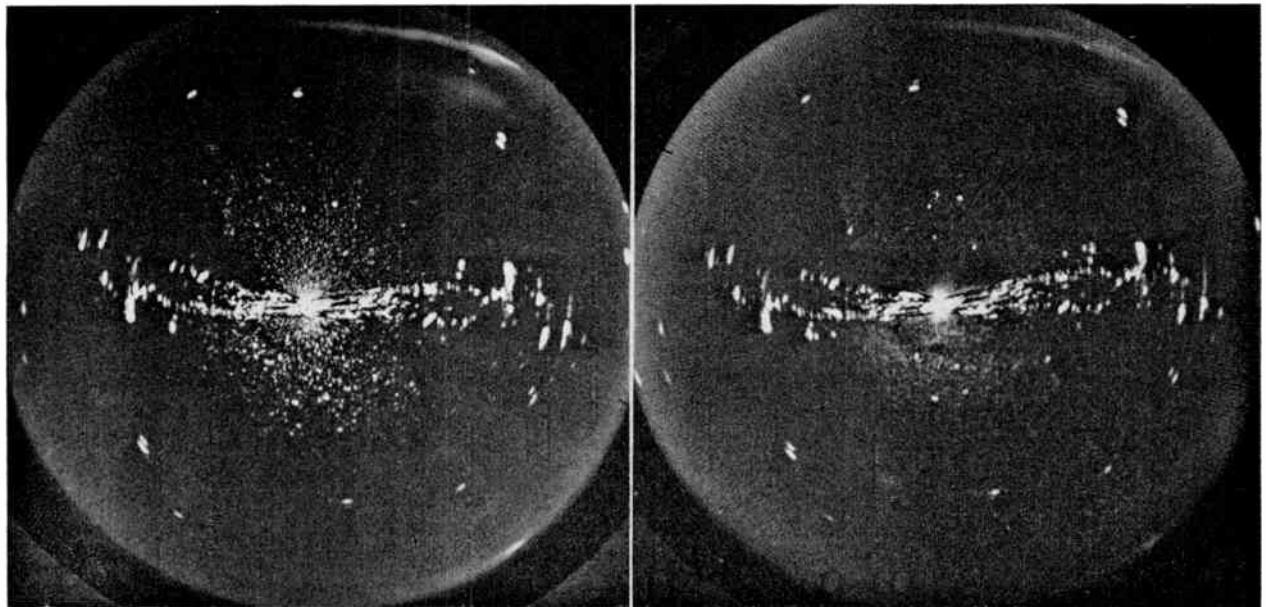
Taking an inverse range cubed law for the clutter/noise power ratio, this would indicate clutter power about 30 times stronger (15 dB) with vertical polarization for this low sea-state; this is in reasonable agreement with data given in Ref. 4. In Fig. 2(b) the fast time-con-



(a) Decorrelated radar. 15 cm (6 in) to 22 cm (9 in) waves. Fast time-constant switched out to show sea clutter. Vertical polarization up, horizontal polarization down.

(b) Fast time-constant switched on to remove clutter.

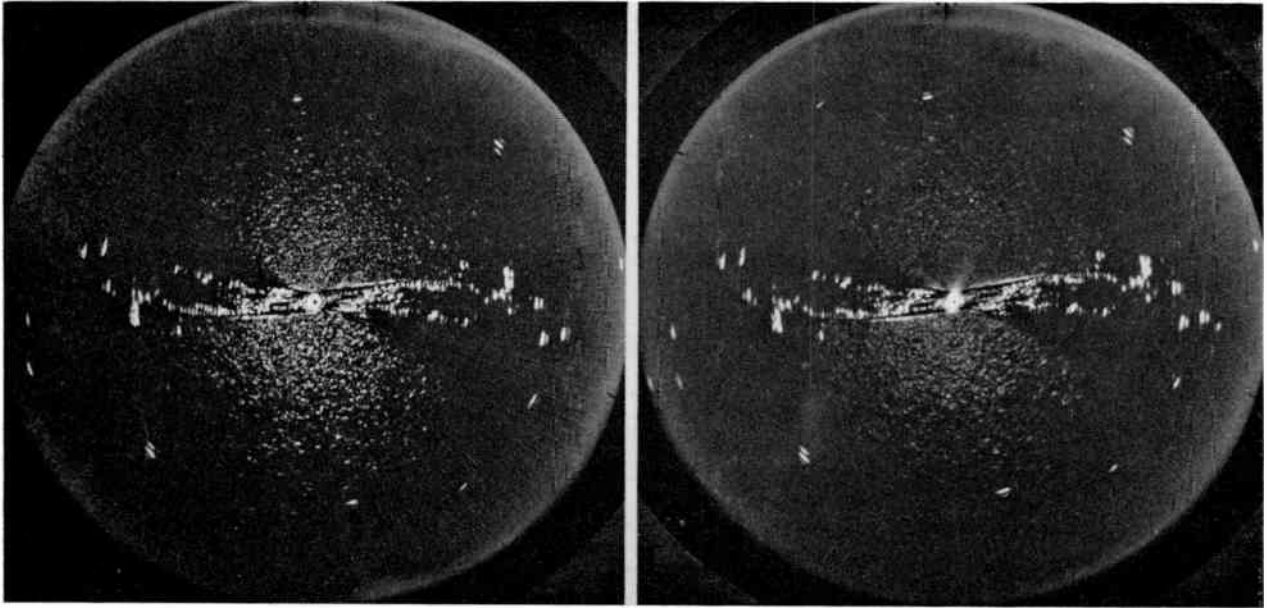
Fig. 2.



(a) Correlated radar (20 rev/min) v.p. up, h.p. down. 30 cm (1 ft) to 45 cm (1.5 ft) waves. 6 seconds exposure.

(b) Decorrelated radar (600 rev/min) 3 seconds exposure.

Fig. 3.



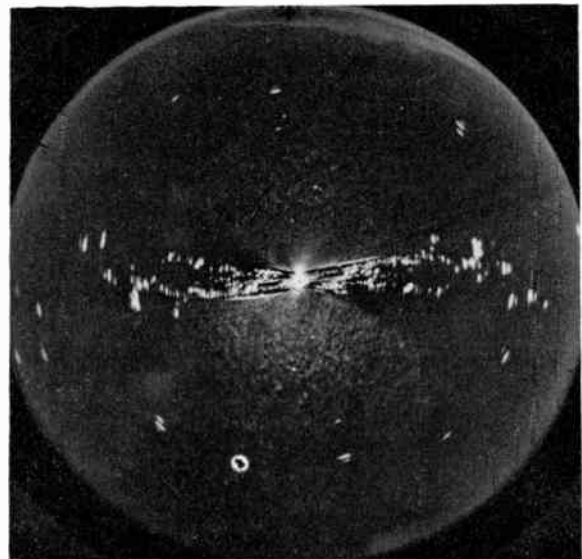
(a) Correlated radar (20 rev/min) v.p. up, h.p. down. 60 cm (2 ft) to 90 cm (3 ft) waves. 3 seconds exposure.

(b) Decorrelated radar (600 rev/min) 3 seconds exposure.

Fig. 4.

stant has been switched into circuit, to reduce the clutter fluctuations to noise level. The arrow shows the echo from a cylindrical mooring buoy just off shore. This buoy was floating with the axis of the cylinder horizontal, this axis lying below the surface in the normal manner of such mooring buoys. Although it may still be clearly detected above the clutter with vertical polarization, the advantage in signal strength above the clutter lies with horizontal polarization. This is of course very much a surface target, the highest generating line of the cylindrical surface would be less than 2 feet above the water surface. Figures 2(a) and (b) were taken with the radar in its decorrelated mode of operation and the extremely effective way in which the clutter is eliminated from the display by the combination of decorrelation and c.f.a.r. receiver is worth noting.

Figures 3(a) and (b) show clutter comparisons with waves 30 to 45 cm (1 to 1.5 ft), in height. The fast time-constant from now on is permanently in circuit to remove the mean level of the clutter. Figure 3(a) shows the results with the 20 rev/min correlated system. There is now nothing to choose between the two polarizations regarding the range to which clutter extends. But we can very definitely see the coarser character of the clutter with horizontal polarization. This evidence of persistent target-like echoes in horizontally polarized clutter has been noticed by numbers of workers.^{3, 4} When we switch to the decorrelated system considerable smoothing of the clutter occurs with both polarizations but the target-like

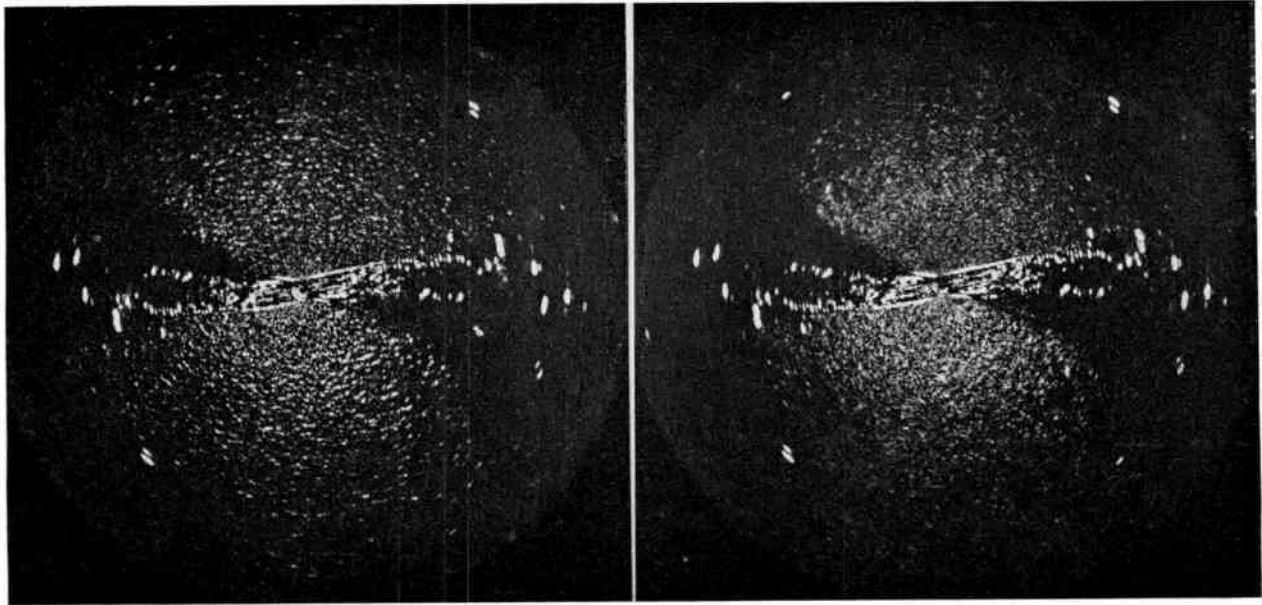


(c) Decorrelated radar (600 rev/min) 60-90 cm (2-3 ft) waves, v.p. up, h.p. down. 24 seconds exposure.

Fig. 4.

echoes are still to be noted in the horizontally polarized clutter. The four genuine targets in the clutter are clearly resolved.

The next comparison is of clutter from 60 cm to 90 cm (2 ft to 3 ft) waves. From Fig. 4(a) there is again no difference between ranges of clutter for the two polarizations. The clutter again is still somewhat finer for the vertical polarization than for the hori-



(a) Correlated radar (20 rev/min) v.p. up, h.p. down. 2.1-3 m (7-10 ft) waves, 3 seconds exposure.

(b) Correlated radar (20 rev/min) 24 seconds exposure.

Fig. 5.

zontal polarization but the difference is less marked. The smoothing of clutter in the decorrelated system is still effective for both polarizations, the slightly increased coarseness of the horizontally polarized clutter being still evident. Three surface targets in the clutter can just be identified in the vertically polarized picture. When in Fig. 4(c) we integrate the picture for 24 seconds (camera integration) they show up much better.

The final comparison was made with waves of 2.1 m to 3 m (7 to 10 ft). Figure 5(a) and (b) are for the 20 rev/min correlated radar system. There is in Fig. 5(a) no noticeable difference, either in range or coarseness of the clutter area, between the two polarizations. In Fig. 5(b) the camera has integrated the display for 24 seconds. The vertical polarization shows perhaps marginally a finer structure. There is no possibility of locating the one genuine target in the clutter region. In Fig. 6(a) we have switched the radar to the 600 rev/min decorrelated mode still with a 24-second integration time (compare Fig. 5(b)). Clutter smoothing is obvious and we can just detect the surface target (5 m²) vertically above and below centre at just over one third full range. The 60 seconds of integration (Fig. 6(b)) brings the target out clearly. Even in this sea there does seem a slight advantage to the smoothing of clutter when vertical polarization is used.

These differences (marked for small seas) in the clutter coarseness for the two polarizations have been explained in Ref. 4 in terms of the lobing structure of the energy over the sea. In Fig. 7 the authors have

computed such a diagram for the conditions of the present trials (antenna height 15 metres) and this shows the lobe structure variation with height at a range of 500 metres. The illumination does not drop to zero at all for vertical polarization, because of 'Brewster angle' behaviour. This means that for small seas, lying mainly in the first lobe, the excursions of illuminating energy for the wave peaks are smaller when using vertical than when using horizontal polarization so that the clutter has less violent modulations. For the higher seas as one goes through several lobes these differences tend to disappear.

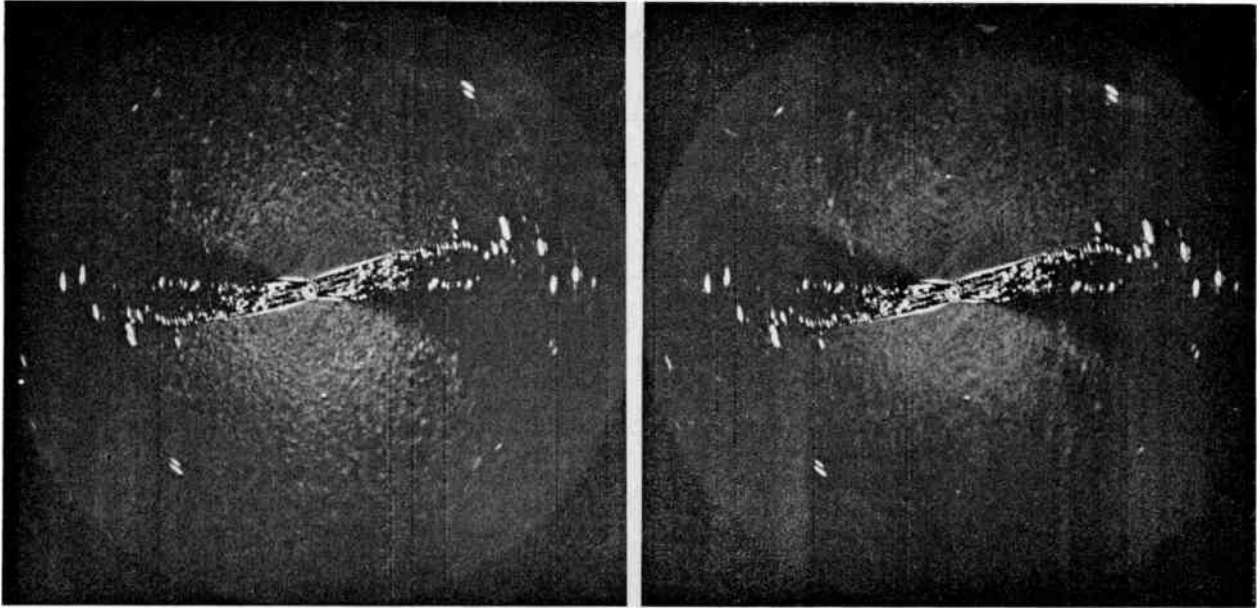
4. C.F.A.R. Behaviour of Logarithmic Receivers

In an earlier paper⁸ the first author has examined analytically the behaviour of a logarithmic receiver to random inputs having a Rayleigh probability amplitude distribution. It was shown that if the receiver could have an idealized logarithmic characteristic given by

$$y = a \log_e (bx) \quad \dots\dots(1)$$

where y is receiver output, x receiver input, and a and b are constants of the amplifier, then the output fluctuations about the mean for a random (Rayleigh distributed) input such as noise or certain types of clutter would be constant and independent of the variations in amplitude of the input fluctuations about the mean, that is,

$$E(y^2) - \{E(y)\}^2 = \frac{a^2 \pi^2}{24} \quad \dots\dots(2)$$



(a) Decorrelated radar (600 rev/min) v.p. up, h.p. down. 2.1-3 m (7-10 ft) waves. 24 seconds exposure.

(b) Decorrelated radar. 60 seconds exposure.

Fig. 6.

Because the idealized characteristics of eqn. (1) mean an output of minus infinity for zero input it cannot be realized. The practical receiver has an output of zero for zero input becoming rapidly logarithmic for positive inputs, i.e.

$$y = a \log_e (1 + bx) \quad \dots\dots(3)$$

In Ref. 9 it was further deduced that because of this departure from the idealized law a correction term would have to be applied to the constant output given by eqn. (2). The corrected output is given by

$$E(y^2) - \{E(y)\}^2 = \frac{a^2}{4} \left[\frac{\pi^2}{6} - \left(\frac{2}{b^2 \sigma^2} \right) \times \left\{ 1 + \log (b^2 \sigma^2) - \gamma + \frac{1}{2b^2 \sigma^2} \right\} \right] \quad \dots\dots(4)$$

- where a is a constant (slope of logarithmic law)
- b is a constant (gain of the amplifier)
- σ = r.m.s. input fluctuations
- γ = Euler's constant

In eqn. (4) the correction term is the bracketed product, which is subtracted from the constant term $\pi^2/6$. This correction term rapidly dwindles as the input random fluctuations (noise, sea clutter) are moved upwards on the logarithmic part of the input-output law of the practical receiver. In Table 1 we show how the output fluctuations vary as a percentage when the input fluctuations (in decibels) move on to the logarithmic curve.

We see that when the noise fluctuations are occupying 20 dB of the logarithmic characteristic the output

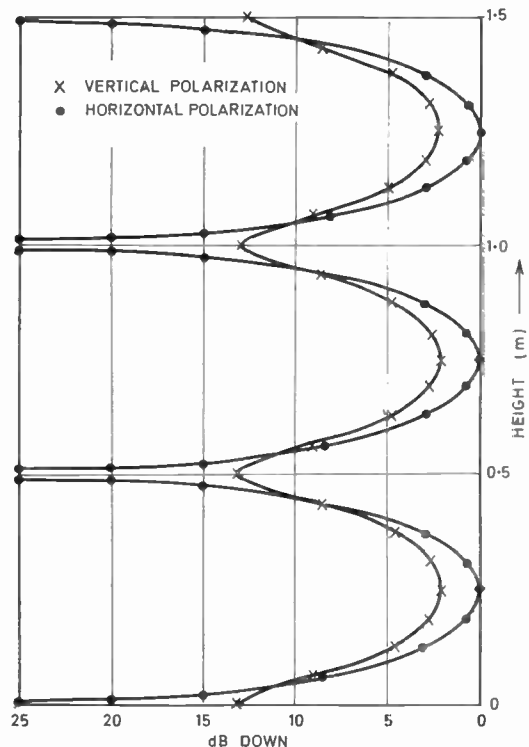


Fig. 7. Illumination over sea at 500 metres range (X-band); Transmitter height 15 metres; for vertical and horizontal polarization.

Table 1

| Input fluctuations dB | Output fluctuations % |
|-----------------------|-----------------------|
| 3 | 19 |
| 6 | 42 |
| 12 | 72 |
| 20 | 94 |
| 30 | 99 |
| 40 | 100 |

fluctuations have reached 94% of their full value. A further 10 dB of fluctuations lifted on to the logarithmic law only contributes a further 5% to the output r.m.s. fluctuations raising them to 99% of full value. This means that looking at the output fluctuations on an A-scope the eye might just notice the increase when the input was raised from 20 dB to 30 dB but beyond that the output would be virtually constant.

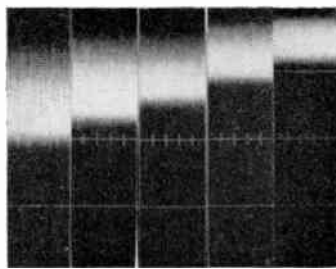
In recent work⁵ A. Reiss and G. W. Carver have examined the behaviour of various types of logarithmic (c.f.a.r.) receivers when fed from a random noise input varying over 70 dB in r.m.s. amplitude. They found that the steeply asymptotic approach to a constant output shown in Table 1 did not in fact obtain in any type of receiver tested. Some designs came nearer to the behaviour of Table 1 than others, but all showed a continuing and obvious increase in magnitude of output fluctuations as the input was increased. This gradual increase in noise output has been called 'noise funnelling', though 'noise divergence' seems a more meaningful term. Amplifiers in which the bandwidth restriction was made mainly before the logarithmic stages, or was provided partly in this position, and partly in a lumped filter half way along the logarithmic chain showed least noise divergence. Those in which the band-pass filters were evenly distributed throughout the logarithmic chain showed most noise divergence. It seemed clear that the divergence behaviour was in some way connected with bandwidth restriction in saturated stages. The authors have examined the noise divergence behaviour of their own true i.f. logarithmic amplifier⁶ used for the clutter comparisons in the first part of the present paper. Two logarithmic amplifiers were used. Amplifier number one had a bandwidth of 10 MHz to the 3 dB points, and 16 MHz to the 40 dB points, centred at 60 MHz. Amplifier number two had a bandwidth of 20 MHz to the 3 dB points, and 37.0 MHz to the 40 dB points, centred at 60 MHz. Amplifier number one therefore, besides having half the bandwidth of amplifier number two, also had a gradient on the skirts of its response curve of nearly three times that of amplifier number two. The amplifiers were fed from a noise-generating amplifier which had a flat response from 53 MHz to 67 MHz, and a band of 50 MHz to 70 MHz between

the 3 dB points. At 40 dB down its band covered 41 MHz to 95 MHz. An accurate decade attenuator was inserted between this noise generator and the logarithmic amplifier.

Divergence of noise in the 10 MHz bandwidth logarithmic amplifier (No. 1) is shown in Fig. 8(a). This amplifier has the steep skirts to its response. From right to left we see the noise fluctuations as we move the noise from 20 dB up the logarithmic characteristic to 60 dB up this characteristic. The amplitude (in accordance with eqn. (4)) should only grow by 6% over this range. It clearly increases some $2\frac{1}{2}$ times.

Divergence of noise in the 20 MHz bandwidth amplifier (No. 2), having the gently sloping skirts to its response is shown in Fig. 8(b). The divergence is much less, the noise amplitude increasing about $1\frac{1}{2}$ times. The noise looks coarse compared with Fig. 8(a), but this is because far fewer traces were photographed (on a faster film) to avoid excessive camera integration, and thus show the noise structure more clearly. The noise divergence, is thought to be caused by generation, in the 'limiting' stages of the logarithmic amplifier of noise components on the skirts of the filters in these stages. It seemed sensible therefore to place a very steep sided filter between the noise generator and the logarithmic amplifier, in order to eliminate input components at the frequencies of these skirts. The bandwidth of the filter was 16 MHz to the 3 dB points and 21 MHz to the 40 dB points, so that the slope of the skirts is approximately 15 dB per megahertz. Figure 8(c) shows the effect of including this filter; the noise divergence effect is significantly reduced. However the use of such a filter at the beginning of a high gain amplifier chain is inadmissible radar-wise. Strong pulses such as the transmitter pulse (ground wave), and reflections (from close range targets) which may be 80 dB or more above noise level, produce very severe range side-lobes (appearing as a 'ringing' response) from such a filter, which would cause extensive clutter responses on the display, and spoil the range discrimination performance and definition. By analogy with antenna aperture distribution theory, the filter which would give no range side-lobes would presumably have a Gaussian amplitude response, but this is not likely to give the rejection characteristic required.

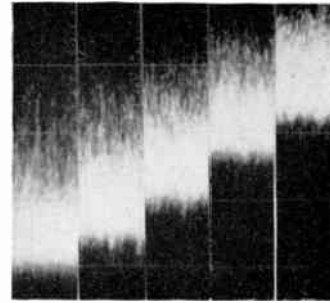
The next experiment was to see if the divergent components of the noise generated in the non-linear stages could be removed by a low-pass filter in the video amplifier circuit. A low-pass filter with a flat response up to 5 MHz and a cut off at 6 MHz (suitable for $\frac{1}{10}$ μ s pulses) was included in the video amplifier stage. The result is shown in Fig. 8(d); clearly the noise divergent components have been entirely removed and are therefore all generated as higher frequency components. This may partly



Number of decibels up logarithmic characteristic

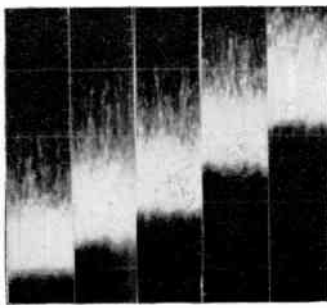
(a) Divergence of noise in 10 MHz bandwidth, 60 MHz centre frequency, logarithmic amplifier (No. 1). This amplifier had steep sides to band-pass (12 dB/MHz). Increase in noise amplitude from right to left may be noted.

This amplitude should be constant. Amplifier was fed through a decibel attenuator.



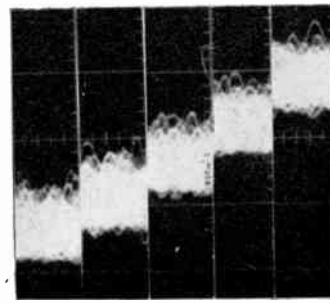
Number of decibels up logarithmic characteristic

(b) Divergence of noise in 20 MHz bandwidth, 60 MHz logarithmic amplifier (No. 2) having less steep sides to band-pass (3.5 dB/MHz). Fewer traces (only 50) have been photographed to show noise structure more clearly by reducing camera integration. Note that relative increase in noise amplitude (right to left) is less than that in Fig. 8(a).



Number of decibels up logarithmic characteristic

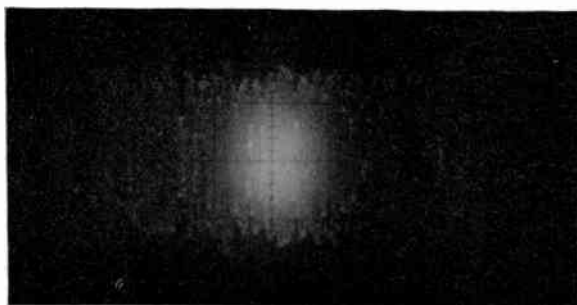
(c) As Fig. 8(b) but with input i.f. filter between noise generator and logarithmic amplifier. This filter had very steep sides (16 MHz bandwidth at 3 dB and 21 MHz at 40 dB). Note reduction of noise divergence.



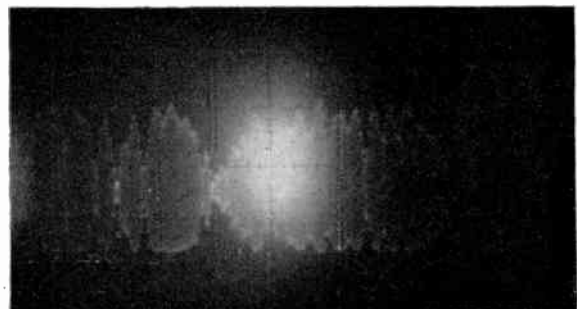
Number of decibels up logarithmic characteristic

(d) As Fig. 8(b), i.e. no input filter, but with a video low-pass filter having a sharp cut-off (full response up to 5 MHz, cut-off at 6 MHz) after detector. Note absence of noise divergence.

Fig. 8.



(a) I.f. output (noise) at output of final i.f. stage of amplifier No. 2 (bandwidth 50-70 MHz at 3 dB points).



(b) Same output as in Fig. 9(a) after passing through a single linear i.f. stage including a steep-sided band-pass filter (53.8-66.2 MHz at 3.0 dB points, 51.5-68.5 MHz at 40 dB points). Note steep modulation of envelope with the filter.

Fig. 9.

explain the losses experienced with logarithmic receivers on weak signals (3 dB or less above noise), when the receiver is followed by a differentiating circuit (high-pass filters), and d.c. restoration.⁸ Even with the receiver noise fluctuations lifted only 20 dB up the logarithmic characteristic, as normally used for clutter rejection, some noise divergent components may be present, and the high-pass filter and restoration circuit will if anything accentuate these compared with the lower frequency components within the pass-band.

It was now argued that an i.f. filter with a similar response (on both sides of the carrier frequency) as the video low-pass filter would, if included in a linear i.f. stage at the end of the logarithmic chain just before the detector, have the same effect in removing the noise divergent components as the video filter. Such a filter, although very steep-sided, would not have its range side-lobes amplified further, and these would therefore be below noise level. The steepest-sided filter we could assemble was therefore tried (53.8 MHz–66.2 MHz at 3 dB down, 51.5 MHz–68.5 MHz at 40 dB down). In fact it seriously increased the noise divergent components. Figures 9(a) and 9(b) attempt to show this. They are pictures which show i.f. noise, with the noise generator turned up so that the noise fluctuations are lifted near to the top of the logarithmic range of the receiver. The pictures are faint because a single trace only was photographed to try to show the presence of the i.f. carrier. Figure 9(a) shows i.f. noise without the filter and the envelope excursions will represent the magnitude of the rectified video noise. Figure 9(b) shows the effect with the filter in position and the increase in the envelope excursions is obvious. Why should this be? One difficulty is that in constructing this filter we arbitrarily centred it at 60 MHz, which was the nominal centre of the response band (to the half power points) of the logarithmic amplifier. In fact at the high input noise level used for the oscillograms of Figs. 9(a) and (b) there can be noise components generated in the wide-band noise generator which, though well down the skirts of the response of the logarithmic amplifier, can be amplified in a non-linear manner, to give high outputs in the later stages of the logarithmic amplifier. Our final filter may now process these in a quite asymmetrical fashion about a 60 MHz carrier which may momentarily be at a very reduced level, and this may well lead to the sort of envelope distortions shown in Fig. 9(b). Clearly the filter in this position would need to hop about in frequency in order to behave in a symmetrical fashion about the strongest carrier present. A video filter of course automatically stays symmetrical to the rectified envelope and we do not have this trouble.

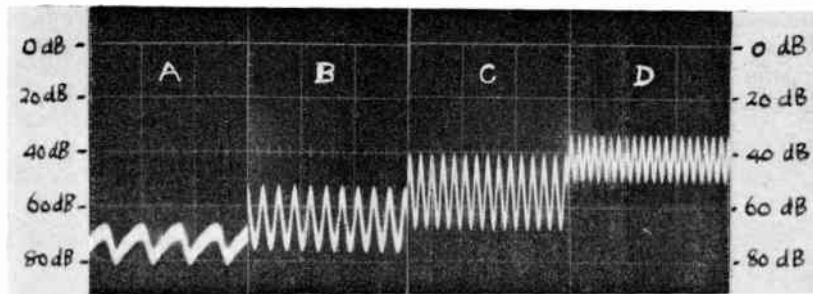
This experiment with the i.f. filter thus confirmed the suspicion that the noise divergence was caused by

noise components generated outside the flat pass band of the logarithmic amplifier, and then asymmetrically amplified on the skirts of its response curve. Of course similar effects will occur in so-called 'linear' amplifiers if bandwidth-extensive noise of steadily increasing amplitude is applied to the input, since non-linear amplification will very soon commence.

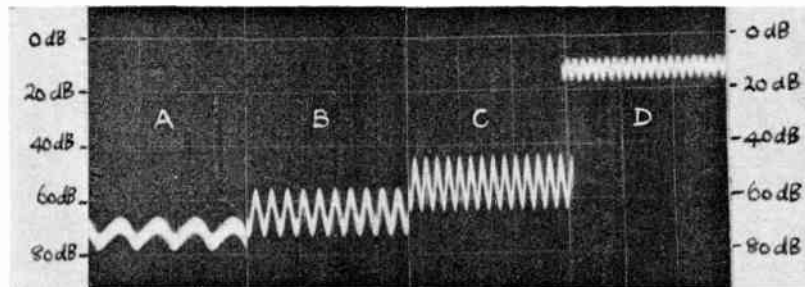
To gain further insight into the effect of noise components amplified on the skirts of the responses of the tuned coupling circuits of the logarithmic amplifier, another experiment was tried. Two signal generators were connected to the input stage of logarithmic amplifier number two (20 MHz bandwidth at 60 MHz). One of them was set to provide a 60 MHz signal, which we shall call the carrier, and at a level which brought it about 70 dB up the logarithmic input-output characteristic, i.e. near to the upper limit of the logarithmic range. This level was left fixed throughout the observations. The signal (called the skirt signal) from the other generator, was set at different frequencies both inside and outside the 3 dB pass band of the amplifier and its amplitude was set to a range of levels from 0 to 12 dB above the carrier. Some of the results are shown in Fig. 10. In Fig. 10(a) we have the skirt signal set to an amplitude 9 dB above the carrier, and reading from left to right the skirt signal is tuned to 57 MHz at A, 53 MHz at B, 50 MHz at C and 46 MHz at D. We note how the resultant signal moves progressively upwards towards the zero line (no output). In Fig. 10(b) the same sequence is taken but for a skirt signal 12 dB above the carrier level and we see that at D the amplifier has almost been completely desensitized by the strong skirt signal well outside its 3 dB pass band. In Fig. 10(b) if, at D, the skirt signal was switched off the carrier trace immediately jumped downwards near to the 80 dB output level from the amplifier. This is akin to demodulation of a weak signal by a strong one.

At this point it is instructive to look back to Fig. 8(b). In this figure again the zero line is at the top, the direction for noise increasing is downwards, which we will call the positive direction. Note that the positive crests of the noise do not change in character from right to left over the 50 dB range of inputs shown. But the negative-going troughs do become very much extended, indeed all the noise divergence is in the negative direction towards the zero line. By analogy with Fig. 10(b), at D, these deep troughs of noise correspond to momentary desensitization of the amplifier to the noise carrier by intense peaks of noise outside the normal pass band, i.e. on the skirts of the amplifier response.

To what extent does all this bother a navigation radar with a logarithmic receiver, differentiating circuit (high-pass filter) followed by d.c. restoration to avoid



(a) Effect of carrier (constant amplitude at 60 MHz) plus skirt signal moved from inside to outside the i.f. band (50–70 MHz). A, skirt signal at 57 MHz; B, skirt signal at 53 MHz; C, skirt signal at 50 MHz; D, skirt signal at 46 MHz. Skirt signal 9.0 dB above fixed carrier level. Note weakening of carrier level from A to D.



(b) As for Fig. 10(a) but with skirt signal 12.0 dB above the same fixed carrier level. Note severe weakening of carrier level at D with strong skirt signal outside the 3.0 dB bandwidth.

Fig. 10.

the shadows of differentiation? First of all, sea clutter and rain clutter signals should not show divergence, because their spectrum is fixed (apart from the tiny Doppler effects) by that of the transmitted pulse; and if the receiver has an adequate bandwidth to accommodate the pulse spectrum there will be no significant components of this pulse on the skirts of the response curve of the amplifier. Receiver noise level is normally and correctly set about 20 dB up the logarithmic characteristic. At this level, with probably only the last two stages of the amplifier noise limited, a small noise divergent component may be present, and this could show itself in a loss of detectability of a weak signal, when the divergent components of noise are processed by differentiation and d.c. restoration. A low-pass video filter before differentiation should reject the divergent components (Fig. 8(d)), and avoid this loss.

Only in the presence of wide bandwidth interference (noise or impulsive) could the noise divergent components become really troublesome. As has already been stressed, severe rejection at the beginning of the amplifier chain by a steep-sided filter having a 3 dB bandwidth approximate to the pulse length is inadmissible because of range side-lobe effects. We therefore cannot eliminate this wide-band interference in this way. The right choice of filter to minimize these

interference effects at the input would be a matter of study. Probably a single tuned circuit, having a 3 dB bandwidth approximate to the transmitted pulse length, backed by a steep-sided filter having about three times this bandwidth at its 3 dB points would be effective. This would give something approximating to a Gaussian response on a rectangular pedestal which by analogy with antenna aperture theory (see Ref. 9, pages 47–50 on the incomplete Gaussian distribution) should give very small range side-lobes.

To sum up then, the design of a logarithmic amplifier to minimize noise divergence effects would conform to the following rules.

(a) Avoid the use of steep-sided filters as couplings between successive amplifying stages. The gentler the slope to the skirts of the overall response the less the noise divergence.

(b) A very steep-sided band-pass filter (3 dB bandwidth matched to pulse length) between the head amplifier and the logarithmic amplifier though it will reduce noise divergence, is to be avoided because of the range side-lobes and pulse distortion introduced on strong pulses and accentuated by the following amplifier.

(c) Where some noise divergence has occurred, it can be eliminated by a low-pass filter with a sharp cut-

off beyond half the 3 dB bandwidth of the i.f. chain, incorporated into the video amplifier before the differentiating circuit (fast time-constant) and d.c. restoration.

(d) To minimize the effects of wide-band noise interference, impulsive interference, and their inter-modulation products, some protection will be necessary at the input. This should be a filter (possibly Gaussian on a rectangular pedestal) which provides strong rejection of interference components two or three i.f. bandwidths away from the transmitter frequency, without causing distortion of strong received pulses and the generation of range side-lobes thereof.

5. Conclusions

For low sea-states (up to 22 cm waves) and angles near grazing incidence the sea clutter returns from vertically-polarized (X-band) radiation are about 15 dB stronger than for horizontally-polarized radiation. The signal/clutter ratio for a surface target lying in the clutter is somewhat greater for horizontal polarization than for vertical polarization and where smooth seas can always be relied upon (e.g. for river radar) there would be a small advantage in using horizontal polarization.

For sea-states giving waves of 30 cm to 45 cm the horizontally-polarized clutter becomes as strong as the vertically-polarized clutter. Moreover, the horizontally-polarized clutter has a much coarser and more target-like appearance than the vertically-polarized clutter. Clutter smoothing techniques such as decorrelation reduce the vertically-polarized clutter to a more noise like background than the horizontally-polarized clutter. It is somewhat easier to detect a genuine target against this background.

For sea-states giving waves of 60 cm to 90 cm, the range of clutter on the two polarizations remains equal. The increased coarseness of clutter using the horizontal polarization is still noticeable compared with vertical polarization but is less marked than with the lower sea-state. Again decorrelation makes the vertically polarized clutter assume a slightly more noise-like background than the horizontally-polarized clutter.

For sea-states giving 2.1 m to 3 m waves the differences between the coarseness of the two types of clutter disappear. Only after decorrelation and prolonged smoothing by integration (60 seconds) can a marginally coarser structure in the horizontal polarization be noticed.

In all sea-states above calm (15 cm to 22 cm waves) the use of vertical polarization, with c.f.a.r. receiver and decorrelation techniques, gives a somewhat better chance of picking out a surface target than do the

corresponding techniques on horizontal polarization. The advantage of vertical polarization gets less as the sea state rises.

Constant false alarm rate receivers which have very accurate logarithmic laws when measured with c.w. or pulse input, may show a wildly erroneous behaviour when measured with a steadily increasing noise input of bandwidth greater than the receiver itself. For such bandwidth-extensive noise inputs the c.f.a.r. behaviour (constant r.m.s. output fluctuation) may never be obtained, the output r.m.s. fluctuations increasing continually. This error, called noise divergence, has been found to be worst in those receivers where a steep-sided band-pass filter is associated with each non-linear amplifying stage. The noise divergence is caused by noise components generated in the relatively wide bandwidth of the head amplifier or input circuit, and then non-linearly amplified on the skirts of the steep-sided filters of the limiting logarithmic stages. Methods of minimizing noise divergence effects are suggested.

Divergence should not be present on sea clutter and rain clutter where the clutter spectrum is limited by the transmitted pulse, provided the receiver has adequate bandwidth to reproduce this spectrum.

6. Acknowledgments

The authors acknowledge valuable discussions with Miss E. A. Killick, Mr. A. Reiss and Mr. J. R. Thomas.

7. References

1. Croney, J., 'Improved radar visibility of small targets in sea clutter', *The Radio and Electronic Engineer*, **32**, No. 3, pp. 135-148, September 1966.
2. Delany, W. and Kyle, R. F., 'Antenna for rapid scan decorrelation radar', *The Radio and Electronic Engineer*, **32**, No. 3, pp. 156-8, September 1966.
3. Macdonald, F. C., 'Characteristics of Radar Sea Clutter, Part 1 Persistent Target-like Echoes in Sea Clutter', N.R.L. Report 4902, March 1957.
4. Boring, J. G., Flint, E. R., Long, M. W. and Winderquist, V. R., 'Sea Return Study' (Final Report Project No. 157-96, 1st August 1957) Georgia Institute of Technology.
5. Reiss, A. and Carver, G. W., 'The response of logarithmic amplifiers to noise signals', *Electronics Letters*, **4**, No. 14, p. 283, 12th July 1968.
6. Woroncow, A. and Croney, J., 'A true i.f. logarithmic amplifier using twin-gain stages', *The Radio and Electronic Engineer*, **32**, No. 3, pp. 149-55, September 1966.
7. Cutliff, M. H., 'The H_{01} mode and the helix waveguide', *Marconi Review*, **28**, No. 156, pp. 1-21, January-March 1965.
8. Croney, J., 'Clutter on radar displays—reduction by use of logarithmic receivers', *Wireless Engr*, **33**, pp. 83-96, April 1956.
9. Ramsey, J. F., 'Fourier transforms in aerial theory: part 6, conclusions', *Marconi Review*, **11**, pp. 45-50, 1948.

Manuscript first received by the Institution on 13th December 1968 and in final form on 12th June 1969. (Paper No. 1280/AMMS.24.)

© The Institution of Electronic and Radio Engineers, 1969

Contributors to this Issue



D. R. Towill is at present Reader in Engineering Production and Head of the Dynamic Analysis Group at the University of Wales Institute of Science and Technology, Cardiff. He was formerly a senior lecturer in automatic control at the Royal Military College of Science, a consulting engineer, and a dynamic analyst at the Bristol Aircraft Company. He served an apprenticeship with Newman Industries Limited and was educated at the Universities of Bristol and Birmingham.

Mr. Towill is the author of many papers on feedback control and production engineering, and was awarded the J. Langham Thompson Premium for the outstanding paper on control engineering published in the Institution's *Journal* in 1966. His book 'Transfer Function Techniques for Control Engineers' is to be published shortly.



Dr. P. Daly is a lecturer in the Department of Electrical and Electronic Engineering at the University of Leeds. He graduated in electrical engineering from Glasgow University in 1959 and after a further three years at Glasgow studying for the Ph.D. degree was appointed to a two-year research fellowship at the California Institute of Technology in Pasadena. He then worked

for a year in the Mathematics Department at Glasgow University before undertaking research fellowships for one year in the Technische Hochschule, Aachen, and for a further year in the Laboratory for Electromagnetic Theory in Copenhagen. He returned to England to take up his present appointment in 1967.



Shamsuddin Ahmed received the B.Sc. degree in electrical engineering from the University of Dacca in 1958 and the M.S. degree in electrical engineering from the University of California in 1963. From 1958 to 1963 he was a lecturer in electrical engineering in Ahsnullah Engineering College, Dacca. In 1963 he was appointed an assistant professor of electrical engineering

in the East Pakistan University of Engineering and Technology and in 1966 he joined the Government College of Engineering, Rajshahi, East Pakistan, as an associate professor of engineering. At present Mr. Ahmed is working for a doctorate in electrical engineering in the Department of Electrical and Electronic Engineering of the University of Leeds.



Malcolm Dudson read physics on an Open Exhibition at Durham University, and after graduating served for two years as a Technical Signals Officer in the Royal Air Force.

He then joined S. Smith and Sons, and having completed a graduate apprenticeship, worked in their Central Research Unit until joining Computer Development Ltd. in 1960, where he was engaged on

the development of the I.C.T. 1301 Computer. He has held various positions in the ICL Group ever since; his present position being Manager of Advanced Technology at the Data Recording Instrument Company Limited. He has represented the Electronic Computer Manufacturers' Association on the U.K. Delegation at I.S.O. meetings on magnetic tape standards.

Biographies of Dr. J. Croney and Mr. A. Woroncow will be printed in the November issue of the *Journal*.

LETTERS

1/f NOISE

I hope Messrs. Bloodworth and Hawkins will forgive me for saying that their paper¹ is an important contribution to knowledge of m.o.s.t.s but does not solve the central problem of 1/f noise. Taking first the common ground, they and I are agreed that the thermodynamic problem is so remote from practical magnitudes that it need not concern us much (but see below), and that there is no future in experimental work aimed at extending the low-frequency range of observations below the 10⁻⁴-10⁻⁵ Hz of experiments which have already been performed.

What I doubt is the implied assumption that a mechanism which is plausible in the particular structure of a m.o.s.t. is 'the mechanism of 1/f noise' and therefore applicable to the great variety of materials and structures which exhibit 1/f noise. In this connexion I maintain that in general the frequency index is nearer to -1 than the authors found. (There is some suggestion that the magnitude of the index increases with the magnitude of band-gap of the semiconducting material, the closest approximation to -1 being obtained in specimens such as carbon films and thin metal films in which the excitation energy is measured in milli-electron-volts, but I am not sure how this would fit in with the tunnelling hypothesis.) It is well established that 1/f noise corresponds to a fluctuation in the total number of mobile charge carriers, and I would cite particularly the following experiments:

- (i) Brophy and Rostoker showed that fluctuations in conductivity were accompanied by fluctuations in Hall coefficient corresponding to the same fluctuation in number of mobile charge carriers.²
- (ii) Brophy showed that also in the absence of a steady current there were fluctuations in the number of carriers present, indicated by fluctuations in the thermoelectric e.m.f.³
- (iii) Hooge has concluded that the 1/f noise which he has recently found in thin gold films is due to fluctuation in number of mobile carriers.⁴

The undoubted effect of surface conditions in germanium and silicon (have surface effects been proved to be important in any other material?) is then viewed as arising from the effect of surface states on the life-time of mobile carriers.

Hooge's observation of 1/f noise in thin films of metallic gold (proved to be continuous metallic films by their conductivity) introduces a striking change in the picture since this type of noise had previously appeared to be a property of semiconductors. The most refined previous set of experiments was that of Bittel and Scheidhauer⁵ who obtained a negative result on both platinum and nickel-chrome wires down to a well-defined limit of size and current-density. Hooge's measurements fall slightly below the lower limit of Bittel and Scheidhauer, so his observations are not inconsistent with theirs. There is, however, the further interesting point that Bittel and Scheidhauer used wire specimens of which the surface/

volume ratio was obviously less than that of Hooge's films. If surface effects were important in this case also, the change in shape might be as important as the reduction in size. But Hooge found the conductivity to be substantially that of bulk gold, and little affected by film thickness, a factor which discourages the idea that surface effects are dominant.

With regard to the queuing theory, the suggestion that it implies the recognition of individual carriers is not to be lightly dismissed; but in this respect how does it differ from generation-recombination noise, the spectrum of which is characteristic of carrier life-time? In addition, the queuing process provides a mechanism for the total number of charge carriers (the size of the queue) to experience fluctuations which are of unlimited extent in both size and duration, as required to produce a 1/f spectrum. But in spite of the comment above that any objection on thermodynamic grounds to 'unlimited' fluctuations is practically irrelevant, there is a possible difficulty. If the process is even approximately ergodic† there will be variations of resistance amongst large numbers of specimens prepared from the same material, which may be large enough to be measurable. It is not a certain difficulty, however, since a solid grown from a completely disordered phase (crystal grown from a liquid or film deposited from a vapour) might thereby be brought into a standard condition initially.

D. A. BELL,

Department of Electronic Engineering,
The University of Hull.
27th August 1969

Authors' reply: We are happy to forgive Professor Bell for not sharing our opinion that recent work on m.o.s. transistors attacks the central problem of 1/f noise, since we may see a different problem! As explained previously,¹ m.o.s. transistors make possible a systematic investigation of noise generation at the silicon-silicon oxide interface present in most semiconductor devices now manufactured, including a study of the effect of the gate field on the physical process involved, whatever that may be. It was suggested (not assumed) that the tunnelling model might be applicable to other semiconductor devices with an insulating surface layer which are less amenable to investigation. Besides planar devices, with a thick oxide layer, we may include 'bulk' samples oxidized in air to a thickness of the order of 200 μm, and thin film components deposited on an insulating substrate. The applicability of the model encompasses virtually every semiconductor specimen of practical importance.

If we may accuse Professor Bell of an implied assumption, it is that surface effects cannot predominate in noise generation when bulk effects predominantly determine the

† 'Ergodic' means that sample values taken from one specimen at many different instants are equivalent to samples taken from many identical specimens at the same instant.

mean conductivity of the specimen. It may or may not be legitimate to apply this assumption to the interesting work of Hooge⁴, which we hope will be published. Similar measurements have been reported (e.g. by Bergero *et al.*⁶) on discontinuous gold films exhibiting non-metallic conduction.

Brophy and Rostoker^{2,3} apparently ignored surface effects and simply assumed that their samples were homogeneous, after subjecting them to sandblasting, etching, and plastic deformation. Their work is typical of many published investigations where the $1/f$ spectrum might have arisen from mobile carriers tunnelling into fixed traps in a surface layer. Where the current is confined at an interface—in the m.o.s. transistor—the cause of the relatively large fractional fluctuation in conductivity is amenable to investigation as well as speculation.

We believe that the conventional analysis of generation-recombination noise does not imply recognition of individual carriers. It is usual to set up the differential equation for the system as a whole, including a random source term (the Langevin equation). The source term represents the essential randomness of the trapping and emission of discrete carriers, and the noise spectrum is the response of the whole system to this stimulus. One usually finds that a small deviation from equilibrium decays exponentially, and Burgess⁷ has shown quite generally that the characteristic time-constant depends on the rates of both capture and emission of carriers. This time-constant is not, in general, equal to the lifetime of individual carriers.⁸

Professor Bell's final comment is ingenious, but we imagine that the number of specimens required to give, say, a 1% fluctuation in conductance in ten years is entirely prohibitive!

G. G. BLOODWORTH
R. J. HAWKINS

Department of Electronics,
University of Southampton.
9th September 1969

References

1. Bloodworth, G. G. and Hawkins, R. J., 'The physical basis of current noise', *The Radio and Electronic Engineer*, **38**, pp. 17–22, 1969.
2. Brophy, J. J. and Rostoker, N., 'Hall effect noise', *Phys. Rev.*, **100**, p. 754, 1955.
3. Brophy, J. J., 'Seebeck effect fluctuations in germanium', *Phys. Rev.* **111**, p. 1050, 1958.
4. Hooge, F. N. (Private communication.)
5. Bittel, H. and Scheidhauer, H., 'Zur Frage des Rauschens metallischer Leiter', *Z. Angew Phys.*, **8**, pp. 417–422, 1956.
6. Bergero, F., Mazzetti, P. and Montalenti, G., 'Current noise in Au thin films', *Nuovo Cimento*, **66B**, pp. 62–69, 1968.
7. Burgess, R. E., 'The statistics of charge carrier fluctuations in semiconductors', *Proc. Phys. Soc. B*, **69**, pp. 1020–1027, 1956.
8. Sah, C. T., 'Theory of low-frequency generation noise in junction-gate field-effect transistors', *Proc. Inst. Elect. Electronics Engrs*, **52**, pp. 795–814, 1964.

WIRED TELEVISION

I was interested to read Mr. J. S. Whyte's written reply to the verbal contribution which I made to the discussion of his paper. 'Telecommunication Services in the U.K.'† I was, however, most surprised to learn that the reason underlying the choice by the Post Office of a v.h.f. television distribution system was 'to give the user freedom to choose any make of receiver for connexion to the system'. The facts of the matter are the other way round, this freedom exists for h.f. but not for v.h.f. systems.

For example, one of the channels which the Post Office are proposing to use in Washington, County Durham, is channel D offset by $1\frac{1}{2}$ MHz. There are, however, many television receivers which will not tune as far as $1\frac{1}{2}$ MHz from a proper channel. Also next year a large number of receivers on sale to the public will be designed for operation on u.h.f. only and it is to be expected that many of the smaller types, particularly imported receivers, will be quite unsuitable for modification for v.h.f. operation. It would be interesting to know what plans the Post Office have for the provision of the necessary conversion equipment in such cases, and who will have to pay for it.

Further, the Post Office system denies the opportunity to the user to choose the cheapest and most reliable set of all, that is the receiver designed for h.f. systems which is available from several alternative sources. There are many tens of thousands of these receivers already in use in the general area of Washington, and people who move into the new town may already own one but are apparently to be denied the opportunity to connect it to the Post Office system.

With h.f. systems on the other hand, any make of television receiver, which is in working order for reception from an aerial, can be connected without modification to the systems by the use of an inverter, and in the case of systems provided under contract with the housing authority, as at Washington, the inverters are provided free of charge to the user.

R. P. GABRIEL

Rediffusion Limited,
Carlton House,
Lower Regent Street, London, S.W.1.
30th May 1969

Author's reply: I share Mr. Gabriel's surprise; there is no reference in his original contribution, to which I was replying, to h.f. systems, and he introduces the subject for the first time in this letter. My reference to freedom of choice to users was made in the context of frequency allocations on v.h.f. systems. If one has decided to provide a v.h.f. system, and the factors involved in this choice are many, then one can give a wide range of choice of receiver to the individual by choosing frequency allocations suitable for the majority of normal commercial receivers. Furthermore, this has the advantage that these receivers can be supplied and maintained through normal retail channels.

J. S. WHYTE

Post Office Telecommunications Headquarters,
Long Range Studies Division,
2–12 Gresham Street, London, E.C.2.
1st July 1969

† *The Radio and Electronic Engineer*, **37**, p. 296, May 1969.

Technical Performance Targets for a PAL Colour Television Broadcasting Chain

Based on papers prepared by the B.R.E.M.A. PAL Working Party†

PREFACE

Workers in the field of colour television broadcasting may find some of the contents of this paper interesting or provocative. The working party has been very conscious all along that many of the figures quoted were open to question either because of lack of solid information or because of differences in techniques and practices, and would therefore welcome comment and contribution. Comments should be sent to the Technical Secretary of the British Radio Equipment Manufacturers' Association.

INTRODUCTION

With the introduction of colour television on BBC-2 it was felt both by industry and by the broadcasters that it was desirable to try to define the technical performance of the various parts of the picture chain between the original scene and the viewer, so that the overall result should be of the best possible quality consistent with the economic factors involved. To this end a Working Party, with members from industry (B.R.E.M.A.), the I.T.A., and the B.B.C., was set up. This Working Party subsequently became Sub-Group XI-A of U.K. Study Group XI.

A considerable programme of work was carried out by the Working Party and this paper sets out the results of this work. It has been found possible to suggest tolerances for many of the parameters of the colour television picture as radiated from the transmitting aerial, even though adequate information for breaking down these tolerances along the transmission chain does not, in every case, exist. Further, performance targets for a colour television receiver will also be set out although, as shown below, on a rather different basis.

The tolerances suggested must be regarded as estimated figures and will undoubtedly be subject to revision as experience and information is accumulated.

The complete chain from the scene to the receiver aerial can be conveniently divided into three parts:

A. Colour television picture origination equipment prior to the colour coder.

B. The colour coder and the distribution and network equipment including transmitters.

C. The receiver.

The figures for the picture equipment are derived from practical measurements on the colour television cameras and telecine machines, etc., at present available, and are a considered judgment of what in fact is achieved on a day-to-day basis most of the time. The figures on the distribution network are derived from routine measurements made over a period of several months on the BBC-2 network between Central Apparatus Room at the B.B.C. Television Centre and several transmitters, including rebroadcast circuits to the more remote transmitters. The specification of the receiver is in terms of a notional reference colour receiver. It is not a design specification, but the performance quoted is representative of good commercial practice.

Only those parameters especially concerned with colour television have been included: it is assumed that characteristics applicable to black and white television have been dealt with according to current black and white television practice. The measurement methods are those of C.C.I.R. Study Group XI, Recommendation 451, or, where applicable, as near to them as possible. The methods of measurement as applied to the receiver are set out in Appendix 1.

The choice of the colour picture tube phosphors is discussed in Appendix 2.

† British Radio Equipment Manufacturers' Association, 49 Russell Square, London, W.C.1.

Part A. PICTURE ORIGINATION EQUIPMENT

A.1. Characteristics Applicable to All Colour Television Picture Originating Sources

A.1.1. Colour Analysis

The colour analysis characteristic of all picture sources will be chosen to be optimum for a typical set of currently used display phosphors and a receiver white point of Illuminant D (6500°K).

The C.I.E. co-ordinates of the display phosphors are assumed to be:

| | C.I.E. (1931) | | C.I.E. (U.C.S.) (1960) | |
|-------|---------------|----------|------------------------|----------|
| | <i>x</i> | <i>y</i> | <i>u</i> | <i>v</i> |
| Red | 0.64 | 0.33 | 0.451 | 0.349 |
| Green | 0.29 | 0.60 | 0.121 | 0.374 |
| Blue | 0.15 | 0.06 | 0.175 | 0.105 |

The above co-ordinates have been chosen as representing the phosphors likely to be used in shadow mask display tubes fitted to receivers manufactured in the U.K. The co-ordinates are different from those of the original N.T.S.C. specification since the latter are no longer of practical significance. However, no alteration of the signal coding has been made to take account of the revised phosphor characteristics used and hence the new co-ordinates are used only in the computation of the colour-analysis characteristics of picture origination equipment. In the event of a new phosphor becoming available and achieving wide popularity because of superior performance, it is expected that the analysis characteristics used in the picture originating sources would be changed to suit the new situation as opportunities arose. It is thought to be unlikely that any changes in display phosphors would be of such a magnitude that the use of new colour analysis characteristics in the picture originating equipment would place other receivers, not having the latest display phosphors, at any serious disadvantage. It is assumed that the RGB display tube gamma is 2.8 (±0.3).

A.1.2. Registration

The mis-registration of colour separation signals in any direction in the picture is expressed in terms of a percentage of the active line. During the major proportion of programme time, registration will not be worse than the following:

| | % active line |
|-----------|---------------|
| in Zone A | 0.1 |
| in Zone B | 0.15 |

Zones are defined as follows:

Zone A is an ellipse touching all four sides of the picture area.

Zone B is the remainder of the picture area.

A.1.3. Resolution

The following resolution capabilities are based upon presenting the picture source with perfect test patterns. During programmes this resolution may not be achieved, particularly with film reproduction where film quality is the limiting factor. All measurements and results are without gamma correction.

A.1.3.1. *Horizontal resolution.* The horizontal resolution of picture sources will be such that a black-and-white square wave grating pattern corresponding to 5.0 MHz (400 lines per picture height) is reproduced by the luminance signal, i.e. matrixed red, green and blue colour separation signals or a separate luminance reproducer, at not less than 50% of the amplitude of a low frequency grating (500 kHz or less).

A.1.3.2. *Vertical resolution.* The luminance signal response to a test pattern having a sharp horizontal transition between black and white will be substantially complete (i.e. the transition will exceed a change from 10% to 90% white signal amplitude) within 1/300 picture height.

A.1.4. Pulse Response

The response of each signal processing channel measured at the input of the coder with a linear contrast law will be: $K \leq 0.5\%$.

A.1.5. Line Streaking

Freedom from line streaking due to optical flare or imperfections in the electrical performance of the picture originating equipment shall be in accordance with current black-and-white television practice, with the additional proviso that the differences between the colour separation channels in this respect shall be negligible.

A.1.6. Low Frequency Response

The low frequency response of each channel shall be in accordance with current black-and-white television practice with the additional proviso that the differences between the three channels shall be negligible.

A.1.7. Signal Amplitude Hum

The level of amplitude hum in the signals shall be in accordance with current black-and-white television practice with the additional proviso that differences in hum between the signals in the three channels shall be negligible.

A.1.8. Positional Displacement due to Hum

Positional displacement of parts of the picture due to hum will be in accordance with current black-and-white television practice with the additional proviso

that differences in positional displacement between the colour separation channels shall be negligible.

A.2. Particular Characteristics of Camera Performance

A.2.1. Contrast Law Correction

Scene brightness to signal output will be related by a pure power law with an error not exceeding $\pm 10\%$ of the correct signal output over a scene contrast of at least 30 : 1. At present the value of gamma lies between 1/2.5 and 1/2.2. All these figures are under active review.

A.2.2. Reproduction of Grey Scale

A neutral grey scale will be reproduced with equal amplitude colour separation signals matched at white level to within 1%. Black level shall correspond to a scene illumination of 1/60th of maximum and at this level the matching of separation signals will be within 0.25% peak white level.

A.3. Particular Characteristics of Telecine Performance

A.3.1. Contrast Law Correction

The signal output will be related to the light transmitted by the film by a pure power law with an error not exceeding $\pm 10\%$ of the correct signal output over a scene contrast of 100 : 1. The gamma correction applied at present by telecine processing lies between 1/3.3 and 1/2.2. Colour films usually have a contrast characteristic with a maximum slope (maximum point-gamma) of between 1.5 and 2.0 and this is not corrected. All these figures are under active review.

A.3.2. Reproduction of Grey Scale

A neutral grey scale will be reproduced by a telecine machine with equal amplitude colour separation signals matched at white level to within 1%. Black level shall correspond to a film density of 2.3 ± 0.05 and at this level the matching of separation signals will be within 0.25% peak white level.

Part B. ESTIMATED TOLERANCES APPLICABLE TO THE RADIATED PAL COLOUR TELEVISION SIGNAL

B.1. Chrominance/Burst Phase

The chrominance signal phase errors with respect to mean phase of the bursts, independent of luminance signal magnitude and measured near to or at blanking level, shall be less than $\pm 10^\circ$.

B.2. Differential Phase

The reference phase of the subcarrier is taken as zero units on the first step (at blanking level) of the staircase waveform. Effects of head banding due to video tape recorders are excluded. The tolerance is $\pm 25^\circ$.

B.3. Differential Gain

The reference amplitude of the subcarrier is taken to be that of the first step (at blanking level) of the waveform. The tolerance is $\pm 25\%$.

B.4. Chrominance/Luminance Gain Inequality

The reference for this is the magnitude of the luminance signal and the tolerance applies to a test signal during the active line scan period, and therefore, does not include the colour burst. The tolerance is $+5\% - 30\%$.

B.5. Ratio Between Chrominance and Colour Burst

The error is the ratio between the amplitude of the chrominance signal and the amplitude of the colour burst should not exceed $\pm 10\%$.

B.6. Chrominance/Luminance Delay Inequality

The tolerance is ± 100 ns.

B.7. Pulse-to-Bar Ratio

The $2T$ pulse to bar ratio will fall between $+5$ and -20% .

B.8. K-Rating of 25 μ s Bar

The K -rating of the 25 μ s bar will not exceed 7%.

B.9. Chrominance/Luminance Crosstalk

This is the percentage increase or decrease of the luminance signal magnitude due to the presence of a chrominance signal excursion equal to that of the picture signal, and extending from blanking level to white level. The reference for calculating the above percentage is the peak-to-peak amplitude of the chrominance signal. The tolerance is $\pm 20\%$.

B.10. Line Time Non-linearity

The line time non-linearity will not exceed 25%.

Part C. A REFERENCE COLOUR RECEIVER

Measurements are made with the test signals applied to the antenna terminals and at such a level that noise does not influence measurements.

| Parameter | Value and Tolerance |
|---|---------------------|
| C.1. <i>Luminance Channel or E'_R, E'_G or E'_B channel</i> ^{(1, 13)†} | |
| (i) 2T pulse/bar ratio % | < 2 |
| (ii) 25 μ s bar K % | < 4 |
| (iii) 50 Hz bar K % | < 5 |
| C.2. <i>Line Time Non-linearity of Luminance Signal before Matrixing</i> ⁽¹⁾ | < 5% |
| C.3. <i>Line Time Non-linearity of Channels carrying Signals E'_R, E'_G or E'_B</i> ⁽¹⁾ | |
| First 'tread' of test signal | < 1% |
| Other 'treads' of test signal | < 3% |
| C.4. <i>Colour Difference Signal, $E'_R - E'_Y$, $E'_G - E'_Y$ or $E'_B - E'_Y$</i> ^(1, 2) | |
| (i) At 35% of its maximum value 10T pulse/bar ratio % | < 5 |
| (ii) At 75% of its maximum value 10T pulse/bar ratio % | < 25 |
| C.5. <i>Line Time Non-linearity of Colour Difference Signals $E'_R - E'_Y$, $E'_G - E'_Y$ or $E'_B - E'_Y$ having maximum amplitude</i> ^(3, 2, 1) | |
| First two 'treads' of test signal | < 5% |
| Other 'treads' of test signal | < 10% |
| C.6. <i>Matrixing Errors</i> ⁽⁴⁾ | |
| E'_R and E'_B at 50% amplitude 100% saturation | < 5% |
| E'_G at 50% amplitude 100% saturation | < 10% |
| E'_R and E'_B at 100% amplitude desaturated by 25% pedestal | < 10% |
| E'_G at 100% amplitude desaturated by 25% pedestal | < 15% |
| C.7. <i>D.C. Component, Loss and Gain relative to E'_Y max.</i> ⁽⁵⁾ | < 10% |
| C.8. <i>Errors of Tube Cut-off Conditions relative to E'_Y max.</i> ⁽⁶⁾ | < 1% |

† The numerals in brackets refer to footnotes.

(1) Measurement is based on C.C.I.R. Study Group XI, Rec. 451, as described in Appendix 1 using a modulated test signal applied to the antenna terminals.

(2) Measurement method is based upon C.C.I.R. Study Group XI, Rec. 451, as described in Appendix 1 but inputs to the coder of the test waveform are such as to give appropriate amplitude of both signs of the three colour difference signals.

(3) Due allowance must be made for the bandwidth of the colour difference channel.

(4) Measurement compares the amplitude of the colour bars with the white bar and the black or grey bars with picture black or grey. The error is the largest deviation expressed

as a percentage of E'_Y max.

(5) Measurement by applying sync and blanking and sync plus 95% E'_Y max. The error is the change in voltage corresponding to black level at the picture tube electrodes expressed as a percentage of E'_Y max.

(6) Each channel, E'_B , E'_G and E'_R is checked at the display tube electrodes for change in grid/cathode voltage corresponding to black level when alternately black level and 90% E'_B max, E'_G max or E'_R max is applied to the encoder input. The three channels are also checked for shift in grid/cathode potential when the test signal is applied to both of the channels not under measurement. The error is the shift of grid/cathode voltage corresponding to black level expressed as a percentage of E'_Y max.

PERFORMANCE TARGETS FOR COLOUR TELEVISION

| Parameter | Value and Tolerance |
|--|--|
| C.9. 'Venetian Blind' Effects with Errors of 30° between Mean Burst Phase and Carrier Chrominance Signal ⁽⁷⁾ | |
| (i) Large areas relative to E'_x max. | < 3% |
| (ii) Colour transitions, measured on a red/cyan transition relative to E'_x max. | < 7% |
| C.10. Demodulation Angle Errors ⁽⁸⁾ | |
| The output of the colour difference demodulators expressed as % of the normal condition signal output, when the same signal is applied in quadrature with respect to the normal condition shall be | < 17 |
| C.11. Auto Chroma Gain (if incorporated) | |
| The change in output of the colour difference demodulators expressed as % of the output under nominal conditions, when the luminance/chrominance including burst ratio is changed by 12 dB shall be | < 10 |
| C.12. Colour Sync. ⁽⁹⁾ | |
| (i) The colour sync and V channel switch sync shall operate without error under signal/noise ratio conditions of 16 dB signal to r.m.s. noise. | |
| (ii) The V channel switch shall be correctly synchronized within a period of 10 lines after a change in the input signal to the alternative burst sequence relative to line scanning sync. | |
| C.13. Chrominance to Luminance Delay Inequality ⁽¹⁰⁾ | 60 ns |
| C.14. Colour Killing | |
| (i) The colour killer circuit shall operate on a transmission without burst present, such that no cross-colour effects are visible on Test Card C, either in the frequency gratings or on transients. This performance is to be maintained in the presence of enhancement of the frequencies in the region of the colour subcarrier relative to low video frequencies by up to 6 dB. | |
| (ii) The colour killer shall enable the chrominance channel down to the point where the sub-carrier is reduced by at least 24 dB relative to the nominal signal specification. It is assumed that the killer circuit fully enables or disables the chrominance channel. | |
| (iii) The colour killer shall both kill and enable the chrominance channel within a period of 10 complete fields following the application or removal of the burst. | |
| (iv) The colour killer shall operate as above at signal input levels to the point where the signal/r.m.s. noise ratio is 16 dB. | |
| (7) The mean phase of the burst is shifted at the coder by $\pm 30^\circ$ relative to the chrominance signal. The signals $E'_R - E'_Y$, $E'_B - E'_Y$ and $E'_G - E'_Y$ are compared on time sequential lines and the difference between them expressed as an equivalent percentage of E'_Y max. | (8) A convenient method of measurement is to cause the U signal to alternate rather than the V signal whilst leaving the burst alternation unaffected. This will automatically provide correct conditions for measurement. |
| Large areas infer a picture component such that neither sideband of the chrominance signal is attenuated relative to the nominal carrier frequency. Colour transition 'Venetian Blind' effects are measured by taking the point on the transition where the effect is greatest. | (9) Test is by viewing the display tube when reproducing a pattern with significant contents in the V axis. |
| | (10) Assessment is by viewing the display tube when reproducing a 600 ns duration red bar on a white background. A variable delay is used in the coder to measure the magnitude of the error. |

| Parameter | Value and Tolerance |
|-----------|---------------------|
|-----------|---------------------|

C.15. *Tuning and Propagation Effects*

- (i) The receiver will either employ a push button tuner with adequate re-set and stability performance or a continuously adjustable tuner with satisfactory indication of correct tuning.
- (ii) The response shall be such that where propagation conditions enhance the chrominance signal relative to low video frequencies by up to 6 dB, the requirements of Appendix 1 are not significantly deteriorated.

C.16. *The Tolerance on Mains Voltage* for the specification shall be ± 6%

C.17. *White Point*^(11, 14)

| | | |
|---|----------|----------|
| The receiver will be set to a white point of | D 6500°K | ± 500°K |
| The chromaticity co-ordinates of the nominal white point are: | <i>x</i> | <i>y</i> |
| | 0.313 | 0.329 |
| | <i>u</i> | <i>v</i> |
| | 0.198 | 0.312 |

C.18. *Display Phosphors*^(12, 14)

| | | | |
|---|-------|----------|----------|
| The C.I.E. co-ordinates of the nominal display phosphors are: | | <i>x</i> | <i>y</i> |
| | Red | 0.64 | 0.33 |
| | Green | 0.29 | 0.60 |
| | Blue | 0.15 | 0.06 |
| | | <i>u</i> | <i>v</i> |
| | Red | 0.451 | 0.349 |
| | Green | 0.121 | 0.374 |
| | Blue | 0.175 | 0.105 |

The tolerances of the phosphor chromaticities of a tube having nominal display phosphors are defined by the quadrilateral given by joining the chromaticity co-ordinates listed below by straight lines:⁽¹⁵⁾

| | | | | | |
|----------|-------|-------|-------|-------|--------|
| | | | RED | | |
| <i>x</i> | 0.643 | 0.628 | | 0.640 | 0.654 |
| <i>y</i> | 0.343 | 0.330 | | 0.319 | 0.332 |
| | | | GREEN | | |
| <i>x</i> | 0.280 | 0.298 | | 0.302 | 0.285 |
| <i>y</i> | 0.610 | 0.634 | | 0.586 | 0.564 |
| | | | BLUE | | |
| <i>x</i> | 0.136 | 0.151 | | 0.155 | 0.148 |
| <i>y</i> | 0.061 | 0.066 | | 0.058 | 0.0542 |
| | | | RED | | |
| <i>u</i> | 0.441 | 0.441 | | 0.461 | 0.461 |
| <i>v</i> | 0.353 | 0.347 | | 0.345 | 0.351 |
| | | | GREEN | | |
| <i>u</i> | 0.115 | 0.119 | | 0.128 | 0.124 |
| <i>v</i> | 0.375 | 0.380 | | 0.373 | 0.368 |
| | | | BLUE | | |
| <i>u</i> | 0.157 | 0.174 | | 0.183 | 0.176 |
| <i>v</i> | 0.106 | 0.113 | | 0.103 | 0.097 |

| Parameter | Value and Tolerance |
|--|---------------------|
| C.19. Purity | |
| The assessment is to be carried out at a viewing distance of 1.5 m and an ambient light level at the tube face of approx. 1 lux and of a colour temperature approx. to Illuminant D with an e.h.t. of 25 kV and a total beam current of 500 μ A. The individual colour rasters should show no colour impurity and when viewing the white raster no significant colour difference should be directly visible and no clear indication of discoloured zones should be seen. | |
| C.20. Mis-registration | |
| The mis-registration of the colour separation signals in any direction is expressed in terms of the active line. | |
| Zone A: an ellipse touching all four sides of the picture area: | |
| red to green | < 0.3% |
| red/green to blue | < 0.4% |
| Zone B: remaining picture areas: | |
| red to green | < 0.6% |
| red/green to blue | < 0.7% |
| C.21. Gamma | |
| The overall receiver gamma is | 2.8 \pm 0.3 |

- | | |
|---|--|
| <p>(11) The white point is measured at a brightness level of 120 apostilbs (11.2 ftL) for a 25-in (63 cm) tube and 73 apostilbs (6.8 ftL) for a 19-in (48 cm) tube. (1 apostilb = 0.3183 candela per square metre.)</p> <p>(12) The above figures for gamma, phosphor co-ordinates are based on current practice and are subject to revision. Such guidance is needed by Broadcasting Authorities, in order that the best overall system performance is achieved.</p> | <p>(13) Connections to the display tube for measurement purposes shall not impose a capacitance load of greater than 4 pF or a shunt resistance of less than 5 MΩ.</p> <p>(14) x and y are C.I.E. co-ordinates 1931. u and v are C.I.E. co-ordinates 1960.</p> <p>(15) The tabulated phosphor chromaticity tolerances were computed for a small range of flesh-tones and are not the manufacturing tolerances of a tube.</p> |
|---|--|

Appendix 1

RECEIVER MEASUREMENT METHODS FOR TELEVISION

Introduction

It is well known that analysis of television equipment performance using impulse response techniques results in a better appreciation of the reproduced picture quality than does d.c. response measurements. Such measurement methods have been applied to studio equipment and networks for a considerable time and the details of these techniques have been internationally agreed in C.C.I.R., Recommendation 451. These measurement methods can be advantageously applied to television receivers but, since Recommendation 451 is orientated towards network performance, certain modifications and additions are necessary. In particular, it is necessary that the measurements are

carried out overall, hence the test signals must be modulated upon r.f. carriers in the normal television bands and channels. The use of these methods of measurement gives unity of procedure for the whole television chain, from picture origination equipment to the reproduced image.

The first part of this Appendix describes measurements of the following parameters:

1. Insertion Gain

This measurement defines the gain of a network or device.

2. Continuous Random Noise

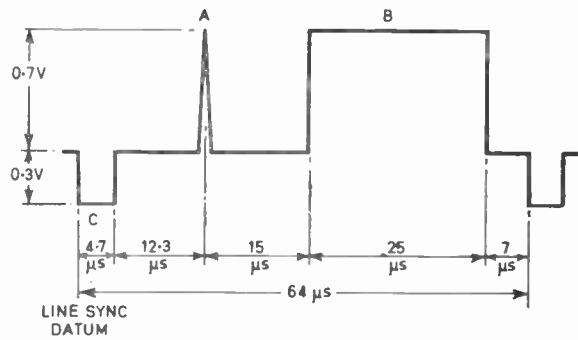


Fig. 1. Pulse-and-bar test signals.†

- A *T* pulse, 2*T* pulse or 10*T* pulse
- B *T* bar, 2*T* bar or 10*T* bar
- C Sync pulse omitted when test signals are applied to a colour coder.

3. Periodic Noise

Examples of this type of noise are sinusoidal interference and power supply ripple.

4. Impulsive Noise

This includes, for example, ignition interference.

5. Non-Linearity Distortion of the Picture Signal

This category includes distortion of the amplitude characteristic.

6. Non-Linearity Distortion of Synchronizing Signal

This assesses the ratio of synchronizing to picture signals.

7. Linear Waveform Distortion

Measurements are given for linear waveform distortions in luminance and chrominance channels.

8. Luminance and Chrominance Inequalities

Inequalities of gain and delay are treated. Measurement methods are provided for assessing delay inequality, both oscillographically and on the reproduced picture tube screen.

9. Matrixing Distortion

Measurements methods are provided for assessing the accuracy with which the colour separation signals are generated from luminance and chrominance components.

10. D.C. Component Distortion

Errors in the reproduction of the d.c. component in both luminance and colour difference channels is measured.

†For the design of the shaping network, see McDiarmid, I.F. and Phillips, J., "A pulse and bar waveform generator for testing television links", *Proc. I.E.E.*, 105B, p. 440, 1958.

Test methods are given in Weaver, L.E., "Sine-squared Pulse and Bar Testing in Colour Television", B.B.C. Engineering Monograph No. 58, August 1965.

A procedure is provided for checking the accurate maintenance of picture tube cut-off under varying picture content conditions.

11. Effects of Chrominance Signal Phase Errors

12. Chrominance Signal Demodulation Angle Errors

13. Colour Sync. Performance

The second part of the Appendix describes measurements of linear waveform distortion, namely on the luminance channel and with chrominance channel before and after demodulation. The third part of the Appendix briefly notes the characteristics of the signals.

Where parameters are subject to variations, short, medium and long period variations are specified as 1 second, 1 hour and 10 hours respectively.

Part 1. Measurement of Chrominance Performance

1. Insertion Gain

Insertion gain is measured using the bar only of the 2*T* pulse and bar test signal shown in Fig. 1. It is the ratio expressed in dB of output to input voltage. (Ref.: Rec. 451-4.1 and 4.2.) When applied to a receiver having a r.f. input and a video signal output, the modulated version of the test signal is used. The insertion gain is the ratio expressed in dB of peak white output voltage to input sync tip carrier r.m.s. voltage.

2. Continuous Random Noise

The signal/weighted-noise ratio for continuous random noise is defined as the ratio expressed in dB, of the nominal peak-to-peak amplitude of the picture luminance signal to the r.m.s. amplitude of the noise measured under the following conditions:

the noise is passed through a specified weighting network, or equivalent;

the measurement is made with an instrument having, in terms of power, an effective time-constant or integrating time of 1 s.

2.1. *Luminance channel*: The nominal frequency range is 10 kHz to 5 MHz. The lower limit is determined by the high-pass member of the junction filter shown in Fig. 2; its purpose is to exclude power supply hum and microphony noise. The upper limit is determined by the low-pass filter shown in Fig. 3. The weighting network is shown in Fig. 4; it has a time constant of 200 ns giving a weighting effect of 6.5 dB for flat random noise and 12.3 dB for triangular random noise.

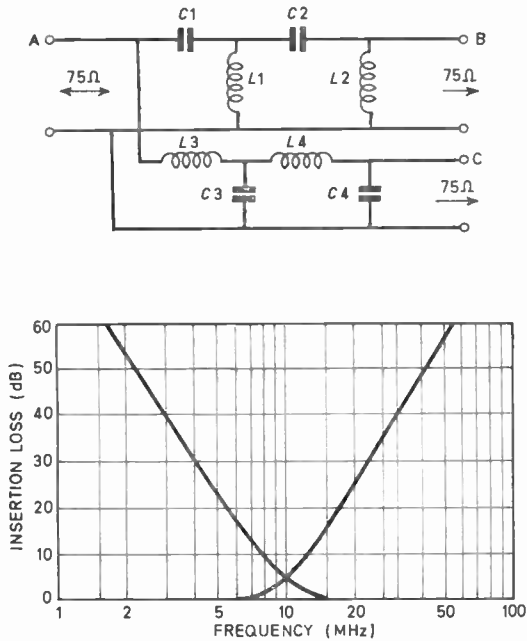


Fig. 2. Junction filter for noise measurements.

- A Input
- B High-pass output
- C Low-pass output

| Table of Values | | |
|-----------------|------------|-----------|
| Component | Value | Tolerance |
| C1 | 139 000 pF | |
| C2 | 196 000 pF | ±5% |
| C3 | 335 000 pF | |
| C4 | 81 200 pF | |
| L1 | 0.757 mH | |
| L2 | 3.12 mH | ±2% |
| L3 | 1.83 mH | |
| L4 | 1.29 mH | |

Note: The *Q*-factor of each inductor should be equal to or greater than, 100 at 10 kHz.

2.2. *Chrominance channel*: The nominal frequency range is 3.5 to 5.5 MHz, determined by the combined bandpass filter and weighting network shown in Fig. 5. For each subcarrier sideband, the filter provides a weighting effect which is approximately equal to that of the luminance weighting network in the 0 to 1 MHz band. (Ref.: Rec. 451-4.3.)

3. Periodic Noise

The signal/noise ratio for periodic noise is defined as the ratio, expressed in decibels, of the nominal peak-to-peak amplitude of the picture luminance signal to the peak-to-peak amplitude of the noise. (Ref.: Rec. 451-4.4.)

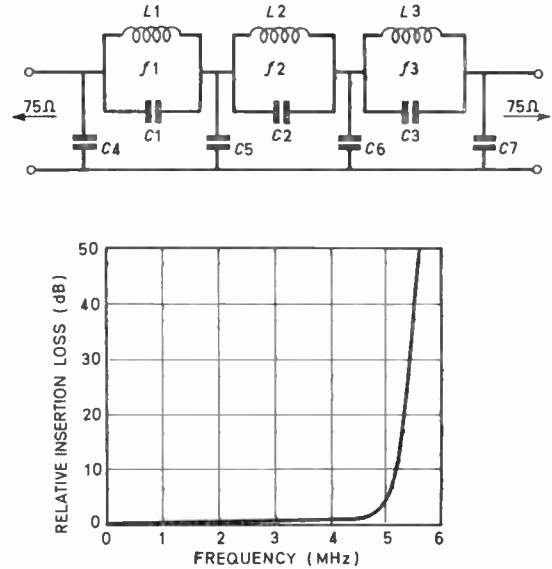


Fig. 3. Low-pass filter for measurements of random noise in luminance channel.

Table of Values

| Component | Value | Tolerance | |
|-----------------------|-----------|---|--|
| C1 | 100 pF | | |
| C2 | 545 pF | | |
| C3 | 390 pF | Each capacitance quoted is the total value, including all relevant stray capacitances, and should be correct to ±2% | |
| C4 | 428 pF | | |
| C5 | 563 pF | | |
| C6 | 463 pF | | |
| C7 | 259 pF | | |
| L1 | 2.88 μH | | Each inductor should be adjusted to make the insertion loss a maximum at the appropriate indicated frequency |
| L2 | 1.54 μH | | |
| L3 | 1.72 μH | | |
| <i>f</i> ₁ | 9.408 MHz | | |
| <i>f</i> ₂ | 5.506 MHz | | |
| <i>f</i> ₃ | 6.145 MHz | | |

The *Q*-factor of each inductor measured at 5 MHz should be between 80 and 125.

4. Impulsive Noise

The signal/noise ratio for impulsive noise is defined as the ratio, expressed in decibels, of the nominal peak-to-peak amplitude of the picture luminance signal to the peak-to-peak amplitude of the noise. (Ref.: Rec. 451-4.5.)

5. Non-linearity Distortion, Picture Signal

5.1. *Luminance channel and chrominance channel before chrominance demodulation*: Line-time non-linearity distortions in the luminance channel and chrominance channel before chrominance demodulation are measured with the test signal shown in Fig. 6, consisting of a 5-riser staircase, with superimposed

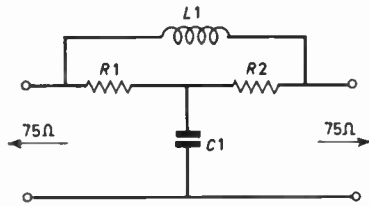


Fig. 4. Weighting network for measurement of random noise in the luminance channel.

Table of Values

| Component | Value | Tolerance |
|-----------|---------|-----------|
| C1 | 2660 pF | |
| L1 | 15 μH | ±1% |
| R1 | 75 Ω | |
| R2 | 75 Ω | |

Note 1: The Q-factor of inductor L1 should be equal to, or greater than, 25 at 8 MHz.

Note 2: Insertion loss = $10 \log_{10} 1 \times (2\pi\tau f)^2$ dB
where $\tau = 200$ ns.

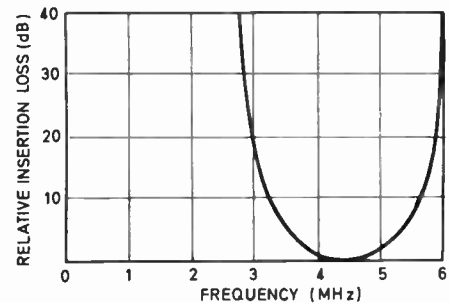
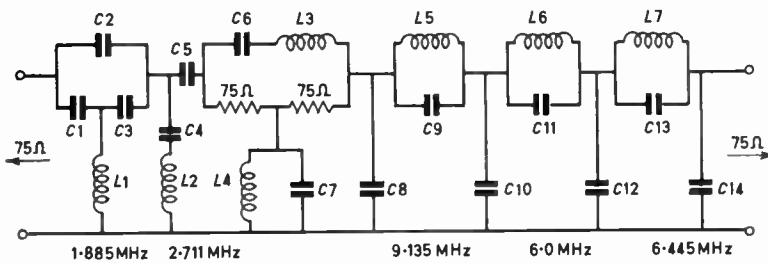
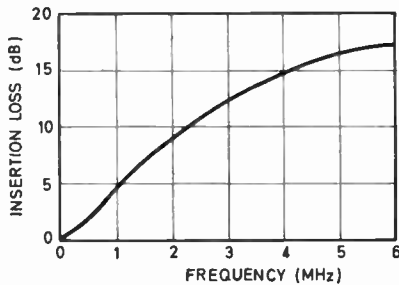


Fig. 5. Band-pass filter and weighting network for measurement of random noise in the chrominance channel.

Table of Values

| Component | Value | Tolerance |
|-----------|-----------|--|
| C1 | 496.0 pF | |
| C2 | 89.47 pF | |
| C3 | 292.1 pF | |
| C4 | 715.8 pF | |
| C5 | 1239.0 pF | |
| C6 | 194.3 pF | |
| C7 | 1182 pF | |
| C8 | 385.7 pF | ±1% |
| C9 | 141.3 pF | |
| C10 | 418.6 pF | |
| C11 | 941.2 pF | |
| C12 | 311.4 pF | |
| C13 | 619.2 pF | |
| C14 | 187.5 pF | |
| L1 | 2.960 μH | |
| L2 | 4.814 μH | |
| L3 | 6.650 μH | L3 is adjusted to resonate with C6, and L4 with C7, at 4.428 MHz. |
| L4 | 1.093 μH | |
| L5 | 2.149 μH | L1, L2, L5, L6 and L7 are adjusted to give maximum insertion loss at the appropriate indicated frequencies |
| L6 | 0.7476 μH | |
| L7 | 0.9846 μH | |

Note 1: The Q-factor of each inductor should be equal to, or greater than, 100 between 3 MHz and 6 MHz.

Note 2: The insertion loss is equal to, or greater than, 35 dB at frequencies above 6 MHz.

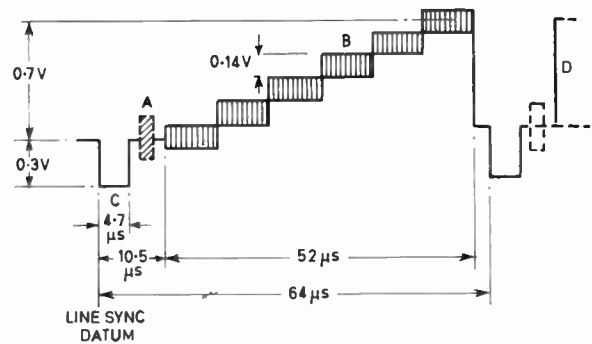


Fig. 6. Staircase test signal.

- A Optional colour burst
- B Superimposed chrominance subcarrier
- C 3 lines at black level or 3 lines at white level
- D Synchronizing pulses

Note: A, B and D are omitted when this test signal is applied to a colour coder.

sub-carrier, in every fourth line. Separate measurements are made with three intermediate lines at black level and white level, and the higher value of distortion is taken as the result.

5.1.1. *Luminance channel:* At the point of measurement, the test signal is passed through a filter whose effect is to eliminate the sub-carrier. The numerical value of the distortion is found by expressing the change in amplitude of the centre point of each 'tread' from its nominal amplitude as a percentage of the amplitude of white level. Where values are not specified for particular 'treads', the largest change in amplitude is taken as the value of the distortion.

5.1.2. *Chrominance channel before chrominance demodulation:* At the point of measurement, the sub-carrier is filtered from the rest of the test signal and its six sections are compared in amplitude and phase. Taking the blanking-level section of sub-carrier as reference, the differential gain is defined as the largest departure from the reference amplitude, expressed as a percentage, and the differential phase is defined as the largest departure from the reference phase-angle, expressed in degrees. (Ref.: Rec. 451-4.6.)

5.2. *Colour separation signal channels and colour difference signal channels:* Line-time non-linearity distortions in the colour separation signal channels and the chrominance channel after demodulation are measured with the test signal shown in Fig. 3 but without superimposed sub-carrier. The test signal is applied to the colour separation signal inputs of a colour coder with the amplitudes as shown in Fig. 7.

5.2.1. *Colour separation signal channels:* Separate measurements are made of the colour separation signal outputs, E'_R , E'_G and E'_B . The method is that of Section 5.1.1.

5.2.2. *Chrominance channel after chrominance demodulation (colour difference signals).* The three colour difference channels: $E'_R - E'_Y$, $E'_G - E'_Y$ and $E'_B - E'_Y$ are measured for positive and negative going signals using the method of Section 5.1.1.

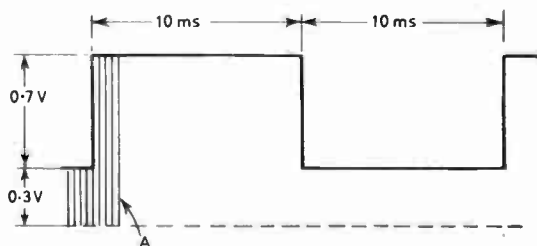


Fig. 8. 50 Hz square wave test signal.

A Sync pulses omitted when applied to a colour coder

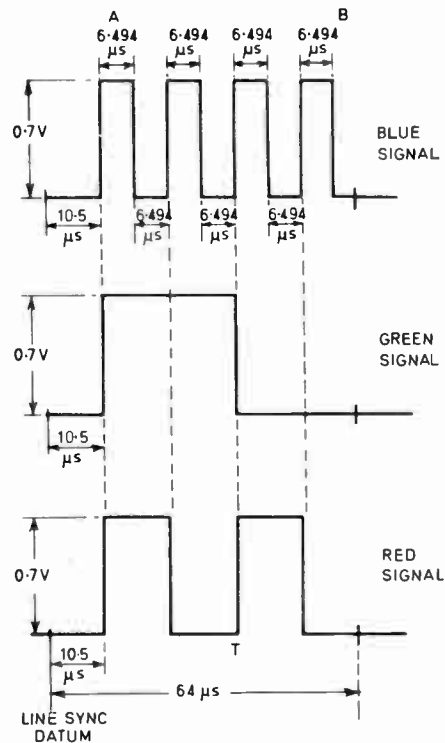


Fig. 7. Three signals to be matched in time within 1 ns.

6. Non-linearity Distortion of Synchronizing Signal

The distortion is expressed in terms of percentage departure of the mid-point amplitude of the line synchronizing pulse from its nominal amplitude, i.e. 0.3 V where blanking level to peak white has an amplitude of 0.7 V. (Ref.: Rec. 451-4.7.)

7. Linear Waveform Distortion

7.1. *Luminance channel:* The short-time, line-time and field-time linear distortions in the luminance channel are found from the waveform responses to the pulse-and-bar and 50 Hz square-wave test signals shown in Fig. 1 and Fig. 8. The result is expressed as a rating factor K by the method described in Part 2.

7.2. *Chrominance channel before chrominance demodulation:* The short-time and line-time linear distortions in the chrominance channel before chrominance demodulation are found from the waveform responses to the pulse-and-bar modulated sub-carrier test signal shown in Fig. 9. The result may be expressed by a rating factor analogous to that of the luminance channel. (Ref.: Rec. 451-4.8.)

7.3. *Chrominance channel after chrominance demodulation:* Short-time, line-time and field-time linear distortions in the chrominance channel after demodulation are found from the waveform responses

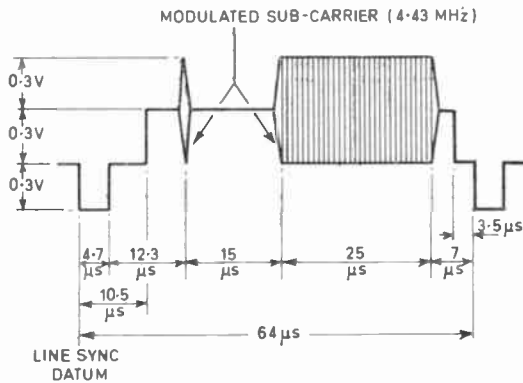


Fig. 9. 10T pulse-and-bar test signals.†

obtained by applying the 10T test signal shown in Fig. 1 or 50 Hz square wave signal shown in Fig. 8 to the colour separation signal inputs of the colour coder. The proportions used are shown in Table 1. The result may be expressed by a rating factor analogous to that of the luminance channel.

8. Luminance/Chrominance Inequalities

8.1. *Gain inequality:* The gain inequality, expressed as the percentage departure of the amplitude of the chrominance element from the amplitude of the luminance element, both measured at the mid-point of the bar.

8.2. *Delay inequality:* The delay inequality is expressed in ns and measured by applying the signals of Fig. 10 to the colour separation signal inputs of a colour coder. The coder is provided with an adjustable calibrated

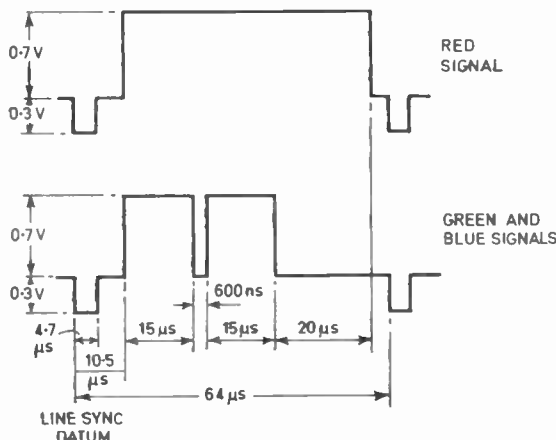


Fig. 10. Signal for luminance/chrominance delay measurements.

Three signals to be matched in time within 1 ns.

† For design of the shaping network see footnote to Fig. 1.

Table 1

Signal proportions for linear waveform distortions, chrominance after demodulation

| Test conditions | Coder Inputs | | |
|-----------------|--------------|---|---|
| | R | G | B |
| 1 | S | O | O |
| 2 | O | S | O |
| 3 | O | O | S |
| 4 | S | S | O |
| 5 | S | O | S |
| 6 | O | S | S |

S = signal of Figs. 1 and 8. O = no signal.

delay in the luminance channel that is varied to cancel any delay inequality, thus indicating the error.

The picture on the receiver under test may be used as an indicator or the colour separation signals may be displayed oscillographically. (Ref.: Rec. 451-4.9.)

9. Matrixing Distortions

9.1. *E_G' - E_V' signal where E_R', E_G' and E_B' are not available:* The accuracy with which E_G' - E_V' is generated from demodulated E_R' - E_V' and E_B' - E_V' is measured at the display tube electrodes and assessed using the waveform of Fig. 11 applied to the coder. The distortion is expressed by the amplitude of the

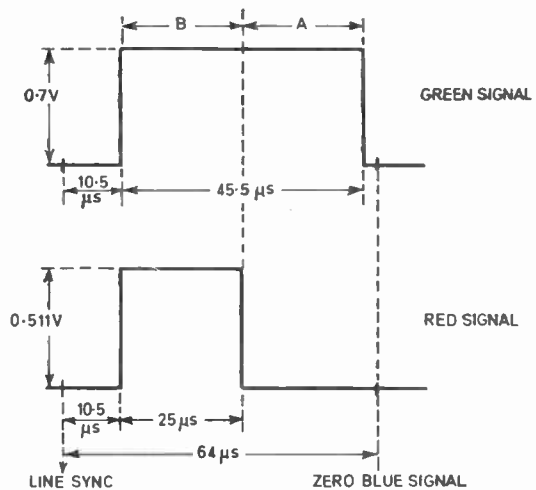


Fig. 11. Signals for matrixing distortions of E_G' - E_V'.

waveform during period B, as a percentage of the waveform amplitude during period A.

9.2. *Colour separation signals E_R', E_G' and E_B':* The colour separation signals of Fig. 7 are applied to the colour coder. The measurement at the display tube electrodes compares the amplitude of the positive signal excursions with amplitude of the waveform

during period *A* and the negative excursions with the amplitude of the waveform during period *B*.

The distortion is taken to be the largest of these two measurements of difference in amplitude and is expressed as a percentage of the signal during period *A* (E'_Y max.).

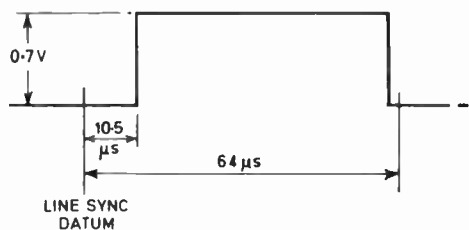
10. D.C. Component Distortion

10.1. Luminance signal E'_Y : The voltage at the display tube electrodes corresponding to black level is measured when applying alternately, 'sync and blanking' and 'sync plus E'_Y max.'

The error is the change in black level voltage for the two signal conditions expressed as a percentage of E'_Y max.

10.2. Colour difference signals $E'_R - E'_Y$, $E'_G - E'_Y$ and $E'_B - E'_Y$: A pair of measurements are provided for each colour difference channel corresponding to the positive and negative excursions of the signals.

The six test conditions are defined by applying the signals shown in Fig. 12 to the three colour separation signal inputs of a colour coder. The error is the change in voltage corresponding to black level at the picture tube electrodes for the two conditions expressed as a percentage of E'_Y max.



| Test Condition | Coder Inputs | | |
|----------------|--------------|---|---|
| | R | G | B |
| 1 | S | O | O |
| 2 | O | S | O |
| 3 | O | O | S |
| 4 | S | S | O |
| 5 | S | O | S |
| 6 | O | S | S |

S = signal of Fig. 12. O = no signal.

Fig. 12. Colour difference signal d.c. component distortion.

10.3. Colour separation channels E'_R , E'_G and E'_B : Two measurements are provided for each of the three colour separation channels. The two conditions are:

- (a) when an input is provided to the colour coder corresponding to the channel under measurement;
- (b) when an input is provided to the colour coder corresponding to the two channels other than that under measurement.

The signals are black level, 75% E'_R max., 75% E'_G max. and 75% E'_B max. The error is the shift of picture tube electrode voltage corresponding to black level that occurs when changing from sync and blanking only to both of the two other test conditions. The errors are expressed as a percentage of E'_Y max.

11. Distortion due to Chrominance Signal Phase Errors

The colour separation signals of Fig. 7 are applied to the colour coder and the mean phase of the burst at the coder is shifted by 30° relative to the chrominance signal. The signals on the display tube electrodes either $E'_R - E'_Y$, $E'_G - E'_Y$ and $E'_B - E'_Y$ or E'_R , E'_G and E'_B are compared on time sequential lines. The distortion is expressed as the maximum difference between equivalent parts of the signal expressed as a percentage of E'_Y max.

11.1. Measurements of large areas infer signal components such that equal energy is contributed by both sidebands of the chrominance signal. This may be allowed for using signal components having a duration of, at least, 3 μs.

11.2. Measurements of small areas are effected by examination of the signal at point T of Fig. 7, corresponding to a green/magenta transition.

Measurement is by taking a point on the transition where the difference between time sequential lines is the greatest.

12. Errors of Chrominance Signal Demodulation Angles

Separate measurements are made for the channels carrying $E'_R - E'_Y$ and $E'_B - E'_Y$.

The output from the colour difference demodulators is measured first with the colour separation signals of Fig. 7 applied normally to the colour coder and second with the same signals applied to the colour coder but with the mean phase of the reference burst shifted by 90° relative to the chrominance channel. The error is the magnitude of the signal under the second condition relative to that under the first expressed as a percentage.

13. Signal/Noise Performance of Colour Synchronization

Colour separation signals corresponding to Fig. 7 are applied to the colour coder. A noise generator is connected to the receiver together with the test signal.

The criterion of performance is that signal to noise ratio at which statistical failure of *V* channel synchronization is just observable, either by viewing the display tube or observing the $E'_R - E'_Y$ or E'_R signals oscillographically.

Part 2

Linear Waveform Distortion Measurements

1. Linear Waveform-distortion, Luminance Channel

The linear transmission performance of the luminance channel is specified in terms of a rating factor K . The precision of this method is adequate but its limitations must be borne in mind since the spectrum of one of the test signals unavoidably extends beyond the nominal limit of the video pass-band, 5.5 MHz.

1.1. Test method: To meet a specified rating factor K , the response to the pulse-and-bar and 50 Hz square-wave test signals shown in Figs. 1 and 8 should fall within the following limits:

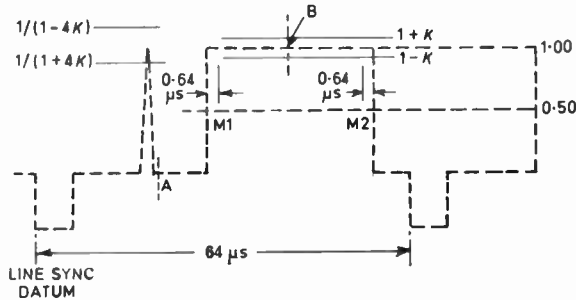


Fig. 13. T , $2T$ and $10T$ response and pulse/bar ratio.

1.1.1. $2T$ bar response: The limits are indicated by the oscilloscope mask shown in Fig. 13. The response should fall within the $\pm K$ limits indicated by the full lines, which extend to $H/100$ from the half-amplitude point of the transition. (H is the duration of one line, 64 μ s.)

1.1.2. $2T$ pulse response: The limits are indicated by the oscilloscope mask shown in Fig. 14. In effect, the oscilloscope is to be adjusted so that:

the sweep velocity corresponds with the time scale indicated;

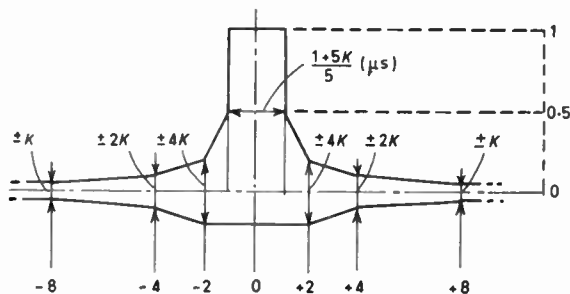


Fig. 14. T , $2T$ and $10T$ pulse response unit intervals 50 ns, 100 ns and 1 μ s respectively.

the 'black' level of the response coincides with the horizontal axis;

the peak of the response falls on the unit-amplitude line;

the half-amplitude points of the response are symmetrically disposed about the vertical axis.

1.1.3. $2T$ pulse/bar ratio: The ratio of the amplitude of the $2T$ pulse response to the amplitude of the $2T$ bar response should fall within the limits $1/(1 \pm 4K)$. The limits are included in the mask shown in Fig. 13.

1.1.4. T pulse response: The rating cannot be specified as rigidly as that for the $2T$ pulse because the spectrum of the T pulse response should not show appreciable ringing at a frequency below 5.0 MHz, irrespective of other aspects. A practical solution to a T pulse rating is found by the insertion, at the output of the pulse generator, of a low pass-filter having a good waveform response. This filter has an insertion loss that is almost constant, up to 5 MHz and then increases by about 3 dB at 5.3 MHz and 20 dB at 5.7 MHz.

The pulse-to-bar ratio of the generator plus filter is 84% and it has been found, in practice, that to meet a specified rating factor K , the T pulse/bar ratio should be between 79% and 89%, as shown in Fig. 13 and Fig. 14.

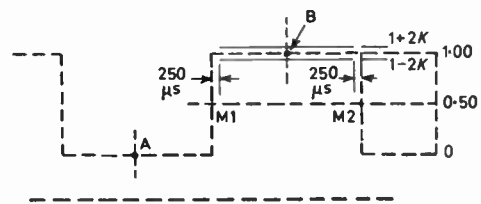


Fig. 15. 50 Hz square wave response.

1.1.5. 50 Hz square-wave response: A suitable oscilloscope mask is shown in Fig. 15. As for the $2T$ bar response, the oscilloscope is to be adjusted so that the waveform passes through the four marked points, the line synchronizing pulses being ignored.

2. Linear Waveform Distortion, Chrominance Channel before Chrominance Demodulation

Linear transmission performance of the chrominance channel before chrominance demodulation is specified in terms of a rating factor K .

2.1. Test method: To meet a specified rating factor K , the response to the pulse-and-bar test signals shown in Fig. 9 should fall within the following limits:

2.1.1. *10T pulse response*: The limits are indicated by the oscilloscope mask shown in Fig. 13. In effect, the oscilloscope is to be adjusted so that:

the sweep velocity corresponds with the time scale indicated;

the 'black' level of the response coincides with the horizontal axis;

the peak of the response falls on the unit-amplitude line;

the half-amplitude points of the response are symmetrically disposed about the vertical axis.

2.1.2. *10T pulse/bar ratio*: The ratio of the amplitude of the 10T pulse response to the amplitude of the 10T bar response should fall within the limits $1/(1 \pm 4K)$. The limits are included in the mask shown in Fig. 13.

3. Linear Waveform Distortion, Chrominance Channel after Chrominance Demodulation

3.1. *Test method*: To meet a specified rating factor K , the response to the pulse-and-bar test signals shown in Fig. 1 and applied in the manner described in Part 1, Section 7.3 and referring to Fig. 13, should fall within the following limits. (For certain measurements the oscilloscope waveform must be inverted.)

3.1.1. *10T bar response*: The limits are indicated by the oscilloscope mask shown in Fig. 13. The response should fall within the $\pm K$ limits indicated by the full lines which extend to $H/100$ from the half-amplitude point of transition.

3.1.2. *10T pulse response*: The limits are indicated by the oscilloscope mask shown in Fig. 14. In effect, the oscilloscope is to be adjusted so that:

the sweep velocity corresponds with the time scale indicated;

the 'black' level of the response coincides with the horizontal axis;

the peak of the response falls on the unit-amplitude line;

the half-amplitude points of the response are symmetrically disposed about the vertical axis.

3.1.3. *10T pulse/bar ratio*: The ratio of the amplitude of the 10T pulse response to the amplitude of the 10T bar response should fall within the limits $1/(1 \pm 4K)$. The limits are included in the mask shown in Fig. 13.

3.1.4. *50 Hz square wave response*: A suitable oscilloscope mask is shown in Fig. 15 and, as for the 2T bar response, the oscilloscope is adjusted so that the waveform passes through the four marked points. The line frequency signal components are ignored.

Part 3

Characteristics of the Signals

The signals used in carrying out the above measurements will conform with Tables 1, 2, 3, 3b and 4 of Report 308 and Part D of Report 407, all of the C.C.I.R. 11th Plenary Assembly, Oslo, 1966.

Where r.f. signals are used, the sound carrier will be present and will be fully modulated with an audio frequency sine wave signal in the region of 400 Hz.

The r.f. signal will correspond to one of the v.h.f. or u.h.f. channels of the system under consideration and will conform substantially to the appropriate vestigial sideband characteristic.

Appendix 2

COLOUR PICTURE TUBE PHOSPHORS

The phosphors initially chosen for the N.T.S.C. system in 1951/2 were the best available at that time and these in turn defined the chromaticities of the analysing primary colours of the system. The selected phosphors were very good from the point of view of the colour gamut, comprising a considerable part of the area enclosed by the spectrum locus. However, these phosphors manifested two most serious deficiencies. In order of importance they are: very low light output, and discrepancies between their individual decay characteristics giving rise to yellowish trailing effects.

In the late 1950s this phosphate group of phosphors was superseded by the sulphide group giving a phenomenal increase in light output together with matched decay characteristics. Whilst the blue phosphor remained as satisfactory as that of the older group, the

red and green phosphors were significantly reduced in saturation and the red phosphor shifted to a shorter dominant wavelength. Further, the dominant wavelength of the red phosphor was in some measure dependent upon excitation level, becoming shorter in wavelength and desaturated with increasing beam current. In spite of these effects, including the attendant reduction in colour gamut, the increase in light output was so significant that these phosphors rapidly became standard. It was recognized that for large differences, picture brightness was always subjectively more desirable than absolute accuracy of colour reproduction.

In 1965, the process of using rare earth activation of the red phosphor was introduced, an example being europium-activated yttrium vanadate. This develop-

ment produced a red phosphor very much nearer to the dominant wavelength and saturation of the original N.T.S.C. phosphate red phosphor. The green and blue phosphors remained the sulphide type.

The situation was now such that whilst various changes had been introduced in the reproducing phosphors, little attention was paid to the resulting effect upon colour reproduction accuracy introduced by leaving the analysing primaries unchanged.

Up to this point, colour television was only in public service in the U.S.A. and Japan and two factors might well have contributed to the lack of consideration given to a possible review of the analysing primaries. These were the masking of other hue inaccuracies by the effects of phase distortion and the presence of a hue control in the N.T.S.C. system in use in those two countries. The other contributing effect was the then virtually universal practice of setting receivers to a white point of 9300°K whilst the signal origination end of the chain was normalized for an illuminant corresponding to Illuminant C.

The commencement of European colour television using systems not requiring a N.T.S.C. type hue control, nor manifesting a sensitivity to distortions giving rise to hue errors, has instigated detailed work on factors that now affect colour reproduction accuracy. In the U.K. this work has been carried out by the Working Party on PAL colour television tolerances.

The relationship between the analysing primaries and the reproduction phosphor chromaticity coordinates was recognized as a major factor affecting accuracy of colour reproduction. As a preliminary, the Working Party agreed to standardize the normalized luminance to Illuminant D6500 and to place due weight on the necessity that receivers be adjusted correctly to this white point. The details associated with the change from Illuminant C to Illuminant D6500 are discussed in the Working Party's CP.10.

Matrixing operations may be employed at the picture originating end of the chain to allow for various factors including discrepancies between analysing and reproducing primary colour chromaticity coordinates. The extent to which this process can be taken is limited by such factors as signal to noise ratio but without knowledge of the reproducing phosphor chromaticity coordinates the process cannot be carried out.

Table 2
Summary of phosphor chromaticity coordinates

| <i>Original N.T.S.C.:</i> | <i>x</i> | <i>y</i> | <i>u</i> | <i>v</i> |
|-----------------------------|----------|----------|----------|----------|
| Red | 0.67 | 0.33 | 0.477 | 0.352 |
| Green | 0.21 | 0.71 | 0.076 | 0.384 |
| Blue | 0.14 | 0.08 | 0.152 | 0.130 |
| <i>Sulphide Group:</i> | | | | |
| Red | 0.663 | 0.337 | 0.464 | 0.354 |
| Green | 0.285 | 0.595 | 0.119 | 0.373 |
| Blue | 0.154 | 0.068 | 0.175 | 0.116 |
| <i>Rare Earth Group:</i> | | | | |
| Red | 0.675 | 0.325 | 0.482 | 0.351 |
| Green | 0.285 | 0.595 | 0.119 | 0.373 |
| Blue | 0.154 | 0.068 | 0.175 | 0.116 |
| <i>Working Party Group:</i> | | | | |
| Red | 0.64 | 0.33 | 0.451 | 0.349 |
| Green | 0.29 | 0.60 | 0.121 | 0.374 |
| Blue | 0.15 | 0.06 | 0.175 | 0.105 |

To resolve the difficulty, discussions took place, following which nominal values for the phosphor chromaticity coordinates were agreed between tube manufacturers, broadcasters and the receiver manufacturing industry. These nominal phosphor chromaticity coordinates of necessity had fairly wide tolerances placed upon them to allow not only for the manufacturing tolerances of picture tubes but to allow foreseeable developments in the picture tube field and realizing improved performance in general. In particular, changes must be considered that give greater light output. In fact, some of these changes have already taken place during the latter part of 1968.

It was agreed that if tube development proceeded such that to obtain very significant improvement in performance, large changes outside the agreed tolerances of phosphor chromaticity coordinates were to take place, then it might prove necessary to reconsider the analysing primaries and to institute a change at a suitable time.

The reference receiver described in Appendix 2 requires a defined performance tube and for this purpose a set of very close phosphor tolerances have been defined. These were computed for a small range of flesh tones and are of course not related to the manufacturing tolerances of any production tube.

Manuscript received by the Institution on 8th August 1969. (Paper No. 1281/Com. 23.)

© The Institution of Electronic and Radio Engineers, 1969

Waveguide Solutions by the Finite-element Method

By

S. AHMED,

B.Sc., M.S., C.Eng., M.I.E.E.†

AND

P. DALY,

B.Sc., Ph.D.†

Summary: A finite-element method based on a function minimization technique is developed for analysing homogeneous waveguide problems. The continuous eigen-value operator as in the Ritz finite-difference method is replaced by a matrix operator. The resulting equations are, however, matrix eigen-value equations. Relevant properties of the matrices are discussed. It is demonstrated that the method has advantages in dealing with awkward boundaries and singularities. A feature of the method is that the error in the eigen-value is a monotonically decreasing function of successive sub-divisions of the cross-section. Minimization of a variational expression assures rapid convergence to the correct eigen-value. Three waveguides are analysed in some detail.

1. Introduction

In recent years numerical techniques have been widely applied to the solution of waveguide problems, where attempts at analytic solutions prove unfruitful. In this paper we formulate a finite-element method for solving propagation problems in uniform homogeneous waveguides. This method has distinct advantages over the standard finite-difference method described by previous authors.¹⁻³ In both cases the waveguide cross-section is divided by mesh lines into a number of polygonal sub-domains or elements. In the finite-difference approach, the linear differential equation satisfied by the fields is replaced by a set of difference equations involving field values at mesh points. Elsewhere, the field is undefined. Using a finite-element technique, one first sets up an integral variational expression for some parameter of the problem—in our case the cut-off wave-number. The field is then defined within any element as a linear algebraic function of the fields at vertices of the elements—triangular elements are chosen in our approach. Minimization of the variational expression with respect to the field values at each vertex is carried out, resulting in a set of linear algebraic equations which are written in matrix form as $\mathbf{A}\phi = \lambda\mathbf{B}\phi$. The coefficients of \mathbf{A} depend on the area of the elements while those of \mathbf{B} also depend on the area moments about the axis. A finite-element method has been used to obtain solutions of Poisson's equation.^{4, 5} The equations described as Ritz finite-difference equations are of the form $\mathbf{A}^1\phi = \theta$ where $\mathbf{A}^1 = \mathbf{A}$ if the boundary conditions are of the Dirichlet type.

In waveguide problems, the advantages of the finite-element method over the finite-difference method are: (i) successive sub-divisions of the cross-section lead to a monotonic decrease of the eigen-value towards its

extremum value, (ii) a more rapid rate of convergence towards the eigen-value is assured, (iii) awkward boundary shapes are easier to handle and do not lead to asymmetric matrices, and (iv) singular points in the waveguide require no special treatment.

The formulation of the finite-element method⁶ for waveguides is developed, the properties of the matrices are discussed and three particular problems are solved to illustrate the procedures.

2. Development of the Finite-element Formulation

The basic requirement of the development of finite-element equations for approximating continuous operator eigen-value equations is to find an extremum functional which can be written in Euler density form. As a first step in the development of this method for waveguide problems, a uniform waveguide which is completely filled with homogeneous and isotropic dielectric and whose boundary walls are perfectly conducting is considered.

It is always possible to obtain pure H or E-type modes in homogeneous waveguides. The extremization functional for electromagnetic wave propagation in a homogeneous waveguide is given by

$$2J(\phi) = \iint_s |\nabla\phi|^2 ds - K^2 \iint_s \phi^2 ds \quad \dots\dots(1)$$

$$2J(\psi) = \iint_s |\nabla\psi|^2 ds - K^2 \iint_s \psi^2 ds \quad \dots\dots(2)$$

where the scalar functions ϕ and ψ correspond to the longitudinal components of \mathbf{H} and \mathbf{E} respectively.

In eqns. (1) and (2) K is the cut-off wavenumber and is given by

$$K^2 = \omega^2\mu\epsilon - \beta^2 \quad \dots\dots(3)$$

It may be shown that the expressions in eqns. (1) and (2) are variational in K^2 .

† Department of Electrical and Electronic Engineering, University of Leeds.

Instead of approximating the true ϕ and ψ by a set of normal mode functions—in many cases we have incomplete knowledge of these functions—the finite-element method employs a set of algebraic functions defined over a sub-section of the whole waveguide cross-section. These sub-sections may be polygonal in shape and are called elements. Thus in the finite-element method the entire domain over which the operator equation is defined is divided into a finite number of elements, on each of which the actual mode function is approximated by a set of continuous algebraic functions which are only defined over the particular element under consideration, and are linearly dependent on the values of ϕ or ψ at the vertices of the element. Hence, if an element has n vertices (for a triangular element $n = 3$), the potential ϕ or ψ within it may be approximated by

$$\phi(x, y)|\psi(x, y) = \sum_k N_k(x, y)\phi_k|\psi_k \dots\dots(4)$$

where ϕ_k and ψ_k are the values of ϕ and ψ at the vertex k and $N_k(x, y)$ is a predetermined algebraic function which is uniquely defined and differentiable over the element and which reduces to zero outside the element.

We consider an arbitrary waveguide cross-section with the scheme of grading into elements as shown in Fig. 1.

For illustrative purposes, the elements, numbering P , are chosen to be triangular. A typical element (the e th) is described by the vertices i, j and m in cyclic order. We assume that ϕ_i, ϕ_j and ϕ_m are the values of ϕ at these vertices. For the element e the functional dependence of $\phi(x, y)$ can be written as

$$\phi^e(x, y) = \alpha_0 + \alpha_1x + \alpha_2y \dots\dots(5)$$

where α_0, α_1 and α_2 are to be determined.

If $(x_i, y_i), (x_j, y_j)$ and (x_m, y_m) are the co-ordinates of the vertices i, j and m then solving for α_0, α_1 and α_2 , we obtain,

$$\phi^e(x, y) = \frac{1}{2\Delta_e} [(a_i + b_ix + c_iy)\phi_i^e + (a_j + b_jx + c_jy)\phi_j^e + (a_m + b_mx + c_my)\phi_m^e] \dots\dots(6)$$

where

$$\begin{aligned} a_i &= x_jy_m - x_my_j \\ b_i &= y_j - y_m \\ c_i &= x_m - x_j \end{aligned} \dots\dots(7)$$

and Δ_e is the area of the triangular element which is given by

$$2\Delta_e = \det \begin{bmatrix} 1 & x_i & y_i \\ 1 & x_j & y_j \\ 1 & x_m & y_m \end{bmatrix} \dots\dots(8)$$

The values of other parameters can be obtained by a cyclic rotation of the suffixes i, j and m .

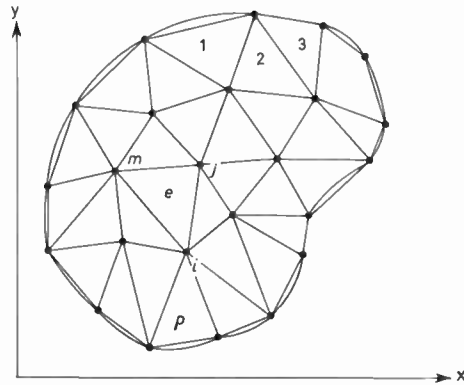


Fig. 1. Divisions of an arbitrary waveguide cross-section into elements.

It is important to note that the functional form of the potential $\phi^e(x, y)$ as described in eqn. (6) for all the elements of the entire domain satisfies the continuity relation throughout the whole region. This continuity of ϕ is essential for the validity of the variational expression. The derivative normal to the line joining two elements is not, however, continuous for such a functional relation. The finite jump in the normal derivative will not introduce any error into the variational formulation, because the contribution of this type of discontinuity in the normal derivative to the net integrated value of the functional J is always zero.

When the functional form of ϕ as given above is substituted into the right-hand side of eqn. (1) and the corresponding integrations are carried out, J will be a function of the variables ϕ_k . If there are in all M vertices, then

$$J(\phi) = F(\phi_1, \phi_2, \phi_3 \dots \phi_M) \dots\dots(9)$$

The optimum value of a set of ϕ_k for a certain functional form of $N_k(x, y)$ may be obtained by minimizing the functional given in eqn. (9) with respect to each of ϕ_k , i.e. equating

$$\frac{\partial J}{\partial \phi_k} = 0; \text{ for } k = 1, 2, 3 \dots M \dots\dots(10)$$

In view of the functional form of $\phi^e(x, y)$ as given in eqn. (6), eqn. (1) can be rewritten as

$$2J(\phi) = \sum_{e=1}^P \iint_{\Delta_e} |\nabla\phi|^2 dx dy - K^2 \sum_{e=1}^P \iint_{\Delta_e} \phi^2 dx dy \dots\dots(11)$$

The minimization of J over the entire cross-section is equivalent to the minimization of J over each of the elements individually. Each element contains only three values of ϕ_i but each ϕ_i is common to those elements for which the vertex i is common. If we minimize the functional J with respect to each ϕ we obtain,

$$\frac{\partial J}{\partial \phi_i} = \sum_{e=1}^P \iint_{\Delta e} \left[\frac{\partial \phi}{\partial x} \frac{\partial}{\partial \phi_i} \left(\frac{\partial \phi}{\partial x} \right) + \frac{\partial \phi}{\partial y} \frac{\partial}{\partial \phi_i} \left(\frac{\partial \phi}{\partial y} \right) \right] dx dy - K^2 \sum_{e=1}^P \iint_{\Delta e} \phi \frac{\partial \phi}{\partial \phi_i} dx dy \quad \dots\dots(12)$$

Substituting for $\phi(x, y)$ into eqn. (12) and taking into account the fact that the vertex i is common to the R elements associated with it, we obtain,

$$\frac{\partial J}{\partial \phi_i} = \sum_{e=1}^R \iint_{\Delta e} \sum_{k=i,j,m} \phi_k^e \left[\frac{\partial N_k^e}{\partial x} \frac{\partial N_i^e}{\partial x} + \frac{\partial N_k^e}{\partial y} \frac{\partial N_i^e}{\partial y} \right] dx dy - K^2 \sum_{e=1}^R \iint_{\Delta e} \sum_{k=i,j,m} N_i^e N_k^e \phi_k^e dx dy \quad \dots\dots(13)$$

where

$$N_k^e = (a_k + b_k x + c_k y) / 2\Delta_e \quad \dots\dots(14)$$

If we consider the minimization of the functional J with respect to every ϕ_i , we obtain a set of linear algebraic eigen-value equations which in matrix notation can be written as follows:

$$\mathbf{A}\phi = K^2 \mathbf{B}\phi \quad \dots\dots(15)$$

In eqn. (15) \mathbf{A} and \mathbf{B} are square matrices of order M and ϕ is a $(M \times 1)$ column matrix whose components are the values of ϕ_k at the M vertices. Thus we see that the solution of the above matrix eigen-value problem will give us an approximate solution to all waveguide problems for which eqn. (1) is the appropriate variational functional. Though we have considered above the case of H-mode propagation, the same formulation will hold for E-mode propagation, if the potential function ϕ is replaced by ψ .

3. Evaluation of the Components of the Matrices A and B

Consider the partial contribution to the components of the matrix \mathbf{A} by a typical element e . This is obtained from eqn. (13) as follows

$$\frac{\partial J^e}{\partial \phi_i^e} \Big|_{\mathbf{A}} = \iint_{\Delta e} \sum_{k=i,j,m} \phi_k^e \left[\frac{\partial N_k^e}{\partial x} \frac{\partial N_i^e}{\partial x} + \frac{\partial N_k^e}{\partial y} \frac{\partial N_i^e}{\partial y} \right] dx dy \quad \dots\dots(16)$$

The above equation can be written in compact form by using matrix notation as

$$\frac{\partial J^e}{\partial \phi_i^e} \Big|_{\mathbf{A}} = \iint_{\Delta e} \left[\frac{\partial}{\partial x} (\mathbf{G}) \cdot \Phi^e \frac{\partial N_i^e}{\partial x} + \frac{\partial}{\partial y} (\mathbf{G}) \cdot \Phi^e \frac{\partial N_i^e}{\partial y} \right] dx dy \quad \dots\dots(17)$$

where

$$\frac{\partial J^e}{\partial \phi_i^e} \Big|_{\mathbf{A}}$$

indicates that part of

$$\frac{\partial J^e}{\partial \phi}$$

which contributes to matrix \mathbf{A} and \mathbf{G} is a row matrix whose components are given by

$$\mathbf{G} = [N_i^e N_j^e N_m^e] \quad \dots\dots(18)$$

and Φ^e is a column matrix given by

$$\Phi^e = \begin{Bmatrix} \phi_i^e \\ \phi_j^e \\ \phi_m^e \end{Bmatrix} \quad \dots\dots(19)$$

Substituting the values of $N_i^e(x, y)$, $N_j^e(x, y)$ and $N_m^e(x, y)$ into eqn. (17), we obtain,

$$\frac{\partial J^e}{\partial \phi_i^e} \Big|_{\mathbf{A}} = [S_{ii}^e S_{ij}^e S_{im}^e] \{\Phi^e\}. \quad \dots\dots(20)$$

where

$$S_{ij}^e = \frac{1}{4\Delta e} (b_i b_j + c_i c_j) \quad \dots\dots(21)$$

In the same way, if J is minimized with respect to ϕ_j^e and ϕ_m^e , we obtain,

$$\begin{bmatrix} \frac{\partial J^e}{\partial \phi_i^e} \\ \frac{\partial J^e}{\partial \phi_j^e} \\ \frac{\partial J^e}{\partial \phi_m^e} \end{bmatrix} \Big|_{\mathbf{A}} = \begin{bmatrix} S_{ii}^e & S_{ij}^e & S_{im}^e \\ S_{ji}^e & S_{jj}^e & S_{jm}^e \\ S_{mi}^e & S_{mj}^e & S_{mm}^e \end{bmatrix} \cdot \begin{bmatrix} \phi_i^e \\ \phi_j^e \\ \phi_m^e \end{bmatrix} \quad \dots\dots(22)$$

or

$$\frac{\partial J^e}{\partial \phi^e} \Big|_{\mathbf{A}} = \mathbf{S}^e \Phi^e \quad \dots\dots(23)$$

where

$$\frac{\partial J^e}{\partial \phi^e} \Big|_{\mathbf{A}}$$

is a column matrix given by the left-hand side of eqn. (22).

If the elements are right triangles with vertex i at the corner of the right angle, then the submatrix \mathbf{S}^e becomes

$$\mathbf{S}_e = \begin{bmatrix} 1 - \frac{1}{2} - \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

Similarly, we can derive an expression for the partial contribution to the components of the matrix \mathbf{B} subscribed by the element e . This will be

$$\begin{bmatrix} \frac{\partial J^e}{\partial \phi_i^e} \\ \frac{\partial J^e}{\partial \phi_j^e} \\ \frac{\partial J^e}{\partial \phi_m^e} \end{bmatrix} \Big|_{\mathbf{B}} = \begin{bmatrix} F_{ii}^e & F_{ij}^e & F_{im}^e \\ F_{ji}^e & F_{jj}^e & F_{jm}^e \\ F_{mi}^e & F_{mj}^e & F_{mm}^e \end{bmatrix} \begin{bmatrix} \phi_i^e \\ \phi_j^e \\ \phi_m^e \end{bmatrix} \quad \dots\dots(24)$$

or

$$\frac{\partial J^e}{\partial \phi^e} \Big|_B = F^e \cdot \Phi^e \quad \dots\dots(25)$$

where

$$F_{ij}^e = \frac{1}{4\Delta_e^2} \iint_{\Delta_e} (a_i + b_i x + c_i y)(a_j + b_j x + c_j y) dx dy \quad \dots\dots(26)$$

and on simplification F^e becomes

$$F^e = \frac{\Delta_e}{6} \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix} \quad \dots\dots(27)$$

Now combining eqns. (23) and (25) we get the partial contribution to the minimization of the function J subscribed by the approximated function over the element e as

$$\left[\frac{\partial J^e}{\partial \phi^e} \right] = S^e \cdot \Phi^e - K^2 F^e \cdot \Phi^e \quad \dots\dots(28)$$

The matrices S^e and F^e as obtained above may be called the element sub-matrices of matrices **A** and **B** respectively. Though they are written for simplicity as (3×3) square matrices, each of them is a $(M \times M)$ square matrix with at most nine non-zero components. The matrices **A** and **B** are the sum of all the element sub-matrices generated by all the elements.

4. Properties of the Matrix Operators A and B

The properties of the matrices **A** and **B** will depend on the nature of P element sub-matrices as discussed above. Referring back to Section 3 it may be seen that

$$S_{ij}^e = S_{ji}^e \quad \text{and} \quad F_{ij}^e = F_{ji}^e$$

Therefore, each of the element sub-matrices is a symmetric matrix. The matrices **A** and **B** are the matrix sum of P symmetric element sub-matrices and thus they are also symmetric. Since each of the vertices is common to a certain number R of the elements which is much less than the total number P , the generated matrices **A** and **B** will be sparse band matrices.

It can be shown that the matrix **B** as given above is always a positive-definite and diagonally dominant matrix. Equation (27) shows that the components of the element sub-matrix F^e are all positive and the algebraic sum of the off-diagonal components of any row is equal to the diagonal component. For those vertices which are adjacent to the electric or the magnetic wall in waveguide problems, the boundary conditions on the corresponding potential will generate some rows for which the sum of the off-diagonal terms will be less than the diagonal terms. Therefore, the matrix **B** is always diagonally dominant.

Finally, let us consider the quadratic form $\Phi^T F \Phi$ for an element e :

$$\Phi^T F \Phi = F_{ii} \phi_i^2 + 2F_{ij} \phi_i \phi_j + F_{jj} \phi_j^2 + 2F_{im} \phi_m \phi_i + 2F_{jm} \phi_j \phi_m + F_{mm} \phi_m^2 \quad \dots\dots(29)$$

Substituting the values of F_{ij} as given in eqn. (27) into eqn. (29), we obtain

$$\Phi^T F \Phi = \frac{\Delta_e}{12} [(\phi_i + \phi_j + \phi_m)^2 + \phi_i^2 + \phi_j^2 + \phi_m^2] > 0 \quad \dots\dots(30)$$

for all values of non-zero ϕ . Therefore the element submatrix F^e as well as matrix **B** is always positive-definite.

Similarly, it can be shown that the diagonal components of all the element submatrices S^e are positive and the algebraic sum of the off-diagonal terms of any row is exactly equal in magnitude to the diagonal term. Also it can be shown that for triangular elements with interior angles not exceeding 90° , the off-diagonal terms are all negative. Again the boundary conditions on the appropriate potential will make the resultant matrix **A** diagonally dominant.

If we now consider the quadratic form $\Phi^T S \Phi$ for an element submatrix S^e we obtain,

$$\Phi^T S \Phi = S_{ii} \phi_i^2 + S_{jj} \phi_j^2 + S_{mm} \phi_m^2 + 2S_{ij} \phi_i \phi_j + 2S_{im} \phi_i \phi_m + 2S_{jm} \phi_j \phi_m \quad \dots\dots(31)$$

But,

$$\begin{aligned} S_{ii} &\geq -(S_{ij} + S_{im}) \\ S_{jj} &\geq -(S_{ij} + S_{jm}) \\ S_{mm} &\geq -(S_{mi} + S_{jm}) \end{aligned}$$

Substituting the above inequalities into eqn. (31) we obtain

$$\Phi^T S \Phi \geq -S_{ij}(\phi_i - \phi_j)^2 - S_{jm}(\phi_j - \phi_m)^2 - S_{im}(\phi_i - \phi_m)^2 \quad \dots\dots(32)$$

Since S_{ij} , S_{im} and S_{jm} are negative,

$$\Phi^T S \Phi > 0 \text{ for all non-zero column vector } \Phi.$$

Hence the element submatrices S^e and matrix **A** are also positive-definite. These properties of matrices **A** and **B** guarantee⁷ the convergence of iterative solutions.

5. Particular Solutions

The numerical techniques developed in the previous sections are used to calculate the wave number of three typical waveguide problems. These problems are chosen to demonstrate the form of the finite-element equations in special cases, the method of handling awkward boundaries and singularities in the waveguide cross-section.

5.1. Circular Waveguide

If in a two-dimensional boundary value problem, the functional variation of the fields with respect to any

one of the dimensions is known, then the corresponding extremum functional can be written in terms of only one variable. For a circular waveguide the extremum functional is given by

$$J(\phi) = \iint_S \left(\frac{\partial\phi}{\partial r}\right)^2 r dr d\theta + \iint_S \frac{1}{r} \left(\frac{\partial\phi}{\partial\theta}\right)^2 dr d\theta - K^2 \iint_S \phi^2 r dr d\theta \dots(33)$$

where ϕ is the longitudinal component of the appropriate field vector. In a circular waveguide, the azimuthal variation of the fields takes the form $\cos n\theta/\sin n\theta$. If we insert the above form of θ -dependence for ϕ into eqn. (32) and integrate with respect to θ , then the resultant integrals will be functions of r only. We now divide the radius into a number of line segments as shown in Fig. 2, each of these segments being the element in a one dimensional problem. We assume a linear variation of the field over each element and following the procedure discussed earlier, we obtain a set of linear algebraic equations. The solution for the lowest eigen-value of this set of equations will correspond to the wave numbers for H_{n1} or E_{n1} mode. To illustrate this, a few of the typical finite-element equations for E_{01} , E_{11} and H_{11} modes, corresponding to the scheme of grading as shown in Fig. 2, are given below. Matrices **A**, **B** and ϕ are given for the case $N = 4$.

(a) E_{01} mode

If p is an interior vertex then

$$-(2p-1)\phi_{p-1} + 4p\phi_p - (2p+1)\phi_{p+1} = \frac{K^2 h^2}{6} \times [(2p-1)\phi_{p-1} + 8p\phi_p + (2p+1)\phi_{p+1}] \dots(34)$$

and for the vertex at the centre of the guide

$$\phi_0 - \phi_1 = \frac{K^2 h^2}{6} (\phi_0 + \phi_1) \dots(35)$$

where h is the length of each element.

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 4 & -3 & 0 \\ 0 & -3 & 8 & -5 \\ 0 & 0 & -5 & 12 \end{bmatrix}; \mathbf{B} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 8 & 3 & 0 \\ 0 & 3 & 16 & 5 \\ 0 & 0 & 5 & 24 \end{bmatrix}; \phi = \begin{bmatrix} \phi_0 \\ \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix}$$

(b) E_{11} mode

For E_{11} mode all vertices are interior vertices and for any vertex p we obtain

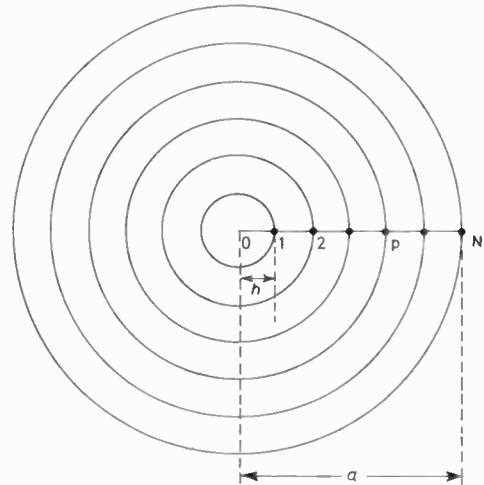


Fig. 2. Grading of a circular waveguide into one dimensional elements.

$$- \left[p(p-1) \ln \frac{p}{p-1} \right] \phi_{p-1} + \left[(p-1)^2 \ln \frac{p}{p-1} + (p+1)^2 \ln \frac{p+1}{p} \right] \phi_p - \left[p(p+1) \ln \frac{p+1}{p} \right] \phi_{p+1} = \frac{K^2 h^2}{12} \times [(2p-1)\phi_{p-1} + 8p\phi_p + (2p+1)\phi_{p+1}] \dots(36)$$

$$\mathbf{A} = \begin{bmatrix} 4 \ln 2 & -2 \ln 2 & 0 \\ -2 \ln 2 & \ln 2 + 9 \ln \frac{3}{2} & -6 \ln \frac{3}{2} \\ 0 & -6 \ln \frac{3}{2} & 4 \ln \frac{3}{2} + 16 \ln \frac{4}{3} \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} 8 & 3 & 0 \\ 3 & 16 & 5 \\ 0 & 5 & 24 \end{bmatrix}; \phi = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix}$$

(c) H_{11} mode

For all interior vertices the finite-element equations will be the same as eqn. (35), but for the vertex N , at the guide wall, this becomes,

$$- \left[N(N-1) \ln \frac{N}{N-1} \right] \phi_{N-1} + \left[1 + (N-1)^2 \ln \frac{N}{N-1} \right] \phi_N = \frac{K^2 h^2}{12} \times [(2N-1) \cdot \phi_{N-1} + (4N-1)\phi_N] \dots(37)$$

$$\mathbf{A} = \begin{bmatrix} 4 \ln 2 & -2 \ln 2 & 0 & 0 \\ -2 \ln 2 & \ln 2 + 9 \ln \frac{3}{2} & -6 \ln \frac{3}{2} & 0 \\ 0 & -6 \ln \frac{3}{2} & 4 \ln \frac{3}{2} + 16 \ln \frac{4}{3} & -12 \ln \frac{4}{3} \\ 0 & 0 & -12 \ln \frac{4}{3} & 1 + 9 \ln \frac{4}{3} \end{bmatrix}$$

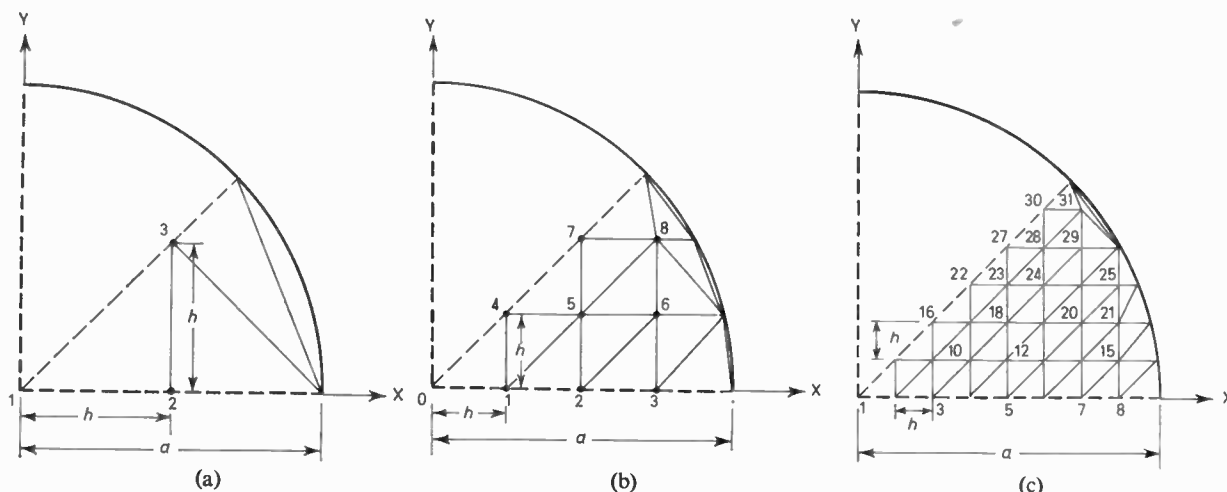


Fig. 3. Scheme of grading a sector of a circular waveguide into triangular elements.

$$B = \begin{bmatrix} 8 & 3 & 0 & 0 \\ 3 & 16 & 5 & 0 \\ 0 & 5 & 24 & 7 \\ 0 & 0 & 7 & 32 \end{bmatrix}; \phi = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{bmatrix}$$

Numerical computations of Ka for a circular waveguide sustaining E_{01} , E_{11} and H_{11} modes are given in Tables 1, 2 and 3.

5.2. Boundary Matching

The theory developed in this paper can be suitably applied to problems with arbitrary boundaries. In the finite-element method, an arbitrary boundary is approximated by a set of piece-wise linear boundaries. The common points between successive linear boundaries are considered as the boundary vertices. The resultant eigen-value operators remain symmetric. As an illustration, a circular waveguide is analysed

Table 1: E_{01} mode in a circular waveguide

| $\frac{h}{a}$ | Ka | $\frac{h}{a}$ | Ka | $\frac{h}{a}$ | Ka |
|---------------|-------|---------------|-------|---------------|-------|
| 1 | 2.450 | $\frac{1}{2}$ | 2.410 | $\frac{1}{3}$ | 2.407 |
| $\frac{1}{2}$ | 2.420 | $\frac{1}{3}$ | 2.408 | $\frac{1}{4}$ | 2.406 |
| $\frac{1}{3}$ | 2.414 | $\frac{1}{4}$ | 2.408 | $\frac{1}{5}$ | 2.405 |

First root of $J_0(Ka)$ is $Ka = 2.405$

Table 2: E_{11} mode in a circular waveguide

| $\frac{h}{a}$ | Ka | $\frac{h}{a}$ | Ka | $\frac{h}{a}$ | Ka |
|---------------|-------|---------------|-------|---------------|-------|
| $\frac{1}{2}$ | 4.078 | $\frac{1}{3}$ | 3.872 | $\frac{1}{4}$ | 3.847 |
| $\frac{1}{3}$ | 3.942 | $\frac{1}{4}$ | 3.859 | $\frac{1}{5}$ | 3.844 |
| $\frac{1}{4}$ | 3.894 | $\frac{1}{5}$ | 3.852 | $\frac{1}{6}$ | 3.841 |

First root of $J'_0(Ka)$ is $Ka = 3.832$

numerically in cartesian co-ordinates for the E_{01} mode, and the curved boundary is divided up as shown in Figs. 3(a), (b) and (c). Numerical results are shown in Table 4. Comparing the computed results with the exact result we find that the rate of convergence is fairly rapid.

5.3. Dominant Mode in a Single Ridged Waveguide

Exact analytical solutions in compact form cannot be obtained for ridged waveguides. In writing the finite-difference equations, various types of assumptions are necessary in taking into account the presence of the re-entrant corner. No special treatment is necessary for such a point in the finite-element method. The finite-element equation for a typical vertex 0 and for the vertex at the singular point s of a ridged waveguide graded into triangular elements as shown in Fig. 4 are given below.

Table 3: H_{11} mode in a circular waveguide

| $\frac{h}{a}$ | Ka | $\frac{h}{a}$ | Ka | $\frac{h}{a}$ | Ka |
|---------------|-------|---------------|-------|---------------|-------|
| 1 | 2.000 | $\frac{1}{2}$ | 1.849 | $\frac{1}{3}$ | 1.844 |
| $\frac{1}{2}$ | 1.875 | $\frac{1}{3}$ | 1.848 | $\frac{1}{4}$ | 1.843 |
| $\frac{1}{3}$ | 1.856 | $\frac{1}{4}$ | 1.845 | $\frac{1}{5}$ | 1.842 |

First root of $J_1(Ka)$ is $Ka = 1.841$

Table 4: E_{01} mode in a circular waveguide

| $\frac{h}{a}$ | Fig. No. | Number of equations | Ka |
|---------------|----------|---------------------|-------|
| $\frac{1}{2}$ | 3(a) | 3 | 2.631 |
| $\frac{1}{3}$ | 3(b) | 9 | 2.511 |
| $\frac{1}{4}$ | 3(c) | 31 | 2.460 |

Extrapolated $Ka = 2.420$. First root of $J_0(Ka)$ is $Ka = 2.405$

For a typical vertex 0

$$4\phi_0 - \phi_1 - \phi_2 - \phi_3 - \phi_4 = \frac{K^2 h^2}{12} (6\phi_0 + \phi_1 + \phi_2 + \phi_3 + \phi_4 + \phi_5 + \phi_6) \dots\dots(38)$$

and for the vertex at the singular point s

$$3\phi_s - \phi_p - \phi_r - \frac{1}{2}\phi_t - \frac{1}{2}\phi_u = \frac{K^2 h^2}{12} (4\phi_s + \phi_p + \phi_q + \phi_r + \frac{1}{2}\phi_t + \frac{1}{2}\phi_u) \dots\dots(39)$$

Table 5 gives the numerical computation of the dominant mode wave number for such a guide. A study of these results shows that on successive element sub-division the wave number monotonically approaches the value given by Pyle.⁸

Table 5: Dominant mode in a single ridged rectangular waveguide

| $\frac{h}{a}$ | $\frac{1}{8}$ | $\frac{1}{16}$ | $\frac{1}{32}$ | $\frac{1}{64}$ | $\frac{1}{128}$ |
|-----------------|---------------|----------------|----------------|----------------|-----------------|
| No. of vertices | 18 | 60 | 216 | 816 | 3168 |
| Ka | 2.300 | 2.274 | 2.257 | 2.252 | 2.250 |

6. Conclusions

By means of the finite-element approach, a uniform treatment of the boundary conditions and the differential equation is possible without any special consideration of the detailed nature of the boundary. The coefficients of the matrices A and B are developed from the moments of the element area about the axes and are thus insensitive to the variations of the local slope of the boundary. Under all circumstances, the matrices remain symmetric, positive-definite, sparse and are band-type; these properties are of great practical value when solutions for very large matrices are required via the computer and also ensure convergence of an iterative solution using over-relaxation procedures.

The rate of convergence is dependent on the over-relaxation factor used during the computation. This factor is calculated⁹ via the Rayleigh coefficient after every few iterations and thus an optimum rate of convergence is assured. Since the computed eigen-value is an upper bound to the exact value and is also derived from the variational expression, successive mesh sub-division leads rapidly to the correct eigen-value. If successive approximations to the exact eigen-value are known, eigen-value extrapolation may also be used successfully.

The method can be extended to deal with inhomogeneous waveguide problems.¹⁰

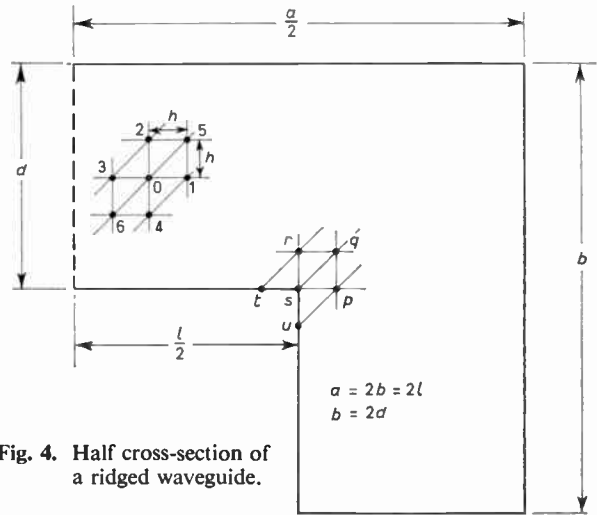


Fig. 4. Half cross-section of a ridged waveguide.

A recent paper which has come to our attention is also concerned with the derivation of a matrix eigen-value equation from the wave equation by means of finite-element methods.¹¹

7. References

- Collins, J. H. and Daly, P., 'Calculations for guided electromagnetic waves using finite-difference methods', *J. Electronics and Control*, **14**, pp. 361-80, April 1963.
- Davies, J. B. and Muilwyk, C. A., 'Numerical solution of hollow waveguides with boundaries of arbitrary shapes' *Proc. Instn Elect. Engrs*, **113**, pp. 277-84, February 1966 (I.E.E. Paper No. 4929E).
- Hannaford, C. D., Ph.D. Thesis, University of Leeds, 1967.
- Kellogg, R. B., 'A Ritz finite-difference approximation to the neutron diffusion equation', *Bettis Technical Review*, WAPDA-BT-31, pp. 51-57, 1964.
- Zienkiewicz, C. O. and Cheung, Y. K., 'Finite elements in the solution of field problems', *The Engineer*, **220**, pp. 507-10, 24th September 1965.
- Ahmed, S., 'Finite-element method for waveguide problems', *Electronics Letters*, **4**, pp. 387-89, September 1968.
- Fox, L., 'Numerical Solution of Ordinary and Partial Differential Equations', p. 288 (Pergamon, Oxford, 1962).
- Pyle, J. E., 'The cut-off wavelength of the TE₁₀ mode in the ridged rectangular waveguide of any aspect ratio', *I.E.E.E. Trans. on Microwave Theory and Techniques*, **MTT-14**, pp. 175-83.
- Silvester, P., 'Modern Electromagnetic Fields', p. 275 (Prentice-Hall, Englewood Cliffs, N.J., 1968).
- Ahmed, S. and Daly, P., 'Finite-element methods for inhomogeneous waveguides', *Proc. I.E.E.*, **116**, pp. 1661-4 October 1969.
- Arlett, P. L., Bahrani, A. K. and Zienkiewicz, O. C., 'Application of finite elements to the solution of Helmholtz's equation', *Proc. I.E.E.*, **115**, pp. 1762-66, December 1968.

Manuscript first received by the Institution on 7th March 1969 and in final form on 3rd June 1969. (Paper No. 1282/CC56.)

News of the Profession

Rationalization of Qualifications and Titles for Engineers

The Council of Engineering Institutions whose constituent members are the fourteen principal chartered engineering institutions in Great Britain is to establish a registration organization for qualifications and titles at all levels in the engineering community.

At a meeting held in London on 4th September, the following resolution was passed unanimously:

'The Council of Engineering Institutions will, in collaboration with other interested parties and subject to the agreement of the Privy Council, initiate the formation of an organization to create and administer a composite register covering the principal sections of the engineering community, currently Chartered Engineers, Technician Engineers, and Engineering Technicians.'

It was further agreed that a Working Party, consisting of one representative from each of the fourteen Institutions within C.E.I., and under the Chairmanship of Sir Arnold Lindley, would be set up forthwith to implement the resolution.

The first duty of this Working Party will be to prepare a submission to the Privy Council to get agreement to such modifications to the C.E.I. Charter and By-laws as may be necessary and then to determine which other interested parties should be invited to collaborate.

Sir Arnold Lindley who is Immediate Past President of The Institution of Mechanical Engineers and a Vice-President of the I.E.R.E., is Chairman of the C.E.I. Membership Committee.

Golden Jubilee of the Womens' Engineering Society

The presentation of a silver rose bowl was made on behalf of the Council of Engineering Institutions by its Chairman, Sir Leonard Drucquer, to the Womens' Engineering Society to mark the Society's 50th anniversary. The presentation was accepted on behalf of the W.E.S. by its President, Dr. Elizabeth Laverick, C.Eng., F.I.E.E., at an informal reception at the Institution of Civil Engineers during the Annual Conference of the W.E.S.



The occasion was also marked by the showing of the new film 'The Engineer is a Woman', produced by the Central Office of Information for the Ministry of Technology with the cooperation of the Society. Intended to present engineering as a career for girls, the interesting fact is presented that whereas in Britain only one engineer in five hundred is a woman, the corresponding figure for France is one in twenty-eight, for Norway one in ten and for Russia one in three.

As a further commemoration of its Golden Jubilee, the Womens' Engineering Society is establishing an annual lecture for fifth and sixth formers, to be known as the 'Verena Holmes Lecture', after the fifth President of the Society, a distinguished Chartered Mechanical Engineer.

Canadian Engineers and the C.E.I.

During August, Mr. T. C. Keefer, P.Eng., Registrar of the Association of Professional Engineers of the Province of Ontario, visited the United Kingdom to continue the discussions initiated three years ago by the Director of the I.E.R.E., Mr. Graham D. Clifford, during his last Canadian tour. Subsequently, the then Secretary of C.E.I. had talks in Canada with the various Associations of Canadian Professional Engineers, and in February of 1968 Mr. L. M. Nadeau, Director General of the Council of Canadian Professional Engineers came to London. As reported in the March 1968 issue of *The Radio and Electronic Engineer*, Mr. Nadeau's meetings on that occasion included informal discussions with the President and senior members of the I.E.R.E. Council.

Progress is steadily being made in reconciling the requirements of the Canadian regulations for professional engineers with the qualifications and experience that is represented by Corporate Membership of the British Chartered Institutions. A brief note on the present position was given in the July 1969 issue of *The Radio and Electronic Engineer* when reporting the activities of the Chartered Engineering Institution's Information Committee. Mr. Keefer's discussions with individual Institutions were supported by a meeting at the C.E.I. offices in Little Smith Street, Westminster, with Institution and C.E.I. officers.

After lunch Mr. Keefer talked to C.E.I. and Institution officers. The Director of the I.E.R.E., Mr. Clifford, is on Mr. Keefer's right and others listening are (right to left) Sir Leonard Drucquer, Chairman of C.E.I., Mr. J. G. Watson, Secretary of the Civils, Mr. A. G. Higgins, Secretary of the Gas Engineers, and Mr. M. W. Leonard Secretary of C.E.I.

Limitations in Magnetic Disk Storage

By

M. F. DUDSON,

B.Sc., A.Inst.P., C.Eng., M.I.E.E.†

Presented at a Joint I.E.R.E.-I.E.E. Computer Group Colloquium on 'Modern Trends in Mass Storage', held in London on 10th April 1969.

Summary: If the fundamental consideration of domain size were the dominant factor, it would theoretically be possible to store information on magnetic disks with a packing density of at least 10^{10} bits per square inch.

To be of practical use, the fundamental design limits of the data channel and its components, the practical magnetic design of the media and recording heads, the mechanical tolerances and wear properties of the entire combination of medium, head and transports, have to be considered. Under practical conditions, current files operate at a density of 2×10^5 bit per square inch.

This paper reviews the overall interaction of these limiting parameters, and discusses some of them in detail.

1. Introduction

A magnetic disk store is a device for storing digital information in a manner in which it is easily accessible to a computer system. By storing the information on discrete disks, any point on which can be accessed by a magnetic head, the device is capable of providing access to any piece of information in milliseconds compared to the seconds of even high-speed type systems.

There are two main types of disk store, fixed disk store, (f.d.s.) and exchangeable disk store (e.d.s.). The e.d.s. allows the user both to be able to remove disks, and to read on one machine disks which were written on another, and hence is the more versatile and widely used device, although a penalty is paid in keeping within the tolerance limits.

The main functional components of an exchangeable disk store are mounted upon a deck plate (see Fig. 1). They consist of a precision spindle upon which the disks (in pack form) can be mounted, and a comb of heads which are moved to the appropriate radial position by means of an actuator. The functions of the other components are subsidiary to these.

At present, information is recorded at a linear packing density of 2200 bits per inch. The track density is 100 tracks per inch giving an effective packing density of about 2×10^5 bits per square inch. In fact, the range on exchangeable disk stores is from 0.5×10^5 to 2×10^5 , while, perhaps surprisingly at first sight, that on fixed disk stores is usually somewhat lower.

The theoretical limit of information storage in magnetic recording was discussed at the Intermag

Conference in 1966 and it was indicated that progress was asymptotic to an improvement of 10^{14} . The difference between this highly theoretical figure and that which will eventually be obtained in practice is due to limitations in pure scientific techniques and practical engineering problems.

2. Improvements

2.1. The Market Requirements

In order to define the general aims for improvements, the user's viewpoint must be studied and the following requirements will be found:

Greater storage capacity,
Faster access time,
Higher data rate and
Greater reliability.

All these will be required for the same or less cost.

2.2. The Engineering Parameters

The four main parameters which can be improved in order to meet these objectives are:

Higher packing density,
Greater number of tracks per inch,
Faster rotational speed and
Shorter access time.

However, when the parameters affecting these major ones are considered in a little more detail (Fig. 2), it can be seen that there are a great many interrelationships affecting the development of the device.

Almost any two of these can be taken and optimized, but then they have an effect upon the remainder of the network. It is of interest to note that mechanical tolerances have a great effect upon the design of the

† Data Recording Instrument Co. Ltd., Hawthorne Road, Staines, Middlesex.

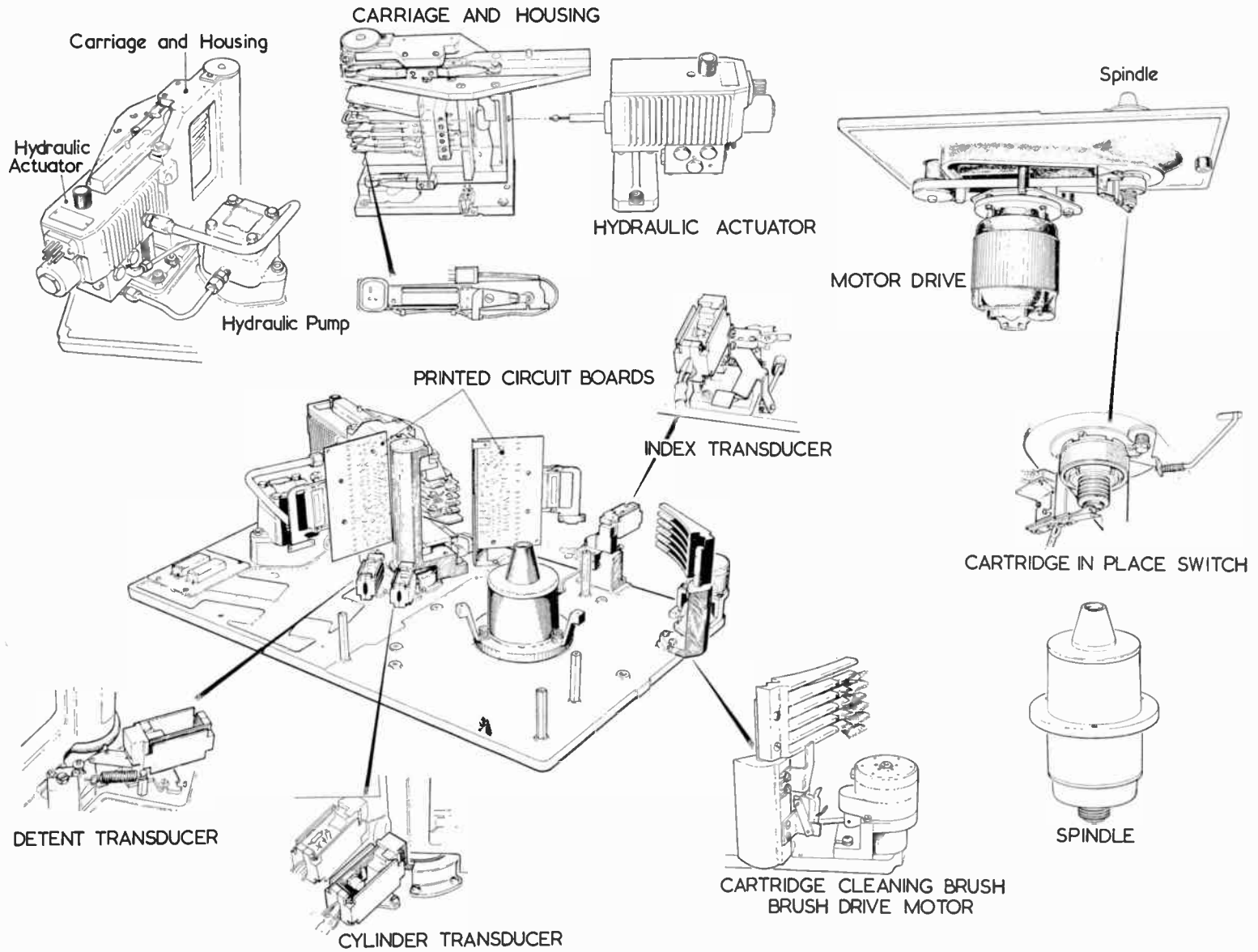


Fig. 1. The main functional components of an exchangeable disk store.

essentially electronic parts of the system. The parameters affecting the design of the data channel are a particularly good example of this. It must also be remembered that the network is only shown here in two dimensions, and that the whole design problem is weighted by cost, time-scale and engineering feasibility considerations.

3. Detailed Consideration of Some of the Interrelations

In order to show the sort of considerations involved, some of the interrelations shown in Fig. 2 which can be treated at least semi-analytically will be examined in detail in the following Sections.

3.1. Considerations of Using Oxide or Plated Disks

The width of an isolated flux transition gives an indication of the performance of the medium as far as resolution is concerned.

It has been shown¹ that the transition width is proportional to the quantity a where

$$a = \frac{B_r C}{2\pi H_c}$$

where C is the thickness of the coating, B_r is the retentivity of the medium and H_c is the coercivity of the medium.

Typical values of B_r and H_c are:

| | B_r (gauss) | H_c (oersteds) |
|--------|---------------|------------------|
| Oxide | 1 000 | 250 |
| Cobalt | 10 000 | 750 |

A method of comparing the effectiveness of the two media would be to compare the output from coatings of the different media which were of the correct thickness to give an identical pulse width. This is done by substituting the above values in the expression for a , and finding the appropriate thicknesses of the respective media. If these thicknesses are then substituted in the formula for the amplitude of an isolated pulse, it is shown that the high-coercivity, high-remanence cobalt plating gives an output which is at least three times greater than that of the oxide (see Appendix 1). However, to realize this advantage, particularly in cases where the head-to-media separation is small compared with the transition width and coercivity becomes the dominant term affecting the output, the medium has to have a high coercivity. This could present difficulties in the design of a suitable head.

3.1.1. The 'write' process

The intensity of the x component of the magnetic field H_x relative to H_g in the gap at a distance y from the gap is given approximately by²:

$$H_x = \frac{2}{\pi} H_g \arctan \frac{g}{y}$$

where g is half the length of the 'write' gap. In order to switch the medium, the horizontal field in the plane of the medium must be at least 20% greater than the coercivity of the medium, which means

$$H_x \geq 1.2H_c$$

$$1.2H_c \leq \frac{2}{\pi} H_g \arctan \frac{g}{y}$$

The form of the arctan curve is well known and shows that at a separation of g the field has fallen to $H_g/2$, while at a separation of $2g$ (the gap length) the field has fallen to $H_g/3$.

Clearly in order to be able to magnetize high coercivity media it is necessary to make y as small as possible. This means that the head must either run in contact with the disk or fly at a very low flying height. Both these conditions cause severe mechanical design problems, and also demand strict operator discipline in the field. These will be discussed later, but it will be recognized that the theoretical calculations have now started to impinge on one of the most mundane, but yet most important of subjects—operator discipline.

3.1.2. The 'read' process

In disk files it is usual, because of head geometry problems, to use a common gap for both 'writing' and 'reading'. It has been shown above that for the 'write' process it is advantageous to have a long gap, although a finite limit is set to this by current rise-times, rotational speeds and core geometry.

The read pulse width in the isolated pulse condition at the 50% point is¹

$$P_{50} = 2\sqrt{(a+d)^2 + g^2}$$

It should be noted that the longer the gap, the worse the resolution on reading. Clearly therefore a head which uses a common gap for 'read' and 'write' is a compromise design.

3.2. The Merits of 'Head per Track' Configuration and Actuator Mechanism

This is a consideration where economics are possibly the dominant factor. If a file has one head per track, access from one track to the next can be accomplished in the switching time of the head circuits or at the very worst in the latency period of the disk rotation. If a single head and a head positioner is used, then the movement time of the positioner is the major term in the access time. Clearly there is an economic equation, the cost of a positioner mechanism and its control on the one side being considered against the cost of many extra heads plus a more complex switching network on the other. The improvement in performance can be considered in this context, and

an optimum specification for any particular set of circumstances can be decided.

The economics would be more favourable if it were possible to produce really cheap heads in, for example, a printed circuit or deposited ferrite form. The ideal would be a strip of such devices which could be 'bolted on'. However, considerations of head flying at present demand tight tolerances on disk flatness only in the circumferential direction. Flatness in the radial direction can be far less, because as the head pad is only about 0.5 in wide, it only interfaces with the disk in discrete 0.5 in bands, and the flatness of each individual band is the important criterion.

If a rigid head strip were mounted across the whole disk surface, then the whole surface would need to be of extreme radial as well as circumferential flatness. In achieving this type of flatness, the cost of manufacturing the substrates, and also of mounting them would greatly increase. Again a compromise is indicated.

3.3. Considerations of Track-width, Packing Density, Surface Speed and Head Output

The output from a head is proportional to the rate of cutting the lines of flux, i.e. $e_0 \propto n \cdot d\phi/dt$, where n is the number of turns. The resonant frequency of the head is inversely proportional to the number of turns wound on the head core, i.e. $f_r \propto 1/n$.

Now suppose that the packing density on a track is doubled. Ignoring pulse crowding effects, if the surface velocity is kept constant, the time it takes a recorded transition to pass the head will be the same, i.e. $d\phi/dt$ has not changed.

The number of transitions has however been doubled, and so the data rate which the head is handling has been doubled. Hence, in order to maintain the same relationship between the resonant frequency and the data rate, the resonant frequency must be doubled. Therefore, the number of turns must be halved, halving the output. This argument, although not strictly true, gives an appreciation of the problems involved and leads to the following relationship.

$$\frac{e_1}{e_2} = \frac{(\text{packing density})_2}{(\text{packing density})_1} = \frac{PD_2}{PD_1}$$

Similar reasoning shows that it is not possible to regain the original output by increasing $d\phi/dt$ (increasing rotational speed). Furthermore, altering this parameter would cause the flying height of the heads to increase, and so the previous magnetic considerations would be completely altered. The resonant frequency of the head again affects the track

width in accordance with the relation

$$L = K_1 n^2 w$$

where w is the width of the track. It can be shown (see Appendix 2) that the relationship between output, packing density and track-width lies between

$$\frac{e_1}{e_2} = \frac{PD_2 w_1}{PD_1 w_2}$$

and

$$\frac{e_1}{e_2} = \frac{PD_2}{PD_1} \sqrt{\frac{w_1}{w_2}}$$

The latter is true only in the ideal case.

As the track-width diminishes and frequency increases, it becomes more difficult to design a head geometry which meets the ideal case.

3.4. Noise Problems

Having considered the effect of various parameters on output amplitude, it is pertinent to consider signal/noise ratios. As the track width reduces, the output tends to reduce directly as the reduction in track width. However, noise power is proportional to track-width, therefore the noise voltage will only diminish by the square-root of the ratio. Hence the signal/disk noise ratio worsens as track-widths reduce.


It is also pertinent to note that the heads are at present switched by diodes, which may be passing a noise current of the order of 0.1 μA into an amplifier of 1 k Ω input impedance.

It has been shown in the previous Section how the output voltage will inevitably diminish as packing densities increase and track widths decrease, and the resultant noise voltage of 100 μV is quite large in terms of the signal output under consideration.

3.5. Head Skew and Positioner Accuracy

Attention has been drawn to the fact that there is a very close relationship between mechanical and electronic design. A good example of this is a consideration of the relationship of head skew and positioner accuracy. The previous discussion on track-width and signal output assumed that the 'read' head was always situated accurately over the written track. If this is not the case, as may well occur, then there will either be a loss of signal or a pick-up of unwanted signal from an adjacent track, depending upon the particular head geometry used and the amount of the off-track error. This will give rise to timing errors in the data chain.

Even if this is minimized, there is a condition of forced loss of signal brought about by the gap between the erase and the read/write head which is common in

AREAS INDICATED THUS  etc. SHOW EXAMPLES OF PARTICULAR INTEREST.

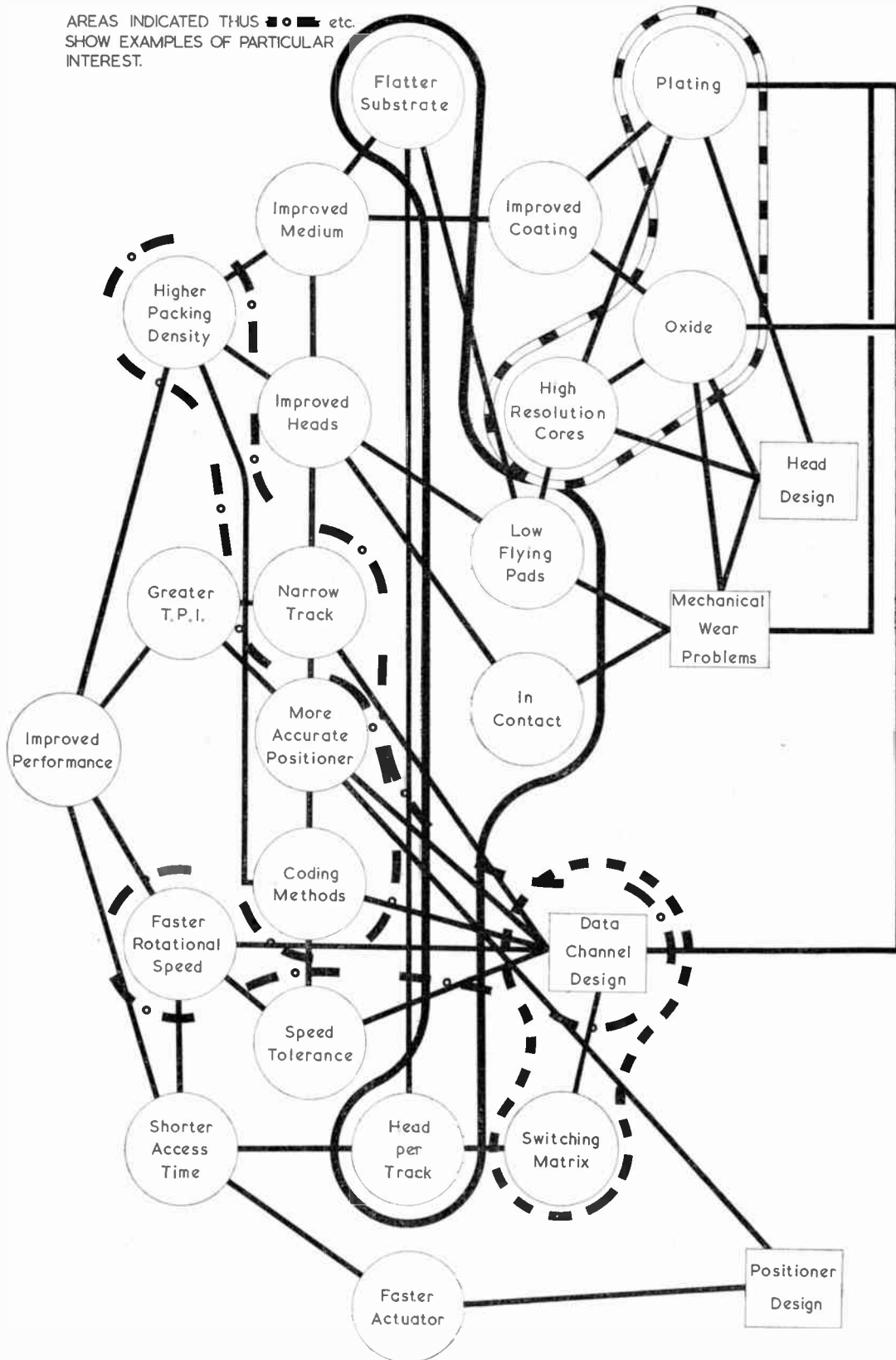


Fig. 2. Interrelationship of various parameters affecting the four main engineering parameters.

many head designs. It can be seen (Fig. 3) that even when head adjustment is optimized, the head can only be correctly aligned at one track. This is improved greatly by the 'straddle erase' heads which are being introduced with the gaps in very close proximity.

There still remains, however, the accuracy of the setting up tools plus the accuracy of the field engineer. These two factors combine with the tolerance of the head gap width to give a practical limit to the number of tracks per inch which can be achieved.

4. User Aspects

In future, the overall performance of disk files can be expected to improve, and some of the engineering problems which will be encountered have been discussed. The requirement for exchangeability of written packs is a dominant factor in all these problems, particularly those concerned with track position tolerances.

As the performance of disk files in general improves and particularly as the on-line data storage capacity increases, it must be questioned whether the requirement for interchangeability justifies the additional engineering constraints.

It has been shown that head to disk separation is a critical factor in all aspects of the magnetic design, and a great deal of research effort has been devoted to minimizing this. As research achieves this, and as track widths become narrower, one of the greater practical problems met in the field to date assumes even more significance. This is the effect of the ingress of dirt into the disk pack area, and this occurs particularly each time a pack is changed or a disk file opened. It is surely significant to end this paper by noting that there were two main design changes involved in a recent doubling of the packing density of a disk file. The first was the development of completely new heads, the second was the introduction of a new air filtering system.

5. Acknowledgments

The views expressed in this paper are the result of many discussions with colleagues in Data Recording Instrument Co. Ltd. and International Computers Ltd. as part of an exercise to formulate and quantify the limiting factors in disk file design.

6. References

1. Bonyhard, P. I., Davies, A. V. and Middleton, B. K., 'A theory of digital magnetic recording on metallic films', *Trans. Inst. Elect. Electronics Engrs on Magnetics*, MAG-2, pp. 1-5, March 1966.
2. Daniel, E. D., 'The influence of some head and tape constants on the signal recorded on a magnetic tape', *Proc. Instn Elect. Engrs*, 100, pp. 168-75, May 1953. (I.E.E. Paper No. 1507.)

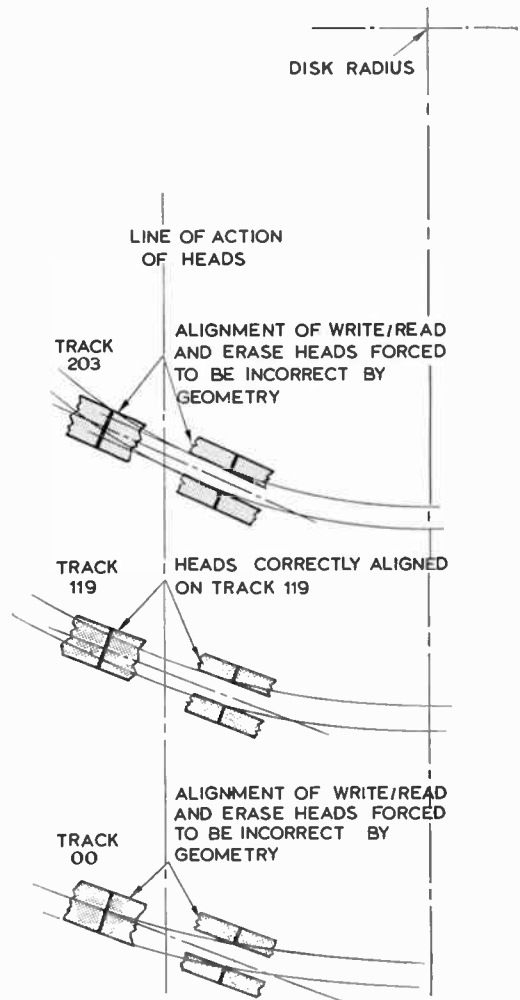


Fig. 3. 'Write/read' head misalignment due to a distance between 'erase' and 'write/read' heads.

3. Speliotis, D. E., Morrison, J. R. and Judge, J. S., 'A correlation between magnetic properties and recording behavior in metallic chemically deposited surfaces', *Trans. I.E.E.E. on Magnetics*, MAG-1, pp. 348-51, December 1965.

7. Appendix 1

Calculations of Output for Equivalent Pulse Widths on Cobalt and Oxide Media

The quantity *a* which is proportional to transition width in the isolated pulse condition gives an indication of the goodness of the medium as far as resolution is concerned.

$$a = \frac{B_s C}{2\pi H_c}$$

where *C* is the coating thickness.

The ratio of coating thickness which will give the same resolution can be found by using typical values

of B_r and H_c for oxide and cobalt.

$$\frac{1000C_{\text{oxide}}}{2 \times 250} = \frac{10\,000C_{\text{cobalt}}}{2 \times 750}$$

$$\frac{C_{\text{oxide}}}{C_{\text{cobalt}}} = \frac{10}{3}$$

Using thickness of value $3C$ and $10C$ respectively, it is now possible to find the outputs for the same potential resolution.¹

$$e_0 = Knvw \frac{B_r C}{d+a}$$

i.e.

$$e_0 \propto \frac{B_r C}{d+a}$$

where d is the head to medium separation.

Three cases can be considered:

(i) *Large separation* $d \gg a$

$$e_0 \propto \frac{B_r C}{d}$$

(ii) *Small separation* $d \ll a$

$$e_0 \propto \frac{B_r C}{a} \text{ i.e. proportional to } H_c$$

(iii) *Practical cases*³ The original formula put forward in this paper for the practical case was

$$e_0 \propto \sqrt{B_r C H_c}$$

although this was subsequently somewhat modified by Dr. Speliotis in a discussion at the 1969 Intermag Conference.

In all these cases, if the relevant parameters and thickness are fed into the formula the output from the cobalt is at least 3 times greater than that from the oxide, i.e. an expected improvement in the signal/noise ratio of 10 dB.

8. Appendix 2

Detailed Considerations of Track Width, Packing Density Surface Speed and Head Output

From the resonance considerations of the head

$$r \propto \frac{1}{\sqrt{L}}$$

$$L \propto n^2$$

Therefore $f_r \propto \frac{1}{n}$

If the packing density on the track is doubled, and the surface velocity kept constant, the data rate is doubled.

Therefore f_r must be doubled, i.e. n must be halved. This will reduce the output by a factor of 2. Now suppose that in order to compensate for this, the surface speed is doubled in order to increase $d\phi/dt$. In this case again the data rate is doubled and as before n must be halved.

Therefore, $\frac{e_1}{e_2} = \frac{PD_2}{PD_1}$

Therefore no increase in output can be obtained by increasing the surface speed.

As the track-width is reduced, so the output is reduced in proportion, so that

$$\frac{e_1}{e_2} = \frac{w_1}{w_2}$$

However, reducing the track width is reducing the width of the ferrite core in the head and so affecting the head inductance. It can be shown that if flux leakage is ignored

$$L = K_1 n^2 w$$

To maintain a constant inductance as the track width reduces

$$n_1^2 w_1 = n_2^2 w_2$$

i.e.

$$\frac{n_1}{n_2} = \sqrt{\frac{w_2}{w_1}}$$

This shows as a compensating term indicating the amount by which the turns ratio may be increased. To be able to make use of this, however, there are many problems in designing the actual head geometry.

Manuscript first received by the Institution on 21st May 1969 and in final form on 23rd June 1969. (Paper No. 1283/Comp. 123.)

Conference on Electronic Engineering in Ocean Technology*

Three years ago, the I.E.R.E. organized the first Conference to be held in Europe on 'Electronic Engineering in Oceanography'. Since 1966, many of the developments discussed at that Conference have found application, as have others in related fields, and the discipline appropriately called 'Ocean Technology' has become established. We are now coming to know more about all aspects of the oceans and in the search for knowledge, electronic techniques have played and will continue to play a vital part.

In order to provide an opportunity for assessing some of the new techniques which are coming forward, the I.E.R.E., with the association of the Institution of Electrical Engineers, is to hold a second conference—'Electronic Engineering in Ocean Technology'—which has the theme of 'The gathering, transmission, processing and display of information from the sea'.

The Conference will be held at the University College of Swansea, and will last four days, from Monday, 21st September to Thursday, 24th September, 1970. The excellent facilities of the College will provide the opportunity of residential accommodation for those attending the Conference and it is planned to make a feature of demonstrations.

The Organizing Committee, which is under the Chairmanship of Professor D. G. Tucker and includes members from leading U.K. organizations concerned with ocean technology, is now inviting papers for the Conference.

These should describe new work in the following fields:

Sensors and Recording

For instance, measurement of salinity, temperature, currents, tides and ship movement; visual and acoustic observation (e.g. viewing systems for low level illumination and acoustic holography).

Communications and Telemetering

For instance, adaptive systems, channel characteristics and signal design for underwater transmission, processing of oxyhelium speech.

Signal and Data Processing

For instance, the use of shipborne computers and new electronic developments in special signal and data processing.

Display Systems } which have specific application within
Power Sources } the theme of the Conference.

As well as full length papers of up to 6000 words the Committee will welcome shorter contributions in the region of 1500–2000 words in length. Papers will be pre-printed and will therefore be required in final form by 15th May, 1970.

Synopses of proposed contributions are invited and should be sent as soon as possible, preferably not later than 31st December, 1969 to:

The Secretary, Organizing Committee for the Conference on 'Electronic Engineering in Ocean Technology', The Institution of Electronic and Radio Engineers, 9 Bedford Square, London, W.C.1.

It is essential that a synopsis should be long enough to enable the Organizing Committee to assess fully the proposed content of the paper and thus, typically, should be about 200 words in length.

STANDARD FREQUENCY TRANSMISSIONS—September 1969

(Communication from the National Physical Laboratory)

| Sept. 1969 | Deviation from nominal frequency in parts in 10 ¹⁰ (24-hour mean centred on 0300 UT) | | | Relative phase readings in microseconds N.P.L.—Station (Readings at 1500 UT) | | Sept. 1969 | Deviation from nominal frequency in parts in 10 ¹⁰ (24-hour mean centred on 0300 UT) | | | Relative phase readings in microseconds N.P.L.—Station (Readings at 1500 UT) | |
|------------|---|------------|-------------------|--|-------------|------------|---|------------|-------------------|--|-------------|
| | GBR 16 kHz | MSF 60 kHz | Droitwich 200 kHz | *GBR 16 kHz | †MSF 60 kHz | | GBR 16 kHz | MSF 60 kHz | Droitwich 200 kHz | *GBR 16 kHz | †MSF 60 kHz |
| 1 | -300.0 | -0.1 | +0.1 | 554 | 479.1 | 17 | -300.2 | 0 | 0 | 568 | 484.3 |
| 2 | -300.0 | 0 | +0.1 | 554 | 478.7 | 18 | -300.1 | 0 | 0 | 569 | 486.5 |
| 3 | -300.1 | 0 | +0.1 | 555 | 479.1 | 19 | -300.0 | -0.1 | +0.1 | 569 | 484.1 |
| 4 | -300.1 | -0.1 | +0.1 | 557 | 476.3 | 20 | -300.0 | 0 | +0.1 | 569 | 484.3 |
| 5 | -300.0 | 0 | +0.1 | 565 | 476.3 | 21 | -300.1 | -0.1 | 0 | 570 | 485.1 |
| 6 | -299.9 | 0 | +0.1 | 564 | 476.3 | 22 | — | 0 | +0.1 | — | 485.2 |
| 7 | -299.9 | 0 | +0.1 | 563 | 476.3 | 23 | — | 0 | +0.1 | — | 485.2 |
| 8 | -300.1 | -0.1 | +0.1 | 564 | 475.2 | 24 | — | 0 | +0.1 | — | 485.2 |
| 9 | -300.1 | -0.1 | +0.1 | 565 | 476.7 | 25 | — | 0 | +0.1 | — | 485.2 |
| 10 | -300.0 | -0.1 | 0 | 567 | 477.4 | 26 | — | 0 | +0.1 | — | 485.6 |
| 11 | -300.0 | 0 | 0 | 567 | 477.4 | 27 | — | 0 | +0.1 | — | 485.8 |
| 12 | -300.0 | 0 | 0 | 567 | 477.5 | 28 | -300.1 | -0.1 | +0.1 | 571 | 486.6 |
| 13 | -300.1 | -0.1 | 0 | 568 | 479.7 | 29 | -300.1 | -0.1 | +0.1 | 572 | 487.2 |
| 14 | -300.0 | -0.1 | 0 | 568 | 480.3 | 30 | -300.1 | — | +0.1 | 573 | — |
| 15 | -299.9 | 0 | 0 | 567 | 479.8 | | | | | | |
| 16 | -299.9 | -0.1 | 0 | 566 | 484.0 | | | | | | |

All measurements in terms of H.P. Caesium Standard No. 334, which agrees with the N.P.L. Caesium Standard to 1 part in 10¹¹.

* Relative to UTC Scale; (UTC_{NPL} - Station) = + 500 at 1500 UT 31st December 1968.

† Relative to AT Scale; (AT_{NPL} - Station) = + 468.6 at 1500 UT 31st December 1968.

Optimum Transfer Functions for Feedback Control Systems with Plant Input Saturation

By

D. R. TOWILL,
M.Sc., C.Eng., M.I.Mech.E.†

Summary: A method for the determination of optimum transfer functions for systems controlling plant with input saturation is presented. The procedure is demonstrated on a specific plant well studied in the literature, and the general conclusions verified by simulator and digital computer results. It is established that the desired system bandwidth should be obtained with a dominant system transfer function of low $\frac{\prod |\text{system poles}|}{\prod |\text{system zeros}|}$. The order of the dominant system transfer function is usually physically obvious from the asymptotic Bode plot, and as in modern control theory, the Butterworth array is found to be particularly useful. Provided the right order of transfer function is chosen, the actual pole-zero geometry is not critical. Such a design is optimal with regard to small transient excursions into the non-linear mode, and with regard to the minimization of noise effects on the plant input level, yet simultaneously meets the customary deterministic performance criteria.

In the comparison of series compensation designs, it is shown experimentally that some marginal reduction of sensitivity is achieved by reducing the order of the dominant transfer function, but this reduction is easily offset by the use of feedback compensation. Far-off poles added to filter high frequency noise are found to have little effect on the choice of optimum transfer function. Velocity constant requirements are met by the addition of integrating dipoles. A useful theorem providing a quantitative measure of the deterioration in non-linear performance caused by the choice of excessive bandwidth is presented and is shown to result in a power law.

By expressing the conclusions as s plane constraints for the dominant poles and zeros, the results can be easily incorporated in any standard design procedure, manual or automated, and should lead to considerable reduction in the time required to obtain a satisfactory design.

List of Symbols

| | | | |
|-------------|---|------------|---|
| θ_i | system command variable | k | Lagrange multiplier |
| θ_o | system controlled variable | $s =$ | d/dt |
| θ, E | system error | ω_B | system bandwidth, the highest excitation frequency at which the amplitude ratio is $1/\sqrt{2}$ |
| θ_p | plant input stimulus | λ | bandwidth ratio |
| U | disturbance appearing at plant input | ω | stimulus excitation frequency |
| $T(s)$ | system transfer function | t | time |
| $G_c(s)$ | feedforward compensation transfer function | ζ | damping ratio of complex system poles |
| $H(s)$ | feedback path transfer function | i.d.s. = | $\int_0^\infty [u(t) - u'(t)]^2 dt$ |
| S_K^T | classical sensitivity function | n | number of effective system poles |
| $u(t)$ | system step function response with nominal plant dynamics | q | number of effective system zeros |
| $u'(t)$ | system step function response with changed plant dynamics | Π | product symbol |
| K | plant gain | $J =$ | $\int_0^\infty [\theta(t)^2 + k^2 \theta_p(t)^2] dt$ |
| | | * | denotes convolution |

† University of Wales Institute of Science and Technology, Department of Mechanical and Production Engineering, Cardiff.

1. Introduction

This paper is concerned with the optimization of feedback control systems which are designed to operate in the linear mode for the major part of their working life. Optimum in this sense is taken to mean that in the linear mode the 'following' characteristics are adequately met, whilst the effect of disturbances, noise, changes in load dynamics, and physical limitations are kept within acceptable bounds.

In the linear synthesis of such systems, the s plane reveals in general:

- (a) there are many different system pole-zero arrays which adequately satisfy the usual requirements for following the input command signal;
- (b) some system pole-zero arrays will reject disturbances and noise significantly better than others;
- (c) some system pole-zero arrays will result in much better performance in the presence of signal saturation, thus making best use of equipment limitations, and are optimal in this sense;
- (d) for such an optimal system pole-zero array there will exist a canonical configuration which is optimal with regard to sensitivity considerations.

The above four problem areas reveal the many considerations which have to be balanced against each other in the successful evolution of a realistic control system, and partly account for the *ad hoc* approach inherent in control system design. It is probable that the above list is in the sequence of increasing complexity, for example many undergraduate texts concentrate on area (a), whereas few texts at any level consider areas (c) and (d).

2. Current Research into Optimization of Feedback Control System Performance

Current research is extremely diversified. Historically, effort was concentrated on the determination of transfer functions of a given order which resulted in adequate response of the system to the command signal.¹ This approach has been formalized^{2, 3} by the use of performance indices of the form

$$\int_0^{\infty} f(\text{error, time}) \cdot dt$$

and is indeed still of major interest since algorithms are readily developed for incorporation in digital computer hill climbing procedures.⁴⁻⁷ An interesting development is the computerized hill-climbing procedure of Davison,⁸ in which the system poles are driven as far as possible to the left subject to constraints which may be added at the designer's discretion.

An extremely practical method of optimization involves the judicious choice of a small number of well-

established compensating elements the numerical values of which are optimized on the basis of analogue model studies which may implicitly consider many performance criteria simultaneously.⁹ Clearly much reduction in design time is achieved by intuition based on previous experience, and there are advantages in the designer remaining in close contact throughout the synthesis procedure.

A number of recent contributions have automated traditional design procedure based on the domination of system response by a complex pole pair,¹⁰ some of this work utilizing the parameter plane concept.¹¹⁻¹³ Sensitivity considerations have been included in this work, but are restricted to the movements of dominant poles. An attraction of the parameter plane approach is the inherent emphasis on the compromise between two adjustable system parameters (such as rate feedback and acceleration feedback coefficients) and various performance criteria of interest. As yet, there appears to have been no attempt to shape the system configuration to take account of saturation limitations likely to be met in practice.

Finally the modern control theory has developed design procedures based on the minimization of a quadratic performance index, the components of the index being system error and plant input.¹⁴⁻¹⁸ Similar considerations can also be given to the use of quadratic indexes for imposing sensitivity constraints.¹⁹ Modern control theory is ideally suited for computerization, but perhaps a major benefit of its existence is that due emphasis is at last placed on plant input, a very necessary consideration in the presence of plant input saturation. Conceptually, there is much in common between the customary modern control theory based on the response of systems to deterministic inputs, and the earlier work of Newton.²⁰

An interesting feature of optimization using modern control theory is the regularity with which the optimal solution is found to approximate to the Butterworth filter of the same order as the effective part of the plant dynamics, that is, the plant poles and zeros between zero frequency and some small multiple of system bandwidth. A full description of the properties of Butterworth filters is to be found in Reference 2.

3. Contribution of this Paper

Comparatively little effort has been directed at simultaneously understanding the effect of choosing a particular form of transfer function on two or more of the various problem areas defined in Section 1. By considering all these problem areas in turn, this paper establishes the conditions under which particular transfer functions are optimum, and by implication defines the system limitations and advantages of the second-order dominant mode philosophy so implicit

in much of the literature. In particular, a simple measure of the effectiveness of the design in the presence of plant input saturation will be developed in terms of system pole zero geometry and presented in a form for ready incorporation in many of the design techniques referred to in the literature survey.

The paper first considers in detail the optimum control of the plant studied in a paper which stimulated the present author to undertake the research work described herein. Specific conclusions are then generalized, and lead to a fundamental optimization philosophy.

4. Optimization of Performance in the Presence of Plant Input Saturation

Following the work of Tustin *et al.*,²¹ it is considered that optimization of performance in the presence of plant saturation should be considered under two separate headings:

- (a) deterioration of performance, and possible instability due to step-like command signals of large magnitude;
- (b) deterioration of performance due to high frequency noise saturating the plant, thus derating the response to the command signal.

In fact these two requirements lead to identical constraints on the system pole-zero arrays. However, reference²¹ suggests that if excessively large step inputs are likely, such as during synchronization in fire control systems, special techniques such as the dual-mode system of Bowler²² or the non-linear shaping approach of Vaughan and Foster,²³ are called for. Thus a unique mode system should either avoid entering into the saturated region, or should perform well for small excursions into the saturated region. For applications where the noise rejection requirement is unimportant, there may be some advantage in utilizing the full torque available irrespective of the system error, thus leading to various types of switching systems which can then be optimized for specific command signals.²⁴

Prior to the examination of a suitable family of system transfer functions for optimality in the presence of plant input saturation, the family is determined on a basis of identical bandwidth and a low overshoot step response. This is a common, but not infallible comparison, since bandwidth, whilst a quantitative measure of the transition region between command signal following and noise signal rejection, is not a direct measure of the ability of the system to follow the command signal. This is because in a well designed system, command signal frequencies occur in the spectrum at a range corresponding to a few percent of bandwidth.²¹ A better measure of command signal following is given by the error constants particularly

the velocity constant, hence when considered advisable for generalization of the conclusions in this paper, integrating dipoles are added to the system transfer functions so that identical velocity constants and bandwidths are achieved simultaneously.

We shall assume that noise occurs at the system input, as frequently happens in fire control systems, and has high frequency components. Little loss in generality results from the assumption of white noise which is often a realistic approximation to experimental data. In contrast, it is usually found that disturbances arising at the plant input, such as wind loadings etc., are of very low frequency.

5. Fundamental Transfer Functions

For a unity feedback control system, if we consider three basic transfer functions, the nominal plant transfer function $P(s)$, the designed system transfer function $T(s)$, and the necessary compensation transfer function $G_c(s)$, then we may write the following transfer functions, which are direct measures of the goodness of system design,

$$\frac{\theta}{\theta_i} = 1 - (\text{system transfer function}) \quad \dots\dots(1)$$

$$\frac{\theta_p}{\theta_i} = \frac{\text{system transfer function}}{\text{plant transfer function}} \quad \dots\dots(2)$$

$$\frac{\theta_0}{U} = \frac{\text{system transfer function}}{\text{compensation transfer function}} \quad \dots\dots(3)$$

$$S_K^T = \frac{(\partial T/T)}{(\partial K/K)} = 1 - (\text{system transfer function}) \quad \dots\dots(4)$$

It is immediately apparent from equations (1) to (4) that the structure of the unity feedback system imposes a severe constraint on the goodness of the system because the sensitivity and load disturbance rejection capabilities are completely defined by the choice of system transfer function. Horowitz²⁵ argues that these restrictions are so severe that the use of unity feedback designs can result in performance which is better than a non-feedback design only with regard to disturbance rejection. Fortunately, the use of more general forms of feedback (such as velocity and acceleration feedback in addition to position feedback, for example) means that equations (3) and (4) no longer apply. The system transfer function may then be chosen to meet the low-frequency command signal requirement and constrain the work done by the plant whilst some form of feedback compensation is then used if necessary to achieve the required level of sensitivity reduction and disturbance rejection.

To illustrate the argument, consider the three structurally different designs shown in Fig. 1. All

three system transfer functions are identical, hence the low-frequency error characteristics are identical and the noise filtering characteristics are identical. However, the non-feedback design, not sensing load disturbances, is clearly inferior in this respect. Reference 25 uses the logarithmic gain of the classical

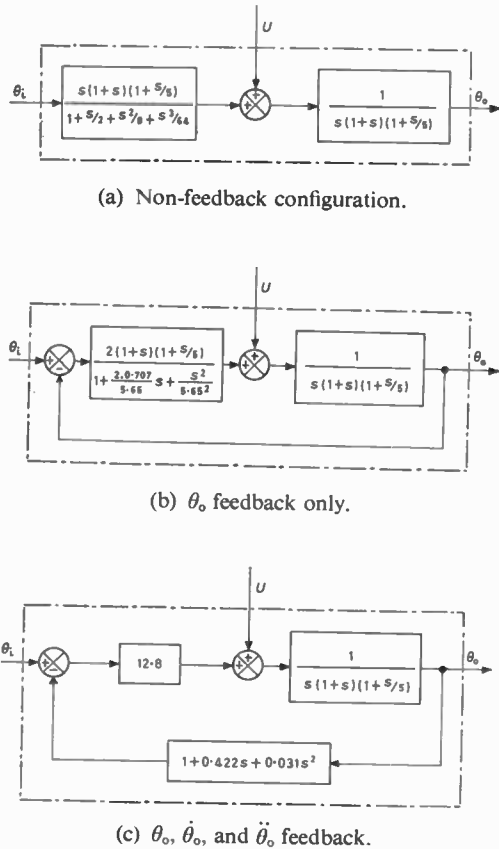


Fig. 1. Three structurally different designs with system transfer function

$$T(s) = \frac{1}{1+s/2+s^2/8+s^3/64}$$

sensitivity function to compare the usefulness of such designs, as shown in Fig. 2. The frequency domain integral of the log of classical sensitivity function is zero for both for non-feedback and unity feedback designs, which is the proof offered by Horowitz on the need to carefully consider the structure of the design if feedback is to be used intelligently. It may be seen by inspection that the classical sensitivity function for the more general feedback design is significantly less than for the unity feedback design for a reasonable frequency range, thus offering considerable potential for sensitivity reduction.

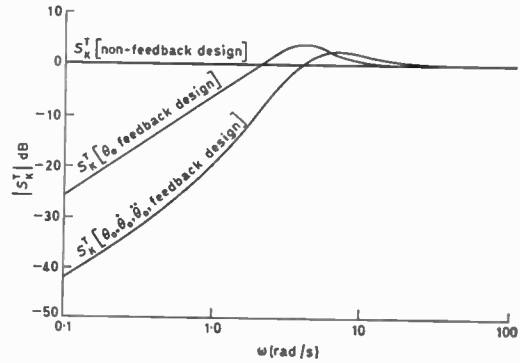


Fig. 2. Classical sensitivity function comparison of the three structurally different designs with

$$T(s) = \frac{1}{1+s/2+s^2/8+s^3/64}$$

This paper is directed at the properties of the system transfer function, but it was felt necessary to show the effect of choice of structure for a given system transfer function on sensitivity, since manifestly it is of fundamental importance to the system designer.

6. Control of Specific Plant Dynamics

To explore the meaning of optimization of feedback control system performance, we shall consider the control of the plant dynamics defined by

$$P(s) = \frac{1}{s(1+s)(1+s/5)} \dots\dots(5)$$

which is the plant extensively studied in optimization procedures using $\int_0^\infty tE^2 dt$ as the performance index.⁴

For the purpose of this paper, the system is required to have a bandwidth of 4 rad/s (deliberately chosen so that all three plant poles are significant), and an overshoot to the unit step in the region of 5 to 10%. If it is assumed *pro tem.* that either cancellation compensation or feedback compensation is permitted, then the system will be third-order with no system zeros. Using the coefficient plane²⁶ it is a simple matter to evaluate a set of system pole-zero arrays which satisfy this specification, and a suitable locus is shown in Fig. 3.

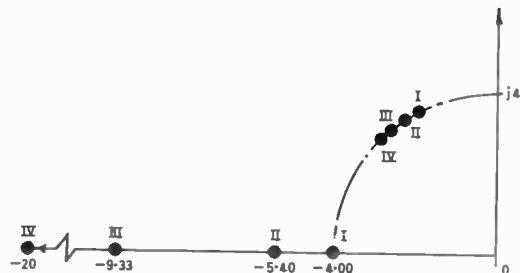


Fig. 3. Some system pole configurations resulting in 4 rad/s bandwidth and a satisfactory transient response.

The limiting designs shown are the third-order and second-order Butterworth; the choice of third pole for the latter at $s = -20$ (i.e. five times bandwidth) is somewhat arbitrary. These two designs will be studied in detail in the quest for optimality.

Using coefficient matching, and assuming cancellation compensation, the third-order Butterworth design requires that

$$G_c(s) = \frac{2.0(1+s)(1+s/5)}{1 + \frac{2.0 \cdot 707}{5.65} s + \frac{s^2}{(5.65)^2}} \quad \dots\dots(6)$$

and the second-order Butterworth design requires that

$$G_c(s) = \frac{2.47(1+s)(1+s/5)}{(1+s/18.7)(1+s/6.94)} \quad \dots\dots(7)$$

For generality a further design using non-cancellation single-stage lead compensation is also considered. After some trial and error typifying the frequency domain procedure, it is found that a suitable compensator is

$$G_c(s) = \frac{1.84(1+s/0.81)}{(1+s/8.1)} \quad \dots\dots(8)$$

The non-cancellation design results in a (1, 4) system with the fourth pole well to the left. Lightly damped complex poles are attenuated in their effect by the real pole near the origin, the real pole contribution in turn being limited by the real zero.

Important transfer functions resulting from these three designs are shown in Table 1.

7. Frequency Domain Properties of the Three Basic Designs

For the three designs of Table 1, all having acceptable (but not identically equal) transient responses, and similar bandwidths, it is important to compare the basic properties defined in equations (1) to (4). Frequency domain results are summarized in Fig. 4.

Primarily due to the higher feedforward gain, the second-order Butterworth design has a significantly smaller low-frequency error, rejects low-frequency load disturbances better, and does not drive the plant any harder at command signal frequencies. The latter result verifies the statement by Tustin *et al.*²¹ that increasing feedforward gain is unlikely to lead to plant saturation problems at command signal frequencies, because the error at these frequencies is correspondingly reduced. As expected, the rate of cut-off at bandwidth is smaller for the second-order Butterworth design, thus increasing error due to any noise in this region of the frequency spectrum. An interesting feature of Fig. 4 is the great similarity between the characteristics of the third-order Butterworth design

and the non-cancellation design. By coincidence they are also equally good in the presence of plant input saturation. Whilst many system pole-zero arrays would result in adequate agreement with regard to $T(s)$, the reason for the goodness of the non-cancellation design with regard to non-linear operation will not become apparent until later, but there is no doubt that for noise at frequencies above 1 Hz, the probability of plant input saturation occurring due to this source is greatly reduced by choice of the third-order Butterworth design or the non-cancellation design.

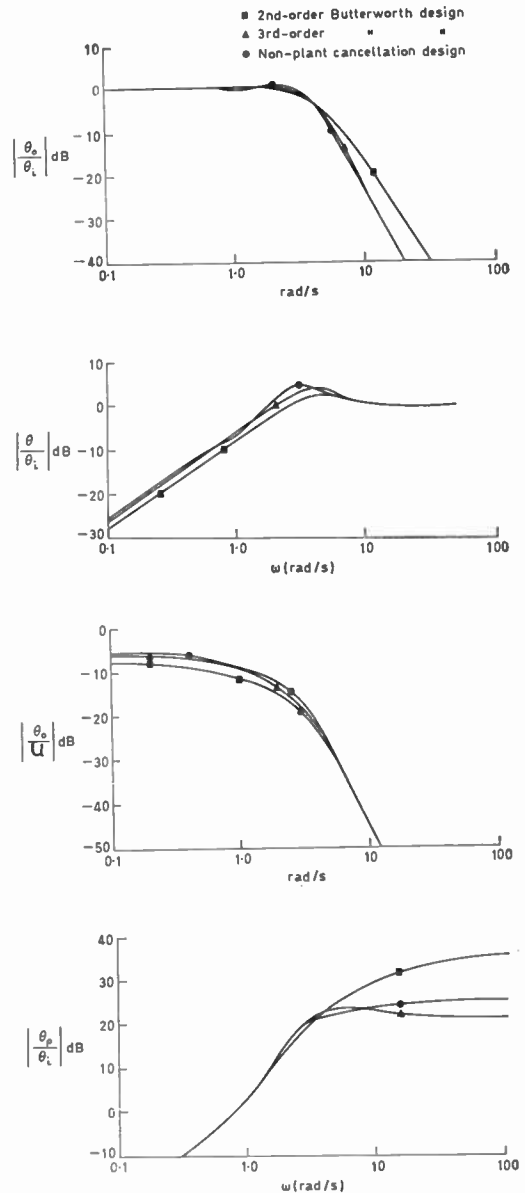


Fig. 4. Frequency response of important transfer functions for three designs with $\omega_B = 4$ rad/s and acceptable transient response.

Table 1
Important transfer functions for the three basic designs

| Transfer function | System | | |
|-----------------------------|--|--|---|
| | Approximate second-order Butterworth | Third-order Butterworth | Non-cancellation design |
| $\frac{\theta_o}{\theta_i}$ | $\frac{1}{\left(1 + \frac{s}{20}\right) \left(1 + \frac{2.0 \cdot 707}{4} s + \frac{s^2}{4^2}\right)}$ | $\frac{1}{\left(1 + \frac{s}{4}\right) \left(1 + \frac{2.0 \cdot 5}{4} s + \frac{s^2}{4^2}\right)}$ | $\frac{\left(1 + \frac{s}{0.81}\right)}{\left(1 + \frac{s}{0.75}\right) \left(1 + \frac{s}{10}\right) \left(1 + \frac{2.0 \cdot 48}{3 \cdot 17} s + \frac{s^2}{3 \cdot 17^2}\right)}$ |
| $\frac{\theta}{\theta_i}$ | $\frac{s \left(1 + \frac{s}{18.7}\right) \left(1 + \frac{s}{6.94}\right)}{2.47 \left(1 + \frac{2.0 \cdot 707}{4} s + \frac{s^2}{4^2}\right) \left(1 + \frac{s}{20}\right)}$ | $\frac{s \left(1 + \frac{2.0 \cdot 707}{5.65} s + \frac{s^2}{5.65^2}\right)}{2 \left(1 + \frac{2.0 \cdot 5}{4} s + \frac{s^2}{4^2}\right) \left(1 + \frac{s}{4}\right)}$ | $\frac{s(1+s) \left(1 + \frac{s}{5}\right) \left(1 + \frac{s}{8.1}\right)}{1.84 \left(1 + \frac{s}{0.75}\right) \left(1 + \frac{s}{10}\right) \left(1 + \frac{2.0 \cdot 48}{3 \cdot 17} s + \frac{s^2}{3 \cdot 17^2}\right)}$ |
| $\frac{\theta_o}{u}$ | $\frac{\left(1 + \frac{s}{18.7}\right) \left(1 + \frac{s}{6.94}\right)}{2.47 \left(1 + \frac{s}{20}\right) \left(1 + \frac{2.0 \cdot 707}{4} s + \frac{s^2}{4^2}\right) (1+s) \left(1 + \frac{s}{5}\right)}$ | $\frac{1 + \frac{2.0 \cdot 707}{5.65} s + \frac{s^2}{5.65^2}}{2 \left(1 + \frac{s}{4}\right) \left(1 + \frac{2.0 \cdot 5}{4} s + \frac{s^2}{4^2}\right) (1+s) \left(1 + \frac{s}{5}\right)}$ | $\frac{\left(1 + \frac{s}{8.1}\right)}{1.84 \left(1 + \frac{s}{0.75}\right) \left(1 + \frac{s}{10}\right) \left(1 + \frac{2.0 \cdot 48}{3 \cdot 17} s + \frac{s^2}{3 \cdot 17^2}\right)}$ |
| $\frac{\theta_p}{\theta_i}$ | $\frac{s(1+s) \left(1 + \frac{s}{5}\right)}{\left(1 + \frac{2.0 \cdot 707}{4} s + \frac{s^2}{4^2}\right) \left(1 + \frac{s}{20}\right)}$ | $\frac{s(1+s) \left(1 + \frac{s}{5}\right)}{\left(1 + \frac{2.0 \cdot 5}{4} s + \frac{s^2}{4^2}\right) \left(1 + \frac{s}{4}\right)}$ | $\frac{s(1+s) \left(1 + \frac{s}{5}\right) \left(1 + \frac{s}{0.81}\right)}{\left(1 + \frac{s}{0.75}\right) \left(1 + \frac{s}{10}\right) \left(1 + \frac{2.0 \cdot 48}{3 \cdot 17} s + \frac{s^2}{3 \cdot 17^2}\right)}$ |

If an integrating dipole is added to the third-order Butterworth design to increase the velocity constant to 2.47, thus providing a more competitive system, analysis shows that restriction of the dipole contribution to percentage overshoot to the traditional 5% results in reduction at *very low frequency only*, the effect disappearing at about 2% bandwidth. However the low-frequency error characteristics can be further improved if the 5% dipole dispersion is allowed to increase, which in the present author's opinion would not be unreasonable since the basic third-order Butterworth design has a low overshoot.

8. Sensitivity to Changes in Plant Dynamics

For a series-compensated system, it is well known that the classical sensitivity function is equal to the error/input transfer function, that is,

$$\left[\frac{\partial T/T}{\partial K/K} \right] = \frac{\theta}{\theta_i} \dots\dots(9)$$

which immediately places a restriction on sensitivity if the system transfer function is fixed. In the systems considered in Section 6 the transfer functions are not identically equal, and therefore it is possible that some marginal improvement may be obtained by opting for a particular design. Because of equation (9) it follows that the classical sensitivity functions are immediately available in Fig. 4, and favours at low frequency the second order Butterworth design. However, a better understanding of the relative merits (if any) with regard to sensitivity reduction is achieved by consideration of the sensitivity of the step-function response

$$\frac{\partial u(t)}{\partial K/K}, \text{ where } \frac{\partial u(t)}{\partial K/K} = [h(t)] * [1 - u(t)] \dots\dots(10)$$

Using the method of reference 27, equation (10) has been evaluated using a standard digital computer program, the results being shown in Fig. 5. The plots of $[u(t) + \partial u(t)/\partial K/K]$ are an aid to interpretation of the sensitivity functions, and are the transient responses that would be obtained if the plant gain were increased by 100%, and the sensitivity function were linear for such a large change in plant dynamics. It can be seen that the second-order Butterworth design is slightly less sensitive in the region of maximum percentage overshoot and, furthermore, addition of an integrating dipole to the third-order Butterworth design as suggested in Section 5 as a means of increasing the velocity constant makes very little difference to the step response sensitivity properties of the design.

To generalize this conclusion, a number of analogue computer experiments were carried out to determine the actual variations in step response resulting from changes in plant gain and two plant time-constants. Typical results are shown in Fig. 6. These experimental

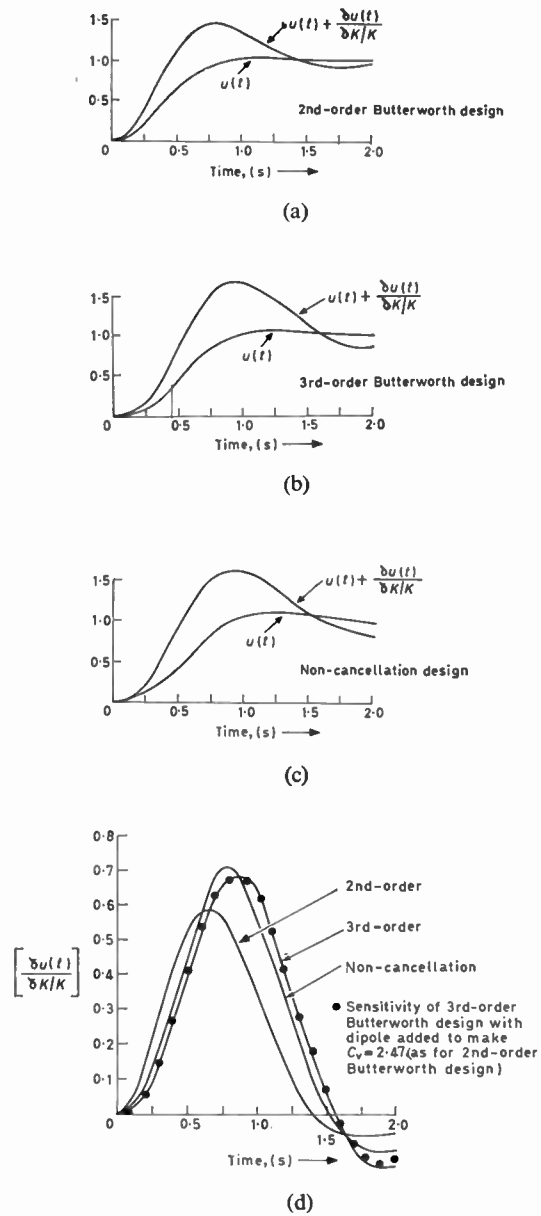


Fig. 5. Step response sensitivity ($\omega_B = 4$).

step responses were analysed numerically using the 'integral of deviation squared' index, defined as follows,

$$i.d.s. = \int_0^\infty [u(t) - u'(t)]^2 dt \dots\dots(11)$$

where $u(t)$ is the nominal response for a particular design and $u'(t)$ is the actual step response obtained when controlling non-nominal plant dynamics. Such an index has been used in the same context in Reference 28, but it must be emphasized that use of the index is

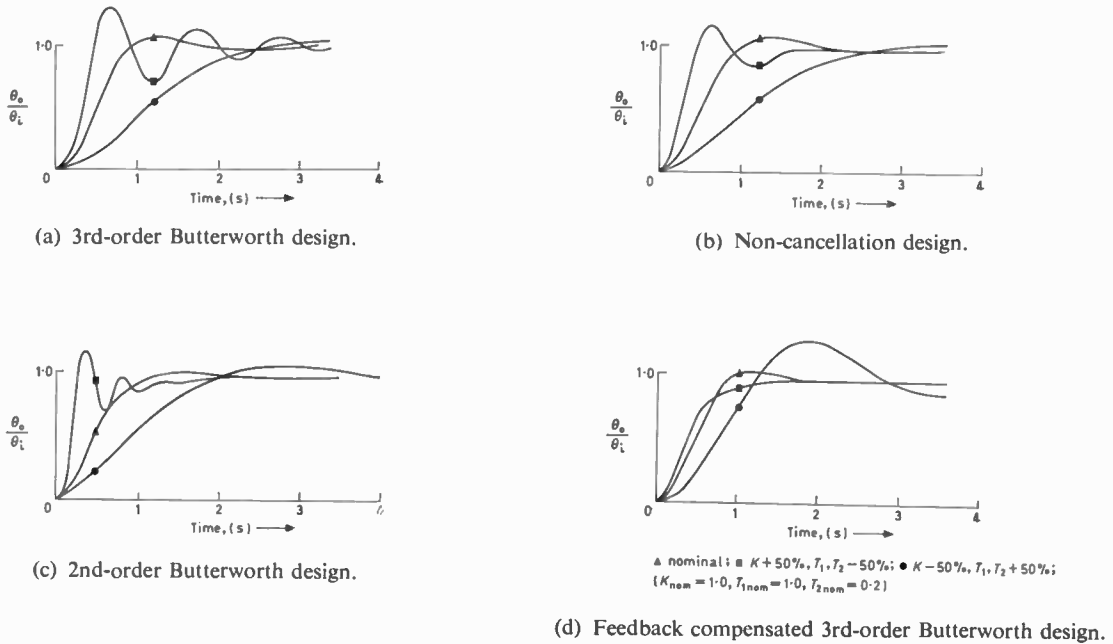


Fig. 6. Typical simulator results showing effect of load changes ($\omega_B = 4$).

still in a state of infancy compared with, for example, the i.t.a.e. index applied to system error. Wide variations of i.d.s. were found to occur for the experimental conditions, and it is considered helpful to treat, for each plant condition, the i.d.s. obtained for the second-order Butterworth design as the norm, and to express the results obtained for the other designs as a ratio. Table 2 summarizes the results, and shows that the second-order design, on average, is best of the three unity feedback designs. The index reasonably quantifies the designs in the same order of preference as the theoretical sensitivities and visual inspection of the whole family of experimental results also leads to the same ranking. An important feature of the results is the sensitivity improvement of the second-order Butterworth design compared with the non-cancellation design, thus showing that *if series compensation is used, nothing is gained in sensitivity reduction by avoiding the cancellation of nominal plant poles.*

9. Sensitivity Reduction by State Variable Feedback

If direct state variable feedback (θ_o , $\dot{\theta}_o$ and $\ddot{\theta}_o$), is used to obtain the desired third-order Butterworth transfer function, then, as shown in Fig. 1(c), the gain becomes 12.8, a feedback transfer function,

$$H(s) = 1 + 0.422s + 0.031s^2 \dots\dots(12)$$

is required, and the classical sensitivity function argument of Fig. 2 predicts an improvement in rejection of changes in plant dynamics. Typical simulator results for this feedback compensation system are included in Fig. 6 and Table 2, and show

conclusively that for the samples covered in the experimental investigation, this design is less sensitive than the series-compensated second-order Butterworth design. A further advantage in using this feedback configuration is the enormous benefit in low-frequency disturbance rejection by a factor of 6.4 : 1.

Table 2

'Integral of deviation squared' analysis of experimental transient responses obtained when controlling various plant dynamics

| Case | Experimental conditions | | Integral of deviation squared referred to second-order Butterworth as norm | | | |
|------|-------------------------|----------------------|--|--------------------------------|-------------------------|---|
| | Plant gain | Plant time constants | Second-order Butterworth design | Third-order Butterworth design | Non-cancellation design | Feedback compensated third-order Butterworth design |
| 1 | 1.5 | 1, 0.2 | 1.00 | 1.57 | 1.35 | 0.29 |
| 2 | 0.5 | 1, 0.2 | 1.00 | 0.88 | 1.27 | 0.31 |
| 3 | 1.0 | 1.5, 0.2 | 1.00 | 0.80 | 1.64 | 0.40 |
| 4 | 1.0 | 0.5, 0.2 | 1.00 | 2.76 | 1.73 | 0.27 |
| 5 | 1.0 | 1, 0.3 | 1.00 | 2.65 | 1.03 | 0.36 |
| 6 | 1.0 | 1, 0.1 | 1.00 | 1.42 | 0.85 | 0.45 |
| 7 | 1.5 | 0.5, 0.1 | 1.00 | 1.27 | 1.06 | 0.11 |
| 8 | 0.5 | 1.5, 0.3 | 1.00 | 1.88 | 1.66 | 0.81 |

10. Non-linear Transient Responses

Sample experimental transients responses obtained from an analogue computer study are shown in Fig. 7, the 20 V input representing a reasonable excursion into the non-linear mode. It is clear from these results that the second-order Butterworth design is vastly inferior in the presence of plant input saturation. There is a certain similarity between the non-linear transients recorded for the third-order Butterworth design and the non-cancellation design despite their different pole-zero geometries.

Previous experimental work on non-linear transient responses such as that by Chang,¹⁴ has been assessed simply by visual inspection of the responses. However, the comparison may be quantified by using the i.t.a.e. criterion; work by Gupta²⁹ suggests that we should attach significance only to relatively large changes in the integral, for example a factor of two. The results shown in Table 3 have been normalized by dividing the integral by the step demand, and confirm the goodness of the third-order design for excursions into the non-linear mode, whilst rating the non-cancellation design better than the second-order design, even for the largest input.

11. Why is the Third-order Butterworth Design Optimum?

It is clear from Figs. 4 and 7 that the third-order Butterworth design is much better than the second-order Butterworth design in the presence of plant input saturation, since the transient response is degraded less and there is a much lower probability of high-frequency noise saturating the plant and thus derating performance in following the command signal even when θ is small. Both system transfer functions result in acceptable transient responses in the linear mode and identical bandwidths, so what other property of $T(s)$ is important in this context? Clearly as $\omega \rightarrow \infty$, $|\theta_p/\theta_i|$ approaches a constant which is a measure of the high frequency noise transmitted from system input to plant input, whilst in the time domain the initial value of θ_p will measure the tendency to saturate in response to transient demands. When θ_i is the unit step, and the system transfer function is of the same order as the plant transfer function, then it may be shown that³¹

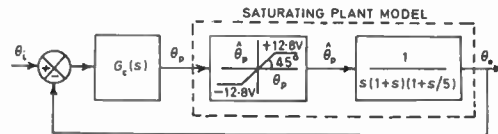
$$\frac{\theta_p}{\theta_i}(0) = \left. \frac{\theta_p}{\theta_i}(j\omega) \right|_{\omega} = \frac{1}{K} \frac{\prod |\text{system poles}| \prod |\text{plant zeros}|}{\prod |\text{system zeros}| \prod |\text{plant poles}|} \dots (13)$$

where K is the plant gain. Equation (13), though physically obvious, is of tremendous practical importance, because it means that systems designed using modern control theory and optimized on the basis of

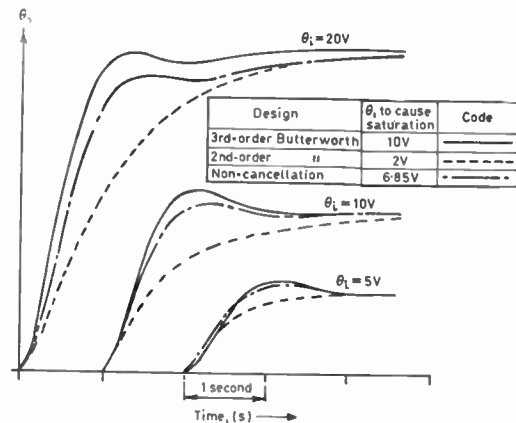
Table 3

Normalized i.t.a.e. $\left(\frac{1}{\theta_i} \int_0^\infty t|\theta|dt \right)$ for experimental non-linear transients of 3 basic designs

| Design input | Third-order Butterworth | Non-cancellation | Second-order Butterworth |
|--------------|-------------------------|------------------|--------------------------|
| 5 | 0.22 | 0.21 | 0.39 |
| 10 | 0.21 | 0.30 | 0.96 |
| 20 | 0.30 | 0.78 | 1.86 |



(a) Block diagram of non-linear plant model used in analogue computer studies.



(b) Analogue computer results.

Fig. 7. Experimental non-linear transients.

the step response will be implicitly optimized in reduction of system input high-frequency noise appearing at the plant input. Values of equation (13) for the three designs are shown in Table 4, and the linear mode step responses for θ_p/θ_i are shown in Fig. 8. The reason for the relative goodness of the non-cancellation design is now readily apparent, although the goodness is coincidental because no attempt was made to control the $\prod |\text{system poles}| / \prod |\text{system zeros}|$ ratio. From Fig. 8 it is evident that the correct choice of $T(s)$ can lead to such a reduction in the plant input demand that in many circumstances extra non-linear compensation will be unnecessary. In particular, we

note through equation (13), the tremendous practical improvement in design obtained by making the system transfer function of same effective order as the plant transfer function.

Table 4

$\theta_p(0)$ for $\theta_i =$ unit step, and $\left| \frac{\theta_p}{\theta_i}(j\omega) \right|_{\omega \rightarrow \infty}$ for 3 basic designs

| Design | $\frac{\theta_p}{\theta_i}(0) = \left \frac{\theta_p}{\theta_i}(j\omega) \right _{\omega \rightarrow \infty}$ |
|--------------------------|--|
| Second-order Butterworth | 64 |
| Third-order Butterworth | 12.8 |
| Non-cancellation | 18.6 |

12. Fundamental Theorem Relating to Optimum Transfer Functions

It is clear from the evidence of Figs. 4, 7 and 8 that we can use equation (13) to postulate an important theorem,

‘An optimal system transfer function for the control of plant with input saturation will have a small Π |system poles|/ Π |system zeros| ratio.’

An important practical consideration is that a given bandwidth is achieved with a small Π |system poles|/ Π |system zeros| ratio by choosing a Butterworth design (or thereabouts) of appropriate order.

13. Fundamental Theorem Relating System Bandwidth to Performance in the Presence of Plant Input Saturation

It is well known that any feedback system transfer function can be rewritten in such a form that the highest and lowest coefficients in the denominator are simultaneously equal to unity. Transient and frequency response shapes are fixed by the coefficients of the normalized transfer function, whilst true bandwidth is proportional to the normalization frequency ω_0 . Let two designs of order $(n-q)$, with identical pole-zero geometries, controlling the same plant differ by a bandwidth ratio λ . It follows immediately from equation (13) that,

$$\left. \begin{aligned} \left[\frac{\theta_p}{\theta_i}(0) \right] &= \lambda^{(n-q)} \left[\frac{\theta_p}{\theta_i}(0) \right] \\ \text{(system with} & \quad \text{(system} \\ \text{bandwidth } \lambda\omega_B) & \quad \text{bandwidth } \omega_B) \\ \left[\frac{\theta_p}{\theta_i}(j\omega) \right]_{\omega \rightarrow \infty} &= \lambda^{(n-q)} \left[\frac{\theta_p}{\theta_i}(j\omega) \right]_{\omega \rightarrow \infty} \\ \text{(system with} & \quad \text{(system with} \\ \text{bandwidth } \lambda\omega_B) & \quad \text{bandwidth } \omega_B) \end{aligned} \right\} \dots\dots(14)$$

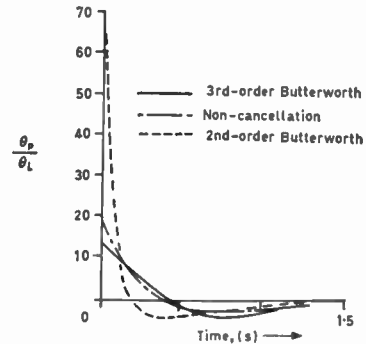


Fig. 8. Plant input demanded by system step input (linear mode response).

Equation (14) leads to the second theorem, ‘For a given system pole geometry, the tendency to saturation is increased by the ratio $\lambda^{(n-q)}$ if the bandwidth is increased by λ .’

This theorem quantifies the expected deterioration in real behaviour due to an arbitrary demand for an increased system bandwidth, without corresponding consideration of system pole-zero geometry.

14. Significance of Plant Dynamics in the Choice of Dominant System Pole-Zero Array

We have seen that for a specific example, an optimum design was achieved by choosing a dominant $T(s)$ of the same effective order as the plant, that the optimality of the design was related to the achievement of a small Π |system poles|/ Π |system zeros|ratio, and that one $T(s)$ directly satisfying these requirements is the Butterworth array. Equally encouraging results have also been obtained by direct choice of the Butterworth array of the same order as the plant, and of the required bandwidth, for a wide variety of plant transfer functions, including control of highly oscillatory loads such as the systems studied by Chang,¹⁴ and Towill, Cooper, and Lamb.³⁰

To appreciate the use and limitations of the two theorems viewed from a practical engineering viewpoint let us study the effect of varying the system bandwidth, for a fixed plant transfer function. Clearly we would wish for sharp transitions in system order as the bandwidth is increased, so that a clear mandate on the desired order of the dominant $T(s)$ can be built into design procedures. Figure 9 shows the frequency domain plot of $|\theta_p/\theta_i|$ as the bandwidth is varied. At a bandwidth of 1 rad/s, only two plant poles are in or near the bandwidth circle in the s plane, yet Theorem I would still indicate advantages for the third-order Butterworth design. Are these advantages significant? It is clear from Fig. 9 that the peak values in $|\theta_p/\theta_i|$ are not the high-frequency asymptotes for the case

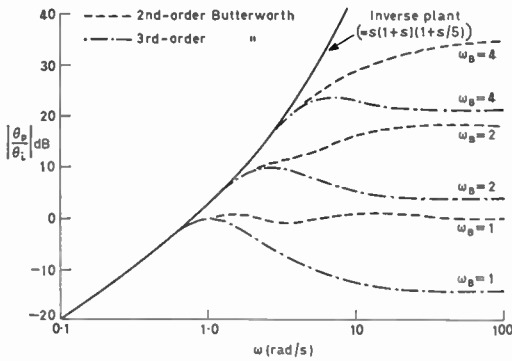


Fig. 9. Plant input as a function of bandwidth.

where $\omega_B = 1$, therefore little practical advantage is gained by choosing the third-order Butterworth design when the plant is second-order if the input noise is flat and band-limited just beyond system bandwidth. Clearly, specific peaks in the noise spectrum would also influence the choice. As the system bandwidth is increased, the peak values of plant input are more directly measured by the $\Pi|\text{system poles}|/\Pi|\text{system zeros}|$ ratio.

In the time domain, in order to study transitions in the order of $T(s)$, we make use of the results of modern control theory, which uses as a measure of goodness,

$$J = \int_0^\infty [\theta(t)^2 + k^2\theta_p(t)^2] dt \quad \dots\dots(15)$$

where $\theta(t)$ and $\theta_p(t)$ are the system error and plant input respectively in response to a system step demand, and k is a Lagrange multiplier.

For the plant dynamics defined by equation (5) the modern control theory answer is given by constructing the root square locus (using conventional locus rules) for the equation,

$$1 + \frac{k^2}{s(1+s)(1+s/25)} = 0 \quad \dots\dots(16)$$

as shown in Fig. 10. To interpret the results in physical terms, the system pole-zero array has been constructed and then via the coefficient plane calibrated in terms of system bandwidth. The system with bandwidth 4 rad/s has a pole product of $(4 \cdot 195^2 \times 5 \cdot 83) = 102$, hence $\theta_p(0)$ would be 20.4, which is slightly higher than for the third-order Butterworth design. Further work has confirmed the broad equivalence between the root square locus optimum design, the third-order Butterworth design and the non-cancellation design: also the vast superiority of all three compared to the second-order Butterworth design.

If the system performance specification requires a much smaller bandwidth, the root square locus selects a system dominated by only two poles. In contrast, Theorem I would still select a third-order Butterworth

design of appropriate bandwidth because there is still a vast difference in $\theta_p(0)$ for second- and third-order designs, but the practical advantage is lost because of the interaction between plant and system dynamics, which tends to make for little difference in the peak values of $\theta_p(t)$, as will be illustrated in Section 16. The root square locus solution has recognized this and used the opportunity to reduce the integral of error-squared contribution in eqn. (15) by opting for a two-pole design with faster rise-time.

The root square locus indicates that the transition from one order to the next is a rapid one, such that plant poles at 2.5 times bandwidth influence optimal transfer functions only marginally, whilst plant poles at 1.25 times bandwidth definitely influence the order of the dominant transfer function. Had the plant possessed a zero near the origin, the root square locus would have driven one pole to form a dipole, leaving the dominant $T(s)$ one order lower.

15. Estimation of the Optimum Transfer Function Directly from the Open-loop Bode Plot

Having established that the actual system pole-zero geometry is not critical (provided the $\Pi|\text{system poles}|/\Pi|\text{system zeros}|$ ratios are compatible) by comparing the properties of the root square locus optimum design, third-order Butterworth, and non-cancellation designs, for $\omega_B = 4$, it is frequently worth while extending Leake's procedure³² to determine rapidly not only the order of the dominant transfer function, but also a suitable transfer function.

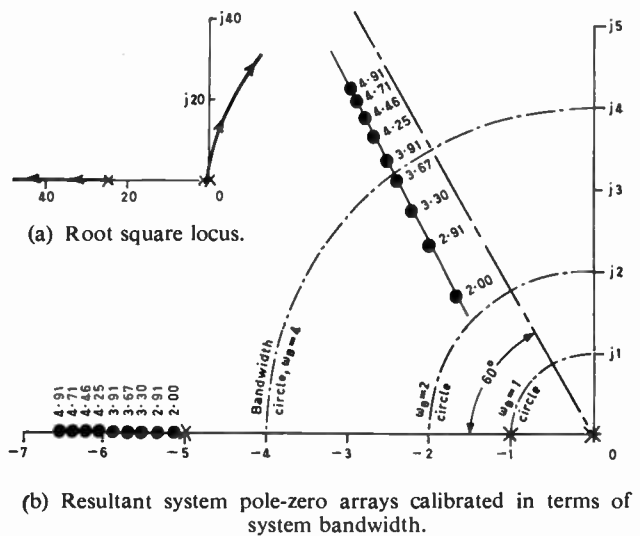


Fig. 10. Root square locus determination of optimum control of

$$P(s) = \frac{1}{s(1+s)(1+s/5)}$$

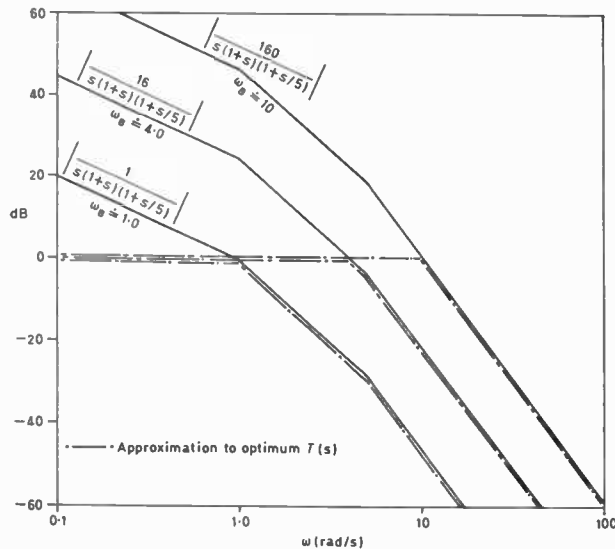


Fig. 11. Rapid approximation to optimum control of

$$P(s) = \frac{1}{s(1+s)(1+s/5)}$$

direct from asymptotic Bode plot.

If we plot the open-loop Bode diagram with feed-forward gain such that the plot crosses the 0dB line at the desired bandwidth, an approximate expression for a suitable $T(s)$ is obtained as shown in Fig. 11 by fitting $T(s)$ beyond bandwidth to the properties of the open-loop response, whilst below bandwidth $T(s)$ is made unity. At bandwidth, $|T(j\omega)|$ must be made -3dB which makes the choice of the Butterworth filter obvious in many cases.

For example, Fig. 11 shows the approximate form of $T(s)$ required for bandwidths of 1, 4 and 10 rad/s, for the plant defined by equation (5). Table 5 tabulates suitable optimum transfer functions.

Rapid transition in the order of the dominant transfer function is again apparent as a function of desired bandwidth. Either of the two systems obtained for a bandwidth of 4 rad/s would be much better than a system dominated by two poles only. Since the system with $\zeta = 0.55$ has a lower $\Pi|\text{system pole}|/\Pi|\text{system zero}|$ ratio it is to be preferred.

16. Deliberate Addition of Far-Off System Poles to Increase High Frequency Noise Rejection

The optimization theorem of Section 12 was developed specifically for systems with the same number of system poles and plant poles, and the same number of system zeros and plant zeros. However, if there is an excess of system poles,

Table 5

Optimum transfer functions determined from open loop Bode plot

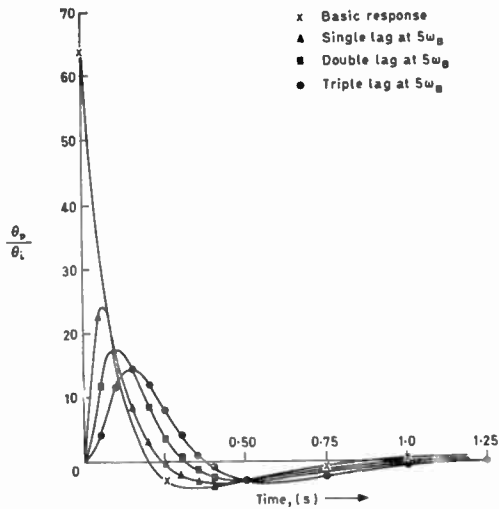
| Bandwidth | Approximate optimum $T(s)$ |
|-----------|---|
| 1 | $\frac{1}{\left(1 + \frac{2.0 \cdot 707}{1} + \frac{s^2}{1^2}\right) (1+s/5)}$ |
| 4 | $\frac{1}{\left(1 + \frac{2.0 \cdot 55}{4} s + \frac{s^2}{4^2}\right) (1+s/5)}$ (bandwidth obtained by adjustment to ζ) |
| 4 | $\frac{1}{\left(1 + \frac{2.0 \cdot 707}{5.33} s + \frac{s^2}{5.33^2}\right) (1+s/5)}$ (bandwidth obtained by adjustment of ω_n) |
| 10 | $\frac{1}{1 + 2 \left(\frac{s}{10}\right)^2 + \left(\frac{s}{10}\right)^3}$ |

$$\theta_p(0) \text{ and } \left. \frac{\theta_p}{\theta_i}(j\omega) \right|_{\omega \rightarrow \infty}$$

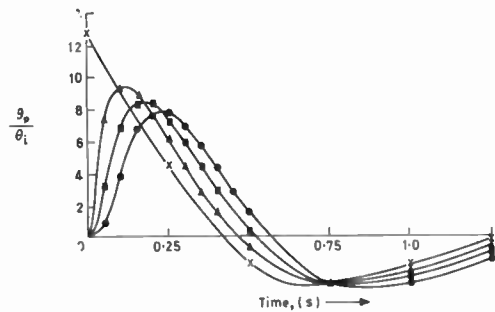
will both be zero.

Does the theorem still apply in defining dominant system poles and zeros under these conditions? To investigate this problem, for the same plant dynamics, the plant input transient responses have been evaluated when a number of simple lags are added to the feed-forward path such that real system poles exist at five times bandwidth. Figure 12 shows the responses for a system bandwidth of 4 rad/s, and although the effect of the extra poles is to reduce the peak of $\theta_p(t)$, the advantage of the third-order design (and hence the theorem) is in no way invalidated. In contrast, Fig. 13 shows that reduction of bandwidth to 1 rad/s results in $\theta_p(0)$ no longer being the peak value of $\theta_p(t)$ for the third-order design and presence of the extra poles reduces still further the difference between the maximum values of plant input for the two designs. It is therefore clear that little advantage is gained in transient considerations by opting for the third-order design, a fact predicted by Figs. 10 and 11.

As a practical example of the deliberate addition of such lags, the blind-landing autopilot proposed by Gill,³³ and optimized directly from analogue computer studies, we find not only the addition of three far-off poles for noise rejection, but also more



(a) Second-order Butterworth design.



(b) Third-order Butterworth design.

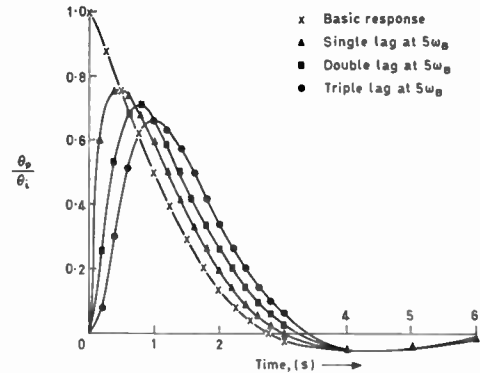
Fig. 12. Effect of additional noise filters on plant response to system step input

$$\left(\omega_B = 4, P(s) = \frac{1}{s(1+s)(1+s/5)} \right)$$

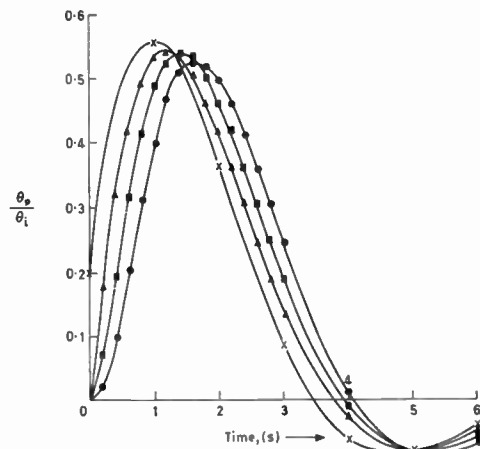
dominant system poles than effective plant poles; in fact the system poles very nearly match the third-order Butterworth filter! The extra dominant pole is introduced to increase rate of cut-off at bandwidth, since noise is significant at this frequency as shown by the design specification which is unusually detailed and explicit.

17. Rational Synthesis Procedure

The contributions of this paper may be summarized in the procedural flow chart of Fig. 14. For traditional and algebraic methods of synthesis, whether automated or manual, definition of the dominant transfer function is achieved by matching the order of this



(a) Second-order Butterworth design.



(b) Third-order Butterworth design.

Fig. 13. Effect of additional noise filters on plant response to system step input

$$\left(\omega_B = 1, P(s) = \frac{1}{s(1+s)(1+s/5)} \right)$$

transfer function to the effective plant dynamics by choice of the Butterworth filter of the same net order as the plant, whilst for hill-climbing procedures and the like it is achieved by the incorporation of checks based on $\Pi|\text{system poles}|/\Pi|\text{system zeros}|$ ratio.

18. Conclusions

It has been established theoretically and verified experimentally that optimization of feedback control system performance in the presence of plant input saturation is achieved by choosing a dominant system transfer function with at least the same number of dominant poles as the net order of the plant dynamics

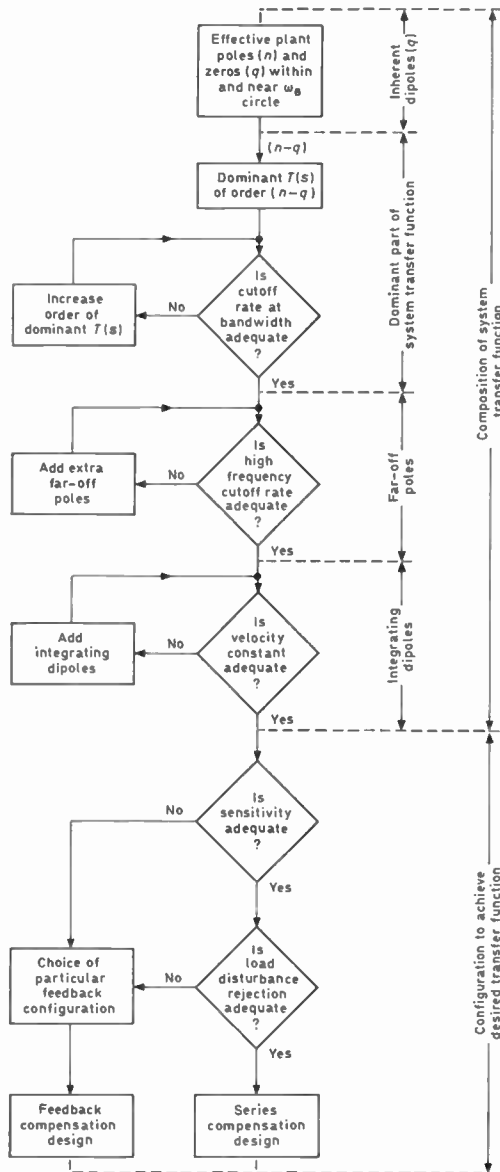


Fig. 14. Philosophy for optimization in the presence of plant input saturation.

in the region enclosed by, and adjacent to, the bandwidth circle. A Butterworth filter of this order will result in effective use of plant capability. Such a system is then slightly more sensitive to changes in plant dynamics compared with a system of similar bandwidth dominated by a lower order transfer function, but this can be rectified by the use of state variable feedback. Generalization of conclusions is justified partly by the range of systems studied (not all reported herein), and the broad equivalence of the procedures described to the modern control theory approach.

A fundamental theorem shows that optimality

under certain circumstances is measured by the $\Pi|\text{system poles}|/\Pi|\text{system zeros}|$ ratio, and a second theorem quantifies the harmful effect of indiscriminate increase in bandwidth in the linear mode on the non-linear performance. The first is easily incorporated as a check in any design procedure. As a by-product, it is seen that no reduction in system sensitivity is achieved by designs which deliberately seek to avoid cancellation of unwanted plant dynamics. Addition of far-off system poles to increase noise rejection does not invalidate the conclusions reached on the order of the optimum dominant transfer function.

19. Acknowledgments

The author wishes to acknowledge the assistance of members of the Dynamic Analysis Group, and thanks are also due to the Science Research Council for supporting the project.

A fundamental question posed by Professor C. Holt Smith gave the investigation considerable impetus.

20. References

- Whiteley, A. L., 'The theory of servo systems, with particular reference to stabilization', *J. Instn Elect. Engrs*, **93**, p. 353, 1946.
- Graham, D. and Lathrop, R. C., 'The synthesis of optimum transient response criteria and standard forms', *Trans. Amer. Inst. Elect. Engrs*, **72**, Pt. 11; *Applications and Industry*, No. 9, p. 273, November 1953.
- Susini, A., 'Filters, Amplifiers, and Servomechanisms', p. 233 (Heywood, London, 1963).
- Bach, R. E., Jr., 'A practical approach to control system optimization', Proc. 1965 IFAC Symposium, Tokyo, 1965, pp. 129-135.
- Diamesis, J. E., 'Time domain calculation of system performance criteria', *I.E.E.E. Trans. on Education*, p. 51, March 1967.
- Craxford, S. C., 'The Optimization of a Function Subject to a Number of Constraints'. Royal Military College of Science, Shrivenham, 2ASC. Project Study, 1967.
- Williams, J. E., 'A note on absolute value criteria', *I.E.E.E. Trans. on Automatic Control*, **AC-12**, p. 615, No. 5, October 1967.
- Davison, E. J., 'An automatic way of finding optimal control systems for large multivariable plants', Proc. of 1965 IFAC Tokyo Symposium, pp. 367-372.
- Musker, G. and Henman, Marian, 'A Computer Study of Automatic Control on the ILS Glide Path'. R.A.E. Technical Report No. 66057, February 1966.
- Breslin, M. J., 'Computer Techniques for System Synthesis'. M.S. Thesis, U.S. Naval Post Graduate School, 1967.
- Walker, J. A., 'Computer Techniques for Implementing Linear Control System Design Using Algebraic Methods', M.S. Thesis, U.S. Naval Post Graduate School 1968.
- Glavis, G. O., 'Frequency Response in the Parameter Plane', M.S. Thesis, U.S. Naval Post Graduate School, 1968.
- Moore, J. B., 'Control System Design Extensions of the Parameter Plane Concept', Ph.D. Thesis, University of Santa Clara, 1966.

14. Chang, S. S. L., 'Synthesis of Optimal Control Systems, Chapter 2' (McGraw-Hill, New York, 1961).
15. Freeman, E. A. and Abbott, K. M., 'Design of optimal linear control systems with quadratic performance indices', *Proc. Instn Elect. Engrs*, 114, No. 8, p. 1180, August 1967.
16. Jacobs, O. R. L., 'Damping ratio of an optimal control system', *I.E.E.E. Trans.*, AC-10, p. 473, October 1965.
17. Rynaski, E. G. and Whitbeck, R. F., 'Theory and Application of Linear Optimal Control', Tech. Report CAL 1H-1943-F-1, Cornell Aeronautical Laboratories, October 1965.
18. Tyler, J. S., Jnr., and Tuteur, F. B., 'Use of a quadratic performance index to design multivariable control systems', *I.E.E.E. Trans.*, AC-11, No. 1, p. 84, January 1966.
19. Kreindler, E., 'Closed loop sensitivity reduction of linear optimal control systems', *I.E.E.E. Trans.*, AC-13, No. 3, p. 254, June 1968.
20. Newton, George C., 'Compensation of feedback control systems subject to saturation', *J. Franklin Inst.*, 254, p. 281, 1952.
21. Tustin, A., Allanson, J. T., Layton, M. M. and Jakeways, R. J., 'Design of systems for automatic control of the position of massive objects', *Proc. Instn Elect. Engrs*, 105, Part C, Supplement No. 1, p. 1, 1958 (I.E.E. Paper No. 2651M).
22. Bowler, P., 'A compensation technique for optimizing large and small signal servo response'. Proc. I.E.E. Conf. on Servocomponents, London, 1967, p. 133.
23. Vaughan, D. R. and Foster, W. C., 'Feedback design of systems with saturation constraints', *Instrum. Soc. Amer. Trans.* 5, p. 165, 1966.
24. Towill, D. R., 'Performance advantages of certain non-linear servomechanisms, Part II,' *Control*, 10, p. 45, January 1966.
25. Horowitz, I. M., 'Fundamental theory of automatic linear feedback control systems', *Trans. I.R.E.*, AC-3, pp. 5-19, December 1959.
26. Towill, D. R., 'Analysis and synthesis of feedback compensated third-order control systems via the coefficient plane', *The Radio and Electronic Engineer*, 32, No. 2, pp. 119-31, August 1966.
27. Liou, M. L., 'A novel method of evaluating transient response', *Proc. I.E.E.E.*, 54, No. 1, pp. 20-23, January 1966.
28. Thompson, J. G. and Kohr, R. H., 'Modelling and compensation of non-linear systems using sensitivity analysis', *Amer. Soc. Mech. Engrs Trans., J. of Basic Engineering*, 90, Series D, p. 187, June 1968.
29. Gupta, S. C., 'Performance of Process Models', IFAC Conference, London, 1966, Paper 5, Session X.
30. Towill, D. R., Cooper, J. S. and Lamb, J. D., 'Dynamic analysis of fourth-order feedback control systems', *Intl. J. Control*, 10, No. 2, August 1969.
31. Towill, D. R., 'Optimum Transfer Functions for Feedback Control Systems'. University of Wales Institute of Science and Technology, Cardiff, DAG TN 22, December 1968.
32. Leake, R. J., 'Return difference Bode diagram for optimal system design', *I.E.E.E. Trans.*, AC-10, p. 342, July 1965.
33. Gill, F. R., 'The Integrity of a Civil Blind Landing System with Particular Reference to the Azimuth Channel'. R.A.E. Tech. Report No. 65022, February 1965.

Manuscript first received by the Institution on 11th February 1969 and in final form on 16th June 1969. (Paper No. 1284/IC14.)

Radio Engineering Overseas . . .

The following abstracts are taken from Commonwealth, European and Asian journals received by the Institution's Library. Abstracts of papers published in American journals are not included because they are available in many other publications. Members who wish to consult any of the papers quoted should apply to the Librarian giving full bibliographical details, i.e. title, author, journal and date, of the paper required. All papers are in the language of the country of origin of the journal unless otherwise stated. Translations cannot be supplied.

PHASE LOCKING OF GUNN OSCILLATORS

The phase locking of Gunn oscillators operating in the steady state has been investigated experimentally at the Research Department of AEG-Telefunken and the measured dependence of the locking range on the injected r.f. power compared with theory. It is shown that in pulsed operation, oscillation build-up and interpulse coherence could be improved by injecting an r.f. signal as weak as 50 dB below the output power. Triggering of the Gunn instability by an injected r.f. signal has been utilized to divide the frequency of the input signal. The power gain, when synchronizing to the first sub-harmonic of the external signal, was a maximum of 11.5 dB.

'Injection phase locking of Gunn oscillators', H. Pollmann and B. G. Bosch, *Nachrichtentechnische Zeitschrift*, 22, p. 174, March 1969.

PARAMETRIC FREQUENCY DEMODULATOR

A Soviet paper studies the frequency modulation of electromagnetic waves theoretically and experimentally in an extended (distributed) diode with a p-n junction. The equivalent circuit of such a diode is discussed. An analytic solution for the waves is obtained in the presence of a slowly-varying modulating voltage with allowance for re-reflection at the line ends. Some characteristics are presented of germanium diode specimens. The conversion of the wave spectrum in such a diode was experimentally observed; it is shown that at signal frequencies of 100-400 MHz, it is possible to obtain deep frequency modulation in such diodes.

'A parametric frequency demodulator using a distributed diode', A. I. Vesmitskiy, N. S. Stepanov and V. N. Shabanov, *Radio Engineering and Electronic Physics* (English language edition of *Radiotekhnika i Elektronika*), 14, No. 2, p. 277 February 1969.

COMPUTER MODELS FOR FIELD EFFECT TRANSISTORS

In non-linear circuits it is frequently impossible to investigate analytically transient processes at high frequencies. One can, however, obtain numerical solutions even for very complex networks by means of the two computer-application programs 'Continuous System

Modelling Program' (CSMP) and 'Electronic Circuit Analysis Program' (ECAP).

Using an f.e.t. mixer as an example the Swiss extension of IBM have shown how computer models can be developed for both programs. By computing the build up transient of the oscillating voltage at the i.f. filter, the conversion transconductance can be found as a function of gate voltage, source impedance, load and frequency. This provides information on the optimal choice of operating point and oscillator voltage.

Finally, the advantages of the two programs are compared and notes concerning application fields are given.

'Computer models for calculating non-linear phenomena in field effect transistors', J. S. Vogel, *Archiv der Elektrischen Übertragung*, 23, p. 331, July 1969.

P.C.M. MICROWAVE CIRCUITS

In the practical design of microwave transmission systems using p.c.m. signals it is important to apply a design technique which ensures that the equipments selected will, when connected as a system, satisfy the given circuit performance requirements. In other words, a system-design method must be applied. A Japanese paper describes work carried out at the Nippon Telegraph and Telephone Co. of Tokyo, on a systematized design technique using the following approach:

- (1) Establish reference circuit and specify a bit error rate as the circuit standard.

- (2) Define the causes of quality degradation in the transmission path (thermal noise, intersymbol interference and frequency interference) and determine their individual contributions to the bit error rate as well as their combined effects.

- (3) Substitute a signal/noise ratio as the circuit performance standard, applying a method of allocating various amounts of noise to the difference degrading effects.

Application of this method is illustrated by the design of a 4 GHz microwave p.c.m. circuit.

'A method of designing p.c.m. microwave circuits', K. Tachikawa, *Electronics and Communications in Japan* (English language edition of *Denshi Tsushin Gakkai Ronbunshi*), 51-B, No. 11, p. 54, November 1968.