THE RADIO AND ELECTRONIC ENGINEER

The Journal of the Institution of Electronic and Radio Engineers

FOUNDED 1925 INCORPORATED BY ROYAL CHARTER 1961

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

VOLUME 28 | SEPTEMBER 1964 | NUMBER 3

INTERNATIONAL EXCHANGE

Although there are now numerous channels for the exchange of international opinion, the engineer, in comparison with members of other professions, does not seem to take such an active role at international meetings. It is true that increasing numbers of engineers travel abroad but in the main their journeys are for prescribed purposes, such as giving 'on the spot' technical advice. Little time can be given to gaining understanding of other ways of living or even to have discussions with overseas members of the same profession who are not employed within identical firms or government departments.

Shortage of qualified technical staff is often advanced as a reason for cutting down an engineer's trip to the minimum time. The engineer therefore misses the opportunity of expanding his knowledge and understanding of matters essential for ultimate managerial responsibility.

There is increasing opportunity for the exchange of technical information at an international level. The plethora of meetings for the presentation and discussion of papers either on a broad range of subjects or exploring a specialized subject in depth may be gauged from the monthly calendar of 'Conferences and Exhibitions' published in this *Journal.* There appears, however, to be a growing tendency for organizations to confine attendance at overseas meetings to engineers who are presenting papers. Thus the prospective author has often to advise the conference organizers that his attendance will be conditional on his paper being included in the programme. This can lead to the submission of rather ineffectual papers or repetition of already published material—factors not easy to assess from a synopsis. This often results in unsatisfactory representation of both company and country.

Attendance at international exhibitions is valuable not only to the engineer concerned primarily with commercial matters but also to the research and development engineer. The prior assessment of the immediate worth of such a visit is difficult, particularly with the ever-increasing number of exhibitions, but an overseas visit can be commercially valuable, particularly in the garnering of technical information.

Conferences concerned with standardization and other international agreements in the technical field are usually organized by international agencies, for example, the International Electrotechnical Commission or the Comité Consultatif International des Radio-communications. Delegates may be sponsored by government departments or, particularly in the standardization field, by industry associations. Occasionally full representation at a conference has to be supported from other sources and the Institution has helped finance British delegates to several I.E.C. meetings. A more enlightened view of the vital nature of these international meetings ought to be taken by industry and government.

Participation in any international function will only be of the greatest benefit to the engineering profession if more individuals travel abroad rather than the same group being sent more frequently! International exchanges demand variation of representatives in order to gain the advantages of fresh minds.

Evidence of the fact that these comments do not apply just to only one or two countries is apparent from the delegate list of almost any conference of interest to the radio and electronics engineer.

G. D. C.

INSTITUTION NOTICES

Conference on Components and Materials

A Conference on "Components and Materials used in Electronic Engineering" is being organized by the I.E.R.E., the I.E.E. Electronics Division, and the U.K. and Eire Section of the I.E.E.E. The Conference will be held in London from 17th to 21st May 1965 and will be concurrent with the R.E.C.M.F. Components Exhibition.

The subjects to be discussed will include recent developments in active and passive components, including integrated circuits, and materials used in their manufacture:

Resistors, dielectrics and capacitors;

Inductors, magnetic materials and transformers;

Piezoelectric, ferro-electric and magnetostrictive materials and devices;

Microwave ferrites and associated components;

Electro-mechanical devices (relays, plugs, etc.);

Computer components (storage devices etc.);

Printed wiring, potted circuits, automatic assembly techniques and methods of connection;

Semiconductor devices;

Integrated circuits, thin film circuits and devices;

Reliability and the effects of extreme operating conditions.

Contributions of length not exceeding 1500 words are invited and synopses of approximately 200 words on any of the above subjects should be submitted to the Secretary of the Organizing Committee, c/o The I.E.E., Savoy Place, London, W.C.2, immediately.

Further information and registration forms will be available shortly and may be obtained from the I.E.R.E., 8–9 Bedford Square, London, W.C.1.

International Conference on "Medical Electronics and Biological Engineering"

The Sixth International Conference on "Medical Electronics and Biological Engineering" will be held in Tokyo from 22nd to 27th August 1965, and is being organized by The International Federation for Medical Electronics and Biological Engineering. The President of the Conference is Mr. W. J. Perkins (Member) who is with the National Institute for Medical Research.

The scope of the Conference will include: applications of instrumentation; telemetering; television; computers; isotopes; radiation; ultrasonic and optical techniques; mathematical analysis; simulation; bionics; biological control; artificial organs; ergonomics; human engineering and hospital automation.

A scientific exhibition of instruments and accessories relevant to the Conference and a commercial exhibition for manufacturers and traders are being arranged to be held concurrently. Full details of the Conference and Exhibitions may be obtained from Professor K. Suhara, Secretary General of ICME & BE, c/o Japan Society of Medical Electronics and Biological Engineering, Old Toden Building, 1–1 Shiba-Tamura-cho, Minato-ku, Tokyo, Japan.

Symposium on "Electronics in Industry—The Next Five Years"

The West Midland Section of the Institution of Electronic and Radio Engineers and the South Midland Centre (Electronics Section) of the Institution of Electrical Engineers propose to hold a symposium on "Electronics in Industry—The Next Five Years" in Birmingham on 6th April 1965.

The symposium will provide the opportunity for the designers and users of Industrial Electronic Equipment to exchange views on the present trends of development and the future applications of electronics in industry.

Papers will be grouped in three sessions devoted to Materials and Components; Circuits and System Equipment; System Design and Application. Offers of papers in these broad categories are invited.

Further details may be obtained from D. P. Howson, M.Sc., A.M.I.E.R.E., Electronic and Electrical Engineering Department, University of Birmingham, Edgbaston, Birmingham, 15.

Guthrie Lecture

The 1964 Guthrie Lecture of the Institute of Physics and the Physical Society entitled "Radio Telescopes" will be delivered by Professor Martin Ryle (University of Cambridge) at 4.45 p.m. on 1st October 1964, in the Physics Department of the University of Keele. The lecture is being held in conjunction with the Conference on "Electron Emission" and for those not attending the conference admission will be by ticket, obtainable from the Administration Assistant, I.P.P.S., 47 Belgrave Square, London, S.W.1.

Thomson Lecture

This year's Thomson Lecture of the Society of Instrument Technology will be presented by Sir Gordon Sutherland, Sc.D., LL.D., F.R.S., Director of the National Physical Laboratory, whose subject will be "The Function of a National Physical Laboratory". The Lecture will be given at The Royal Institution at 6 p.m. on Thursday, 22nd October 1964.

Admission will be by ticket only and a number of tickets have been reserved for members of the I.E.R.E., who should apply in writing to the Secretary of The S.I.T., 20 Peel Street, London, W.8.

U.D.C. 621.387.422

A 1-Mc/s Bidirectional Gas-filled Counting Tube

By

K. APEL, Dr. rer. nat.†

Reprinted from the Proceedings of the Symposium on "Cold Cathode Tubes and their Applications" in Cambridge 16th–19th March 1964.

Summary: The study of the transfer mechanism of glow discharge counting tubes has resulted in the development of a new counting principle and the construction of corresponding tubes. The main features realized are high-speed operation up to 1 Mc/s, good behaviour under stand-by conditions and low pulse drive requirements that make direct transistor drive feasible. High-speed behaviour, output characteristics and a new reset method are described. Life-test results are given and a typical transistor drive circuit is discussed.

1. Introduction

During the last ten years a great many gas-filled switching and counting tube designs have been introduced. Most of them have, however, certain shortcomings regarding reliable bidirectional high-speed operation. The objective of the new design was to produce a small tube for bidirectional counting up to as high a frequency as possible, using transistor drive and having long life and good behaviour under stand-by conditions.

2. Transfer Considerations

In the interests of reliability and counting security the transfer mechanism had to be as simple and speedy as possible. The transfer in all tubes developed so far is characterized by making one cathode (an auxiliary cathode) negative to the index cathode which is covered with glow. It is well known that this potential drop cannot be allowed to occur rapidly without disturbing the counting process. The reason lies in the relatively weak coupling of the glow and the anode potential. The drift current between glow and anode is the result of diffusion processes of electrons and ions. Additionally the anode-to-cathode stray capacitance has to be charged during the transfer time. Thus the anode potential follows after a certain time. As the 'information' takes a path via the anode, this process takes some time and involves instabilities.

Between two transfers the auxiliary cathodes have to be positive (about 50 V) to the main cathodes to keep the voltage to the anode below the maintaining voltage. For a secure transfer the stepping pulse must be so high that the auxiliary cathode is driven sufficiently negative to the main cathode. Normally it must be in the order of 100 V. Therefore a direct transistor drive is considered inconvenient.

The counting direction is determined by one of two governing conditions: either there are two or more symmetrical auxiliary cathodes, which receive transfer pulses one after another, or there is just one auxiliary cathode between two main cathodes, which is shaped in such a way that transfer occurs in either direction. The first condition takes more time; the second one depends on the geometrical position of the discharge and can therefore be influenced by ageing.

Attempts were made to find a transfer mechanism which avoids any large variations of the anode voltage but which effects a variation of the current distribution between the participating cathodes. This transfer mechanism must be good for low-voltage pulse drive; it should not need more than one auxiliary cathode between two main cathodes and the counting direction should only be determined by electrical means. Extended studies and experiments confirmed the validity of the above assumption. The counting rate of tubes using this transfer principle, during life and reliability tests, was less than one failure per 10^{13} impulses in 1000 hours.

3. Counting Principle

The tube contains a cylindrical anode and four star-shaped electrodes (two cathodes K1, K2 and two auxiliary cathodes G1, G2) (Fig. 1). Each electrode has five limbs referred to in clockwise notation as K11, G21, K21, G12, K12 and so on. An output electrode S was placed near the end of each cathode limb.

The main cathodes are driven by two anti-phase square-waves from the output of a bistable multivibrator. These square-wave voltages are also applied to the auxiliary cathodes via differentiating capacitors C_{g1} and C_{g2}.

Assuming cathode K11 carries a discharge in a stable condition, limbs of K2 and the auxiliary cathodes between this limb and the adjacent limbs of K1, are all positive to K11, thus preventing any trans-

† Elesta A.G., Bad Ragaz, Switzerland.

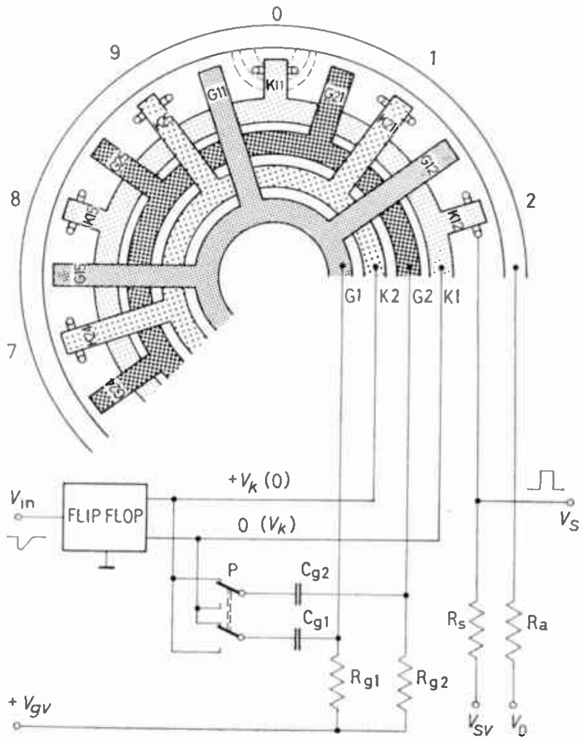


Fig. 1. Arrangement of electrodes and driving circuit.

fer of discharge. An input pulse switches the flip-flop into its second stable position (Fig. 2). The voltage of K1 rises from zero to V_K; the voltage of K2 decreases from V_K to zero and G21 gets a short negative-going pulse while G11 gets a short positive-going pulse. The discharge follows the most negative potential in the neighbourhood and transfers from K11 via G21 to K21. This produces one counting step. The anode voltage varies only according to the 'overlapping failure' of the flip-flop output pulses; the variations in practice are of the order of 5 to 10 V. By changing over the switch P the relative phasing of potentials at the cathodes and auxiliary cathodes will be reversed, thus causing transfer of the discharge to take place in the opposite direction.

The tube is filled with hydrogen. Due to the short ionization and de-ionization times a very high counting frequency is possible. With a hydrogen-filled tube there is a severe tendency to produce instabilities and uncontrolled oscillations. The design of the cathode limbs is intended to lead to the highest stability of discharge. These are made of molybdenum in the shape of a U, thus forming a kind of hollow cathode. According to the counting direction the cathode limbs are symmetrical. To obtain a wide anode current range the stability of the discharge (i.e. the hollow-cathode effect) must be maintained over a wide current range. Present tubes count normally up to 200 kc/s in the range from 1 mA to 5·5 mA.

The dimensions of the cathodes become critical when counting frequencies of several megacycles are reached. The upper frequency limit is determined by imperfections of the edges, slight changes in the geometry during outgassing and by production tolerances. With carefully built experimental tubes counting frequencies of about 3 Mc/s were reached although standard production models with an upper speed limit of 1 Mc/s are manufactured.

A considerable difference in the counting behaviour at high frequencies exists between periodical impulses and statistically incoming impulses. Assume the discharge glow is positioned at K11 for some time (Fig. 1). An input pulse switches the flip-flop and the discharge transfers from K11 via G12 to K21. When the next pulse arrives after 1 microsecond the discharge has to switch to cathode limb K12. This step will be made only when the ionization rate at K11 has decreased below the ionization rate which is initiated by the auxiliary cathode G12 at the cathode limb K21. If this condition is not fulfilled, the glow discharge will return to the main cathode limb K11 and the counting is not correct. The ionization due to the stationary discharge is a function of time and current and can be reduced by reduction of the anode current. Thus useful anode current range is limited. In the case of periodical input pulses the discharge stays only a short time at each cathode. The ionization does not reach its stationary level, so the tube will count correctly

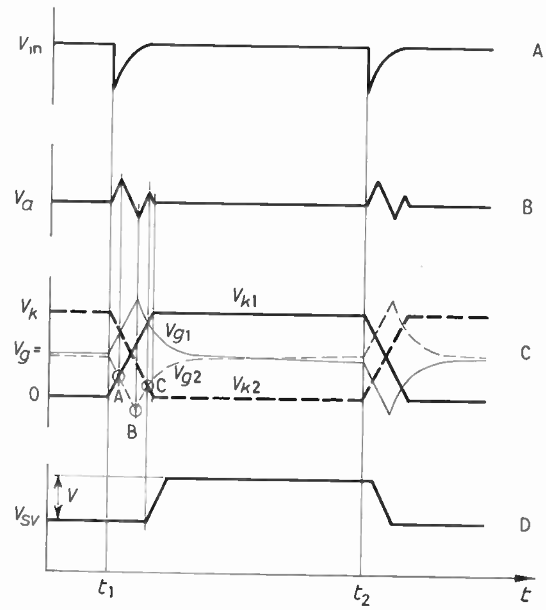


Fig. 2. Waveforms of input, output and electrode voltages.
A—Input; B—Anode voltage; C—Cathode and auxiliary cathode voltages; D—Output.

even at a higher anode current. Figure 3 shows the current range for sine-wave driving (A), periodical pulse driving (B) and statistical pulse driving (C) for a typical average tube.

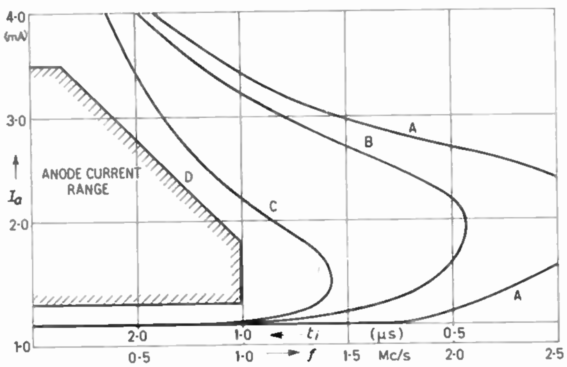


Fig. 3. Typical anode current range vs. counting frequency A—Sine wave drive; B—Periodical pulse drive; C—Statistical pulse drive; D—Recommended range.

4. Output

For an electronic read-out a pair of wire output electrodes are placed at the end of each main cathode limb. These electrodes act as probes which have necessarily different potentials with respect to the surrounding space potential thus effecting heavy distortions of the space charge. By numerous experiments a probe position was found which gives optimal output voltage with negligible influence on the counting process. Only the anode-cathode burning voltage and the ion current entering the auxiliary cathodes are affected by the space charge distortions and these are varied about 5%.

The internal resistance, on which depends the obtainable output power, cannot be influenced significantly without complicating the production and without affecting the counting behaviour. An average value is about 600 kΩ.

5. Reset

For reset on zero there are two possibilities. The first one is to apply a negative pulse to the selected output electrode. As soon as it becomes sufficiently negative to the glow-covered cathode it will take over the current, the anode voltage decreases and the glow on the index cathode declines. When the probe finally is made positive the anode voltage rises and reaches the striking voltage of the primed main cathode. This transfer is the same as that used in other counting tubes to transfer the glow from one position to the next. The disadvantage of this means that high peak-voltage impulses in the order of 200 V are needed and must be kept away from the low-voltage circuits by using high-voltage fast switching diodes. At the same time the flip-flop must be reset into the corresponding position.

It was found to be simpler and more reliable to use another reset method. The reset generator was designed as a self-starting pulse generator. The reset pulses are applied via a gate diode to the input of each of the counting stages in parallel. This gate is closed by the positive-going output voltage when the predetermined counting position is reached. The apparent disadvantage of the necessary time involved can easily be tolerated in most known applications. In addition it is possible to use this reset mechanism in a simple self-checking circuit for counting systems.

6. Operating Data

In the following section the main data for counting tubes of the type described are given to illustrate the previous section. The burning voltage of 305 V is due to the hydrogen filling, while the striking voltage is a little higher, 345 V. The driving voltage must be 50 V p-p or more and the direct voltage of the auxiliary cathodes must be between both cathode levels at 25 V. The guaranteed current range is between 1·8 and 3·5 mA with 2·7 mA as a recommended value. The voltage of the output pulse depends on the load resistor, and reaches 160 V for a load of 1 megohm. The differentiation capacitor for the auxiliary cathodes depends on the wanted upper counting frequency limit and on the rise and fall times of the driving voltage. For 100 kc/s a 220 pF capacitor is recommended, and the rise and fall time of the rectangular waves has to be less than 1·5 microsecond. The auxiliary cathode resistors have to be 10 kΩ each.

7. Mechanical Stress

The mechanical structure is very compact and rigid (Fig. 4). Mechanical measurements made on the tube showed that the tolerable mechanical stress under conditions of vibration at 50 c/s is 5 *g* over 100 hours in each of three positions at an acceleration of about 250 *g* during 1 minute and shock acceleration of up to 500 *g* in three different positions of the tube. The range of permissible ambient temperature is −100° C to +100° C.

During the tests the tube was operating, and the output pulses were compared with the output pulses of another tube operating under normal conditions. No counting errors could be observed.

8. Operating Life and Reliability of Count

Extended life tests were run to investigate the life expectancy and counting accuracy of the tube. They are divided into low-frequency tests, high-frequency tests and stand-by tests.

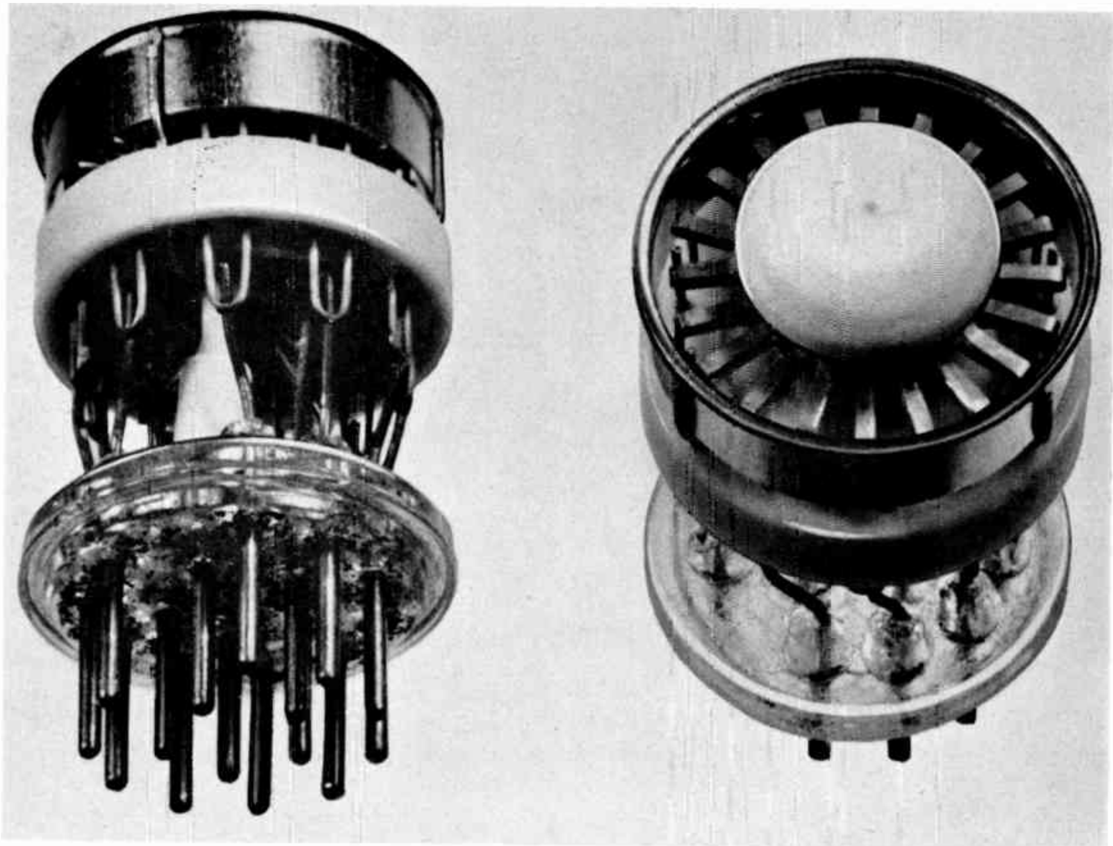


Fig. 4. Physical arrangement of the tube.

It is known that many troubles in using gas-filled counting tubes arise from burning voltage changes of adjacent electrodes when the glow remains on one cathode for a long period. The order of magnitude of the rate of the change is about 2 V per day for a helium-neon mixture of 95 : 5 and nickel cathodes according to Acton and Swift.[2]

By achieving extreme purity of the metal and ceramic parts, by careful mounting, and by flashing and outgassing of the tubes, this value could be reduced to about 5 V in 2500 hours This does not noticeably affect the counting behaviour. In the life test, tubes are set up with the glow discharge in one position for a long time (average is 1000 hours). After that time they receive a series of pulses at 100 kc/s for a few seconds and the accuracy of counting is checked. A restoration of any abnormal burning voltage is thus reduced to a minimum because each cathode glows only for a few tenths of a second.

The discharge is then set to the same position as before and stays there for the next 1000 hours. These kinds of tests are continued for up to 10 000 hours. On the basis of the present results a useful operating life well in excess of 25 000 hours may be confidently expected.

9. Typical Counting Circuits

Figure 5 shows a typical transistor drive circuit for speeds of up to 100 kc/s. The negative-going edge of the output pulse triggers the next counting stage. On this pulse the much steeper negative-going edge of the flip-flop output is coupled to the diode D3. The time delay between the input pulse and the output pulse of a counting stage is reduced to 400 ns by the same arrangement. A circuit for counting frequencies up to 1 Mc/s is shown in Fig. 6. To get optimum pulse shapes, the differentiating network is of the *LCR* type.

10. Read-out Circuit

Although a direct read-out of the position of the discharge is possible some designers of modern counters prefer numerical displays. A very simple circuit has been suggested, using positively biased output electrodes. The positive voltage is set a few volts less than the burning voltage and is not sufficient for an ignition between output electrode and cathode. When the discharge is switched over to the corresponding cathode the voltage is lowered close to the voltage of the plasma around the output electrode. This voltage drop is in the order of 25 to 30 volts. A direct

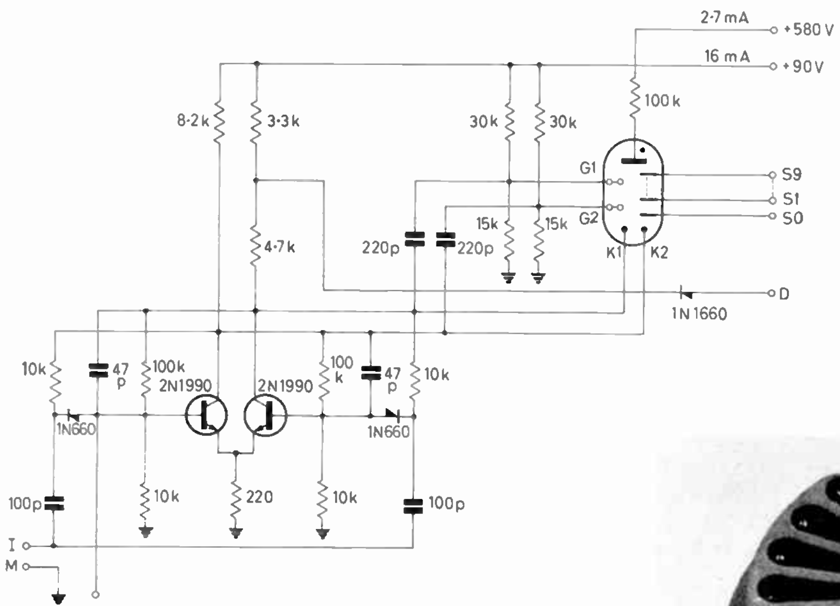


Fig. 5. Counting stage for frequencies up to 100 kc/s.

drive to a 'Nixie' tube is possible but gives less or more background glow due to the small voltage difference between burning and non-burning numerals. In addition the counting process may be upset.

Thus a special numerical display tube was designed. It consists of a ten-chamber gas-amplifier system and a 'Nixie' type display system in one envelope (Figs. 7 and 8).

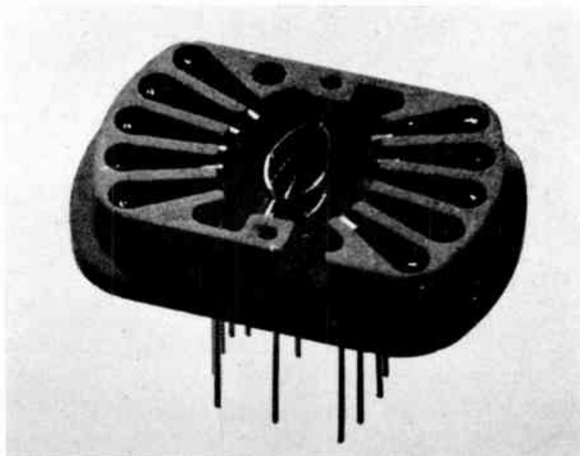


Fig. 7. Ten-chamber amplifier.

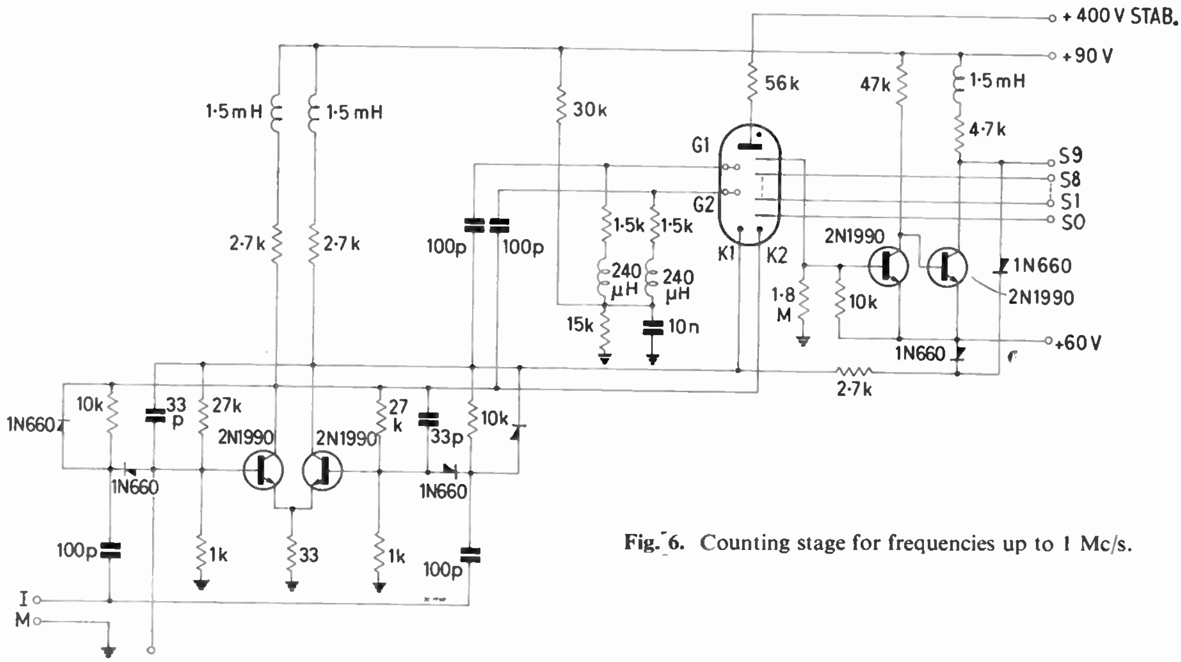


Fig. 6. Counting stage for frequencies up to 1 Mc/s.

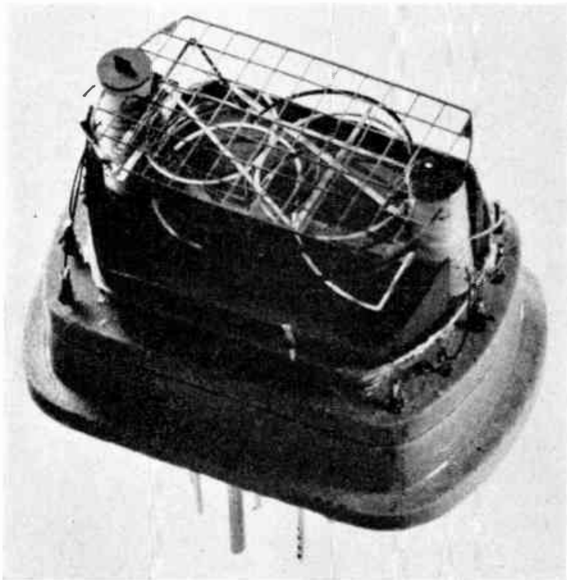


Fig. 8. Layout of a typical numerical display tube.

In the centre of the ten-chamber system an auxiliary discharge is burning and produces ions and electrons. Each chamber has a grid-controlled entrance and an anode which is connected to one of the ten numerals.

Normally the grid is slightly negative to the reference electrode (about −3 V). When the striking voltage of the chamber anode is reached (about 250 V) the chamber is closed (Fig. 9). A small fraction of the positive output pulse of the counting tube is sufficient to open the chamber; the chamber anode takes over the discharge and its voltage drops. Thus the striking voltage between the connected numerical cathode and the main anode is reached and a glow occurs.

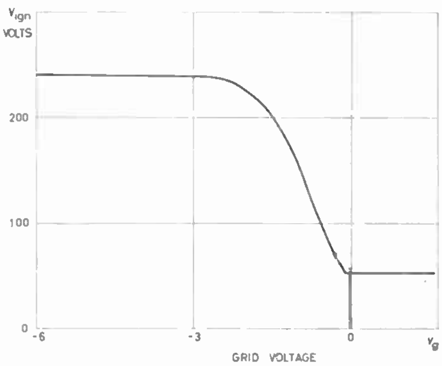


Fig. 9. Typical graph of striking voltage of a chamber anode vs. control grid voltage.

11. Conclusion

Transfer considerations and experiments led to a new transfer mechanism with one auxiliary cathode between two main cathodes and without the anode stray capacitance charging. They resulted in a high-speed counting tube with an upper speed limit of 1 Mc/s. One advantage is the low drive voltage requirements which enable a simple direct transistor drive to be used with it. As the gas filling is hydrogen special care had to be taken in the design of the cathode limbs to yield a stable discharge. At high counting speeds the useful anode current range is reduced, especially when statistical events must be counted.

The tube is read out by aid of ten probes positioned near the ten main cathodes. The voltage output could be optimized with negligible reaction on the counting process. Life tests show very reliable counting and good stability in stand-by operation over prolonged periods.

12. References and Bibliography

1. S. Flügge, "Handbuch der Physik", Vol. XXII, pp. 120ff. (Springer-Verlag, Berlin, 1956).
2. J. R. Acton and J. D. Swift, "Cold Cathode Discharge Tubes", Ch. 18, Dekatrons (Heywood, London, 1963).
3. K. Apel, "ECT 100—Ein neues Bauelement für die Zähltechnik", *Elektronische Rundschau*, **15**, pp. 57–61, 1963.
4. W. Stedtnitz, "Betriebseigenschaften der Zählröhre ECT 100 bei hohen Frequenzen", *Elektronische Rundschau*, **15**, No. 9, pp. 466–8, 1963.
5. Technical Document on Counting Tube ECT 100, (Elesta A.G., Bad Ragaz, Switzerland).

Manuscript received by the Institution on 7th February 1964. (Paper No. 924.)

DISCUSSION

Mr. J. M. Glackin: With regard to life-testing, the author has stated that very good life-test results have been obtained during static operation. The experience of manufacturers of devices similar to the one described in the paper is that the most critical life-test condition is operating on two adjacent cathodes alternately for approximately 8 hours for each period. Has the ECT 100 been life-tested in a similar manner?

Dr. Apel (*in reply*): Similar observations have been made on conventional counting tube designs with gas mixtures where sputtering and contamination of neighbouring cathodes takes place. In the ECT 100 with its hydrogen filling and the hollow cathodes these effects are very small. In addition the improved counting mechanism with its high inherent security is not influenced by slight changes of the maintaining voltage of neighbouring cathodes. Similar tests to those mentioned by Mr. Glackin have not shown results significantly different from those obtained with static operation.

U.D.C. 621.396.965.4 + 534.88

The Use of an Effective Transmission Pattern to Improve the Angular Resolution of 'Within-pulse' Sector-scanning Radar or Sonar Systems

By

D. C. COOPER, Ph.D.†

Reprinted from the Proceedings of the Symposium on "Signal Processing in Radar and Sonar Directional Systems", held in Birmingham on 6th–9th July, 1964.

Summary: The paper describes the application of a double transmission technique to linear arrays in order to obtain a directional response with a small beamwidth. An ideal system is analysed and experimental results obtained with a model of such a system are given.

The theoretical and experimental results show that a reduction of the array beamwidth by a factor of two or four may be achieved by the use of additive or multiplicative processing respectively of the received signals.

A practical time spaced transmission scheme is proposed and its application to within-pulse scanning radar systems is considered.

List of Symbols

A	amplitude of transmitted carrier
α	attenuation factor
B	amplitude of received carrier at the output of one array element
d	element spacing along array
f_s	array scanning frequency
E	energy radiated during one transmission cycle
N	number of elements in the array
n	element identification index
P_N	noise power at receiver output
R	range of target from element 1
S_1	array output due to transmission from element 1 only
S_N	array output due to transmission from element N only
S_A	output of the additive processing circuit
S_M	output of the multiplicative processing circuit
x	bearing parameter $= \dfrac{\pi d \sin\theta}{\lambda}$
θ	bearing of target from normal to array
λ	operating wavelength
ω	operating angular frequency

1. Introduction

In recent years a number of authors[1,2,3] have described various electronic methods for rapidly scanning the directional response patterns of linear arrays. The high scanning rates, which can be achieved with these techniques, have made possible a new approach to the design of pulsed radar or sonar systems[1,2]; this approach involves the use of a receiving array which is caused to scan the complete sector of interest repeatedly at a rate which is sufficient to enable all angular positions in the sector to be examined at least once during any time interval equal to the transmitted pulse duration.

The transmitting array of such a system is designed to radiate the pulse energy over the entire sector of interest. The returns from the various parts of the sector are pulses which are sampled at least once by the receiving array. Thus target range is available in the usual form of a time delay, with some ambiguity due to the sampling process, and the angular information is provided by the array scanning system.

For the basic 'within-pulse' scanning system outlined above it is obvious that the angular resolution (i.e. the correct identification of multiple targets at approximately the same range and bearing) is entirely defined by the directional properties of the receiving array. This leads to some loss of resolution capability when the within-pulse scanning system is compared with a slowly scanning system of the conventional type for which the directional properties are given by the product of the transmitting and receiving array patterns. Obviously it is desirable to attempt to improve the angular resolution properties of the within-pulse scanning system by whatever means are available. This paper is concerned with the use of

† Electronic and Electrical Engineering Department, University of Birmingham.

separate, and identifiable, transmissions from the end elements of a linear array and the subsequent processing of the separate returns from any target to obtain the desired improvement in resolution.

The double transmission system will be described in two stages. First the performance of an idealized system will be analysed and some experimental results will be given. This will be followed by a description of a suggested practical system in which some of the obvious limitations of the idealized system are avoided.

2. The Idealized Double Transmission system

2.1. *The Use of Orthogonal Polarizations*

For the purposes of the analysis given in this section we will suppose that the separate transmissions from the end elements of an array can be identified and received separately by virtue of their polarization.

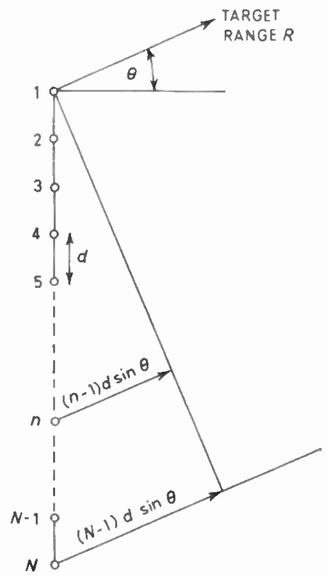


Fig. 1. Geometry of the N-element array.

Thus we consider that the transmissions are orthogonally polarized and assume that all targets reflect energy polarized in the plane of the incident polarization with a reflection coefficient which is independent of polarization.

It must be noted that identification by means of the plane of polarization cannot be postulated as having any significance in the sonar field and in addition the idealized system using polarization has many obvious and serious limitations as a practical radar technique. The idealized system serves to introduce the principles involved in the use of double transmission in a simple way, and we will consider a practical system afterwards.

The geometry of the idealized system using an N-element array is illustrated in Fig. 1. Elements 1 and N are assumed to radiate a carrier waveform represented by $A \cos \omega t$, and it is assumed that a single target is positioned at a range R from element 1 on a bearing θ measured with respect to the normal to the array. The range R is assumed to be very much greater than the length of the array.

In the receiving role all the elements of the array will be used and their outputs will be added without the introduction of time or phase shifts. This means that in the analysis it is assumed that the receiving array is intended to receive the two orthogonally polarized signals arriving from a normal direction. The complications introduced for the purpose of rapidly scanning the direction of maximum response will be considered later.

The array outputs which originate from the separate transmissions can now be evaluated.

The output arising from transmission by element 1 is,

$$S_1 = \sum_{n=1}^{N} B \cos \left[\omega t - \frac{2\pi}{\lambda} \{2R + (n-1) d \sin \theta\} \right] \quad \ldots(1)$$

and the output arising from transmission by element N is,

$$S_N = \sum_{n=1}^{N} B \cos \left[\omega t - \frac{2\pi}{\lambda} \{2R + (N+n-2) d \sin \theta\} \right] \quad \ldots\ldots(2)$$

where d = element spacing
B = constant depending on transmitted amplitude A, the attenuation in the medium, target echoing area, etc.

If we put

$$x = \frac{\pi d \sin \theta}{\lambda}$$

and

$$\phi = \frac{4\pi R}{\lambda} + 2(N-1)x$$

we obtain the simple expressions,

$$S_1 = B \cos[\omega t - \phi + (N-1)x] \frac{\sin Nx}{\sin x} \quad \ldots\ldots(3)$$

$$S_N = B \cos[\omega t - \phi - (N-1)x] \frac{\sin Nx}{\sin x} \quad \ldots\ldots(4)$$

Thus the separate outputs resulting from the orthogonal transmissions have the same amplitude and this has the expected $\sin Nx/\sin x$ dependence on the bearing parameter x, but the phases of the outputs differ by $2(N-1)x$ radians which is again to be expected since the path lengths involved are not equal unless θ, and therefore x, is zero. It will now be apparent that although we have two signals which each have an amplitude dependent on x that corresponds to the normal sensitivity pattern for the linear array we have in addition a phase difference, dependent on x, which can be utilized in a number of ways to give the overall system a narrower directional response pattern. In referring to the width of the directional response pattern we will use the term 'beamwidth' and adopt the usual practice of measuring this to the -3 dB points of the response.

2.2. *Additive Signal Processing*

The direct addition of the signals which have been separately transmitted and received is a very simple method for obtaining a narrower beamwidth for the double transmission system.

Thus the summed output is

$$S_A = S_1 + S_N = 2B\cos(\omega t - \phi)\cos(N-1)x\frac{\sin Nx}{\sin x} \qquad \ldots\ldots(5)$$

The form of this directional response is indicated in Fig. 2, for which a nine-element array has been assumed. It will be seen that the beamwidth for the double transmission system with additive signal processing is approximately one half of that for the array used in the usual receiving manner. The directional response of the array is shown in Fig. 2 for comparison purposes.

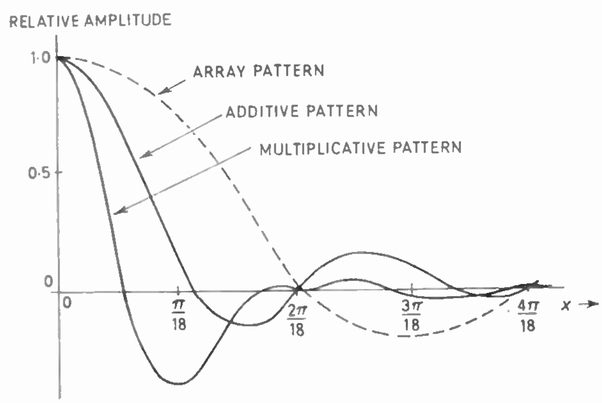


Fig. 2. Theoretical response patterns for an idealized double transmission system.

It will be noticed that the directional response of the double transmission system with additive signal processing comprises the product of the $\sin Nx/\sin x$ receiving pattern of the array with the term $2\cos(N-1)x$ which is the transmission pattern that would result from the use of similar polarizations for the transmissions. Thus the same result could be obtained quite simply by letting the addition process be the result of the superposition of radiation in the medium. However, in that case there will be regions in the medium which will not receive any appreciable transmitted energy whereas in the orthogonal polarization scheme the transmitted energy will be distributed fairly uniformly over the sector of interest. This means that by having separate transmissions the direction of maximum response may be changed, that is the beam may be scanned, and the improved response pattern can be maintained without any adjustment being made to the relative phase of the transmissions.

Thus all phasing adjustments can be made on reception and this is an essential feature of any scheme which is to be applied to a within-pulse scanning system.

2.3. *Multiplicative Signal Processing*

An alternative to additive processing is the multiplication of the separate received signals and the removal of high frequency components by the use of a low-pass filter. In order to find the output of the filter we formulate the expression $S_1 \times S_N$ and extract the zero frequency terms from the result.

The zero frequency component in the product of S_1 and S_N is found to be

$$S_M = B^2\frac{\sin^2 Nx}{\sin^2 x}\cos 2(N-1)x \qquad \ldots\ldots(6)$$

The directional response pattern indicated by the above expression is shown in Fig. 2 for a nine-element array. It will be seen that multiplicative processing of the signals received from an isolated target gives a directional response pattern with a beamwidth of approximately one quarter of that for the array used in a receiving role only. This narrow beam is accompanied by first side-lobes of negative polarity with a relative amplitude approaching 40%, but the remaining side-lobes are very small indeed. The pattern which results from multiplicative processing in this case will be seen to have some similarity to the output of the multiplicative array processing systems described by Welsby and Tucker.[4]

3. The Experimental Model of the Idealized System

Experimental results have been obtained using orthogonally polarized transmissions at a frequency in the microwave X band with an arrangement illustrated in the block diagram of Fig. 3.

In this arrangement the transmitting elements were two waveguide horns, one giving vertically- and the other horizontally-polarized radiation. The transmitting horns were placed above the ends of two slotted waveguide arrays which were themselves placed one above the other. One of the arrays was

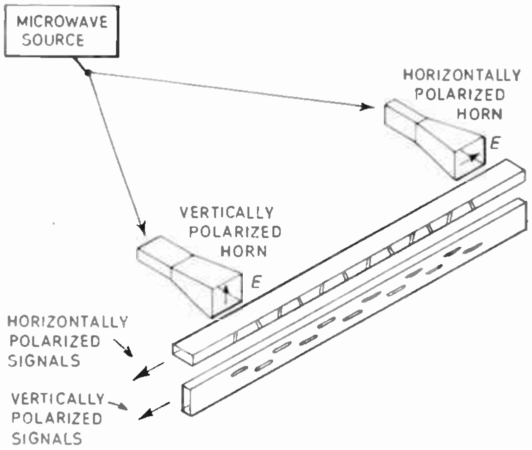


Fig. 3. Arrangement of the experimental arrays.

designed to receive vertically-polarized signals while the other received horizontally-polarized signals. In this way the individual polarizations were received substantially independently and the correct geometrical arrangement was realized for the measurement of the directional response of the system in the horizontal plane.

The beamwidth of the radiation from the transmitting horns was large compared with the beamwidth for the receiving arrays, all beamwidths being those measured in the horizontal plane. The radiated power density could be considered as being independent of bearing in the region covered by the experimental results since these were taken only in the small range determined by the narrow directional response of the complete system.

The experimental results were obtained by using a single target and changing the bearing of this target while maintaining its range constant. When additive signal processing was required, the separate outputs of the two arrays were added by using a 3 dB waveguide directional coupler. The phases of the signals were adjusted for maximum response in the direction in which the receiving arrays were 'looking'. The summed output was then detected and this detected output was recorded as the target bearing was changed. Multiplicative processing was performed at an intermediate frequency of 40 Mc/s after separate mixing of the array outputs with a common local oscillator. Again the phase settings were adjusted to give maximum response in the direction in which the arrays were 'looking' and the d.c. output of the multiplier was recorded as a function of the target bearing.

The results obtained with additive and multiplicative processing are shown in Fig. 4, together with the directional response of an array used in the receiving role only. It can be seen that the theoretical predictions are supported by the experimental results and the measured 3 dB beamwidths are as follows:

Beamwidth of one array used solely for reception $= 3{\cdot}9°$

Beamwidth of double transmission system with additive processing $= 1{\cdot}9°$

Beamwidth of double transmission system with multiplicative processing $= 1{\cdot}1°$

4. A Practical Double Transmission System

4.1. *The Use of Time-spaced Pulses*

It has already been pointed out that the use of orthogonal transmitted polarizations will not produce the desired results in the radar situation owing to the complex reflecting properties of targets, and in the case of sonar systems orthogonal polarizations cannot be propagated.

We will now consider a double transmission system in which two phase-locked carrier pulses are radiated, one from each transmitting element, at different times. In this case the signal received by a single array will be in the form of two time-spaced pulses when an isolated single target is considered. Obviously for a target on the normal to the array the first pulse received will be that resulting from the first transmitted pulse and this will also be true for targets well away from the normal providing that the propagation time for a distance equal to the length of the array is much less than the pulse spacing.

4.2. *Additive Processing in the Double Pulse System*

If we wish to apply additive processing to the received pulses we must obtain approximate time coincidence and exact phasing between them when they are arriving from the direction in which we wish to maximize the system response. Obviously this

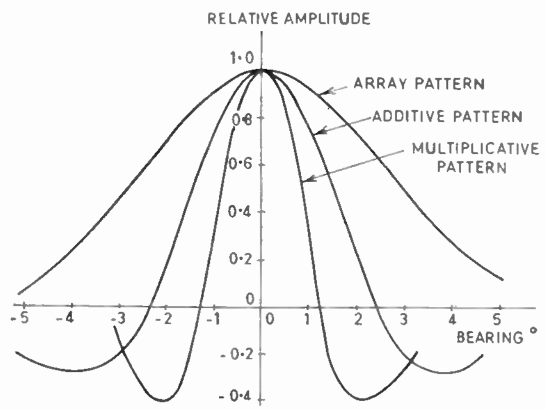


Fig. 4. Experimental response curves.

direction should coincide with the maximum response direction for the receiving array and initially we will consider that this direction is the normal to the array.

The appropriate time and phase compensation can be obtained by the use of the circuit arrangement shown in Fig. 5, in which the delay introduced must be such that the first received pulse emerges from the delay line and is applied to one input of the adder, in phase with and approximately in time coincidence with, the second received pulse which is applied immediately to the other input of the adder. The output pulse which is thus formed will have an r.f. amplitude which depends on the target bearing in exactly the same manner as does the output of the idealized system with additive signal processing, and hence eqn. (5) is applicable in this case.

However the output of the circuit of Fig. 5 will consist of three separate pulses when two are received and the first and last of the three output pulses will simply be the first received pulse and the second after delay.

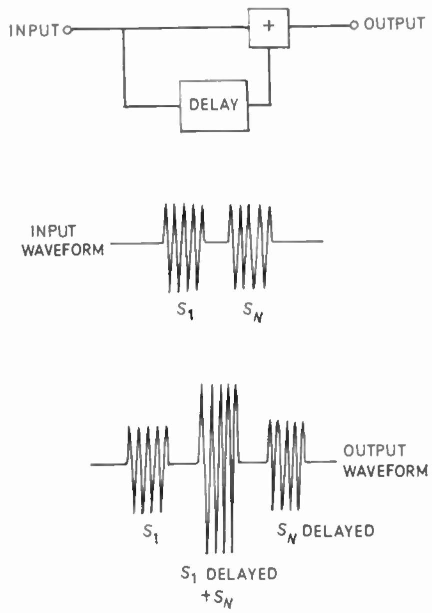


Fig. 5. Double-pulse processing circuit and waveforms.

Hence only the centre pulse will have the desired narrow directional response to target bearing and the other two will be defined by the directional response properties of the receiving array alone.

The method of additive processing just described has in fact introduced what might be called range side-lobes whose peak amplitudes are one half of the peak amplitude of the required central pulse and whose beamwidths are approximately twice that for the central pulse. The way in which these outputs vary with target bearing is best illustrated by a three-dimensional model and such a model is sketched in Fig. 6.

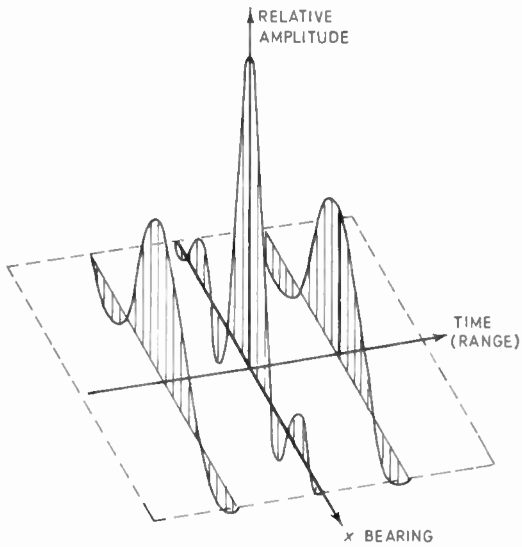


Fig. 6. Response of double-pulse system with additive processing.

It will be realized that in a multiple target situation, interaction between returns from targets at different ranges can now occur and the bearing resolution will depend on the relative phases and amplitudes of the interacting signals. However, it is reasonable to suggest that some improvement in resolution will be obtained when the number of targets contributing to any interaction is small.

4.3. *Multiplicative Processing in the Double Pulse System*

A multiplicative processing circuit can be obtained by replacing the adder of Fig. 5 by a multiplier and low-pass filter. In this case only one output pulse is formed when one isolated pair of pulses is received and the amplitude of the output pulse changes with change of target bearing in a manner which corresponds to that obtained with multiplicative processing in the idealized system. Thus eqn. (6) determines the amplitude of the output pulse and the beamwidth for the system is approximately one quarter of that for the receiving array alone.

The absence of range side-lobes in the case of multiplicative processing applies only to the case of an isolated target. When a number of closely-spaced targets are present, interaction will occur if delayed pulses from one target coincide with pulses received from other targets. The situation will be affected by any relative motion of the targets and in this event the resolution capabilities of the system will be dependent on the number of statistically independent observations that can be made.

Shaw and Davies[5] have investigated the bearing resolution capabilities of multiplicative and additive processing systems for the case of two closely spaced targets of similar echoing areas. The comparison between the processing systems assumes that the additive system is followed by square-law envelope demodulation so that the overall response of both systems is then of a square-law nature. These authors have shown that under the stated conditions the multiplicative system is never worse than the additive one and it gives an average improvement in bearing resolution of approximately 25%.

5. The Application of Within-pulse Scanning to the Double-pulse System

The double-pulse system is obviously inefficient if it is designed to 'look' in one direction only. This is because the energy radiated from the transmitting elements will be distributed over a large sector and a large part of this energy will be wasted.

However, within-pulse scanning of the receiving system can make effective use of all the radiated energy and we will now consider the modifications required to permit the double-pulse system to scan a sector during one pulse duration.

Scanning a complete sector in a time equal to the duration of one transmitted pulse means that the two pulses of the double-pulse system are actually received during different, and preferably successive, scans of the receiving system. Successive scans are preferable since this means that the transmitted pulses are contiguous and the spread of any range side-lobes is then kept to a minimum. Since we wish to utilize the phase information contained in the pulse pair it is apparent that any scanning system adopted must be arranged to repeat its scans in a very precise manner, and the system described by Cottony and Wilson[3] appears to be more amenable to the application of precise control than does the earlier system described by Tucker, Welsby and Kendal[1] or by Davies.[2]

Hence we will consider that the receiving array pattern is scanned by the application of fixed frequency shifts to the output of each array element, the nth element being shifted by a frequency $(n-1)f_s$, where f_s is the required scanning frequency. In this way a phase shift which increases linearly with time is applied to the output of each element and when the incremented frequencies are phase locked the incremental phase shifts are such that the directional pattern of the array is swept repeatedly over a sector. As a consequence of the phase locking the scans of the system will be repeated in a precise manner.

The application of this scanning technique to the double-pulse system will enable two sampled received pulses to be obtained which originate from the separate transmitted pulses, and the relative phase information of the sample pulses will not have been destroyed by the scanning operation.

This now leaves the problem of getting approximate time coincidence and correct phasing of the samples. Time coincidence can be obtained by the use of a delay circuit such as is shown in Fig. 5 and all that remains to be done is to apply the phase correction; this will of course depend on the direction in which the system is intended to be looking at the particular instant. This correction corresponds to the difference between the lengths of the propagation paths of the individual pulses from their transmitting elements

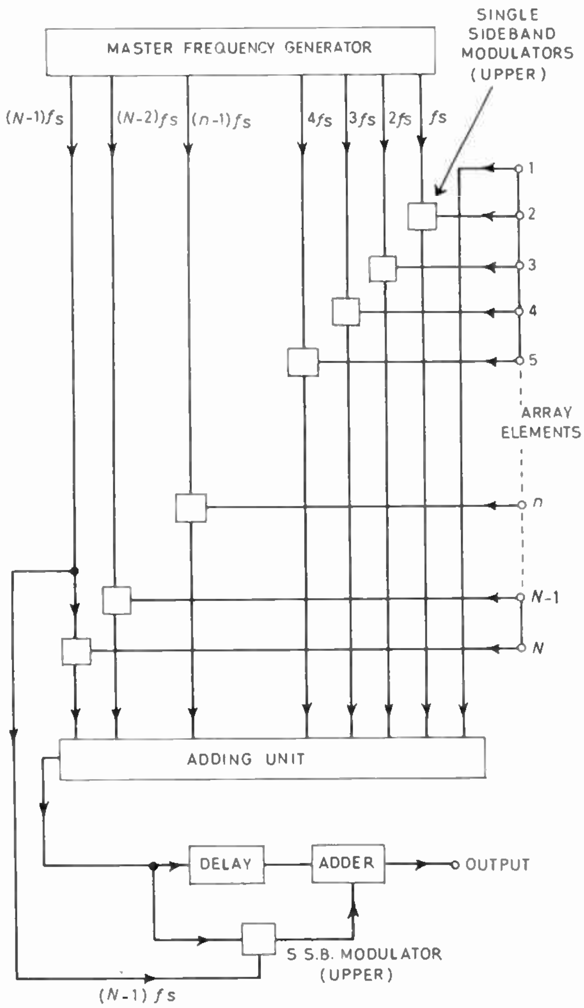


Fig. 7. Block diagram of double-pulse scanning system.

to a target situated in the direction under consideration, the return propagation path being the same for each pulse.

We have already noted that the phase correction must vary in sympathy with the scanning of the directional pattern of the receiving array and we see that the appropriate time varying phase shift has already been provided between the end elements of the array when used in its receiving role. Hence the phase correction can be provided by shifting the frequency of one of the inputs to the adder, or multiplier, in the circuit of Fig. 5, by an amount $(N-1)f_s$; this frequency has already been used for scanning the linear array. In this way the improved beamwidth of the double-pulse system can be maintained while the entire sector covered by the transmitted energy is scanned at a rate equal to the reciprocal of the duration of one pulse. Obviously the same technique can be applied irrespective of whether or not the final processing is additive or multiplicative.

A block diagram of the proposed double-pulse scanning system is shown in Fig. 7. For convenience in explanation this particular arrangement of frequencies has been used, but economy of equipment would result from the use of a symmetrical arrangement using only half the number of local frequencies, lower sidebands being taken on one side of the array, and upper sidebands on the other.

6. Signal/noise Performance of Within-pulse Scanning Systems

6.1. *The Additive Double-pulse System compared with the Single-pulse System*

For the purposes of the comparison we wish to make we will assume that the same linear array is used in either system and concentrate attention on the signals received from a single target in the centre of the 'beam' and at constant range.

The total energy transmitted per pulse in the single-pulse system and per pulse pair in the double-pulse system will be assumed to be E. On reception we will then have a single pulse of energy αE in one case and a pair of pulses of total energy αE in the other case, where α represents the transmission and scanning loss, etc.

If the noise power at the output of the receiver in the single-pulse case is P_N this will also be the noise power output in the double-pulse case providing all the pulses are of the same length and amplifier bandwidths are therefore similar. It is reasonable to assume that there will be no correlation between the noise outputs during the reception of the separate pulses in the double-pulse system and in this case the noise contributions are additive on a power basis.

Taking the pulse duration after reception and scanning as τ we obtain the following expressions:

Signal/noise power ratio at receiver output for single-pulse system

$$= \frac{\alpha E/\tau}{P_N} = \frac{\alpha E}{P_N \tau} \qquad \ldots\ldots(7)$$

Signal/noise power ratio (peak) for additive processing in the double-pulse system

$$= \frac{\left(2\sqrt{\frac{\alpha E}{2\tau}}\right)^2}{P_N + P_N} = \frac{\alpha E}{P_N \tau} \qquad \ldots\ldots(8)$$

Thus the signal/noise performance for the additive double-pulse case is the same as that for the single-pulse system.

6.2. *The Multiplicative Double-pulse System compared with the Single-pulse System*

When we wish to consider the performance of the multiplicative double-pulse system we obviously must define the signal/noise ratio of the video output in some suitable fashion and compare this signal/noise ratio with that produced in the single-pulse system when the output has been obtained in video form by using a suitable demodulation process. It will be convenient to assume that a square-law envelope demodulator is used for this purpose since it is well known[6] that this is the optimum law for detection purposes, and in addition we will obtain a square-law overall response for both systems.

If we define the video signal/noise voltage ratio as

$$\frac{\text{change in mean level due to presence of signal}}{\text{r.m.s. noise in the absence of signal}}$$

it can be shown that for the single pulse system with square-law envelope demodulation this is equal to the r.f. signal/noise power ratio.

In the case of multiplicative processing it has been shown[7] that the video signal/noise ratio we have adopted is $\sqrt{2}$ times the input r.f. signal/noise power ratio at each input to the multiplier, providing the noises at the inputs are uncorrelated.

Using the above results we obtain the following expressions:

Video signal/noise voltage ratio of single-pulse system

$$= \frac{\left(\frac{\alpha E}{\tau}\right)}{P_N} = \frac{\alpha E}{P_N \tau} \qquad \ldots(9)$$

Video signal/noise voltage ratio of double-pulse system with multiplicative processing

$$= \sqrt{2}\,\frac{\left(\frac{\alpha E}{2\tau}\right)}{P_N} = \frac{1}{\sqrt{2}}\,\frac{\alpha E}{P_N \tau} \qquad \ldots\ldots(10)$$

It will be observed that the double-pulse multiplicative system requires the energy transmitted to be increased by a factor of $\sqrt{2}$ in order to give the same video signal/noise ratio as the single pulse, or the additive double-pulse system. This corresponds to increasing the average power level by 1·5 dB.

7. Conclusions

It has been shown that double transmissions can be utilized to obtain narrower directional response patterns for within-pulse scanning radar and sonar systems. A comparison of the beamwidths obtained in the two versions of the double-pulse system with the beamwidth of the single-pulse within-pulse scanning system has shown that the additive processing version gives a 2 : 1 reduction and the multiplicative version gives a 4 : 1 reduction.

The improved beamwidth should give a 2 : 1 improvement in bearing resolution for the case of additive processing although the range side-lobes which are present at the output will cause some loss of range resolution. For the multiplicative double-pulse system an improvement of 4 : 1 in bearing accuracy will be obtained, but an improvement of the same amount in bearing resolution is unlikely to be obtained in most practical situations, where it would appear that the system will provide an average improvement ratio of the order of 2·5 : 1. Range side-lobe effects are also present in the multiplicative version and range resolution will again be affected.

The signal/noise performances of the double-pulse systems are very similar to that of the single-pulse arrangement and it would appear that the use of the double-pulse technique is one way in which an improved bearing resolution may be obtained at the cost of some loss of range resolution. However, the range resolution will be returned to its original value if the duration of the individual pulses of the double-pulse system can be made one half of that for the basic single-pulse system which is our reference. Hence bearing resolution can be improved without loss of range resolution by the use of a greater transmission bandwidth.

8. Acknowledgments

The author wishes to record his appreciation of the work of Mr. J. W. Lyons (a student on the post-graduate course in Information Engineering), who conducted the experimental investigations described in this paper, and many thanks are due to Professor D. G. Tucker and Dr. D. E. N. Davies for helpful discussions concerning the subject matter of the paper.

9. References

1. D. G. Tucker, V. G. Welsby and R. Kendal, "Electronic sector scanning", *J. Brit.I.R.E.*, **18**, No. 8, p. 465, August 1958.
2. D. E. N. Davies, "Radar systems with electronic sector scanning", *J. Brit.I.R.E.*, **18**, p. 709, December 1958.
3. H. V. Cottony and A. C. Wilson, "A high-resolution rapid-scan antenna", *J. Research Nat. Bur. Stand.*, **65D**, No. 1, pp. 101–10, January–February 1961.
4. V. G. Welsby and D. G. Tucker, "Multiplicative receiving arrays", *J. Brit.I.R.E.*, **19**, p. 369, June 1959.
5. E. Shaw and D. E. N. Davies, "Theoretical and experimental studies of the resolution performance of multiplicative and additive aerial arrays", *Proceedings of the Symposium on Signal Processing in Radar and Sonar Directional Systems*, Birmingham 1964, Paper No. 7. *The Radio and Electronic Engineer*, **28**, No. 4, October 1964 (to be published).
6. J. I. Marcum and P. Swerling, "Studies of target detection by pulsed radar", *Trans. Inst. Radio Engrs (Information Theory)*, **IT-6**, No. 2, pp. 59–308, April 1960.
7. D. G. Tucker, "Signal/noise performance of multiplier (or correlation) and addition (or integrating) types of detector", *Proc. Instn Elect. Engrs*, **102C**, pp. 187–90, 1955. (I.E.E. Monograph 120R, February 1955.)

Manuscript received by the Institution on 9th January 1964. (Paper No. 925/RNA 30.)

DISCUSSION

Under the Chairmanship of Dr. E. V. D. Glazier

Dr. R. Benjamin: Non-linear processes in the amplitude domain generally destroy some information and so degrade the signal/noise ratio. However, where the signal/noise ratio is even marginally above an operationally acceptable threshold, non-linear or logical operations may improve the apparent contrast and so produce a more satisfying or intelligible display or permit the use of simpler data-extraction circuits.

Analogous considerations apply to multiplicative or other non-linear or logical operations in the angular domain.

Dr. Cooper (*in reply*): I agree entirely with the comments made by Dr. Benjamin, but I would like to emphasize the fact that the radar or sonar system described in the paper obtains extra bearing information by the use of two transmissions instead of the usual one.

Dr. Benjamin's remarks apply to the processing of this information and the further enhancement of bearing resolution which can be obtained by the use of multiplicative instead of additive processing.

U.D.C. 621.396.969.36

Studies of a Twin-channel Frequency-modulated Echo-location System

By

D. E. N. DAVIES, M.Sc., Ph.D. †
(*Associate Member*)

AND

B. S. RAO, M.Sc., Ph.D. ‡

Summary: This paper describes theoretical and experimental studies of a proposed twin-channel echo-location system using frequency-modulated transmissions, in terms of its possible application to primary-radar systems or target-tracking systems. The proposed system can measure the range and bearing of a target from a fixed aerial system, although in the presence of multiple targets, ambiguities of target position restrict its performance. The system may be adapted, however, for accurate tracking of a target containing a suitable signal source.

List of Principal Symbols

f_{R_A}	range frequency channel A.
f_{R_B}	range frequency channel B.
$T_s = 1/f_s$	modulation period of frequency-modulated transmission.
τ_1	time delay between the transmitted signal and the arrival of the reflected signal at A.
τ_2	time delay between the transmitted signal and the arrival of the reflected signal at B.
Δf	total peak-to-peak frequency deviation of frequency-modulated transmission.
D	distance between receivers A and B.
R_A	range of target from A.
R_B	range of target from B.
C	velocity of electromagnetic waves.
θ	bearing of target relative to normal to array of aerials.
f_b	frequency indicating bearing of target.
ΔR	range resolution.
N	total number of cycles in one modulation period of waveform.

1. Introduction

Studies by Griffin[1] have shown that certain species of bat transmit frequency-modulated ultrasonic waves, receive the reflected signals and process them in order to avoid collision with objects, to find their way about and to catch their prey. The ability of this animal to negotiate difficult and complicated objects and to operate in the presence if high background noise levels, tends to suggest that this type of echo-location system would be of interest to radar engineers. Kay[2] has proposed a possible echo-location system using frequency-modulated transmission which he put forward as a possible explanation of bats' echo-location acuity. In many instances it is not reasonable to discuss complicated adaptive biological processing systems in terms of simple electrical circuits; nevertheless Kay's proposal appears to merit some detailed analysis since it is also being evaluated in the form of an ultrasonic aid for the blind.[3]

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

‡ Birla Engineering College, Pilani, India. Formerly University of Birmingham.

This paper studies this proposed echo-location system with particular consideration of the possible relevance to radar techniques and some of the experimental measurements have been made in the microwave band.

2. Principle of Operation

The principle of operation is an extension of the principles of frequency-modulated ranging systems; it provides as well range information from an aerial system consisting of one transmitting and two receiving aerials. Figure 1 shows a schematic diagram of the simplified form of the double-channel radar. Saw-tooth, frequency-modulated, continuous-wave signals are radiated from the transmitting aerial which is of wide beamwidth covering the entire sector to be investigated. Signals reflected from the targets in this sector are received separately at the two receiving aerials where they are treated in the normal way for frequency-modulated ranging systems. The received signals are modulated with a small sample of the transmitted signal and the difference frequency signal filtered off.

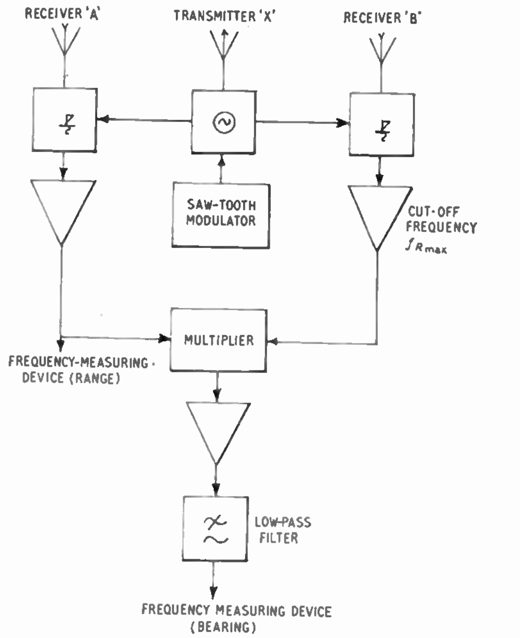


Fig. 1. Block schematic of twin-channel frequency-modulated radar.

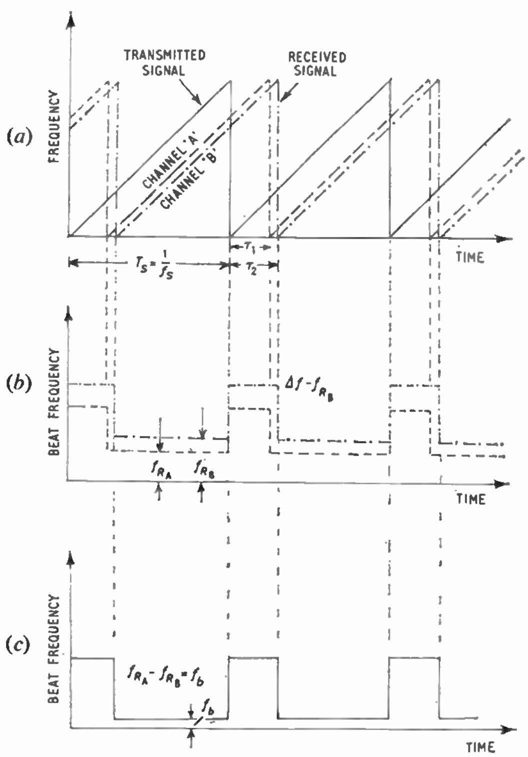


Fig. 2. Twin-channel f.m. radar operation.
(*a*) Transmitted and received signals.
(*b*) and (*c*) Instantaneous frequency differences.
(*b*) Range.
(*c*) Bearing.

Figure 2 shows the form taken by these signals on a frequency-time plot. It can be seen that for a single target the difference frequency signals are f_{R_A} for channel A and f_{R_B} for channel B during most of the sweep duration T_s. For the remaining period $(T_s - \tau_1)$ and $(T_s - \tau_2)$ the instantaneous frequencies in the two channels are $\Delta f - f_{R_A}$ and $\Delta f - f_{R_B}$ as shown. Here Δf is the total frequency excursion of the transmitter. These two higher frequencies may be removed by filtration leaving the range frequencies f_{R_A} and f_{R_B} which are given by:

$$f_{R_A} = \frac{\tau_1}{T_s} \Delta f \qquad \ldots\ldots(1)$$

$$f_{R_B} = \frac{\tau_2}{T_s} \Delta f \qquad \ldots\ldots(2)$$

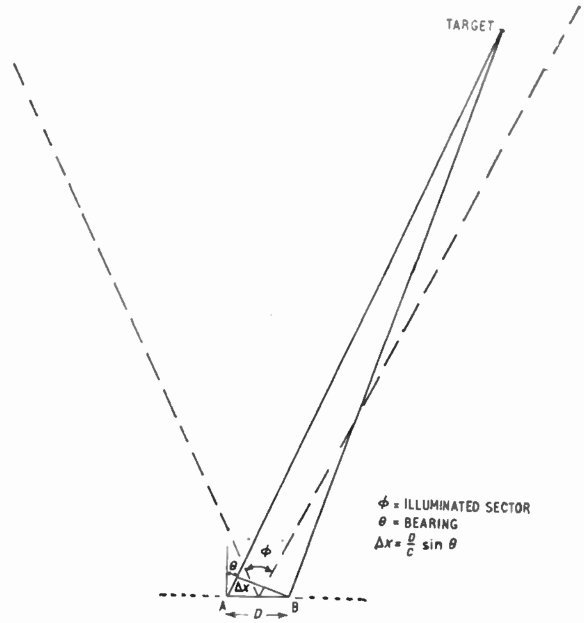


Fig. 3. Bearing evaluation.

Thus the values of these frequencies directly proportional to the time delay between the transmitted signal and its reception at the two receivers. In the case of conventional ranging, the transmitter and receiver will be close together and the frequency f_R is therefore a direct measure of range. In the case concerned we consider the two receiving aerials separated by a distance D with the transmitter located centrally between them. (This is not a necessary condition but a convenient case to consider.) If the range R of the target greatly exceeds D, then it is still reasonable to consider f_{R_A} and f_{R_B} as representing the range of the targets from the two receiving aerials A and B. These ranges will therefore be given by:

$$R_A = \frac{1}{2}\frac{Cf_{R_A}}{\Delta f}T_s \qquad \ldots\ldots(3)$$

$$R_B = \frac{1}{2}\frac{Cf_{R_B}}{\Delta f}T_s \qquad \ldots\ldots(4)$$

However, it can be seen from Fig. 3 that the difference in range $R_A - R_B$ is given by:

$$R_A - R_B = D\sin\theta \qquad \ldots\ldots(5)$$

where θ is the bearing of the target relative to the normal to the array of three aerials. Now if the two bearing frequencies are mixed together in some non-linear device such as a multiplier, then the resultant difference frequency component $f_{R_A} - f_{R_B}$ will be proportional to the sine of the bearing of the target.

$$f_b = f_{R_A} - f_{R_B} = \frac{D}{C}\frac{\Delta f}{T_s}\sin\theta \qquad \ldots\ldots(6)$$

It is relevant to note the approximations in this approach. In the first place it has been assumed that R_A and $R_B \gg D$ and that f_{R_A} and f_{R_B} indicate the range of the targets from the receiving aerials. A more accurate representation of the locus of constant τ_1 is not an arc of a circle centred on A, but an arc of an ellipse with A and X as foci. Secondly it has been assumed that the bearing of the target can be taken to be the same relative to all three aerials. An important simplification in the previous analysis is that due to the use of the concept of instantaneous frequency. The frequencies f_{R_A} and f_{R_B} are actually spectra with their maxima at f_{R_A} and f_{R_B} and their lines spaced at intervals of $f_s = 1/T_s$ on the frequency scale. Detailed analysis of these spectra have been given by Hymans and Lait[4] and have been shown to possess envelopes of a $(\sin x)/x$ form.

It can therefore be seen that the spectral analysis of the difference frequency f_R corresponds to the video waveform of a pulse radar. The bearing information of the proposed twin-channel radar is obtained as two ranges, the range from aerial A and from aerial B. Thus the range information is available from spectral analysis of either f_{R_A} or f_{R_B} and the bearing information from spectral analysis of the bearing frequency f_b.

2.1. *Bearing Measurement*

It is convenient to discuss the performance of the system using the approximate concept of instantaneous frequency at this stage, the significance of the more detailed analysis will be given later.

One of the first problems of bearing measurement is determining the sign of the bearing angle relative to the normal to the aerial system. Since f_b is proportional to $\sin\theta$ the problem of deciding which side of the normal the target is on is that of distinguishing between f_b and $-f_b$, that is finding out which

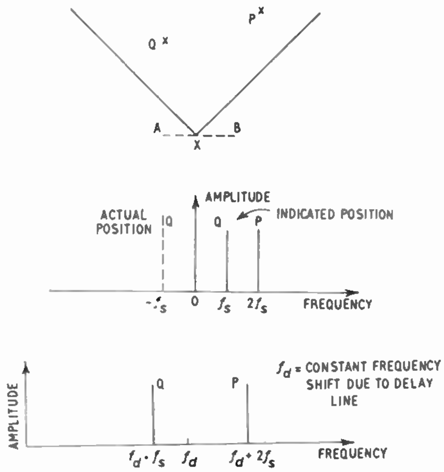


Fig. 4. Bearing ambiguity.

channel contains the higher frequency. The solution to this problem is to frequency-translate one of the input channels to the multiplier by an amount f_d which is greater than the maximum value of f_b. The required frequency translation may be obtained either by conventional single sideband techniques or merely by including an extra delay line between one of the receiving aerials and the mixer. The effect of the frequency translation is shown in Fig. 4.

The spectrum corresponding to the range frequency of a single channel is shown in Fig. 5. The maximum of the spectrum is at f_R and the spectral lines are at multiples of the modulation frequency f_s. The expression for the Fourier spectrum of this signal as analysed in reference 4 is:

$$\sum_k \frac{\sin\frac{\tau}{2}[k\omega_s - \Delta f f_s(T_s-\tau)]}{\frac{\tau}{2}[k\omega_s - \Delta f f_s(T_s-\tau)]} \times$$

$$\times \cos\{k\omega_s t - [\omega_0\tau - \tfrac{1}{2}k\omega_s(T_s-\tau)]\} \quad \ldots(7)$$

where k is an integer, ω_0 is the centre frequency of the transmitter and $f_s = 1/T_s$. The half-power width of

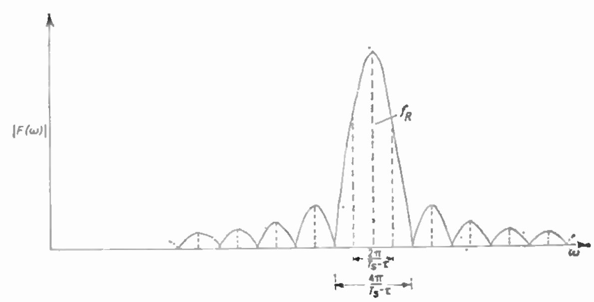


Fig. 5. Beat signal spectrum.

the main lobe of the spectrum is $1/(T_s-\tau)$ and this width is a function of τ. In most frequency-modulated ranging systems it is usual to make $\tau_{max} \ll T_s$ and in this event the half-power width of the spectrum $= 1/T_s = f_s$. The range resolution of such a system is the minimum frequency separation of two such spectral lobes corresponding to two targets adjacent in range. Accurate specification of the separation criteria is very complicated owing to the problem of interpolating the complex spectrum from very few spectral lines. However, it is not unreasonable to stipulate a minimum separation of f_s for $\tau_{max} \ll T_s$. This therefore corresponds to a range resolution given by

$$\Delta R = \frac{C}{2\Delta f} \qquad \ldots\ldots(8)$$

The importance of the above relationships to bearing measurements is that since bearing is measured as the difference between two ranges, and range is measured in units of ΔR, the bearing accuracy will clearly be dependent upon this quantity. If we consider that the range frequencies can be represented by multiples of the modulation frequency:

$$f_{R_A} = K_A f_s \quad \text{and} \quad f_{R_B} = K_B f_s \qquad \ldots\ldots(9)$$

then the bearing frequency will also be a multiple of f_s.

$$f_b = K f_s \qquad \ldots\ldots(10)$$

Hence the bearing accuracy θ_a will be of the order of the angle corresponding to $f_b = f_s$, which from eqn. (6) is given by:

$$\theta_a \simeq \sin^{-1}\left[\frac{C}{\Delta f D}\right] \qquad \ldots\ldots(11)$$

Discussing this effect in terms of the spectrum of the bearing signal f_b; it can be seen from Fig. 2 that the form of the bearing signal will be an interrupted constant frequency for a single target. Therefore the spectral components can be calculated from the results of the range spectra and clearly will also have the form of a $(\sin x)/x$ envelope. In this case the bearing accuracy which is mainly a problem of interpolation will be of the order of the separation of the lines f. which gives the same result as in eqn. (11).

2.2. *Bearing Resolution*

In the previous sections the authors have referred to the range resolution of a frequency-modulated radar in terms of the ability to separate two or more targets closely separated in range but have not referred to resolution in bearing. At first the problem would appear to be analogous to that of range resolution and merely depend upon suitable separation between the bearing spectral responses of the two targets. However, this is unfortunately complicated by the presence of bearing ambiguities in the form of unwanted cross-product terms in the bearing spectrum.

Consider the case of two targets P and Q in the medium. Channel A will indicate their presence by two peaks in the range spectrum at $f_{R_{AP}}$ and $f_{R_{AQ}}$, for instance. Similarly channel B will have peaks at $f_{R_{BP}}$ and $f_{R_{BQ}}$. These two spectra are then fed to the two inputs of a multiplier or modulator and the resultant difference frequency components will be selected by filtration. It can at once be seen that there are four components given by:

$$\begin{aligned} f_{b11} &= f_{R_{AP}} - f_{R_{BP}} \\ f_{b22} &= f_{R_{AQ}} - f_{R_{BQ}} \\ f_{b12} &= f_{R_{AP}} - f_{R_{BQ}} \\ f_{b21} &= f_{R_{AQ}} - f_{R_{BP}} \end{aligned} \qquad \ldots\ldots(12)$$

It can be seen that the first two terms represent the wanted bearing frequencies corresponding to the correct bearings of the targets and given by the self-products. The remaining two terms are cross products and represent incorrect bearings of the targets. At first it might be assumed that the unwanted terms are due to the use of a non-linear device such as a multiplier in order to obtain the bearing spectrum, but this is not so. The multiplier is not being used as a demodulator but merely to obtain frequency differences. The cross products are inherent in the two-channel ranging system and occur in the same way if pulses are used to obtain range information. In fact since it is most convenient to explain this effect in terms of pulse transmissions this approach will be used. Also, since certain periodic pulse and frequency-modulated wave-

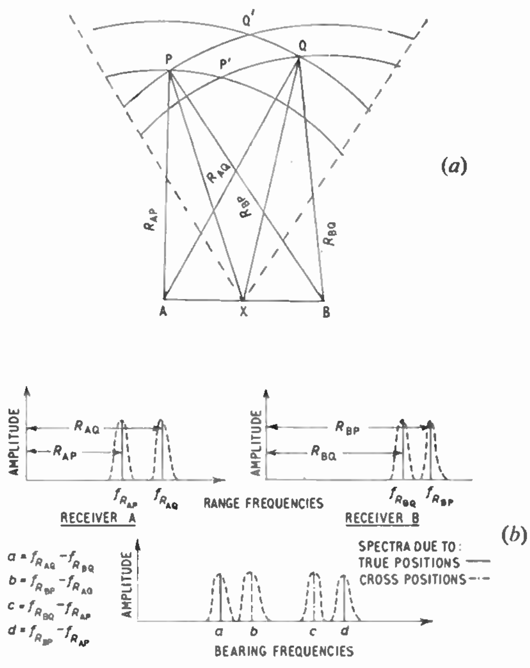


Fig. 6. Bearing resolution (ambiguity).

forms of the same bandwidth differ only in the phases of the various frequency components, this approach is quite valid and general.

Figure 6 shows two targets P and Q and the corresponding range and bearing spectra. It is just the same to consider the range spectra as video pulse waveforms. The problem is that in general there is no way of knowing which pair of frequencies (or pair of pulses) should be compared in order to obtain bearing information. It can also be seen that the cross product frequencies correspond to the image positions P′ and Q′ in the target space. Since there are four possible bearings for two targets, it would appear that there are $^4C_2$ ways of choosing the correct ones. This, however, is not so; if the real targets had been at P′ and Q′ it would have resulted in the same bearing spectrum, but if at some other combination such as P and Q′, there would have been only two frequencies in the bearing spectrum. This is because both P and Q′ are at the same range relative to B.

It can therefore be concluded that if there are n targets in the medium there will be n different range frequencies in each of the two channels (assuming that the targets are separated by at least one range resolution) and the bearing spectrum will contain up to n^2 frequencies, corresponding to n correct bearings and $n(n-1)$ incorrect bearings. Clearly some method of distinguishing the correct bearing frequencies is essential if the system is to work in a multiple-target environment.

2.3. *Methods of Resolving Multiple-target-bearing Ambiguity*

If we first consider the simple cases of two or three targets it is clear that target amplitude could be used to distinguish the self-product from the cross-product terms of the bearing signal provided that the target echoes are not of the same signal strength.

The same sort of conclusions may be drawn about the study of the Doppler spreading of the components of the bearing spectrum in order to determine the self-product components of the spectrum. In suitable cases for a few targets with suitable velocities it is possible to distinguish the true bearings by examination of the Doppler shifting and spreading of the resultant spectra.[5]

However, for any realistic multiple target environment it is clear that the above approaches would be of little or no practical use. This result holds true irrespective of whether the ranging is carried out by frequency-modulated or pulse techniques.

An alternative approach to the elimination of the bearing ambiguities is to consider the effect of tilting the base line of the three aerials as shown in Fig. 7. First of all, the signals from the two targets P and Q

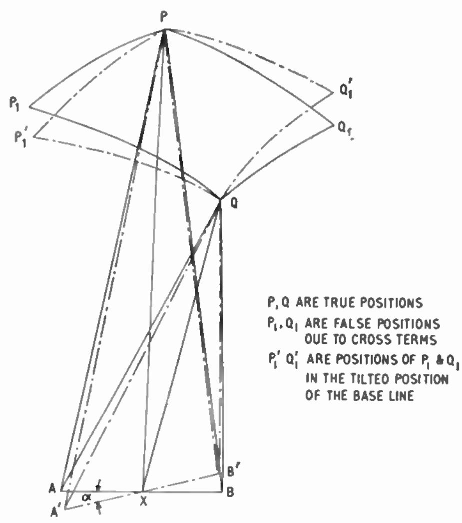


Fig. 7. Elimination of position ambiguities by tilting the base line.

are processed and indicate the presence of four possible positions of the targets P, Q, P_1, and Q_1. The base line of the aerial system is then tilted from AXB to A′XB′ through an angle α. If the transmissions are then repeated and the four possible positions of the targets are again computed taking account of the move of the base line this will result in the 4 positions being P, Q, P_1' and Q_1' as shown in Fig. 7. It can therefore be concluded that P and Q are the correct positions since they did not change with the angle of the base line.

However, although it is theoretically possible to resolve the ambiguities in this way no realistic form of processing to achieve this result is known to the authors, save the laborious process of storing the whole information tilting the base line and comparing the results. Moreover, it is clearly necessary to move the base line to such an extent that each cross-product term moves at least one unit of range resolution. Also since these changes are obtained by taking the difference between two successive range measurements the range accuracy is very important and it is not always reasonable to make the approximations associated with $R \gg D$. A serious disadvantage in this method is the need to move the aerial system, particularly since in order to obtain good bearing accuracy the base line might well be of considerable length and this therefore destroys one of the principal attractions of the system, a stationary aerial. One method of overcoming the disadvantage is to adopt the aerial system of Fig. 8. In this case the base line has been tilted about receiving aerial A by switching channel B alternately between two receiving aerials at B and B′. For convenience the transmitter may also be situated at A.

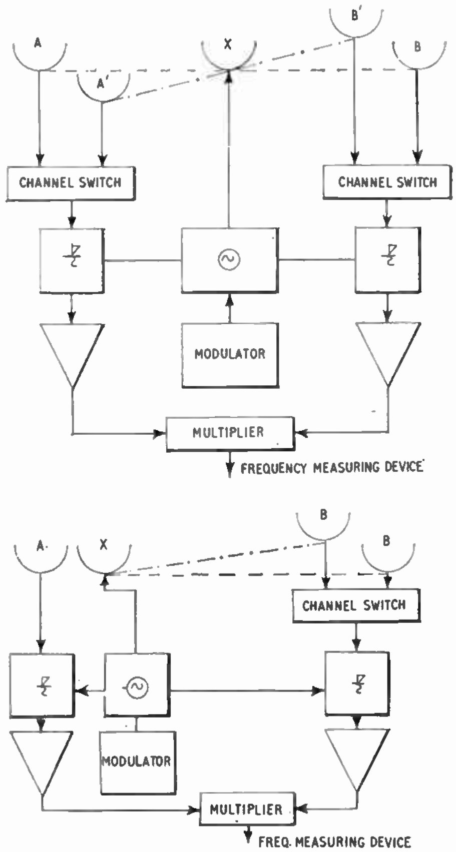


Fig. 8. Base-line tilt arrangements for resolving bearing ambiguities.

3. Experimental Studies

To verify some of the previous conclusions and to study certain practical problems associated with the proposed twin-channel echo-location system, two experimental equipments were constructed and tested. The first was a form of the twin-channel radar operating on an experimental basis for short-range working and operating in the 3-cm microwave band. The second equipment consisted of a simulated ranging system using a tapped delay line to simulate targets.

3.1. *Microwave Experimental Equipment*

The schematic form of the microwave experimental twin-channel radar is shown in Fig. 9. The three aerials are parabolic reflectors with a beamwidth of about 8 deg in both planes. The radar site used for measurements was crowded with buildings and the narrow beamwidth was used to restrict interest to a single target or group of targets. The following represents a brief specification of the equipment:

carrier frequency	9375 Mc/s
total frequency sweep	$\Delta f = 20$ Mc/s
repetition rate of sweep	$f_s = 1$ kc/s
transmitted power	$P_t = 100$ mW

Separation between receiving aerials D is variable up to 25 metres.

The transmitted signal is frequency modulated in a repetitive saw-tooth manner, and a portion of this signal is fed to the crystal mixers of the receivers and the amplified and filtered outputs of these mixers represent the range spectra of the two channels. These spectra are analysed using conventional laboratory spectrum analysers. The two range signals are also fed to the inputs of a multiplier employing a ring diode circuit. The difference frequencies corresponding to the bearing spectrum at the multiplier output are also analysed with a spectrum analyser.

One of the problems of saw-tooth frequency modulation is the high degree of frequency linearity required. In this application departures from linearity

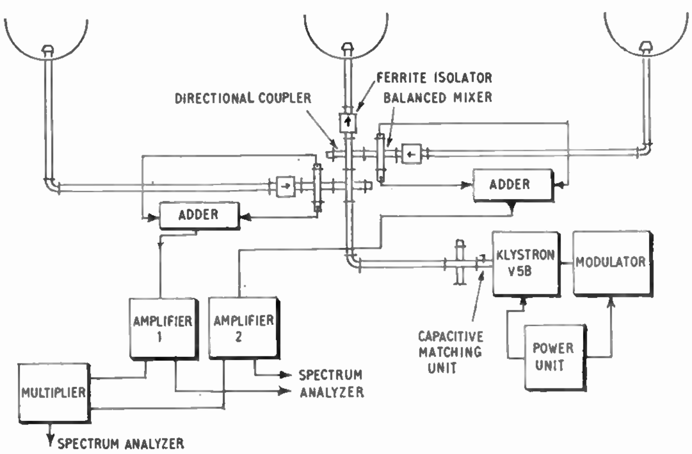
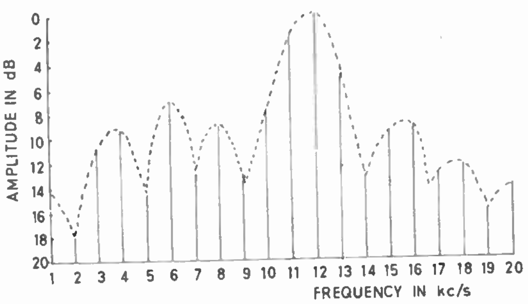
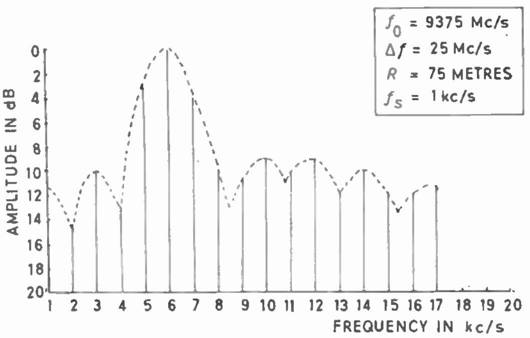


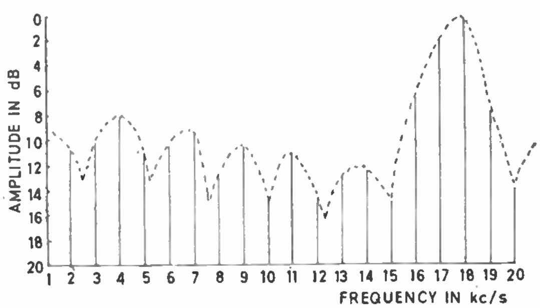
Fig. 9. Experimental set-up.



(*a*) Range signal spectrum (channel 1).



(*b*) Range signal spectrum (channel 2).



(*c*) Bearing signal spectrum.

Fig. 10. Single target operation with balanced mixers and susceptance matching.

cause severe distortion of the range spectra as does any amplitude modulation of the transmitted signal. This restricted the total frequency swing of the klystron well below the 3-dB mode bandwidth. It was also found that an improved linearity and reduced amplitude modulation resulted from suitable susceptance matching of the valve. For short-range working balanced mixers also helped to balance out unwanted products and provide improved spectra.

Another problem with the double-channel system is that small amounts of transmitted signal coupled into the receivers also find their way to the receiving aerials and due to imperfect matching some signal is reflected back to the receivers. Since it may be necessary to make D of the same order of magnitude as a short range target, multiple reflections of this form can simulate targets at short ranges in the receivers. The effect can be reduced to negligible proportions by the use of isolators in the transmitter and receiver aerial feeds as shown.

Two typical range spectra and the corresponding bearing spectrum for reflections from a single target are shown in Fig. 10. This corresponds to a building at a range of about 75 metres from the aerial system. It can be seen that the spectral envelopes resemble the theoretical $(\sin x)/x$ form though the height of the subsidiary lobes is higher than the theoretical pattern. Apart from the reasons previously given for spectrum distortion a further contributing factor is the large number of buildings and other unwanted targets on the site. Thus despite the use of narrow-beam aerials these unwanted targets contribute to the spectra via the side-lobes of the aerials. Since it was not possible to produce a carefully controlled target situation on this site, to study the effect of multiple targets it was found necessary to resort to a simulated system for this part of the study.

3.2. *Simulation System*

Since the measurement of range is entirely effected by the measurement of time delay it is simple to simulate ranging systems by the use of delay lines. Figure 11 shows a schematic diagram of the simulation system. A frequency-modulated oscillator is used to simulate the transmitter and a tapped-cable delay line

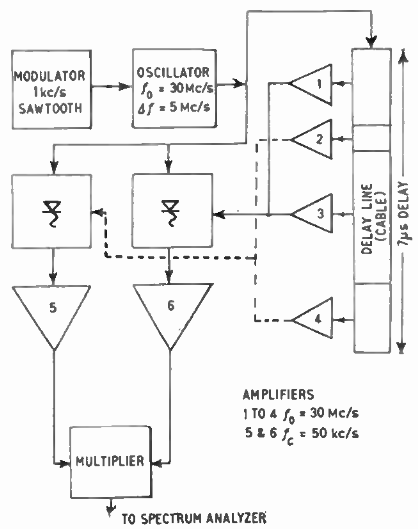
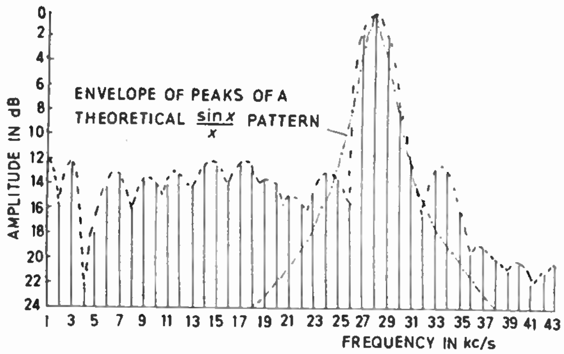
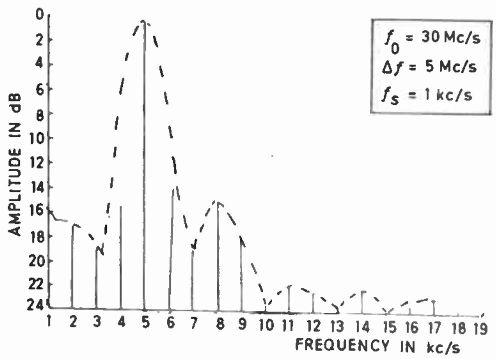


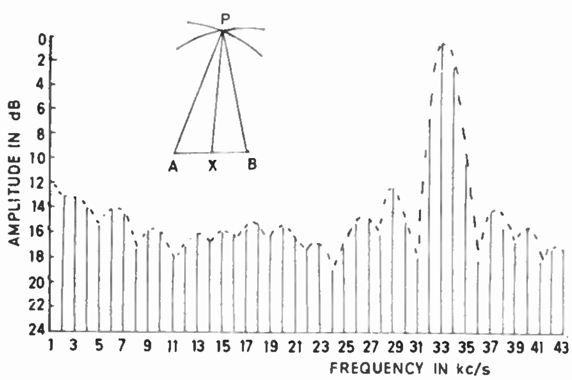
Fig. 11. Block schematic of twin-channel frequency-modulated radar simulation.



(*a*) Range signal spectrum (channel 1).



(*b*) Range signal spectrum (channel 2).



(*c*) Bearing signal spectrum.

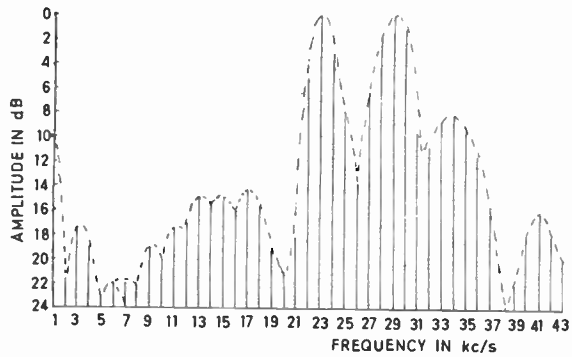
Fig. 12. Single target operation (simulation).

provides the appropriate delays. The parameters of the system are given below:

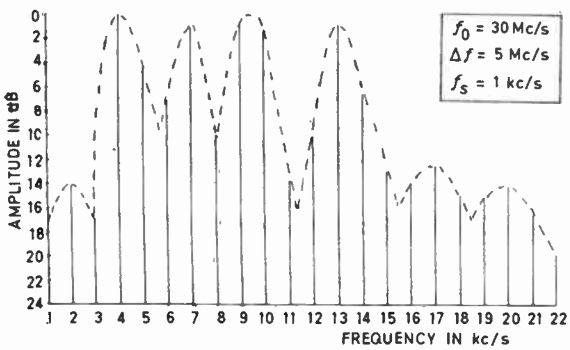
Centre frequency of oscillator	$f_c = 30$ Mc/s
Frequency sweep	$\Delta f = 5$ Mc/s
Repetition rate	$f_s = 1$ kc/s
Maximum delay line delay	$\tau_{max} = 7$ µs

The single-target and two-target performances of the system is demonstrated in the results given in Figs. 12, 13 and 14. Figure 12 shows the measured single-target performance using only two tapping points on the delay line to represent the two ranges from the two receivers. The two range frequencies can be seen to be 28 kc/s and 33 kc/s respectively, and the centre of the bearing spectrum is 5 kc/s representing the difference frequency.

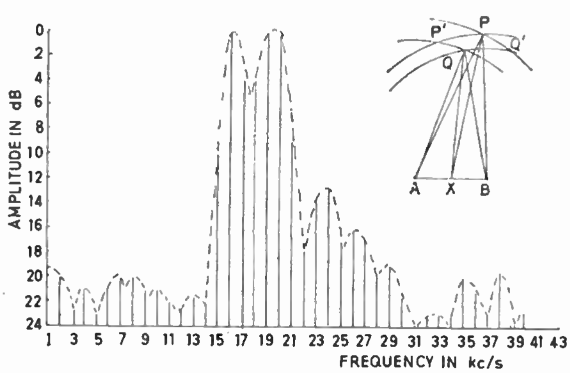
Figure 13 shows the general case of two targets at different ranges relative to both receivers. Therefore



(*a*) Range signal spectrum (channel 1).

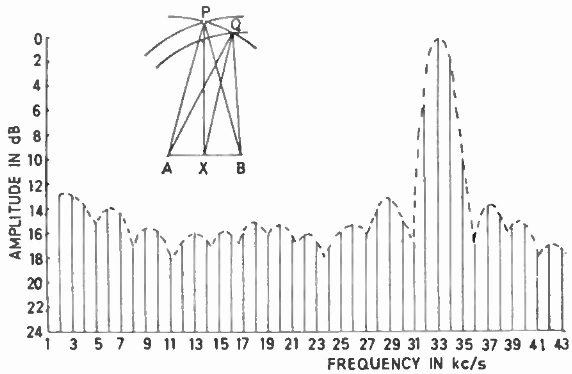


(*b*) Range signal spectrum (channel 2).

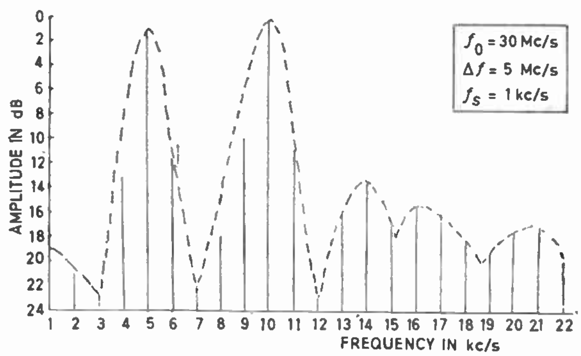


(*c*) Bearing signal spectrum.

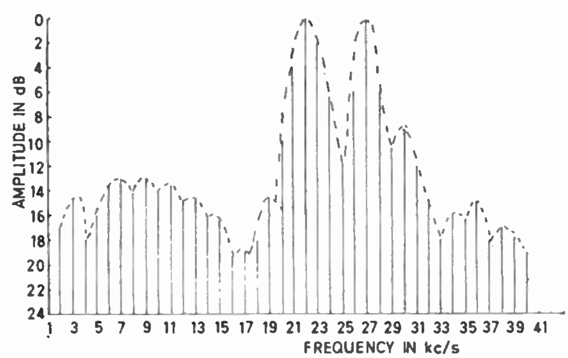
Fig. 13. Simulated two target operation.



(*a*) Range signal spectrum (channel 1).



(*b*) Range signal spectrum (channel 2).



(*c*) Bearing signal spectrum.

Fig. 14. Simulated two target operation (two targets at same range from (*a*)).

both range spectra have two principal maxima and consequently the bearing spectrum contains four principal maxima which in the case of an echo-location system would correspond to two correct target positions and two incorrect 'image' positions due to cross-product bearing frequencies. If the two targets are at different ranges relative to one receiver but at the same range from the other; the resultant bearing spectrum contains two principal maxima as shown in Fig. 14.

4. Discussion of Performance of Twin-channel Radar

The previous sections have described some aspects of a study of a particular form of echo-location system put forward in reference 2. It is therefore reasonable to comment briefly on the relevance of the results of this study to:

(*a*) a radar or sonar system based on these principles.

(*b*) the echo-location ability of certain species of bats.

(*c*) the possible application of the system to a proposed aid for the blind.[3]

4.1. *The Application to Radar or Sonar*

It has been established for several years that the directional properties of a receiving array or aerial system may be improved by the use of wide-bandwidth signals[6] and the twin-channel system under consideration is one example of how this may be accomplished. There are several applications in both radar and sonar which call for a high degree of directional accuracy in the location and tracking of targets; also in many such instances the physical problems associated with mechanical movement of the directional arrays are formidable and this often leads to the study of systems which can determine the range, bearing and sometimes the velocity of targets from a stationary directional array. Several such systems have been built in the field of both radar and sonar, but the principal attraction of the twin-channel system is its simplicity—in particular the need to employ only two receiving aerial elements.

Unfortunately, as can be seen from the material in Sections 1 and 2, the price that was paid for this was loss of multiple-target resolution properties, unless facilities for tilting the base line may be introduced as described in Section 1.5. Even in cases where this might be practical, such as by the use of switching between two aerial systems, the practical problems of sorting out the correct target positions from the false positions is fairly prohibitive.

On the question of signal/noise ratio the twin-channel system is not efficient since there is a loss of aerial gain owing to the need to employ wide-beamwidth aerials covering the entire sector of interest.

Considering all these factors it would appear that the application of the twin-channel system to any situation involving a complex multiple-target environment is rather limited. There are also very few applications involving one, or very few targets; one example of the latter is satellite tracking, but this situation is rapidly changing. If the aerials concerned are all narrow-beamwidth tracking aerials, then the frequency-modulated returns may be used to obtain accurate bearing information within this beamwidth. In this case the multiple-target ambiguities are restricted to the region of the narrow primary beamwidths of these aerials.

4.2. *The Echo-location Ability of Bats*

The results of this study do not appear entirely consistent with the known remarkable ability of certain species of bats to avoid fixed obstacles and to catch their prey, as discussed in Section 1.1. Their ability seems to suggest a high degree of both range and bearing resolution in the presence of a complex target environment. They also seem to possess the ability to function against very high background-noise levels. The results of the study seem to indicate that the bat could only sort out a multiple-target situation by some complicated process using the relative strength of the returns and the relative Doppler shifts of the returns or by utilizing the change in the returned signals when its head was inclined at a different angle (or a different flight heading). It should not be impossible to devise experiments to try to test these effects. It has been observed that certain species of bats move their ears in conjunction with the sonic transmissions; this appears to be similar to the base line tilt system for removing ambiguities.

It is, of course, possible that the authors have overlooked some property of the twin-channel system which explains how multiple targets may be resolved. However, if we assume that this is so the resolution in bearing would presumably be proportional to the frequency sweep. In such an event if the frequency sweep were to be doubled, both the range and bearing information rate would be increased and from an information theory standpoint this seems rather unlikely.

An alternative explanation is that Kay's proposed model for the echo-location system of the bat is invalid. However, all forms of ranging system can be considered as different ways of performing the same correlation process. Therefore the results of Section 1 would apply to any system using one transmitter and two receivers with a band-limited transmitted signal.

The authors feel that any discrepancies probably lie in the assumption that it is reasonable to describe biological phenomena in terms of fairly simple electrical engineering concepts involving linear systems with time-stationary components. An animal is able to make extensive use of *a priori* information and moreover represents an adaptive and time-variant system. Nevertheless, simple proposals for explaining such complicated phenomena may be of considerable scientific and engineering interest provided that the simple electrical analogues are not expected to possess exactly the same properties as the biological process.

4.3. *The Ultrasonic Blind Aid*

For the same reason that the analysis did not appear to help in assessing the echo-location performance of the bat it would be unreasonable to condemn the possible use of the aid for blind guidance based upon the principles of the twin-channel system. This sort of problem illustrates the differences between the conventional analysis of a system with constant parameters and the operational measurement of performance of the same sort of system but with facilities for adaptive processes and self-optimization.

5. The Use of Twin-channel Receivers for Tracking an Active Source

A different possible application of the proposed system is the tracking of the trajectory of a co-operating target, such as a satellite or rocket, either to obtain telemetered information or to make accurate tracking observations.

For this purpose the target is equipped with an active source to assist tracking and this immediately reduces the situation to a single target problem. The most conventional tracking system uses a monopulse aerial which produces bearing error signals from a split-beam directional pattern. These error

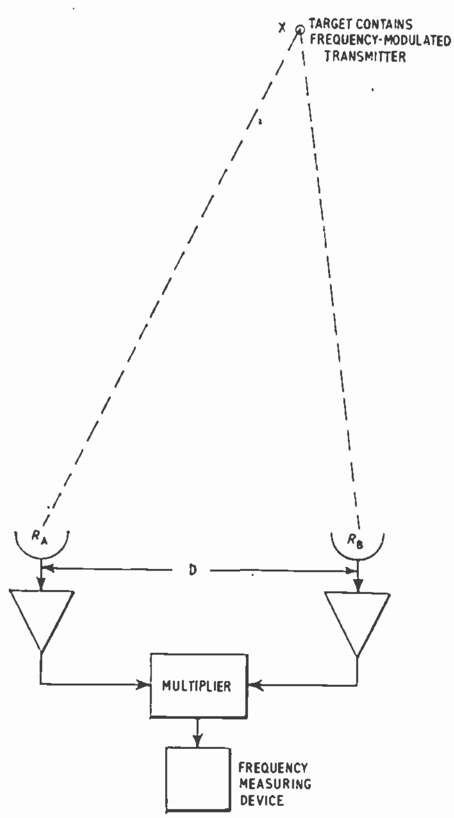


Fig. 15. Twin-channel interferometer.

signals may be fed into a servo system which orientates the aerial until it is facing the target. For such a system the bearing accuracy is directly proportional to the size of the aerial in wavelength. Clearly such a system provides directional and not range information, in the absence of transmitted signals from the ground.

For very high accuracy tracking two aerials, separated by a long base-line may be used, the outputs of the two aerials being combined to form an interferometer. Radio astronomy interferometers utilize wide bandwidth inputs to the correlator to improve the directional characteristic, although this is not a single-target application.

The twin-channel frequency-modulated radar may be converted into a tracking system by using the frequency-modulated signals as the active source on the target to be tracked. The two separated receiving aerials are connected to a multiplier as shown in Fig. 15. The difference between the instantaneous frequencies arriving at the two aerials will be proportional to sin θ where θ is the bearing of the target relative to the base line.

$$f_b = \frac{D}{C}\frac{\Delta f}{T_s}\sin\theta \qquad \text{......(13)}$$

This is a method of measuring the range difference of the target from the two receivers, that is measuring $D \sin\theta$.

Let

$$\tau_1 = \frac{D}{C}\sin\theta = \frac{T_s f_b}{\Delta f} \qquad \text{......(14)}$$

Thus the accuracy of bearing measurements is the accuracy of measuring τ_1 and since T_s and Δf are fixed, this involves measuring f_b which will, of course, take the form of a spectrum centred on this frequency. Thus for very accurate tracking $\Delta f/T_s$ would have to be maintained to a very high degree of accuracy.

However, since this is a single-target application an important simplification is possible. If we consider a simple single-channel frequency-modulated ranging system the range frequency (as shown in Fig. 2) is always at one of two values, either f_{RA} or $(\Delta f - f_{RA})$. It can be shown that the total number of cycles N, in one modulation period of this waveform is a direct indication of range and is given by:[7]

$$N = \frac{2}{\pi}\Delta f T_s \sin\left(\frac{\pi}{T_s}\right)\tau \qquad \text{......(15)}$$

so that for single-target applications it is only necessary to measure the average frequency of this waveform (including the high and low frequency portions) using cycle counting techniques. A further important property is that the form of the frequency modulation does not affect the result which is dependent only upon the total frequency sweep. This removes the restrictions of frequency linearity of the frequency modulation and sinusoidal modulation may be used resulting in a sinusoidal frequency variation about zero frequency for the output signal from the multiplier. In this case f_b refers to the average number of cycles of the waveform per modulation period and

$$f_b = \frac{N}{T_s} = \frac{2}{\pi}\Delta f \sin\left(\frac{\pi\tau_1}{T_s}\right) \qquad \text{......(16)}$$

and if the modulation rate is chosen such that $\tau \ll T_s$

$$f_b = \frac{2\Delta f}{T_s}\tau_1 \qquad \text{......(17)}$$

In the case of the determination of f_b by spectrum analysis, the accuracy was determined by the interpolation of the centre of a spectrum from the line components. In the case of cycle counting the accuracy is limited by the need to measure integral numbers of cycles. Therefore the change in τ_1 due to one cycle change in N is given by:

$$\delta\tau_1 = \frac{1}{2\Delta f} \qquad \text{......(18)}$$

so the minimum detectable change of bearing is given by:

$$\delta\theta = \frac{Cd_1}{D\cos\theta} = \frac{C}{2D\Delta f\cos\theta} \qquad \text{......(19)}$$

Putting some reasonable numbers into this equation we may take Δf = 3 Mc/s then for a base line of D = 100 km, the tracking accuracy would be about 1·5 minutes of arc. However, since most co-operating targets move in a reasonably predictable manner it is possible to smooth the output and thus obtain an accuracy greater than that corresponding to one cycle change of f_b. This is the case for frequency-modulated radio altimeters,[8] where the accuracy due to smoothing is improved threefold.

It may also appear that the accuracy is low compared with long base-line interferometers; this is because such interferometers give multiple ambiguities, but the frequency modulated system provides no ambiguities (provided that $\tau < T_s$). The ambiguity of sign of θ may be removed by a delay line in the same way as described in Section 1.3. The proposed technique may therefore be of value as a tracking system particularly at low frequencies where directional tracking aerials are more expensive.

6. Conclusions

This paper has discussed the properties of a proposed twin-channel echo-location system using frequency-modulated transmission. It has been shown that this system is capable of measuring range and bearing of a target from an aerial system consisting of one transmitting and two receiving aerials. However, in the presence of multiple targets it is shown that position ambiguities arise such that n targets correspond to n^2 possible positions. The ambiguities may be resolved or partially resolved by making successive measurements after moving the base line. Comments are made on the possible relevance of this analysis to the echo-location ability of bats.

Finally it is shown that the same approach may be applied to a tracking system for tracking active frequency-modulated signal sources. Since this is a single-target application there are no problems from unwanted multiple-target position ambiguities.

7. Acknowledgments

The authors wish to thank Professor D. G. Tucker for discussions on this work and helpful criticism of this paper.

8. References

1. D. R. Griffin, "Listening in the Dark". (Yale University Press, New Haven, Connecticut, 1958).
2. L. Kay, "A plausible explanation of the bats echo-location acuity", *Animal Behaviour*, **10**, No. 1–2, January/April 1962.
3. L. Kay, "Auditory perception and its relation to ultrasonic blind guidance aids", *J. Brit.I.R.E.*, **24**, No. 4, p. 309, October 1962.
   See also the article "Ultrasonic Aid for Blind Persons", *The Radio and Electronic Engineer*, **25**, No. 2, p. 166, February 1963.
4. A. J. Hymans and J. Lait, "Analysis of a f.m. continuous-wave ranging system", *Proc. Instn Elect. Engrs*, **107**, Pt. B, p. 365, July 1960, (Paper No. 3264E).
5. B. S. Rao, "Twin-channel Frequency-modulated Radar for the Measurement of Range and Bearing", Ph.D. Thesis, University of Birmingham, 1964.
6. H. Nodtvedt, "The correlation function in the analysis of directive wave propagation", *Phil. Mag.*, **42**, p. 1022, 1951.
7. M. Capelli, "Radio altimeter", *Trans. Inst. Radio Engrs*, (*Aerospace and Navigational Electronics*), **ANE-1**, No. 2, p. 3, June 1954.
8. J. S. Shayler, "Radio guidance elements of the B.L.E.U. automatic landing system for aircraft", *J. Brit.I.R.E.*, **21**, No. 1, p. 17, January 1961.

Manuscript received by the Institution on 4th June 1964. (Paper No. 926/RNA31.)

U.D.C. 621.396.677.3

Planar Arrays with Unequally Spaced Elements

By

M. I. SKOLNIK, Dr. Eng.†

AND

J. W. SHERMAN, M.S.‡

Reprinted from the Proceedings of the Symposium on "Signal Processing in Radar and Sonar Directional Systems", held in Birmingham on 6th–9th July, 1964.

Summary: The application of the optimization technique known as dynamic programming to the design of thinned planar arrays with unequally spaced elements is described. Two different approaches were investigated. In one method the planar aperture was considered to consist of a number of concentric ring arrays and dynamic programming was applied to determining the radii of the rings which yielded radiation patterns with the minimum peak side-lobe. In the other method the planar array was divided into an even number of equal sectors with elements identically located in each sector. Both methods were selected to avoid excessive calculation times. With the aid of large digital computers designs were obtained and radiation patterns computed for 40λ diameter circular planar arrays and for various degrees of thinning. The results differed from those obtained by other unequally spaced array design methods in that the side-lobes of the radiation pattern tended to decrease with increasing distance from the main beam.

1. Introduction

Unequally spaced planar arrays are of interest in applications where it is desired (1) to avoid an amplitude taper, or (2) to operate with significantly fewer elements than required for a conventional array with half-wavelength spacings between elements. The former is called a *density-tapered* array; the latter, a *thinned* array. The conventional array is termed a *filled* array.

The degree to which an unequally spaced array is thinned is generally expressed as the percentage of elements removed from a filled array occupying the same physical area. That is, if the filled array contains 10 000 elements and if the thinned array consists of 1000 elements the degree of thinning is 90%. Since the gain of an array, neglecting mutual coupling, is proportional to the number of elements, a thinned array suffers a proportionate loss in gain and a corresponding rise in the average side-lobe level, although the half-power width of the main beam remains relatively unchanged. The object in design is to select the element spacings such that objectionably high peak side-lobes do not appear. The thinned array has application when a narrow beam is desired with a given number of elements. It is useful as an economical expedient for demonstrating the operation of large arrays when only a limited number of elements can be provided.

The density-tapered array designed to yield radiation patterns comparable to those obtained with amplitude tapering are thinned on the order of 40% to 60%. Thus, the gain is not significantly reduced compared to the conventional amplitude tapered array. The side-lobe level, especially the peak side-lobe, can be maintained to a reasonably low value. Density-tapered arrays are of practical advantage in applications where it is convenient that each element be identical. This is of special importance in large transmitting arrays where a power amplifier is used at each element.

Several methods have been described in the literature for designing unequally spaced planar arrays. These include the placement of elements on the intersection of rings and radials,[1] random removal,[2] deterministic density taper,[3,4] and statistical density taper.[5] In general, the design methods for a density-tapered array differ from those of a severely thinned array. This paper is concerned primarily with the application of *dynamic programming* to unequally spaced planar arrays. Dynamic programming is an optimization procedure widely used in operations research.[6–8] It permits an N-dimensional optimization problem to be treated as N one-dimensional optimization problems and, when it can be applied, considerably simplifies the necessary calculations as

† Formerly with Electronic Communications, Inc., Maryland; now with the Institute for Defense Analyses, Washington, D.C., U.S.A.

‡ Research Division, Electronic Communications, Inc., (now known as Adtec), Timonium, Maryland, U.S.A.

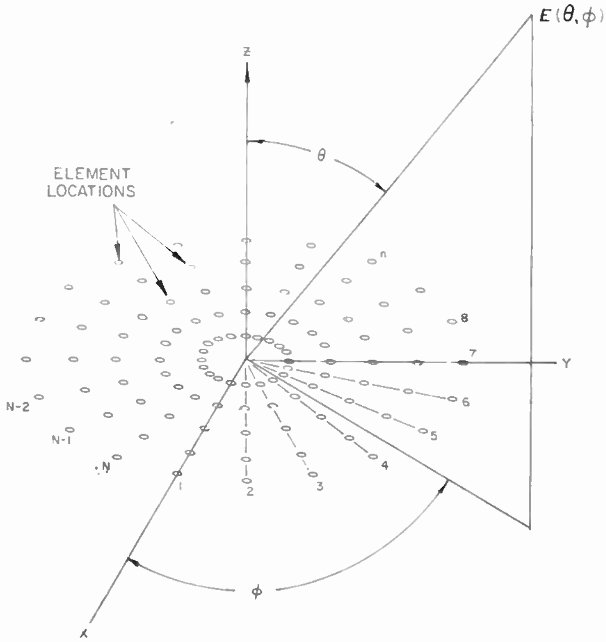


Fig. 1. Configuration of an array of elements placed at the intersection of rings and radials. Array co-ordinate system in terms of the angles θ, ϕ is also shown.

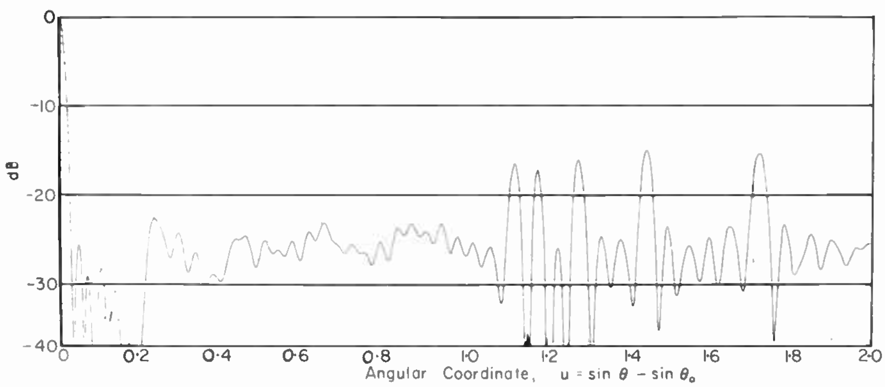
compared with the total enumeration of all possible solutions.

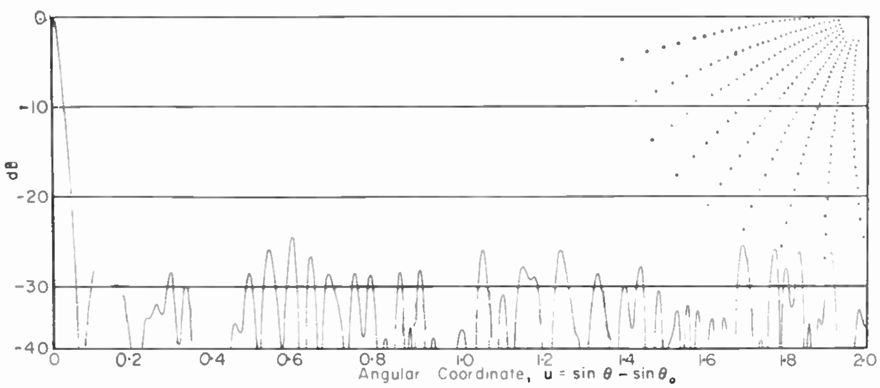
Before proceeding to the subject of array design using dynamic programming the several reported methods for specifying the element locations of unequally spaced planar arrays are briefly described in the next section. In this paper, it is always assumed that the elements are all of equal weight, i.e. there is no amplitude taper of the aperture illumination.

2. Methods of Design

The design of thinned arrays by placing elements at the intersection of rings and radials (Fig. 1) is an empirical procedure evaluated mainly by trial-and-error.[1] An example of the array factor for a 90% thinned array with one degree beamwidth is shown in Fig. 2(a). The 1008 elements within the 56-wavelength diameter aperture were located on 28 equally spaced rings at the intersection of radials equally spaced at 10 deg intervals. The abscissa is $u = \sin\theta - \sin\theta_0$ where θ is the angle measured with respect to the array normal and θ_0 is the angle to which the beam is steered in the θ direction. Within the visible region ($u < 1$) the maximum side-lobe level is $-22{\cdot}6$ dB. If the beam is steered ($u > 1$) the peak lobe is $-14{\cdot}9$ dB at $u = 1{\cdot}45$. The best

(a) Elements located at the intersection of equally spaced rings and equally spaced radials in a 56λ diameter aperture. Ring separation = 28/31λ. radial separation = 10 deg, 90% thinning, pattern computed in the plane $\phi = 0$ deg.





(b) Elements displaced to form spiral arms repeated every 10 deg as shown in upper right-hand corner. Ring locations taper-matched to a 35 dB Taylor distribution with $\lambda/2$ minimum ring separation.

Fig. 2. Pattern of an array of 1008 elements.

radiation patterns with this technique occurred for rings whose spacings were selected such that the density of rings was proportional to the amplitude of the aperture illumination of a conventionally designed array and where the elements on the rings were displaced so as to form spirals rather than radials. An example of one of the better patterns obtained in this manner appears in Fig. 2(b). The element locations are sketched approximately in the upper right hand corner. The peak lobe is $-24\cdot4$ dB and occurs at $u = 0\cdot58$. A comparison of Figs. 2(a) and 2(b) illustrates that if a low peak side-lobe is to be obtained the side-lobes should be of more or less uniform magnitude. Low side-lobes in some portions of the u-region and high side-lobes in other regions indicate that it should be possible to reduce the high lobes at the expense of raising the low lobes. The side-lobes may also be reduced if the null width of the main beam is increased. The null width of the low side-lobe design of Fig. 2(b) is almost double that of the design of Fig. 2(a), although half power widths are not too significantly different. In any of the design techniques thus far reported there is little, if any, control of the main beam null width. Since the aperture distributions repeat every ten degrees the patterns of Fig. 2 apply for $\theta = 0°$, $10°$, $20°$, etc. Slightly higher lobes might sometimes appear in the patterns taken through the intermediate planes.

A conventional planar array filled with elements equally spaced on a half-wavelength grid can be readily converted to an unequally spaced array by removing elements at random; that is, independent of the location within the aperture. (Although elements could be added at random to an empty aperture, the analysis is more convenient with a model which removes elements from a filled array.) Maher and Cheng[2] analysed this technique for a linear array. It would seem that similar effects occur for a planar array. Since the actual arrangement of elements in the thinned array depends on a random process, the exact radiation pattern of any particular design is unpredictable and must be described in statistical terms. The (ensemble) average pattern can be shown to consist of the pattern of the original filled array plus a statistically 'omnidirectional' pattern. The statistical portion of the pattern causes the peak side-lobe to be greater than the already high peak side-lobe of a conventional array with uniform amplitude distribution. Therefore, the relatively high near-in peak side-lobes mean that the random removal of elements is not generally a satisfactory array thinning technique, even though it is a simple design procedure to implement in practice.

Although the random removal of elements independent of position within the aperture does not always produce satisfactory results, the non-uniform removal of elements can be made to produce reasonably good patterns. This technique involves making the density of unequally spaced elements within the aperture proportional to the amplitude of the aperture illumination that would be used with a conventional equally spaced array. It has proved to be a satisfactory design technique. Its theoretical justification is that the density taper is an approximation to the amplitude taper. The degree of approximation is limited by the number of elements and the practical requirement that no two elements can be spaced closer than some minimum distance, usually one half-wavelength.

There are two methods for determining the locations of elements in a density-tapered planar array. In one method the element locations are juggled by a trial-and-error placement until a satisfactory density taper is found.[3,4] (In a linear array the design procedure can be made more definite by using the cumulative distribution of the aperture illumination.[9]) This may be termed the *deterministic* approach as contrasted to the other method which is a *statistical* approach.[5] In the statistical method the probability that an element is placed at a particular location is made proportional to the amplitude that the conventional antenna aperture illumination would have at

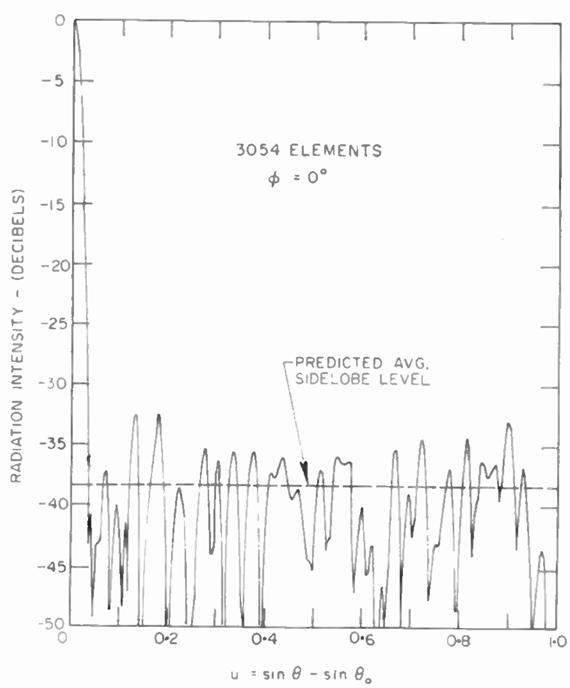


Fig. 3. Computed radiation pattern of a statistically designed density-tapered array, 50λ diameter, using the 40 dB Taylor circular aperture distribution as the model for A_n; 62% thinning, 3054 elements.

the same location. An example of a computed pattern of a statistically designed density tapered array is shown in Fig. 3. The statistical (ensemble) average power pattern is

$$\overline{|E(\theta,\phi)|^2} = k^2|E_0(\theta,\phi)|^2 + \sum_{n=1}^{N} kA_n(1-kA_n) \quad \ldots\ldots(1)$$

where $|E_0(\theta, \phi)|^2$ is the power pattern of the N element equally spaced array designed by conventional means with aperture illumination function given by A_n, and k is the probability an element will appear at the point where the aperture illumination is a maximum, usually the centre of the array. The second term is independent of angle and represents a statistically 'omnidirectional' pattern. This term generally dominates except in the region near the main beam. It is interesting to note that the average pattern expression of eqn. (1), even though it is based on statistical considerations, is one of the few analytical expressions available for predicting the general pattern behaviour of an unequally spaced array, other than determining the actual spacings and computing the patterns.

A conceivable procedure for designing an unequally spaced array is to try all possible combinations of element locations and select that configuration which results in the 'best' radiation pattern. If there are m available element locations for n elements, there are $m!/n!(m-n)!$ unrepeated combinations of element-spacing configurations, assuming all the elements are identical. This can be a large number. For example, there are approximately 10^{14} different combinations to choose from if 25 elements were to be placed within 50 possible element locations. Even with modern high-speed computers the design of an array by the total enumeration of all possible combinations is a prohibitive procedure in all but the simplest of cases.

Under certain conditions, it is possible to perform the equivalent of total enumeration with a considerable reduction in the required computations by use of the procedure known as *dynamic programming*. Dynamic programming will yield the same solution as total enumeration when it can be proved that the *principle of optimality* applies to the problem. Although we have not been able to prove the principle of optimality for the design of unequally spaced arrays, dynamic programming was applied, nevertheless, to this problem and obtained interesting and useful results. Fortunately there are similar precedents in antenna analysis such as the scalar Kirchhoff theory, where useful results are obtained in spite of theoretical objections.

Dynamic programming has been applied successfully to the design of linear arrays of unequally spaced elements and has achieved results equal to or superior to arrays designed by other methods.[10–12] The remainder of this paper describes the application of dynamic programming to circular planar arrays.

3. Dynamic Programming Applied to Planar Arrays

In the computer programs written for the linear array work reported previously[10,11] arrays of 50 unequally spaced elements could be handled comfortably on large machines such as the IBM 7094. One hundred elements could be handled by the IBM *Stretch* computer. Planar arrays generally contain considerably more elements: 500 to 4000 unequally spaced elements might represent the typical range of interest. Even with large capacity, high-speed computers and the vast savings in computational effort made possible by dynamic programming, the design of such arrays by this technique results in prohibitively long machine run-times with the presently constituted program when the array is large. To keep the problem within bounds and avoid lengthy calculations, advantage was taken of the symmetry of the circular array. Two different approaches were investigated. In one method the planar aperture was considered to consist of a number of concentric ring arrays and dynamic programming was applied to determine the radii of a given number of rings which minimized the side-lobe radiation over a specified angular region. This technique had the advantage that it could utilize the computer program previously developed for the linear array if certain simple modifications were made. In the other method the planar array was divided into a number of equal sectors with elements identically located in each sector. An element located in one sector automatically placed an element in a similar position in all the other sectors, thus reducing the number of decisions that had to be made. Dynamic programming selected those element locations within each sector which minimized the side-lobe level over the specified angular region. These two procedures were investigated so as to keep the problem within the capabilities of existing computers.

4. Ring Array Approach to Planar Array Design

In this method, the circular planar array consists of a number of concentric ring arrays. On each ring, elements are equally spaced at approximately one half-wavelength. By taking advantage of the almost circular symmetry the planar array radiation pattern can be simplified and expressed in a form similar to that of a linear array.

The radiation pattern of an array of equally radiating isotropic elements located on M concentric rings with $[4\pi r_m]$ elements on each ring may be represented as

$$E(\theta,\phi) = \sum_{m=1}^{M}\sum_{n=1}^{[4\pi r_m]} \exp\left\{jkr_m\left[\sin\theta\cos\left(\frac{2\pi(n-1)}{[4\pi r_m]}-\phi\right)-\sin\theta_0\cos\left(\frac{2\pi(n-1)}{[4\pi r_m]}-\phi_0\right)\right]\right\} \quad \ldots(2)$$

where $k = 2\pi/\lambda$, λ = wavelength, r_m = radius of the mth ring, (θ, ϕ) = direction in which the far field is being determined, (θ_0, ϕ_0) = direction to which the main beam is steered, and the brackets [ ] about $4\pi r_m$ denote the largest integer less than $4\pi r_m$. This expression can be simplified if the radiation pattern is examined only in the plane in which the beam is steered; that is $\phi = \phi_0$. With this assumption and writing $u = \sin\theta - \sin\theta_0$, eqn. (2) becomes

$$E(u,\phi_0) = \sum_{m=1}^{M}\sum_{n=1}^{[4\pi r_m]} \exp\left\{jkur_m\cos\left[\frac{2\pi(n-1)}{[4\pi r_m]}-\phi_0\right]\right\} \quad \ldots\ldots(3)$$

If the ring of radius r_m contains $[4\pi r_m]$ elements, its radiation pattern can be approximated by the zero order Bessel function $J_0(kur_m)$ with a weighting proportional to the number of elements within the ring. Thus the radiation pattern of the array of rings is independent of ϕ and can be written

$$E(u) \simeq \sum_{m=1}^{M}[4\pi r_m]J_0(kur_m) \quad \ldots\ldots(4)$$

This is similar to the radiation pattern of a linear array of symmetrically located elements in which

$$E(u) = 1+2\sum_{n=1}^{N}\cos 2\pi d_n u \quad \ldots\ldots(5)$$

where d_n is the distance, measured in wavelengths, of the nth element with respect to the centre of the array. To use the computer program developed previously for the linear array[10-12] the constant term is dropped, the cosine function is changed to the Bessel function, and an amplitude factor is included to account for the variation in the number of elements per ring.

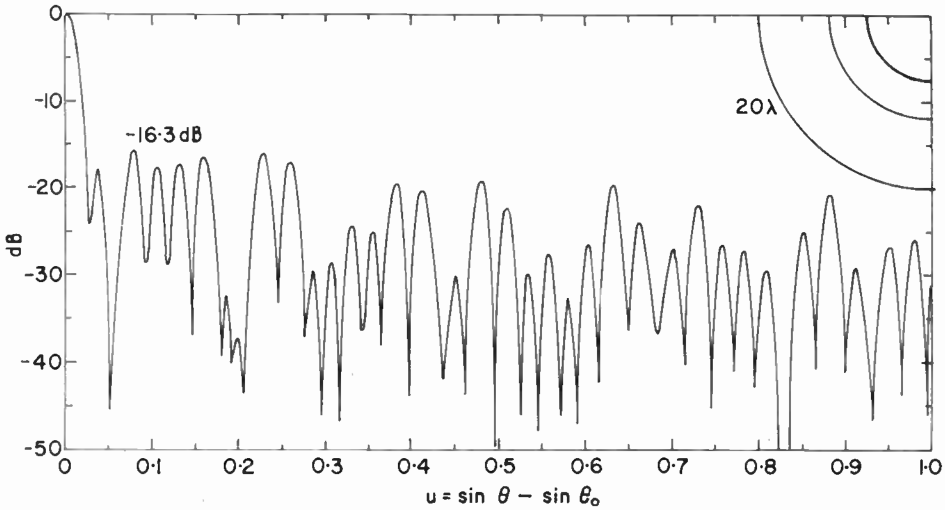
In this formulation of the planar array design problem, dynamic programming is used to find the optimum radii of a specified number of rings to be placed within a planar aperture of fixed radius, in this case $r_{max} = 20\lambda$ (λ = wavelength). There is no control over the number of elements. This is determined only after the optimum ring radii are found.

Although the ring array is a convenient model to analyse, it is generally not as easy to provide steering commands as with an array of elements equally spaced on a rectangular grid. It has been found that relocating the elements of a ring array to the nearest position on a half-wavelength grid affects the pattern but little.[1] Thus the designs achieved by assuming the ring array model may be readily applied to practical arrays.

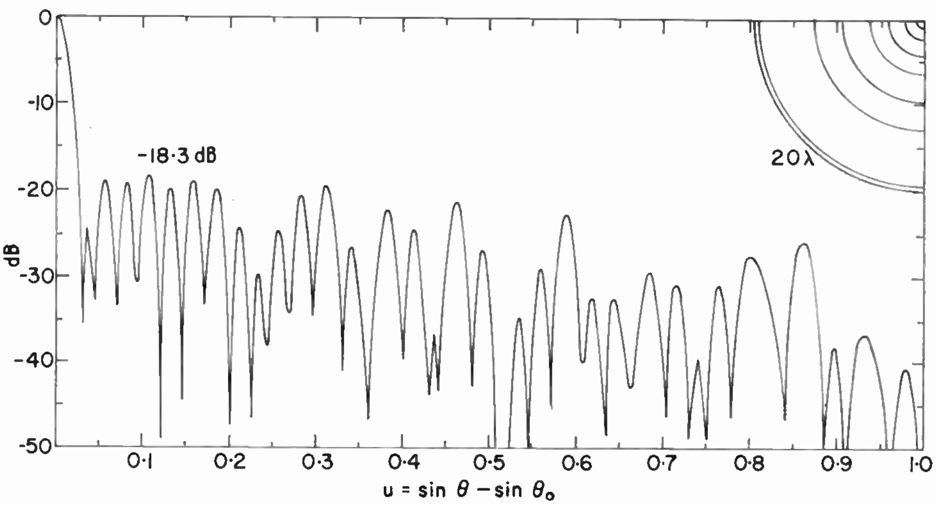
A criterion must be selected with which to judge the 'best' pattern. Lacking specific requirements as to what is considered a desirable pattern a 'minimax' criterion was used in which dynamic programming selected that configuration of element locations which minimized the maximum side-lobe within some specified angular region. Since the pattern expression of eqn. (4) is independent of the angle ϕ, the pattern as a function of the parameter $u = \sin\theta - \sin\theta_0$ only need be considered. The maximum value of u was generally chosen to be unity, corresponding to $\theta = 90°$ when the beam is unsteered. The minimum value of u was chosen by trial-and-error. Too small a u_{min} might place a portion of the main beam within the region of optimization with the result that the maximum of the pattern $E(u)$ is determined by some large value on the main beam and not by the side-lobes. Too large a u_{min} might result in the appearance of an excessively high side-lobe between the main beam and u_{min}.

A uniformly illuminated circular aperture of diameter D has a null width equal to $1{\cdot}22\lambda/D$. With a 40λ diameter aperture this value is equal to 0·03. Thus a first choice for u_{min} would be 0·03. Lower side-lobes than obtained with uniform illumination require a broader null width and therefore, larger values of u_{min} are necessary. In examining the designs of 40λ diameter ring arrays, patterns were obtained for u_{min} = 0·03, 0·04, and 0·05. The outer ring was fixed at $r_{max} = 20\lambda$ and designs were determined as a function of the number of rings M ranging from 3 to 20. The program was constrained so that adjacent rings would not be spaced closer than one half-wavelength. It was found that the value of u_{min} depends on the number of rings. With few rings, u_{min} of 0·03 or 0·04 appears satisfactory, but with many rings u_{min} = 0·05 produced the best results. In this paper the results will be presented for u_{min} = 0·04. Although it was not done here, it would seem worthwhile to include directly in the computer program some means for permitting the computer to select the proper value of u_{min} as it performs the dynamic programming calculation.

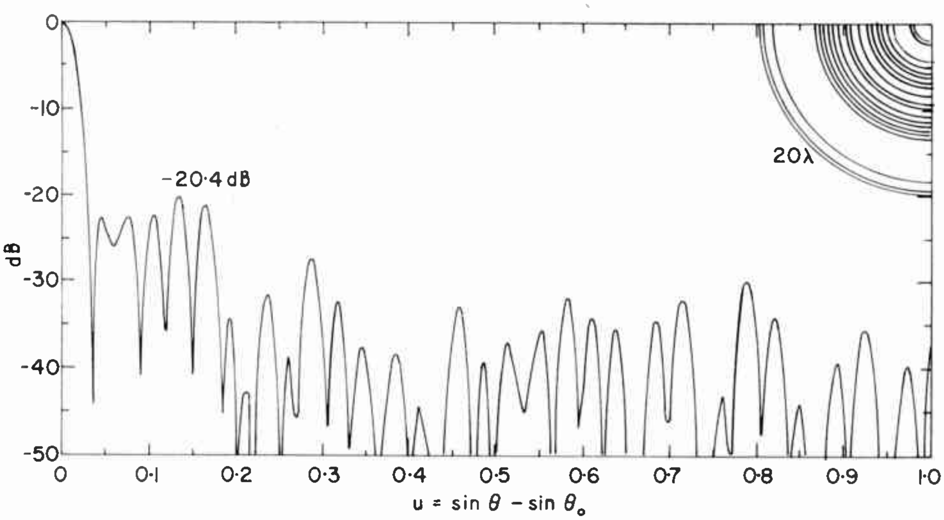
Examples of the type of radiation patterns obtained with ring array antennas are shown in Fig. 4(a), (b), (c), for 3 rings, 8 rings, and 20 rings. At the upper right-hand corner of each plot is the configuration of rings. The side-lobes tend to decrease with increasing u. This is unlike the dynamic programming results achieved with the linear array or the density-tapered planar array in which the side-lobe level was more or less independent of u over the region of optimization. With the linear array, the summation



(a) Array of three rings of radii 7·5, 12·0 and 20·0λ as found by dynamic programming. $u_{min} = 0{\cdot}04$, $u_{max} = 1{\cdot}0$. Number of elements, if spaced at half-wavelength intervals along each ring, equals 495 and corresponds to about 90% thinning. Filled array would contain 5130 elements.



(b) Array of eight rings of radii 1·0, 2·5, 4·5, 7·0, 9·5, 13·0, 19·5 and 20·0λ. Number of elements = 975, 81% thinning.



(c) Array of twenty rings of radii 2·0, 2·5, 4·5, 5·5, 6·0, 6·5, 7·0, 7·5, 8·5, 9·5, 10·0, 11·0, 11·5, 12·0, 12·5, 13·0, 13·5, 18·5, 19·5, and 20·0λ. Number of elements = 2516, 50% thinning.

Fig. 4. Pattern of a circular planar array.

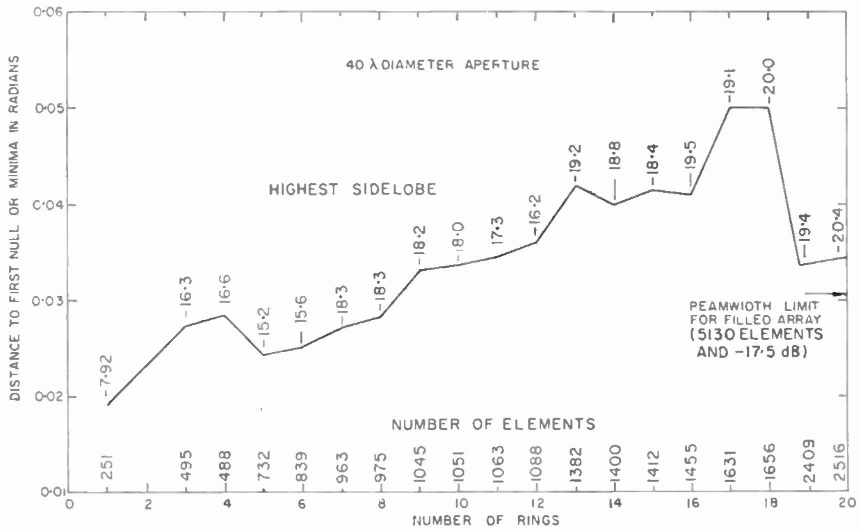


Fig. 5. Summary of ring array designs with $u_{\min} = 0{\cdot}04$, $u_{\max} = 1{\cdot}0$, 40λ diameter circular aperture.

involved (eqn. (5)) consists of cosine functions whose local maxima are independent of the angular co-ordinate u. On the other hand, the summation for the array of rings consists of zero order Bessel functions (eqn. (4)) whose local maxima decrease with increasing u since for large x

$$J_0(x) \rightarrow \frac{2}{\sqrt{\pi x}} \cos\left(x - \frac{\pi}{4}\right).$$

It is not surprising, therefore, that the dynamic programming designs of ring arrays have peak values at the lower values of u.

The results obtained for the designs with $u_{\min} = 0{\cdot}04$ are summarized in Fig. 5. Shown in this figure are the distance to the first null (a measure of the beam width), the peak side-lobe level, and the number of elements contained in the array. The number of elements tends to increase with the number of rings, as would be expected, but there are exceptions where the total elements decrease as a ring is added. Similar data were obtained for $u_{\min} = 0{\cdot}03$ and $0{\cdot}05$. In general, these results indicate that the lower the side-lobe level, the wider the main beam. It is as though there is a certain amount of energy that is to be distributed throughout space and if the side-lobes are to be lowered, the removed energy must be 'swept' under the main lobe.

When compared with planar arrays of equivalent thinning designed by density taper or by rings-and-radials, the peak side-lobes with the ring arrays are generally higher. It is believed that this is not an indication of a failure of the dynamic programming method but is a result of the radiation pattern of an individual ring decreasing with increasing u. Furthermore, the large amplitude factor $[4\pi r_m]$ of the outermost ring has a large effect on the radiation pattern which must be compensated by the inner rings of lesser weighting. The other design techniques for planar arrays produce side-lobes which are more or less uniform with angle. As can be observed from the ring configurations of Fig. 4, an array of rings does not resemble the density-tapered array and, therefore, similar results should not be expected. The decreasing side-lobes of the ring array design may be of advantage in the same way that the Taylor modified $(\sin \pi u)/\pi u$ distribution with decreasing side-lobes may be preferred in certain instances of conventional array design to the Chebyshev distribution.[13]

5. Sector Approach to Planar Array Design

Unequally spaced arrays designed using a density taper produce satisfactory radiation patterns, but do not yield element locations which can be considered as circularly symmetric as in the case of the ring arrays. Ideally there should be no restriction placed on the permitted locations of the elements (except for the half-wavelength minimum separation) in the design of an array by dynamic programming. With no restrictions the computer time for the design of a 40 wavelength diameter array using the 'minimax' criterion would be prohibitive.

In an attempt to provide more freedom in the location of elements than provided by ring arrays, but to avoid the excessive computations of the full planar aperture, the array was divided into a number of identical sectors, one of which might be as shown in Fig. 6. The possible element locations in each sector are arranged on a polar coordinate grid instead of a rectangular grid so as to simplify the calculations. The grid and the element locations must be selected to guarantee that adjacent elements be at least a half-wavelength apart. If an element is located at some point P in the basic sector, elements are similarly located in all the other sectors as at P′, P″, etc.

The requirement to maintain a half-wavelength minimum spacing and the use of a polar coordinate grid results in a hole in the centre of the array, the size of which depends on the sector size. With a 5-deg sector the hole is 11 wavelengths, with 10 deg it is 5 wavelengths, and with 15 deg, it is 0·5 wavelength. A correction for the hole is necessary when the smaller sector sizes are used. This may be accomplished in a number of ways but the method employed here was simply to fill the hole with a continuous illumination whose amplitude was made a variable.

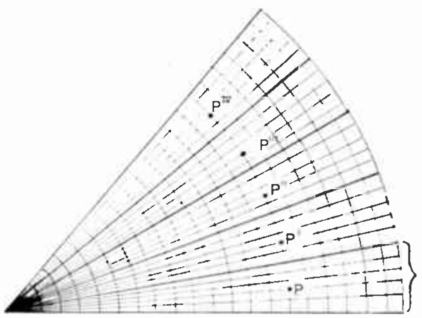


Fig. 6. Division of a planar array into identical sectors and the grid on which elements may be located.

Since the major interest is to explore and understand the potential of dynamic programming it seemed worthwhile to fill the hole in a manner which least complicated the computational problem. The radiation pattern (array factor) of an array with the hole filled with a continuous illumination is

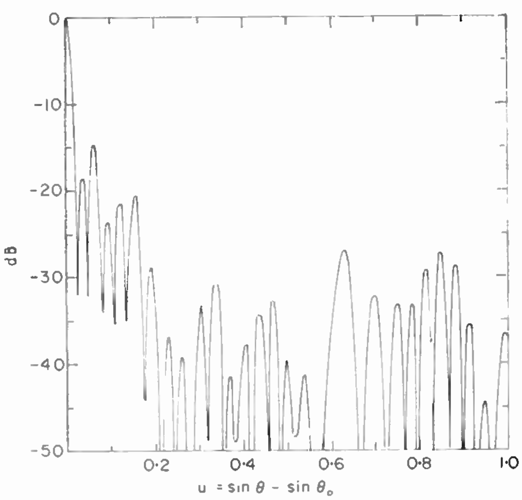
$$E(u,\phi) = \frac{A2J_1(au)}{au} + \sum_{n=1}^{N}\sum_{k=1}^{K/2}\cos\left\{2\pi r_n u\cos\left[\frac{2\pi}{k}(k-1)+\psi_n-\phi\right]\right\} \quad \ldots(6)$$

where A = amplitude factor of the illumination within the hole, $a = \pi d/\lambda$, d = diameter of the hole, $J_1(x)$ = first order Bessel function, (u, ϕ) = angular coordinates, N = number of elements within a sector, K (an even number) = number of sectors, and r_n, ψ_n defines the location of the nth element at a radius of r_n and angle ψ_n. The first term of eqn. (6) might be approximated in practice by a uniformly random array of $2A$ elements within the hole, or by an array of ring arrays of $2A$ elements, or by a half-wavelength equally spaced array with element illumination reduced as compared with the other elements of the array The exact manner in which the elements are placed in the hole is not believed to be too important to the radiation pattern, except at large angular displacements from the array normal (large values of u) where in most instances the side-lobes are generally sufficiently low not to be of any major concern.

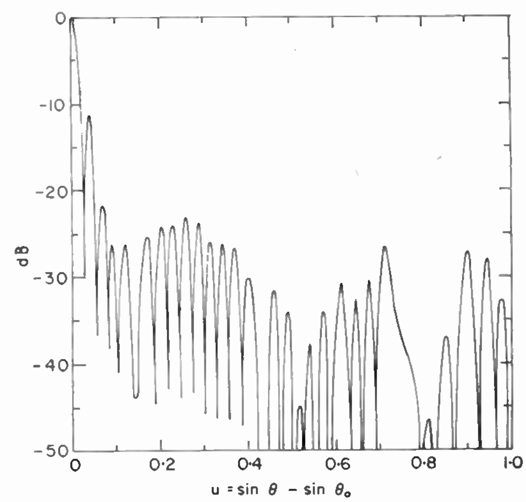
The computation time is proportional to $(N-1)m^2p$, where N is the number of decisions (element locations to be found), m is the number of possible locations each element may occupy, and p is the total number of values of the radiation pattern to be examined in determining the desirability of the pattern. To keep the computation time to a minimum most of the runs were taken with the 5 deg sector. The calculation time of a 5 deg sector with the CSC modified Univac 1107 required between 9 and 17 minutes per run depending on the input conditions. A few runs were made with the 10 deg sector at an average time of about 1¾ hours. A 15 deg sector was programmed and was estimated to run approximately eight hours. Since the machine did not have a re-start capability (the ability to re-start the problem in the middle if an error occurs) it was felt that the reliability of the results could not be guaranteed and shorter runs were submitted. This was unfortunate since the hole size of a 15 deg sector was small enough to have omitted the first term of eqn. (6).

The radiation pattern of a 40 wavelength circular planar array determined with a 5 deg sector and with a hole of 11 wavelengths diameter is shown in Fig. 7(a) ($A = 0$ in eqn. (6)). There is a high first side-lobe of −10·8 dB next to the main lobe. The second side-lobe is −21·5 dB. All other side-lobes are less than −23·8 dB. The high first side-lobe apparently is a result of the hole and is analogous to the high first side-lobe that results from aperture blocking in a conventional reflector antenna. Figure 7(b) is the radiation pattern obtained under the same conditions as Fig. 7(a) but with the hole completely filled ($A = 186$). The difference in the number of elements results from the additional 372 elements in the hole. The first side-lobe is considerably reduced but the second lobe whose value is −14·4 dB is the maximum. Beyond $u = 0{\cdot}15$ the peak side-lobe is less than −23·8 dB.

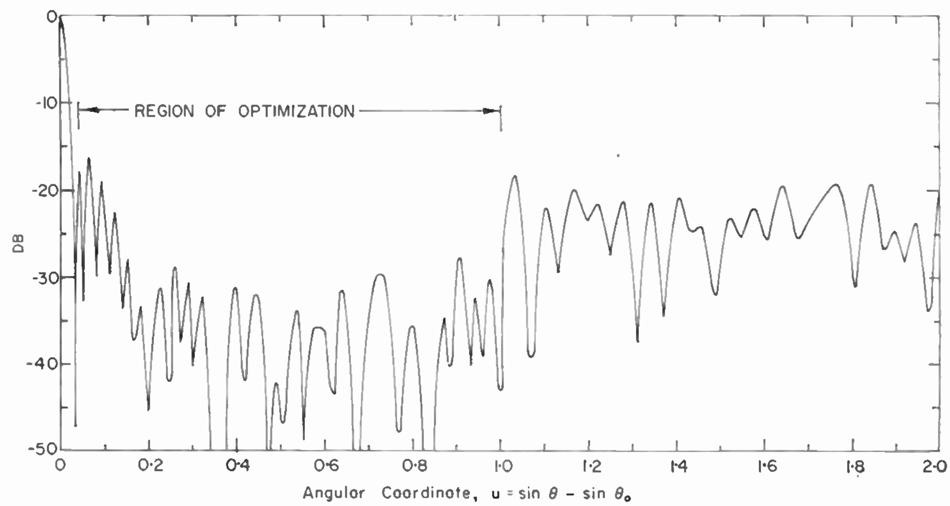
It would seem that since the first side-lobe was the maximum for $A = 0$ while the second side-lobe was maximum for $A = 186$ there ought to be some intermediate value of A which yields better patterns. Several other values of A were examined of which Fig. 7(c) was the best obtained. The near-in side-lobes are still relatively high. Although other choices of A and perhaps $u_{\min}$ might possibly provide improved results, it seems that the requirement for repeating elements every 5 deg produces an effect like that observed with the ring arrays. The side-lobes are



(a) Number of elements = 1008, 80% thinning. 11λ diameter hole in the centre (A in eqn. 6 equals 0).



(b) The hole in the array centre is completely filled ($A = 186$). Number of elements = 1380, 73% thinning.



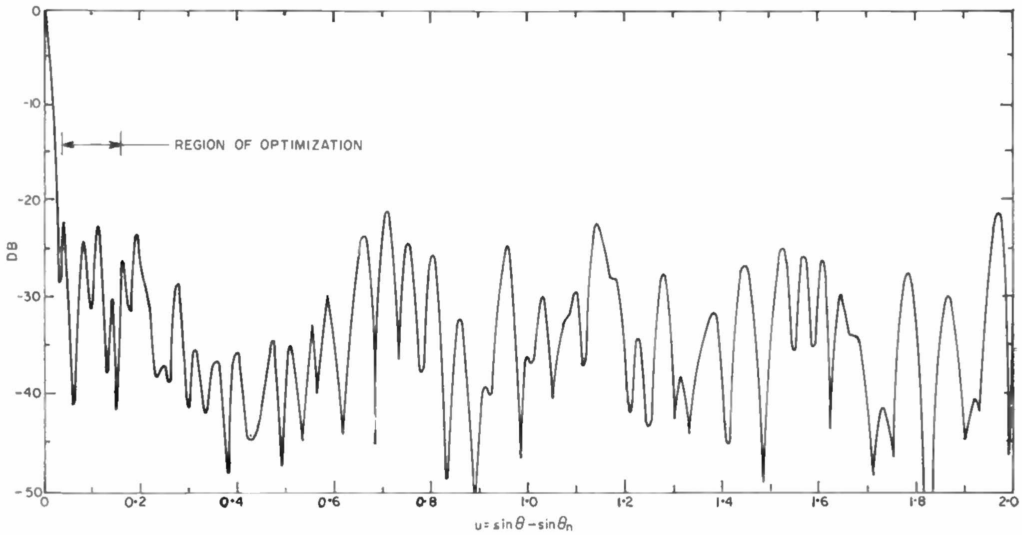
(c) Except that the hole in the array centre is only half-filled ($A = 93$). Number of elements = 1194, 76·5% thinning.

Fig. 7. Pattern of a 40λ diameter circular planar array designed by dynamic programming using the sector approach. Sector size = 5 deg, $u_{\min} = 0{\cdot}04$, $u_{\max} = 1{\cdot}0$.

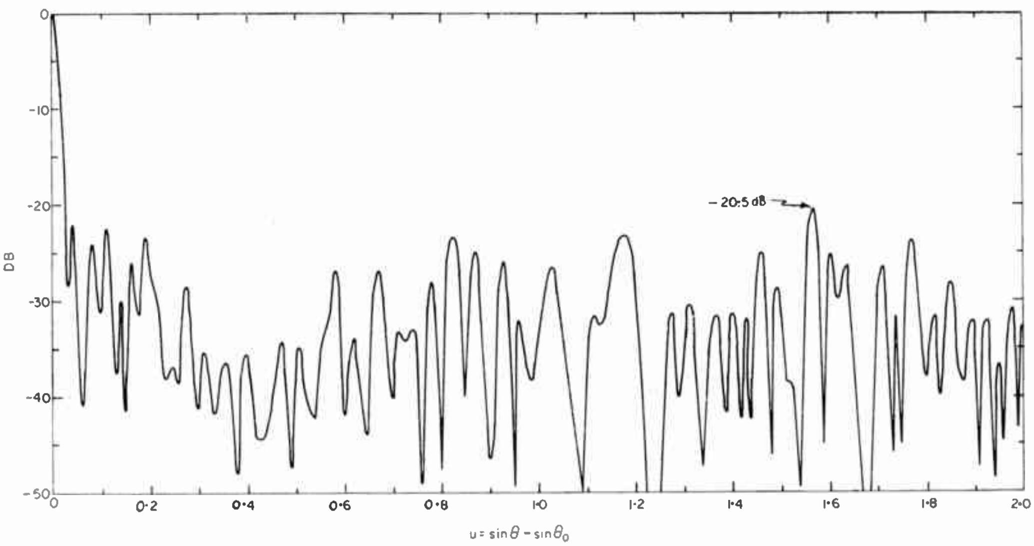
high near the main beam but decrease with increasing u. The larger the sector the less pronounced this effect should be.

An example of the radiation pattern obtained for a dynamic programming design with a 10 deg sector and $A = 53$ is shown in Fig. 8. The fourth side-lobe is the highest and is -18 dB. All other side-lobes at greater values of u are less than $-23{\cdot}5$ dB. Although it is slightly better than the similar example for the 5 deg sector it seems as though an even larger sector should be used if possible.

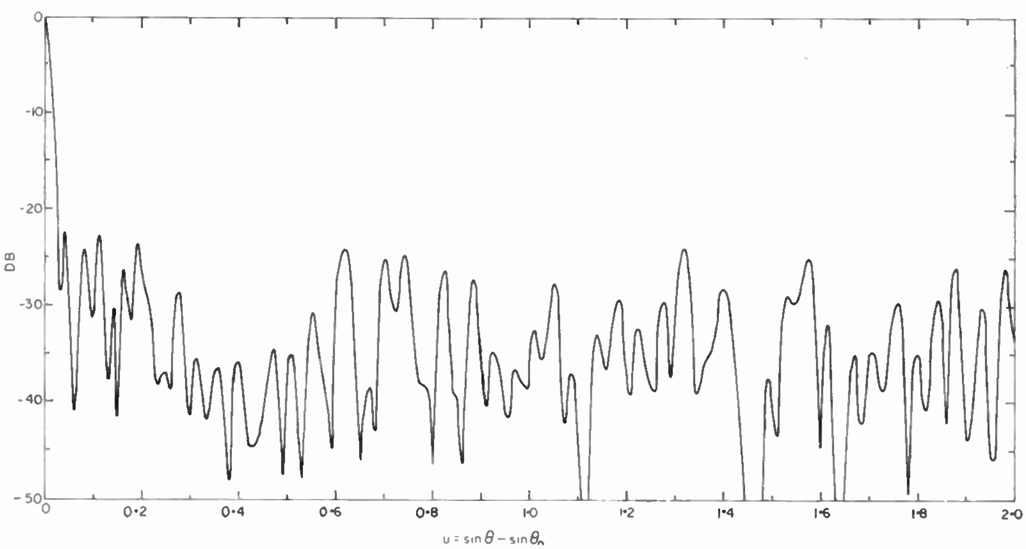
The 'minimax' criterion was employed in the above curves over the region $0{\cdot}04 \le u \le 1{\cdot}0$. Similar results were obtained for $u_{\min}$ of 0·03 and 0·05. In all the cases examined, the peak side-lobes occurred for u less than 0·15. For this reason, several designs were tried with $u_{\max}$ equal to 0·15 instead of 1·0. The best result achieved is shown in Fig. 9(a). The peak side-lobe is $-21{\cdot}5$ dB. Figures 9(b) and 9(c) show the patterns in two different planes within the 5 deg sector. The pattern is observed to be quite symmetrical within the region $u < 0{\cdot}4$. The peak side-



(a) Pattern steered and computed in plane $\phi = \phi_0 = 0$ deg.



(b) Pattern steered in the plane $\phi =$ 1·667 deg.



(c) Pattern steered in the plane $\phi =$ 3·333 deg.

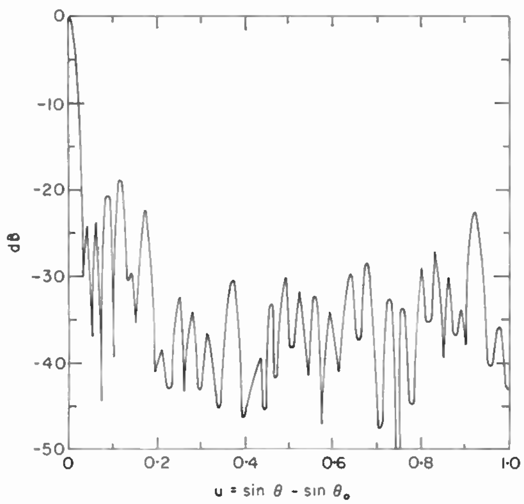


Fig. 8. Pattern of a 40λ diameter circular planar array designed by dynamic programming using the sector approach. Sector size = 10 deg, u_{min} = 0·04, u_{max} = 1·0. Hole in the centre partially filled (A = 53). 1114 elements, 78% thinning.

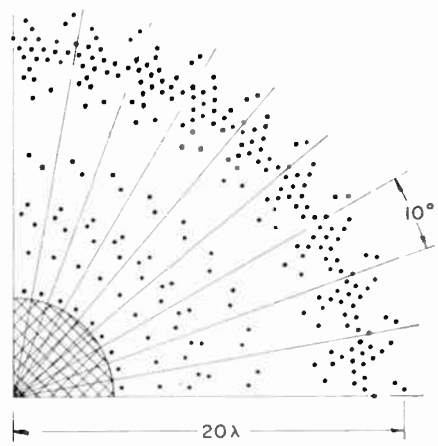


Fig. 10. Element locations for the array whose pattern is shown in Fig. 8. Sector size = 10 deg. Only one quadrant shown.

lobes in the visible region ($u < 1$) range from −22 to −25 dB. One side-lobe of −20·5 dB is found when the array is steered to $u = 1·57$. The element locations for the array whose pattern is given in Fig. 8 is shown in Fig 10.

6. Conclusions

The examples presented in this paper illustrate the applicability of dynamic programming as a tool for designing thinned, unequally spaced planar arrays. Although the theoretical validity of dynamic programming in this application is in doubt because of the inability to prove the principle of optimality, the results achieved provide sufficient justification for its consideration.

The work reported here was not intended to supply optimum designs. Instead, its purpose was to show the potential of dynamic programming and to indicate how it might be applied to planar array designs. In actual application to practical arrays it is important to eliminate as many of the constraints as possible. For instance, the restriction to the sector size imposed here for convenience of calculation should be as broad as possible and if feasible the computer should be programmed to select on its own the optimum value of u_{min}.

A great simplification could be achieved if the 'minimax' criterion could be replaced with some criterion of goodness which does not require the full radiation pattern to be computed. It has been observed that low side-lobes are obtained when the main beam null width is broad. It might be possible to use this fact as a criterion and reduce the length of the calculations.

Although dynamic programming has the potential for producing designs as good as or better than other known methods, it is an expensive method to implement for the large planar array. When cost is of importance, a cheaper technique such as the statistical density taper might prove satisfactory for many applications even though it might not yield an array with as low a side-lobe level.

Fig. 9. Pattern of a 40λ diameter circular planar array designed by dynamic programming using the sector approach. Sector size = 5 deg, u_{min} = 0·03, u_{max} = 0·15. Hole in the centre partially filled (A = 60), 840 elements, 83·5% thinning.

7. Acknowledgments

The work reported here was performed under Contract AF 30(602)–3057 with the U.S. Air Force Rome Air Development Center. Further details of the dynamic programming technique, additional array designs, the element locations for each design, and the computer program are given in the two reports issued under this contract[14, 15] which are available to qualified requestors from the Defense Documentation Center, Alexandria, Virginia.

We wish to acknowledge the assistance of Mr Michael Thomas in helping bridge the gap between the antenna theory and the computer programming and to Dr. George Nemhauser for his advice on dynamic programming. The computations of the ring array designs were performed on the U.S. Weather Bureau *Stretch* Computer located in Washington, D.C. The sector array calculations were performed by the Computer Sciences Corporation of

El Segundo, California. Mr. John Potenza was the RADC technical monitor for this project.

8. References

1. J. W. Sherman and M. I. Skolnik, "Thinning planar array antennas with ring arrays", *I.E.E.E. International Convention Record*, **11**, Part 1, pp. 77–86, 1963.
2. T. M. Maher and D. K. Cheng, "Random removal of radiators from large linear arrays", *I.E.E.E. Trans. on Antennas and Propagation*, **AP-11**, pp. 106–12, March 1963.
3. F. C. Ogg, "Steerable array radars", *I.R.E. Trans. on Military Electronics*, **MIL-5**, pp. 80–94, April 1961.
4. R. E. Willey, "Space tapering of linear and planar arrays", *I.R.E. Trans. on Antennas and Propagation*, **AP-10**, pp. 369–77, July 1962.
5. M. I. Skolnik, J. W. Sherman and F. C. Ogg, "Statistically designed density-tapered arrays", *I.E.E.E. Trans. on Antennas and Propagation*, **AP-12**, No. 4, pp. 408–17, July 1964.
6. R. Bellman, "Dynamic Programming". (Princeton University Press, Princeton, N.J., 1957.)
7. R. Bellman, "Adaptive Control Processes". (Princeton University Press, Princeton, N.J., 1961.)
8. R. Bellman and S. Dreyfus, "Applied Dynamic Programming". (Princeton University Press, Princeton, N.J., 1962.)
9. A. L. Moffett, "Array factors with non-uniform spacing parameter", *I.R.E. Trans. on Antennas and Propagation*, **AP-10**, pp. 131–6, March 1962.
10. M. I. Skolnik, G. Nemhauser, L. C. Kefauver, and J. W. Sherman, "Thinned, unequally spaced arrays designed by dynamic programming", I.E.E.E. PTGAP International Symposium, Boulder, Colorado, 11th July 1963.
11. M. I. Skolnik, G. Nemhauser and J. W. Sherman, "Dynamic programming applied to unequally spaced arrays", *I.E.E.E. Trans. on Antennas and Propagation*, **AP-12**, pp. 35–43, January 1964.
12. J. W. Sherman and M. I. Skolnik, "Effect of the number of elements on unequally-spaced arrays designed by dynamic programming", in preparation.
13. T. T. Taylor, "One Parameter Family of Line Sources Producing Modified Sin $\pi u/\pi u$ Patterns", Hughes Aircraft Co., California, Tech. Memo. 324, Contract AF 19(604)-262-F-14, 4th September, 1953.
14. "Pseudo-random Array Investigation", Technical Documentary Report No. RADC-TDR-63-380, September 1963. Electronic Communications, Timonium, Maryland.
15. "Pseudo-random Array Investigation", Final Report, 31st January 1964. Electronic Communications, Timonium, Maryland.

Manuscript received by the Institution on 2nd April 1964. (Paper No. 927.)

DISCUSSION

Under the Chairmanship of Dr. E. V. D. Glazier

Dr. R. Benjamin: Do you benefit in the realization of your computed pattern with drastic thinning, by reduction in mutual couplings between the element?

Dr. Skolnik (*in reply*): I would think so, although there is not universal agreement on this point. Experimentally measured radiation patterns of scale models of large planar arrays with pseudo-random spacing agreed quite well with the theoretical patterns. These measurements were made with a scale modelling technique known as the 'holey plate'. This indicates to us that, at least for the type of thinned arrays studied, mutuals do not cause significant effects. On the other hand, J. Allen of M.I.T. Lincoln Laboratory has reported that mutuals seriously affected the patterns measured of a small unequally spaced linear array. The question as to the effect of the mutual coupling was circumvented in the ESAR array described in the paper of Molz,† by providing the full number of elements and terminating those not used with a matched load.

Mr. P. R. Wallis: Why is it necessary to thin the array in this way? In practice a radar or sonar often requires a bunch of receiving beams (e.g. a monopulse cluster, or all-round in azimuth); it follows that the transmitting beam can be rather wider than those used for receiving. This suggests that rather than thinning, one should merely pack the smaller number of transmitting elements into a smaller area at the usual $\lambda/2$ spacing. This would give the same transmitting gain, perhaps a decibel or two better. Instead of transmitting a high average level of all-round radiation it is concentrated around the main beam.

Dr. Skolnik (*in reply*): What you suggest is not an unreasonable method for the design of an array and has merit in many applications.

† K. F. Molz, "Phased array radar systems", *Proceedings of the Symposium on Signal Processing in Radar and Sonar Directional Systems*", Birmingham, 1964, Paper No. 12.

U.D.C. 621.316.1

The Frequency of Occurrence and the Magnitude of Short Duration Transients in Low-voltage Supply Mains

By

J. H. BULL, B.Sc.†

AND

W. NETHERCOT, M.A., B.Sc.†

Presented at a Symposium on "The Operation of Electronic Equipment under Conditions of Severe Electrical Interference" in London on 15th April 1964.

Summary: Statistical information on the relationship between amplitude and frequency of occurrence of short duration transients in low-voltage supply mains has been obtained using multi-channel recording equipment. The equipment responds to pulses having durations exceeding 0·3 μs and amplitudes greater than certain predetermined levels. Either 4 or 6 level recorders have been installed for about 3 months at each of 19 sites chosen from industrial and domestic premises, and Area Board sub-stations. Analysis of the 36 000 transients which have been recorded show that the frequency of occurrence decreases approximately by 10 to 1 for a 2 to 1 increase in amplitude. Transients of up to 50 V peak occur frequently but one of 500 V or more is rare and is unlikely to occur at or near to sub-station transformers.

1. Introduction

Short-time voltage variations in low-voltage supply systems had in the past aroused relatively little attention since most apparatus is immune to their effects, either because of the long time-constant inherent in the apparatus (e.g. rotating machines) or because of the insulation margin provided to meet the requirements drawn up partly with a view to withstanding deleterious conditions (e.g. moisture deposits) and partly to ensure reliability. The much wider use of semiconductor devices in the last decade and the increasing emphasis on their reliability have drawn attention to the existence of variations of very short duration but of considerable magnitude.

Though otherwise reliable, many semiconductor devices are particularly sensitive to overvoltage and in some cases even a single overvoltage surge may cause breakdown.[1, 2] In addition to destructive damage which may be caused by these surges, there is the possibility of false operation of apparatus such as computers, which can arise from coupling between the apparatus and mains supply wiring.

Literature on the measurements of transients in low-voltage systems is sparse and up to 1961 at least little statistical information was available. Because of this the Electrical Research Association was asked to investigate the problem and to obtain the necessary data on the magnitude and frequency of occurrence of these short duration transients.

Methods of detecting these transients had been described in the literature but when this work was started there appeared to be no equipment for the automatic recording of the transients.[3]‡ It was therefore necessary for the E.R.A. to develop equipment specially for this purpose.[4, 5]

Measurements of transients have been made over a period of about two years at 19 selected sites; in all some 36 000 transients have been recorded.

2. Approach to Present Work

For a survey of the occurrence of transients on the public supply mains it is desirable to have equipment which will automatically record all transients above certain pre-determined levels. The equipment is of more general application if it can be left unattended for long periods and does not require frequent calibration.

Though apparatus connected to the mains is subjected to the voltage stress resulting from the combined effect of the supply voltage and the superimposed transient, a more accurate assessment of the transient itself is possible in the absence of the supply voltage. The voltage sensitivity of an instrument to record total voltage excursions cannot be chosen to give response to less than the peak of the supply wave (about 350 V on the systems being investigated) so that quite large amplitude transients which are in opposition to the instantaneous value of supply voltage will not be recorded. The counters are therefore preceded by a filter which rejects the components of the supply frequency and the first few harmonics (Fig. 1).

† The Electrical Research Association, Leatherhead, Surrey.

‡ Commercial equipment for the automatic recording of transients is now available.

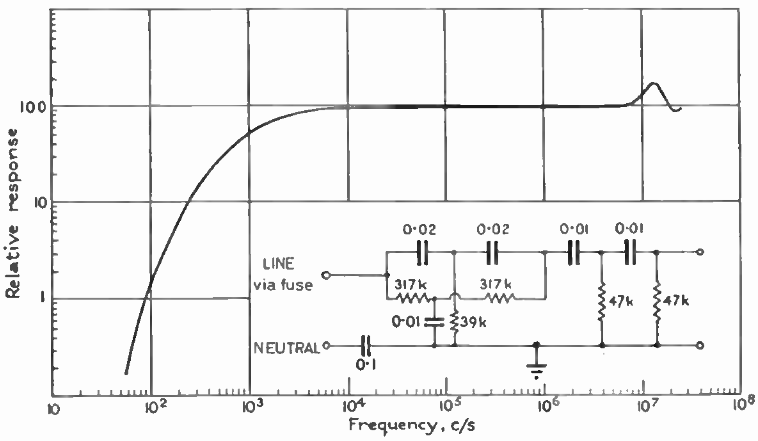


Fig. 1. Circuit and frequency characteristic of filter.

Attention is concentrated on the line-to-neutral signal as it is thought that surges in this circuit will affect most types of equipment which are sensitive to voltage transients. As short transients will not generally be coincident on two phases the situation line-to-line will be similar to that being investigated i.e. line-to-neutral. In order that one side of the terminals may be earthed, the neutral is connected to earth through a large capacitor at the instrument.

Since the positive and negative excursions of the mains supply are equally numerous and transients may be expected to occur in either half of the wave with equal probability, it is assumed that positive and negative transients will also occur with equal probability and, for simplicity, individual counting units are arranged to respond to one polarity only.

A very limited oscillographic investigation indicated the existence of transients lasting for only a few microseconds and there is evidence to suggest the possibility of damage to semiconductor devices by transients of this order of duration. The counters have therefore been designed to have constant sensitivity for transient durations down to less than 1 μs. The mains frequency rejection filter introduces some attenuation of the components below 1 kc/s (Fig. 1).

Present interest is centred mainly on the larger and less frequently occurring transients, so a high rate of counting is not necessary and a mechanical register is suitable for keeping a cumulative record.

3. Valve-operated Register Unit

This unit has a conventional power supply system energized from the low voltage a.c. mains supply which is being monitored for transients. An inductor in the live lead to the power unit isolates this as regards short duration transients. The voltage-sensitive detector circuit consists of two r.f. pentodes connected as a Schmitt trigger and the grid of the normally non-conducting valve is coupled by a capacitor to the input attenuator. Only positive pulses will operate the trigger and the sensitivity remains constant for pulse lengths down to less than 1 μs. The variation of sensitivity with pulse length for pulses with very short rise-time is shown in Fig. 2. A stable operating condition of the trigger circuit is obtained with a potential difference of about 50–100 V between grid and cathode of the input valve, the critical pulse amplitude being approximately equal to this bias.

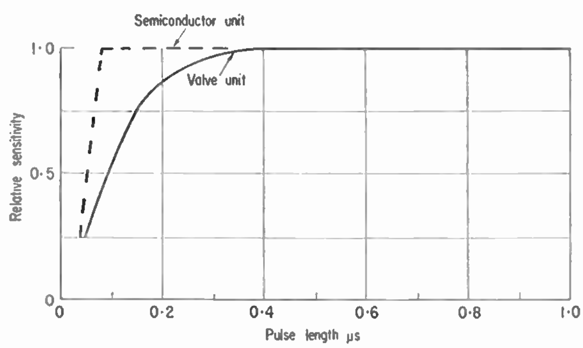


Fig. 2. Variation of sensitivity with pulse length.

A Post Office type of electromagnetic register is used for counting the pulses and is capable of up to ten operations per second. To reduce the size of the power unit required for a multi-channel recorder the supply to the register circuits is restricted so that the maximum repeated counting rate is reduced to about one per second. Several rapid operations of each register are however possible before temporary paralysis occurs.

The counters will continue to operate for a brief period after the failure of mains supply and therefore record any transients associated with the failure, but if the loss of supply persists the transients occurring during renewal will not be recorded.

A block diagram is given in Fig. 3.

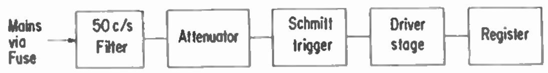


Fig. 3. Block diagram of valve operated unit.

4. Development of Unit using Semiconductor Devices

Whilst quite satisfactory in performance, the general bulk and power consumption of the valve instrument and its necessary reliance on continuity of the power supply makes for some inconvenience. A recorder using semiconductor devices in place of valves and operated from battery supplies has evident advantages.

A transistor version of the Schmitt trigger circuit is not sufficiently stable to give reliable operation for long periods owing to variation of characteristics with temperature, but another semiconductor device, the tunnel diode, provides a trigger with an operating current which does not change significantly with normal variations of ambient temperature. The tunnel diode switches extremely rapidly when the critical current is exceeded, and may be made to operate at a defined voltage by the addition of a resistor in series. The resistor must have low values of series inductance and shunt capacitance to maintain a constant sensitivity to pulses down to very short durations.

Following the tunnel diode another trigger circuit is required which may conveniently take the form of a pair of complementary transistors which in turn switch a power transistor to operate the electromagnetic register. A prototype unit has been constructed using high-speed switching transistors in the trigger circuit. The response of the unit to short pulses is shown in Fig. 2. With a 12 V battery a current of about 100 mA is required for about 50 milliseconds each time the register operates but the quiescent current is less than 5 μA per channel at room temperatures. A multi-channel unit will operate for several months from a small battery capable of providing the peak current.

The power supply to the registers in this equipment does not need to be restricted so that the counting rate is limited only by their speed of operation, and counting rates of up to 10 pulses per second may be achieved.

An additional facility is provided in this instrument by including circuits which cause it to respond in relation to the duration as well as the amplitude of the input pulse. A block diagram is given in Fig. 4.

The tunnel diode trigger circuit with one terminal of the battery earthed responds to negative polarity pulses. By interchanging the input connections and taking certain measures to reduce the capacitance to earth of some components, pulses of positive polarity may be detected. A modification cannot be made to the valve equipment to obtain a similar response to negative pulses owing to the large capacitance existing between the case of the instrument and many components. Several of these instruments are being constructed.

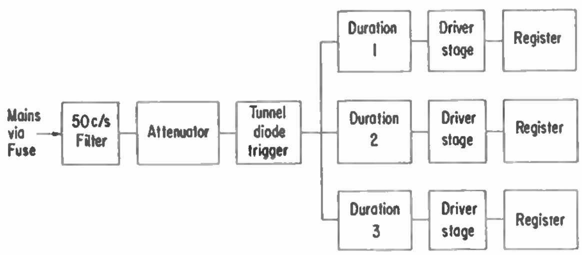


Fig. 4. Block diagram of semi-conductor device unit.

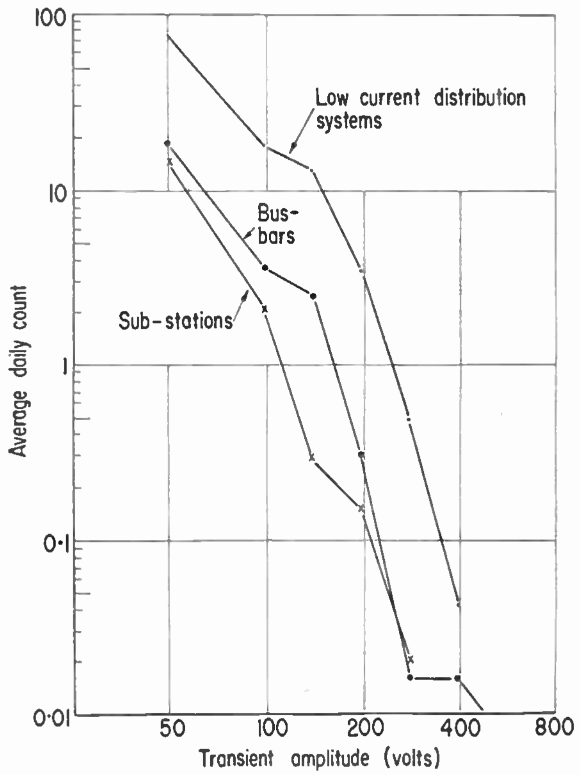


Fig. 5. Variation of frequency of occurrence with amplitude for short duration transients.

5. Results at Test Sites

The valve recorders have been used to obtain records of transients occurring at 19 locations. At each site either a 4- or a 6-amplitude channel unit was installed for several weeks. The test sites were chosen in order to obtain as much information as possible on the occurrence of transients in different parts of the supply system.

The accumulative summary of results is given in Table 1. The test sites have been divided into three classes determined by the types of distribution system between the recorder and the sub-station from which the supply is obtained. The average rate of occurrence of transients at each amplitude for each class of site is calculated from the total number of transients at the level in question and the total number of recording days at that level. The variation of the average rate of occurrence of transients at each amplitude for the three classes of test sites is shown in Fig. 5.

5.1. *Sub-stations*

For the first class of site, inside sub-stations, the recorders were placed as close as possible to the transformers, often connected by leads not longer

Table 1

Average Number of Counts per Day for Various Transient Amplitudes

Location of Site	Duration of Test (days)	50 V	100 V	140 V	200 V	280 V	400 V	560 V
Sub-stations—short connection to transformer l.v. terminals:								
1. L.E.B.	74	2·6	0·23		0		0	
2. M.E.B.	122	24·6	3·7		0·46		0	
3. Steel works 1	59			0	0	0	0	0
4. Steel works 2	43			0·14	0·02	0	0	0
5. Woolwich Arsenal	47	6·2	0·79		0·04		0	
6. S.E.E.B.	108		1·9	0·50	0·09	0·04	0	0
Average		14·4	2·03	0·29	0·15	0·02	0	0
Bus-bars—short distance from sub-station:								
7. E.R.A. Laboratory, Perivale	77	8·6	2·7		0·36		0·03	
8. Joseph Lucas Ltd. 1	96	5·1	2·4		0·29		0·06	
9. Joseph Lucas Ltd. 2	112	55	1·3		0·30		0·03	
10. Supermarket	129	3·5	0·15		0		0	
11. Car body factory	104		2·6	0·82	0·21	0·01	0·01	0·01
12. Steel works 3	59		1·1	0·29	0·03	0·03	0·02	0·02
13. L.E.B. Sub-station	57	18·6	2·5		0·44		0	
14. S.E.E.B. Sub-station	158		11·5	4·4	0·76	0·01	0·01	0
Average		18·7	3·66	2·50	0·31	0·016	0·016	0·007
Low current distribution systems:								
15. E.R.A. Laboratory, Leatherhead	52	69·1	2·0		0·33		0·02	
16. Domestic premises	57	84·5	9·1		0·26		0	
17. Wired-television amplifier station	100		3·8	2·25	0·94	0·08	0·01	0
18. S.E.E.B. Sub-station (local supplies)	87		14·8	8·4	3·7	0·59	0·14	0
19. Artificial fibre factory	83		54·8	31·4	10·4	0·86	0·04	0
Average		77·3	18·0	13·2	3·5	0·48	0·045	0

than 10 feet. However in one or two cases connection had to be made at a point on a distribution board which might be as much as 30 feet from the transformer. For this reason it cannot be assumed that a lower recorded rate of occurrence of transients at a particular sub-station necessarily means that the site is more free from voltage disturbances than the others. The three Area Board sub-stations in this class of site are all located in suburban areas with a little light industry in the district.

At sites 2 and 5 the 11 kV supply to the transformers also supplies power to nearby arc furnaces. This type of load, which has become notorious as a source of 'flicker', does not appear to cause abnormal transient variations. As a further investigation at site 2 the recorder was connected as the only load to the l.v. terminals of a 300 kVA transformer fed from the same 11 kV feeder that energized the transformer used in the previous test. The rate of occurrence of 50 V transients was about 0·15 per day with no other l.v. load compared with about 24·6 per day with the normal load, and no transient above 100 V was recorded whereas previously the rate had been about 3·7 per day. This is a definite indication that the majority of the transients recorded at the l.v. terminals of this sub-station transformer are generated in the low-voltage system rather than coming from the high-voltage side of the transformer.

As another test of the susceptibility to load variations, uncharged capacitors of several microfarads capacitance were connected to the l.v. terminals of the sub-station transformer at site 1 and also at the E.R.A. laboratory sub-station at Leatherhead. At both sites transients with amplitudes exceeding 100 V were produced and the waveform observed on an oscilloscope was a damped sinusoidal oscillation, with a frequency in the region of 10 kc/s corresponding to the series resonance of the capacitor and the connecting leads. The amplitude of transients produced by connection of the capacitor indicates a higher source impedance at the transformer at this elevated frequency than would be the case at 50 c/s.

5.2. *Bus-bars*

The second class of site consists of locations where the recorders were connected to the main low voltage feeders at the point of entry to the factory or building concerned, generally by wiring into a fuse-box on the main distribution panel. At the two Area Board sub-station sites connection was made in a building a short distance (about 100 feet) from the transformer. The shop (site 10) is supplied from the transformer used at site 1 in the previous class. At site 12 the distance from the transformer is much greater than at sites 3 and 4, although all three are located in the same area of the steelworks. For sites 8 and 9 the connections are to similar transformers in the factory sub-station which have different types of load; at site 8 largely light machinery and at site 9 chiefly welding equipment.

5.3. *Low-current Distribution Systems*

At each of these sites the recorder was connected to a local distribution circuit with about 300 feet of cable between the point of connection and the transformer supplying the system. A very high rate of occurrence of the lower amplitude transients is evident in most cases but at the higher levels, particularly at 400 V and above, the situation is not markedly different from that at the bus-bar sites.

6. Discussion of Results

During the period of more than two years that recorders have been installed at the 19 test sites a total of over 36 000 transients have been detected. The accumulation of this data by any means other than automatic recording equipment would have been totally uneconomic.

All the evidence suggests that the majority of transients detected on the low-voltage supply mains are generated in the low-voltage systems and in many cases may be attributed to the switching of appliances close to the recorder. The l.v. terminals of a sub-station transformer with a normal load are not free from transients and some with amplitudes of 200 V may be expected. It is evident that the transformer presents a considerably higher impedance to the system at the elevated frequencies of the chief components of the short transients being detected than is the case at 50 c/s.

At points on the mains supply system remote from the sub-station the occurrence of transients evidently depends critically on the local load situation. For long lengths of low-current wiring the impedance to be considered is that of the equivalent transmission line, probably in the region of 100 ohms in many cases, so that the switching of low-current non-inductive loads can still produce short duration transients, the amplitude of which are comparable with that of the supply voltage. It was noticed that switching on a 100 watt tungsten filament bulb in the domestic premises at site 16 produced a transient of more than 100 V on many occasions.

Although the distribution curves in Fig. 5 show some variations from straight lines on the logarithmic scales of frequency of occurrence and amplitude, a reasonable approximation is an inverse power law giving a 10 to 1 reduction in numbers for a 2 to 1 increase in amplitude. The extrapolation of the distribution curves cannot be relied upon to predict the likelihood of large transients with any great accuracy, but for the bus-bars class of site a transient of 1 kV might be expected perhaps once in three years at a typical site. At sub-stations the probability of a 1 kV transient is considerably less. For low-current distribution systems the variations in local conditions can be so great that an average prediction is valueless.

As a supplement to the present method of assessment where the transient alone is detected, recordings could be taken in the absence of the mains filter using a modified form of the present register units. In this case the response would be to the total voltage excursion, which is the most important parameter for some types of equipment, e.g. a.c./d.c. television sets with silicon rectifiers. It is in fact proposed to operate both types of recorder simultaneously from the same supply. With the addition of coincidence counters, it will be possible to determine whether or not the occurrence of the transient is random relative to the phase of the 50 c/s supply voltage.

7. Conclusions

Short duration transients, superimposed on the supply voltage, occur frequently in low-voltage public supply systems but the frequency of occurrence falls rapidly with increase in the amplitude of the transient. The frequency of occurrence decreases approximately by 10 to 1 for a 2 to 1 increase in amplitude.

At any given site, transients of up to 50 V peak occur at rates of up to 100 per day but transients of 400 V occur only about once a month or less.

Transients of nearly 600 V have been recorded but they are rare and are very unlikely to occur at or near to sub-station transformers.

Most of the transients are generated in the low-voltage systems themselves. They can be expected to occur more frequently in the low current sections, that is at points some distance from the sub-station.

8. Acknowledgment

The authors wish to thank the Director of the Electrical Research Association for permission to publish this paper.

9. References

1. D. B. Corbyn and N. L. Potter, "The characteristics and protection of semiconductor rectifiers", *Proc. Instn Elect. Engrs*, **107A**, p. 255, 1960.
2. M. Kubat, V. Maly and J. Zika, "Protective equipment for circuits with semiconductor power rectifiers", *Elektrotech. Obzor.*, **49**, p. 92, 1960. (In Czech. English translation: E.R.A. Translation Ref. Trans./IB1818.)
3. G. O. Crowther, C. F. Hill, J. P. Osborn and S. Zilkha, "Portable mains overvoltage detector", *Mullard Tech. Commun.*, **16**, p. 12, 1961.
4. S. F. Pearce and J. H. Bull, "High-frequency Transients in Low-voltage Supply Systems", E.R.A. Report Ref. V/T152, 1962.
5. S. F. Pearce and J. H. Bull, "High-frequency Transients in Low-voltage Supply Systems. The Development of Detection Circuits", E.R.A. Report No. 5010, 1963.

Manuscript first received by the Institution on 12th March, 1964 and in final form on 7th July 1964. (Paper No. 928.)

U.D.C. 621.3.013.7: 621.382: 621.311.29

Measuring Transient Overvoltage on Instrument Feeders

By

H. J. TURNER, B.Sc.(Eng.)†

Presented at a Symposium on "The Operation of Electronic Equipment under Conditions of Severe Electrical Interference" in London on 15th April 1964.

Summary: The problem of obtaining information on unwanted transient voltages which may damage semiconductors is considered in relation to electricity generating and switching stations. It is concluded that the collection of data from actual sites by an unattended monitoring instrument would yield valuable information with the minimum effort. The factors affecting the specification of such an instrument are discussed and an instrument developed for this purpose is described.

1. Introduction

Instrumentation in the generating and switching stations of the Central Electricity Generating Board has, in the past, relied mainly on electro-mechanical devices backed in some instances by electronic equipment containing thermionic valves. With the advent of nuclear stations it was necessary to use very many more electronic devices to provide adequate control and protection, and again much of the early instrumentation used thermionic valves.

Since transistors have many advantages over valves their introduction to modern power station instrumentation and control was inevitable. However, transistors are more liable to damage from short duration overloads because of their lower breakdown voltages and the much lower thermal capacity of their functional parts. Since the C.E.G.B.'s stations handle extremely large amounts of electrical energy, it is thought that large voltages of relatively short duration may occur on many circuits connected to electronic instruments.

The new interest is directed towards the shorter duration voltages which would not have damaged more robust thermionic components, but which could cause widespread damage to semiconductor devices. In addition to these large, infrequent overvoltages which damage the devices, there is also the more frequent possibility of lower level voltages which can cause momentary false outputs from instruments, causing, for example, spurious tripping of a reactor. The use of redundancy to minimize the effects of unwanted signals is possible, but is not yet really practicable for the case of damaging overvoltages where really massive redundancy would be required. Nevertheless, it is felt that the most difficult problem at the moment is that of the extra large voltage which by its rare occurrence will be hard to trace, and it is the need for information relating to such voltages which is dealt with in this paper.

† Central Electricity Research Laboratories Leatherhead, Surrey.

2. Sources

Transient overvoltages are coupled to the instrument wiring inductively or capacitively, by radiation or through some common impedance, commonly that of the earth system. The highest degree of coupling will generally occur at high frequencies so that the resulting waveforms on the instrument wiring will generally be proportional to the rate of change of voltage or current of the originating surge. The inductances of the common impedances and the values of coupling capacitance need not be large when the rate of change of voltage or current at the source is very high. Switching operations on overhead lines are often a source of high rates of change of both voltage and current; their effects can be coupled into the low voltage wiring by any of the above methods, e.g. by the capacitance across mains transformers and instrument voltage transformers and current transformers. Figure 1 shows the voltage at the top of the earthed supporting frame of a capacitor voltage transformer with respect to a remote earth point when a 275-kV line was charged from another. A less frequent and even less predictable rise in local earth potential will be caused by the onset of a fault on any heavy power consuming circuit.

Since most heavy power circuits are three-phase systems, the resultant short duration surges due to switching or faults will depend on the relative switching instants of the three phases and the high frequency symmetry of the circuits.

Lightning strikes are another source of fast rising voltages and currents of relatively short duration, and

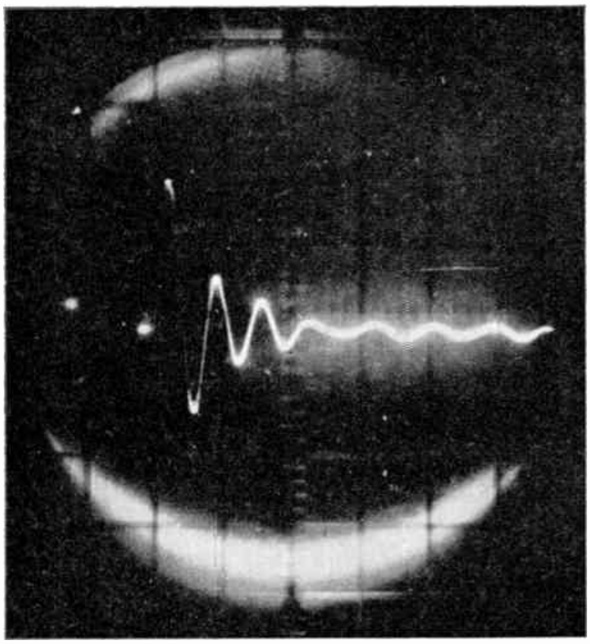


Fig. 1. Voltage between the earthed voltage transformer tank and a remote earth at the instant of charging a 275-kV overhead line.

Scales: Vertical, 500 V per major division.
Horizontal, 10 μs per major division.

the effects of these may reach low voltage wiring via direct strikes on the overhead lines or less frequently, though more seriously, by a direct strike on a building causing high currents in the metal framework.

3. Relation to Semiconductor Devices

The most common cause of damage to a semiconductor device is due to overheating of some part of it. Sometimes this occurs in the connecting leads causing fusing, and thermal cycling can cause fatigue fractures of the soldered connections to the crystal element; where connecting leads are more robust and improved methods of making contact are used, the overheating occurs within the semiconductor itself, either internally or quite commonly at a point on its surface. In many devices the biggest voltage drop occurs across the depletion layer, hence the major part of the power absorbed is converted into heat in a small region within the crystal. The amount of energy which may be absorbed as heat in this region for a given temperature rise is a function of the duration of the input. If the region remains constant in size and position the permitted energy input would be constant for durations considerably less than the time constant of this region, i.e. the product of its thermal capacitance and its thermal resistance to the case. For durations longer than this the permitted energy increases steadily until the rate of increase reaches its maximum for durations longer than the bulk time-constant, i.e. the product of the same thermal resistance and the total thermal capacity of the crystal element. A curve showing the maximum temperature rise caused by a rectangular power pulse of constant energy against pulse duration for a simple thermal model is shown in Fig. 2.

In practice the volume of the heated region will vary with current density and applied voltage and in

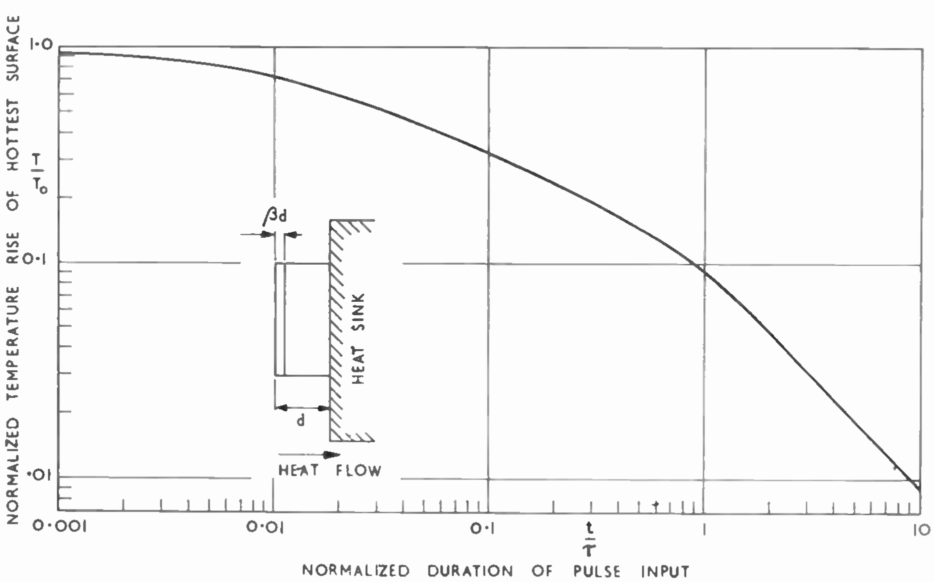


Fig. 2. Temperature rise as a function of duration of a constant-energy rectangular-pulse input for a simple thermal model of a transistor.

some circumstances can change position entirely, thus changing its thermal capacity or thermal resistance or both and hence altering the temperature rise characteristics for a given power input. It is thus very difficult to predict the thermal behaviour of a device unless something is known of the way it behaves electrically to large currents and voltages. Except in a few cases it is at the present almost impossible to predict the current and voltage distributions under overload conditions.

4. Electrical Breakdown of p–n Junctions

When a p–n junction or combination of junctions is subjected to increasing electrical stress, a point is reached at which there is an abrupt change in the mode of conduction which may consist of a sudden increase in current for a small rise in voltage or a drop in voltage as the current is increased. These changes in themselves are usually reversible provided the junction is not permanently damaged by the overheating which may occur. When a voltage breakdown occurs the current flowing will generally be limited only by the impedance of the external circuit, with the result that there will be a sudden increase in power input to the device which, if large enough, will destroy the junction by overheating.

There are at least five known modes of breakdown and these are conveniently divided into two groups—those dependent on voltage alone and those which depend on the energy input.

The three types of voltage breakdown are: (*a*) avalanche; (*b*) Zener; (*c*) punch through. These are described more fully in refs. 1 and 2 and are characterized by an abrupt change in dynamic resistance to a very low value at a critical voltage; where there is a combination of junctions avalanche breakdown can also exhibit a negative resistance region, the voltage across the junction falling to a lower value than that required to initiate the effect.

The other types of breakdown are:

(*d*) thermal runaway, and (*e*) second breakdown (see refs. 3 and 4).

Thermal runaway is fairly slow to develop, taking several seconds, so that it is not very important when short duration overloads are being considered. Second breakdown is still time-dependent but develops much more quickly. Little is known of the mechanism but a local melting over a region of about 5×10^{-4} cm radius has been suggested by A. C. English[4], who has observed a red glow from junctions exhibiting this phenomenon. The time-constant associated with this effect can be about 60 microseconds, although Oka and Oshima[5] have observed a similar effect with a time-constant of less than 1 microsecond.

The important characteristic of this effect from the transient overload point of view is that it exhibits a negative resistance region, the voltage across the device falling to a low value at a critical current. This means that under certain load conditions a device can be permanently switched into a region of excessive power dissipation by a relatively short duration overload which is not damaging in itself.

5. Methods of Assessment

The ideal approach to this problem would be one which allows the prediction of the worst voltage which can occur at the relevant points in an instrumentation system. In order to do this it is necessary to know all the nodes of a system at which unwanted energy may be injected, together with its magnitude. In a large power station there are very many possible points of injection whose source powers may be difficult to estimate. Even if the sources are known, it still remains to calculate or measure the transfer function between each pair of injection nodes and vulnerable instruments. In carrying out such tests or calculations consideration would have to be given to any possible non-linearities such as dielectric breakdown or iron saturation. In the C.E.G.B.'s stations the numbers of paths to be considered would be extremely large and the possible injected powers largely unknown so that a meaningful assessment by such methods would require a very large effort.

Another approach which is much less expensive in manpower is to keep a continuous watch over a considerable period on a large number of vulnerable points, with unattended instruments, which will record the numbers of damaging overvoltages that occur. For this approach to have much advantage over the previous one the numbers of instruments and duration of the experiment must be sufficient to ensure a reasonable chance that some unusual condition will occur.

The relatively common transients due to line switching and operation of circuit breakers, etc., of which a limited number can be made to order, can be examined in more detail using a suitable cathode-ray oscilloscope. Since only a few such operations will be permitted in an operational station it is necessary to take special precautions to ensure that full details of such transients are recorded. For such recordings, control must be exercised over the point on the cycle at which contact breakers are opened or closed. Very often the most important part of the transient is preceded by unimportant signals which will spuriously trigger the oscilloscope time-base, or the transient may consist of several 'bursts' of short duration, the first of which may not be the largest, spaced over relatively large time intervals. In addition, the order of the amplitude to be expected is unknown. It is therefore necessary to have an oscilloscope capable of resolving a wide range of time and amplitude on a single trace, or alternatively to use several instruments in parallel with differing time and voltage scale settings.

6. Requirements of a Monitoring Instrument

The fullest information would be obtained from an oscillogram, together with a measure of the source impedance, but since a large number of points must be monitored over a considerable period of time, such measurements would be very expensive. A practical instrument must be simple so that it is reliable, cheap, and requires the minimum of maintenance; in fact, once installed the only attention should be that of reading.

These requirements severely limit the amount of information which can be obtained from each transient since it is only practicable to measure one characteristic per instrument, the choice of which will be governed by the simplicity of the circuitry required and the relative importance of those possible.

The three characteristics which are essential when assessing the harmful effects of transients to semiconductor devices are: amplitude, duration and source impedance. The most important of these is amplitude, since the voltage breakdown usually initiates the damaging overload. However, since a single characteristic is desirable, one which combines two or more of the fundamental ones should be considered. Source impedance can only be dealt with by actual measurement and, since this in many cases will be determined by the surge impedances and capacitances of cables, it should remain constant in most cases. Since energy is the ultimate factor causing damage, some quantity proportional to this such as $\int v^2 dt$ would give useful information, but this is difficult to measure, although a suitably thermally isolated thermistor of very small thermal capacity could be used over a limited range. A more easily measured factor is the time integral of voltage or pulse area.

The measurements may be recorded in two ways: either the actual value of the characteristic, or the number of times it exceeds a preset value may be recorded. The first method requires a relatively large storage capacity unless only the highest value is recorded. The second method gives a less precise value, but since only the approximate levels are needed, it is capable of giving much more information for a given storage capacity. This method is also the easiest to engineer and has the advantage that the relative risk to a variety of devices can be more easily assessed.

The instrument should be capable of correctly recording, irrespective of the rise-time and duration, pulses longer than about 1 microsecond. If voltage is to be measured the input range is determined by the range of p–n junction breakdown voltages which vary from a few volts to the order of a thousand volts. For energy the range should begin at about 10^{-4} joules and an upper limit would be set at about one joule by limiting the pulse duration to be measured to say 100 microseconds. For pulse area a range of from about 10^{-4} volt-seconds to several volt-seconds would be suitable.

It is desirable that such an instrument should be capable of being used across isolated pairs of leads, which may have transient common mode voltages with rise-times of only a few nanoseconds, or when it is not known which lead will have a transient voltage with respect to earth. The instruments will not always be able to be situated where stray capacitances to it are not important, so that it is desirable that the outer case should be capable of being earthed locally so that there is no fear of increasing the possibility of pick-up at the point of connection. This is particularly important when protection and safety circuits are being monitored. Also in such cases the instrument should be incapable of injecting signals into the point of connection, nor should its input impedance ever become low enough to cause maloperation of any circuit. The instrument should be capable of dealing with pulses of either polarity.

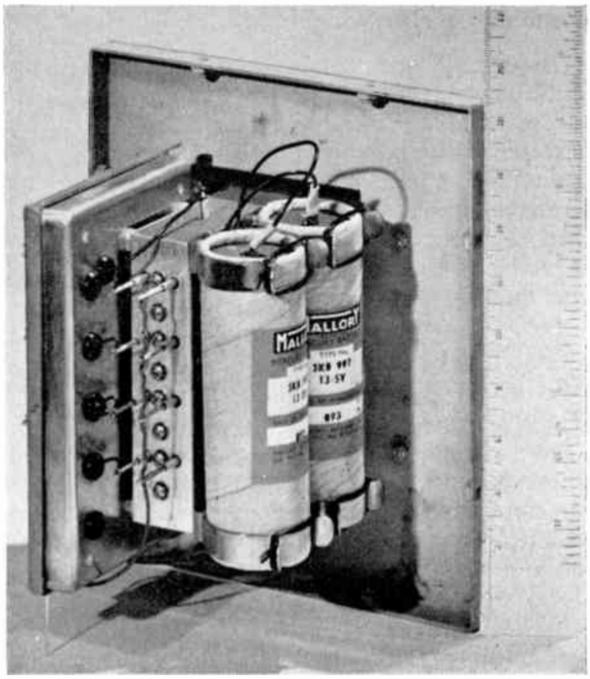


Fig. 3. The C.E.R.L. surge amplitude monitor with outer case removed. (Early model; batteries are now screened, see Fig. 4.)

7. The Instrument Designed

In consideration of the foregoing an instrument was designed to record the number of pulses which exceed each of four preset voltage levels. A photograph of the instrument is shown in Fig. 3 and the circuit diagram of one channel in Fig. 4. Four levels were chosen to keep the instrument physically small and yet allow a voltage range of the order of 30 : 1

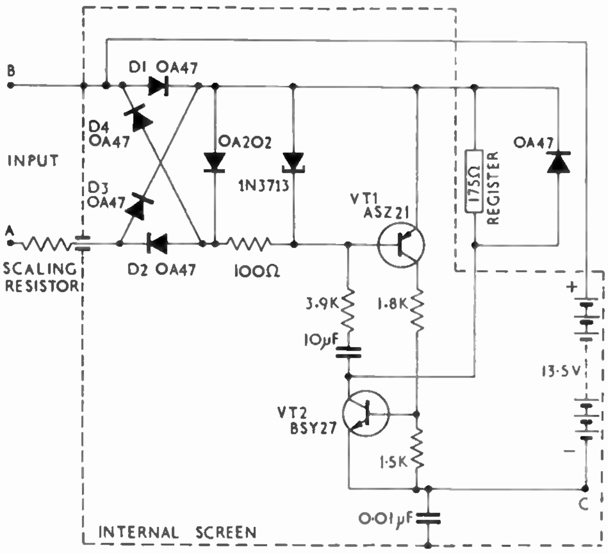


Fig. 4. One channel of the 4-channel surge amplitude monitor.

to be broken into reasonable groups. The instrument has a negligible quiescent current drain so that it can be battery operated and have no external controls. This latter requirement was considered necessary so that the instruments can be left unattended in places where they may be disturbed accidentally. The battery life is about two years or 10^6 counts when operated in an ambient temperature range of $-10°C$ to $+50°C$.

The voltage discrimination is obtained by passing a current, proportional to the voltage, through a tunnel diode which is connected across the base emitter junction of one transistor of a monostable pair. The increase in static resistance of the tunnel diode when the current through it reaches its critical value causes a fraction of the current to be diverted to the base emitter circuit of the transistor VT1, so turning the two transistors on for a period determined by the time constant of the feedback path or the duration of the pulse, whichever is the greater. An electro-mechanical counter in series with the collector of the second transistor VT2 registers the occurrence of a voltage greater than the critical level. The critical voltage level is determined by the value of the scaling resistor in series with the tunnel diode which has a critical current of 1 mA.

The use of a fast switching transistor in the second stage enables a rectifier bridge input to be used in conjunction with a common battery supply for any number of circuits without any appreciable difference in the response to positive- or negative-going short pulses, which would otherwise occur due to the negative feedback caused by the emitter current of VT1 flowing through one diode of the bridge.

It was considered very important that the counter should not register abnormally high voltages in the presence of pulses with rise-times of a few nanoseconds. Since the circuit is inherently sensitive to extremely small amounts of electrical energy it was found necessary to screen completely the transistor tunnel diode circuits and battery from the scaling resistors. This screen then becomes one terminal of the input so that the whole has to be encased in a further metal screening box. The scaling resistor layout has been kept simple so that the ranges may be changed readily. This has led to the reduction of the circuit time response to about 1 microsecond to ensure that the attenuator will be suitably compensated when fast rising pulses are to be measured. Figure 5 shows the theoretical response of the counter taking into account the response due to the scaling resistors alone. The points shown are those measured with 3×200 kΩ resistors in the scaling network. The initial pulse of current due to the stray shunt capicitance of the resistors lasts for only a few nanoseconds and its amplitude is always less than the final value.

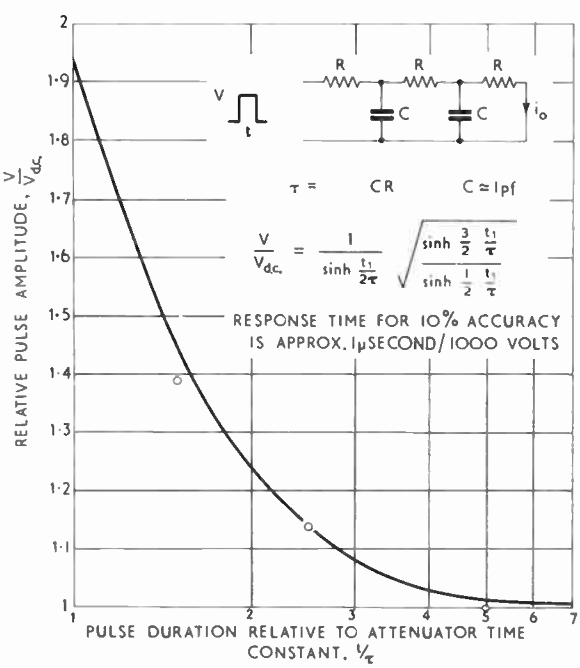


Fig. 5. Theoretical response of counter with three type NJ65 oxide scaling resistors.

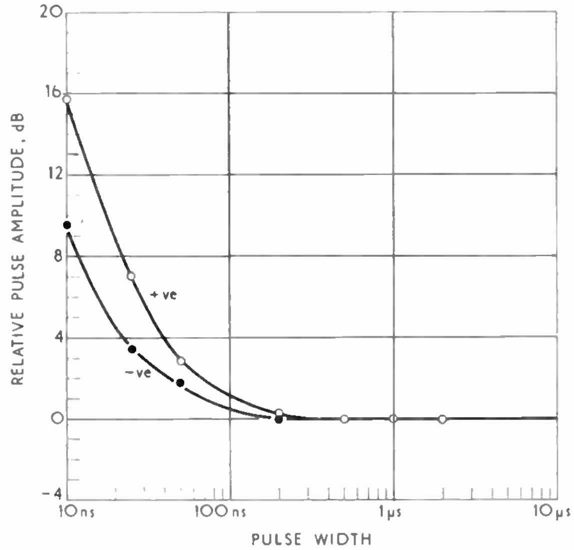


Fig. 6. Response of the counter to a short pulse at low input voltage (5 V).

Figure 6 shows the inherent response of the basic circuit taken with a single 4·7 kΩ scaling resistor which does not modify the circuit response. The difference in response at short durations is due to the negative feedback previously mentioned. When the instrument is used under conditions where the inner case stays near earth potential, the voltage levels can be set up to 1·5 kV with no special precautions. If the instrument is to be used on isolated pairs when it is not known which lead will have a transient potential difference with respect to earth, it is necessary to screen the scaling resistors from the outer case.

The instrument will not falsely register if a common mode signal of 1000 V with an exponential rise-time constant of 5 nanoseconds is applied between the joined input terminals and the outer case.

The input impedance is determined by the parallel combination of the four scaling resistors. The stray capacitances are: about 3 pF between the two isolated terminals, about 3 pF between the scaling resistor terminal and outer case, and about 50 pF between the inner and outer screens.

To summarize, an instrument costing about £70 is available which can be made in the large quantities required for a statistical survey of the amplitudes of pulses from 10 V to 1500 V, or higher with an external attenuator, of durations greater than 1 microsecond. The pulses may be of either polarity or they may be counted irrespective of polarity. The instrument may be connected across isolated pairs of lines or between a line and earth and may be left unattended for long periods in almost any environment found in the C.E.G.B.'s stations.

8. Deployment of Counters

The C.E.G.B. has about 240 generating stations plus about 500 substations, so that a complete coverage with say only 20 instruments per station would require an excessive number of counters.

Of the generating stations, only about 40 are continuously operating on base load, with about 60 which have at least some generators on for a large percentage of the time. About 50 stations are only used for peak load conditions and the remainder only very infrequently used. Stations of the first two groups have the highest generating capacity. It therefore seems reasonable to limit the survey as far as generating stations are concerned to the 40 base load stations.

The initial experiment will involve the installation of about thirty counters on various circuits in West Thurrock generating station, together with about twelve counters in a substation, initially Chessington. This should give some indication as to which circuits are

Table 1

Results obtained at Chessington 132 kV Switching Station from January to May 1964, using the C.E.R.L. surge counter

Circuit	Surges recorded
Overcurrent C.T. phase to neutral measured via a 50 c/s filter +ve and −ve (one instrument per phase).	20 over 90 V 1 over 250 V 1 over 600 V 0 over 1500 V
V.T. phase to neutral measured via a 50 c/s filter (one instrument) + ve and − ve.	9 over 90 V 0 over 250 V
V.T. neutral to local earth (one instrument) + ve and − ve.	8 over 50 V 0 over 150 V
†110 V d.c. battery at relay room. Capacitor coupled to measure increases in p.d. only.	55 over 50 V 5 over 150 V 0 over 500 V
†110 V d.c. battery centre tapped by 2 × 0·1μF capacitors to local earth. + ve and − ve.	8 over 50 V 0 over 150 V
Across contact breaker trip coil. Single polarity connected to measure induced voltage only.	37 over 150 V 6 over 300 V 0 over 600 V
Across contact breaker trip contacts—capacitor coupled. Single polarity connected to measure induced voltage only.	90 over 50 V 1 over 150 V 0 over 500 V
†50 V d.c. battery in control room capacitor coupled to measure increases in p.d. only.	583 over 25 V 61 over 50 V 0 over 100 V

† The instrument would record pulses of the battery level on each restoration of supply after an interruption. Records show that not more than 10 such interruptions had taken place over this period.

the most vulnerable so that, say, five counters could then be installed in each of about twenty of the base load stations on these circuits. The counters in the main survey stations will be moved on to different circuits at intervals of not less than six months. In this way it is hoped that about 200 counters operating continuously for several years will give a reasonably reliable survey of the amplitudes and frequency of short duration voltage pulses on the C.E.G.B.'s instrumentation wiring. Preliminary results obtained over five months at the Chessington sub-station are shown in Table 1.

9. Conclusions

The introduction of the more fragile semiconductor device in place of the relatively robust valve and relay in an environment where large amplitude, short duration voltages are expected to occur has made a survey of that environment essential. These devices are less tolerant to the overheating that occurs when they are overloaded because of their much lower thermal capacities. Voltage breakdowns which lead to overloading generally occur at much lower voltages. In view of the complexity and size of the C.E.G.B.'s system, the most convenient survey that can be made is by a continuous monitoring of a large number of vulnerable points with about 200 cheap unattended instruments which count the number of occurrences of voltages above four preset levels.

As far as semiconductor devices themselves are concerned, very little is known of their ability to withstand short duration overloads other than knowledge of the electrical breakdown voltages. It seems likely that some devices in which breakdown takes place uniformly over the whole junction area will be capable of withstanding much larger overloads than those in which breakdown is localized. The conditions under which the various types of breakdown take place require more study so that devices can be characterized by their relative ability to withstand short duration overloads.

10. Acknowledgments

The work described in this paper carried out at the Central Electricity Research Laboratories, Leatherhead. This paper is published by permission of the C.E.G.B.

The author also wishes to acknowledge the help of Messrs. G. W. Bowdler, S. J. Chesney and B. J. Terry, of C.E.R.L., in the development of the surge amplitude monitor, and the co-operation of Mr. G. S. H. Jarrett, Technical Engineer, South Thames Division, for facilitating the installation of the monitors at Chessington.

11. References

1. S. C. Ryder-Smith, "Voltage breakdown of transistors", *Electronic Technology*, **38**, pp. 348–56, October 1961.
2. L. van Biljon, "Transistor avalanche voltage", *Electronic Technology*, 37, pp. 72–6, February 1960.
3. H. A. Schafft and J. C. French, "'Second breakdown' in transistors", *Trans. Inst. Radio Engrs (Electron Devices)*, **ED-9**, No. 2, pp. 129–36, 1962.
4. A. C. English, "Mesoplasmas and 'second breakdown' in silicon junctions", *Solid State Electronics*, 6, pp. 511–21, 1963.
5. H. Oka and S. Oskima, "Breakdown in silicon power diodes", *Mitsubishi Denki Laboratory Reports*, **3**, No. 2, pp. 165–81, April 1962.
6. S. F. Pearce and J. H. Bull. "High-frequency Transients in Low-voltage Power Systems", Electrical Research Association, Report No. 5010.
7. Associated Electrical Industries Ltd., Provisional Patent No. 35913/61.

Manuscript received by the Institution on 3rd March 1964. (Paper No. 929.)

Instruction in Electronic Drawing for Students of Radio and Electronic Engineering at Professional Level

A Report of the Education and Training Committee of the Institution

There has been recently a good deal of discussion as to whether the traditional subject of Engineering Drawing should be included in the educational requirements of a radio and electronic engineer. Although there is among electronic engineers a strong feeling that engineering drawing (as normally understood) is undesirable, there has been some pressure to include the fundamental principles in the syllabus in the interests of uniformity of initial training among the various types of engineer and also because it is an important means of communication of ideas. Indeed, it is included in the syllabus for Part I of the Engineering Institutions Examination to which the I.E.R.E. subscribes.

There can be little doubt that the manual skills acquired as a result of training in engineering drawing are of great use to the electronic engineer. However, the engineering background and knowledge through which these skills are acquired, being traditionally associated with machine parts such as steam stop valves, frequently leads the electronic engineer to regard the subject as largely uninteresting and irrelevant.† Thus such training is not only highly inefficient in the training of the electronic engineer, but it may be actually harmful.

However, illustrations taken from the electronic field (e.g. microwave components, crystal structures) are perfectly effective and much more acceptable in giving the student an opportunity to demonstrate his ability to reason in three dimensions. The new Part I examination syllabus does not preclude such examples but it may require the setting of optional questions.

There would probably be general agreement that the standard of drawing among professional radio and electronic engineers, even when only simple electronic circuits are concerned, is generally poor. Such low standards of drawing are generally accepted in the same way as is bad writing, namely because supervisors have not been trained to expect anything better. Unless the professional engineer has the background knowledge of drawing and layout, even draughtsmen and tracers concerned with electronic drawing will produce inferior work. There is a common experience that the channel of communication between the engineer and draughtsman is inadequate and skill in graphical representation is essential if the engineer is to convey information to a drawing office or to a model shop. At the same time, good graphical work serves to clarify his own ideas. Thus there seems to be an overwhelming case for an improvement in the standard of drawing skill and understanding among professional electronic engineers and this would raise the standard of drawing at all levels.

It is the view of the Education and Training Committee that the basic principles of Engineering Drawing as tested in the revised syllabus of the Engineering Institutions Part I examination (Appendix 1) should be part of the curriculum of the radio and electronic engineer, although as previously stressed, examples of simple engineering components must be chosen judiciously. In addition, it is considered that professional electronic engineers do not pay sufficient attention to the following principles and practices in circuit and equipment drawing, many of which are dealt with in the various British Standard specifications:

(a) Laying out a circuit diagram. The fundamental principle, which should be observed, but frequently is not, is that the layout should show the function of the various parts of the circuit and of the circuit overall. It is very useful when a circuit is laid out as a more detailed version of a block schematic. Clarity should not be sacrificed for the sake of neatness or appearance.

(b) Preparing functional circuit diagrams. These are diagrams in which only bare essentials are retained in order to make the fundamental principles very clear. For example, in dealing with thermionic valves, the heaters, screens and suppressor grids, and power circuit decoupling arrangements would all be omitted.

(c) Representing circuit elements correctly and unambiguously and distinguishing between crossings and junctions in the wiring using the correct British Standards symbols where they exist.

(d) Preparing wiring schedules. Often diagrams are made where a tabular presentation would be preferable.

(e) Drawing component layouts on chassis and printed circuit panels.

(f) Sketching special fittings, microwave components, etc.

† This tradition is not only found in text-books, but also in the British Standard 308 : 1953. Engineering Drawing Practice.

It is the Committee's view that there is a need for instruction in electronic drawing, based on the types of work set out above, in academic courses at all levels. At present only traditional engineering drawing instruction is given at professional level.

A provisional syllabus has been prepared (Appendix 2) and it is considered that an adequate time in professional courses for the electronic drawing alone might be five to ten periods, each of two- to three-hours duration, combining lectures and drawing practice; these could be given about half-way through the courses. A 'pass' mark should be required on class performance.

For students obtaining professional qualification by means of the Institution's own examination, one possibility is that a two-hour examination paper in electronic drawing (which would necessarily involve design) should be introduced at one of the intermediate stages of the Graduateship examination.

One difficulty in implementing these proposals may well prove to be the absence of suitable textbooks, as the only ones written in English are, in fact, of American origin and do not describe British practice; moreover, their price puts them beyond the reach of most students in this country. It is therefore recommended that steps should be taken to see that a suitable small textbook is produced. Until courses and textbooks are available, it is clearly impracticable for the Institution to include a compulsory paper in this subject in its Graduateship examination, or to insist on its being included in exempting qualifications. It might, however, be possible to give encouragement to the subject by including a compulsory question involving skill in communicating information, including a drawing, in one of the electronics papers.

Bibliography

1. N. M. Raskhodoff, "Electronic Drafting Handbook" (The Macmillan Co., New York, 1961).
2. G. Shiers, "Electronic Drafting" (Prentice Hall, Englefield Cliffs, N.J., 1962).
3. L. H. Bainbridge-Bell, "The standardization of symbols and the arrangement of electronic circuit diagrams", *J. Brit.I.R.E.*, **13**, pp. 339–53, June 1953.
4. C. E. Williams, "The utility factor in circuit diagrams", *Proc. Instn Radio Engrs, Australia*, **13**, pp. 345–49, September 1952.
5. L. H. Bainbridge-Bell, "Improving circuit diagrams", *Electronic Engineer*, **20**, pp. 175–77, June 1948.
6. 54 IRE 21 S1. I.R.E. Standards on Graphical Symbols for Electrical Diagrams, 1954. Reprinted in *Proc. Inst. Radio Engrs*, **42**, pp. 965–1020, June 1954.
7. Interservices Graphical Symbols for use in Telecommunications Engineering. Army Reference E.M.E.R. Tels. A301.
8. BS 530: 1948. Graphical Symbols for Telecommunications.
   BS 530: Supplement No. 1: 1950. Component References.
   BS 530: Supplement No. 2: 1951. Graphical Symbols used in Waveguide Technique.
   BS 530: Supplement No. 3: 1963. Additional Symbols for Electronic Tubes and Valves, including gas switches.
   BS 530: Supplement No. 4: 1956. Miscellaneous Recommendations and Symbols, including Transistors.
   BS 530: Supplement No. 5: 1962. Logic and Functional Symbols (mainly for use in automatic switching, computers and similar applications).
   BS 530: Supplement No. 6: 1959. Services' Preferences and additional symbols.
   BS 530: Supplement No. 7: 1962. Symbols for Semiconductor Devices.
9. BS 3238: Part 1: 1960. Graphical Symbols for Components of Servomechanisms.
10. BS 108: 1951. Graphical Symbols for General Electrical Purposes (power and lighting).
11. BS 1646: 1950. Graphical Symbols for Instrumentation.
12. IEC 117–1. Recommended Graphical Symbols.
13. BS 308: 1953. Engineering Drawing Practice.

Appendix 1: Suggested Revised Syllabus in Engineering Drawing for the Engineering Institutions Part I Examination

The principles of Engineering Drawing. First-angle and third-angle projection.

The preparation of working drawings and freehand sketches in orthographic and pictorial projection of common engineering components and simple assemblies.

The properties, with the associated constructions, of the common-plane geometrical figures including conic sections and other loci. The projection of lines of inclined and oblique planes. Auxiliary views and sections. Intersections of planes and interpenetration of solids. The development of surfaces.

The examination is intended to test the candidate's ability to make well-proportioned sketches of engineering components; to read, interpret, and construct an engineering drawing according to standard practice (B.S. 308: 1953); and to apply the principles of practical geometry.

Appendix 2: Electronic Drawing

A Provisional Syllabus for a Short Course of Instruction and Practice in Drawing for Electronic Engineers

The suggested syllabus which follows is based on the assumptions that:

(a) Only about 15–30 hours of instruction can be provided or justified.

(b) The students are being educated as professional engineers and not as technicians.

(c) Some previous acquaintance with geometrical drawing has been obtained, for example, as in the proposed Engineering Institutions Part I syllabus and in most University first-year courses in engineering. In spite of the recommendations in the report these will probably have been of the traditional type.

(d) Some knowledge of electrical and electronic circuits and apparatus has been obtained, say to the level of second-year degree or H.N.D. courses.

Introductory matters, for example the purpose of electronic drawing, clarity, function, absence of ambiguity, standard practices, recommendations in BS 308 : 1953, and so on.

Circuit diagrams, standard symbols BS 530, indication of the method of operation, labelling of components and indicating their values on diagrams in a separate table. Crossings and junctions, simplified functional circuit diagrams, detached-contact technique, etc. Special conventions for thermionic valves, transistors, variable and other components, currents and voltages, e.m.f. and current sources, and so on.

Application of functional flow principles to block schematics and circuit diagrams.

Layout diagrams, chassis, assembly and wiring diagrams. Wiring and cabling schedules, use of colour, printed circuits.

It is thought that the inclusion of some discussion on the design and construction of electronic equipment in relation to specifications and the proposed environment and manner of use would be very advantageous. For this purpose the "Guide to the Design and Construction of Electronic Equipment" published by the Electronic Engineering Association is very helpful.

Approved for publication by the Council on 26th May 1964. (Report No. 28.)

STANDARD FREQUENCY TRANSMISSIONS

(Communication from the National Physical Laboratory)

Deviations, in parts in 10^{10}, from nominal frequency for **August 1964**

August 1964	GBR 16kc/s 24-hour mean centred on 0300 U.T.	MSF 60 kc/s 1430–1530 U.T.	Droitwich 200 kc/s 1000–1100 U.T.	August 1964	GBR 16 kc/s 24-hour mean centred on 0300 U.T.	MSF 60 kc/s 1430–1530 U.T.	Droitwich 200 kc/s 1000–1100 U.T.
1	− 151·3	− 150·8	+ 4	17	− 151·5	− 150·4	− 18
2	− 151·8	− 151·2	− 10	18	− 150·0	− 150·2	− 18
3	− 151·3	− 151·7	− 10	19	− 149·6	− 150·0	− 16
4	− 151·3	—	− 13	20	− 149·4	− 150·7	− 12
5	− 149·8	− 149·7	− 11	21	− 149·8	− 151·2	− 13
6	− 150·0	− 150·3	− 27	22	− 150·7	− 150·6	—
7	− 149·5	− 149·4	− 25	23	− 150·8	− 151·0	—
8	− 149·3	− 149·9	− 24	24	—	− 151·9	− 12
9	− 148·5	− 148·6	− 23	25	− 150·4	− 151·2	− 14
10	− 148·7	− 149·7	− 22	26	—	− 150·0	− 13
11	− 149·6	− 149·6	− 21	27	− 149·2	− 150·4	− 17
12	− 149·9	—	− 21	28	− 149·5	− 149·7	− 10
13	− 150·4	− 149·6	− 21	29	− 149·9	—	− 9
14	− 150·0	− 151·4	− 19	30	− 149·2	− 149·7	− 10
15	− 151·0	—	− 19	31	− 149·4	− 150·0	− 7
16	—	− 151·7	− 19				

Nominal frequency corresponds to a value of 9 192 631 770 c/s for the caesium F,m (4,0)–F,m (3,0) transition at zero field.

Note: the phase of the GBR/MSF time signals will be retarded by 95·5 milliseconds at 0000 UT on 1st September, 1964.

U.D.C. 621.327.6

A Study of Coding with Multiple-anode Glow Discharge Tubes

By

H. E. SEIFERT, Dipl. Phys.†

Reprinted from the Proceedings of the Symposium on "Cold Cathode Tubes and their Applications" held in Cambridge, 16th–19th March 1964.

Summary: For optical indication of numerals or symbols glow discharge tubes or fluorescence devices are often utilized. These have a multitude of luminous elements and the symbol is obtained by paralleling some of these elements. A new tube having several equal anodes has been developed to switch such indicators in a simple way. The number of the anodes is the same as the number of the lighting elements in the optical device. Utilizing the GT-mechanism, it is possible to operate the tube with a signal voltage of about 5 V only. Circuits for controlling both a glow discharge indicator tube and a fluorescence plate, are shown.

1. Introduction

The term 'coding' has different meanings, depending on the special technique of which we are speaking. In this case 'coding' is the transformation of single signals, which enter the coder on different lines at different times, into a number of signals, which leave the coder on different lines simultaneously. The number and the composition of the output lines will depend on the specific input line.

A simple example will make this clear.

There are devices on the market for optical read-out from electronic measuring instruments, which contain a number of electrodes, which may be switched on to light up. By selecting and combining the single electrodes, the desired symbol or number is obtained.

In the present state of the technique, optical indication may be obtained by a glow discharge tube and selection and combination is made efficiently by a transistorized matrix. But there are also two disadvantages: matrix and indicator are separate units and normally it is necessary to insert an amplifier in between.

This paper describes the development of a new device, combining the coder and the glow-indicator which may be driven by the output signal of a 6-V transistor circuit.

2. Mechanism of Operation

The function of cold cathode tubes may be compared with electro-mechanical relays and therefore we shall first consider an example utilizing relays. Figure 1(a) shows a glow indicator tube (I), with four cathodes, which are arranged in the form of a cross. The cathodes are electrically insulated from each other and each one is connected to the contacts of relay RL1. In addition, two of the cathodes are connected to the contacts of another relay RL2. When RL1 is energized, the glow cathodes will show a light in the form of a cross, by closing RL2 only, the light will form a horizontal line.

In Fig. 1(b) the electromechanical relays are replaced by cold cathode switching tubes. This gives rise to some fundamental problems, which will be discussed in relation to Fig. 2.

Switching and indicator tubes form two glow discharges, connected in series. In the indicator tube there are several glow discharges, with different cathodes and a common anode. In the switching tube there are also several glow discharges, in this case with several anodes and a common cathode.

When two glow discharges are connected in series, it can be expected that the potential of the middle electrodes (these are the anodes of the switching and the cathodes of the indicator tube) is not well defined. Measurements have shown, that in all cases a leakage current of about 10^{-10} A exists. This is sufficient to fix the potential of the 'floating' electrodes. The exact value of this potential is determined by the supply voltage (V_0), applied across the cathode of the switching tube and the anode of the indicator tube.

If the supply voltage is lower than twice the striking voltage of the indicator tube, the potential of the 'floating' electrodes is half of the supply voltage:

$$(V_A)_S = (V_C)_I = \tfrac{1}{2} V_0$$

where $$V_0 \leqslant 2\,(V_{St})_I$$

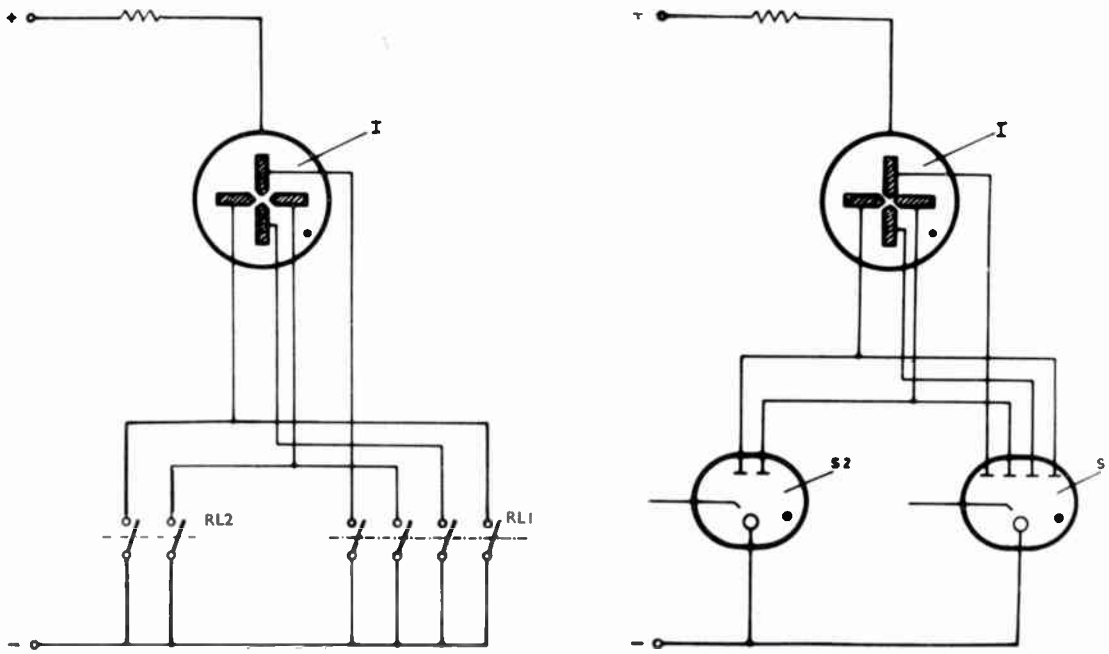
If the supply voltage is higher than twice the striking

† Formerly with Cerberus A.G., Männedorf, Switzerland.

voltage of the indicator tube, a glow current in this tube can be observed, having the same value as the leakage current (10^{-10} A), but producing a constant voltage drop, corresponding to the sustaining voltage (V_{Su}). The potential of the 'floating' electrodes will then correspond to the supply voltage less the sustaining voltage of the indicator tube:

$$(V_A)_S = (V_C)_I = V_0 - (V_{Su})_I$$

where $$V_0 > 2\,(V_{St})_I$$



(a) Switching with relays.

(b) Switching with cold cathode tubes.

Fig. 1. Glow indicator with a switching tube.

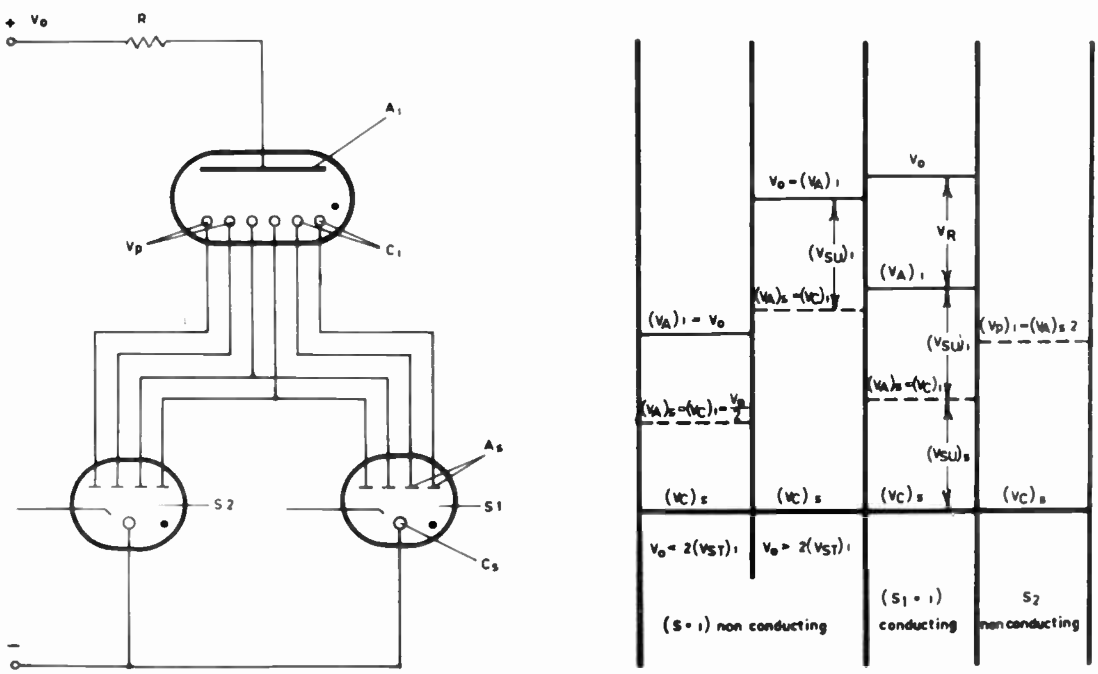


Fig. 2. Switching and indicator tubes in series.

The constant value of $(V_{Su})_I$ means that all supply voltage fluctuations are directly effective at the anode of the switching tube:

$$V_0 \pm \Delta V_0 = [(V_A)_S \pm \Delta V_0] + \text{constant}.$$

Generally speaking, the ignition properties of two glow discharge tubes in practical circuits will be determined by the characteristics of the tube having the higher striking voltage.

When switching only a part of the glow discharges in the same glow tube, the other cathodes and the anodes to which they are connected will adopt a probe potential, situated between the voltage of the glow cathodes and the anode:

$$(V_C)_I < (V_{CP})_I = (V_{AP})_S < (V_A)_I$$

The anodes of non-conducting switching tubes may have two different voltages. If they are in connection with glowing cathodes of the indicator tube, their voltage corresponds to the sustaining voltage of the switching tube. If they are connected to non-glowing cathodes, they will have the probe potential, as defined above. From this it is possible to deduce two conditions for proper 'coding':

For good insulation between the anodes it is necessary that the striking voltage between these anodes is higher than the striking voltage of the indicator tube which in turn must be higher than the difference between the probe voltage of a floating cathode and the sustaining voltage of the switching tube:

$$(V_{St-AA})_S > (V_S)_I > (V_{CP})_I - (V_S)_S$$

For the non-self-firing condition the striking voltage of the switching tube must be higher than the sum of the two sustaining voltages of the switching and indicator tube:

$$(V_{St})_S > (V_S)_S + (V_S)_I$$

3. Preliminary Experiments

To establish experimentally the considerations set out above, a demonstration unit has been assembled for the optical indication of digits. There are two basic factors:

(*a*) It needs as many switching tubes as there are digits to be presented.

(*b*) Every switching tube needs as many anodes as it needs illuminated elements in the indicator to compose the corresponding digit.

In the particular case under consideration there are ten switching tubes with a maximum of seven anodes.

The driver signal was derived from a commercial transistor counter and decoder, having ten output connections each giving a signal of 6 V.

When utilizing switching tubes with grid control, as described in a previous paper,† it is possible to control these tubes directly with the signal coming from the transistors.

To determine the supply voltage it is necessary to know the ignition characteristic of the switching tube and the maintaining voltage of the indicator tube, as described in the previous section. Figure 3 shows the characteristics of the tubes used. With these data it is easy to determine the absolute minimum and the permissible maximum voltage as a function of the grid voltage. The most suitable supply voltage is between these two extremes, and in the present example is at about 300 V r.m.s.

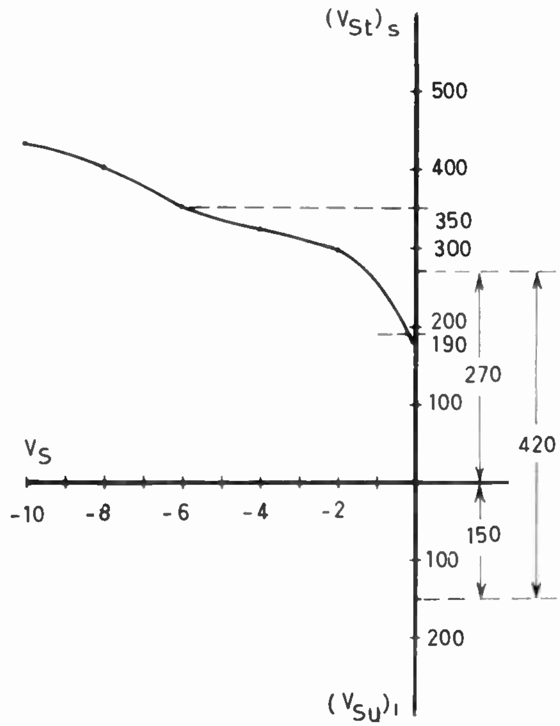


Fig. 3. Ignition characteristics of a typical switching tube.

After the supply voltage has been determined, it is easy to build the demonstration unit. Figure 4 gives the circuit of a proved arrangement.

The cathodes of the switching tubes are grounded. The current for each auxiliary cathode is about 250 μA. Each grid is connected by a 100-kΩ resistor to one output of the transistorized decoder.

The manufacturers of glow-indicator tubes recommend a current-limiting resistor for each glow cathode. But it has been found that one resistor only, connected to the anode of the tube, is sufficient. The increase of the sustaining voltage with current

† H. E. Seifert, "Das Glimmthyratron", *Elektronik*, **11**, p. 237, August 1962.

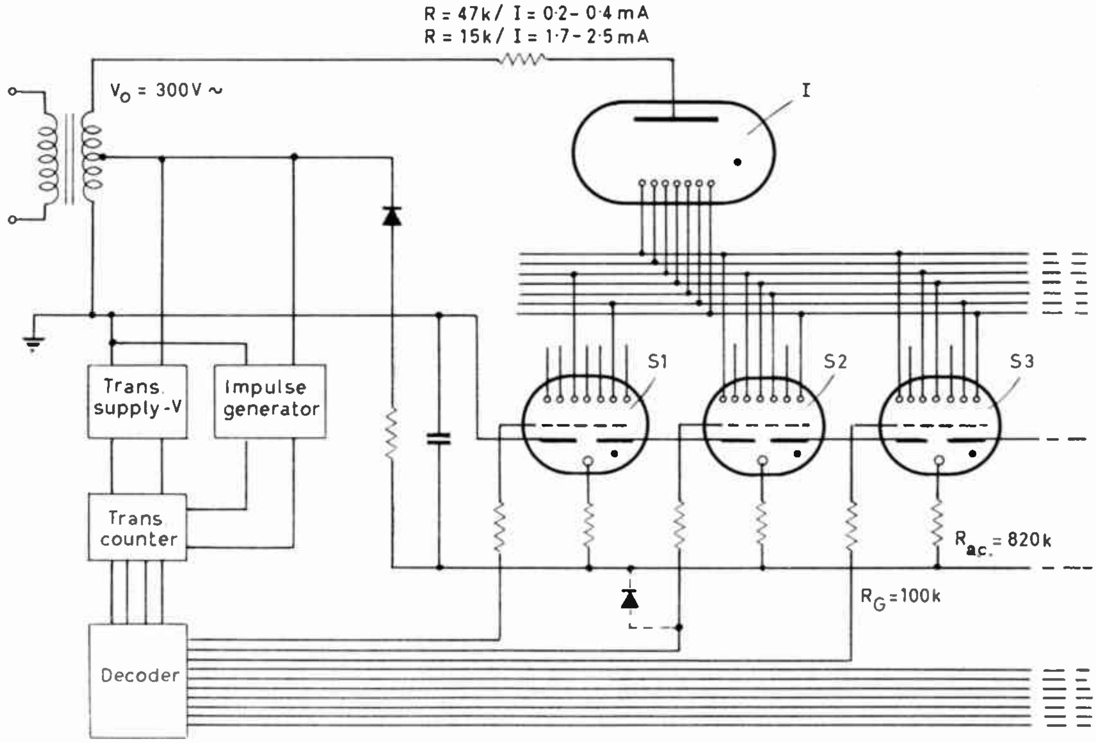


Fig. 4. Circuit diagram of the multiple-anode switching system.

density acts as a self-regulator of the total current in the indicator tube.

It is important that the grid of a switching tube acts as a probe in the main-discharge, as soon as this is fired. The grid will carry a relatively high potential of about 100 V. Then the grid resistor, together with the internal resistance of the decoder, forms a voltage-divider. Depending on the ratio of the two resistances, it may occur that the output line of the decoder will reach a positive voltage, which may disturb the signals of all other lines. In this case, it is necessary to connect a diode across the grid line and earth, in order to short-circuit every positive signal. This is shown by a broken line in Fig. 4.

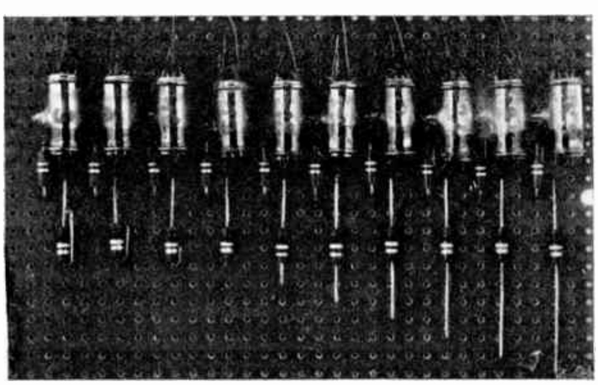


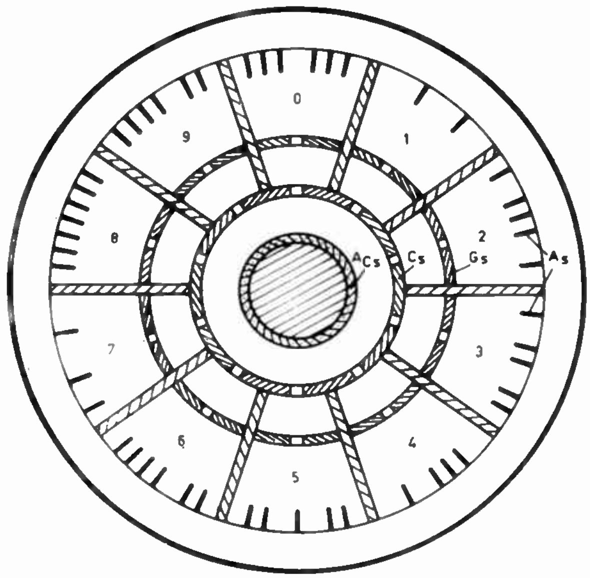
Fig. 5. Ten switching tubes mounted on printed circuit board.

The part of the circuit in Fig. 4, containing the ten switching tubes, can be easily mounted upon a printed circuit board, as shown in Fig. 5.

If, instead of employing a glow indicator, a luminescence panel is used, the circuit does not change substantially. The most important condition to be fulfilled in this case is also that the leakage current through the panel is about 10^{-11} A and always sufficient to polarize the anodes of the switching tube to assure reliable operation.

The fluorescence panel is energized by a current of about 400 c/s and of 20 μA per element. Frequency and low current do not disturb the switching tube, but it is important to have well-shielded connections between panel and switching tubes, all capacitive couplings of the conductors, not only between each other but also against ground gives an unwanted glow of non-insert elements.

The fluorescence panel develops its full brightness with alternating current. But the switching tube, like all glow discharge tubes acts as a half-wave rectifier. To overcome the resulting loss of light, the insertion



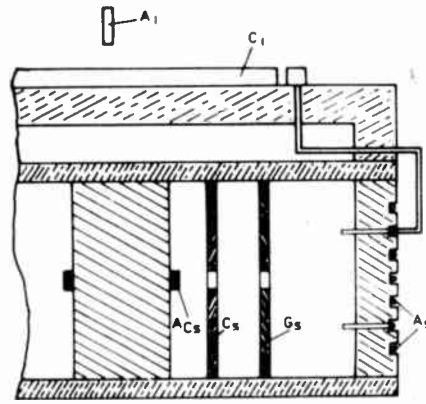


Fig. 6. Construction of the multi-channel switching tube.

of a full-wave rectifier between the 400 c/s generator and the panel is recommended.

4. Multi-channel Switching in a Single Tube

Having seen that the unit is operating satisfactorily, the next and logical step was the unification of the ten switching systems in a single tube. For such a construction it is possible to use a common cathode and a common auxiliary cathode. Only the ten discharge channels between cathode and anodes have to be well separated. A possible and practically proved construction is shown in Fig. 6.

Two disks of insulating material are separated and mechanically linked by an insulating axle. In the middle of this axle is a metallic electrode, the auxiliary cathode, surrounded by a cylindrical main cathode. The latter has ten holes for the passage of electrons. The space outside the main-cathode is subdivided into ten sectors by insulating sheets. In each sector there is one control grid. The whole construction is enclosed by an insulating cylinder. On its outer side there are seven metallic rings, which serve as a collector for short anodes, entering the discharge sectors by small perforations in the cylinder. Every collector ring is connected to a glow cathode in the indicator tube.

The possibility for connecting one, some or all collector rings with the discharge in one, some or all sectors allows the glow discharge to be used as a switch for predetermined electrodes of the indicator.

5. Switching and Indicating Tube Unification

The next and last step in this development is evident: it will be the unification of the switching tube and indicator tube, but there are considerable problems.

The switching tube should have a low sustaining voltage, a large difference between sustaining and striking voltage and sufficient cathode sputtering for an active cathode and auxiliary cathode.

The indicator tube should have a high sustaining voltage and a small difference between sustaining and striking voltage, little or no cathode sputtering and a bright, luminous glow light.

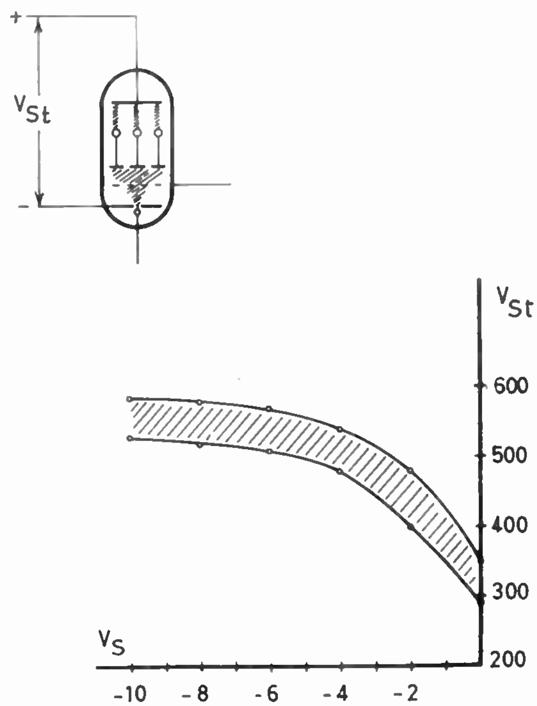


Fig. 7. Range of anode striking voltages of the ten switching circuits.

Regardless of these difficulties it was possible to realize some laboratory samples of this kind of tube, permitting the expected electrical characteristics to be verified giving an idea of the feasible size.

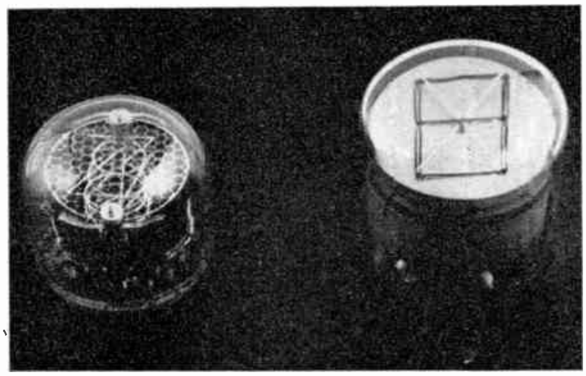


Fig. 8. Final laboratory model (right) compared with the 'Nixie' tube.

Figure 7 shows the range of the anode striking voltage of the ten switching systems, mounted in the described manner with the indicator in the same bulb. The range is relatively small and corresponds quantitatively with the characteristic for separated tubes shown in Fig. 3.

Figure 8 shows a laboratory sample, comparing its size with the well known 'Nixie' tube.

Industrial production or application for such a tube are still far off, but this preliminary study has shown that there is a certain chance of realizing the tube in practical form.

Manuscript received by the Institution on 11th March 1964. (Paper No. 930.)

DISCUSSION

Mr. G. O. Crowther: How do you ensure that current is fed uniformly to the various segments of the number tube. It seems that when displaying the numeral 8, for instance, there is a possibility that the discharge may fail to cover one of the segments due to variations between the anode-to-cathode breakdown voltages of the number-tube section. If one section fails to ignite an erroneous read-out may result.

Mr. H. E. Seifert (*in reply*): If one segment of the number tube is not covered, the corresponding anode in the switching tube will accept a probe potential in the discharge occurring between the main cathode and the other anodes. This probe potential will be less than the maintaining voltage of the other anodes and therefore the corresponding segment will have a greater voltage difference to the main anode than the covered segments. For this reason, the segments will cover correctly as long as the difference between the maintaining voltages in the indicator tube does not exceed the anode fall in the switching tube. This may be assured by an accurate selection of the material of the electrodes and the filling gas.

U.D.C. 621.387.332

A New Gas-filled Trigger Tube for Logical Functions

By

G. KERR, B.Sc.(Eng.)†

Reprinted from the Proceedings of the Symposium on "Cold Cathode Tubes and their Applications" held in Cambridge, 16th–19th March 1964.

Summary: The tube has been designed with a view to large-scale mechanized production with the least possible handling and this has affected the mechanical form and electrical characteristics. A study of its basic operation has led to a new way of presenting its characteristics, enabling the equipment designer to calculate more exactly the requirements for ignition and extinction. In particular, the build-up of current between trigger and cathode after the application of an overvoltage has been studied, and from the results are derived the requirements for pulse ignition of the tube and the effect of different pulse shapes. The requirements for extinction of the anode-cathode discharge either by an externally produced pulse or in a self-extinguishing circuit are also discussed.

1. Introduction

The increasing use over the past few years of small trigger tubes for counting and other logical applications has made evident the need for further development in order to satisfy the requirements of modern equipment and new circuit techniques. When such a development was being considered, the following points were proposed by the applications and manufacturing groups:

1. Higher frequencies were required than is possible with most of the present simple trigger tubes.
2. Voltage levels should be as low as possible.
3. A high level of stability over a long life period was essential.
4. The physical size should be small.
5. A simple geometry with as few components as possible was desirable and the assembly of the components by mechanical means should be realizable.
6. During the development period a detailed study would be made of the characteristics of the tube in order to be able to present as fully and as logically as possible the characteristics required for accurate dimensioning of circuits.

Most of the foregoing electrical and physical requirements are of course interdependent but will as far as possible be treated separately in the text.

2. Considerations of Frequency

In a bistable device such as a trigger tube the maximum frequency of operation is determined by the sum of the times required to switch the device from the 'off' state to the 'on' state and then to return to the 'off' state. In a trigger tube the switch-on or ionization time is generally short compared with the switch-off or deionization time. Furthermore, the switch-on time can be reduced by careful design of circuits and voltage levels. Thus in the design of a fast trigger tube, the most important factor to be considered is the deionization time, which may be roughly defined as the time required after extinction of the tube before the anode supply voltage may be re-applied. The process involved and circuit requirements are discussed in that part of the paper dealing with tube characteristics.

2.1. *Effect of Gas-filling*

In most high-speed gas-filled tubes, hydrogen, or a mixture of hydrogen with one of the inert gases is used since due to the quenching of the metastables by the hydrogen, very fast deionization can be realized. Using this technique, decade ring-counter tubes for frequencies up to 1 Mc/s have been developed. The use of hydrogen was not however acceptable in this case since it conflicted with two of the other requirements laid down for the development. Firstly, for any fixed geometry the breakdown voltage across a gap in hydrogen or hydrogen mixtures is high compared with most other gases; secondly, the presence of hydrogen tends to equalize the work functions of most common metals, making it difficult to sputter only the cathode and prepare a clean and reproducible cathode surface which, although possible, would entail more complicated processing during manufacture. The next gases considered were helium with 5% neon, a mixture which is also widely used in decade ring counters, and pure argon, a gas used in many coated-cathode tubes. Both of these were tried in tubes and no significant difference in deionization time was noticed, but a pronounced hollow cathode

† N. V. Philips' Gloeilampenfabrieken, Eindhoven, Holland.

effect in the helium-neon caused fast sputtering in the area of the cathode slot. Argon was thus chosen, having the further advantages of low voltage levels and insensitivity to contamination.

2.2. *Gas Pressure*

Since the physical size of the tube was required to be kept as small as possible, a relatively high pressure had to be used in order to obtain a satisfactory anode-cathode breakdown voltage. This had the further advantage that deionization due to recombination would be favourably influenced since the probability of recombination is theoretically proportional to the square of the pressure in the tube. Diffusion to the walls at higher pressures becomes less important and later measurements on the tube in fact confirmed that deionization depended almost completely on recombination in the gas volume.

3. Stability of Characteristics

Molybdenum is generally accepted in the cold cathode field as the cathode surface giving the most stable and reproducible characteristics and the author's experience on other small trigger tubes has amply confirmed this. The relatively low rate of sputtering also gives an important advantage over nickel, its closest competitor, since much longer life times can be obtained with a cathode of the same area and thickness.

Since molybdenum is easily contaminated during operation particularly by sputtering of material from the trigger electrode, and to a lesser extent from the priming discharge, it was decided to make all electrodes from the same material. Although not eliminating contamination this would minimize its effects and further would allow easier cleaning of the tube during processing.

4. Physical Size and Geometry

The initial requirement was that the external dimensions of the tube should be as small as possible to facilitate mounting on printed-circuit boards. It was thus decided to use as a bulb glass tubing of 6-mm external diameter and approximately 4-mm internal diameter, since this standard size was already in use for several other small tubes. The pinch-sealing technique used for the other tubes was not adopted since it imposes severe geometrical restrictions, and a powdered-glass base was developed to allow more freedom in the placing of electrodes.

The Paschen curve for pure argon, shown in Fig. 1 determined the operating points for the trigger-cathode and anode-cathode gaps, and a gas pressure of 100 mm was chosen giving spacings of 0·075 mm and 1·2 mm respectively. According to the graph these dimensions would give a trigger-cathode ignition voltage of about 135 V and an anode-cathode ignition voltage of about 400 V. These values apply of course to a parallel plate geometry and it will be shown later in the text that the actual values achieved were lower.

Since the maximum permissible distance between trigger and cathode is only 75 μ, it will be realized that very careful control of tolerances on components and assembly jigs was required in order to prevent short-circuiting. The geometry shown in Fig. 2 was chosen since, with suitable control of slot size and material thickness, it is relatively easy to ensure that

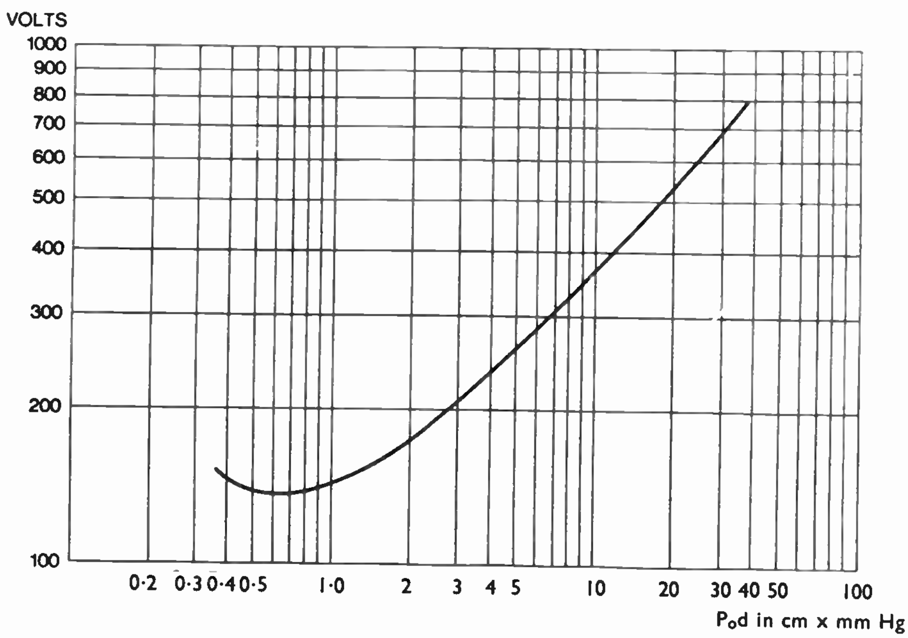


Fig. 1. Breakdown voltage as a function of pressure × distance for flat parallel molybdenum electrodes in pure argon.

the distance between the cathode and one edge of the trigger will always be equal to or less than the required value.

To reduce the number of different components to a minimum, it was decided to use the same geometry for the anode-primer assembly as for the cathode-trigger thus giving a tube which, before the final electrical processing, is symmetrical. The cut-away view of the tube in Fig. 3 shows the relative positions of the four electrodes.

To avoid the effects experienced in previous trigger tubes with a metallic sputtered layer on the glass wall, the sizes of the various electrodes were chosen to ensure that any charge on the wall layer would cause breakdown preferentially to the trigger or primer

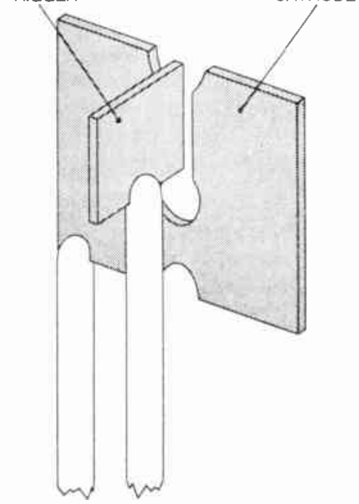


Fig. 2. The trigger-cathode geometry.

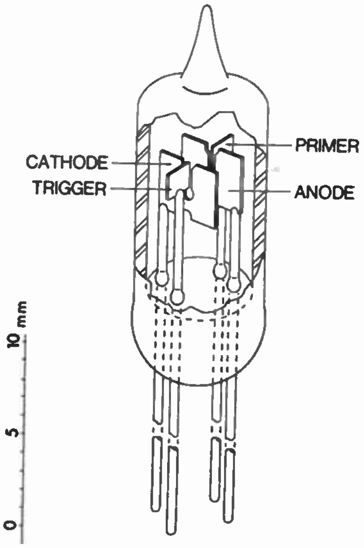


Fig. 3. A cut-away view of the tube.

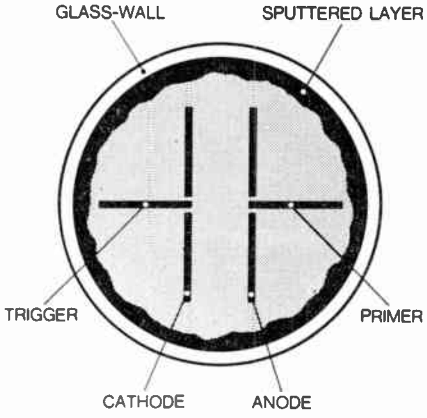


Fig. 4. A cross-section through the electrode assembly, showing the position of the various electrodes with respect to the sputtered layer.

rather than to an edge of the cathode. Such a discharge is remote from and effectively shielded from the main discharge gap and will not cause any adverse effects, while a discharge to an edge of the cathode would almost certainly ensure breakdown between anode and cathode. This is made clear by the cross-section in Fig. 4.

In order that mechanical handling of the components should be as simple as possible, it was decided to use a relatively large cathode which would not be fully covered by the discharge at normal tube currents. This has the advantage of giving satisfactory life at the high peak currents often used in self-extinguishing circuits.

For reliable triggering the trigger-cathode discharge should take place on the edge or front of the cathode; that is on the side of the cathode facing the anode. Any discharge occurring on the other side of the cathode will be effectively shielded from the main anode-cathode gap. To prevent this happening, the back of the cathode is plated with nickel, the higher work function giving satisfactory preference to a discharge starting on the front or edge of the cathode. The effectiveness of this process can be seen during sputtering of the tube, when even at peak currents of some hundreds of milliamperes, the discharge does not spread to the back of the cathode.

5. Primary Electrical Characteristics

As mentioned in Section 4, the breakdown voltages for the main and auxiliary gaps are lower than those values deduced from the Paschen curve for argon, the actual values being about 340 V between anode and cathode and 125 V between trigger and cathode. Due to the very small gas volume, relatively large pressure differences may occur from tube to tube and this results in a rather large spread in anode to cathode voltage since the Paschen curve at this point is steep. Since the trigger-cathode gap is calculated to operate at or near the minimum of the curve, the spread here is much smaller and measurements on large batches of tubes show a total initial spread not exceeding 10 V.

The breakdown characteristics of a typical tube are shown in Fig. 5. It is clearly seen that the ignition voltage of the trigger-cathode gap is independent of the voltage between anode and cathode. Since the trigger projects slightly in front of the cathode, the breakdown voltage between anode and cathode will obviously be influenced if the trigger is made negative with respect to the cathode.

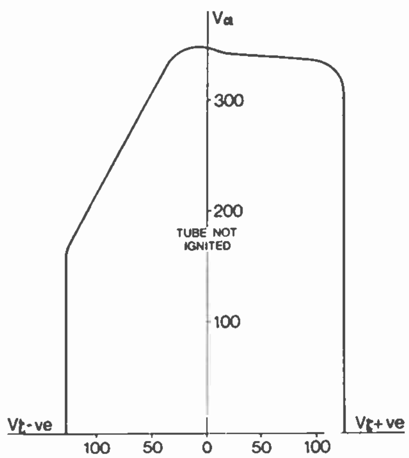


Fig. 5. Breakdown characteristics for a typical tube in the first and second quadrants.

The second quadrant characteristic shows clearly that the reduction in anode hold-off voltage is greater than the negative voltage on the trigger, this being due to the increased field strength over the small trigger area.

The tube is of course recommended for operation in the first quadrant, the second quadrant characteristic being required only for the prevention of ignition if the trigger becomes negative. This may occur in pulse circuits, due to differentiation of a square pulse by the trigger circuit components, and care must be taken in the design of such circuits to avoid spurious breakdown.

The tube is however extremely insensitive when d.c. or level firing is used on the trigger. This can be explained by referring to Fig. 6, which shows a cross-section of the trigger-cathode gap after processing of the tube. This figure shows that although the nickel plating is successful in preventing the main anode-cathode discharge from spreading on to the back of the cathode, a little of the nickel is actually removed from the edges during processing. Breakdown between

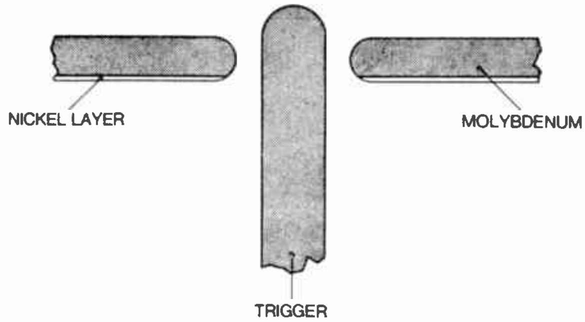


Fig. 6. The trigger-cathode gap after sputtering.

trigger and cathode will normally occur across the shortest gap, provided that the molybdenum has been properly cleaned, but as the current increases, the discharge may spread from this point towards the front or the back of the cathode, i.e. it may or may not affect the space charge between anode and cathode. Thus ignition of the main gap in these circumstances will not be reliable. The firing of the tube is further discussed in Section 6.

Since only one side of the cathode is used for the main discharge, the cathode is effectively covered at a current of about 6 mA. Thus, although the tube will in many applications be used at a current of 2 or 3 mA, up to 6 mA may be drawn before the abnormal glow region is approached. Above this current the rate of sputtering will increase giving a shorter tube life, and the maintaining voltage will go outside the normal limits. The use of higher peak currents is not of course precluded since under dynamic operating conditions the maintaining voltage is normally not so critical and the increase in sputtering during conduction is offset by the lower duty cycle.

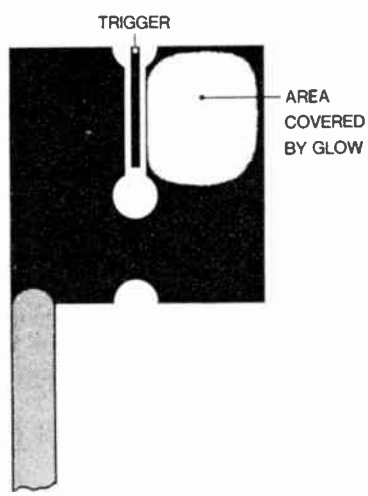


Fig. 7. The cathode discharge at 2 mA.

The difficulties of stable operation of the tube when the cathode is only partly covered by the glow discharge have been overcome by the use of the slot geometry. At low currents, the discharge will cover one side or part of one side of the usable area (see Fig. 7). Thus, although other parts of the cathode will be contaminated by sputtered material during conduction, part of the cathode edge close to the trigger will be covered by the discharge and will remain uncontaminated allowing normal reignition between trigger and cathode. Life testing has shown that the tube may be operated with the cathode only partly covered by the discharge even when a considerable amount of negative trigger current is being drawn, i.e. when the trigger is acting as cathode with respect to the anode and thus is also being sputtered.

6. Pulse Ignition of the Tube

It has already been shown that, due to its construction the tube cannot be triggered reliably by a direct voltage between trigger and cathode. The normal d.c. transfer characteristic as shown in Fig. 8 does however exist, but when the trigger current is allowed to increase slowly the final position of the discharge may vary greatly from tube to tube and large variations in sensitivity are found. For reliable operation the trigger must be pulse fired, enabling the required trigger current to be reached before the discharge is able to spread from the edge of the cathode. The precise requirements for triggering by this method are described in the following sections.

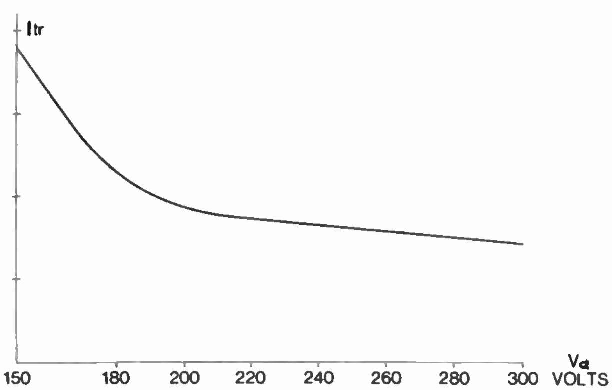


Fig. 8. Typical d.c. transfer characteristic.

6.1. *Growth of Current in a Gap*

When the static ignition voltage between two electrodes in a tube is exceeded, ignition will occur provided that an electron is available within the gap to initiate the process. Since the tube contains a primer, an adequate source of ionization is always present and multiplication begins as soon as the required

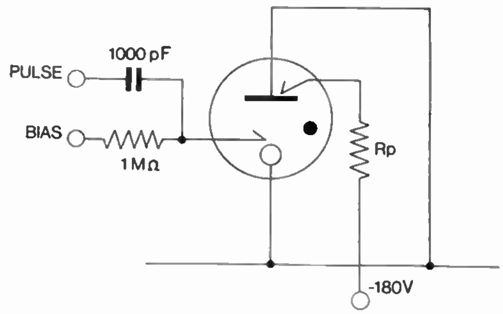


Fig. 9. Measurement circuit for growth of current between trigger and cathode.

voltage is reached. To study the growth of current in the trigger-cathode gap, the circuit of Fig. 9 was used, anode and cathode being connected together to permit normal use of the primer without the possibility of an anode-cathode discharge taking place.

With a fixed bias just less than the static trigger ignition voltage, a rectangular pulse was applied via a capacitor to the trigger, the pulse height being sufficient for the sum of pulse and bias to exceed the static ignition voltage. The measurements were carried out for several values of pulse height, and results are given in Fig. 10, the voltage given for each line being the actual value of overvoltage. The extrapolation below 10^{-4} A is justified by the fact

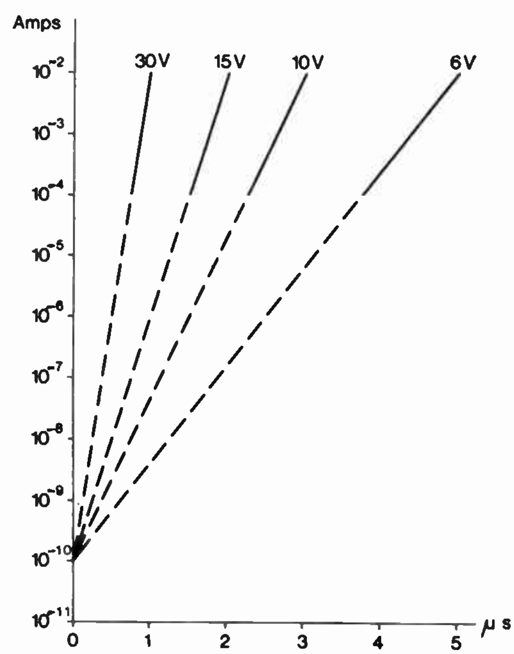


Fig. 10. Increase of current between trigger and cathode—effect of different values of overvoltage.

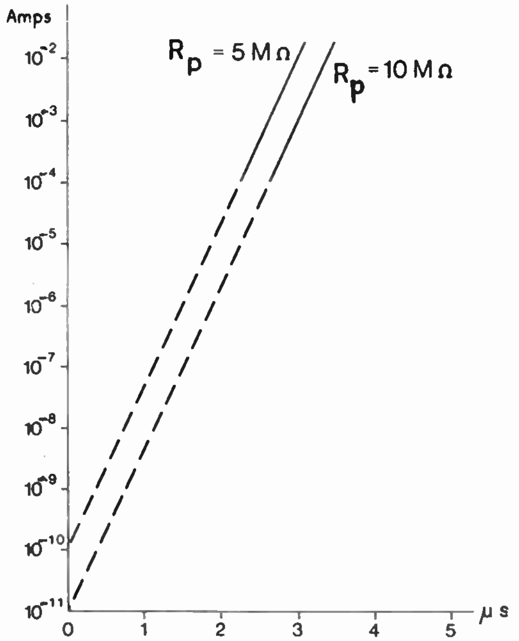


Fig. 11. Increase of current between trigger and cathode—effect of different primer currents.

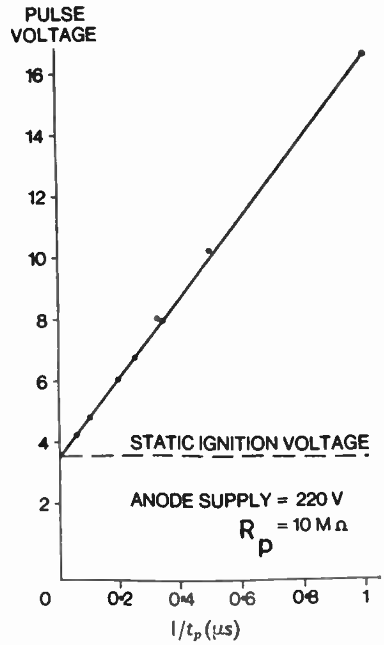


Fig. 12. Relation between pulse height and pulse length required for anode transfer.

that for the four different overvoltages, one starting point is found. This point represents the current, due to ionization from the primer discharge, that flows between trigger and cathode before the static ignition voltage is exceeded. Attempts to measure this statistically have not been successful since, in addition to the current between trigger and cathode, each of these electrodes draws a probe current several orders higher than the inter-electrode current. The effect of different values of primer current is shown in Fig. 11 where the build-up of current with an overvoltage of 10 V is given, the measurement having been done with two different values of primer resistor. It will be noted that the slope of the line remains unchanged, only the starting point or initial current being altered.

From the foregoing measurements the following conclusions may be drawn.

1. In the presence of some initial ionization, multiplication processes begin at the instant of application of an overvoltage.
2. The current increases exponentially with time.
3. The relative rate of increase of current is directly proportional to the overvoltage applied.
4. An increase in primer current increases the initial level of ionization, allowing the multiplication processes to begin at a higher level.

Thus for a given primer current the time required to reach any specified value of trigger current is inversely proportional to the overvoltage.

6.2. *Transfer Requirements*

Although the tube cannot be used reliably with d.c. on the trigger a definite value of trigger current is required to cause ignition of the main gap. When pulse ignition is used the trigger discharge will remain on the edge of the cathode and the required value of trigger current will be reached in a time depending on the pulse overvoltage. Thus for ignition of the tube a fixed product of pulse overvoltage and pulse length is required and this will now be referred to as the ionization product. Figure 12 shows that this product is in fact constant, since the graph of pulse height as a function of the reciprocal of pulse length is a straight line cutting the pulse axis at the static trigger ignition voltage of the tube.

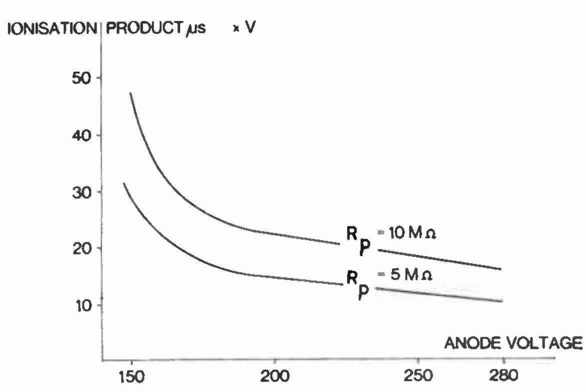


Fig. 13. Ionization product as a function of anode voltage.

The effect of the anode voltage on the required ionization product is shown in Fig. 13 which can be explained by two factors. Firstly, at high anode voltages less trigger current is required to initiate the main discharge and secondly at higher voltages the primer current is also higher, giving a higher initial trigger-cathode current. Figure 13 also shows that a useful gain in sensitivity can be obtained by using a low primer resistor; this factor will often reduce pulse requirements considerably.

6.3. *Calculation of the Required Pulse*

The foregoing measurements on the ionization product have been made using pulses between 1 μs and 20 μs. Below 1 μs it is difficult to obtain high enough pulses in practical circuits and above 20 μs the requirement of the discharge staying on one point of the cathode is no longer satisfied. For non-rectangular pulses the increase in current is of course no longer exponential but is a function of the pulse shape and can be calculated or determined graphically. However for simplicity it can be taken that the tube will always be fired if the required rectangular pulse can be drawn within the boundaries of the non-rectangular pulse.

It is intended that the ionization product shall be published as a characteristic of the tube, being given as a function of primer current and anode voltage. The static ignition voltage of the trigger will be given only for the purpose of determining the maximum bias level and the pulse requirements. It will thus be possible to calculate from the given data the required pulse height for a pulse of a defined width, or vice versa. The total pulse height must include the whole spread of static ignition voltage values for all tubes, since the bias must never exceed the lower limit of static ignition voltage; the pulse-plus bias must exceed the upper limit by the amount required to provide the upper ionization product.

7. Extinction of the Tube

When a gas-filled trigger tube has been ignited, the trigger electrode has little further effect on its operation and in particular cannot be used to extinguish the discharge. In order to accomplish this it is necessary to reduce the voltage across the tube to a value below that value required to maintain the discharge and to keep the voltage reduced until the tube is able to 'recover' by deionization. The two methods commonly used for this purpose are the self-extinguishing circuit and forced extinction by means of negative or positive pulses applied respectively to the anode or the cathode from an external source. The requirements of these two methods will be treated separately.

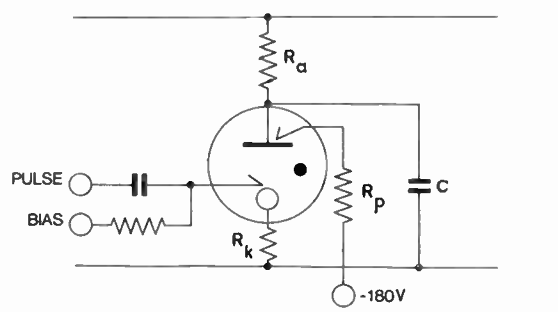


Fig. 14. The self-extinguishing circuit used for measurements.

7.1. *The Self-extinction Characteristics*

All the measurements on self-extinguishing have been made by using the circuit shown in Fig. 14, the primer being connected to a negative supply to ensure that the priming discharge was not extinguished. The tube was ignited by pulsing the trigger at a frequency of 100 c/s.

Although it is not intended to give a detailed description of the operation of this circuit, it is of interest to note that the capacitor C is charged up to or near the supply voltage and, on ignition of the tube, discharges to a voltage lower than the normal maintaining voltage of the tube. Thus the tube is extinguished and will remain so provided that the recharging capacitor does not reach the voltage required for reignition during the recovery period of the tube.

The required product of R_a and C is not constant but is dependent on the value of C, R_k and the supply voltage, since these three factors determine the current through the tube before extinction and the value of the rest voltage, or voltage to which the capacitor discharges. In addition to the above factors and the spread from tube to tube, each tube has a small hysteresis effect which allows it to self-extinguish continuously with a lower value of C than was originally required to start the tube extinguishing.

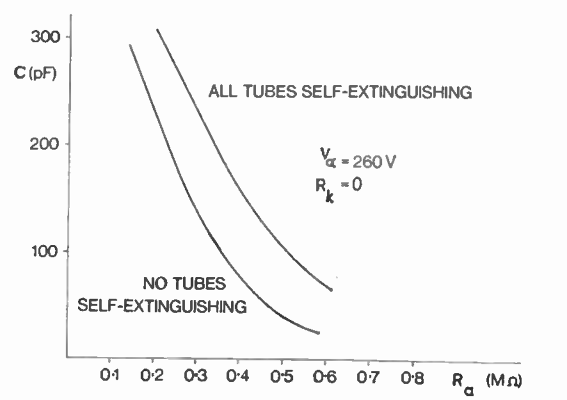


Fig. 15. Values of R_a and C required for self-extinguishing.

For a large population of tubes there is a limiting series of values of R_a and C, above which all tubes will always extinguish, and a separate series below which no tube will ever extinguish. Between these two limits the operation of an individual tube is unpredictable. A sample of the measured results with no cathode resistor and a supply voltage of 260 V is given in Fig. 15. For published data complete families of curves for the two variables (R_k and V_a) must be given in order to predict the operation of any tube with any given circuit values.

7.2. *Forced Extinguishing*

Figure 16 shows the basic circuit used for the forced extinguishing measurements. As in the case of the self-extinguishing circuit, the primer resistor was connected to a negative supply of -150 V. Preliminary measurements showed that the recovery time was not appreciably affected by variations in primer current and the value of the primer resistor was thus kept fixed at 10 MΩ.

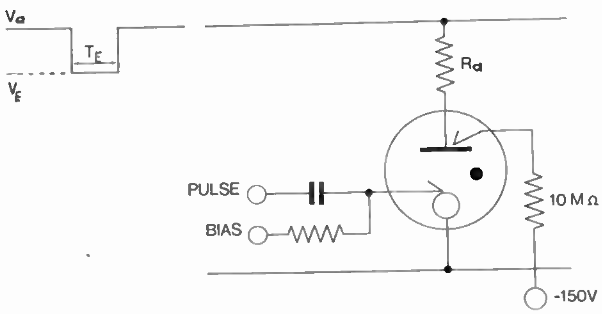


Fig. 16. The measuring circuit for forced extinguishing.

The tube was ignited as in the previous case using a bias voltage and a superimposed pulse on the trigger, and extinguished by a negative pulse on the anode supply voltage, the lower level V_E of this pulse being adjustable. After extinction the tube was reignited, the process being repeated at a frequency of 100 c/s. The anode resistor was adjusted to give the required current at different values of V_a. During measurement the pulse length T_E was increased until the tube had recovered without reignition occurring.

Recovery in a trigger tube is normally accomplished by recombination of ions and electrons in the gas and by recombination at the walls after diffusion. Since the second process is affected by the various potentials in the tube the actual recovery time is normally dependent on the voltage across the anode-cathode gap. Figure 17 shows the recovery time (T_E) as a function of this voltage, V_E, and it is clear that the dependence, if any exists, is very slight. This can be explained by the fact that, in a tube with a relatively high pressure, recombination in the gas proceeds at

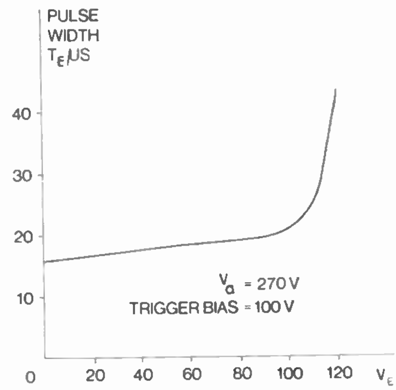


Fig. 17. Recovery time as a function of rest voltage V_E.

such a rate that the diffusion, relatively slow at such pressures, plays little or no part in the deionization, i.e. the slowly diffusing ions are neutralized before arriving at the boundaries of the tube.

Having shown that the rest voltage has little effect on the recovery of the tube, measurements were made to find the influence of the magnitude of the current before extinction, and of the supply voltage.

Since the measurements were made at a low frequency and the time between extinction and reignition was very short, the measured tube could be considered to be running continuously at a steady current. Figure 18 shows the results and, as would be expected, the recovery time increases rapidly with increasing current or with the voltage re-applied after extinction. It will always be advantageous where the highest possible operating speed of a circuit is desired, to limit the current to the lowest possible value.

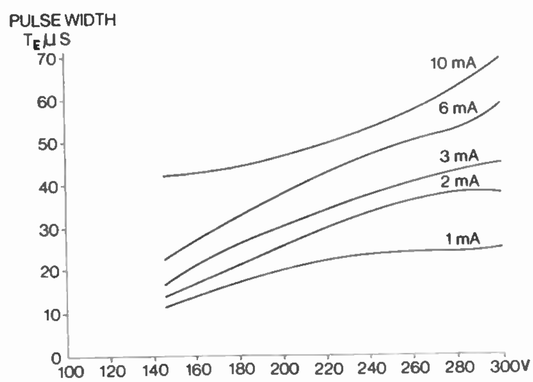


Fig. 18. Recovery time as a function of supply voltage after operating at various currents.

8. Conclusions

Although measurement work on the tube is not yet complete, results show that its basic characteristics are satisfactory for use in logical circuits where pulse ignition is used and at frequencies up to 25 kc/s with normal tolerance components. In addition a large life-test programme is being carried out to prove the reliability of this type of device. One of the most important tasks now to be faced is the presentation of the very considerable data now available in a manner which will enable circuit designers to achieve the best possible results.

9. Acknowledgments

The author wishes to thank the directors of N.V. Philips Gloeilampenfabrieken for permission to publish this paper. He further thanks Messrs. Frouws and Horseling who initiated this development and the members of the measurements and application groups whose work with the tube has produced most of the electrical data and characteristics.

Manuscript received by the Institution on 5th March 1964. (Paper No. 931.)

DISCUSSION

Mr. N. Kitz: I would like the author to give some information about the capacitance effects between the sputtered molybdenum layer of the walls of the tube and the trigger which on similar designs have been known to cause spurious breakdown.

Mr. Kerr (*in reply*)*:* Difficulty in this and other tubes has been caused, not by capacitance effects between the sputtered layer and the trigger, but by a discharge between the sputtered layer and one of the electrodes, usually the cathode. Since the sputtered layer is in fact a floating electrode, it becomes charged to a positive potential in the presence of the priming discharge. This potential is easily altered by capacitive effects in the tube or externally. For example, a positive pulse on the anode will cause a corresponding capacitive rise in the wall potential. Similarly, for a fixed amount of charge on the wall, a decrease in the capacitance between the wall and its surroundings will cause an increase in its potential. In either case a discharge may occur between the positively charged wall and one of the electrodes in the tube, and this may cause spurious ignition. In the tube described in the paper, the electrode geometry is such that a discharge from the wall layer will occur by preference to the trigger at a point where it will not initiate the main discharge. However, difficulty has been experienced due to the fact that the relative trigger-to-wall and cathode-to-wall distances are determined by the internal diameter of the glass tubing used—a factor which is difficult to control in production. Some slight geometrical changes may thus be required to ensure that there is sufficient preference for the discharge to take place to the trigger, even in the presence of the maximum permissible trigger bias.

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. Information on translating services will be found in the Institution publication " Library Services and Technical Information."

MONITORING NAVIGATIONAL AIDS

All radio aids to navigation must be continuously monitored to ensure the radiated signals conform to internationally observed standards. In Australia the primary air-route guidance and precision approach aids are continuous-wave devices and hence the automatic monitoring equipment must effectively measure voltages or currents for long periods without adjustment. Due to the manner in which these aids operate, the sensitivity and accuracy of the monitors are quite often affected by changes in the signals being monitored. The conventional monitoring systems for radio ranges and instrument landing systems, and a new type of system put into service in Australia have been described at the Radio and Electronics Engineering Convention of the Australian I.R.E. in Melbourne. This new system, known locally as 'alarm-point monitoring' has certain basic advantages which make it extremely useful. Problems encountered in the design and adjustment of monitoring units have been considered together with possible solutions in some cases. One of the major difficulties, appears to be the maintenance of air–ground relationships.

"Automatic monitoring of continuous-wave navigational aids", J. J. Richards and H. B. O'Keeffe. *Proceedings of the Institution of Radio and Electronics Engineers Australia*, **25**, No. 3, p. 157, March 1964.

DISTORTION CORRECTION USING RLC NETWORKS

The conventional method of correlating signal distortion uses RLC networks for offsetting the frequency response of amplitude and phase of the communication channel. This type of compensation implies that the transmitted signals are either sine and cosine functions or can be represented by them. However, signals may be represented just as well by any other orthogonal system of functions. Each representation leads to its own theory of compensating circuitry. For the transmission of telegraph signals it is advantageous to forgo representation and to equalize the communication channel only for the signals actually transmitted, e.g. only for rectangular pulses. Thus distortion, for which a classical approach to correction involving a complicated frequency response of amplitude and phase would be necessary, can be compensated.

"On the correction of signal distortion using RLC networks" H. Harmuth. *Archiv der Elektrischen Übertragung*, **18**, No. 3, p. 174, March 1964.

IMPLOSION-PROOF PICTURE TUBE

With conventional picture tubes the chance of implosion although extremely remote, nevertheless exists, and for this reason a protective window is usually fitted in front of the tube. Dutch engineers have described investigations aimed at preventing implosion. The possibility of implosion is attributable to the atmospheric pressure producing tensile stresses around the rim of the bulb, the region between face-plate and cone. The object was achieved by applying a reinforcement band around this region. This band, normally free from stresses, limits the cracks once they have formed, inhibits expansion of the bulb edge, and thus preserves cohesion in the bulb in the event of the glass fracturing. In a version that has now been put into mass production, the reinforcement consists of a metal band and polyester resin. Tests show that the method is entirely effective, the principal advantages being: the protective window is dispensed with, assembly in the cabinet is simpler, the weight and dimensions of the television set are reduced, and new cabinet designs are made possible.

"An implosion-proof picture tube for television", F. de Boer, P. Cirkel, W. F. Nienhuis and C. J. W. Panis. *Philips Technical Review*, **25**, No. 4, p. 81, 24th March 1964. (In English.)

TELEVISION BAND COMPRESSION

A new method of television band compression, 'multi-mode interpolation', has been developed by Japanese engineers. The experimental results of the computer simulation by this method show that the information rate can be reduced to 2·7 bits per picture element, and the signal can be transmitted through a digital channel with a channel capacity less than one half of that of the conventional p.c.m., for which 6 bits per picture element are required for the faithful transmission of the input picture. Since the television signal is to be observed as a picture on a kinescope, some distortion of the signal waveform, imperceptible to the human eye, can be taken advantage of in this method of television band compression. The input video signal is sampled with a Nyquist interval, and every fourth sample is selected as a main pulse. The main pulses are transmitted faithfully, while the samples between two adjacent main pulses are encoded by using the values of these main pulses.

"Television band compression by multi-mode interpolation", K. Fukushima and H. Ando. *Abstracts from the Journal of the Institute of Electrical Communication Engineers of Japan*, **47**, No. 1, p. 9, January 1964. (In English.)