The Journal of the

BRITISH INSTITUTION OF RADIO ENGINEERS

FOUNDED 1925 INCORPORATED 1932

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

VOLUME 21 APRIL 1961 NUMBER 4

# TELEVISION STANDARDS

EVER SINCE the nineteen-thirties the problem of the future of British television has been cropping up at intervals, and yet, with the exception of a brief experimental period of 30-line low definition television, and an even briefer 240-line transmission, the 405-line Marconi EMI system inaugurated by the B.B.C. in 1936 has so far prevailed. It says much for the technical foresight of A. D. Blumlein and others that this system has been in existence for so many years, and has provided an acceptable picture in so many homes without having been superseded.

One may go further, and say that the present 405-line standard is capable of producing appreciably better definition than can be obtained on the average domestic receiver to-day, as is well evidenced by comparing such a receiver with one that has been properly engineered without having to consider the cost of production.

It is interesting to speculate on the ultimate value, in the viewer's home, of a 625-line receiver, which it is generally agreed would give only a marginal improvement in definition. It is surprising to find how little the general public appreciates picture quality, and it is doubtful whether viewers, in the mass, would even be aware of the difference between a 405-line and a 625-line picture, as seen on the *average* commercial receiver.

On the other hand, if one judges the matter purely from the engineering point of view, there is no doubt that the 625-line standard offers advantages for future worth-while improvements in definition and better picture quality on the larger tubes.

To look at another aspect of the television standards problem, it has been argued in some quarters that if Britain had a 625-line system it would be possible to export receivers to the Continent and to other 625-line countries. This might indeed be so, but competition would be severe and our home market would, in turn, be opened to the import of receivers of continental manufacturers.

The foregoing are but a few aspects of the problem which emphasize the difficulties facing the Pilkington Committee, which has the task of sifting the voluminous and widespread opinions which are reaching them from interested parties in all fields.

It has to consider the problem both from the engineering and from the commercial and economic points of view. It must have due regard for the provision of additional stations and services in the future, which could well mean the use of frequency bands whose suitability for complete coverage of the British Isles cannot accurately be assessed. It has to consider the factor of obsolescence of existing receivers, of the possibility of dual standards for some years, and, in addition to all these problems, the vexed question of colour TV, compatability, and its economic feasibility. A side issue, but one which has its own difficulties, is that of the possible introduction of local radio stations.

The Institution recognizes full well the problems of the Committee and that it would be unrealistic to expect a final verdict which would satisfy all schools of thought. However, it is essential that a firm and final decision should be reached before long so that the period of uncertainty can be ended and development can take place on whatever lines are laid down.

W. E. M.

# INSTITUTION NOTICES

### British Productivity Council

The Director of the British Productivity Council has invited the Institution to nominate a representative to attend meetings organized by the Council to discuss quality control and reliability methods.

The Council has nominated Mr. F. G. Diver, M.B.E., Chairman of the Technical Committee, to represent the Institution.

### Income Tax Relief on Subscriptions

The Council wishes to remind members that under Section 16 of the Finance Act 1958, the whole of the annual subscription paid by a member is a permissible deduction from the emoluments assessable for income tax purposes.

Applications for relief can only be accepted where the subscription is defrayed out of the emoluments of the office. A further condition is that membership should be relevant to the office or employment of the claimant.

Members who have not already made application to their H.M. Inspector of Taxes should apply for form P.358. The claim may be made retrospectively to the 1st April 1959.

Members resident in the Channel Islands may obtain similar relief under Rule 2 of Case II as amended. Members should apply in writing to the Comptroller, Income Tax (Jersey), States Office, Jersey, C.I.

### Computer Control of Air Traffic

The Computer and Radar Groups of the Institution are organizing jointly a symposium of papers on the computer control of air traffic. The meeting will be held on Wednesday, 3rd May at the Gustav Tuck Lecture Theatre, University College, London, W.C.1, commencing at 3 p.m.

The six contributions will be as follows:

"General Problems of Air Traffic Control"—P. C. Haine.

"Input Requirements for Data Processing Complex"—L. G. Payne, Ph.D.

"An Experimental Data Processing Complex"—A. St. Johnston, B.Sc. (Member).

"The Output Display of Data from the Data Processing Complex"—K. H. Simpkin (Associate Member).

"Real Time Inputs and Dynamic Display for Data Processing Systems"—R. F. Hansford.

"Design Principles of Computers for Air Traffic Control"—H. Scholten.

Members are asked especially to note that, because of the limited accommodation, *admission will be by ticket only*, obtainable from the Institution.

### Institution Dinner

A Dinner of the Institution will be held on Thursday, 8th June 1961, at the Savoy Hotel, London, W.C.2. Admiral of the Fleet Earl Mountbatten of Burma, K.G. (Vice-Patron and Past President) will preside.

Application forms for tickets have been sent to all members in the United Kingdom and, as accommodation is limited to 550 members and guests, those wishing to attend are advised to return their completed forms as soon as possible.

### Waverley Gold Medal Essay Competition 1961

The scientific journal *Research* is this year sponsoring The Waverley Gold Medal Essay Competition for the ninth year in succession. The Competition is designed to encourage the scientist in the laboratory and the engineer in the production plant to express his views and translate his work into an essay that will be readily understood by other scientists and those interested in science and technology.

The Waverley Gold Medal, named after and bearing the coat of arms of the late Lord Waverley, together with £100 will be awarded for the best essay of about 3,000 words describing a new project or practical development in pure or applied science, giving an outline of the scientific background, the experimental basis and the potential or actual application of the idea to industry or their importance to society. The essays will be judged by specialists in the subject for technical content, for clarity of presentation and for style. The entry should be written as an essay which should interest a well-informed layman. In assessing these essays particular attention will therefore be paid (*a*) to the logical presentation, (*b*) to the style, and (*c*) adherence to these terms of reference.

A second prize of £50 will be awarded and also a special prize of £50 for the best entry from a competitor under the age of thirty on 31st July 1961. If the first prize is awarded to a competitor under the age of thirty, the special prize will go to the next best entry. Entry Forms can be obtained from the Editor of *Research*, 88 Kingsway, W.C.2. *The last date of entry is 31st July 1961.*

### Correction

The following amendment should be made to the paper "The Relative Magnitudes of Modulation Products in Rectifier Modulators and some Effects of Feedback" published in the March *Journal*:

Page 278: Equation (23) *should read*

$$i_0 = \frac{(r_b + R) - s(r_b - r_f)}{r_b - r_f} \cdot \frac{\pi i_1}{\sin \pi s}$$

and the following line *should read* "so that as $s$ tends to 0 or 1, . . . ."

U.D.C. 621.397

# Television Anomalies—Past, Present and Future

*By*

L. H. BEDFORD,

C.B.E., M.A., B.Sc.(Eng.), F.C.G.I.,

(*Past President*)†

*Address given at the Inaugural Meeting of the Institution's Television Group held in London on 1st March 1961.*

Television is a subject which bristles with anomalies and a review of some of these seemed to me to be suitable subject matter for this occasion. I think it will be of greater interest if I select from this rather wide field largely those subjects in which I can claim to have had a personal hand. For this reason, and for the intrinsic practical importance of the subject, I propose to devote my first and perhaps major emphasis to the telecine problem.

What then is the anomaly of telecine? I would say first that that which should be the easiest problem of television technique has in fact proved to be one of the most difficult. Secondly, we seem to have come full circle back to the flying spot method of film scanning, which I personally abandoned in 1938 as being obsolete in view of the overwhelming advantage of the storage method.

Let me cast your mind back to 1936 when the B.B.C. opened the first seriously-to-be-considered television service in the world.

On the recommendation of the Selsdon Committee we were inflicted with two systems using different standards. I would like to remark in retrospect that I consider that it should have been possible for the Committee to see, on the current evidence, that the system based on the storage type camera tube, not to mention the 50 c/s field rate, was bound to prevail, and that we ought to have been spared this painful period of dual standards. However, in the present connection I am glad that we were made to suffer in this way because the Baird system transmissions on 240 line 25 frames/second did include some of the finest film transmissions that I have seen to this day. (It is perhaps only fair to remark that Fernseh A.G. were producing strictly comparable results on an experimental basis in Germany, as is well known to those who were fortunate enough to visit those memorable Funkausstellungs in Berlin.)

Why were these 1936 vintage film transmissions so outstandingly good? Fundamentally because the light conditions were adequate to provide a very good signal/noise ratio. Indeed so fundamental does this light advantage with film prove to be that a number of parties (including myself at one time) took the view that one could reasonably contemplate a television service based exclusively on film. The limit of this philosophy was expressed in the Fernseh-Baird Zwischen-film process on which the earliest so-called "direct transmissions" were put out over the Baird system. The results were quite unacceptable, and the public, and the press in particular, were quite unmoved by the really gallant technical efforts which lay behind all this and saw only the humorous side. I recall the following piece of press notice:

" 'The next conjuring act by Mr. Blank', said the announcer, 'will surprise you.' It did. His head vanished, followed by a gradual fade out of his entire body."

Any one who like myself knows the agony of being committed to demonstrate on a sticky wicket will sympathize with Baird at this time. But the Zwischen-film process was a poor stool on which to seat a Canute act of holding back the tide of the storage camera tube.

Reverting, however, to the subject of film scanning *per se*, perhaps I may be allowed to discuss for a short time my own work on this subject, carried out in collaboration with Puckle and Stevens. It is incidental that we were at that time exploring a method of transmission which we called "velocity modulation", independently paralleling the work of Von Ardenne, which he called "liniensteurung". We succeeded with or without velocity modulation in producing a reasonably good picture source starting with film subject matter and scanning with a cathode-ray tube.

The difficulties were of two kinds. The primary difficulties were concerned with scanning a stationary film and were the associated difficulties of screen after-glow and signal/noise ratio.

† English Electric Aviation Ltd., Luton

Perhaps we were the first to hit the problem of after-glow compensation. With the screen materials to which we had access the large amount of high frequency compensation caused us to be pushed for signal/noise ratio. The screen specialists could not help us much since after-glow to them was something measured in seconds and we were talking in micro-seconds.

The time scale was therefore something quite foreign to them and though we could provide the means of measurement they lacked a clear picture of the mechanism of fluorescence which would enable them to make progress. In fact to this day progress has not been altogether startling as the short time-constant materials usually suffer from low fluorescent efficiency. One way or another, however, progress has been just sufficient to transition the system from a state of marginal to spectacular acceptability.

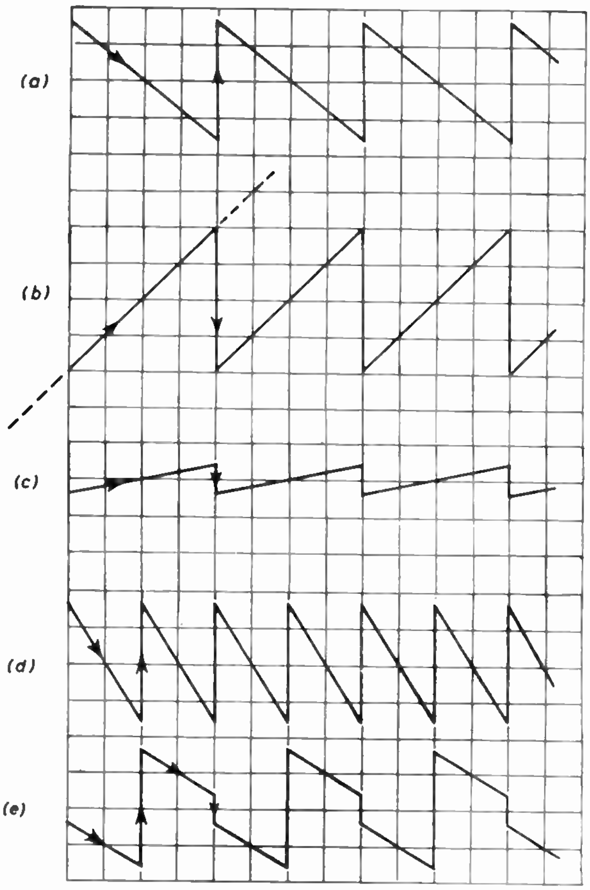



Fig. 1. Scanning patterns with a uniformly moving film.
(*a*) Frame scan of one nett picture height.
(*b*) Film motion over one gross picture height.
(*c*) (*a*) and (*b*) superposed.
(*d*) 50 c/s field rate with 25 c/s film frame rate.
(*e*) Composite scanning and chasing waveforms.

The secondary difficulties relate to the problems of scanning moving film. There is a natural inclination from the outset to discard intermittently moving film owing to the 25% pull-down time occupied by the standard 90 degree Maltese cross movement.

Consider then the problem of scanning a uniformly moving film. My preferred way of regarding this problem is to consider that one starts with a raster scanning a stationary film frame and then applies to it a "chasing" motion to cause it to follow the film motion.

Figure 1 illustrates the dynamics of the affair and shows the relevant events as seen at the tube face. It is assumed that the line scan is horizontal and that frame scan is from top to bottom (the argument, however, applies generally to any scanning convention whatever including the Chinese with which we actually started!). Figure 1(*a*) shows the frame scan of one nett picture height. Figure 1(*b*) shows the film motion at a rate of one gross picture height per frame time; but since we discontinuously change attention from one frame to the next once per frame time, the appropriate chasing wave form becomes a saw-tooth of amplitude one gross picture height. (It is convenient to co-phase the discontinuities of the scanning and chasing waveforms.)

The result of superposing motions (*a*) and (*b*) is shown in (*c*) and here we have to thank Providence for the finite difference between gross and nett picture height. Had this been zero, the scanning raster would have collapsed to a single line resulting in a screen loading which no material could stand!

Figure 1(*d*) shows the case of a 50 c/s field rate with a 25 c/s film frame rate, and Fig. 1(*e*) shows the composite scanning and chasing waveform. We now arrive at a double scanning raster of favourable total area. Unfortunately the process is riddled with snags. First of all it is hardly necessary to say that one would not think of producing waveform (*e*) by actual superposition of waveforms (*d*) and (*b*), for any non-linearity in the chasing waveform would cause discrepancies in the information content of the two interlaced half rasters resulting not only in loss of effective resolution but in 25 c/s flicker.

The above is mentioned as typifying the problem which accompanies multiple field scanning, and in one form or another it obtrudes into any solution. For example, if we overcome linearity difficulties by the obvious procedure of producing waveform (*e*) by the superposition of a 50 c/s saw-tooth of reduced amplitude and a square wave of 25 c/s we find ourselves now suffering from differential luminosity between corresponding elements in the tube screen. The better solution is then to use a single scanning raster at field frequency and introduce the square wave motion optically, e.g. with dual lenses and obscuration.

An alternative approach pursued by Puckle and myself may be worth recording as it contained an element of boldness. Perhaps this is a wrong description for a bold step is one which comes off; if it does not come off it is sheer foolishness.

This particular approach was prompted by the fact that the prime experimental requirement for us was the scanning of stationary film and we were impressed by the superior results of this as compared with even our best results on moving film. This suggested a reversion to intermittent film motion, speeding up of the pull-down time, accommodating the high speed part of the film movement in the blanking time and the chasing out of the relatively slow film movement not contained in the blanking time.

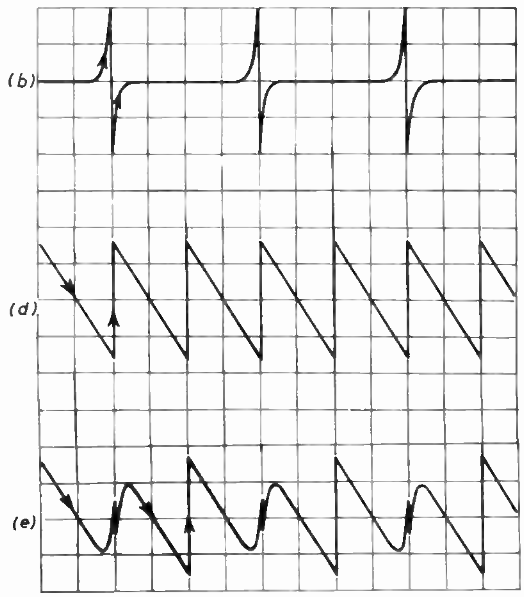



**Fig. 2.** Scanning patterns with intermittently moving film. The patterns correspond to the similarly labelled items of Fig. 1.

Figure 2 illustrates the waveforms resulting. The problem is to provide a chasing waveform corresponding to the film movement, Fig. 2(*b*); this waveform is to be applied as such with sufficient accuracy in waveform, amplitude and phase relationship to the film movement. We used a light-cam driven directly from the constant speed 25 c/s wheel driving the Maltese cross. We shortened the pull-down time by making up a 120 deg Maltese cross thus reducing the pull down angle to 60 deg (6·6 milliseconds). Even with the most careful adjustment the results were barely acceptable, and here we may note a minor anomaly in passing. As engineers we were determined not to stretch the blanking period beyond its standard value of 1·4 milliseconds maximum. However, we note that at the present day it is common to transmit "wide screen" format, which is ruthlessly realized in practice by extending the blanking period to an extent which would have swallowed up all our difficulties and would apparently have met with unqualified public acclamation!

All the above work refers to 35 mm film. At a later date when the same problem arose for 16 mm film I attempted the same sort of solution starting off with the advantage that the standard pull-down time is smaller (4 milliseconds approx.).

This time we attempted to follow the movement of the film itself rather than that of its driving member; this we did by tracking the movement of a sprocket hole. The result of this was a failure, but it played some part in leading to a device which was quite successful, namely the 16 mm fast pull-down mechanism, which is in actual use in Marconi flying spot equipment.

What determines the extent to which one can speed up the pull-down of an intermittent motion? In a mechanism designed for minimum acceleration, the latter is inversely proportional to the square of the pull-down time.

For 1 millisecond pull-down the minimum acceleration assumes the following rather terrifying values:

35 mm film : 7800 $g$
16 mm film : 3400 $g$

The Marconi fast pull-down mechanism is essentially a standard heart-and-square movement driven by a member whose rotational speed is not uniform, but which has a maximum speed in the vicinity of pull-down. This ingenious design, due to Kingston, features easy adjustment of the pull-down time; this is normally set to 2 milliseconds.

The lesson is perhaps not to be too rigidly bound by either television or film industry standards (the latter at one time being regarded as absolutely sacred). A little bending of each is capable of producing a useful result.

In order not to over-weight this subject of flying spot scanning I have left out the evidently painful problem of applying this process on American standards—namely 60 c/s field rate, a 24 c/s film frame rate.

I should here point out that many of the problems discussed above under flying spot telecine occur in the inverse problem of film telerecording. I know of only two effective solutions. The first is the fast pull-down mechanism which solves all the problems for both telecine and telerecording for both 50 c/s and 60 c/s field frequencies by invoking the philosophy of departing slightly from accepted standards in the two arts. The second, which as far as I know has not been applied, being ruled as offside before being started, is to abandon totally the idea of telerecording on accepted film standards but to treat the film as a data recording medium divorced from film standards. Here then one might contemplate a continuously moving film (which could incidentally be at somewhat reduced speed) recording from a standard interlaced raster, consecutive half rasters thus occupying consecutive portions of the film, and the signal would be read out flying spot-wise in exactly the same manner.

I appreciate the commercial objections to the situation implied here of having separate equipment for the scanning of normal film and tele-film, but I am convinced that this has been given too much weight, especially when one considers the rather horrific solutions that have been advocated such as suppressed field recording. Indeed with the relative increase of importance of tele-film I am sure events would have moved along this direction but for the complete change in the situation occurring in a somewhat unexpected manner, the introduction of video tape recording.

Of this I would only like to say that any initial scepticism that one may have felt, and I personally felt a lot, has been completely confounded. Indeed so great is the success of this technical *tour-de-force* that I am now in the embarrassing position of being unable to recognize a recording from a direct transmission—a situation I could never have complained about in film tele-recording, but which now causes me the greatest indignation.

Before discussing the apparent overwhelming advantages which lie with the storage method of telecine perhaps a few words about the storage tube camera tube *per se* would be appropriate. Undoubtedly it is the greatest concept which ever came to the aid of the television art. Owing to the facts that the development is foreshadowed by some long thinking by Campbell Swinton and that parallel work proceeded largely independently in America and in this country, I am not sure as to the strictly legal paternity of the storage camera tube, but I believe that the first practical tube was the iconoscope of Zworykin. It gives me great satisfaction to record that the Institution has honoured Zworykin by the award of Honorary Membership.

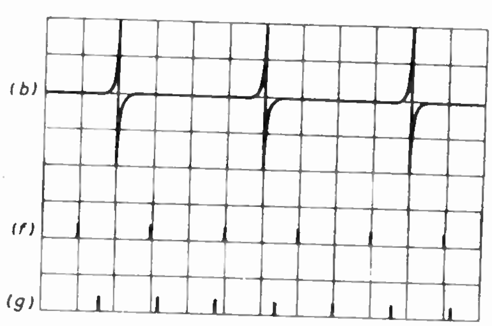



Fig. 3. The scanning patterns with a storage tube.
(*b*) Film motion over one gross picture height.
(*f*) Film movement with 50 c/s field rate and 25 c/s film frame rate.
(*g*) Film movement with 60 c/s field rate and 24 c/s film frame rate.

My first acquaintance with the iconoscope proposal was in the form of that most unfavourable of all technical publications, the patent specification. I reported that the idea was a master stroke but that the tube was most unlikely to work. A defective piece of prophesy if ever there was one, but one which could have been fully salvaged if only I had had the forethought to add the words "as described". For indeed, as was later elucidated, the *modus operandi* of the iconoscope was considerably different from what the inventor had in mind.

Now from all I have been saying above it is clear that a camera tube which can store a picture as charge pattern from pulse-like illumination disposes *in toto* of all the film movement problems described above. Figure 3 makes this clear.

All we have to do is to illuminate during field blanking and then we have all the time in the world for film movement. Figure 3(*f*) shows a field rate of 50 per second and film frame rate of 25 per second. Figure 3(*g*) shows a field rate of 60 per second, and a film frame rate of 24 per second. This last figure shows that pull-down time must be less than 1/5th film frame time, i.e. 72 deg or 8 ms approx. Furthermore the iconoscope belongs to the class of tube in which both signal/noise ratio and grey scale improve with increase of light, so, considering the superlative studio pictures which the tube was producing, it seemed inevitable that the iconoscope would outclass all competition as a telecine device.

But such was not the case, and for some years the standard of film transmission lagged significantly below that of direct studio transmissions. The reason may have been clear to those directly concerned with the problem, but to the rest of the engineering fraternity it was something of a puzzle, which only became clear when the curious storage characteristic of the iconoscope began to be understood. I refer to the phenomenon now known as line storage, which can be briefly described by saying that the main storage effect occurs only during the scanning of an adjacent line. The practical result of this was that not only was the storage on intermittent illumination very inefficient but the so-called shading effects (tilt and bend in B.B.C. parlance), already a major problem in the continuous illumination case, became virtually unmanageable.

A solution was found in terms of the Mechau projector; a throw-out, if I may so put it, from the film industry. The Mechau projector was an extremely elegant attempt to provide projection without intermittent motion or shutter.

The Mechau projects a continuously moving film by use of an optical chasing system comprising a rather complicated mirror drum. It does in fact superpose adjacent frames fading smoothly from one frame to the next. It was, I imagine, conceived as a means of permitting a reduction in frame frequency for there is theoretically a complete absence of flicker at any speed. It was rejected in the film industry for reasons concerned with light efficiency and delicacy of adjustment. The first of these was inoperative in the television case and only the second appears to have defeated the scheme, but this only after it had given an excellent run for its money.

The next logical move in the evolution of storage tube telecine was a reversion to intermittent motion, which despite the barbarity of its concept gives a substantially faultless performance of high practical reliability, and to change to a more perfect storage tube. I use the word perfect as qualifying the storage process as such. But, to everybody's surprise and chagrin, no really satisfactory device could be found. My own experience was limited to work with the image orthicon, and I well remember some of the shocks we sustained in trying to deploy the tube in this role. For example, we rapidly discovered that the storage mechanism in the image orthicon was, like the Black Sea, "not all it's cracked up to be". One disconcerting experiment which we made was to pulse-illuminate at 25 c/s while scanning at 50 c/s field rate when it was found that while the first field after the illumination came through intact the second had largely disappeared. Adding to imperfect storage the other characteristics of the image orthicons (3-in.) then available, e.g. limited signal/noise ratio and the dislike of the normal film gamma, the overall results were worse than disappointing.

It was in this phase of the art that the flying spot method of telecine staged its come-back under the inspired leadership of Nuttall. Just what an advance in picture quality was achieved is best expressed by the reaction of our North American visitors, who described our telecine as something of a totally different order.

I would express the opinion that flying spot telecine is here to stay, except for the fact that in this still rapidly advancing art one can only obtain a reputation for wisdom by keeping one's mouth shut; for already the problems of *vidicon* telecine are being overcome to an extent which could still reverse the situation in favour of the storage method.

In this connection I want to make a species of plea which arises from the fact that one of the attractions of vidicon telecine is its relatively low cost (e.g. compared with the flying spot). My plea is that the cost comparison is quite irrelevant and that the choice should be made on technical considerations. This is because however lavishly a broadcaster may invest in engineering equipment the total costs are still quite trifling compared with the other station running costs on the programming and artistic side. For I am one of those who believe that television is *only just good enough at its best*, and that the philosophy "any old thing will do for television" should be read in the vernacular interpretation of this phrase.

My next two major anomalies are inextricably bound up together. I refer to the questions of *standards* and *colour*. On the first issue, without wishing to imply that our present 405 line standard is technically outmoded, for I have yet to be shown a finer quality of television service than our own anywhere in the world, we have to accept that we are standards-wise in a state of isolation.

Here we are perhaps paying the penalty of the pioneer, namely that we had to determine our standards in an early stage of the art. I think it a great credit to Blumlein and the rest that these standards have stood up so remarkably well over a period of 25 years.

When television was resumed after the interruption of the war years, we had a chance to reconsider the question of standards. U.S.A. had already settled for 525 lines, 60 fields/second 30 frames/second interlaced, and opinion in Europe was hardening towards something equivalent but adjusted to suit the European mains frequency, eventually 625 lines 50 fields/second 25 frames/second. There was a strong school of thought† that since the tendency of the art must be to advance and since this would probably be our last chance for a reconsideration of standards, we should go forward with Europe. Against this were (*a*) the consideration of implied responsibility to owners of pre-war sets, now seen to be trivial, and (*b*) the anxiety to get back into business with the least trouble and delay. These considerations won the day, so that we now find ourselves committed to a splendid isolation in standards with apparently no way out. Does this matter? Clearly it may do from two points of view. First it greatly increases our export problem both for transmitting and receiving apparatus, although the latter is probably the less important item since no major television consuming country is likely to remain for long an importer of receivers. Secondly, it puts an impediment in the path of what may well be television's most important role, namely international programming. For I can think of no medium so well

† See "Post-war development report. Part 1", *J.Brit.I.R.E.*, **4**, pp. 135–50, October 1944.

adapted to make the nations aware of each other's culture and mode of life.

While acknowledging the already excellent efforts that have been made in the direction of international television programming, I think this aspect has so far had nothing like the emphasis that it deserves. The B.B.C. crest features the motto "Nation shall speak peace unto nation". This phrase was invented for radio but could well find its true fulfilment around television.

But any discussion of the future of television and of television standards immediately gets itself snarled up in another issue, that of colour. Broadly speaking, would the introduction of colour provide the chance which was previously thrown away of reaching international standardization? Conversely, however, one must ask the question "if we adopt new standards for colour have we automatically damned it to failure by the absence of compatibility—a feature which was so basic to the American approach?"

Before pursuing this argument, we should perhaps pause to consider certain features of the American scene. Going back into history you will recall that for a short period the relatively easy solution of a colour field sequential system (Goldmark) was formally adopted by the F.C.C. as the standard American system. The fact that the receivers as launched in the simplest embodiment comprised a mechanical colour wheel may have been objectionable but was not fundamental. Two other factors were fundamental:

(*a*) In order to achieve colour the system had either to demand more bandwidth or to cede an undue amount of resolution.

(*b*) The absence of compatibility.

I need not embark on a formal discussion of this latter issue. It suffices to say what actually happened. Every time a station went on the air with colour it became inundated with panic telephone calls from viewers complaining that their picture had gone wrong. No amount of education availed; the set was O.K. on the other channels so the transmitter must be at fault!

The logical conclusion was not so much the necessity for compatibility but rather that an incompatible transmission was inadmissible on existing channel allocations.

The nett result of all this was that the field sequential system was withdrawn and the F.C.C. referred the colour problem back to the industry with the following delightful formula:

(1) The desired signal must be compatible in the sense that monochrome receivers would work from it without degradation and without any need for modification.

(2) There was to be no increase in channel bandwidth.

(3) The colour picture was to have the same resolution as existing monochrome receivers.

The cynic may say that here was the let out, since the resolution of the average American television set at the time was down on the theoretical by a large factor! But, discounting this somewhat frivolous interpretation, the F.C.C. was in fact posing to the industry a strictly insoluble problem.

Insoluble problems have a way of getting solved, sometimes by the introduction of an intelligent element of compromise. This is what in fact the American engineers achieved with a brilliance which has my fullest admiration.

Let us then attempt a summary of the American break through:

(1) My friend and namesake, A. V. Bedford of R.C.A., elucidated the principle of "mixed highs". He showed that the resolution of the eye as to colour was much less than its resolution as to luminance. If, then, it were possible to transmit separate luminance and chrominance signals the latter could be of substantially reduced bandwidth. Moreover the luminance signal should constitute an ideal "compatible" signal for monochrome reception.

(2) Dome among others expounded the principle of "frequency interlacing", or more strictly "line spectrum interlacing". The spectrum of a steady picture is essentially a line spectrum at multiples of the line frequency. A line spectrum made up of odd multiples of half the line frequency can be interlaced with this.

(3) Sziklai and others introduced the concept of handling the chrominance information in terms of hue and saturation, these two quantities to be transmitted as phase and relative amplitude of a sub-carrier. The selection of the sub-carrier at an odd multiple of half line frequency allows the above frequency interlacing process to be applied.

These are very powerful principles indeed.† I will pass from them to a compressed account of how they are translated into practice.

The system starts off as a three-colour simultaneous system by which I mean that we produce concurrently signals $r(t)$, $g(t)$, and $b(t)$, which we require to transmit simultaneously. However, as indicated above, we prefer to transform these quantities prior to transmission. We transmit the quantities $I(t)$, $Y(t)$, and $Q(t)$ obtained from the primary quantities by the linear

† Both the importance and the validity of the Dome principle can be held arguable. Its influence on the N.T.S.C. system is that it suggests a device which proves useful in a somewhat different realization.

transformation

$$\begin{bmatrix} I \\ Y \\ Q \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & & \\ d_{31} & & \end{bmatrix} \begin{bmatrix} r \\ g \\ b \end{bmatrix}$$

Here the units for $r$, $g$, $b$ are chosen such that these quantities are numerically equal for "white".

This transformation has nine degrees of freedom of which we may immediately discard three as merely scaling factors. The basic requirement of the system is that one of the signals, say $Y$, should constitute a "luminance signal". This introduces the concept of the luminance factors $l$, $m$, $n$ of the primaries, namely the contribution of each primary in lumens to one lumen of white light. For the N.T.S.C. primaries $(l, m, n) = (0{\cdot}30, 0{\cdot}59, 0{\cdot}11)$.

It is shown in the Appendix that the condition for $Y$ to carry the whole luminance signal and for $I$ and $Q$ to carry no luminance signal is $d_{21} : d_{22} : d_{23} :: l : m : n$. This is the Hazeltine contribution, which they curiously described as a "constant luminance system". This condition absorbs a further two degrees of freedom, leaving four for disposal. It is next convenient to make the $I$ and $Q$ signals vanish on white, which involves

$$d_{11} + d_{12} + d_{13} = 0$$
$$d_{31} + d_{32} + d_{33} = 0$$

This leaves two degrees of freedom for disposal. There is thus plenty of room for manoeuvre; the N.T.S.C. choice is to my way of thinking somewhat arbitrary:

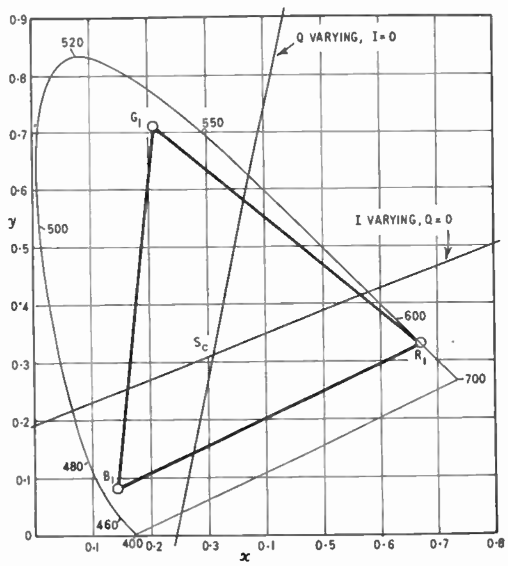



**Fig. 4.** C.I.E. chromaticity diagram for N.T.S.C. system.

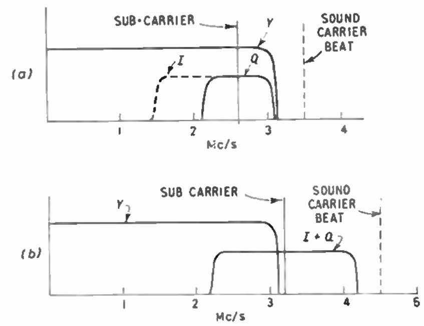



**Fig. 5.** Disposition of $I$ and $Q$ signals in the video channel.

$$\begin{bmatrix} I \\ Y \\ Q \end{bmatrix} = \begin{bmatrix} 0{\cdot}60 & -0{\cdot}28 & -0{\cdot}32 \\ 0{\cdot}30 & 0{\cdot}59 & 0{\cdot}11 \\ 0{\cdot}21 & -0{\cdot}52 & 0{\cdot}31 \end{bmatrix} \begin{bmatrix} r \\ g \\ b \end{bmatrix}$$

whence conversely (for the receiver)

$$\begin{bmatrix} r \\ g \\ b \end{bmatrix} = \begin{bmatrix} 0{\cdot}95 & 1 & 0{\cdot}62 \\ -0{\cdot}28 & 1 & -0{\cdot}64 \\ -1{\cdot}10 & 1 & 1{\cdot}73 \end{bmatrix} \begin{bmatrix} I \\ Y \\ Q \end{bmatrix}$$

Figure 4, for what it is worth, presents these transformations on the C.I.E. chromaticity diagram.

Now consider how these signals are actually disposed in the channel. I will translate to British 405 line standards for ease of comprehension (Fig. 5).

The sub-carrier is placed at 2·6 Mc/s approximately, the exact frequency being an odd multiple of half line-frequency. The sub-carrier is provided in two phases, "in phase" and "quadrature", and the $I$ and $Q$ channels are transmitted as modulations of these two carriers. To separate these at the receiver requires a synchronous oscillator at sub-carrier frequency tightly phase-locked to the transmission. This is achieved by the "colour burst" transmitted at the beginning of each line: a major cause for concern at first sight, which has fortunately proved to be ill-founded. The synchronous detection process is a successful and powerful technique which may well prove a valuable by-product of the colour enterprise.

The above placing of the sub-carrier leaves room for a double sideband channel of $\pm$ 0·4 Mc/s only, an undesirably low figure. To improve this one would like to place the sub-carrier lower, but one is debarred from so doing by the resulting dot pattern. It is possible to use an asymmetric sideband transmission on one channel and this is done as shown. The advantage is, however, partly illusory for two reasons.

First, dot patterning accompanying the higher frequency part of the chrominance signal introduces objectionable spuriosities into the picture. Secondly, it is not entirely accepted, and certainly not by me, that there exists this obligingly accommodating colour "direction" which is chosen for the narrow band channel.

By and large the various dot patterning which results from interlacing the chrominance channel with the luminance channel is an expression of imperfect realization of the Dome principle. This is not surprising, for the latter would seem to imply the application of quite unrealizable "comb filters". In the absence of these a frequency interlace gives rise to a dot pattern whose geometrical pitch is minimized by the interlacing effect of the odd multiple half line frequency device; but unfortunately the dot interlacing is at half frame frequency.

At this stage perhaps we should ask the question, to what extent Britain should accept the N.T.S.C. colour system. This question breaks down into the following sub-questions:

(1) Should we start a 405 line N.T.S.C. colour service? This of course is by far the easiest thing for us to do; indeed the practical difficulties of doing anything else are so great that I am forced to the opinion that this is the most likely event. Moreover, such a service would undoubtedly "get by" on a short term basis. But I have the most serious doubts as to its acceptability as an all-time standard, and, make no mistake, this is what it would become if we once put it into service.

(2) If we consider that an advance on the above described standards is required, what should this be? I would like to answer this first of all with reference to a 405 line standard. My proposal would be to follow the N.T.S.C. approach but to shift the sub-carrier outside the 3 Mc/s luminance band, $I$ and $Q$ channels to be each double sideband and of about $\pm$ 1 Mc/s bandwidth (Fig. 5(*b*)). This gives us a total of approximately 4 Mc/s video channel width, which should not put an insuperable problem to the communication authorities or, as the Americans so rudely put it, to the "Common Carrier". The difficulty is that the proposal requires a change in the spacing of the sound carrier, namely to at least 4·5 Mc/s from the vision carrier. Such a procedure is not possible in Band I. It probably also is not possible in Band III. I mention for the record that when Band III was being allocated I was among the few who made a strong plea for the reservation of wider channels in reserve for colour, but so great was the greed for channels that we could get no hearing.

So it seems that any such system is forced into Bands IV and V or off the air altogether and on to cable.

I would remark here, at the risk of being mistakenly considered a dyed-in-the-wool 405 line advocate, that this proposal would result in an anomalous but highly favourable situation. The problem to the Common Carrier would be no worse than that for the American case, but we should secure an infinitely superior picture. The situation may be likened to that of video tape where already the British pictures are superior to the American because the demands of the lower line standard are better matched to the capabilities of the medium.

(3) In the event of a decision to transmit colour only in Bands IV and V, and the practicality of such a decision is certainly suspect, then the time would be ripe for a reconsideration of line standards, principally for the sake of standardization with Europe. In fact international standards for colour transmission in Bands IV and V would seem to be a worthy object of our serious and immediate attention. I would plead that the artificial though fruitful ruling of the F.C.C. that the colour channel width should not exceed that of monochrome is not applicable.

We should be grateful to the F.C.C. and the N.T.S.C., but should be prepared to depart from them in order to meet Nature half way!

But European standardization brings us head on to another anomaly. The odd man out is now likely to be U.S.A., and this at a time when satellite communication opens up for the first time the possibility, indeed probability, of transatlantic television.

(4) Finally, we may ask, is the inauguration of a British colour service a matter of urgency? I think that channel provisioning is a matter of urgency, but that the service itself is the *reverse*. This brings me to my last and greatest anomaly of the television scene, the colour receiving tube or more particularly the R.C.A. tricolour tube. Of this most remarkable achievement, if I say that it was boldness to the point of folly to expect this tube to be either workable or makeable, and that nobody but a madman would have attempted to put it into large-scale production, then the laugh is on me. It *is* in large-scale production; I am told at a rate of 1000 per day!

It is rather difficult to square this figure with another (of equal but low authenticity), namely the present rate of colour television receiver production at 100 000 per year. Either figure sounds extremely high by British standards, and would suggest that colour television has now made its long heralded break-through. But such is not the case. Even at this figure, the proportion of the television market intruded by colour is extremely small. It is also by no means certain that colour television sets in the field are good propaganda; a high proportion of them, and especially those most seen by the public, namely those in bars and saloons, are certainly not.

When the American colour television programme was formulated, I had grave misgivings as to whether it would "go". This on three grounds:

(*a*) The circuitry associated with brilliant N.T.S.C. system appeared as possibly too complex for the home. On this point most of us are now reassured.

(*b*) The relative cost of the extra information, namely the chrominance, seemed disproportionate.

(*c*) The tricolour tube, even if manufacturable at an acceptable cost, gave undue opportunities for maladjustment in the home.

It is my belief that this last point is still valid and that despite the brilliance and boldness of the work leading to the tricolour tube this item is still the dead weight which is holding back the progress of colour television. I go so far as to say that the introduction of a service is premature so long as the tricolour tube remains the only practical receiving tube in view.

All in all, I am tempted to end with a quotation:


(THE MIKADO)

## Appendix

Proof that the condition $d_{21} : d_{22} : d_{23} :: l : m : n$ renders the $I$ and $Q$ channels non-luminance-bearing.

The transmitted quantities $I$, $Y$, $Q$ are defined by

$$\begin{bmatrix} I \\ Y \\ Q \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & & \\ d_{31} & & \end{bmatrix} \begin{bmatrix} r_1 \\ g_1 \\ b_1 \end{bmatrix}$$

where $r_1$, $g_1$, $b_1$ are the primary intensities at the camera; and the receiver performs the inverse operation

$$\begin{bmatrix} r_2 \\ g_2 \\ b_2 \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & & \\ d_{31} & & \end{bmatrix}^{-1} \begin{bmatrix} I \\ Y \\ Q \end{bmatrix}$$

$$= \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & & \\ c_{31} & & \end{bmatrix} \begin{bmatrix} I \\ Y \\ Q \end{bmatrix}$$

The receiver luminance $Y_2$

$$\begin{aligned} &= lr_2 + mg_2 + nb_2 \\ &= (lc_{11} + mc_{21} + nc_{31})\, I + \\ &\quad + (lc_{12} + mc_{22} + nc_{32})\, Y + \\ &\quad + (lc_{13} + mc_{23} + nc_{33})\, Q \end{aligned}$$

We have to show that the condition

$$d_{21} : d_{22} : d_{23} :: l : m : n$$

causes the coefficients of $I$ and $Q$ in the above expression to vanish.

Writing $D$ for the determinant

$$\begin{vmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & & \\ d_{31} & & \end{vmatrix}$$

consider the coefficient of $I$:

$$c_{11} + mc_{21} + nc_{31}$$

$$= \frac{1}{D}\left[ l \begin{vmatrix} d_{22} & d_{23} \\ d_{32} & d_{33} \end{vmatrix} + m \begin{vmatrix} d_{23} & d_{21} \\ d_{33} & d_{31} \end{vmatrix} + n \begin{vmatrix} d_{21} & d_{22} \\ d_{31} & d_{32} \end{vmatrix} \right]$$

$$= \frac{1}{D} \begin{vmatrix} l & m & n \\ d_{21} & d_{23} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{vmatrix}$$

which vanishes for the conditions stated. Similarly the coefficient of $Q$ vanishes.

*Manuscript received by the Institution on 1st March 1961*
*(Address No. 25.)*

# PROCEEDINGS OF THE COUNCIL

*Reports and recommendations from the various Committees of the Institution are discussed at Council Meetings. The work of the Membership, Programme and Papers, Group, Convention, Technical, Local Section and other Committees is regularly featured in the* Journal. *A digest of other matters of immediate and particular interest to all members will be given in these notes from time to time.*

**Institution Accommodation.**—The matter of acquiring adequate and suitable headquarters for the Institution has been repeatedly discussed since 1951 by successive Councils. It is now impossible to effect any further expansion within 9 Bedford Square. Further accommodation is thus a matter of extreme urgency in order to meet the needs of additional staff and to ensure that services to members shall continue to increase.

The programme of meetings arranged by the specialized Groups, in addition to the normal monthly meetings of members, is dependent upon the availability of Lecture Theatre accommodation in nearby University buildings. Regular meetings of Standing and Group Committees alone now necessitate the use of the Institution's Meeting Room every day, and the Council is also anxious to provide for extensions to the Library and reading rooms.

Every endeavour continues to be made to acquire a suitable site or building in London for the future permanent headquarters of the Institution. In the meantime, and as a temporary measure, the Council is considering taking a lease on No. 8 Bedford Square in addition to the existing tenancy of No. 9.

**9th Edition of List of Members.**—The Council approved the publication of the 9th Edition of the List of Members. It was agreed that the handbook should include the names and addresses of all corporate members, Companions, Associates and Graduates as at 31st December, 1960. The usual features of a topographical list and other relevant Institution information will be included in the new edition.

Printing arrangements should enable the handbook to be circulated to all members, other than Registered Students, early in May 1961.

**Membership Subscriptions.**—In the report of the Finance Committee given at the 34th Annual General Meeting in 1959, unanimous approval was given to a recommendation that subscriptions be increased. For the reasons given at the Annual General Meeting held on 11th January, 1961, the Council has not taken the necessary steps to convene an Extraordinary General Meeting in order to implement that agreement.

Council has now decided that an Extraordinary General Meeting shall be held within the next six months. In order to avoid the necessity to hold such special meetings to approve fluctuations in subscription rates, the Council will recommend that the appropriate Article or Bye-Law of the Institution shall stipulate *maximum* subscriptions beyond which the rates may not be raised without the approval of corporate members.

Under this arrangement it is proposed that the subscriptions shall be as follows:

*Full Members and Companions:*
Not to exceed £20.0.0d. per annum and for the present to be £9.0.0d.

*Associate Members, Associates and Graduates over 35 years:*
Not to exceed £16.0.0d. per annum and for the present to be £7.10.0d.

*Graduates (25–35 years):*
Not to exceed £12.0.0d. per annum and for the present to be £6.0.0d.

*Graduates (under 25 years):*
Not to exceed £10.0.0d. per annum and for the present to be £5.0.0d.

*Students:*
Not to exceed £10.0.0d. per annum and for the present to be:

Over 25 years £5.0.0d.
Under 25 years £3.10.0d.
Under 21 years £2.10.0d.

The Council is giving further consideration to the rate of subscriptions which shall be paid by members outside the United Kingdom, and for the present does not propose to recommend any alteration to the rates of Entrance Fees.

**Indian Advisory Committee.**—The Council considered a recommendation from the Indian Advisory Committee that an Institution Convention be held in India in 1962–63. This will be the first such venture undertaken by the Indian Advisory Committee, and the Committees of the five Sections (Bombay, Delhi, Bangalore, Calcutta and Madras) are assisting the main Advisory Committee in planning the scheme of meetings. It is expected that the Convention will be held in Delhi.

U.D.C. 621.316.86

# Tin Oxide Resistors

By
R. H. W. BURKETT, B.Sc.†

*Presented at the Symposium on New Components held in London on 24th–27th October, 1960.*

**Summary :** The methods of depositing thin films of tin oxide are briefly described and the physical, mechanical and electrical characteristics of the films are then reviewed. Manufacturing techniques for resistors are discussed, in particular the use of a ceramic substrate. Precision, power and miniature film resistors are described and typical performance figures given.

## 1. Introduction

The application of thin films of various oxides to the surface of vitrified materials is not a new technique. It has been used traditionally for the decoration of glassware to imitate the iridescence present on the surface of glass which has been weathered. The use of tin oxide as a resistive film is fairly recent, and the first applications were for its use as a means for heating and de-icing of wind shields,[1] and for stress distribution on high-voltage insulators.[2] The use of pure tin oxide produces a relatively unstable film with an uncontrolled temperature coefficient but, for the above purposes, this lack of quality is not very important. It has been found that the addition of other elements to the tin oxide film improves stability and temperature coefficient. The first process employed the addition of antimony oxide,[3] and since then other materials have been found to be equally effective.[4]

Tin oxide in the pure form is an insulator and, therefore, the deposition of a film of the material is not an obvious way of producing a resistive element. Clearly the process of deposition must produce some degree of semi-conduction in the film and tests confirm that the film is an *n*-type semiconductor. It appears that the film is inherently semi-conductive because of oxygen deficiency, though there is a suggestion that the addition of other elements may produce an electron excess or deficiency.[5] It is possible that both mechanisms apply, though the majority of carriers arise from oxygen deficiency.

The process has been exploited in the manufacture of resistors by application of the semi-conducting films to glass and ceramics. The first commercially available resistor was based upon glass, but there are certain disadvantages in this material as a substrate and later applications of the process have generally utilized a ceramic substrate. The latter is a common material used in the construction of film resistors because it has certain advantages, particularly greater mechanical strength, and also because it can be fabricated into shapes which may be specially suitable for resistors.

† Welwyn Electric Ltd., Bedlington, Northumberland.

## 2. The Process of Deposition of Tin Oxide Films

The film is highly crystalline and Fig. 1 is an electron micrograph which shows the crystalline structure of the film most convincingly.

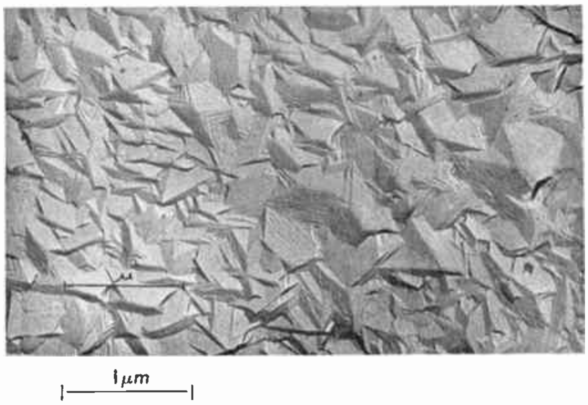



Fig. 1. Electron micrograph of tin oxide films.

There are several methods of application of tin oxide to the substrate. In most cases a vapour or dispersion of a tin compound together with such additives as the process requires is applied to the heated substrate. It is found that the nature of the dispersion, the ambient temperature, the temperature of the substrate and the application time all have an effect upon the properties of the film. Combining these variables with possible variations of the composition of the vapour or dispersion it can be seen that there is a most complex system to be studied in order to select the optimum process.

Processes which have been described in the literature involve the application of a mist of tin chloride, which has been dissolved in a solution, to a heated rod or tube, in one case directly drawn from a glass melting furnace. Other processes involve heating the bodies after fabrication and subjecting them to a fine spray or vapour. At least one application involves immersing the heated articles to be coated into a non-aqueous solution of tin and other materials.

There is much to be said for the use of ceramic for the substrate since it is possible to raise this material to temperatures well over 1000° C. It has been found that some of the physical properties of the film have a marked temperature dependence and the great range of temperatures which ceramics withstand gives very much more scope for study. For example, it is found that both temperature coefficient and resistivity are functions of the deposition temperature of the film.

### 3. Properties of Tin Oxide Films

In one commercial process resistivities of 1000 ohms per square can be achieved with film thicknesses as high as 10 000 Å. This compares with a film thickness of 250 Å for carbon films of the same resistivity, while in the case of metal films the corresponding film thickness would be less than 50 Å. This fact alone would account for the robustness of this tin oxide film.

A further feature is that the film is so hard that it is necessary to modify helical grinding techniques which have hitherto proved satisfactory for pyrolytic carbon resistors. The hardness appears to be little less than that of diamond and this characteristic would also account for the great robustness of the film.

Further advantages can be shown. For example, the film is essentially an oxide and therefore capable of running at high temperatures in an oxidizing atmosphere. This is in contrast to carbon and metal films which are alternative materials for film resistors. The upper limit of working temperature in a carbon film resistor is around 200° C, for metal film resistors it is 250 to 300° C. In the case of oxide films the figure is probably 450° C. The limit is in all cases very much bound up with the stability and reliability required. Thus, fault conditions which can give rise to high temperatures of short duration cause little or no damage to an oxide film resistor. Furthermore, any irregularity in the film which gives rise to a hot spot will not lead to progressive deterioration and eventual thermal run-away. This property, together with the robustness of the film, very largely explain the great rarity of an open circuit resistor.

The material is also highly resistive to the effects of moisture. It is possible to boil oxide film resistors for long periods with very little change of resistance, in fact the resistance seems to be asymptotic to some limit. This is unlike the behaviour of metal and carbon films under similar conditions. Study of the behaviour of freshly deposited tin oxide films shows that there is an initial period during which the resistance of the film is affected by oxygen and moisture. Observations show that the effect of oxygen and water are inter-related but that the film rapidly achieves a stable condition, and after this settling down period they have little further effect.

The film is damaged in the presence of moisture and a voltage gradient. The electrolysis which occurs in a surface film of moisture produces nascent hydrogen at the cathode with consequent reduction of the oxide. This corresponds to the deterioration in metal and carbon films where oxidation occurs at the anode by the formation of nascent oxygen. It is important to note that the reaction rate is very much slower in the case of oxide films. The ratio under identical conditions for metal, carbon and oxide films being of the order of 1000 : 100 : 1. It is also necessary to remember that because of its high resistivity tin oxide is present in a greater amount for a given resistance value so that the effect of moisture is even further reduced. This fact shows the reliability of oxide films under conditions which make it necessary to use heavy protection or even hermetic sealing for carbon and metal films.

The temperature coefficient of resistance is very much a function of the composition of the film and the physical conditions of deposition, in particular the temperature. Accordingly, there are likely to be substantial differences between manufacturers with regard to this parameter. It is found that the general pattern will be similar though there will be differences in detail.

For one type of tin oxide film the relationship of temperature coefficient with surface resistivity is shown in Fig. 2. There is some scatter about the mean value. This scatter of temperature coefficient is generally acceptable over most of the range, but in order to meet a requirement of $\pm$ 350 parts in $10^6$ per deg C, greater control is required at the highest resistivities.

The relationship between resistance and temperature is as shown in Fig. 3. This applies to typical values of surface resistivity.

Oxide films tend to have a fairly high contact resistance and it is therefore necessary to keep this

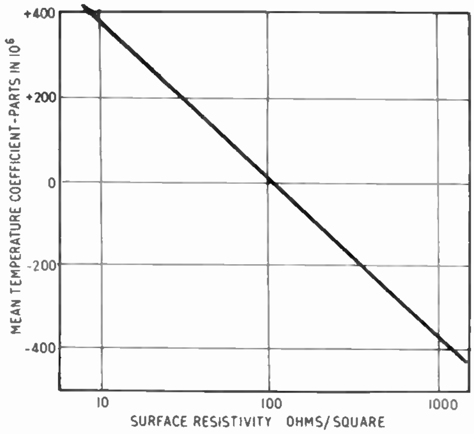



Fig. 2. Relationship between temperature coefficient and resistivity of tin oxide films.

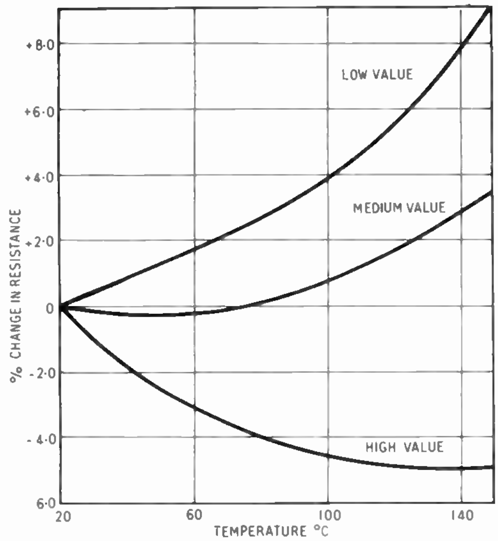



Fig. 3. Relationship between resistance and temperature for tin oxide films.

constant or to overcome it. For this reason in the conventional form of resistor very tight fitting caps are required, and this can, of course, be best achieved by the use of "machine-turned" caps rather than the simple "pressed" type. In addition, the ends can be prepared to reduce contact resistance, for example by the use of intermediate metal coatings. A particularly satisfactory method of making contact is to coat the ends with metal by electroplating or a similar process or by firing on a metal to which the usual wire terminal leads are soldered or brazed.

As the film is a semi-conductor it will give rise to the typical current-noise of such materials. This noise is difficult to measure except on the higher resistance values. The electrical noise developed by the film can easily be masked by the noise generated by a poor contact to the specimen. This danger is a direct result of the high contact resistance mentioned in the previous paragraph. With a good contact the noise level is extremely low and is less than that of a pyrolytic carbon resistor of similar value. Even for the highest resistance values on small size specimens the average figure is well below 0·5 $\mu$V/V in a decade.[6]

Similarly it is to be expected that the film would exhibit some degree of non-linearity of resistance with voltage. This effect is often referred to as "voltage coefficient". Such a description is misleading since the rate of change of resistance with applied voltage is not constant but is greater for small voltages than for higher voltages. The effect is a function of the electrical stress in the film and will depend upon the dimensions of the test piece and so is a function of the size and shape of the resistor. In general the change of resistance is less than 0·1% over 100 volts.

### 4. Properties of Tin Oxide Film Resistors

The film is applied to rods or tubes of a vitrified material such as glass or ceramic and can be fabricated into resistors in a way which is already well established in the field of carbon and metal film components. The customary method is to grind or otherwise cut a helix through the film into the substrate to increase the aspect ratio of the film. This enables high values of resistance to be obtained and, at the same time, enables the operator to adjust the value of resistance very closely.

Having produced a resistor it does not necessarily follow that the properties are exactly those of the initial film. The process of cutting the helix and the application of a protective coating can modify the properties of the film. In addition, the method of application of terminations will also enter into the performance of the component.

The operating temperature of the resistor will have some effect upon its performance. As the hot spot temperature varies with the size of the resistor (most resistor specifications are illogical in this respect) the stability of these resistors varies with rating. The stability of the film is also dependent upon the nature of the substrate, particularly the condition of the surface and on the presence of mobile ions at high temperature.

Tables 1 to 4 show typical performances of components of different sizes for different applications.

#### 4.1. *Precision Film Resistors* (*Tables 1 and 2*)

These are intended to be used for purposes where the initial tolerances and long-term stability are of the order of 1%, and are generally regarded as competitive with pyrolytic carbon films.

In order to achieve the desired quality special care is taken during manufacture; the caps are of better quality and precautions are taken regarding contact resistance. They will probably have better quality helical grinding and will be probably subject to some sort of measurement of temperature coefficient. They may also be subject to an initial ageing process which is directed towards improving stability.

**Table 1**

Properties of Typical Precision Film Resistors

| Resistance | Climatic Test to R.C.S. 114 | Stability at 70° C | | Temperature Coefficient (parts in 10⁶ per deg C) |
|---|---|---|---|---|
| | | Initial | Long Term per 1000 hr | |
| Low | 0·3% | 0·3% | 0·02% | −100 to +400 |
| Medium | 0·6% | 0·5% | 0·05% | −250 to +250 |
| High | 1·0% | 0·6% | 0·1% | −400 to +100 |

The resistance values corresponding to the resistivities in column 1 are set forth in Table 2.

**Table 2**

| | RO1C | RO1B | RO1A |
|---|---|---|---|
| Low Resistivity | ⩽ 10 kΩ | ⩽ 20 kΩ | ⩽ 50 kΩ |
| Medium Resistivity | 10–40 kΩ | 20–80 kΩ | 50–200 kΩ |
| High Resistivity | 40–200 kΩ | 80–400 kΩ | 200 kΩ–1 MΩ |
| Dimensions (inches) | $\frac{5}{8} \times \frac{13}{32}$ | $1\frac{1}{16} \times \frac{3}{8}$ | $2\frac{3}{16} \times \frac{3}{8}$ |
| Rating (watts) | $\frac{1}{2}$ | 1 | 2 |

4.2. *Power Resistors* (*Table 3*)

These components are intended to replace wire wound resistors, generally those protected with cement or vitreous enamel. Because there will be substantial temperature rises, i.e. up to 250° C, and temperature coefficient is not negligible, changes on load may be as high as 10%. It is thus not usual to supply them to better than ±5% tolerance. Furthermore there is no strong incentive for supplying a quality which will give a stability better than 1% or 2%. Among the applications for such resistors there are many relating to the domestic electrical or electronics field, and a low price is frequently a firm requirement so that the construction tends to be simpler with lower inspection standards.

**Table 3**

Properties of Power Resistors

| Rating (watts) | Maximum length (inches) | Hot spot temperature rise (degrees C) |
|---|---|---|
| 4 | $\frac{15}{16}$ | 190 |
| 6 | $\frac{15}{16}$ | 200 |
| 8 | $1\frac{3}{4}$ | 250 |
| 10 | $2\frac{1}{8}$ | 250 |

All resistors are $\frac{3}{8}$ in. maximum diameter.

Such resistors are usually required in low resistance values only. The apparent limitation of range of resistance is due to the nature of the demand rather than because of limitations of the film.

4.3. *Miniature General Purpose Resistors* (*Table 4*)

In this country film resistors are generally regarded as precision components and the use of films for general purposes has not been common. Tin oxide film has now been applied to a range of resistors intended to be used for general purposes. The ability of the film to operate at high temperatures means that the sizes of such resistors for a given rating are rather smaller than the carbon composition type of resistor which has heretofore served this purpose.

**Table 4**

Properties of Miniature Film Resistors

| Rating watts | Load stability at 40° C | | Climatic | MIL–R–11 size (inches) |
|---|---|---|---|---|
| | Initial | Long term per 1000 hours | | |
| $\frac{1}{2}$ | Up to 2% | 0·01–0·05% | 3% | 0·162 × 0·416 |
| 1 | Up to 1% | 0·02–0·05% | 2% | 0·240 × 0·593 |
| 2 | Up to 1% | 0·02—0·1% | 2% | 0·336 × 0·728 |

## 5. Conclusions

The foregoing shows the general properties of three ranges of resistors utilizing tin oxide as a conducting element. It has been shown that this film has some advantages over those previously used. The particular properties which make this component appeal to the electronic engineer are:

(*a*) Freedom from catastrophic failure.
(*b*) Good moisture resistance.
(*c*) Resistance to mechanical damage.
(*d*) Ability to operate at a high temperature.
(*e*) Low temperature coefficient.

One is tempted to make some prediction as to the ultimate quality which may be achieved by further development work on these resistors. It is probable that further improvements will be possible by modifications to the composition of the films as well as by closer control of the process. The following list of requirements is by way of a target and there is some reasonable expectation that these qualities will ultimately be achieved.

(*a*) Initial stability of the order of 0·1 to 0·2% according to resistance value.
(*b*) Long term stability 0·01 to 0·02% per 1000 hours according to resistance value.
(*c*) Control of mean temperature coefficient to ±150 parts in $10^6$.
(*d*) Control of scatter of temperature coefficient to ±100 parts in $10^6$.
(*e*) Extend the range of stable resistivity up to 10 000 ohms per square.

With these improvements this type of component would seem to meet all the general requirements for resistors for electronic equipment.

## 6. References


1. U.S.A. Patent No. 2,429,420.
2. British Patent No. 604,878.
3. British Patent No. 639,561.
4. British Patent No. 814,674.
5. L. Holland, "Vacuum Deposition of Thin Films", p. 496 (Chapman & Hall, London, 1956).
6. P. L. Kirby and R. H. W. Burkett, "Units for current noise", *Electronic Engineering*, **32**, pp. 412–3, July 1960.


*Manuscript received by the Institution on 2nd August 1960 (Paper No. 624).*

U.D.C. 621.396.963.3: 621.396.677

# A Fast Electronically Scanned Radar Receiving System

By

D. E. N. DAVIES, M.Sc., Ph.D.†

*Presented at a meeting of the Radar and Navigational Aids Group held in London on 1st February, 1961*

**Summary:** This paper describes some of the properties of an electronically scanned receiving system at very high scanning rates. The basic principles of the system are applicable to either an electromagnetic or an acoustic receiving aerial system, but the particular system described here relates to a microwave aerial in the 3 cm band, and is capable of scanning rates up to 1 Mc/s. The possible application of these principles to "within-pulse" scanning echo-ranging systems is also discussed.

## 1. Introduction

It is well known that the transmitting or receiving beam of a linear array may be deflected from the normal position by means of linear progressive phase-shifts along the length of the array. It is convenient to refer to the main lobe in the directivity pattern of a receiving array as a receiving beam although there is no physical quantity in the medium other than a "beam of sensitivity". The system described here scans a receiving beam by suitable variation of the phases of the array elements. These phase-shifts are produced in a delay line carrying a frequency-modulated i.f. signal. This system was first described for an underwater acoustic echo-ranging system[1] and a later paper discussed briefly the possible applications to radar.[2]

Although the proposed system has been successfully used in an underwater sonar system,[9] this paper is concerned with some of the absolute limitations of the system, which take the form of beam distortions and signal distortions and occur at high scanning rates. The theoretical analysis of these distortions may be considered valid for both radar and sonar systems, though the experimental scanner used to confirm the analysis takes the form of a microwave aerial in the 3 cm band. Although the choice of the medium used to verify the theory is purely a matter of convenience, the use of a radar frequency has provided additional information about the phase accuracies required from the microwave components, and the results obtained with the experimental radar scanner are encouraging.

In principle this system of scanning applies both to transmitters and to receivers, but with present techniques such practical considerations as high transmitter powers limit the application to receivers. This disadvantage can be overcome in certain applications by the use of "within-pulse" scanning of the receiving beam as discussed in Section 3. This clearly calls for very high scanning rates and consequently has led to an investigation of the theoretical and practical limitations involved in this problem.

† Department of Electrical Engineering, University of Birmingham

## 2. Principle of the Proposed Scanning System

Figure 1 shows a schematic diagram of the scanning receiving aerial. The aerial array is divided into a number of sections, each one of which feeds a separate channel terminating in a crystal mixer. The incoming carrier frequency is constant in the medium but is mixed with the output of a frequency-swept local oscillator common to all channels. The difference-frequency modulation product at the output of each mixer is fed to a tapping point on a delay line which has a linear phase characteristic over the i.f. band. The relative phase-shift introduced within this line between adjacent tapping points will be linearly related to the frequency of these signals and hence to the frequency of the local oscillator.

Therefore a frequency-shift of the local oscillator may be used to produce an angular displacement of the receiving beam in the medium. Signals in the medium having the form of an amplitude-modulated carrier are still available in this form at either output from the delay line. It is interesting to note here that the two separate ends of the delay line represent beams deflected to the opposite sides of the normal.

If the array is directly illuminated with a c.w. transmitter, and the beam is scanned continuously in some manner, the delay line output after amplification and detection represents the dynamic polar diagram of the aerial, with the main lobe in the direction of the transmitter for a given modulation rate of the local oscillator. This pattern may be displayed using the control signal of the frequency-swept oscillator as the time base, as shown in Fig. 1.

The directivity pattern of this aerial will be the product of two factors:[1]

(1) The pattern for $n$ point-receivers equally spaced.

(2) The pattern of one individual aerial element.

The diffraction pattern of $n$ point-receivers spaced a distance $d$ apart has principal maxima at angles $\sin^{-1}(m\lambda/d)$ where $m$ is any integer. A linear phase-shift along the array produces a deflection of this pattern relative to the directivity of individual elements. It is convenient to assume that the deflection is linearly related to phase-shift, this being approximately true provided that large angular deflections are not used. There is also a slight increase of beamwidth with

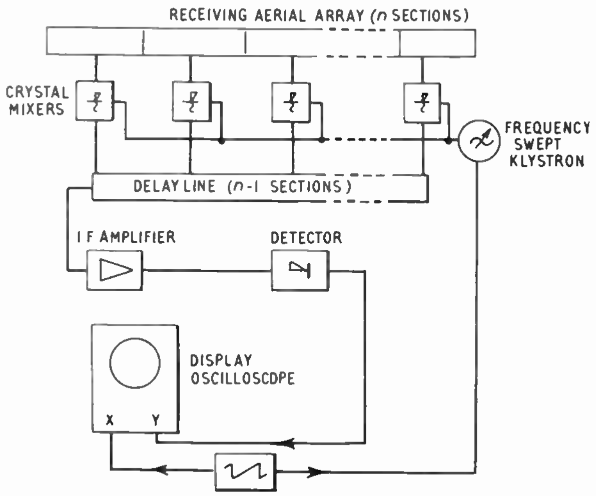



**Fig. 1.** Electronic scanning receiver.

deflection. Therefore the limit of useful scan of one beam corresponds to the angular spacing of the main lobes of the diffraction pattern,[1] i.e. $\pm \sin^{-1}(\lambda/2d)$. This represents $n$ beamwidths, where $n$ is the number of channels, since the array beamwidth is $\sin^{-1}(\lambda/nd)$. This sector has to be reduced to $(n-1)$ beamwidths in most practical cases in order to remove any ambiguity from diffraction secondary lobes.

It was shown in reference 1 that the phase change required to produce a deflection of $\sin^{-1}(\lambda/d)$ is $2\pi$ radians per channel. If the delay per section of the delay line is $t_2$ then

$$t_2 = \frac{d\beta}{d\omega} = \frac{2\pi}{\Delta\omega} = \frac{1}{\Delta f} \qquad \text{......(1)}$$

where $\beta$ = delay line phase-shift section.
and $\Delta f$ = frequency deviation of the local oscillation.

It would appear therefore that the phase-shift required for scanning can be produced either by a large frequency sweep together with a small delay or vice versa. This is true in practice for low scanning rates, but as the scanning rate is increased, restrictions arise imposing a minimum limit to the frequency sweep. These are discussed later.

### 3. "Within-pulse" Scanning

Conventional pulse radar systems scan both the transmitter and receiver together in synchronism, often the same aerial being used for both transmission and reception. Under these conditions there exists a maximum rate of scanning which corresponds to a rotation of one beamwidth in the time taken for a signal to reach maximum range and return. This corresponds to one received pulse per target, per beamwidth of scan. Once this limit has been reached further increase of scanning rate will cause echoes from the maximum range to be lost. It is possible to use a radar transmitter operating at two or more frequencies to improve this situation[7] but this is really a method of combining the outputs of two or more radar systems together to give an increased information rate.

"Within-pulse" scanning is also a method of combining several radar systems together to provide increased information but without the complexity of several transmitters operating at different frequencies. If we consider $n$ separate radar receivers with separate beams fixed in space to cover some sector or solid angle, then the video output of these systems could be sampled in turn by a suitable high-speed sampling switch, and the resultant waveform displayed. Provided that the sampling were carried out at a sufficient rate,[3] this process would combine the independent outputs of $n$ systems and thus increase the information rate $n$ times, for an ideal sampling switch. Under these circumstances the entire sector could be illuminated with one wide-beamwidth pulse of energy.

The proposed method of within-pulse scanning goes a stage further than this in using the receiving beam of an electronically scanned radar system to act as a sampling switch in the medium. This system has been used successfully in underwater acoustic echo-ranging systems to provide an increase of information rate.[9] It involves the illumination of a sector with a wide-beamwidth pulse of energy. A receiving beam, $n$ times narrower than this sector, is then scanned continuously across the sector with the angular deflection having a saw-tooth form. If the duration of each scan is made equal to the duration of the transmitted pulse then this beam will sample each returned echo, and no information will be lost in this process. The duration of each sampled pulse will be $1/n$th of the duration of the transmitted pulse.

Range and bearing information remains unaffected by the sampling since there is no knowledge of the part of the incoming pulse to be sampled. The output waveform after sampling has the form of a dynamic directivity pattern of the aerial array, and therefore the actual sampled pulse consists of the main lobe of this pattern. All the normal theory of arrays (such as the effect of amplitude tapering) still applies and can be used to control the shape of the directivity pattern.

It is now evident that $n$ times as many pulses are received in a given period of time compared with a conventional scanning system, since information is now available concerning $n$ different beamwidths in space for one transmitted pulse. This represents a gain in information rate provided that the signal/noise ratio is not impaired.[4] This latter situation is considered in Section 7.3.

The display problems arising with this form of scanning are not fundamentally different from those in conventional radars except that the bandwidths are necessarily greater. Since the electronic scan is usually restricted to a sector less than 180 deg it may be convenient to employ a B-scan display, because the control signal from the swept oscillator can also be used as the bearing deflection signal of the display.

The radiation pattern of the transmitting aerial must illuminate the entire scanned sector, and careful control of the shape of this pattern can improve the side-lobe level of the resultant system pattern, which is the product of the transmitting and the receiving aerial patterns.

## 4. The Effect of Fast Scanning

### 4.1. *Fundamental Restrictions on Scanning Rates*

At very high scanning rates three distinct types of distortion become apparent. Two of these occur even in the reception of a c.w. signal from the medium, and are associated with malformation of the receiving beam. The other occurs only in the reception of pulsed signals and results in a distortion of the shape of the sampled-pulse waveform.

It is interesting to consider first why any such restrictions upon scanning rates should arise. To do this we consider the scanning receiver as a sampling switch. A sample of duration $\tau/n$ is taken from a pulse of duration $\tau$. Hence the bandwidth of the sampled pulse has increased $n$ times, and this increase must take place somewhere in the system before the output of the delay line.

Since the delay-line is a linear, passive network with constant parameters it cannot increase the spectrum of signals passed through it. Also the bandwidth of a single modulation product of a linear modulator cannot exceed $(W_1+W_2)$ where $W_1$ is the bandwidth of the amplitude-modulated input signals, and $W_2$ is the bandwidth of the frequency-modulated switching function.

Then

$$W_1 + W_2 \simeq \Delta f + \frac{1}{\tau} \qquad \ldots\ldots(2)$$

It is therefore evident that the source of the additional bandwidth must be the swept-frequency oscillator, and consequently

$$\Delta f + \frac{1}{\tau} \geqslant \frac{n}{\tau}$$

Therefore

$$\Delta f \geqslant \frac{(n-1)}{\tau}$$

or using eqn. (1)

$$W_s = \frac{1}{\tau} \leqslant \frac{\Delta f}{(n-1)} \qquad \ldots\ldots(3)$$

where $W_s$ = scanning rate.

The limitation is therefore upon the scanning rate for a fixed value of $\Delta f$.

### 4.2. *Forms of Distortion of C.W. Signals*

#### 4.2.1. Type A distortion

For a c.w. input to the aerial, the signal fed into each section of the delay line will consist of a frequency varying linearly with time during one scan. Additional phase-shift terms may be neglected by considering the received wave to arrive from a direction normal to the array. Let this signal be represented by $\sin(\omega t+xt^2)$.

Then the delay line output will be the sum of $n$ terms:

$$\begin{aligned}&\sin(\omega t+xt^2)\\&+\sin[\omega(t-t_2)+x(t-t_2)^2]\\&+\ldots\ldots\\&+\sin[\omega\{t-(n-1)t_2\}+x\{t-(n-1)t_2\}^2]\end{aligned}$$

The $r$th term of this series can be rewritten

$$\sin[(\omega t-xt^2)-2x(r-1)t_2\,t-(r-1)\omega t_2+x(r-1)^2t_2^2]$$

which is of the form

$$\sin(\omega t+xt^2+\phi_1+\phi_2) \qquad \ldots\ldots(4)$$

where

$$\phi_1 = -(r-1)t_2(\omega+2xt) \qquad \ldots\ldots(5)$$

$$\phi_2 = x(r-1)^2t_2^2 \qquad \ldots\ldots(6)$$

$\phi_1$ represents the required form of phase modulation for scanning the beam, since it varies linearly with the number of the channel and also with time.

Since the aerial elements are equally spaced, $\phi_2$ represents a quadratic phase error across the aerial aperture which is constant with time but varies with scanning rate. It is only significant at high scanning rates, and clearly can be compensated by fixed phase-shifts if the scanning rate is constant. This causes the phase front or equi-phase surface of the aerial to become quadratic in shape. The effect of this on the beam is a slight angular displacement, an increase in beamwidth and a higher relative sidelobe level.

### 4.2.2. Type B distortion

Figure 2 shows the $n$ components of the delay line output in their relative positions in time. Each block represents one period of modulation of the swept oscillator (i.e. one scan of the beam) with a c.w. input to the array. It can be seen from the diagram that between A and B, signals from one scan overlap signals from the adjacent scan, and this phenomenon must result in a distorted output during this time. The physical interpretation of this effect is that the equi-phase surface for the array does not fly back in one movement after the completion of a scan. The "flyback" occurs in $n$ steps as the elements change their phase one at a time. The duration of this distortion can be seen to be $(n-1)t_2$, (the total delay in the delay line) and this time must therefore be made small compared with the modulation period. If we confine this distortion to a period $\tau/k$ per scan, then $(n-1)\,t_2 \leqslant \tau/k$. Therefore, using eqn. (1),

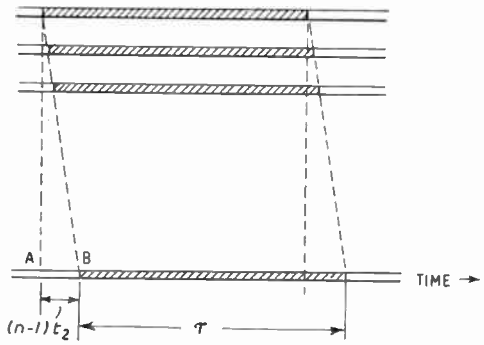



**Fig. 2.** The $n$ components of the delay line output for c.w. input to the array.

$$\Delta f \geqslant \frac{k(n-1)}{\tau} \quad \text{or} \quad t_2 \leqslant \frac{\tau}{k(n-1)} \qquad \ldots\ldots(7)$$

It is interesting to note that this restriction is of exactly the same form as that obtained in Section 4.1 and differs only by a factor $k$. In practice it may be reasonable to confine this distortion to one beam-width of scan since in many cases this angle would have to be sacrificed due to a diffraction secondary-lobe just entering the edge of the sector, as discussed in Section (2). This case results in $k = n$.

This beam distortion may be reduced or even eliminated by disregarding the conditions in eqn. (1), and increasing the phase modulation such that the beam is scanned beyond the normal sector. The distortion can then be arranged to occur when the main beam is outside the illuminated sector and in a position which could not normally be used because of diffraction secondary lobes. This has the effect of reducing the duration of the samples taken from the incoming pulses. The distorted sector may be blanked out on the display.

When "overscanning" of this kind is used to reduce beam distortion at the edges of the sector, eqn. (1) is replaced by

$$\Delta f = \frac{1}{t_2} + x(n-1)t_2 \qquad \ldots\ldots(8)$$

The frequency sweep $\Delta f$ consists of two terms: the first term is that part of the frequency sweep used during the undistorted scanning of the sector; the second term represents the increased frequency sweep needed to keep the distortion outside the required sector.

$$\Delta f = \frac{1}{t_2} + \frac{\Delta f}{\tau}(n-1)t_2$$

$$\Delta f\left[1 - \frac{(n-1)t_2}{\tau}\right] = \frac{1}{t_2}$$

$$\Delta f = \frac{\tau}{t_2[\tau-(n-1)t_2]} \qquad \ldots\ldots(9)$$

The minimum value of $\Delta f$ is given when

$$\frac{\mathrm{d}}{\mathrm{d}t_2}\{t_2[\tau-(n-1)t_2]\} = 0$$

i.e. when $$t_2 = \frac{\tau}{2(n-1)}$$

Substituting in (8) for $t_2$ we have:

$$\Delta f = \frac{4(n-1)}{\tau} \qquad \ldots\ldots(10)$$

It can be seen from the above that the best that can be obtained with overscanning is $k = 4$.

### 4.3. *The Distortion of Signal Pulses*

In any electronic scanning system the aerial array is not always parallel to the wave front. If this aerial is divided into sections and connected to separate channels then there will be a progressive time delay between the excitation of each channel by the incoming signal. In the proposed system this will cause time delays between the signals in the various channels, and these delays also vary with deflection. The function of the delay line is to impart phase-shifts to the signals such that the outputs from the $n$ channels add in phase at a position in the bearing scan determined by the angle of deflection of the beam. As a result of this each channel is subjected to different amounts of delay depending upon the number of delay line sections traversed. Therefore output signals are affected by two sets of delays, the first in the medium and the second in the delay line.

Hence, instead of $n$ pulses being coincident in time at the output of the delay line, they will be displaced relative to one another and the resultant pulse will be elongated at the leading and trailing edges. In the limit, with sufficient delay in the delay line, the output would be $n$ separate pulses with no overlapping. Clearly this situation must be avoided. When the receiving beam is scanning continuously the envelope of the delay line output represents a sample of the received pulse. If this sample is taken from the distorted leading or trailing edges of the pulse then there is a severe risk that the sampled pulse may not be detected. The probability, $p$, of losing a pulse due to this distortion on a given scan, is the ratio of the duration of the distortion to the duration of the scan, that is,

$$p = \frac{(n-1)t_2}{\tau} \qquad \ldots\ldots(11)$$

If we substitute in this formula the value of $t_2$ given by the equality sign in eqn. (7) we get $p = 1/k$. So that if we make $k = n$, as suggested in Section 4.2.2, we lose one pulse in $n$ on the average.

The above approach is really only approximately valid, and a more rigorous study shows that there can be transient distortions during the sidelobe output also. Consider the first component pulse arriving at the delay-line output sometime during a main lobe part of the scan. This means that the wavefront arrives at the array just when the beam "looks" in that direction. At the centre of the main lobe all the signals add together in phase, it can therefore be surmised that the effect of the time-staggering will be to produce a staircase, stepped wave-front to the pattern.

Now consider this stepped effect to start at a zero of the pattern. A zero of a directivity pattern occurs when all the component vectors representing the output of each element, combine to form a closed vector polygon. Therefore, with the exception of special cases of polygons, where vectors are coincident, there cannot be a zero until all the components are present. Therefore if the wave arrives in the region of a zero of the pattern, there will result a transient distortion which can increase the sidelobe level at this point.

The extent of such distortions may be calculated from the vector diagram of the directivity components.[8] Figure 3 shows five joined vectors, if they are given angular velocities $\Omega$, $2\Omega$, . . . $5\Omega$, then the modulus of the line joining O to the point $A_5$ represents the directivity pattern of a 5-element array. Zeros of the pattern are caused by a closed polygon. The transient sidelobe level at any instant may be found from the vector resultant of the appropriate number of vectors in the correct position in the cycle.

For an electromagnetic aerial array the time delays in the medium are very much less than those in the delay line. The two delays become comparable when the frequency sweep is of the same order as the carrier frequency in the medium. Therefore for almost all conceivable radar applications these delays can be neglected. In a sonar system, however, the slower velocity of propagation of acoustic waves makes these delays much more important. Consequently the practical sonar systems built so far have employed frequency sweeps much greater than the carrier frequency in the medium, so that the medium delay is the most significant delay.[1] Some small degree of improvement can be obtained by employing time-varying delay lines for scanning, at the expense of further complication, and this matter is still being investigated.

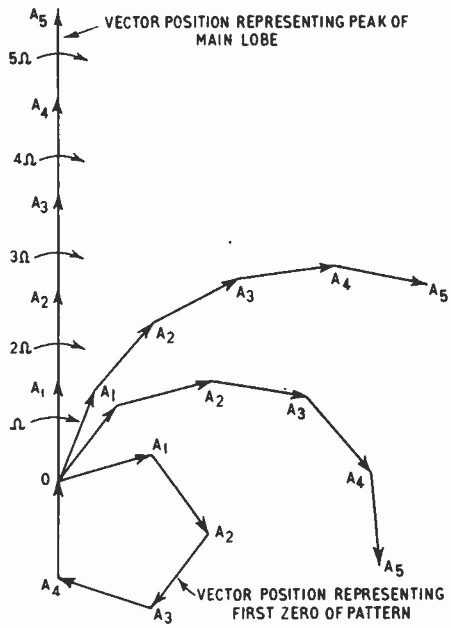



Fig. 3. Vector components of delay line output.

### 4.4. *The Reduction of Transient Distortion*

There are two methods for the complete elimination of the effect of transient distortion due to the delay line. One method is to place the delay line in the feed from the swept oscillator and add all the channel outputs in-phase. This is equivalent to imposing the phase modulation on the switching signal of the mixers rather than on the input signal. There are then no delays in the channels to cause distortions, and this modification does not affect the c.w. distortions.

The alternative method of eliminating transient distortion is to insert delays in each channel prior to the mixers, such that the total delay from the array elements to the delay-line output is constant in each channel.[2] This modification also has no effect upon the c.w. distortions.

### 4.5. *Other Methods of Electronically Scanning a Linear Array*

The technique of within-pulse scanning is clearly not limited to the method described here for scanning a receiving beam. Any method of scanning may be used provided that a sufficiently high scanning rate can be obtained. Scanning systems employing frequency modulation in the medium[16, 17] are subject to similar types of restriction upon scanning rate as the proposed system, since both work on the same basic principles of using frequency modulation and time delays to produce phase modulation. Frequency modulation in the medium will, however, result in a very wide medium-bandwidth, when applied to within-pulse scanning.

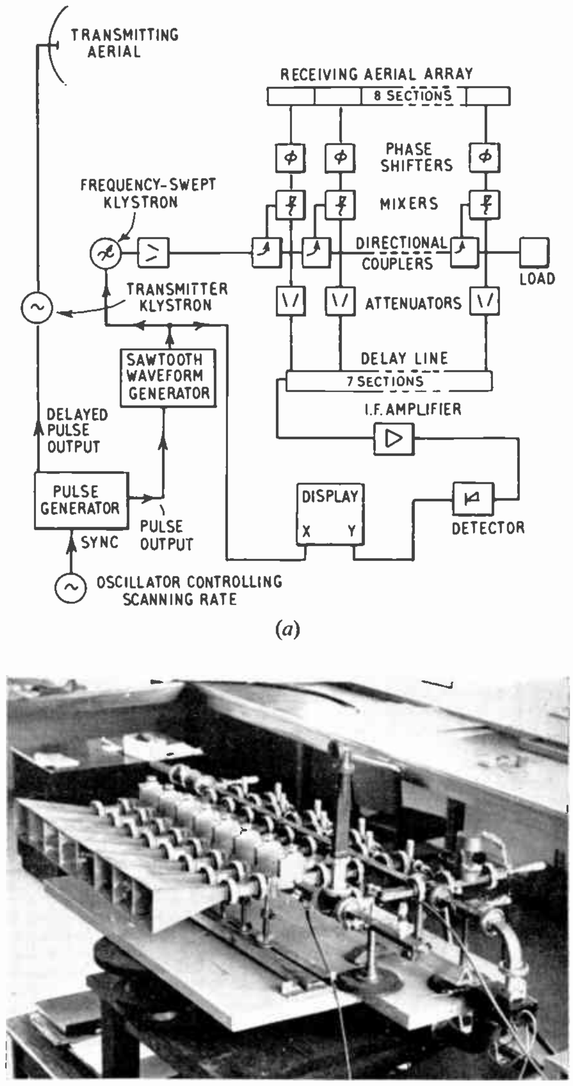
Receiving or transmitting beams may also be scanned by employing travelling-wave tubes to produce an electrically-variable time delay.[18,20] There is very little information in the literature on this method, but it would appear that although high scanning rates can be obtained, such a system must be rather complicated. A more popular method of scanning is to vary the polarizing magnetic field in a ferrite medium propagating a microwave signal.[18, 19] This method applied to within-pulse scanning would probably be limited to scanning rates up to about 1 Mc/s since it is very difficult to modulate ferrites at higher rates than this.[21]

## 5. The Experimental Scanning Receiver

The experimental scanning receiver was built in order to verify the proposed method of scanning, and to investigate practically the extent of the forms of distortion which have been predicted in the theoretical analyses. In order to simplify the experimental work a scanning receiving aerial was constructed and was directly illuminated by a point-source transmitter. This also provided a very convenient method of measuring the distortion of beam shapes at various scanning rates.

The aerial operates in the X-band and employs eight channels and a klystron local oscillator. A photograph of the aerial array and schematic diagram of the research equipment is given in Fig. 4. The aerial aperture is $2{\cdot}84\,\lambda$ by $22{\cdot}6\,\lambda$ at 9455 Mc/s. This represents beamwidths in the vertical and horizontal plane of 21·8 deg and 2·6 deg respectively. The maximum scanned sector, defined as eight beamwidths, is therefore 20·8 deg.

The receiving-aerial elements are horns which are followed by matching sections, dielectric vane phase-shifters and 90 deg twists in order to use broad-face-to-broad-face directional couplers. These directional couplers are spaced two guide-wavelengths apart on a single run of waveguide and enable power from the swept local oscillator to be fed to all the channels in phase. Cruciform slots are used and the coupling values graded in order to feed the same level of power into all channels. The channels terminate in mixers employing CV2154 crystals. The centre frequency of the swept local oscillator is 9375 Mc/s.




(*a*)

(*b*)

Fig. 4. (*a*) Schematic diagram of experimental scanning receiver. (*b*) The aerial array.

In order to ensure a reasonable degree of linearity of the frequency characteristic and as little amplitude modulation as possible, the frequency sweep of the local oscillator was made much smaller than the half-power electronic tuning bandwidth of the klystron.

The valve chosen was the V58 klystron manufactured by Varian Associates which, when operated in the $5\frac{3}{4}$ mode, provides a power output of 80 mW with an electronic tuning range of about 100 Mc/s. The frequency sweep was chosen to be 20 Mc/s.

In a practical system, frequency sweeps in excess of the above figure may be obtained using klystrons, and very large sweeps of the order of several hundred megacycles may be obtained from backward-wave oscillators.[5] The figure of 20 Mc/s was chosen in order to ensure that the forms of distortion under investigation would occur at scanning rates which could be conveniently displayed. Since it becomes difficult to produce saw-tooth waveforms with fast flyback at frequencies higher than 2 Mc/s, it was necessary to display the effects of distortion at a slightly lower scanning rate than this. For type B distortion, taking eqn. (7) with $k = n = 8$ we obtain a maximum scanning rate of 0·36 Mc/s (for $\Delta f =$ 20 Mc/s). Scanning rates of this order enable both forms of c.w. distortion to become apparent, and at about 2 Mc/s the entire pattern is lost. For type A distortion a maximum channel phase error of one wavelength corresponds to a scanning rate of 0·204 Mc/s.

Monitoring facilities enable the f.m. spectrum of the swept local oscillator and the transmitter line spectrum to be observed on a spectrum analyser. A small amount of the local oscillator output is also fed to a detector to observe the extent of the amplitude modulation of this signal. This enables the reflector voltage to be adjusted so that the modulation is symmetrical about the centre of the mode. This is necessary in order to keep both amplitude modulation and non-linearity of the frequency characteristic to a minimum. Facilities are also provided for injecting a signal at transmitter frequency into all the channels in order to simulate the effect of a wave arriving normal to the array.

The difference-frequency outputs from the mixers are fed to tapping points on a cable delay line. These T-junctions employ series resistors so that the 75-ohm cable impedance is not mismatched at the junction points. The double-screened cable has a velocity ratio of 0·73 and is 252 ft long between the first and last tapping points.

$$\text{Total delay} = (n-1)t_2 = \frac{(n-1)}{\Delta f} = 0{\cdot}35\ \mu\text{s}.$$

The delay line is matched at both ends and one end feeds an i.f. amplifier constructed by Decca Radar Ltd. The gain of this is 80 dB over the band 40–120 Mc/s. Although only a part of this band is used at present it is proposed to employ greater frequency sweeps at a later date. The final stage of this amplifier feeds a simple detector with a cathode-follower output. The detected output is fed directly to a display oscilloscope and the saw-tooth voltage used to modulate the klystron is fed to the X-plates of the display. For modulation rates up to 2 Mc/s the oscilloscope used for the display should have a Y bandwidth of about 25 Mc/s. The photographic results for fast scanning were taken on a Tektronix type 545 oscilloscope with a bandwidth of 30 Mc/s.

The output of the saw-tooth generator has a flyback time of 0·08 μs over the frequency range 0·3–2 Mc/s. A fast flyback is necessary in order to distinguish between the overlapping period of distortion and the flyback itself, which at high scanning rates represents a completely distorted scan.

The transmitter consists of a waveguide-fed parabolic dish, radiating about 30 mW continuously from an English Electric K340 Klystron. It is also possible to pulse the transmitter and to synchronize this pulse with the scanning saw-tooth in order to observe the effect of sampling the distorted leading or trailing edge of an incoming pulse as discussed in Section 4.3.

Since the formation of the beam depends upon the phase coherence from channel to channel, care must be taken in order to reduce phase errors between the aerial array and the delay line output. A 10 deg phase error represents 0·48 ft of cable at the centre of the i.f. band. The same phase-shift is produced in 0·05 in. of waveguide at 9375 Mc/s. The waveguide components consist mostly of low-grade, standard, commercial items which did not meet the requirements for phase accuracy, but the inclusion of variable phase shifters enables the phases to be adjusted as required. Amplitude adjustments are also provided in each channel in the form of attenuators in the i.f. inputs to the delay line.

If either the frequency of the transmitter or the centre frequency of the local oscillator sweep, changes due to such effects as temperature, there will be a resultant change of i.f. frequency and this corresponds to a bearing error. In the experimental system this error represents 1·1 deg/Mc/s change of frequency. This figure could be reduced in any such system either by increasing the frequency sweep or by employing automatic frequency control.

It is important to emphasize that the maximum useful scanning rates obtained with the experimental receiver are limited only by the design parameters, and that much higher rates could be obtained using larger frequency sweeps. The large sweeps obtainable from backward-wave oscillators[5] should make scanning rates of the order of 10–20 Mc/s quite feasible.

## 6. Experimental Results

### 6.1. *Continuous Wave Results*

Most of the experimental results consist of photographs of displayed radiation patterns under various experimental conditions. A static pattern is also included for reference. It is important to appreciate that the displayed patterns do not represent true "dynamic directivity patterns" but only approximate to them. The true dynamic directivity pattern must be defined as the relative sensitivity of the aerial in each direction at one instant of time. The measured dynamic patterns represent the variation of sensitivity in one direction with time, during scanning.

Figure 5 shows directivity patterns displayed at 1 kc/s scanning rate. The patterns are shown with the transmitter situated at several different bearings within the scanned sector. This is achieved by mechanical rotation of the electronically scanned receiving aerial. The extent of the pattern shown is about 21 deg, because although much larger deflections are possible this represents the useful limit of scan. It can be seen from these results that the maximum scanned sector is at the most eight beamwidths, and is limited by the presence of a diffraction secondary lobe.

The beamwidth of the displayed patterns can be seen to remain constant over the sector but the amplitude of the signals fall off towards the edge of the sector due to the radiation pattern of the individual aerial elements. This effect may be overcome by the inclusion of an equalizing network at the output of the delay line.

A

B

C

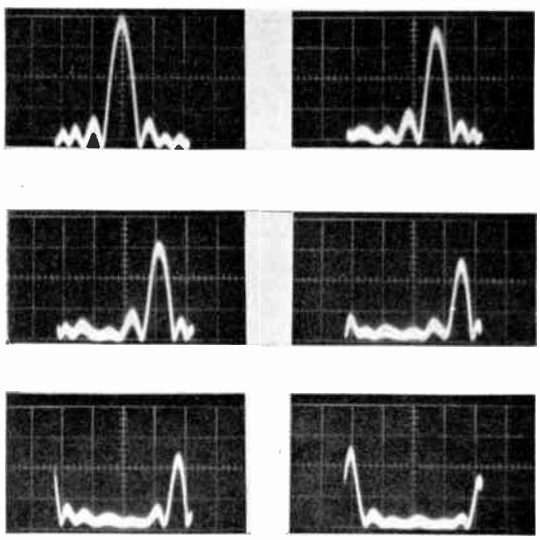


Fig. 5. Directivity diagrams at 1 kc/s scanning rate for several positions of the transmitter within the sector.

It can also be noticed from Fig. 5 that the relative sidelobe level remains constant with deflection. Now this represents a difference between the displayed pattern and the real sensitivity pattern in the medium. The displayed pattern reduces in amplitude as the

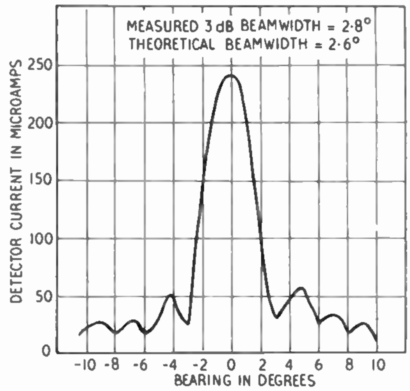



Fig. 6. Measured static directivity diagram.

transmitter moves across the sector, since the received signal falls off due to the directivity pattern of the individual horns. The true pattern in the medium is the product of the horn pattern and the diffraction pattern of $n$ point-receivers. Therefore since the diffraction pattern is deflected relative to the broad-beamwidth horn pattern, this deflection causes a deterioration in sidelobe level.

Figure 6 shows a static directivity diagram measured in terms of detector current. This measurement was made with the local-oscillator frequency constant, at the centre of its normal swept band. The static pattern is a true pattern of sensitivity against bearing, and the sidelobe level of the true pattern should theoretically be better than that of the displayed pattern in the zero bearing position, and worse at the edge of the sector.

#### 6.1.1. Fast scanning results

The effect of fast scanning upon the displayed pattern is shown in Fig. 7. The two types of c.w. distortions can be seen in these photographs, and it is clear that they both increase with scanning rate and eventually render the pattern useless. The quadratic phase taper (type A distortion) makes the main beam wider, causes it to be displaced from its normal position and also increases the sidelobe level. As the scanning rate is increased these three effects become much worse until at scanning rates approaching 1 Mc/s the pattern becomes unrecognizable.

The form of the overlapping c.w. distortion, (type B distortion) does not show up quite as obviously on the photographic results. The distorted sector is the

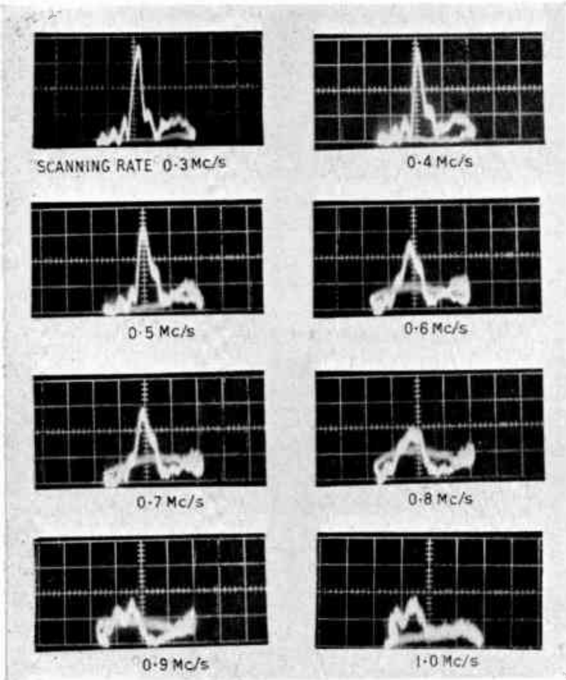



Fig. 7. Directivity patterns at high scanning rates.

extreme right hand edge of the photographs, where no coherent pattern is formed. As the scanning rate is increased this distorted sector intrudes further across the pattern because the duration of the distortion is fixed and equal to $(n-1)t_2+T_f$. That is the total delay in the delay line plus the flyback time $T_f$.

In order to compare the results of the quadratic phase-taper with theory, graphs have been plotted relating loss of height of main lobe, and bearing displacement of main lobe, to increase of scanning rate. Gazey[6] has calculated some beam shapes for a similar sonar system exhibiting the same form of distortion. In order to compare these results with the experimental radar results, the graphs have been normalized by plotting against the coefficient of the quadratic phase-shift term. Also the beam displacement has been measured in terms of original beam-width. These results are shown in Fig. 8.

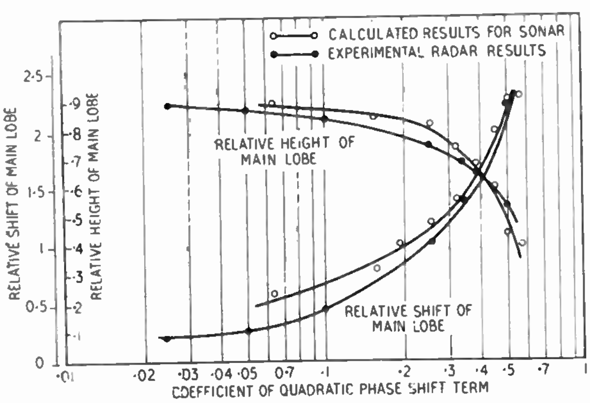



Fig. 8. Quadratic phase error in scanning systems.

### 6.2. *Results with Pulsed Signals*

The results for pulsed signals were obtained by pulsing the transmitter in synchronism with the scanning of the receiver. The transmitter was pulsed once each scan and it was possible to vary both the duration of the pulse and the position of the pulse relative to the start of the scan.

The results shown in Fig. 9 are for scanning rates in the range 0·3–0·5 Mc/s and for pulse lengths in the range 1–1·5 μs. When the pulse starts or stops in the region of the main lobe, stepped leading and trailing edges result as described in Section 4.3. It can also be seen that the directivity patterns remain undistorted between the stepped leading and trailing edges of the pulses.

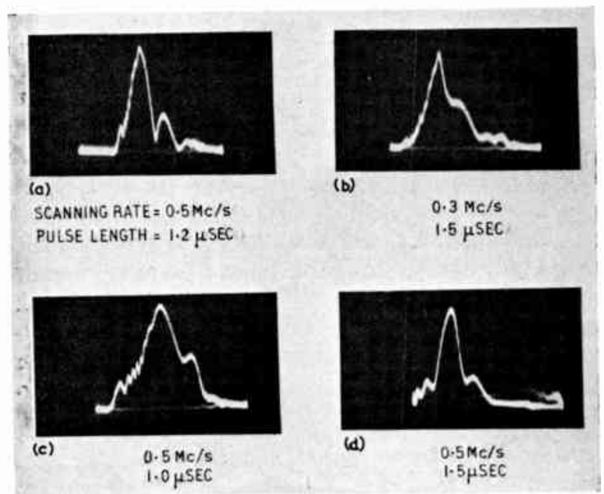



Fig. 9. Directivity patterns with a pulsed transmitter showing transient distortions.

When the pulse is switched on or off during a sidelobe or zero of the pattern, distortion of this part of the pattern results. In Fig. 9(*c*) for example the first sidelobe and first zero are distorted and part of the main lobe is also affected. It can be seen that during the main lobe the distortion takes the form of steps, while during sidelobes the distorted edges are more sawtooth in shape. This is the result that would be expected from considering the vector polygon discussed in Section 4.3.

In Fig. 9(*d*) both the switching on and off has been arranged to occur during the c.w. distorted sector at the edge of the pattern, or during the adjacent flyback period. Consequently hardly any of the transient distortion is apparent since it is masked by the c.w. (type B) distortion at the edge of the sector.

The experimental results indicate that the analysis given in Section 4 is satisfactory in calculating the features of distortion at high scanning rates.

## 7. The Application of Electronic Scanning to a Within-pulse Scanning Radar System

### 7.1. *Within-pulse Scanning and Information Sampling*

A conventional scanning radar system reaches a limit of information rate corresponding to the reception of one pulse per beamwidth of scan. Once this limit has been reached neither the bearing discrimination nor the scanning rate may be increased without sacrificing the other.

The proposed method of within-pulse scanning uses the receiving beam of the radar to act as a sampling switch in the medium. The scanning receiving aerial described in this paper has $n$ channels and a maximum deflection of $n$ beamwidths; we therefore expect the receiver bandwidth to increase $n$ times, but as was shown in Section 4.2.2 the actual increase is $kn$. It is important to realize however, that this wide bandwidth starts at the mixers and the system remains narrow-band in the medium, that is to say it has the same bandwidth as a conventional radar of similar transmitted pulse length. This is because the medium itself can be considered as $n$ separate communication channels.

Although a within-pulse scanning system provides $n$ times more pulses in a given period of time than a conventional radar, this alone does not necessarily give an increase of $n$ times in data rate, unless the probability of detection of the pulses at all ranges remains unimpaired. Consequently, comparisons of r.m.s. signal/noise ratios have been made between a conventional radar and a within-pulse scanning radar. But first it is important to clarify the difference between using within-pulse scanning for improving bearing accuracy and for increasing the scanning rate.

In a mechanical scanning radar of beamwidth $\theta_m$ there will be a maximum scanning rate $W_s$ corresponding to reception of one pulse per beamwidth of scan.

Let a within-pulse sector-scanning radar have $n$ channels and a sector beamwidth $\theta_s$. Then the receiver beamwidth $\theta_r = \theta_s/n$. If we now replace the conventional mechanical scanning radar with the sector scanning system, it may be rotated mechanically while it also executes within-pulse sector-scanning.

Consider the case $\theta_s = \theta_m$.

This situation results in the maximum speed of rotation remaining unchanged at $W_s$, but since the received beamwidth has been decreased $n$ times the bearing discrimination has improved $n$ times.

Now consider the case $\theta_r = \theta_m$.

This results in the bearing resolution remaining the same as for mechanical scanning but since there are $n$ times as many pulses available in a given period of time, the maximum rate of rotation can now be increased to $nW_s$.

The two cases just considered represent two extremes. Clearly the parameters may be chosen so that the gain of $n$ can be divided, theoretically without restriction, between bearing resolution and scanning rate. The distinction between the two cases is important however, since it is shown later that they represent different signal/noise ratios.

### 7.2. *Radar Receiver Noise and Detection*

A significant portion of receiver noise in most radar systems arises in the crystal mixers and the i.f. amplifier. This situation will also hold for electronic scanning systems, with the added disadvantage that the greater bandwidth will increase the receiver noise by a factor $nk$. This increase is due entirely to the bandwidth of the frequency sweep. The extra mixers do not increase the noise power because when matched for maximum power transfer the total noise power remains the same.

The output signal from the delay line has a frequency deviation $\Delta f \simeq kn/\tau$ but the detected bandwidth of a sampled pulse of duration $\tau/n$ need not exceed about $n/\tau$. Therefore the noise after detection can be reduced by means of a post-detector filter. However the spectrum of the noise output of a linear detector changes when a carrier tuned to the mid-band of the noise is added to the noise before detection.[10,11] The presence of the carrier causes the detected noise spectrum to change from a triangular to a rectangular shape.

Since the signal/noise ratio of a radar system is most important at maximum range, it is reasonable to consider the case of the very weak signal and assume a triangular spectrum of bandwidth $kn/\tau$. Therefore since the signal bandwidth is $n/\tau$ a post-detector filter will reduce the noise power after detection by a factor $k/2$, because the filtered portion of the triangular spectrum approximates to a rectangle.

### 7.3. *Signal/Noise Ratio Comparisons*

The comparisons of signal/noise ratios are divided into three cases which are analysed by considering the effect of sampling and filtering on the signals and the noise separately.

Let the receiver noise per unit bandwidth be $N$.

Let the peak transmitter power be $P$.

Let the peak received power be $KP$.

*Case* (1). The first comparison may be more of academic than practical interest, but is important because it represents a comparison between two systems with the same potential information rate. One system consists of $n$ fixed-beam receivers positioned to cover a sector of $n$ beamwidths, the receiver outputs are sampled at video frequencies by an ideal sampling switch and the resulting signal displayed. This system is compared with a within-pulse scanning radar covering the same sector and employing $n$ channels. Therefore the receiving beamwidth is the same for the two cases and the same wide beamwidth transmitter may be used for both systems.

Since the two systems employ the same transmitter and the same aerial gain, the received pulses will be at the same power level. The process of sampling does not affect the signal/noise ratio, except that for the proposed within-pulse scanning system more noise is introduced due to the wide bandwidth required in the r.f. and i.f. circuits. The actual comparison is shown in Table 1 and it can be seen that the signal/noise ratio of the electronically scanned system falls short of the idealized system by a factor $\sqrt{(2n)}$.

**Table 1**

Comparison between the signal/noise ratios of a within-pulse electronically scanned radar system and an idealized within-pulse sampling system.

| | *Within-pulse Scanning System* | *Idealized Sampling System* |
|---|---|---|
| Received power (peak) | $KP$ | $KP$ |
| Receiver noise power (before detection) | $\frac{nkN}{\tau}$ | $\frac{N}{\tau}$ |
| Receiver noise power after detection and filtration | $\frac{2nN}{\tau}$ | $\frac{N}{\tau}$ |
| R.m.s. signal/noise ratios | $\sqrt{\frac{KP\tau}{2nN}}$ | $\sqrt{\frac{KP\tau}{N}}$ |

*Case* (2).—This case compares a conventional mechanically scanned radar system with a within-pulse electronically scanned system under the conditions of equal transmitter power and equal receiving beamwidths (i.e. $\theta_r = \theta_m$).

In the electronic scanning system the transmitter power is spread out over a wider sector of $n\theta_r$ which reduces the power in the received pulses to $KP/n$. But on the other hand, within-pulse scanning gives $n$ times as many returned pulses from each target in a given period of time when compared with conventional scanning. It is therefore possible to integrate or correlate these $n$ pulses which should theoretically improve the signal/noise ratio by a factor $\sqrt{n}$. The comparison made in Table 2 shows that the signal/noise ratio of the within-pulse scanning system is $\sqrt{(2n)}$ times worse than for a conventional mechanically-scanned system under the conditions outlined above.

**Table 2**

Comparison between the signal/noise ratios of a within-pulse, electronically scanned radar system and a conventional mechanically scanned radar under the condition $\theta_r = \theta_m$.

| | *Within-pulse Scanning System* | *Mechanically Scanned System* |
|---|---|---|
| Received power (peak) | $\frac{KP}{n}$ | $KP$ |
| Receiver noise power (before detection) | $\frac{nkN}{\tau}$ | $\frac{N}{\tau}$ |
| Receiver noise power after detection and filtration | $\frac{2nN}{\tau}$ | $\frac{N}{\tau}$ |
| Improvement in signal/noise ratio due to integrating $n$ pulses in within-pulse scanning system | $\sqrt{n}$ | — |
| R.m.s. signal/noise ratios | $\sqrt{\frac{KP\tau}{2nN}}$ | $\sqrt{\frac{KP\tau}{N}}$ |

*Case* (3).—This case compares a conventional mechanically-scanned radar system with a within-pulse electronically scanned system under the conditions of equal transmitter power and equal transmitter beamwidths (i.e. $\theta_s = \theta_m$).

Although these two radar systems may employ identical transmitters, the receiving aerial gain of the electronically scanned system must be $n$ times greater than the aerial gain of the mechanically scanned system. This results in the peak received power for the within-pulse scanning system being $nKP$. The signal/noise comparison for this case is given in Table 3 and it can be seen to be $\sqrt{2}$ times worse than for a conventional mechanically scanned radar system.

The improvement over case (2) is due to the increased aerial gain of the high definition receiving aerial. It may of course be argued that this improvement could have been achieved with a conventional system but such a step would also have necessitated reducing the scanning rate. The justification for the comparison in case (3) is that in certain practical systems, within-pulse scanning could be used to improve bearing accuracy under the condition $\theta_s = \theta_m$. An example of such an application is given in Section 8.

**Table 3**

Comparison between the signal/noise ratios of a within-pulse, electronically-scanned radar and a conventional mechanically scanned radar under the condition $\theta_s = \theta_m$.

| | *Within-pulse Scanning System* | *Mechanically Scanned System* |
|---|---|---|
| Received power (peak) | $nKP$ | $KP$ |
| Receiver noise power (before detection) | $\frac{nkN}{\tau}$ | $\frac{N}{\tau}$ |
| Receiver noise power after detection and filtration | $\frac{2nN}{\tau}$ | $\frac{N}{\tau}$ |
| R.m.s. signal/noise ratios | $\sqrt{\frac{KP\tau}{2N}}$ | $\sqrt{\frac{KP\tau}{N}}$ |

## 7.4. *Methods of Improving Signal/Noise Ratio*

The cases given in the previous section show the signal/noise ratio to be $\sqrt{(2n)}$ times worse than for an idealized system. There are several methods of improving this situation at the price of further complication.

### 7.4.1. Narrow-band amplification in the channels

The bandwidth in each channel up to the input to the mixers is $1/\tau$. Therefore it is possible to employ narrow band amplification at this point so that the signal level is increased before meeting the mixer noise associated with the much wider bandwidth of the frequency sweep. Such amplification could be achieved by means of travelling-wave tubes, or alternatively by the use of two stages of modulation with i.f. amplifiers in each channel between the modulation stages. This would use the second modulation stage for scanning and maintain the first modulation stage and the narrow band of the amplifiers.

The significance of the above modification is that the initial stages of the receiver control the signal/noise performance, but the wide bandwidth associated with any form of sampling system need not be introduced until after amplification. These modifications result in reclaiming the $\sqrt{(2n)}$ loss of signal/noise ratio provided that the noise power per unit bandwidth in the first stages of the system remains the same. This means that in case (3) the signal/noise ratio would be $\sqrt{(nKP\tau/N)}$ which is $\sqrt{n}$ times better than the conventional mechanically-scanned system. The improvement has, of course, come from the increased antenna gain.

### 7.4.2. Remodulation with a swept oscillator

The output from the delay line contains amplitude modulation over a bandwidth $n/\tau$ situated on an f.m. carrier, of bandwidth $\Delta f$. It is therefore quite possible to remove the frequency modulation by means of remodulating with a second frequency swept signal, which may be obtained by frequency translation of the swept local oscillator. This reduces the signal spectrum to a band $n/\tau$ before detection,[8] and such a method has been tried in a sonar system.[12]

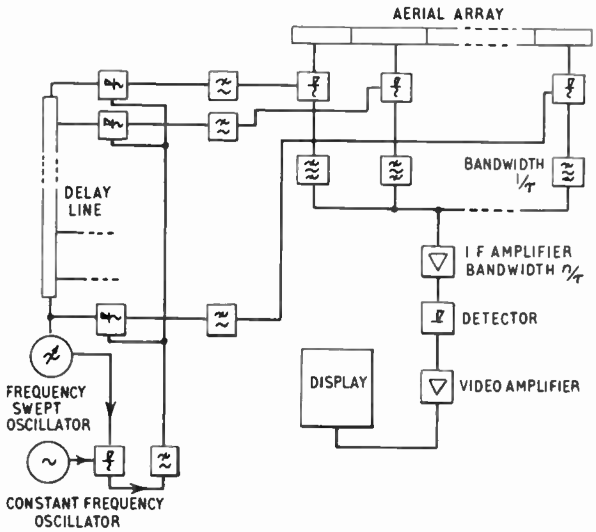



**Fig. 10.** Schematic diagram for remodulation with a swept oscillator.

It is furthermore possible to rearrange the system components so that the remodulation takes place outside the channels in the leads from the swept oscillator as shown in Fig. 10. In this case each channel mixer is switched with a slightly different frequency, but the bandwidth of each mixer and each channel is $1/\tau$. It is only when the channels are added together that the bandwidth increases to $n/\tau$ owing to the different frequencies of the switching signals. This results in the same noise power output as for a single channel radar of bandwidth $1/\tau$, since the total noise power available from the combination of the mixer noise outputs is still no greater than can be obtained from one mixer under matched conditions.

Clearly this modification also, will result in reclaiming the $\sqrt{(2n)}$ loss of signal/noise ratio, and although the system becomes more complex, most of the additional complexity occurs outside the information carrying channels.

## 8. Possible Applications of Within-pulse Scanning to Radar Systems

Conclusions about the application of within-pulse scanning to any specific radar system would require a detailed study of the system requirements. This would be out of place in this paper, and therefore remarks are confined to a brief mention of a few applications which may benefit in some way from within-pulse scanning techniques. These suggested applications are made rather tentatively to demonstrate the most significant features of the system. They are not intended as serious development suggestions which would clearly require far more detailed study of each application.

### 8.1. *Precision Approach Radar*

Precision approach radars are used in conditions of poor visibility to observe and guide an aircraft in its approach to the runway. The requirements of such a radar system are: high range and bearing resolution over a short maximum range, and a fast scanning rate over a narrow sector of search.[13] Clearly these requirements suggest this to be a likely field of application for any electronically-scanned radar system.

The application of a relatively simple form of the proposed within-pulse scanning system to such a radar would seem to be quite feasible. The only complicating factor would be that the high range resolution would result in a high scanning rate, and consequently a large frequency sweep of the local oscillator. Although future precision approach radars will probably represent one link in a complete blind landing system, nevertheless the more continuous nature of the information available from within-pulse scanning should always be of value in any such application.

### 8.2. *Collision Avoidance in the Air*

Aircraft collision avoidance is a potential radar application not yet developed due to severe limitations on weight, aerial gain, and ground clutter; it is also an application which requires a very high scanning rate.[14] Clearly the use of stationary aerials is desirable in this type of application to assist in the mechanical design of radomes, furthermore within-pulse scanning has the ability to increase scanning rates beyond the normal limits.

### 8.3. *High Definition Radar Systems*

High definition radar systems such as airfield surface movement radars often operate in Q band (the 8 mm band) in order to obtain narrow beamwidths from reasonable aerial dimensions. Electronic scanning might improve this situation since there is little restriction on the dimensions of stationary aerials. This might also result in a change to X-band operation with a consequent improvement in bad-weather performance.

### 8.4. *Airfield Surveillance Radar*

Airfield surveillance radars employ fan-shaped beams in the vertical plane which are rotated to find the range and bearing of aircraft at ranges up to about 100 miles. The range resolution is however fairly low, corresponding to pulse lengths of about 3 μsec.[15] The height of the aircraft is usually found by asking the pilot for this information over v.h.f. radio, or by a separate height finding radar. This leads to difficulties in co-ordinating information from the two sources.

If within-pulse scanning could be added to this type of surveillance radar in the vertical plane to provide height information, then the bearing information could still be available as before by mechanical rotation of the aerial. Consider a fan-shaped transmitting beam as employed at present, then a narrow receiving beam could be electronically scanned in the vertical plane illuminated by the fan beam. The height information about the sector would then be available at the output of the electronically-scanned receiver, while the azimuth information could be obtained using the transmitter radiation pattern for receiving as at present.

An interesting feature of this application is that it represents applying within-pulse scanning under the condition $\theta_s = \theta_m$ as discussed in Section 7.1. This is because the system has been used to increase the aerial directivity rather than scanning rate. It was also shown in Section 7.3 that this application (case 3) results in a signal/noise ratio only 3dB worse than for conventional forms of scanning, even with the simple form of the system. However, if either of the two systems suggested for improving the noise factor could be adopted, this would theoretically result in an improvement of signal/noise ratio over the mechanically-scanned system. In this application the signal/noise ratio would depend to a large extent on the resultant change of antenna gain, from the cosec$^2$ pattern to the pencil beam. Since relatively long pulses are used in this type of radar, this would result in quite reasonable scanning rates in the region of 0·3 Mc/s which consequently would not require large frequency sweeps of the local oscillator.

## 9. Summary of the Advantages and Disadvantages of Within-pulse Scanning Radar

*Advantages:*

(1) Convenience of stationary aerials.

(2) Flexibility of form of scanning. It is easy to vary scanning rates or to adapt for a lock-and-follow system.

(3) Compared with other methods of electronic scanning there is no need to scan transmitters.

(4) Aerial aperture amplitude tapers may be easily altered in order to modify beam patterns while scanning.

(5) Higher information rate under suitable circumstances.

*Disadvantages:*

(1) More complexity.

(2) More stringent restrictions on phase and frequency stability.

(3) Worse signal/noise ratio in most applications unless further complication is warranted.

(4) Worse sidelobe performance.

It can therefore be concluded that the application of within-pulse scanning radar of the form described in this paper is most likely to be restricted either to short range applications where a slightly worse signal/noise ratio can be tolerated, or to very large and complex installations where techniques can be exploited to the full.

## 10. General Conclusions

This paper has shown that it is possible to scan receiving beams at very high scanning rates. The analysis revealing the limitations on scanning rate of the proposed system of within-pulse scanning, has been verified by the experimental results. This has shown that a practical within-pulse scanning system is quite feasible. The contents of this paper are intended to assist in the design or prediction of system performance of such a radar system, and with due consideration of the different factors involved, may be applied to an equivalent form of underwater sonar.

## 11. Acknowledgments

The author is indebted to Professor D. G. Tucker for suggesting this research, and for his continued interest in the work. Thanks are also due to Mr. J. T. Allanson under whose direction the work was done, and to several other colleagues in the Department of Electrical Engineering, in particular Mr. B. K. Gazey, for many useful discussions.

The author received a scholarship from the National Research Development Corporation during the early stages of the work.

## 12. References

1. D. G. Tucker, V. G. Welsby and R. Kendell, "Electronic sector scanning", *J. Brit.I.R.E.*, **18**, pp. 465–84, August 1958.
2. D. E. N. Davies, "Radar systems with electronic sector scanning", *J. Brit.I.R.E.*, **18**, 709–13, December 1958.
3. C. E. Shannon, "Communication in the presence of noise", *Proc. Inst. Radio Engrs*, **37**, pp. 10–21, January 1949.
4. P. M. Woodward, "Probability and Information Theory with Applications to Radar", (Pergamon Press, London, 1955.)
5. A. F. Harvey, "Microwave tubes", *Proc. Instn Elect. Engrs*, **107C**, pp. 29–59, March 1960 (I.E.E. Monograph No. 343.)
6. B. K. Gazey, "Electronic Sector Scanning Asdic", M.Sc. Thesis, University of Birmingham, 1960.
7. J. F. Reintzes and G. T. Coate, "Principles of Radar", (McGraw-Hill, New York, 1953.)
8. D. E. N. Davies, "A Radar System with Fast Electronic Scanning", Ph.D. Thesis, University of Birmingham, 1960.
9. D. G. Tucker, V. G. Welsby, L. Kay, M. J. Tucker, A. R. Stubbs and J. G. Henderson, "Underwater echo-ranging with electronic sector scanning: sea trials on R.R.S. *Discovery II*", *J. Brit.I.R.E.*, **19**, pp. 681–96, November 1959.
10. J. R. Ragazzini, "The effect of fluctuation voltages on the linear detector", *Proc. Inst. Radio Engrs*, **30**, pp. 277–88, June 1942.
11. W. R. Bennett, "Response of a linear rectifier to signal and noise", *J. Acoust. Soc. Amer.*, **15**, pp. 164–72, January 1944.
12. B. S. McCartney, unpublished results, University of Birmingham, 1960.
13. G. J. Moorcroft, "Precision approach radar", *Proc. Instn Elect. Engrs*, **105B**, Supplement. 9, pp. 344–50, 1958. (I.E.E. Paper No. 2581R).
14. P. Gaudillère, "Visual and electronic methods of avoiding collisions", *J. Inst. Navigation*, **11**, pp. 40–55, January 1958.
15. E. Eastwood and C. D. Colchester, "Advances in ground radar for civil aviation", *Proc. Instn Elect. Engrs*, **105B**, Supplement, pp. 370–379, 1958 (I.E.E. Paper No. 2580R).
16. J. S. Seeley and J. Brown, "The use of dispersive artificial dielectrics in a beam-scanning prism", *Proc. Instn Elect. Engrs*, **106B**, pp. 93–102, March 1959. (I.E.E. Paper No. 2735, November 1958).
17. A. Bystrom, R. V. Hill and R. E. Metter, "Ground-mapping antennas with frequency scanning", *Electronics*, **33**, No. 19, pp. 70–3, May 6th, 1960.
18. H. R. Serf, "Electronic antenna scanning", *Proc. National Conference on Aeronautical Electronics*, 1958, Dayton, Ohio.
19. F. Reggia and E. G. Spencer, "A new technique in ferrite phase shifting for beam scanning of microwave antennas", *Proc. Inst. Radio Engrs*, **45**, pp. 1510–17, November 1957.
20. R. C. Cumming, "The serrodyne frequency translater", *Proc. Inst. Radio Engrs*, **45**, pp. 175–86, February 1957.
21. A. L. Morris, "Microwave ferrite modulators for high modulation frequencies", *J. Brit.I.R.E.*, **19**, pp. 117–29, February 1959.
22. British Patent Application No. 15215 1956.
23. British Patent Application No. 8402 1958.
24. H. V. Cottony and A. C. Wilson, "A High Resolution Rapid Scan Antenna", National Bureau of Standards Report No. 6723, October 1960.

*Manuscript first received 31st October, 1960, and in revised form on 1st December, 1960. Paper No. 625.*

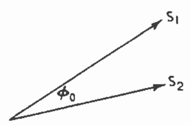
# DISCUSSION

*Under the chairmanship of Mr. R. N. Lord, M.A. (Associate Member)*

**Mr. K. Foster:** It seems that there will be a limitation on the use of this method of scanning to systems in which the received signal is always much greater than the random noise injected by the crystal mixers. I would imagine that in the underwater applications of this method the dominant "noise" is that due to disturbances in the medium and is thus correlated from one receiving element to another whereas in the case of a radar the crystal mixers will inject noise into their receiving channels which is entirely uncorrelated between channels and this would appear to be a possible cause of scanning error or distortion.

In order to clarify the question we can take the case of a simple aerial with only two elements. At any one instant of time the signals in each channel can be represented by two vectors of equal amplitude but of different phases, let us say a difference of $\phi_0$ in the phase (Fig. A). The mechanism described by Dr.

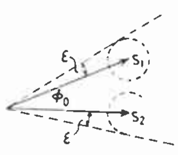
Fig. A.




Davies can be said to determine $\phi_0$ and through what is effectively a one-to-one correspondence allocate an azimuthal scanning angle $\theta_0$ to the signal. This system is clearly very effective in the noise-free use or even in the case when medium noise is present since this noise can be regarded as adding further vectors to $S_1$ and $S_2$ which are themselves equal in amplitude and differing in phase by the same $\phi_0$. In the small signal radar case, however, the noise vectors to be added to $S_1$ and $S_2$ are entirely uncorrelated although it is fair to say that they will have equal r.m.s. amplitudes.

Thus the diagram of Fig. B is relevant in which the

Fig. B.




signal vectors are depicted and the noise vectors to be added are now vectors whose phase is entirely random and whose amplitudes are Gaussian distributed (say) with r.m.s. values given by the dashed circles. The subsequent process will now determine a phase angle $\phi$ between the resultants of the signal vectors plus respective noise vectors which is unique at any instant of time and will generally lie between $(\phi_0+2\varepsilon)$ and $(\phi_0-2\varepsilon)$. (The precise probability distribution of this angle is unimportant for the immediate argument.) Thus at any instant of time the azimuthal angle $\theta$ assigned to the signal would be different from $\theta_0$. The signal, however, is under observation for a definite time interval in each received pulse which will be a function of the "effective beamwidth" of the system and the scanning speeds. We can postulate two differing effects which may be observed.

In the first case, if, over the observation time the correlation functions of the noise from the mixers do not decrease substantially (as could be obtained by a bandwidth limitation), then the resultant vectors between which the phase comparison is made will not change phase substantially and hence for a single received pulse an error in $\theta$ will be observed. For the next received pulse a different value of $\theta$ will be obtained and so on, giving finally over a sequence of many pulses a probability distribution of the angle $\theta$ the mean of which is the true angle $\theta_0$.

The second case assumes that during a time short compared with the observation interval the correlation functions of the noise fall to zero (which would be obtained if the bandwidth of the noise is sufficiently large). In this event the probability distribution in the angle $\theta$ referred to above would be obtained during a single pulse.

Over a long series of pulses the results of both cases will be similar but the distinction is worth drawing because the radar aerial may quite possibly be scanned mechanically in one direction as well as electrically in the other and the former scanning will probably reduce the number of pulses in one sweep to about five to twelve which would be insufficient for the determination of the probability distribution in the first case.

If the signal is small enough to make the uncertainty of $\theta$ corresponding to the uncertainty of phase of $4\varepsilon$ larger than the "effective electrical beamwidth" of the aerial the response of the system under

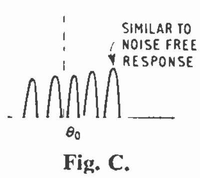



Fig. C.

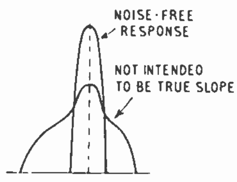



Fig. D.

the two extreme cases above would seem to be as shown in Figs. C and D, drawn for five pulses.

Although this effect has been described for a two-element system I would be interested in any comments Dr. Davies has regarding its applicability to his system.

**Dr. D. E. N. Davies** (*in reply*): The distortion described by Mr. Foster will certainly occur for poor signal/noise ratios and will result in an increase of the bearing uncertainty. The important point however, is that this phenomenon is a basic property of any directional system and the processes of mechanical or electrical scanning do not effect the situation.

The significant part of Mr. Foster's argument is that the bearing uncertainties arise because in the electronically scanned system, uncorrelated noise is added to the signals in each channel before they are added together to form a beam in the medium. In a conventional mechanically-scanned radar system mixer noise is added after the outputs of the array have been combined and therefore cannot affect the phases of the signals from the elements of the array. The argument, however, is not upheld by the principle of superposition. Since the addition is performed by a linear passive network, such as a length of cable, the output waveform may be calculated by the vectorial sum of the signal waveforms and noise waveforms separately. Therefore all the noise sources may be replaced by one noise source situated after the addition process, as in a mechanically-scanned system. Furthermore the signal/noise ratio is theoretically the same for these two systems as previously discussed in the paper.

The concept of noise interfering with the phases of the signals is therefore just another method of regarding the well-known fact that noise reduces target resolution in any radar system. Looked at in a different way, a single target usually appears as a spot on the display, but since the entire face of the cathode-ray tube is scanned by the electron beam, noise signals near this spot will increase the uncertainties about the target position both in range and bearing.

It is probably worth mentioning that the most significant part of medium noise in a sonar system is also uncorrelated from channel to channel provided that the spacing between elements is not less than about one wavelength.

**Mr. T. F. Spriggs** (*Associate Member*): Current practical radar design tends to favour the single horn feed in order to permit the use of efficient rain cancellation units. Does Dr. Davies' multiple horn system make the use of modern polarizers difficult or impossible?

Dr. Davies has told us of the possibility of scanning rates of up to 20 Mc/s. Can he tell us if angle data followers have been developed for use with computers or cathode-ray tube displays.

**Dr. Davies** (*in reply*): In reply to Mr. Spriggs, I see no reason why circular polarization techniques should not be applied to an array of horns either by using circular polarization in each horn or by the use of a polarizing lens across the whole aperture. The former method would probably lead to more complication and the latter method might be restricted to small angles of deflection. It is worth mentioning, however, that the array of horns could be made quite small by reducing the element spacing and using a secondary radiator to maintain the aerial gain.

With regard to angle data followers, I do not know whether these have been developed for very high data rates, but I should point out that the data rate does not increase directly with scanning rate and that within-pulse scanning provides one sample of data each period instead of several integrated pulses once per mechanical scan. Extracting data from the display itself is not affected by the type of scanning and the data rate per target is related mainly to the velocity of the target.

**Dr. J. W. R. Griffiths** (*Associate Member*): In his paper Dr. Davies refers to an important form of distortion of the scanning pattern which arises due to the different times of arrival of the wave front at each section of the transducer when the echo is returning from an object not located on a line normal to the face of the transducer. Although he suggests a method by which this form of distortion may be corrected, I feel he gives the impression that this distortion is fundamental in electronic scanning. This, however, is true only for those methods of scanning which attempt to compensate for a time difference by suitable alteration of the phase of the carrier arriving at each transducer element. It is perfectly feasible to scan the beam by introducing some form of storage element, e.g. a delay line, after each transducer element. By suitable electronic switching between tappings on the delay line the beam can then be scanned without any so-called medium distortion. It is true to say, of course, that apart from the possible increased complexity of such a system, it also has the disadvantage of moving the beam in discrete steps and introducing what may be termed switching noise. This disadvantage can probably be largely eliminated by interpolation methods.

**Professor D. G. Tucker** (*Member*): I think I should point out to Dr. Griffiths that the system he describes is really no different in function from the system using $n$ fixed-beam receivers to cover the sector of $n$ beam-widths, which is referred to by Dr. Davies in Section 7.3, Case (1), of his paper. There is, as I see it, only one difference in performance; this is that, whereas in the $n$ fixed-beam system the directional pattern is identical in shape for each fixed position (apart from the shift of axis), in the system referred to by Dr. Griffiths the transducer sections impose their own directivity on the pattern which is electronically deflected with the result that the effective directional pattern is slightly different for each fixed position.

**Dr. Davies** (*in reply*): The transient distortion mentioned by Dr. Griffiths occurs when a wave arrives at the array from some angle other than the array normal and produces stepped leading and trailing edges to pulses at the output of the receiver. I quite agree with Dr. Griffiths that this distortion is not fundamental to all electronically-scanned systems and can be eliminated for each direction by such means as a fixed delay line. Such a system could consist of several fixed beams formed by independent delay lines, electrical switching could then be used to sample these beams. The similarity between this system and the use of a separate fixed beam receivers has already been pointed out by Professor Tucker, and there is only one small point, that I would add to his remarks concerning the variation of beam shape with deflection. This is not only caused by the directivity of the individual elements but also due to the fact that the array factor is deflected in a sin $\theta$ scale, which causes the beamwidth to increase with deflection from the normal position. This effect is normally only significant for large deflections.

The elimination of transient distortion with fixed delays would seem to suggest that it would be better to scan using time-varying time-delays instead of time-varying phase-shifts. A brief study of this method has led to the rather surprising result that although it eliminates the stepped leading edge to the pulses, the distortion at the trailing edge remains.

**Mr. J. S. Shayler** (*Associate Member*): It is most refreshing to find a new technique being applied to a well-established subject like radar. So-called fast scanning systems have been considered for radar for many years, and within-the-pulse scanning has been under investigation for sonar for a similar period, but this is the first time I have seen within-the-pulse scanning seriously suggested for radar and certainly it is the first time that the factors affecting the performance of such a system have been investigated both theoretically and practically. Dr. Davies is to be congratulated on a most useful as well as a clearly presented paper.

The main advantage of the techniques discussed is that they reduce the limitations introduced by the finite speed of light, although in the simpler system at the expense of signal/noise ratio. It is interesting that the disadvantage can, in theory, be overcome by system complication.

On the subject of application, I think it unlikely that the added complication would pay off in the first three examples, which are all short-range systems where information rate can already be made adequate by more standard techniques. There seems a much stronger case in the author's fourth example, a long range system where there is an information rate problem when using more conventional techniques. An expensive high-power equipment is needed to satisfy this sort of requirement, and so some additional complication (and expense) may more readily be tolerated to improve performance. In addition the long pulse length that can be used eases the circuit requirements of the within-the-pulse scanning units.

**Dr. Davies** (*in reply*): I quite agree with Mr. Shayler's remarks about the application of within-pulse scanning to long range radar systems and I feel that suitable applications of such scanning techniques may also lie in systems requiring searching in two dimensions. The principles of electronic scanning can be extended for independent scanning in two dimensions.

For short range applications, work is in progress at the University of Birmingham on the use of multiplicative arrays (time-average-product arrays) in conjunction with fast scanning. This technique has been tried in a sonar system and has given a remarkable improvement in performance. It can provide a 2-to-1 reduction of beamwidth for a given size of array together with a lower sidelobe level, and the application of this system to electronic scanning radar is at present under investigation.

# APPLICANTS FOR ELECTION AND TRANSFER

As a result of its meeting on 28th March the Membership Committee recommended to the Council the following elections and transfers.

In accordance with a resolution of Council, and in the absence of any objections, the election and transfer of the candidates to the class indicated will be confirmed fourteen days after the date of circulation of this list. Any objections or communications concerning these elections should be addressed to the General Secretary for submission to the Council.

### Direct Election to Member

NOBLE, Spencer Watkins. *Malvern, Worcestershire.*
THONG, Saw Pak, B.Sc., Ph.D., F.Inst.P. *Kuala Lumpur, Malaya.*

### Transfer from Associate Member to Member

CHRISTOPHERS, Lt. Cdr. John Roland, R.N. *St. Albans, Herts.*
HERSEE, George. *Oxted, Surrey.*
MAY, Commander Bernard Calverley, R.N. *Rochester, Kent.*

### Direct Election to Associate Member

DAY, Albert Victor. *Wembley, Middlesex.*
*FOSTER, Hugh Wilson. *Baldock, Hertfordshire.*
HOLFORD, Kenneth. *Reigate, Surrey.*
RAVEN, Francis Edwin, B.Sc. *Monkseaton, Northumberland.*
RODWELL, Francis Ascough, B.Sc. *Durham City.*
TALEKAR, Vinayak Laxman, Ph.D. *Jaipur, India.*
TROTTER, John Ernest. *Birkenhead, Cheshire.*

### Transfer from Graduate to Associate Member

AGARWAL, Chunni Babu, B.Sc. *Mhow, India.*
BEWLEY, William. *Goldington, Bedford.*
CLARKE, Christopher Dowse. *Loughborough, Leicestershire.*
CUNNINGHAM, David Keith. *Shefford, Bedfordshire.*
DIXON, Alex Emmanuel, D.L.C. *Kingston, Jamaica.*
MARKS, Leslie Dryden. *Chippenham, Wiltshire.*
O'NEILL, Daniel. *Shannon Airport, Eire.*
PALMER, Leonard Sidney, *London, S.W.20.*
PUTLEY, Derek Edward. *Cheltenham.*
ROBINSON, Gordon Stanley. *Nottingham.*
STEELE, Michael. *Wallington, Surrey.*
TARNER, Henry Charles James. *Potters Bar, Middlesex.*
WILLIAMS, David Lloyd McNeil, B.Sc. *Carlisle, Cumberland.*

### Transfer from Student to Associate Member

ALLMAN, Royston Anthony. *Woodford Green, Essex.*
GEORGE, Flt. Lt. Kurudamannil Abraham, B.Sc., I.A.F. *Bangalore.*

### Direct Election to Associate

BOSWORTH, Richard William. *Gwelo, S. Rhodesia.*
BRIGGS, William Charles. *Southall, Middlesex.*
COFFEY, William Stanley. *Ipswich, Suffolk.*
DONNE, Major John Lewis, R. Sigs. *Camberley, Surrey.*
ELLIOT, James Murray. *Sutton, Surrey.*
FARRELL, Roy Colin. *Bishops Stortford, Hertfordshire.*
HERRING, William Terence. *Ipswich, Suffolk.*
JANE, Frederick John. *Withernsea, Yorks.*
KIGHTLEY, Anthony John, B.Sc. *Northampton.*
McCULLOUGH, Raymond de Bude. *Chesham, Buckinghamshire.*
MARSHALL, Arthur Leslie. *Hull, Yorkshire.*
ROBINSON, Albert William. *Reading, Berkshire.*
SCOTT, Trevor Louis Brian. *Stevenage, Hertfordshire.*
WILLIAMS, William. *Bolton, Lancashire.*

### Transfer from Student to Associate

MOORE, John Keith. *St. Albans, Herts.*

### Direct Election to Graduate

BUTCHER, John Brian, B.A. *Stevenage, Hertfordshire.*
COWIN, Lt. Anthony Cockshut, R.N., *Swindon, Wiltshire.*
DONOGHUE, James William. *Newcastle-upon-Tyne.*
FERGUSON, Tony. *Salisbury, Wiltshire.*
FRASER, Lt. Alexander Hugh Campbell, B.Sc.(Eng.), R.N. *Portsmouth.*
FRIEND, Beverley William. *Reading, Berkshire.*
HARRINGTON, David George. *Castleford, Yorks.*
JONES, John Francis. *Birkenhead, Cheshire.*
*LIPSCHITZ, Selwyn. *Wolmaransstad, South Africa.*
MADDEN, Donovan Ronald. *Kumasi, Ghana.*
PINCHEN, Anthony Patrick. *West Drayton, Middlesex.*
PROBERT, Peter Ernest. *London, N.13.*
RIGBY, Gerald Peter. *Wigan, Lancashire.*
WHITBOURN, Edward Albert. *Burgess Hill, Sussex.*
WILSON, Malcolm. *Stevenage, Hertfordshire.*

### Transfer from Student to Graduate

BEAUMONT, Antony John. *St. Albans, Herts.*
BRUNSDON, Graham Paul. *Brentwood, Essex.*
DIAS, Cyril Francis. *Bangalore, India.*
DUTTA, Subal Chandra. *Kanpur, India.*
FARUQI, Muzaffar Hussain Shah. *Karachi.*
GIBBONS, Geoffrey Frederick. *Southampton.*
HO KWOK KI. *Hong Kong.*
IBRAHIM, Tipu Mohamed, B.Sc. *Bangalore, India.*
KEON, Ralph Edmund Gaston. *Longfield, Kent.*
KHADKIKAR, Ganesh Dinkarrao. *Jabalpur, India.*
McLEAN, Plt. Off. Donald, R.A.F. *R.A.F. Henlow, Beds.*
MALHOTRA, R. N. *New Delhi, India.*
MENON, Menon Parambill Sethu Madhava, B.Sc. *Bombay.*
NEVILLE, John. *Wirral, Cheshire.*
OKONGWU, Josiah Onyenagolum. *Evesham, Worcs.*
RAO, Krishna Bhaskar, B.Sc. *Ernakulam, India.*
SARGOOD, Alan Richard. *Southampton.*
SHARMA, Gulzari Lal. *Secunderabad, India.*
SHELTON, Roy George Alexander. *Johannesburg, South Africa.*
UNNIKRISHNAN KARTHA, B., B.Sc. *Muvattupuzha, India.*
WIGGINS, John. *London, W.2.*

## STUDENTSHIP REGISTRATIONS

The following students were registered at the 28th February and 28th March meetings of the Committee. The names of a further 30 students registered at the 28th March meeting will be published later.

IJIWOLA, Gabriel Oladele. *Ife, Nigeria.*
JEPHCOTT, Roger E. *Tavistock, Devon.*
KAPOTA, Kenneth James. *Abingdon, Berks.*
KEEL, Ronald. *Wallingford, U.S.A.*
KING, Norman Fredric W. *Berkhamsted, Herts.*
McCONNELL, John Frank. *Cambridge.*
McKAY, Ian Guy. *Ilford, Essex.*
*MACNAMARA, Patrick Colman. *Dublin.*
MUMMERY, Brian D. *Bridgend, Glam.*
MYERS, John Alfred. *Kumasi, Ghana.*
NAFZGAR, Robert. *Hornchurch, Essex.*
*NISSIM, Moshe. *Ramat-Gan, Israel.*
OGUNLELA, Felix Olayinka Adebowale. *Lagos.*
ONWASIGWE, Michael E. *Awka, E. Nigeria.*
*PITAWALA, Abeywardena B. *London, W.2.*
ROBSON, William Peter. *Newcastle-upon-Tyne 7.*
ROGERS, George Brian. *Weston-super-Mare.*
RUDGE, Alan George. *Bury St. Edmunds.*
SALVAGE, Anthony C. *Brampton, Cumberland.*
SEABURG, Bernard Douglas. *Hornchurch.*
SECKER, Robert D. *Taranaki, New Zealand.*
SHAH, Jitendra Shantilal, B.Sc. *London, W.2.*
TAYLOR, David Starkie. *Burnley, Lancs.*
THOMAS, Peter Mansell. *Northwood, Middx.*
TIGHE, Francis Vincent. *Kampala, Uganda.*
VENKITARAMAN, Mullurcara P. *Kerala, India.*

AHIMIE, Edwin O. *Lagos, Nigeria.*
ANDERSON, John. *Keswick, Cumberland.*
ASHEN, David John. *London, E.17.*
ASHWORTH, Arthur Edmund. *Harlow, Essex.*
AYNSLEY, Gerald. *South Shields, Co. Durham.*
BAYLEY, Arthur Ward. *Staines, Middlesex.*
BERZINS, Gunars. *London, W.12.*
BOHTAN, Sardar S., B.A., M.Sc. *Rohtak, India.*
BURNS, Roberts. *London, S.E.19.*
CHEESEMAN, Richard. *Orpington, Kent.*
CHUKUEMEKA, Angus. *Umuahia, Nigeria.*
COETZEE, Andries Petrus. *Johannesburg.*
CULVERHOUSE, Michael J. *Chesham, Bucks.*
DREDGE, Terence Arthur. *Bristol.*
EDMONDSON, Brian. *Bolton, Lancashire.*
EKWELIBE, Michael. *Via Orlu, Nigeria.*
EVANS, John. *Fivemiletown, N. Ireland.*
EZEILO, Godwin Ikechukwu. *Lagos.*
FARROW, James Malcolm. *London, S.W.19.*
FROST, Robert. *Bristol.*
GALE, George D. *Enniskillen, N. Ireland.*
GAMESON, Anthony J. *Tredegar, Mon.*
GIBBONS, Michael John. *Manchester.*
HAKENEY, Joseph Michael. *Birmingham.*
HAN, SOON JUAN. *Singapore.*
HARDING, Leon. *Potters Bar, Middlesex.*
HARROLD, Ronald T. *Chelmsford, Essex.*
HAWTHORNE, Robert J. *Limavady, N. Ireland.*
HENDERSON, Alan J. *Largs, Ayrshire.*
HILL, Michael Edward. *Birmingham.*
HUNG, Victor Chai How, B.Sc. *Hong Kong.*
ISSARANUKORN, Miss S. *London, S.W.3.*
JORDAN, Kenneth Charles H. *Southampton.*
KAMALU, Humphrey. *Lagos, Nigeria.*
KAUSHAL, Jagdish Chandra. *Rugby, Warwicks.*
KENNEDY, Aubrey B. *London, S.W.12.*
KENTUCK, Terrance. *Bedfont, Middlesex.*
KULKARNI, Vilas Mahadeo. *Sangli, India.*
LAIGHT, Basil Critchey P., B.Sc. *London, S.W.1.*
LAWSON, Ronald W. *Hexham, Northumberland.*
LEONCE, Alphonse, B.A. *Karikal, India.*

* Reinstatements

U.D.C. 621-52: 629.135.2: 681.14

# Electronic Simulation and Computer Techniques in the Design of Automatic Control Systems

By

C. SNOWDON†

*Presented at the South Western Section's Convention on "Aviation Electronics and its Industrial Applications" held in Bristol on 7th–8th October 1960.*

**Summary**: Computing techniques for the development of aircraft automatic control systems are described which commence with a completely theoretical investigation on an analogue computer, thence by the introduction of more and more actual components to a complete simulator. Preliminary component specifications are ascertained for any given overall system characteristics. The testing of components which have been developed for use in the actual system is carried out to check the final characteristics of the system. Techniques which have been developed for the inclusion of complex mechanical loads in such simulators, and also the part played by the human operator in a control system are described. The idea of a computer which is capable of automatically setting its own coefficients (i.e. system parameters) so as to give an optimized overall system response is discussed briefly.

## 1. Introduction

The idea of automatic control systems can be traced very far back into the dim and distant past, but there is little doubt that the last war created an urgent need to apply automatic control systems on such a scale, of such complexity, as had never before been considered feasible. The general application of automatic control systems is now permeating the whole of the general industrial field. It should be appreciated however, that it is one thing to put forward the idea of an automatic control system and quite a different thing to make it work in actual practice.

The scheme chosen for any particular application is largely dependent upon the background, experience and ingenuity of the designer, together with the requirements and peculiarities of the particular project under consideration. Putting the idea into hardware form and making it work satisfactorily can be a difficult process; aircraft and missile engineers are continually faced with new complex control systems and much skill and bitter experience has been directed towards the development of the technique which is now considered to be the only satisfactory way of achieving the desired performance.

In the early days, the only way to develop an automatic control system was to do such calculations as were sensible and not too protracted, and then purchase what appeared to be suitable items of control equipment for the construction of an experimental rig. The testing and further development of the system often proved harassing and sometimes dangerous since there was every possibility that the system would be unstable, or generally misbehave for some obscure reason. This form of trial and error development is expensive and time consuming, furthermore it usually involved the operation of the primary machine. When the primary machine is an aircraft or missile it will be readily appreciated that, particularly in the case of the missile, such trial and error methods are likely to lead nowhere very rapidly.

These fundamental differences can be overcome by making full use of the technique of simulation—indeed this technique brings in its train many additional advantages. It is the purpose of this paper to draw attention to this technique, and by way of illustration to describe how it was applied to two major projects involving complex control systems, namely

(1) The conversion of a *Canberra* bomber to fully automatic control so that it could be used as a target aircraft for missile proving on the range.

(2) The development of autostabilizer and automatic landing aid equipment for a vertical take-off jet lift aircraft.

## 2. The Technique of Simulation

The development of a complex control system is a lengthy process involving many repetitive steps, each step retraced being modified and re-examined in the light of subsequent experience. The development begins with detailed appreciation of the problem and a first attempt at a basic control scheme—it is a step

† Short Brothers & Harland Ltd., Belfast.

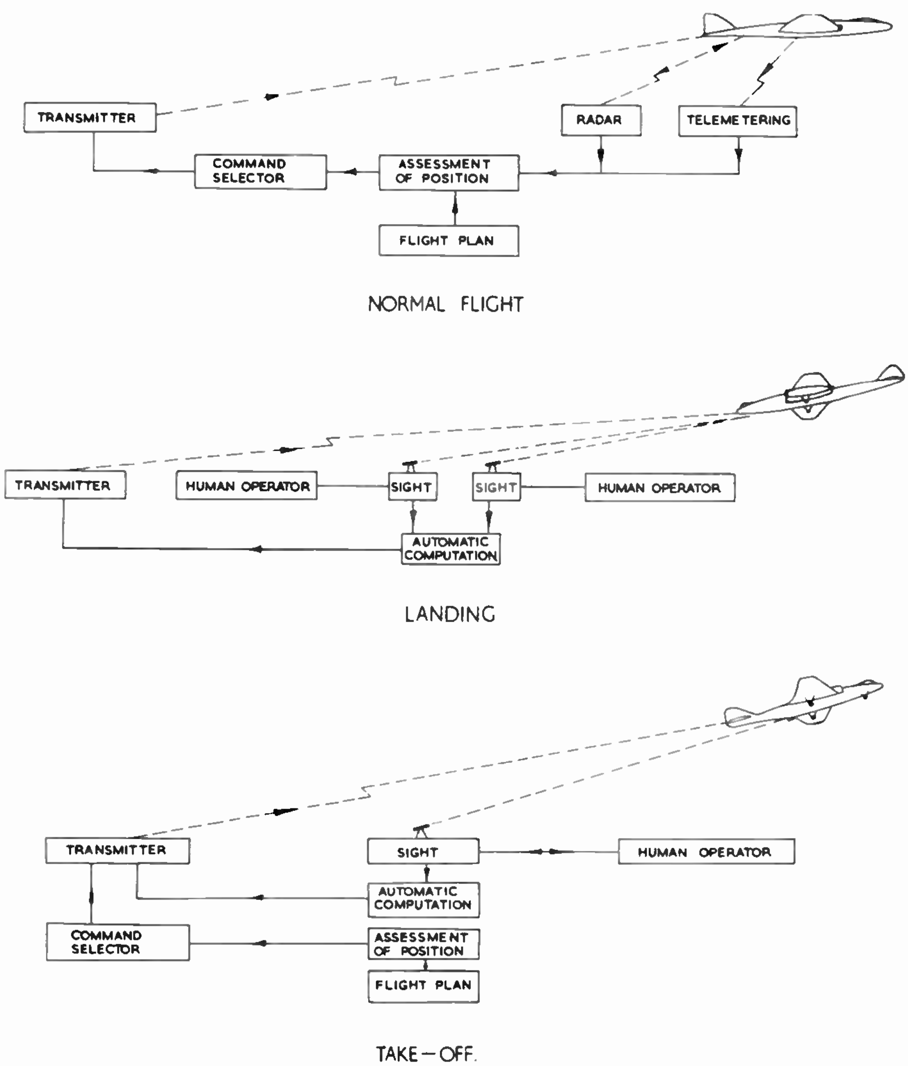



Fig. 1. Pilotless aircraft—control conditions.

which relies almost entirely on the ingenuity and experience of the systems engineer. Having established a tentative scheme the system is examined from a mathematical view-point by means of an analogue computer †.

As experience increases, hardware in the form of real components replaces the appropriate analogue computer elements until the simulator contains the greatest amount of real hardware and the minimum of computer elements. In this way, the system incorporates the idiosyncrasies of the components, e.g. back-lash, hysteresis, non-linearity, noise, etc., whose existence so often mars the performance of an otherwise good system. The simulator can now be expected to give realistic results and should be used to establish final performance specifications and accuracy of the various components together with the overall system performance characteristics.

Simple tests with the primary machine and the control system can now be done to confirm various characteristics as predicted on the simulator, thereby increasing confidence in the ability of the simulator to predict characteristics. Because there was probably an incomplete initial appreciation of either the problem or the requirements, these are likely to undergo some modification as a result of simulator experience; a new system meeting the modified requirements can now be developed in the shortest time and with the maximum confidence that it will perform satisfactorily when installed.

† E.g.: R. J. A. Paul and E. Lloyd Thomas, "The design and applications of a general-purpose analogue computer", *J.Brit.I.R.E.*, 17, pp. 49–73, January 1957.

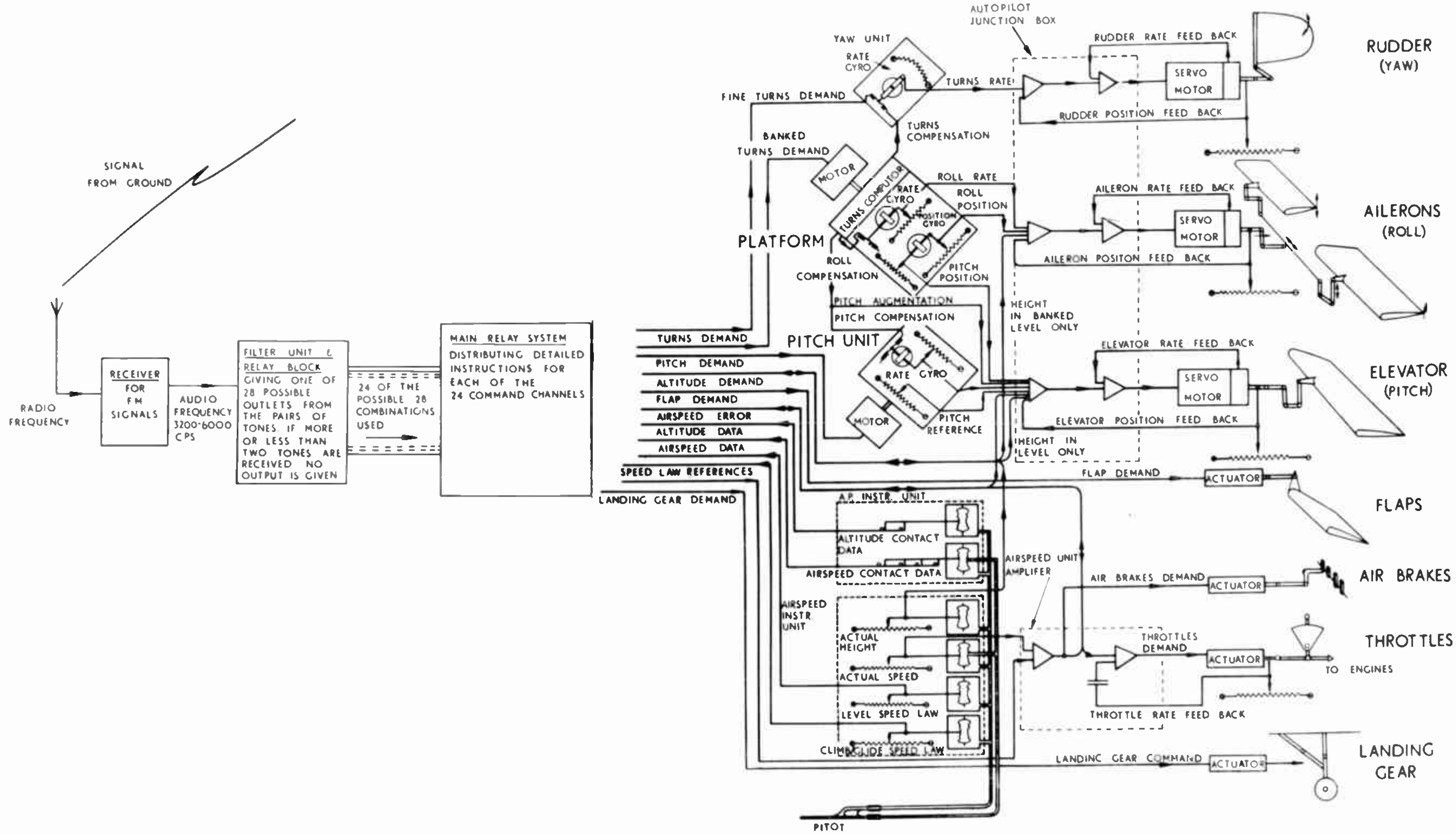



**Fig. 2.** Pilotless aircraft—detailed control system.

Needless to say, the simulator can be used to test components suspected of misbehaviour in the prototype system at all times—this is an invaluable role for the simulator.

### 3. Simulation Techniques as Applied to a Drone Aircraft

The missile industry needed a full-size aircraft which could be used as an unmanned target for the final proving tests of missiles. This automatic role covered take-off from stand-still on the runway, flight for a considerable distance from the drone airfield to the missile range, control of flight over the range by range personnel, and—on the assumption that the missile missed the target—the return of the target aircraft to the drone airfield and its subsequent landing for use at a later date.

Examination of the conditions show that there are two distinct modes of operation of the aircraft:

Phase I: Free flight condition.
Phase II: Landing and take-off condition.

These are illustrated in Fig. 1.

During the free flight condition the aircraft was to be controlled by one man—the Master Controller who was concerned with the geographical position and general attitude of the aircraft as distinct from the flying of the aircraft; the latter can be left to an auto-pilot. The Master Controller therefore monitors the general state of the aircraft and does not in any way contribute to its stability or instability.

During the landing/take-off condition, the Master Controller was to hand over the control of the aircraft to two Ground Controllers, who, via a radio link to the auto-pilot, were to control directly the elevation and azimuth placing of the aircraft relative to the runway with the aid of semi-automatic sighting equipment. Unlike the Master Controller, these two men contribute directly to the stability—or otherwise—of the aircraft, since they are within the control loop. Should it become apparent to the Master Controller that the landing is likely to be unsuccessful he had the authority to over-ride the Ground Controllers and cause the aircraft to climb away on full throttle to return for another attempt at landing.

Diagrammatic representation of these two modes of operation are shown in Fig. 1 whilst Fig. 2 gives a detailed representation of the control system.

A brief examination of the various conditions of operation of the aircraft indicates that the take-off and landing phase would possess the least stability, since not only does it include the auto-pilot–aircraft loop, but also the human operator; sighting equipment and radio command link all introduce lags, and some of these produce discontinuous control data. Of these two cases, the landing phase is likely to possess least stability and prove the most dangerous from the point of view of loss of equipment and the pilot who would be involved in preliminary trials. Auto-pilot development was not likely to be hazardous since more or less suitable auto-pilots were in existence and tests involving the alignment of the auto-pilot to the aircraft could be performed quite safely by skilled test pilots at a safe altitude.

The automatic take-off and landing of the aircraft involved the control of height by variation of throttle-setting—a new and untried form of control; furthermore, the time lags and sensitivity likely to be obtained with Ground Controllers manually operating automatic sights were likely to be unsatisfactory.

A direct analytical approach to the problem of automatic control of the landing and take-off is highly complex, and certainly not amenable to investigation into possible modifications and improvement; on the other hand, direct experiment with an actual aircraft on a trial and error basis is time consuming and could be disastrous. A simulator is the only way to approach this problem and it is essential that it should contain the maximum number of actual components including in particular the human operator. There are other good arguments in favour of simulation, namely:

(1) The use of a real aircraft depends entirely upon the combination of weather and machine serviceability; further, suitable personnel and equipment must be immediately available whenever suitable flight conditions appear.

(2) In the case of the *Canberra*, a lengthy (10 miles) low angle approach path was proposed for the landing technique and in consequence any such operation would be restricted to one or two aerodromes in the United Kingdom. Simulators on the other hand, are housed within one building, if not in one room, and can usually be sited wherever most convenient to the personnel and equipment—in or near a laboratory.

(3) Should instabilities appear, then with a simulator no damage to either man or machine need occur. This leads to a cogent point—the margins of stability can be determined on a simulator—for very obvious reasons this can never be done with real aircraft.

This control system is complex in that it contains two control loops—an inner loop consisting of an auto-pilot for stabilizing the flight path of the aircraft and maintaining it in any demanded condition, and an outer loop consisting of two human beings for controlling the spatial position of the aircraft relative to the runway. Because aircraft auto-pilot technique is a well developed art, and because the system in-

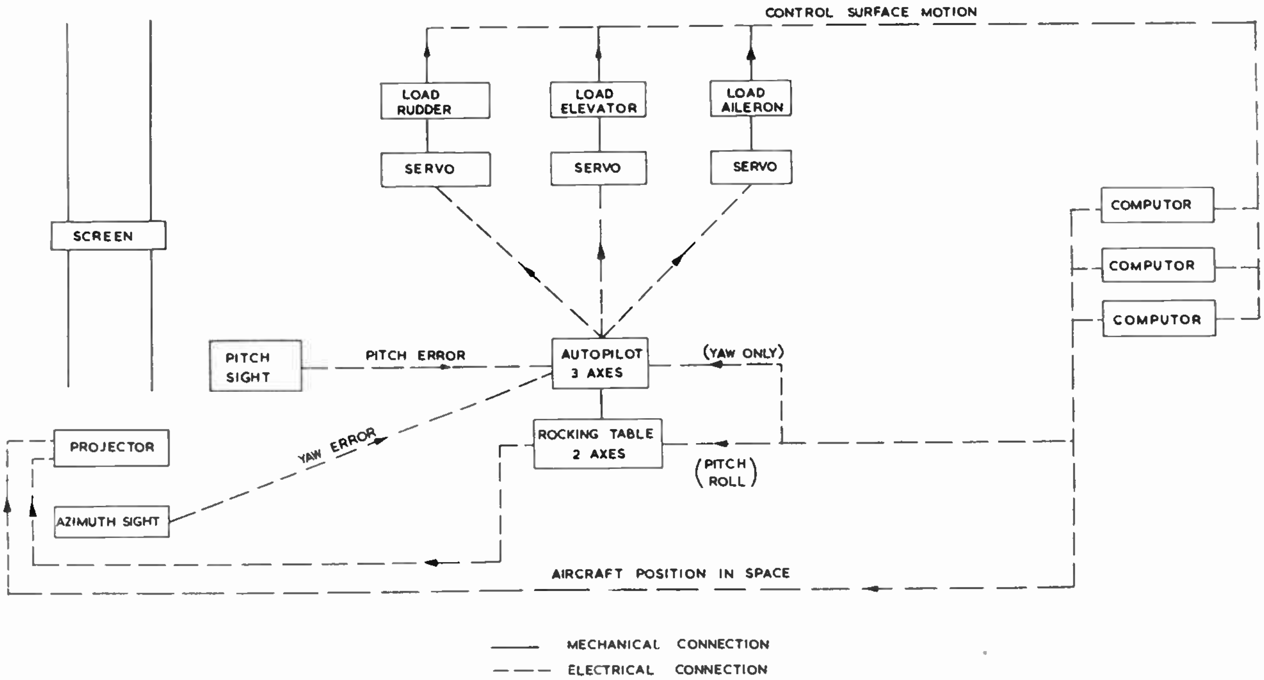



**Fig. 3.** Pilotless aircraft—block diagram of simulator.

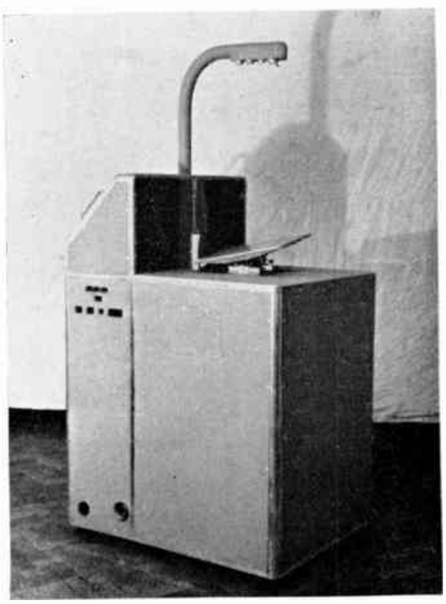


**Fig. 4.** Servo rocking table—rear view.

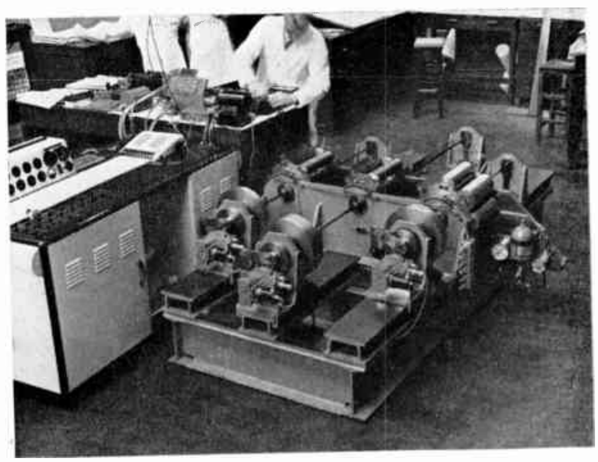


**Fig. 5.** Servo-motor loading rig.

volved the human operator, the initial mathematical phase of the simulation was omitted.

The first task was to build a simulator which would permit the matching of the auto-pilot characteristics to the aircraft thereby achieving satisfactory stability of flight path.

A block diagram showing the set-up appears in Fig. 3. The analogue computer contained the control equations of all three axes of the aircraft and produced output signals proportional to the roll, pitch and yaw angles of the air-frame; it received inputs proportional to the movements of the control surfaces. The roll and pitch signals from the computer operated a high-speed servo-driven two-axis rocking table (Fig. 4) on which were mounted the sensitive elements of the auto-pilot. The auto-pilot servo motors worked into a load having the correct dynamic mechanical impedance, potentiometers on the servo motor loading rig producing signals proportional to the movement of the control surfaces of the aircraft and forming the computer inputs (Fig. 5). Thus the servo driven

Fig. 6. Simulation of pilotless aircraft.

rocking table behaved in the same way as the aircraft platform. It is interesting to note that the decision to include the maximum amount of hardware led to the need for servo driven high speed rocking table; no such equipment was in existence at the time and a rocking table was specially developed for this and similar purposes. It was found to be an invaluable item of laboratory equipment.

Performance of the auto-pilot/aircraft loop was examined and adjusted until satisfactory performance was achieved—this particular auto-pilot had never before been applied to a *Canberra*.

The simulator set-up for the outer loop of the simulation takes the physical form shown in Fig. 6. In this figure will be seen the equipment forming part of the aircraft–auto-pilot loop, and also the means of presenting the motions of the aircraft in space to the two ground controllers operating their actual sighting equipment. Signals from the sighting equipment are fed into the auto-pilot via a command link relay system thereby completing the external loop; actual radio transmission of signals was not included. The visual display equipment received signals from the analogue computer and the displacement of the aircraft on the screen was such that the angular displacements to the ground controllers' eyes were equal to those which would occur with a real aircraft. The visual display system was therefore to scale and the screen moved towards the ground controllers at scale speed. The nett result was that the ground controllers—whose functions were to maintain a graticule in their sighting equipment on the image of the aircraft—were presented with the actual angles, rate of change of angles, etc., which would be subtended by the real aircraft.

Whilst the simulator study was proceeding, some flight test experience was gained of the performance of the actual aircraft–auto-pilot combination with the settings indicated by the simulator. Simulator performance was confirmed, and the predictions of the simulator were then viewed with a reasonable amount of confidence. Eventually, the day arrived for preliminary flight testing of the landing system—with a pilot on board with instructions to revert to manual should he consider that performance was not to his liking. It is gratifying to be able to say that the system behaved satisfactorily and according to prediction—a few minor adjustments were, of course, found to be necessary. Confidence in the simulator was now high —so high in fact that it became a recognized practice to test all units before flight on the simulator.

### 4. Simulation Techniques as Applied to a Research Jet-lift V.T.O.L. Aircraft

The development of an automatic control system for the jet lift vertical take-off aircraft which the author's Company has been developing, presented a more difficult problem in that this was not the conversion of a well-proved aircraft to automatic flight, but the development of automatic control systems for an aircraft which was itself in the very early research stages.

In the hovering state, this aircraft possesses very little inherent stability and early experiments with the Rolls Royce *Flying Bedstead* indicated the need for an autostabilizer which was capable of automatically providing attitude stability in the roll and pitch axes. Experiments with the *Bedstead* indicated the type of control equation which was desirable and the first stage in the development of the autostabilizer was to set up on an analogue computer the proposed control system with probable aircraft characteristics.

Development of the autostabilizer proceeded along the lines suggested earlier; pre-flight investigations were essential on this project and these were performed on the dynamic rig shown in Fig. 7. This rig consisted of a platform having the same inertias and mass distribution as the real aircraft and mounted on

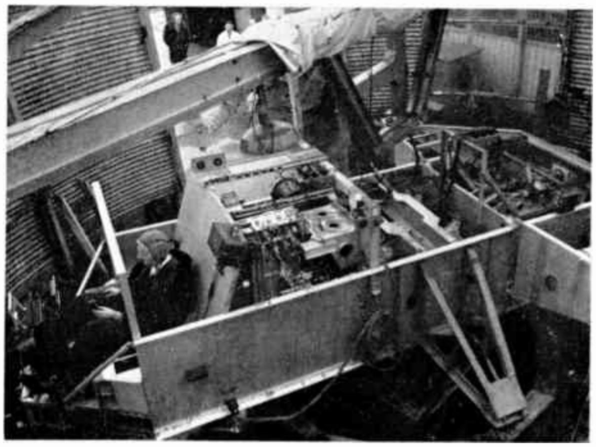


Fig. 7. Dynamic rig for Short S.C.-1 v.t.o.l. research aircraft.

a "frictionless" spherical air bearing. The platform was fitted with jet-reaction moment-producing nozzles in the pitch and roll axes, the nozzles being fed with air from an Avon jet engine. In this way the behaviour of the developed autostabilizer was investigated under extremely realistic conditions and without involving the primary machine.

The results of this work led to the successful design of the SC-1 aircraft. Experiments soon indicated, however, that improvements in performance were both necessary and possible, and the simulator has been used to develop more advanced forms of autostabilizer. When this aircraft is doing a transition from fully wing-borne flight to hovering flight, the aerodynamic lift from the wings falls off with speed and the deficiency must be made good by the jet lifting engines. It was thought that some form of semi-automatic aid was necessary to operate the throttles of the jet lifting engines in such a way that the total thrust (aerodynamic and jet lift) was equal to the weight of the aircraft. A control system using a vertically mounted accelerometer was devised and fine control of the system was left in the hands of the pilot. The system was set up on an analogue computer and the "pilot" was given instruments indicating vertical height and velocity together with a very simple control handle. Early experiments indicated that the pilot was in control of the situation but it was suspected that he would lose control if he was required to pay attention to other tasks. This was actually proved by giving the pilot a very simple distraction. This took the form of a set of random numbers written on a sheet of paper and presented to the "pilot", whilst he was operating the controls, a colleague would call out a number and the pilot was to locate that number on the piece of paper. Under those circumstances, control was poor and it was decided to continue simulation studies in conjunction with a "functional" cockpit in which the pilot would be asked to operate the height control equipment and perform other normal flying tasks. Eventually a suitable system was developed—this is now being installed in the aircraft. This raises the general topic of human operators in control systems.

## 5. Simulation of Human Operator

Many types of control systems, including those already described, incorporate some degree of manual control. The human operator's characteristics are very complex and subject to wide variations; the process of systematically optimizing a manual control system thus involves considerable difficulties. Despite these difficulties it was felt that something useful could be achieved if account were taken of the limited control situation applying generally in any well-conceived system which requires to be optimized. Such a man would operate a standard control from visual information on a standard form of scene and, under these circumstances, it is reasonable to assume that a relatively simplified form of simulation would represent the man with sufficient validity. Of course, even with the assumed form, the characteristics would change from run to run with the same man, and from man to man, but these variations could be introduced by altering the set of constants which determine the precise characteristics of the form of simulation chosen. With the limited set of conditions in which the operator is put for his specific control task, the range of variations for each constant should be less than for the wide set of conditions and tasks to which he responds in his daily life.

There is accordingly the practical possibility of representing a large number of operators and variations in individual operators by changing the simulation constants in sequence. The variation of each constant throughout its range can be provided by switching a suitable number of fixed values. If, for example, each constant is accorded 10 fixed values and there are 10 constants then a sequence of 100 combinations would be sufficient to represent all the characteristics of human operators for the given task. For the simulation to be valid, the probability of each constant having a particular value, according to its deviation from the mean, must be taken into account and this could be done by weighting suitably the results of each run.

Once the analogue man is made, numerous runs can be done to optimize the system and one can rely on no learning effects being present. This is not all, however—the simulation can be operated in condensed time, so that large numbers of runs, ordinarily occupying weeks or months with human operators, can be done in minutes or hours.

The arrangement for optimization of a system is illustrated in the block diagram of Fig. 8. A fixed control task is presented to a simulation of the complete control system under investigation and an "analogue man" operating in condensed time replaces the human operator. The characteristic constants are changed in sequence as described and an evaluation circuit assesses the performance of the system on a statistical basis from the errors obtained in the series of runs. Variations of the control system parameters can then be made until optimum performance is achieved.

For the proposed procedure to be possible it is necessary first of all to have data as to a valid form of analogue for the man and data on the ranges of his "constants". These were not available and it was proposed to obtain the information as follows:

After careful consideration of the researches of the various workers in this field, it was decided to start

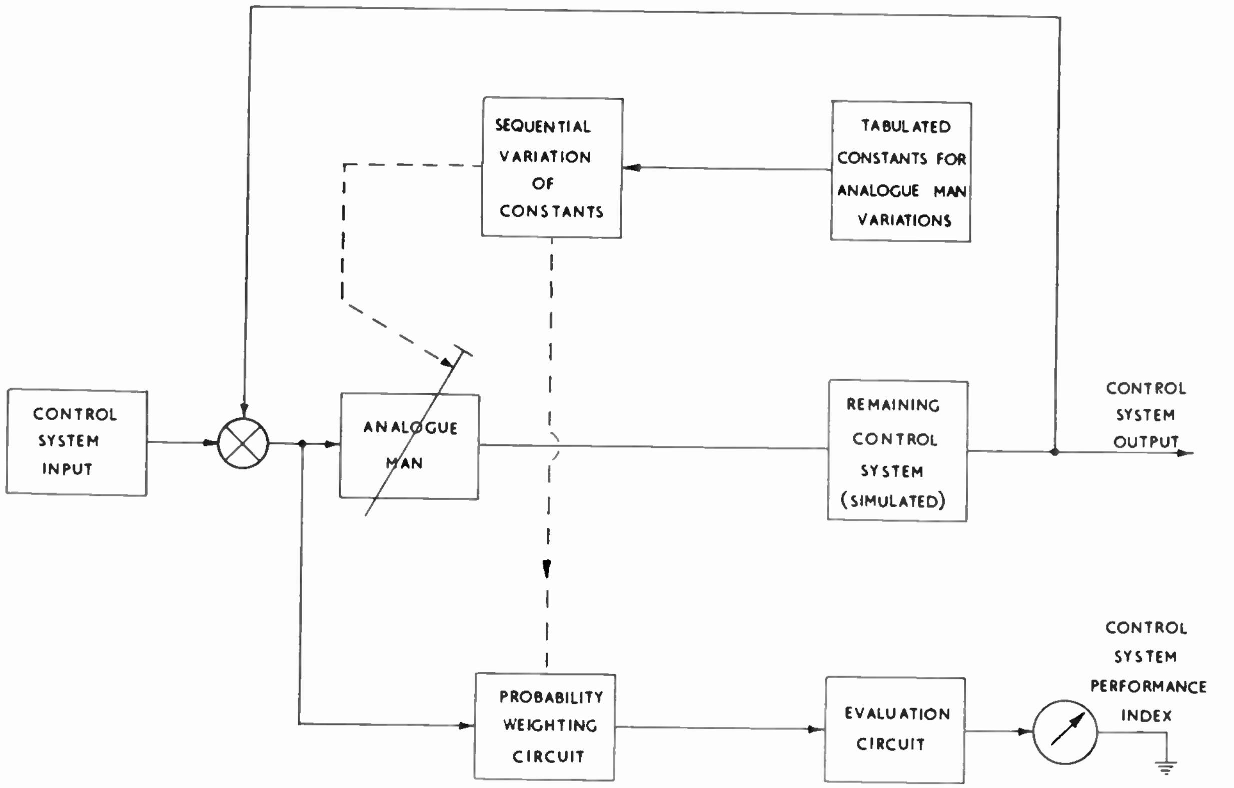



Fig. 8. Control optimization using analogue man.

with the form of simulation shown in Fig. 9 for the man. By presenting this "analogue man" and an actual human operator with the same control task in the same control system simultaneously (see Fig. 10) the excellence or otherwise of the form of analogue could be measured using, for example, cross-correlation methods to compute a "goodness" factor. Then, once the form of analogue is established to be sufficiently valid for the restricted control tasks considered, work could proceed on determining the range of the various constants together with their probability distribution. All of this represents a very considerable programme of work, but once obtained, the data should be applicable to many problems of manually controlled systems.

### 6. Self-adjusting Computers

Usually when a problem has to be solved a complete set of numerical constants is given and it is required to solve for the variables. In some cases, however, the reverse problem may be posed and one may be asked to determine the values of coefficient parameters etc., from values of the variables. Either discrete values may be given at, say, particular instants of time or curves may be supplied. In general a criterion is supplied describing the condition to be reached on the computer for a solution.

The natural method to use is trial and error. An arbitrary set of values of the unknown coefficients or initial conditions is set up and solutions of the variables observed. These are compared with the criteria and readjustment made before the next computer run. This process is continued until the solutions have the correct form.

A simple well-known example is the determination of aircraft flutter speed. Here a number of coefficients dependent on speed are readjusted together on the flutter simulator until the required condition of zero damping is obtained in the solution.

With more complex problems operation is too slow with a human operator and the task of computing the next step of adjustment too difficult. An automatic self-adjusting system is required and this may employ a subsidiary type computer to examine previous solutions and direct the next move. Automatic methods of successive computation have been receiving serious consideration in different countries including France, Russia and the U.K., and various computers have been under development for operation in this manner.

There is more than scientific curiosity involved here and, as one example, the need for a machine capable of investigating atomic pile diffusion problems for

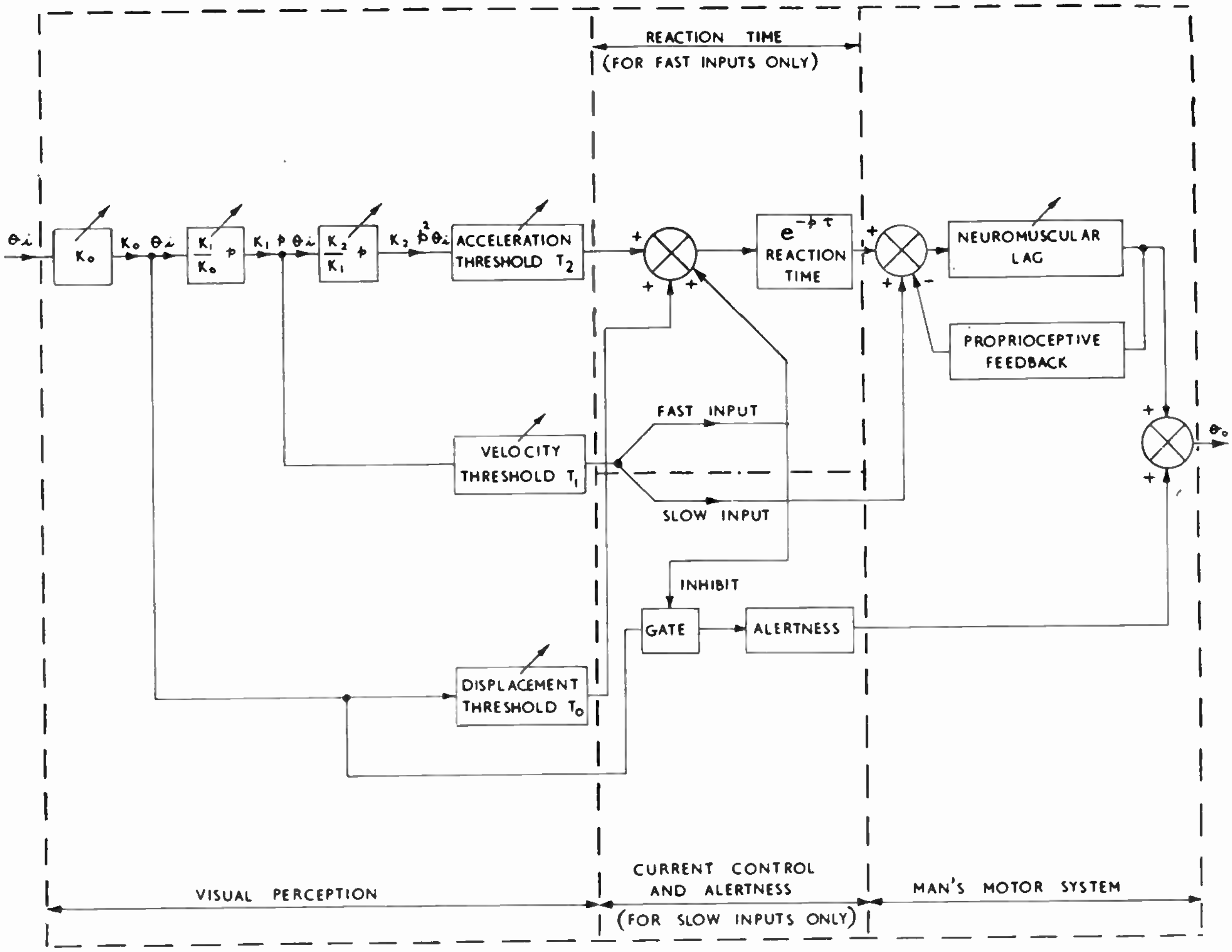



**Fig. 9.** Proposed form of analogue man.

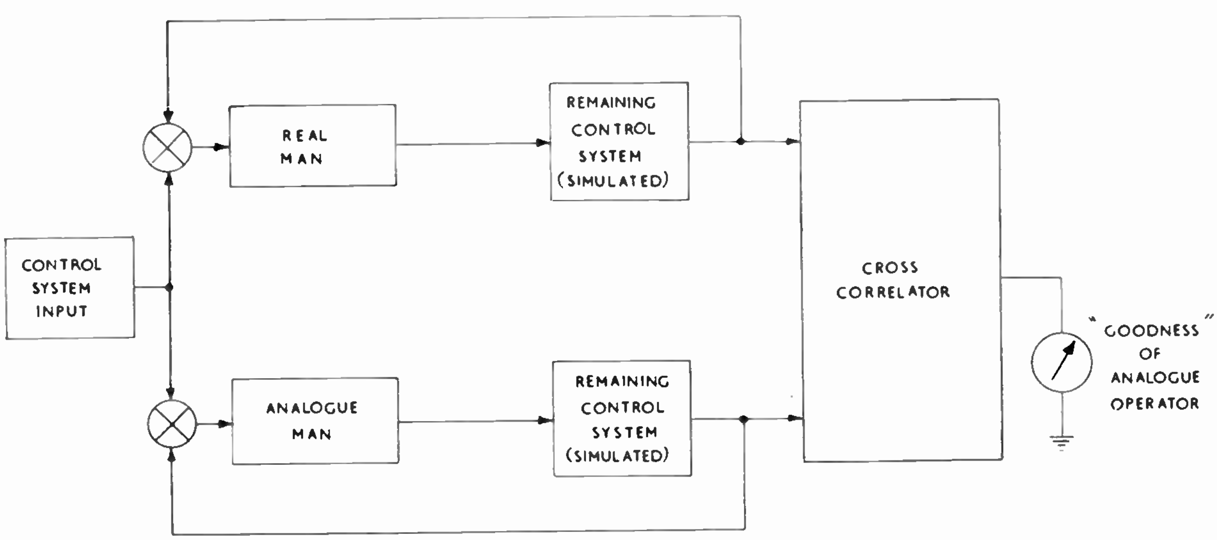



**Fig. 10.** Scheme for matching analogue man to real man.

arrays of general shape may be mentioned. Another application is in the study of flow in large systems of waterways such as found in Holland.

A somewhat different application of the same general principle arises in sympathetic servos. There are two kinds of control system which come to mind in connection with these. In one system the response of the control system may be affected by variations in its parameters during normal operation and vary markedly from a desired response characteristic. By designing the system as a sympathetic servo variations in a parameter which is made deliberately controllable may be used to compensate for other variations and maintain a constant response characteristic. The second type of system is in reality closely related. Instead of a constant response characteristic it may be desirable to have the response vary in a particular manner with conditions and similar methods may be used.

## 7. Simulation of Typical Industrial Problems

Since the differential equations arising from problems in different branches of engineering are very similar, there is a very wide field of application for simulation methods.

An application in the automobile industry receiving attention nowadays is the simulation of vehicle suspension systems. Here the effect of bumps etc., can be tried out and the suspension optimized. Data on the stressing of members can be obtained for use in design. Stresses in beams or shafts from known loading conditions can be obtained quickly and presented as a c.r.t. trace.

A less known use of the analogue computer is its use in conjunction with an electric motor or generator to make the latter present a mechanical load of complex characteristics as indicated diagrammatically in Fig. 11. This in effect is a way of constructing mechanical impedances having complex dynamic characteristics without resorting to the construction of elaborate hardware such as is indicated in the loading rig of Fig. 8. It possesses the added advantage of great flexibility.

Hydro-electric systems involving a series of lakes and waterways demand a knowledge of the effects caused by controlling the flow at a given point on the levels of the various lakes. Such systems have been simulated and the effects studied.

The progress of a chemical reaction may be observed on a computer by simulating the equations relating the rate of growth of each constituent to the instantaneous concentrations of the constituents affecting it. Problems arise also where not only the chemical equations are involved but also, as in the case of an oil refinery installation, the characteristics of the containers, piping and general control arrangements.

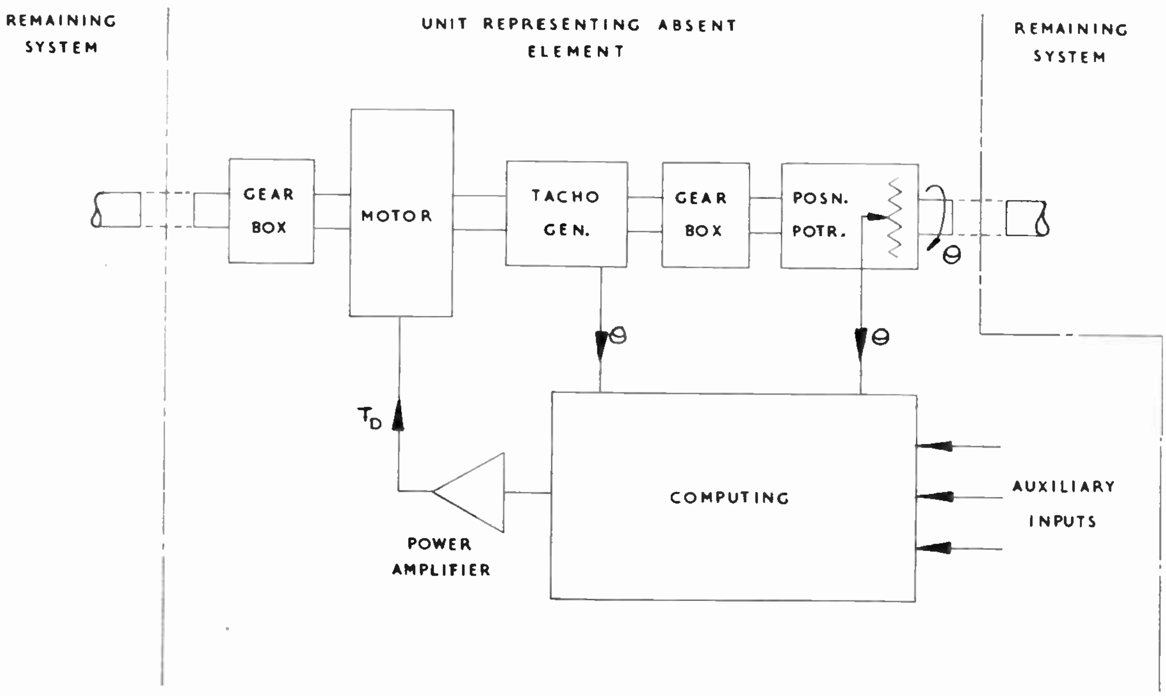



**Fig. 11.** Mechanical load simulation using computer.

Economic problems can be treated which involve say, production and demand and the influence of one upon the other. The stability of such systems may be investigated and the effect of suggested corrective measures studied. One can imagine the long periods required to do the same in the real situation and how impracticable this would be.

Very considerable application for analogue simulation is to be found in the atomic energy field; any or all of the system of a nuclear power station may be required to be represented, the items being, broadly speaking, the nuclear kinetic equations of the reactor, the subsidiary control equations and the heat exchanger characteristics. The advantages, or rather the necessity for investigating the performance of a proposed station before incurring the huge capital expenditure involved does not require to be emphasized.

Operators of a nuclear power station have a responsible task and must be given good training. It is not surprising that special analogue computers have been built as nuclear reactor training simulators so that an accurate representation of actual conditions can be provided.†

An interesting use of the analogue computer as a simulator not only in connection with atomic power stations but also in oil refinery and process control generally, is for prediction purposes. By introducing the present conditions in the actual plant as initial conditions in the simulator and operating it in speeded-up time, an estimate can be made of the conditions at a future time. In process control this would probably be used to give details of the products at the end of the process, thus enabling corrective measures to be taken in advance. The reactor in an atomic power station is subject to poisoning effects which influence the decision as to when the pile must be shut down to change the fuel elements. The simulator can assist in predicting the optimum date for this economically critical operation.

## 8. Conclusions

Since the differential equations arising from problems in different branches of engineering are very similar, there is a wide field of application for the simulation technique. This technique can be applied to such widely varying problems as motor vehicle suspension systems, hydroelectric systems involving waterways and lakes, chemical reactions, nuclear kinetics, and nuclear power stations and economic problems. It is realized that the degree to which hardware can be included is very variable and depends entirely upon the problem, but it does not detract from the general approach given in this paper.

## 9. Acknowledgment

The author wishes to thank the Directors of Short Brothers and Harland Ltd. for permission to publish this paper.

*Manuscript first received by the Institution on 18th May 1960 and in final form on 5th September 1960 (Paper No. 626).*

† E.g.: W. A. Havranek, "A special-purpose analogue computer and its use in reactor engineering", *J.Brit.I.R.E.*, **21**, pp. 225–34, March 1961.

# News from the Sections

## South Midlands Section

Dr. A. E. Karbowiak who read a paper before the Section in Cheltenham on 24th February entitled "Design Aspects and Characteristics of Long-distance Waveguide Communication Systems" began by giving a general outline of a waveguide communication system. He compared its advantages and disadvantages with those of communication systems currently in use. Waveguide properties were discussed with particular reference to signal distortion and large bandwidth handling capacity. The importance of pulse code modulation systems was stressed and a number of essential components as well as special waveguides were discussed in a little more detail. It was concluded that a waveguide communication system was the answer to the growing demands for communication links of increasing capacity.

The thirty-second meeting of the Section was held on 23rd March in the Malvern Winter Gardens, the paper presented being by Dr. A. C. Moore of R.R.E. Malvern on "Thin Film Magnetic Stores". Dr. Moore started by defining "thin" as being less than 1000 Å and "fast" as being less than $10^{-9}$ seconds. He then discussed the properties desired of computer stores. Expressions for switching by domain wall change were developed, and it was shown that switching speeds of the order of microseconds were possible. A mechanical model was used to illustrate the effect. It was then shown that switching speeds of the order of $10^{-9}$ seconds were possible using 90 degree rotation in thin films.

It was shown that the spacing of the storage "bits" could be reduced to below 2·5 mm, giving a possible packing density of 1 megabit per cubic ft. Outputs of 1 V for switching times of the order of 1 nanosecond were obtainable. Dr. Moore then described the deposition of the films on suitable substrates and explained why a chemical deposition gave very satisfactory results both technically and economically. The paper was concluded by descriptions of recent British and American developments in the field of thin films. A selection of apparatus showing switching waveforms and magnetic properties of thin films was demonstrated.

Opening the discussion, Mr. G. Perry spoke as a circuit engineer concerned with the electronics associated with these stores. He compared ferrite rings and films from the point of view of cost, speed, size and associated electronics. Problems of switching, pulse amplification and generation, and readout were mentioned. Recent work with tunnel diodes was described.

In general discussion, a wide range of aspects of the device were dealt with, ranging over the possible use of non-uniform read-in wires, the speed of signalling down lines, the effect of temperature and the effect of the earth's magnetic field.

A. H. M.

## North Western Section

On 23rd November at the College of Science and Technology, Manchester, Mr. P. Denby (Associate Member) gave a paper on "Video Tape Recording". Mr. Denby, who is with the B.B.C., explained that there are two fundamental problems in recording a television picture on magnetic tape—the bandwidth of the signal and the highest frequency which it contains. The effective bandwidth can be reduced by modulating the signal on to a carrier but with a further increase in the highest frequency to be recorded. In order to replay this highest frequency, the wavelength of it on the tape must be longer than the gap width of the replay head. This demands a high head-to-tape speed or a specially narrow gap head.

The mechanical difficulties of either solution are considerable and the situation is eased by multi-track techniques such as were used in VERA (Video Electronic Recording Apparatus).† A more successful solution, however, is provided in the American Ampex and R.C.A. systems which employ four rotating heads to produce recorders having a high head-to-tape speed with a low tape speed.‡

Mr. Denby pointed out that in addition to the usual requirements of the transmission engineer with regard to system performance the telerecording engineer has also to deal with defects in the storage medium and speed variations of the record and replay mechanisms. Two further sources of picture defects that can arise in recorders of this type are due to the presence of a modulation system and of four parallel transmission paths on a time-multiplex basis.

Distortion of modulating signal may occur either due to the modulation/demodulation process itself or because of the restricted pass-band of the amplifying units. In general both will lead to amplitude and phase distortions equivalent to a time-varying low-pass characteristic.

The four head structure of the picture will be apparent if there is discontinuity at the switch or if there is a difference of picture quality before and after the switch. Some of the causes of these defects

† "Magnetic recording of television programmes", *J.Brit.I.R.E.*, **18**, pp. 273–4, May 1958.

‡ A. H. Lind, "Operating facilities in the R.C.A. colour television tape recorder", *J.Brit.I.R.E.*, **20**, pp. 611–20, August 1960.

were described in some detail and illustrated by photographs of B.B.C. test cards. Those arising from mechanical errors in the position of the video heads with respect to the recorded tape were of particular interest.

At the third meeting of the present session on 1st December, Mr. I. M. Waters gave a very comprehensive paper on "Industrial Television". He pointed out that television development up to 1950 had been chiefly directed to broadcast use. This was mainly because the camera tubes available up to that time had necessitated a camera chain too complex and expensive to be applicable to everyday uses.

The appearance of small photo-conductive tubes had made possible the production of an industrial camera chain which was reasonably small, light, simple and reliable and which could be produced at a price within the reach of industrial users.

The circuit practice and operation of a typical chain was described and demonstrated. It incorporated the remote focus and lens change facility and remote pan and tilt for the camera; mention was made of optical attachments commonly used and of the need to protect an industrial camera against adverse operating conditions by means of special housings.

Mr. Waters said that the three main classes of application of such equipment were:

(*a*) To view things which are dangerous directly to observe, or that cannot be seen in any other way. (E.g. in an atomic reactor or for underwater salvage.)

(*b*) Uses in everyday industry where better visual information enables a process to be performed more economically. (E.g. observation of railway waggons in a marshalling yard.)

(*c*) Uses which take advantage of the natural appeal of a picture, in order to communicate information to the general public. (E.g. in teaching).

A film was shown of examples of these various applications and after a useful discussion, members and visitors were able to inspect a camera chain and see it demonstrated within the lecture theatre.

F. J. G. P.

### The North Western Section and the Manchester Federation of Scientific Societies

For a number of years the North Western Section has been a member of the Manchester Federation of Scientific Societies, a body representing scientific societies having branches or sections in the Manchester district and other local societies; eligible societies are Chartered bodies and those which publish original scientific work. The objects of the Federation include co-ordinating arrangements for meetings and publishing a calendar and otherwise providing the participating societies with a means of co-operating in matters of common interest. The Federation is currently worried about the diversion of effort (and reduced attendances) being caused by clashing and duplication of arrangements and to solve these problems the setting up of Sections such as Physical, Engineering, etc., is proposed. The North Western Section is naturally anxious to further the effectiveness of its own meetings in particular and of the Federation in general and it is hoped that a satisfactory solution will be found.

A recent venture on the part of the Federation in collaboration with the *Sunday Times* and the British Association for the Advancement of Science was a Science Fair at which exhibits made by pupils of local schools were displayed; each exhibit illustrated a scientific principle, the aim being that this should be understood by a layman without accompanying demonstration or lecture. This Fair was of the nature of a "pilot" and it is intended to hold a full-scale Fair in 1962.

F. A. M.

### Union of South Africa Section

On 17th November Mr. Leslie Durston, chief architect of the South African Broadcasting Corporation addressed the Section on the history of studio building and described the necessary modifications and acoustic treatment carried out when the S.A.B.C. recently occupied the whole of Broadcast House. A brief description of the new transistor programme input equipment was then given by Mr. J. Vollmer, engineer in charge, Johannesburg studios of the S.A.B.C. Those present were then able to inspect the various aspects referred to in the papers.

At the Section's Meeting on 2nd February, the first of two papers dealing with electronic instruments developed at the National Physical Research Laboratories of the Council of Scientific and Industrial Research was read by Mr. W. W. Schroeder, head of the electro-technical instrumentation unit at the Laboratories. The first instrument described was a direct reading circuit used in conjunction with the Ebert spectrometer. Mr. Schroeder prefaced his paper with an introduction on the applications of spectro-chemistry and then discussed the design considerations of the direct reading circuit. The second part of the paper is to be presented at a later meeting and will deal with time resolved spectra techniques. An informal contribution was then given by Mr. P. du Preez of the Department of Agriculture in which he compared spectro-chemical methods with conventional chemical analysis and pointed out the advantages of the direct reading method in spectro-chemistry in preference to photographic methods.

G. V. M.

# OBITUARY

The Council has learned with regret of the deaths of the following members of the Institution.

**Sidney George Bennett** was killed in a motor accident on 3rd March, 1961.

Born and educated in Liverpool, Mr. Bennett served an apprenticeship with Radiorite Ltd. He joined the Institution as a Student in 1940 and was transferred to the grade of Associate in 1949. During the war he served with the Royal Air Force as a Radar Officer, and prior to his release was on the staff of Headquarters, Air Command, S.E. Asia, Singapore.

He joined E.M.I. Sales and Service Ltd. for a short period, and in 1949 was appointed South Lancashire representative for Ultra Radio and Television Ltd.

Mr. Bennett was 39 years of age, and leaves a widow and one child.

* * *

**Lieutenant-Commander Edward Frederick Sare** died in November 1960, aged 41 years, as a result of an accident whilst swimming.

He was elected an Associate of the Institution in 1948 and was transferred to Associate Member in 1951.

Mr. Sare was born and educated in Coulsdon, Surrey, and on the outbreak of war joined the Royal Navy. He qualified as a Radar Officer, and after the war served for three years on the Instructional Staff of H.M.S. *Collingwood.*

From 1952 to 1953 he held the appointment of Deputy Electrical Engineer (Radio) at H.M. Dockyard, Singapore, and was transferred to the staff of the Director, Radar Equipment, Admiralty, for a term of three years.

From 1956 to 1958 he served as a Squadron Electrical Officer, and at the time of his death he was Assistant Fleet Electrical Officer on the staff of the C.-in-C. Mediterranean.

* * *

**William Samuel Earle** died on 1st April, aged 50 years, after a long period of ill-health.

He was originally trained and worked as a clockmaker, but after studying at the London Radio College he passed the Institution's Graduateship Examination and subsequently qualified as an Associate Member in 1938. He joined Erie Resistor Ltd. in 1939 and worked for this company throughout the war in its research and development laboratories.

In 1949 he was for a short time with Dawe Instruments Ltd. and he then joined Evershed and Vignoles Ltd. as a development engineer.

In 1952 Mr. Earle became seriously ill with tuberculosis which necessitated his first operation. Ever since he had spent long periods in hospitals and sanatoria. In the face of such difficulties he exhibited great courage and periodically returned to light local employment which, although not commensurate with his technical ability, provided some opportunity to show his independence of spirit and desire to support his family.

An operation at the end of 1960 for the removal of a tumour finally confined him to bed. He was admitted to hospital again at the beginning of this year. Mr. Earle leaves a wife and two children. The eldest child, a boy aged 16 years, has been a pupil and boarder at the Royal Wolverhampton School since 1957.

* * *

**George Joseph Bailey** who died suddenly in Montreal on 26th January aged 49 years was elected an Associate of the Institution in 1954. Born in Dublin, Mr. Bailey was educated in Liverpool, and in the years before the war was employed as a test engineer in successively more responsible posts with A. C. Cossor, Philco and Burndept-Vidor.

At the outbreak of war he joined the R.A.O.C., later transferring to R.E.M.E., and spent a period at the Military College of Science as an instructor. In 1944 he was seconded to the Signals Research and Development Establishment, Ministry of Supply, and he was subsequently posted to the Tropical Testing Establishment in Nigeria. From 1946–53 Mr. Bailey was employed by the Nigerian Government as a wireless engineer and assistant superintendent of workshops, being concerned with the establishment of v.h.f. telephone communication networks.

On returning to England he was for about a year with the Plessey Company's Test Gear and Telecommunications Engineering Laboratories and in 1954 he emigrated to Canada where he joined Canadair Limited as a design engineer. At the time of his death Mr. Bailey was a senior engineer in the Company's electronics laboratory concerned with government research and advanced development projects.

U.D.C. 621.385.832.24

# On The Problem of Magnetic Focusing of a Beam of Electrons Emitted with Thermal Velocities†

By

JIŘINA VEJVODOVÁ‡

**Summary:** The current passing through the anode opening is determined for the case of a beam of electrons focused by a longitudinal homogeneous magnetic field and simultaneously accelerated towards the anode by a longitudinal electrostatic field. A point cathode, or a strip or circular plane cathode, is located in the magnetic field. At the cathode a Maxwell-Boltzmann distribution of velocities is assumed. The criterion of sufficient focusing is determined first for a magnetic field and then for an electrostatic field as the variable.

## List of Symbols

$A \quad = \dfrac{m}{8kT}$

$B$ magnetic induction

$D \quad = \sqrt{\dfrac{y}{2(e/m)E}}$

$e$ charge of electron

$E$ electric field

$g \quad = w\sqrt{\dfrac{y}{2\,e/m}}$

$\phi(u) = \operatorname{erf} u = \dfrac{2}{\sqrt{\pi}}\displaystyle\int_0^u \exp(-u^2)\,du$

$i$ current density

$i_0$ current density at the cathode

$I$ current through the aperture

$I_0$ current at the cathode

$k$ Boltzmann's constant

$m$ mass of electron

$n$ number of electrons per second

$T$ temperature in deg K

$\omega \quad = -\dfrac{eB}{m}$

$\omega_m$ root of $D\omega = \tan(D\omega)$

$P \quad = a'b'$ rectangular aperture

$P_0 \quad = ab$ rectangular current source

$P_\omega$ relative current peak

$P_\omega \quad = -\log P_\omega$

$R$ radius of circular aperture (with point source)

$R$ radius of circular source (with flat source)

$R_s$ radius of aperture (with flat source)

The MKS system of units is used throughout the paper.

*Editorial note:* The original paper employed the Gaussian system of units but with the kind co-operation of the authoress the appropriate MKS units have been substituted as this system is more familiar to members.

## 1. Introduction

The initial velocities of electrons emitted from a cathode have a defocusing effect on the electron beam. Langmuir[1] showed that the initial velocities limit the current density in the demagnified image of the cathode (even when the effect of space charge is neglected). More recently Cutler and Hines[2] have concerned themselves with the influence of thermal velocities on the focusing of an electron beam, and have determined the ratio of the current density in a beam produced by a Pierce cathode in the absence and in the presence of thermal velocities. The effect of thermal velocities is particularly important in focusing by a longitudinal magnetic field, which influences their radial components. Stewart[3] has considered this problem for an idealized travelling-wave-tube structure. Pierce and Walker[4] derive a method for the determination of that part of a magnetically focused electron-beam current lying within a certain radius under the simultaneous influence of space charge and thermal velocities, but they introduce a little-justified assumption regarding the distribution of space charge.

† First published (in Czech) in the *Czechoslovak Journal of Physics*, **6**, pp. 656–67, 1956. Translation made by Dr. Charles Süsskind (Associate Member), University of California, Berkeley, California.

‡ Chair of Electronics and Vacuum Physics, The Mathematical-Physical Faculty, Charles University, Prague, Czechoslovakia.

It is the purpose of this contribution to determine the distribution of the current density in a beam focused by a homogeneous longitudinal magnetic field and simultaneously accelerated by a longitudinal electrostatic field, and to determine the dependence of the total current passing through the anode aperture on the intensity of the focusing magnetic field. The source of electrons is located in the magnetic field.

## 2. Beam Formed by Electrons Emitted from a Point Source

The point source of electrons is located at the origin of a rectangular coordinate system, the homogeneous electric accelerating field is directed along $-y$, and the magnetic field along $+y$. We determine the current density by first determining the electron trajectories, which obviously also depend on initial velocities; the initial velocities are then expressed as functions of the coordinates, and the derived dependence is substituted into the Maxwell-Boltzmann distribution of velocities. In cylindrical coordinates the equations of motion take the form

$$m(\ddot{r}-r\dot{\phi}^2) = eBr\dot{\phi} \qquad \text{......(1a)}$$

$$\frac{m}{r}\frac{\mathrm{d}}{\mathrm{d}t}(r^2\dot{\phi}) = -eB\dot{r} \qquad \text{......(1b)}$$

$$m\ddot{y} = eE \qquad \text{......(1c)}$$

where $m$ is the mass of the electron, $e$ is the electron charge, $B$ is the intensity of the magnetic field, and $E$ is the intensity of the electrostatic field.

By integrating eqn. (1) with the initial conditions $r_{t=0} = 0$, $\dot{r}_{t=0} = v = \sqrt{(v_x^2+v_z^2)}$, $\dot{y}_{t=0} = v_y$, we obtain

$$r = \frac{2v}{\omega}\sin\tfrac{1}{2}\omega t \qquad \text{......(2)}$$

$$y = \tfrac{1}{2}\frac{e}{m}Et^2+v_y t \qquad \text{......(3)}$$

With regard to the present problem, we are particularly interested in current-density distribution in a plane $y$ = constant, in which the electrons have been accelerated sufficiently to proceed through the anode aperture into the field-free region. In that plane, the electrons have a $y$-directed velocity corresponding to not less than $10^2$ eV; we are justified in neglecting the initial velocity $v_y$ (99% of all electrons emitted from the cathode at a temperature $T = 2500°$ K have a component of velocity $v_y$ smaller than that corresponding to about 1 eV), so that

$$r = \frac{2v}{\omega}\sin\left(\omega\sqrt{\frac{y}{2\frac{e}{m}E}}\right) \qquad \text{......(4)}$$

Let the number of electrons emitted from the source per second be $n$ (total emitted current is $I_0 = ne$), so that a number of electrons

$$\mathrm{d}n = \frac{nm}{2\pi kT}\exp\left[-\frac{m}{2kT}(v_x^2+v_z^2)\right]\mathrm{d}v_x\,\mathrm{d}v_z$$

$$= \frac{nm}{2\pi kT}\exp\left[-\frac{mv^2}{2kT}\right]v\,\mathrm{d}v\,\mathrm{d}\phi \qquad \text{......(5)}$$

have velocity components between $v_x$ and $v_x+\mathrm{d}v_x$ and between $v_z$ and $v_z+\mathrm{d}v_z$, where $v_x = v\cos\phi$ and $v_z = v\sin\phi$.

In the plane $y$ = constant, these d$n$ electrons per second pass through an element of area $\mathrm{d}P = r\,\mathrm{d}r\,\mathrm{d}\phi$, so that the current density there is given by the expression (from eqns. (4) and (5))

$$i = e\frac{\mathrm{d}n}{\mathrm{d}P} = \frac{I_0}{\pi}\alpha\exp(-\alpha r^2) \qquad \text{......(6)}$$

where

$$\alpha(\omega) = \frac{m}{8kT}\,\frac{\omega^2}{\sin^2\left(\omega\sqrt{\frac{y}{2\frac{e}{m}E}}\right)}$$

It follows from eqn. (6) that at $r = 0$, i.e. on the $y$ axis, the current density becomes infinite at

$$\omega = \frac{N\pi}{\sqrt{\frac{y}{2\frac{e}{m}E}}} \qquad \text{......(7)}$$

which is, of course, caused by the fact that the emission takes place from a point source of a finite

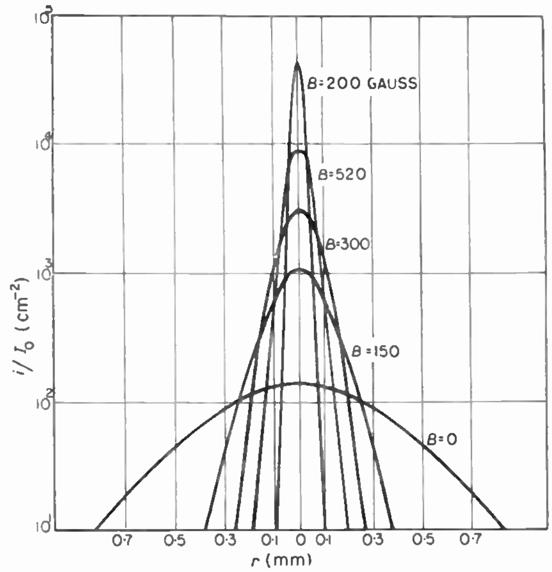



**Fig. 1.** Distribution of current density in an electron beam emitted from a point source ($y = 5$ mm, $E = 200$ V/cm, $T = 2500°$ K).

emission current (in the source, current density is infinite) and at points on the axis $y = 2(N\pi/\omega)^2 E(e/m)$ the source is imaged with unity magnification (a well-known property of a homogeneous magnetic field). Figure 1 shows the function $(\alpha/\pi)\exp(-\alpha r^2)$ for several values of magnetic-field intensity and for fixed values of $y$ and $E$. (For $B = 212$ gauss, an image of the source is formed.)

The total current flowing through a circular aperture of radius $R$ located on the axis at a distance $y$ is

$$I = \int_0^R \int_0^{2\pi} \alpha \frac{I_0}{\pi} \exp(-\alpha r^2) r \, \mathrm{d}r \, \mathrm{d}\phi = I_0[1 - \exp(-\alpha R^2)] \qquad \ldots\ldots(8)$$

From eqn. (8) it is evident that the maximum current passes through the aperture for $\alpha = \infty$, i.e., $\omega^2 = (2EN^2\pi^2/y)(e/m)$, and is equal to $I_0$, i.e., to the total emission current of the point source. The minimal current passing through the aperture is readily determined from the expression

$$\mathrm{d}I/\mathrm{d}\omega = (\mathrm{d}I/\mathrm{d}\alpha)(\mathrm{d}\alpha/\mathrm{d}\omega) = 0$$

and the corresponding values $\omega_m$ are given by the relation

$$\tan(D\omega_m) = D\omega_m \qquad \ldots\ldots(9)$$

provided $D = \sqrt{\dfrac{y}{2eE/m}}$. Table 1 shows several values of the roots of eqn. (9) (the expression $1/\cos^2 D\omega_m$ will be used later).

Figure 2 shows a graphical representation of the relation (8) for two different values of the aperture radius. Although the total character of the dependence of the current through the aperture on the magnetic-field intensity remains preserved for various values of radius $R$, it is evident that for a sufficiently large radius of the aperture the minima will in practice not occur: the intensity of the current increases with increasing magnetic field to its maximum value and then, in practice, does not decrease and increase—only the maximum diameter of the beam passing through the aperture decreases. If the focusing magnetic field is also used to regulate the current after the anode aperture, it is advantageous to work in the vicinity of the maximum, where the current is most sensitive to variations in magnetic-field intensity. If, on the other hand, it is important that the current should not be sensitive to variations in magnetic field, it is more advantageous to work in the vicinity of the minimum.

**Table 1**

| $m$ | $D\omega_m$ | $\dfrac{1}{\cos^2(D\omega_m)}$ |
|---|---|---|
| 1 | 0 | 1 |
| 2 | 4·4934 | 21·19 |
| 3 | 7·7253 | 60·65 |
| 4 | 10·9041 | 119·97 |
| 5 | 14·0662 | 198·93 |

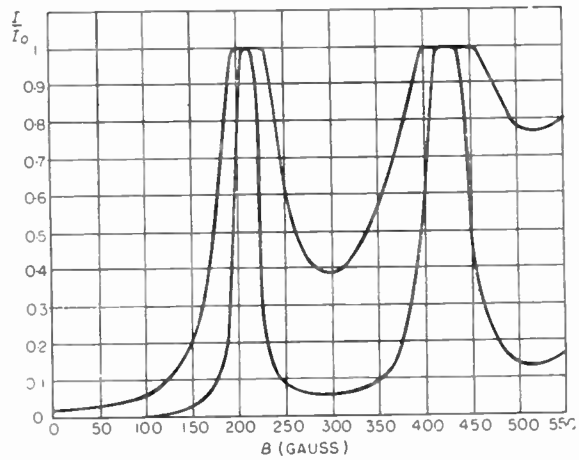



Fig. 2. Dependence of the current transmitted through the anode aperture on the magnetic-field intensity ($y = 5$ mm, $E = 200$ V/cm, top curve for $R = 0{\cdot}1$ mm, lower curve for $R = 0{\cdot}02$ mm).

In the evaluation of the focusing of a given beam it is advantageous to introduce the relative current peak $P(\omega_m)$ as the ratio of the difference between the current at the maximum and at the minimum to the maximum current,

$$P(\omega_m) = \frac{I_0 - I_{\omega_m}}{I_0} = \exp\left(-\frac{A}{D^2}\frac{R^2}{\cos^2 D\omega_m}\right) \ldots\ldots(10)$$

where $$A = \frac{m}{8kT}$$

If it is desired that the intensity of the current passing through the aperture should be in practice independent of changes in the intensity of the magnetic field, one can readily determine from eqn. (10) the corresponding magnetic-field intensity for the appropriate minimum. If it is required that the relative current peak should not exceed a certain value, i.e. that $P(\omega_m) \leqslant 10^{-p}$ ($p = 1$ for 10% current peak, $p = 2$ for 1% current peak, etc.), then

$$\frac{1}{\cos^2 D\omega_m} \geqq 2{\cdot}3026 \frac{p}{R^2}\frac{A^2}{D} \qquad \ldots\ldots(11)$$

The appropriate value $1/\cos^2 D\omega_m$ can be found in Table 1 and $\omega_m$ can then be determined from it.

In practice it is also often required that the total current through the anode aperture should not depend on the intensity of the electrostatic field (i.e. that the current intensity in the region beyond the anode should not depend on the velocity of electrons). If we express, from eqn. (8), the total current through the anode aperture as a function of the intensity of the isolating field we obtain

$$I = I_0\left[1-\exp\left(-A\omega^2R^2/\sin^2\frac{g}{\sqrt{E}}\right)\right] \quad ......(8')$$

where $$g = \omega\sqrt{\frac{y}{2e/m}}$$

From eqn. (8′) we can again determine the values of the electrostatic field for which current is maximum and minimum and the corresponding currents:

$$E_{max} = \frac{\omega^2 y}{2\frac{e}{m}\pi^2(N+1)^2}, \qquad I_{max} = I_0$$

$$E_{min} = \frac{\omega^2 y}{2\frac{e}{m}\frac{\pi^2}{4}(2N+1)^2},$$

$$I_{min} = I_0[1-\exp(-AR^2\omega^2)]$$

The relative current peak $P(E)$ is here

$$P(E) = \exp(-AR^2\omega^2) \quad ......(12)$$

and does not depend on the intensity of the electrostatic field. For a given aperture size and for a required anti-logarithm of the relative current peak, the required intensity of the focusing magnetic field is obtained from eqn. (12),

$$\frac{eB}{m} = \omega = \sqrt{\frac{p}{AR^2}}\cdot 2{\cdot}3026 \quad ......(12')$$

### 3. Beam Formed by Electrons Emitted from a Flat Source of Rectangular Shape

A rectangular source of area $P_0 = ab$ is located symmetrically on the axis $y$ in the plane $y = 0$. The source is in a homogeneous magnetic field directed along $y$ and electrons are accelerated towards the anode by a homogeneous electrostatic field $E$ in the direction $-y$. The equations of motion of the electrons in this configuration are

$$m\ddot{x} = e\,\dot{z}B$$
$$m\ddot{y} = eE \quad ......(13)$$
$$m\ddot{z} = -e\,\dot{x}B$$

By their integration we obtain

$$x-x_0 = \frac{v_x}{\omega}\sin\omega t - \frac{v_z}{\omega}\cos\omega t + \frac{v_z}{\omega} \quad ......(13a)$$

$$z-z_0 = \frac{v_z}{\omega}\sin\omega t + \frac{v_x}{\omega}\cos\omega t - \frac{v_x}{\omega} \quad ......(13b)$$

$$y = \frac{eE}{2m}t^2 \quad ......(13c)$$

Under the initial conditions for $t = 0$

$$x = x_0, \quad \dot{x} = v_x$$
$$z = z_0, \quad \dot{z} = v_z$$
$$y = 0, \quad \dot{y} = 0$$

If we consider eqns. (13a) and (13b) as a system of two linear equations for $v_x$ and $v_z$, we obtain by their solution a relation between the component of initial velocity, the coordinates of the location of the source from which the electron is emitted, and the coordinates of the general plane, $y$ = constant, through which the electron will pass at time $t$:

$$v_x = \frac{\omega}{2}\left[(x-x_0)\frac{\sin\omega t}{1-\cos\omega t} - (z-z_0)\right] \quad ......(14a)$$

$$v_z = \frac{\omega}{2}\left[(x-x_0) + \frac{\sin\omega t}{1-\cos\omega t}(z-z_0)\right] \quad ......(14b)$$

From the elemental area of the source $dP_0 = dx_0\,dz_0$, $dn'$ electrons per second are emitted with velocity components between $v_x$ and $v_x+dv_x$ and between $v_z$ and $v_z+dv_z$:

$$dn' = n\frac{m}{2\pi kT}\exp\left[-\frac{m}{2kT}(v_x^2+v_z^2)\right]dv_x\,dv_z\,dx_0\,dz_0 \quad ......(15)$$

The current density in any plane $y$ = constant is then again given as $i = e\,dn/dP = e\,dn/(dx\,dz)$. By use of the relationships (15), (14b), and (13c) we then obtain for the current density

$$i = \frac{i_0}{\pi}\alpha\int_{-\frac{1}{2}a}^{+\frac{1}{2}a}\int_{-\frac{1}{2}b}^{+\frac{1}{2}b}\exp\{-\alpha[(x-x_0)^2+(z-z_0)^2]\}\,dx_0\,dz_0 \quad ......(16)$$

where $\alpha = \alpha(\omega)$ is identical with the expression (7) and $i_0$ is the density of the emission current from the source.

After integration we obtain

$$i' = \frac{i_0}{4}\{\Phi[\sqrt{\alpha}(x+\tfrac{1}{2}a)]-\Phi[\sqrt{\alpha}(x-\tfrac{1}{2}a)]\}\times$$
$$\times\{\Phi[\sqrt{\alpha}(z+\tfrac{1}{2}b)]-\Phi[\sqrt{\alpha}(z-\tfrac{1}{2}b)]\} \quad ......(16')$$

where $$\Phi[\xi] = \frac{2}{\sqrt{\pi}}\int_0^{\xi}\exp(-\xi^2)\,d\xi$$

This general expression for the current density of a beam of rectangular cross-section gives little information. In practice there are beams of rectangular cross-section or strip beams whose one transverse dimension is larger by one or two orders of magnitude than the other, i.e. the electrons are emitted from a narrow, long cathode. For such a case it is permissible to consider the cathode as infinitely long and (16′) then becomes

$$i = \lim_{a\to\infty} i' = \tfrac{1}{2}i_0\{\Phi[\sqrt{\alpha}(z+\tfrac{1}{2}b)] - \Phi[\sqrt{\alpha}(z-\tfrac{1}{2}b)]\} \qquad \text{......(17)}$$

which is illustrated in Fig. 3.

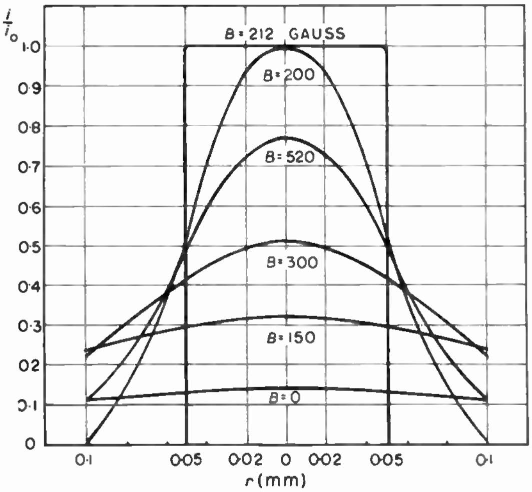



**Fig. 3.** Current density distribution in a strip beam ($y$ = 5 mm, $E$ = 200 V/cm, $b$ = 0·1 mm).

Extreme values of current density as a function of the magnetic-field intensity occur at magnetic-field intensities determined by the same conditions as in the case of the point source (in which, of course, the total current through the aperture is considered), i.e. for $\sqrt{\alpha} = \infty$ and its minimum values for $D\omega = \tan(D\omega)$. In the plane $y$ = constant, for distances from the axis in the interval $-\tfrac{1}{2}b < z < \tfrac{1}{2}b$, the current density assumes its maximal values for $\sqrt{\alpha} = \infty$ and its minimum values for $D\omega = \tan(D\omega)$. For distances from the axis within the above-mentioned interval, the current density is minimum for $\sqrt{\alpha} = \infty$ and maximum for $D\omega = \tan(D\omega)$.

The condition $\sqrt{\alpha} = \infty$, i.e. $D\omega = N\pi$, is the condition for the appearance of an image of the source in the plane $y$ = constant. Since in a homogeneous magnetic field the source images with unity magnification, the current density in all points of the image is equal to the current density of the source, and in all other points of the image plane must be equal to zero:

$$i_0 = \tfrac{1}{2}i_0\{\Phi[\infty] - \Phi[-\infty]\} = i_0 \quad \text{for } -\tfrac{1}{2}b < z < \tfrac{1}{2}b$$

$$i_0 = \tfrac{1}{2}i_0\{\Phi[\pm\infty] - \Phi[\pm\infty]\} = 0 \quad \text{for } \tfrac{1}{2}b < z < \infty$$

The dependence of current density on the intensity of the magnetic field is shown for three values of the coordinate $z$ in Fig. 4. Of particular importance is the current density on the axis of the configuration $i = i_0\Phi[\tfrac{1}{2}b\sqrt{\alpha}]$, since for a width of an aperture much smaller than the width of the source this value in practice determines—after multiplication by the appropriate factor—the total current. The over-all character of this dependence is the same as the character of the function (8). It is substantially different from the dependence of current density on magnetic-field intensity for a point source (6), since the current density of a beam emitted from a plane source is everywhere finite.

In analogy to the concept of relative current peak for a point source it proves advantageous to introduce the relative current density peak on the axis of the configuration (since relative current peak would here lead to expressions too complex and of too little importance for practical utilization, as will be seen in the derivation of the expression for the total current through a strip aperture),

$$P_i(\omega) = \frac{i_0 - i_{\min}}{i_0} = 1 - \Phi\left[\frac{\sqrt{A}}{D}\,\frac{1}{\cos D\omega}\,\frac{b}{2}\right] \qquad \text{......(18)}$$

One can determine the value of $1/(\cos D\omega)$ from this expression and from Table 1 one can find the next higher value of $1/(\cos D\omega_m)$ from which the required magnetic field can be determined. For the case of a variable electrostatic field the relative current density peak on the axis of the configuration is given by the expression

$$P_i(E) = 1 - \Phi\left[\sqrt{A}\omega\,\frac{b}{2}\right] \qquad \text{......(19)}$$

from which the required magnetic-field intensity can be readily determined.

The total current through a rectangular aperture in the anode is evidently

$$I = a'\,\frac{i_0}{2}\int_{-\frac{1}{2}b'}^{\frac{1}{2}b'} \Phi[\sqrt{\alpha}(z+\tfrac{1}{2}b)] - \Phi[\sqrt{\alpha}(z-\tfrac{1}{2}b)]\,dz$$

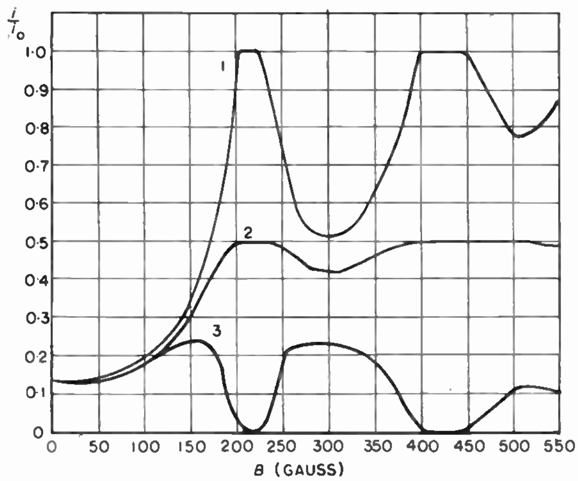



**Fig. 4.** Dependence of current density in a strip beam on the intensity of the magnetic field ($y$ = 5 mm, $E$ = 200 V/cm, $b$ = 0·1 mm); (1) $z$ = 0, (2) $z$ = 0·02, (3) $z$ = 0·1 mm.

if $P = a'b'$ is the area of the aperture. If the indicated integration is carried out and the values of aperture and source areas and of total emitted current are substituted, one obtains

$$I = \frac{I_0}{2}\frac{P}{P_0}\left\{\left(\frac{b}{b'}+1\right)\Phi\left[\sqrt{\alpha}\left(\frac{b+b'}{2}\right)\right] - \left(\frac{b}{b'}-1\right)\Phi\left[\sqrt{\alpha}\left(\frac{b-b'}{2}\right)\right] + \frac{\exp\left[-\left(\frac{b+b'}{2}\right)^2\right] - \exp\left[-\left(\frac{b-b'}{2}\right)^2\right]}{\frac{1}{2}b'\sqrt{\pi\alpha}}\right\} \qquad \ldots\ldots(20)$$

where $I_0$ is the emitted current, $P_0$ is the area of the source, $I$ is the current passing through the anode aperture, $P$ is the area of the aperture, $b'$ the width of the aperture, and $b$ the width of the source.

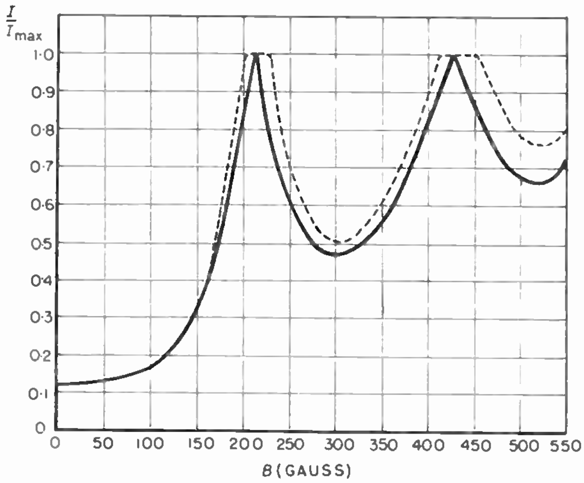



**Fig. 5.** Dependence of the current through an anode aperture of rectangular cross-section on the intensity of the magnetic field for a strip beam. Dashed curve for $b' = 0{\cdot}2b$; solid curve for $b' = b$.

In the case in which the widths of the aperture and of the source are the same ($b = b'$), eqn. (20) reduces to the form

$$I = I_0\left\{\Phi[b\sqrt{\alpha}] + \frac{1}{\sqrt{\pi\alpha}b}\left[\exp(-\alpha b^2)-1\right]\right\}$$

The shape of the curve of total current through the aperture as a function of the magnetic-field intensity will be similar to that for the current density on the axis of the configuration. Maxima and minima will be determined by the same conditions as follow from (17) for the case of $z = 0$ and from eqn. (20). Figure 5 shows the dependence of $I/I_0$ on the magnetic-field intensity for the case in which the width of the aperture is equal to the width of the source. A comparison with Fig. 4, which represents the dependence of current density (on the axis) on the magnetic-field intensity, shows that in the least favourable case, i.e. if the width of the aperture is equal to the width of the source, $I/I_0$ differs very little from $i_z/i_0$; the main difference is that the maxima of $I/I_0$ are sharper and that the minima reach somewhat lower values than for $i/i_0$. In practice, therefore, it is possible to utilize relative current density peak on the axis in place of relative current peak in the determination of the required magnetic field.

### 4. A Beam of Circular Cross-section Formed by Emission from a Plane Source

In considering the point source a difficulty arose as a result of the fact that the current density at the image points assumes an infinite value. This purely mathematical difficulty did not arise in the investigation of the total current through the aperture. However, a further difficulty is associated with the utilization of the theory of a point source, and that of a practical nature. This is the question under what circumstances, in view of a given aperture size, one can consider an actual source to be a point. One can say approximately that this will be the case when the source area is smaller by at least an order of magnitude than the area of the aperture. The theory of the point source cannot be, therefore, utilized for a very small aperture. If we replace the point source by a plane circular source, we obtain a current density from (16)

$$i = \frac{i_0}{\pi}\alpha\int_0^R\int_0^{2\pi}\exp[-\alpha(\rho^2+r^2)\exp(2\alpha r\rho\cos\phi)]\,r\,dr\,d\phi \qquad \ldots\ldots(21)$$

Here we have introduced polar coordinates by means of the relationships

$$x = \rho\cos\theta, \qquad x_0 = r\cos\phi$$
$$z = \rho\sin\theta, \qquad z_0 = r\sin\phi$$

and the limits of integration of the appropriate circular area; $R$ is the radius of the source.

(*Note:* in place of $\cos(\theta-\phi)$ we write simply $\cos\phi$ in the exponent, since the value of the integral between the limits 0 and $2\pi$ does not depend on the value of the constant $\theta$, as can be easily shown.)

The integral

$$\int_0^{2\pi}\exp(2\alpha r\rho\cos\phi)\,d\phi = \mathscr{I}_0(2\alpha r\rho)$$

(Bessel function of zero order and imaginary argument), so that eqn. (21) becomes

$$i = i_0 . 2\alpha\exp(-\alpha\rho^2)\int_0^R\exp(-\alpha r^2)\,\mathscr{I}_0(2\alpha r\rho)r\,dr \qquad \ldots\ldots(21')$$

If we express $\mathscr{I}_0(2\alpha r\rho)$ by means of an infinite power series and integrate eqn. (21′) term by term, we obtain

$$i = i_0 \exp(-\alpha\rho^2) \sum_{m=0}^{\infty} \frac{\alpha^m \rho^{2m}}{(m!)^2} \times$$
$$\times\{1-\exp(-\alpha R^2)[(\alpha R^2)^m + + \sum_{k=1}^{m} m(m-1)\ldots(m-k+1)(\alpha R^2)^{m-k}]\} \quad \ldots\ldots(22)$$

The relation (22) thus expresses the current density emitted by a plane circular source in the form of an infinite series arranged according to powers of $\alpha\rho^2$. For values of the magnetic field corresponding to $\alpha = \infty$, the current density in the plane $y$ = constant should assume values $i = i_0$ or $i_0 = 0$ according to whether $\rho < R$ or $\rho > R$. It has not been possible to determine the limits of eqn. (22) for $\alpha \to \infty$ in the general case. Only for the case $R \to \infty$ is it possible to show that $i = i_0$; one can therefore estimate that for $R$ large compared with $\rho$, the current density behaves in the same way as on the axis of the configuration, since

$$i_{\rho=0} = \frac{i_0}{\pi}\alpha \int_0^R \int_0^{2\pi} \exp(-\alpha r^2)\, r\, \mathrm{d}r\, \mathrm{d}\phi = i_0[1-\exp(-\alpha R^2)] \ldots\ldots(23)$$

In the determination of the total current through the circular aperture located on the axis of the source one can distinguish the following cases:

(a) The radius of the aperture is smaller by an order of magnitude than the radius of the source;

(b) The radius of the aperture is larger by an order of magnitude than the radius of the source;

(c) Both radii are of the same order of magnitude.

In case (a) it is possible to consider the current density to be independent of the distance ($\rho$) from the axis and equal to the current density on the axis (23), so that the current $I$ through the anode aperture is given by the relation

$$I = I_0 R_s^2 \frac{1-\exp(-\alpha R^2)}{R^2} \quad \ldots\ldots(24)$$

where $R_s$ is the radius of the aperture and $I_0$ is the total emission from the source.

In case (b) it is possible to consider the source to be a point and the current through the aperture is determined by expression (8), so that

$$I = I_0[1-\exp(-\alpha R_s^2)] \quad \ldots\ldots(25)$$

In case (c) it is necessary to refer to the general expression (22) so that

$$I = 2\pi i_0 \sum_{m=0}^{\infty} \frac{m}{(m!)^m} \left\{1-\exp(-\alpha R^2)\left[(\alpha R^2)^m + \sum_{k=1}^{m} m(m-1)\ldots(m-k+1)(\alpha R^2)^{m-k}\right]\right\} \times \int_0^{R_s} \exp(-\alpha\rho^2)\rho^{2m+1}\,\mathrm{d}\rho$$
$$= \frac{I_0}{\alpha R^2} \sum_{m=0}^{\infty} \frac{1}{(m!)^2} \left\{1-\exp(-\alpha R^2)\left[(\alpha R^2)^m + \sum_{k=1}^{m} m(m-1)\ldots(m-k+1)(\alpha R^2)^{m-k}\right]\right\} \times$$
$$\times \left\{1-\exp(-\alpha R_s^2)\left[(\alpha R_s^2)^m + \sum_{k=1}^{m} m(m-1)\ldots(m-k+1)(\alpha R_s^2)^{m-k}\right]\right\} \ldots\ldots(26)$$

In practice it will be necessary to limit oneself to the first few terms of expression (26). For large values of $\alpha$ it would be necessary to take a considerable number of terms but that is not possible in practice since the use of even the first few terms leads to complex expressions and tedious computations. (If it is considered that each term also contains the sum of a finite series, it is seen that the higher-order terms are progressively more complex.)

If we choose in the zeroth approximation expression (24) for the current flowing through the anode aperture it is evident that it will be always larger than the actual value, since the current density is at its maximum on the axis. On the other hand, an approximation by means of a finite number of terms of the series (26) will always give smaller values than the actual value. The actual current intensity will thus lie between the zeroth and the utilized $n$-th approximation of the series (26). It has been determined that for values $R^2 = 1$ (when we also have $R_s^2 = 1$), even the second approximation $I^{(2)}$ gives good results and differs only slightly from the first approximation $I^{(1)}$:

$$I^{(1)} = I_0 \frac{1-\exp(-\alpha R_s^2)}{\alpha} \frac{1-\exp(-\alpha R^2)}{R^2} \quad \ldots\ldots(27)$$

$$I^{(2)} = I^{(1)}\left\{1 - \frac{[1-\exp(-\alpha R_s^2)-\alpha R_s^2 \exp(-\alpha R_s^2)][1-\exp(-\alpha R^2)-\alpha R^2\exp(-\alpha R^2)]}{[1-\exp(-\alpha R_s^2)][1-\exp(-\alpha R^2)]}\right\} \quad \ldots\ldots(28)$$

In Fig. 6 the zeroth and second approximations are shown for the case $R_s = R/2$. It is evident that second approximations gives good results for values $\alpha R^2 \leqq 1$.

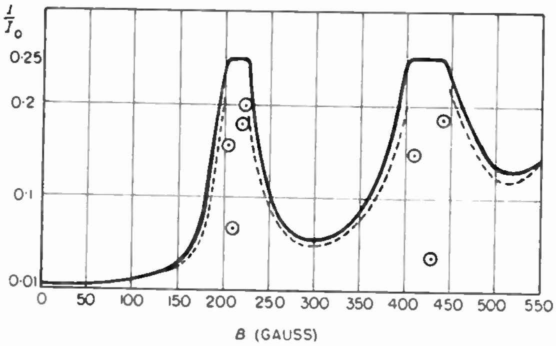



**Fig. 6.** Dependence of the current through an anode aperture on the intensity of the magnetic field for a beam of circular cross-section produced by emission from a plane source ($y = 5$ mm, $E = 200$ V/cm, $R = 0{\cdot}05$ mm, $R_s = 0{\cdot}5R$).

For values $\alpha R^2 > 1$ this approximation is insufficient and gives wrong values (shown in the graph by circles). For such values little can be done except to be satisfied with the zeroth approximation, which differs from the actual dependence of the current intensity on magnetic intensity particularly in that it shows less sharp maxima.

*Translation received by the Institution on 4th February 1960 (Paper No. 627).*

## 5. Conclusions

In this paper we have determined the distribution of current density in beams focused by a longitudinal homogeneous magnetic field and we have determined the total current flowing through an anode aperture, and that as a function of the intensity of the magnetic field for a beam of circular and rectangular cross-section. It was found that this dependence has a characteristic shape with minima and maxima. Their location was determined and an expression for the corresponding intensity of the magnetic field was found. It was found useful to define a relative current peak (or relative current density peak on the beam axis) as the ratio of the difference between maximum and minimum current intensity to the maximum intensity. The negative logarithm of this quantity is a measure of the focusing of given configuration and permits a ready determination of the required intensity of the focusing field.

## 6. References

1. D. B. Langmuir, "Theoretical limitations of cathode ray tubes", *Proc. Inst. Radio Engrs*, **25**, pp. 977–91, August 1937.
2. C. C. Cutler and M. E. Hines, "Thermal velocity effects in electron guns", *Proc. Inst. Radio Engrs*, **45**, pp. 307–15, March 1955.
3. J. L. Stewart, "Electron flow through small tubes with magnetic focusing", *J. Appl. Phys.*, **24**, pp. 1236-40, September 1953.
4. J. R. Pierce and L. R. Walker, "'Brillouin flow' with thermal velocities", *J. Appl. Phys.*, **24**, pp. 1328–30, October 1953.

# Science and Parliaments

The Parliamentary and Scientific Committee celebrates this year the twenty-first anniversary of its foundation. It was celebrated at the annual meeting and subsequent lunch of the Committee held on 16th February last at the Savoy Hotel, London. It also was marked by the election of the Right Hon. The Marquess of Salisbury, K.G., as President of the Committee. In his address, Lord Salisbury said:

"The Parliamentary and Scientific Committee occupies a quite unique position in the world today. It is composed of politicians, scientists and those who are engaged in the practical application of science. It is completely unofficial and yet it has been blessed by the heads of all political parties.

"The news that a European Parliamentary and Scientific Conference is going to be held here in the next few weeks may give us hope that an attempt is going to be made by other countries to set up similar organizations in their own lands. If so, I know it will be our most sincere hope that they will flourish.

"Some institution of this kind has no doubt become an absolute necessity for any country which wishes to be thought progressive in this uncharted world into which we are moving. For after all, as we are often told, we live in an age of change unparalleled in history, an age when science and the practical applications of science have come into their own as never before and are entirely transforming our whole existence."

Referring to the advances being made in space research and other spheres of science, Lord Salisbury said:

"We must assume that this is only the beginning—the beginning of perhaps even vaster progress ahead. In such circumstances it is vitally important that the closest contact should exist here in this country between the scientist, those concerned with the appropriate applications of science, and politicians and the Government, contacts which enable them to meet each other, to learn from each other, enabling the scientists to tell the politicians what *ought* to be done, and the politicians to tell the scientists what *can* be done.

"And now we have been provided with an even more highly placed channel of communication between Science and Parliament, by the creation of a Minister for Science, who is a member of the Cabinet."

In his speech The Lord President of the Council and Minister for Science, The Rt. Hon. Viscount Hailsham, Q.C., said that because Parliamentary Government operated on a Party basis it was essential that there should be special facilities, in matters where party politics were not appropriate, to see that an objective approach was maintained. This was being done through the Parliamentary and Scientific Committee.

Lord Hailsham continued:

"This country cannot afford *not* to concentrate upon science and technology: we are confronting difficulties created by the inadequacy of the scientific and general education of the period between the two wars, something from which we shall not recover at once and for which we are now paying a heavy price."

Lord Hailsham felt that the time had now come to put a little more emphasis on what was being done about the application of science. "It is not that the pure and fundamental work ought not to go on, or ought not to go on increasing. It is in the field of applied science that we now want to look for the most rapid growth. Have we really developed the means of adopting sufficiently quickly known scientific techniques? I would think probably not. Are not our scientists concentrating too much in a few forward-looking industries, leaving the great mass of industry not quite so well served by scientists and engineers? I would think that this can be established to be the case. Do brains of the first quality apply sufficient attention to the problems of applied science in this country, as distinct from problems of pure science? I think we have to remember that the two things go hand in hand.

"It is, for instance, very good to know that the Cambridge scientists have made remarkable studies in the field of the ultimate origin of the universe—pure science if ever there was pure science. But how many of those discoveries, and how much of that work at Cambridge, or at Jodrell Bank, could have been done if it had not been for the work of the applied scientists in constructional engineering, electronics and radio engineering? I would feel that fundamental and applied science were no more members of two separate cultures than the scientist and the artist ought to be. These things must walk hand in hand. I would like you to help to give an impetus, therefore, to the movement to use science in our everyday lives.

"I would like to feel that you were supporting the D.S.I.R. in their study of the manpower and research needs of individual industries, and in the attempts which they have been carrying out to publicise locally and at their regional centres, for the benefit of smaller and medium-sized firms, the discoveries of science. I would like to feel that you were also supporting the drive for more generously financed Research Associations."

He continued: "I would also like us all to devote a little thought to the theme of international co-operation. We have made a number of tentative approaches: the experiments of the International Atomic Energy Agency, and also the newly formed European Space Research Organization. But this is essentially a question of the right development of limited manpower and limited resources. There are some which are better done on the basis of the individual institution, whether it is the individual firm or the individual university. There are other things which are better done on a national scale and rightly made the subject of national pride when they are done well. Increasingly over the whole field of science the research scientist is demanding expensive apparatus. Governments are faced with the alternative either of not going into a number of potentially fruitful fields at all—an expedient which must at best be a terrifying gamble, in as much as we all of us know that in principle the success of particular experiments cannot be predicted—or clubbing together in order to provide the ultimate discoveries upon which future wealth and future technology can be based.

"On the question of space research, I would say that it might very well be that we could go on for another generation of space research experiments entirely alone. £60 million for the launcher, £12 million for the satellite and whatever number of million pounds you like to say for each shot, are not perhaps beyond our resources, but what about the third generation of experiments which will thereafter be generated? How much longer can individual countries go on? Sooner or later the time will come when nations will have to work together if they want to pursue the scientific work which the scientists require."

Lord Hailsham said he would end by sounding a different note. He was profoundly disturbed by what he called the flight from reason in the modern world. It showed itself in many emotional attitudes not usually connected with science but repugnant to it—excessive nationalism, suppression of evidence, violence, unfriendliness and want of personal integrity in society. "If science is going to be the influence for good that it could be, it cannot leave morality entirely to the churches or to the humanists."

## The First European Parliamentary and Scientific Conference

Appropriately in the twenty-first year of the Parliamentary and Scientific Committee, the first European Parliamentary and Scientific Conference was held in London from 20th to 23rd March, 1961. Representatives from seventeen European countries attended and were provided with a programme of activity to show the work of the British Committee.

The conference was organized in response to the request of the Council of Europe and the Organization for European Economic Co-operation.

Visits to research associations and attending meetings of working parties of the Parliamentary and Scientific Committee were featured in the arrangements made for the European delegates.

The President, Professor E. E. Zepler, and other Officers of the Brit.I.R.E. gave a luncheon in the River Room of the Savoy Hotel on Tuesday, 21st March, which provided an opportunity for informal discussions on technical education and the development of radio and electronics in some European countries.

One of the most important meetings, and attended by all delegates, was a general meeting of the Parliamentary and Scientific Committee held in the Grand Committee Room, Westminster Hall. The meeting was attended by His Royal Highness The Duke of Edinburgh, K.G., who, in welcoming the delegates, said:

"I am only an Honorary member of this Committee but I think perhaps the other members would allow me to offer our welcome to all of you who have come here to this conference. Whatever you learn here I hope will be valuable to you when you get home. I believe this Parliamentary and Scientific Committee is doing a very useful job here, and I hope you will use what you discover here to establish the same sort of thing, or something similar, at home, particularly because it seems to me absolutely vital, at this particular point in history, that those people who are at the centre of the administration of affairs in public life should keep in close touch with what is going on in the scientific world, and this seems to me a very good way of doing it. I hope you all have a very interesting Conference."

It is understood that the next European Parliamentary and Scientific Conference may be held in Strasbourg.

U.D.C. 621.373.421

# A Transistorized Frequency Synthesizer

By

G. HUSSON, B.Sc.†

AND

B. N. SHERMAN, B.Eng.†

*Presented at a Symposium on Stable Frequency Generation, held in London on 25th May, 1960. (The paper was read on behalf of the authors by Mr. R. L. J. Awcock.)*

**Summary:** The synthesizer provides 30 000 discrete frequencies between 2–32 Mc/s in steps of 1 kc/s with the stability of the driving frequency standard, using the arithmetic processes of addition, subtraction, multiplication and division, with the most economical use of a set of readily available basic numbers. Use is made of the phase-locked loop principle. A variable frequency oscillator covers the 2–32 Mc/s range and, for each desired frequency, a suitable arrangement of frequency conversion and filtering allows the oscillator to lock on a fixed frequency directly derived from the standard. Spurious signals and noise generated respectively by mixers and harmonic generators, are greatly attenuated, since the phase-locked loop acts as an "active filter". In particular the "catching range" of an automatic phase control loop is examined and a method is shown whereby this catching range can be greatly increased by very simple and economical means.

## 1. Introduction

In this discussion, the term "frequency synthesizer" is reserved exclusively for equipment generating output frequencies which are completely dependent on a single frequency standard. The basic elements used in frequency synthesizers are:

1. Frequency standard
2. Frequency dividers
3. Frequency multipliers
4. Frequency changers
5. Controlled oscillators

### 1.1. *Frequency Standard*

Precision quartz crystal controlled oscillators are presently available with frequency stabilities of the order of one part in $10^9$ per day. Higher stabilities are available with the use of atomic and molecular resonance as a reference standard.

### 1.2. *Frequency Dividers*

Many varieties of circuits are available as frequency dividers, ranging from relaxation oscillators such as multivibrators and blocking oscillators, to capacitive step charging circuits. Digital counters, regenerative modulation frequency dividers, and locked oscillators may also be used as frequency dividers.

The choice depends on the particular system requirements such as upper frequency limit, required reliability, etc.

### 1.3. *Frequency Multipliers*

Non-linear tuned amplifiers have long been used as frequency multipliers. Later techniques such as the "drift-cancelled oscillator", the "impulse-governed oscillator", and the "automatic phase control" loop offer attractive advantages for high multiplication ratios.

### 1.4. *Frequency Changers*

Conventional frequency changers or mixers, coupled with appropriate filters, are widely used to translate frequencies suitable for synthesizing the required output.

### 1.5. *Controlled Oscillators*

The impulse-governed oscillator, and the closely related automatic phase control (a.p.c.) loop are examples of integral control mechanisms widely studied under the class of servo-mechanism control systems. The most important feature of this class of controlled oscillator is that the system requires no steady-state error of the controlled variable, but instead utilizes an error in the integral of the controlled variable, i.e. an error in phase difference.

### 1.6. *System Design*

The design of all frequency synthesizers is based on the combination of these basic elements discussed above. Many combinations are possible, and the ultimate choice depends on the specific requirements such as frequency range, frequency increments, spurious output, jitter, ease of operation, size, weight, reliability and economic factors.

† Canadian Marconi Company, Montreal, Canada.

The specific equipment described in the subsequent section is designed to cover the frequency range of 2–32 Mc/s in 1 kc/s increments with low values of spurious output. Small size, light weight, ease of operation and reliability were prime requirements. These requirements were met through:

(1) Complete transistorization of the equipment.

(2) Use of the minimum number of controlled oscillators.

(3) Use of simple non-critical passive filters in the frequency selection process.

(4) Strict utilization of reliable basic circuit elements.

## 2. Frequency Generation Scheme

The frequency generation scheme shown in Fig. 1 can best be understood by reading the block diagram from left to right.

A variable frequency oscillator, VFO1 is tuneable over the frequency range of 2–32 Mc/s. The process of setting up the five-decade digits to the desired frequency automatically adjusts the output oscillator to the correct frequency within the first three significant digits. For example, a selected frequency of 11·987 Mc/s results in the oscillator being set to approximately 11·9 Mc/s. The precise tuning of the oscillator is achieved with a voltage controlled reactance as part of an automatic phase control (a.p.c.) loop.

Simultaneously the appropriate harmonic of the frequency standard and its sub-multiples (derived from three decade dividers) are selected and injected into five series connected frequency changers, FC1 to FC5 inclusive. The final frequency thus obtained is compared in phase with the reference frequency of 26 kc/s directly derived from the frequency standard.

The specific choice of filter frequencies and the selected harmonics, are primarily determined by spurious and image frequency considerations, plus the simple realization of the filters involved. However, the levels of spurious signals thus achieved at the two inputs of the phase discriminator are normally much greater than the desired spurious content at the final output. The a.p.c. loop locks VFO1 to the reference frequency, and in addition functions as a highly selective filter, thus greatly minimizing the output spurious content.

## 3. Phase Locked Loop Characteristics

Automatic phase control systems have been used extensively in telemetry equipment to track signals in the presence of noise, and it is in fact one of their most important characteristics that they can be used as active filters of extremely narrow bandwidth.

An a.p.c. loop can be completely defined if we know its noise bandwidth and its holding range, i.e. the tuning range of the oscillator over which the latter remains synchronized. It has been shown[1, 2, 3] that this holding range is proportional to the gain of the loop according to the following formula:

$$\sin\phi = \frac{\Delta\omega}{K}$$

where $K$ is the gain, $\Delta\omega$ the oscillator detuning, and $\phi$ the steady-state phase error.

This equation shows that the holding range cannot exceed $K$.

The noise bandwidth of the loop is defined as:

$$B = \int_{-\infty}^{+\infty} \left[\frac{\phi_o}{\phi_i}(j\omega)\right]^2 d\omega$$

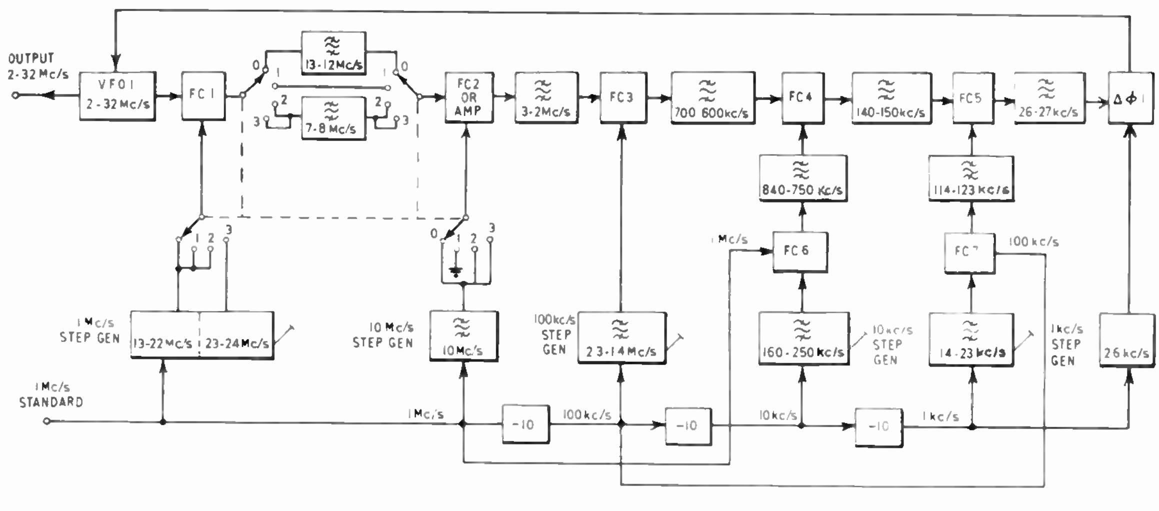



**Fig. 1.** Block diagram of the frequency synthesizer.

Calculations show that a loop with a single RC network following the phase discriminator has a noise bandwidth proportional to the gain, i.e.

$$B = \pi K$$

Therefore the requirements for narrow noise bandwidth and large holding range are incompatible. This contradiction may be overcome by the use of an integrating network in the loop as shown in Fig. 2.

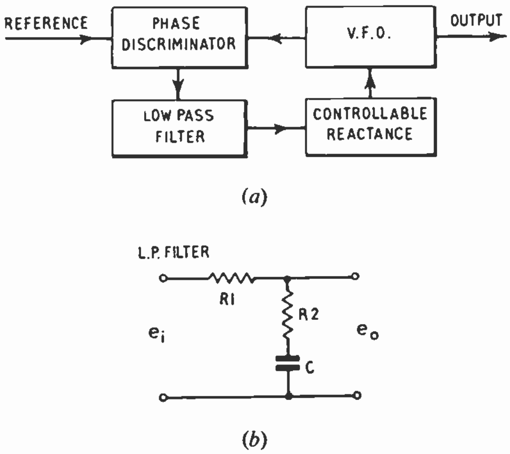



**Fig. 2.** (*a*) A typical automatic phase control loop.
(*b*) The low-pass filter.

Its transfer function is given by $\dfrac{e_o}{e_i} = \dfrac{1 + t_1 p}{1 + t_2 p}$
where $t_1 = R_2 C$ and $t_2 = (R_1 + R_2)C$.

In this case, the noise bandwidth, $B = 2\pi\omega_n$, where $\omega_n$ is the natural resonant frequency of the system, in the absence of damping.

Also $\omega_n^2 = \dfrac{K}{t_2}$ where $t_2$ is the filter lag time-constant.

Therefore by keeping $K$ constant (i.e. keeping the holding range constant) one can theoretically make $\omega_n$ (hence the noise bandwidth) as small as possible by simply increasing the time-constant $t_2$.

The lead time-constant $t_1$ of the integrating network is then given by

$$t_1 = \frac{1}{\omega_n} - \frac{1}{K}$$

The above formulae are based on an optimized value of damping ratio

$$\xi = \frac{1 + K t_1}{2\omega_n t_2} = 0{\cdot}5$$

Hence the use of an integral control network gives the freedom to choose a convenient holding range irrespective of the desired narrow noise bandwidth.

In particular, in a frequency synthesizer as the one described above, rejection of spurious frequencies and noise is achieved by this low-pass filter.

However, two important factors remain to be considered: the catching range and the synchronizing time of an a.p.c. system. The catching range is that within which the oscillator can be pulled into synchronism. Both phenomena depend essentially on the time-constant of the low-pass filter and on the gain of the a.p.c. system. Here again a compromise should be made between an arbitrary narrow bandwidth, the desired catching range, and the synchronizing time. It can be shown that for a broadband a.p.c. system (i.e. by ignoring the low-pass filter), the catching range is equal to the holding range. The insertion of an integrating network reduces the bandwidth and decreases the catching range. It has been shown[3] that

$$|\Delta\omega|_{\text{catching}} = \sqrt{\omega_n K}$$

if we take the optimized case where the damping ratio is equal to 0·5.

This expression can be written:

$$\frac{|\Delta\omega|_{\text{catching}}}{|\Delta\omega|_{\text{holding}}} = \sqrt{\frac{\omega_n}{K}}$$

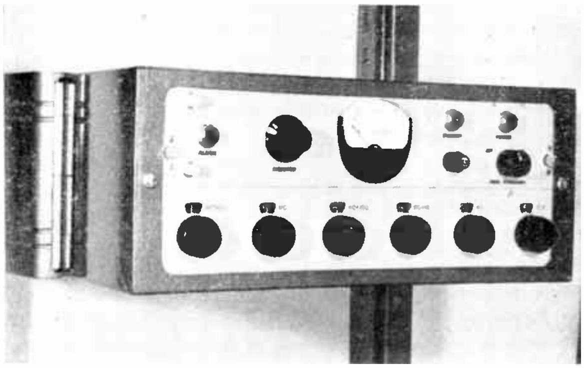
For a large gain and a narrow bandwidth ($\omega_n$ small) this ratio is very small and the catching range becomes many times less than the holding range. This can be tolerated only when the free-running frequency of the oscillator to be controlled is close enough to the reference at all times. In a practical frequency synthesizer this cannot be achieved because limitations exist as to the number of mechanically tuneable frequency steps that may be performed by the frequency selection dials.

Locking of the controlled oscillator can hence only be achieved with a catching range large enough to cover these frequency increments. This imposes two new requirements:

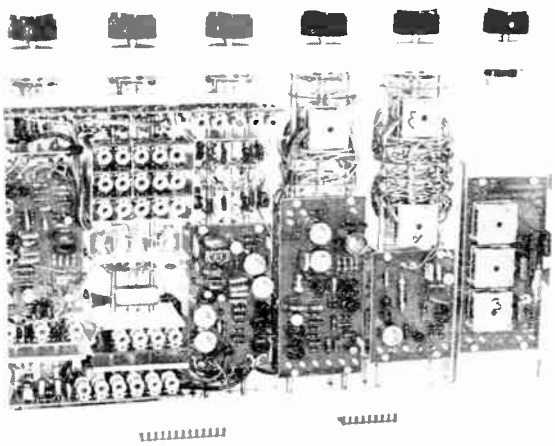
(1) The variable reactance must have a dynamic range sufficiently large to cover the required frequency change.

(2) The catching range of the a.p.c. loop becomes excessively small when the loop bandwidth is restricted to minimize spurious outputs.

The following method is employed to increase artificially the catching range. Capacitor C of the low-pass filter (Fig. 2) is momentarily short circuited with each selected switch digit. At that particular moment, the system is unlocked and a gating signal is made available so as to short circuit resistor R1 in the low-pass filter. Capacitor C now charges to a fixed potential through resistor R2 only, and the controllable reactance therefore sweeps rapidly through a predetermined frequency range. As the VFO output frequency approaches the selected frequency within 10 kc/s a signal appears at the output of FC4. This



(a)



(b)

Fig. 3. (a) External appearance of the frequency synthesizer. (b) Internal construction of the frequency synthesizer.

signal is used to restore resistor R1 in the low-pass filter to its original value. Thus, the required narrow bandwidth is restored, and the sweeping rate of capacitor C (hence controllable reactance) is considerably slowed down. This process continues until the frequency is within the catching range of the a.p.c. loop.

## 4. Conclusion

The most difficult problem in frequency synthesis is the suppression of spurious signals and noise. The a.p.c. loop enables the designer to reduce these spurious levels to values which cannot practically be achieved by conventional filter techniques. It also simplifies the requirements of the basic elements which make up the synthesizer, and makes possible the generation of frequencies in very small increments over a broad frequency range.

Transistorization of frequency synthesizers for the h.f. band can be easily achieved, with the usual resultant advantages of low power consumption, light weight and reliability.

## 5. References


1. Donald Richman, "The d.c. quadricorrelator: a two-mode synchronization system", *Proc. Inst. Radio Engrs*, **42**, pp. 288–99, January 1954.
2. H. T. McAleer, "A new look at the phase-locked oscillator", *Proc. Inst. Radio Engrs*, **47**, pp. 1137–43, June 1959.
3. W. J. Gruen, "Theory of a.f.c. synchronization", *Proc. Inst. Radio Engrs*, **41**, pp. 1043–48, August 1953.


*Manuscript received by the Institution on 23rd May 1960 (Paper No. 628)*

# 1961 Convention

## "RADIO TECHNIQUES AND SPACE RESEARCH"—OXFORD, 5th–9th JULY

### *Synopses of Papers to be presented at the Convention*

This is a selection of some of the papers which have been accepted for presentation during the seven sessions of the Convention. The sessions to which the papers have been *provisionally* allocated are indicated and will be confirmed later. The Convention Committee is considering over 50 papers at present and further lists of titles and synopses will be published in the May and June issues of the *Journal*.

**Critical Engineering Factors in the Design and Development of Space Systems** SESSION 1

J. M. BRIDGES. (*Office of the Director of Defense Research and Engineering, Washington, U.S.A.*)

This paper will deal with the basic concepts underlying the design and development of systems for space operations—space probes, orbital systems, or interplanetary flights. Following a consideration of experience gained in the past few years, the state-of-the-art and requirements for the next decade, the central point of discussion will be the importance of weight and reliability as system parameters, particularly in regard to cost, performance and development time schedules. How do the peculiar needs of a space system affect the design philosophy, the ground test programme and the developmental approach to an optimum system? Some suggestions will be made regarding the techniques and procedures to be used in design and development to ensure that the system will operate satisfactorily on its mission in the space environment. The paper will also point out certain work that is needed in basic research before we can design the reliable space systems that will enable our conquest of space.

**Engineering Aspects of Satellites and their Launching Rockets** SESSIONS 2/3

G. K. C. PARDOE. (*De Havilland Aircraft Company Limited.*)

The paper discusses the different engineering problems associated with satellites and probes, and their launching rockets. The different set of operating conditions encountered in these two main areas of space technology are examined and the motive power and control problems which arise are dealt with.

The basic requirements for the multi-stage launching rockets are discussed; rocket equipment at present under development is then reviewed, noting the limitations imposed by present rocket techniques and launching facilities. Future trends of rocket development are mentioned in relation to possible payloads

The composition of a typical satellite or space probe is then considered, with some attention being given to each of the different systems, such as structure, power, attitude and speed control, instrumentation and communication equipment, etc. Critical design criteria such as weight, and electrical power requirements are highlighted. The operating environment is then related to the choice of design of the various systems. Particular consideration is then given to several specific types of satellites and space probes, examining the differences which arise from the varied requirements, and assessing what integration may be achieved by the design of common systems for the various applications.

**The Reliability of Components in Satellites** SESSIONS 2/3

G. W. A. DUMMER, M.B.E., MEMBER. (*Ministry of Aviation, Royal Radar Establishment, Malvern.*)

The environmental conditions to which space vehicles are subjected place severe stresses on the components and the severity of these conditions is discussed, i.e. shock, vibration, acceleration, temperature extremes, low pressure, solar radiation, cosmic rays and aurorae, ozone content of space, dissociated gases, etc. The probable effect of these conditions on average components is discussed and hermetic sealing and protection problems are examined. The present programme of special quality components for guided weapon use is described and some of its difficulties outlined. The development of high quality components for use in under-sea cables is also relevant and is described. Finally, the possible effect of microminiaturization techniques is discussed and in particular the effect of these environments on solid circuits. A summary is given of the most reliable types of components for use in satellites.

**Power Supplies for Space Vehicles** SESSIONS 2/3

K. E. V. WILLIS, B.SC., A.R.C.S. (*National Research Development Corporation, London.*)

The requirements for electrical power in space vehicles over the next five years are reviewed and estimated figures given with particular reference to power-to-weight ratio. Vehicle life duration and its effect on choice of supply are discussed. Problems encountered in space vehicles due to environmental conditions, e.g. temperature, radiation, zero gravity etc., are then pointed out. The available generation methods are: conventional batteries; thermo-electric generation; thermionic generation; fuel cells; magneto-hydrodynamic methods; nuclear systems; solar power.

**Ultra-Violet Astronomy from Rockets and Satellites** SESSION 4

D. W. O. HEDDLE, PH.D. (*University College, London.*)

The astronomical observations which can be made from the surface of the earth are severely limited by the presence of the earth's atmosphere. These limitations can be overcome by conducting observations from a vehicle at a sufficient height above the earth. The inaccuracies introduced by atmospheric refraction (scintillation and apparent displacements) are extremely small. The vehicle stability required to offer significant advantages for astrometric observations is perhaps too great for serious consideration at the present time, but ultimately a system with a short-term drift of less than 0·01 seconds of arc per minute will require study.

At the present time, the most significant advantage to be obtained from the use of rockets lies in the extension of the observable spectrum to wavelengths less than 3000Å. The experiments so far carried out have involved the installation in rockets of detectors sensitive to wavelengths within a band of a few hundred angstroms centred in the 1000Å to 2000Å region. The limitations of this technique lie in the short observing times dictated by the roll rate of the vehicle, and a vehicle with a certain degree of stability is urgently required. With the "pre-determined" vehicle it will be possible to obtain crude spectra of selected bright stars in the short time offered by a sounding rocket flight. The more stable vehicle would enable spectra of good resolution to be obtained from a satellite, or detailed broadband photometry to be carried out from a rocket.

**Overall System Requirements for Low Noise Performance** SESSIONS 6/7

C. R. DITCHFIELD. (*Ministry of Aviation, Royal Radar Establishment, Malvern.*)

With the use of low noise amplifiers the limit of system performance may well be not the receiver noise itself, but rather the ambient noise background against which the signal is sought and the noise which is generated by the aerial, transmission lines or other components of the complete system. For a ground based system receiving signals from a satellite, considerations of galactic noise and atmospheric attenuation lead to a choice of wavelength in the centimetric band, as the optimum with regard to source noise. The optimum frequency for satellite to satellite communication would be much higher. The practical limits which may be achieved in the various components of a microwave system are discussed and the overall performance is assessed.

**The Advantages of Attitude Stabilization and Station Keeping in Communications Satellite Orbits** SESSIONS 6/7

W. F. HILTON, PH.D. (*Hawker Siddeley Aviation Limited.*)

Random distribution of non-oriented satellites have been suggested for communications purposes. Such a system may require 50 satellites as against 3 attitude stabilized satellites spaced evenly round their orbit. Attitude stabilization which will direct the satellite radio antenna at the earth may cost 200 lb of complicated payload in order to gain a factor of say 10 on power consumption, by not sending radio to empty space. We are thus replacing costly solar cells (£1M per kW) by less costly attitude stabilization. Clearly attitude stabilization will pay dividends in high radio power satellites, and be uneconomic for low radio powers.

An intangible disadvantage of the non-attitude stabilized satellite is the zero axial field strength from a dipole. This may be eliminated for example by using a pair of phased and perpendicular dipoles, while incurring a further weight and power penalty of double. Alternatively, by spinning the satellite with its axis horizontal the radiated signal will always meet the earth's surface, and eliminate rotational period fading. This use of spin stabilization may well limit the choice of orbit for spin stabilized satellites, as the electromagnetic braking from the earth's magnetic field will quickly damp out the spin in many orbits.

No such limitations apply to the orbits of attitude stabilized satellites, which can be biased to cover given areas, say the Northern hemisphere, or to cover given periods of time, say the sunlit hours as the sunlight advances westward round the world. This may imply the use of elliptical orbits.

**Television Communications using Earth Satellite Vehicles** SESSIONS 6/7

L. F. MATHEWS, MEMBER. (*Associated Television Limited, London.*)

Although much work has still to be done before public service transmission of radio and television using satellites can be expected, many experiments are now proceeding, particularly in the United States. The first part of the paper will consist of considerations on the choice of an orbit, and the signal carrier frequency best fitted for television satellite relays or transmissions. The types of equipment likely to be employed, such as travelling wave tubes, masers and micro-module components, are discussed. Following a survey of experience so far gained with operational weather satellites, including *Tiros I* and *II*, the paper concludes with a speculation on further developments towards television transmission and relays, using the space vehicle, and random dipoles in orbit.

U.D.C. 537.228: 621.372.54

# Piezoelectric Ceramic Transformers and Filters

By

ALAN E. CRAWFORD,
(*Member*)†

*Presented at the Symposium on New Components held in London on 26th–27th October 1960. The work described formed part of a longer paper read at a meeting of the Southern Section in Southampton on 14th December 1960.*

**Summary:** The development of improved piezoelectric ceramics has enabled these materials to be used in the design of electronic components. Two new devices are described embodying piezoelectric principles of operation and relying on electromechanical resonance to obtain efficient energy transformation. The ceramic transformer provides a high voltage source without involving insulation problems or magnetic fields. The Transfilter is used as a miniature filter or an interstage impedance matching transformer suitable for radio receivers. Both are essentially solid-state in character and possess all the advantages of such systems. They represent an entirely new approach to component development and can be considered as forerunners of other devices using similar principles.

## 1. Introduction

The use of components employing piezoelectric principles of operation has, with one notable exception, been restricted to generator or motor functions. The former is well illustrated by a gramophone pick-up element while an example of the latter is the ultrasonic generator transducer. The exception is the quartz unit which is used as a filter or frequency standard.

A circuit component must generally possess high stability in terms of temperature and time, and until recently the only piezoelectric materials that have satisfied these conditions has been quartz and certain water soluble single crystals. Unfortunately, the piezoelectric characteristics of this group of stable materials are not always suitable for a wider exploitation of other possible devices. However, the recent development of the polycrystalline piezoelectric ceramics has provided the component designer with an adaptable material and this paper describes two new components employing the piezoelectric effect found in lead zirconate titanate ceramics.

## 2. Piezoelectric Principles

When a piezoelectric material is subjected to mechanical strain an electric charge appears across selected electrode faces. Similarly there is a motor effect when voltage is applied to the electrodes, producing an alteration to the physical dimensions of the material. The former is known as the direct piezoelectric effect and the latter is known as the converse piezoelectric effect.

A large number of single crystals exhibit piezoelectric properties to a greater or lesser degree. The direction of maximum activity varies with different crystals and is usually defined in terms of the crystallographic axis. A plate cut from the crystal with reference to a particular axis will therefore be mainly limited to this direction when used as a transducer. There are certain exceptions such as the 0 deg Z cut of tourmaline where a volume response is obtained, but generally other planes of activity are minor in character.

The class of materials known as polycrystalline ceramics possess the property of preferential directions of activity. This means that the direction in which the piezoelectric effect occurs is decided during manufacture by choice of electrode areas and the polarizing process.

Unlike a single crystal which possesses an inherent polarization the polycrystalline body of ceramic consists of a number of randomly oriented domains. The domains can be aligned by the application of an electrostatic field under certain controlled conditions. The effect is permanent after removal of the field and the polycrystalline mass then acts as a single crystal so far as the piezoelectric effect is concerned. This was first noticed by Gray in barium titanate[1] and exploitation of this material either in its basic form or with various additives has been very rapid. Barium titanate however possesses a number of disadvantages. It has a relatively low Curie point, that is, the temperature at which the domains revert to random distribution. The electromechanical conversion efficiency is quite low and there is also a large dependence of parameters on temperature.

† Brush Crystal Co. Ltd., Hythe, Southampton.

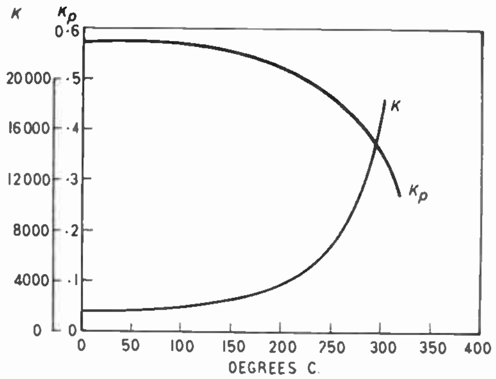



**Fig. 1.** Temperature dependence of radial coupling factor and dielectric constant of LZ-4a piezoelectric ceramic.

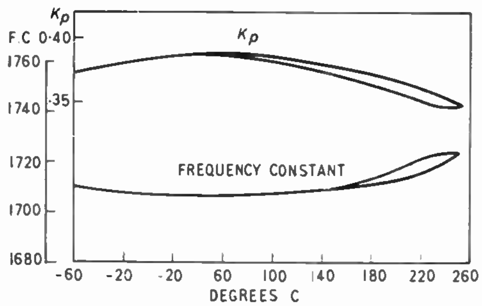



**Fig. 2.** Temperature dependence of frequency constant and radial coupling factor of LZ-6 piezoelectric ceramic (after stabilization).

More recent work has revealed two further classes of piezoelectric ceramics; those based on niobate compositions[2] and the solid solution series lead zirconate-lead titanate.[3] The former have high temperature properties but are somewhat inefficient with relatively high dielectric losses. The latter group represent a major advance as they are capable of a wide variation in characteristics by the use of various additives. They are characterized by a high Curie point, very good conversion efficiencies, low dielectric losses, and can be operated over a wide range of temperature. Figure 1 shows the temperature dependence for the radial coupling factor ($k_p$) and dielectric constant of the Brush type LZ-4a. This grade finds its major use in the manufacture of high power transducers where the high mechanical $Q$ can be fully exploited. The temperature dependence of LZ-6 is shown in Fig. 2, and it will be seen that the frequency constant shows little variation over the range 0° C to 100° C. This material is being used for electrical filters where high stability is required.

All mechanical structures possess resonant modes of vibration corresponding to their physical properties and structural dimensions. At these resonances the strain produced by the applied vibrating force rises to a maximum due to the standing wave distributions. When a piezoelectric resonator is energized by an applied voltage of a frequency corresponding to the natural resonant frequency, the maximum mechanical deformation will be produced.

Combining this principle with a piezoelectric system it will be obvious that a system can be devised consisting of two piezoelectric elements physically coupled together in a form where strain can be introduced into one element by electrically energizing the other. The stress is converted back to an electrical output by the direct piezoelectric effect. The input and output characteristics can be varied as required by modification of the physical and piezoelectric factors inherent in the elements.

While it is possible to use single crystal elements manufactured from materials such as quartz and Rochelle salt the limitations in the direction of activity necessitate a complex assembly of separate elements. The use of piezoelectric ceramics enables composite resonators to be built from an homogeneous piece of material by preferentially polarizing separate sections of the structure. Similarly complex shapes can be initially formed and then polarized in any preferred direction. The ceramic transformer and the device known as a Transfilter use this method of construction and several systems are possible.

## 3. Piezoelectric Transformers

Three practical systems have been studied and each possess certain advantages and disadvantages. Apart from the fundamental principles of resonant operation and double piezoelectric conversion they are somewhat different in operation.

### 3.1. *Ring Type*

The construction of this system is shown in Fig. 3. It consists of two bars of piezoelectric ceramic cemented end to end, electrodes being provided at each end of the composite bar with a common electrode at the joint face. An input connection is made to one end and the central electrode, while the output is obtained from the other end and the common centre. The bar is energized in the fundamental length mode as a half-wave resonator and a free-free mounting obtained by clamping at the nodal point.

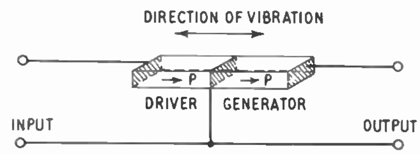



**Fig. 3.** Ring-type piezoelectric transformer.

The open circuit amplification is independent of the geometry, being a function of the mechanical $Q$ and the electromechanical coupling coefficient $k_{33}$ of the material. With available piezoelectric ceramics the open circuit voltage amplification is between 15 and 30. The range of input impedance relative to the output impedance is limited by the symmetrical geometry and the level of input impedance is high compared to other types of ceramic transformers.

A specific version of this system is found in the overtone Transfilter and used as an i.f. transformer.

### 3.2. *Transverse Type*

Figure 4 shows the construction of a transformer using transverse principles. A long bar of ceramic is given electrodes on two major faces over half the length and then polarized in the thickness direction. The remaining half is polarized parallel to the length by applying an end electrode and temporarily commoning the thickness electrodes.

The bar is excited into length resonance by using the thickness electrodes for energizing in the transverse mode. This produces a mechanical strain in the second half of the bar which in turn is piezoelectrically transformed into a voltage appearing across the end electrode. The wave form of the driving voltage is preferably sine wave, but square or saw tooth wave forms will also energize the bar. Single voltage pulses with short rise times can be used to ring the resonator and produce amplified voltage pulses.

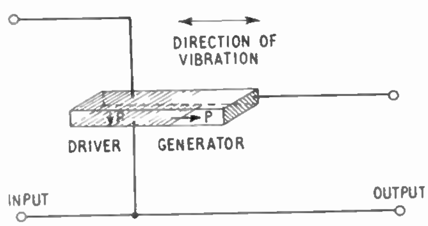



**Fig. 4.** Transverse type piezoelectric transformer.

The transverse transformer has a number of advantages over the ring type. In the latter case the input to output voltage ratio cannot be determined by the geometry of the bar, while with the transverse transformer there is a direct relationship between the dimensions and the voltage ratio. The input impedance of the transverse system is very much lower and since dimensional changes bear a direct relationship to electrical parameters it is possible to design transformer bars to a given specification.

### 3.3. *Hybrid Type*

Recent developments in piezomagnetic materials such as the magnetostrictive ferrites[4] have enabled a composite transformer to be constructed and this is shown in Fig. 5. A suitably shaped and proportioned block of ferrite is bonded to a length polarized bar of piezoelectric ceramic. The necessary d.c. polarization of the magnetostriction element is supplied by an insert permanent magnet. Alternating current energizes the magnetostrictor at its resonant frequency and the strain introduced into the piezoelectric

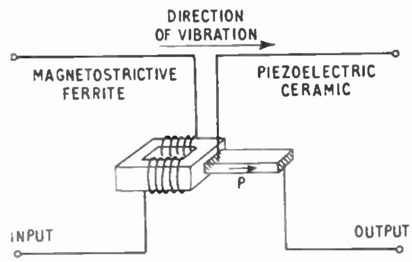



**Fig. 5.** Hybrid type piezoelectric transformer.

element produces a voltage across the electrodes. The ability to provide a very low impedance input is somewhat offset by the lower electromechanical efficiency and strain limitations imposed by the ferrite materials. However the provision of a resistive input rather than a capacitive input may be of use in some applications and the low impedance would facilitate its use with simple transistor circuits.

### 3.4. *Design Considerations*

Of the three possible systems the transverse type appears to offer the greatest advantages and a design study has been made to define operating parameters. A study of piezoelectric ceramic characteristics shows that lead zirconate titanate ceramic exhibits very high values for $k_{33}$ the linear mode efficiency, and $k_{31}$ the transverse mode efficiency. The low dielectric loss and the high temperature Curie point also enables this ceramic to be operated at high power levels without harmful effects. These factors combine to make the ceramic the most suitable for transformer manufacture.

Although it is possible to operate the unit at harmonic modes the complications inherent in high frequency operation generally limit the mode of vibration to the fundamental or second harmonic. The lower frequency limit is somewhat influenced by the nuisance value of the audible noise and also by the large size of the bar at frequencies below 20 kc/s. If the first condition can be tolerated it is possible to mass load the bar at the ends to reduce its resonant frequency. Operation at 5 kc/s could then be used with reasonable lengths of bar. The choice between

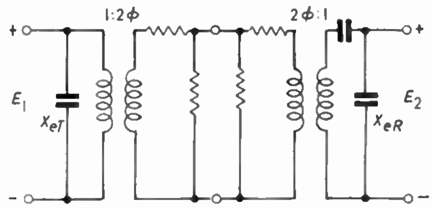



**Fig. 6.** Equivalent circuit of transverse transformer.

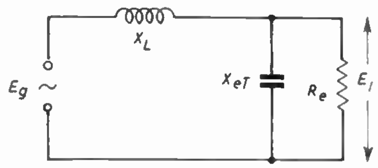



**Fig. 7.** Simplified equivalent circuit of transverse transformer with resonating inductance input.

fundamental or second harmonic operation is decided primarily on the input level requirements for matching the driving source and electrical load.

The theoretical calculations for operating parameters are mainly based on equivalent circuit analysis, the circuit being shown in Fig. 6. It is not proposed to give the derivations or practical formulae as they have been fully covered in published papers.[5, 6] The length, thickness and width of the bar are initially calculated with substituted specification parameters. The section of bar comprising the generator requires modification to its physical dimensions to match the acoustic impedance of the two sections and also to ensure that each section is a $\frac{1}{4}$ wavelength in length. The difference is determined by the linear coupling factor $k_{33}$, the length $L'$ being determined from

$$L' = \frac{L}{(1-k_{33}^2)^{\frac{1}{2}}}$$

and the cross-sectional area $W'T'$ as

$$W'T' = WT\,(1-k_{33}^2)^{\frac{1}{2}}$$

The input represents both a resistive and reactive impedance. In practice it is desirable to neutralize the reactive part and this can be done by resonating it with an inductor (Fig. 7). This usually results in an increased voltage amplification. The required inductance is calculated from the resistive and reactive components as

$$L_e = \frac{X_L}{\omega} = \frac{1}{\omega} \cdot \frac{X_{eT}}{1+\left(\frac{X_{eT}}{R_e}\right)^2}$$

where

$$X_{eT} = \frac{1}{\omega\varepsilon_{33}^T(1-k_{31}^2)} \cdot \frac{T}{WL}$$

and

$$R_e = \frac{\pi}{C^E Q_m Y_3^E\, d_{31}^2} \cdot \frac{T}{W}$$

It will be realized that due to the high output impedance the piezoelectric transformer is essentially a voltage device and is capable of only limited current. The ratio of input to output voltage can be designed to be between unity and about 500 : 1 in the case of the transverse systems and considerably higher with hybrid transformers.

Typical performance curves for a parallel bar 40 kc/s transverse transformer are shown in Figs. 8 and

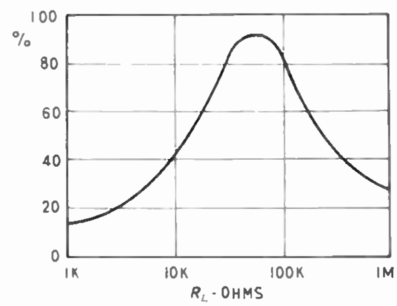



**Fig. 8.** Variation of efficiency with output load for 40 kc/s parallel element transverse transformer.

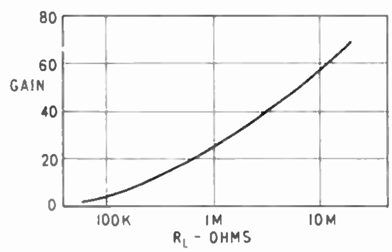



**Fig. 9.** Variation of gain with output load for 40 kc/s parallel element transformer.

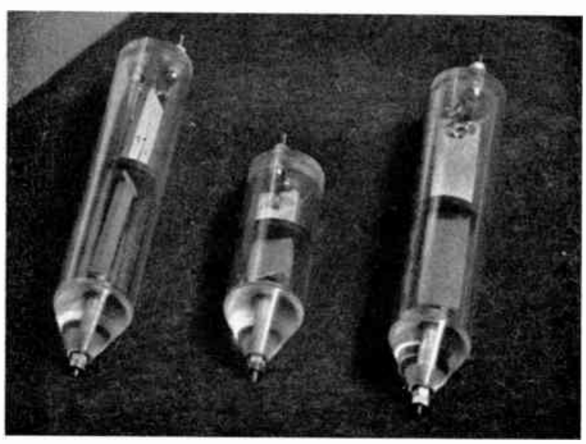


**Fig. 10.** Experimental piezoelectric transformers. (Left to right) 20 kc/s stepped bar, 40 kc/s parallel bar, 20 kc/s parallel bar.

9. It will be noted that the efficiency can reach a very high value with suitably matched output loads. The gain shows a high dependence on the current drawn from the output section, but this is largely due to the shape factor used in this sample. A considerable improvement can be made by correctly proportioning the bar. The transformers are illustrated in Fig. 10.

Suggested applications include cathode-ray tube high voltage supplies, Geiger tube voltage sources, discharge tube power supplies and high voltage pulse generation. The main advantages in the use of these devices are the absence of a magnetic field, elimination of insulation problems, light weight and simple high-frequency operation.

## 4. Transfilters

Using a temperature and time stable piezoelectric ceramic possessing a reasonably high mechanical $Q$ the ring-type transformer can be employed as a three-terminal filter element. A disc operating in the radial mode is used as this mode is relatively free from spurious responses. The overtone frequencies are also not harmonically related. Other configurations such as thickness-expander or thickness-shear elements are susceptible to spurious resonances and their suppression is often difficult.

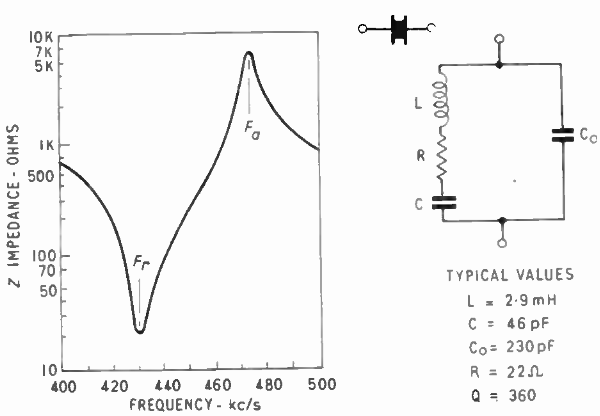



**Fig. 11.** Impedance variation and equivalent circuit of fundamental resonator.

Economic production of filter elements requires a piezoelectric material with high stability characteristics combined with simplicity of manufacture. Although some single crystal materials can satisfy the first condition it is necessary to use a polycrystalline ceramic to enable easy mass production. The development of the previously mentioned lead zirconate titanate ceramic has solved these problems and LZ-6 gives an order of stability unsurpassed by any other piezoelectric material except quartz. Over a temperature range from $-40°$ C to $+85°$ C the frequency constant varies by only $\pm 0{\cdot}1\%$. The ageing properties are also very good, the frequency constant varying by less than $0{\cdot}2\%$ in 10 years.

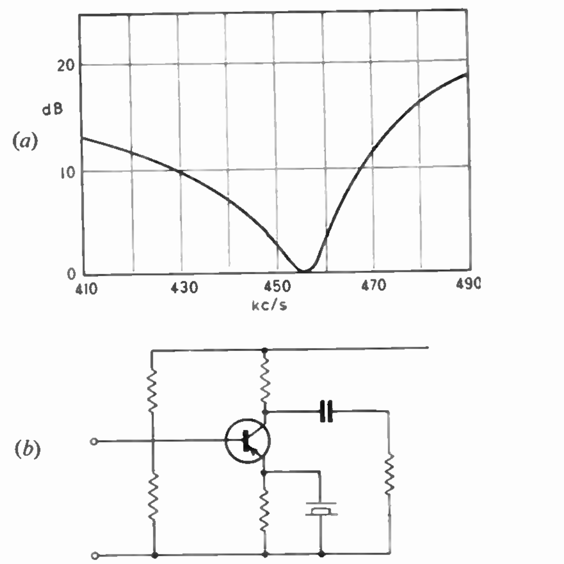



**Fig. 12.** Emitter bypass circuit and response curve using a fundamental Transfilter

### 4.1. *Fundamental Resonator*

A fully silvered disc can be used in the fundamental frequency as a simple filter. Figure 11 shows the impedance variation of such a resonator near its resonant and anti-resonant frequencies and it should be noted that the impedance change from $F_a$ to $F_r$ is greater than 300 to 1. The equivalent circuit is also shown with typical values for a disc 0·030 in. thick and resonant at 430 kc/s.

The value of each element of the equivalent circuit can be calculated from the following relationship:

$$\gamma = \frac{F_a^2 - F_r^2}{F_r^2} = \frac{C}{C_0} \text{ capacitance ratio.}$$

$C_M = C + C_0$ in $\mu F$ where $C_m$ is measured capacitance at 1 kc/s

$$L = \frac{1}{\omega^2 C} \text{ in mH}$$

$$Q = \frac{\omega L}{R} = \frac{1}{\omega CR} \text{ where } R \text{ is measured mechanical resistance at } F_r.$$

$$F_r = \text{Series resonant frequency } \omega L = \frac{1}{\omega C}$$

$F_a$ = Parallel resonant frequency

$$\omega L - \frac{1}{\omega C} = \frac{1}{\omega C_0}$$

With these values known it is possible to design band-pass amplifiers using ceramic resonators which take full advantage of their improved selectivity characteristics. The fundamental radial resonator is free from spurious response in the vicinity of its fundamental resonance, the nearest overtone response being 2·6 times the fundamental.

The impedance of ceramic resonators is determined by the area of the electrodes, the thickness of the disc and the dielectric constant of the material. The maximum electrode area is inversely proportional to the resonant frequency and this places a limit on the minimum impedance. This impedance can be increased by making the electrode area less than the total area of the disc. One quarter of the total area is the limit as below this value the disc is insufficiently excited.

The thickness provides a second effective parameter for impedance control. The minimum thickness is limited to about 0·015 in. due to fabrication problems, while the maximum thickness is governed by the diameter of the disc. The minimum ratio of diameter to thickness is 4 to 1 as smaller ratios will introduce interference due to other adjacent modes of vibration.

The dielectric constant $K$ is about 1000 for LZ-6. This enables a highly selective tuned circuit to be built. A typical application circuit is shown in Fig. 12(a), the Transfilter being used as an emitter by-pass. The response of the circuit in terms of attenuation is given in Fig. 12(b).

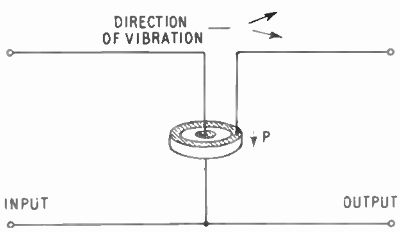



Fig. 13. Arrangment of electrodes on overtone Transfilter.

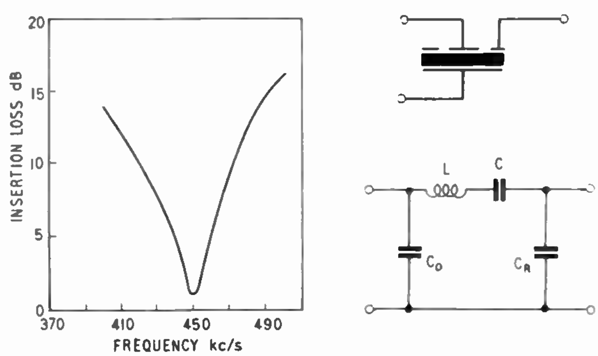



Fig. 14. Insertion loss and equivalent circuit of overtone Transfilter.

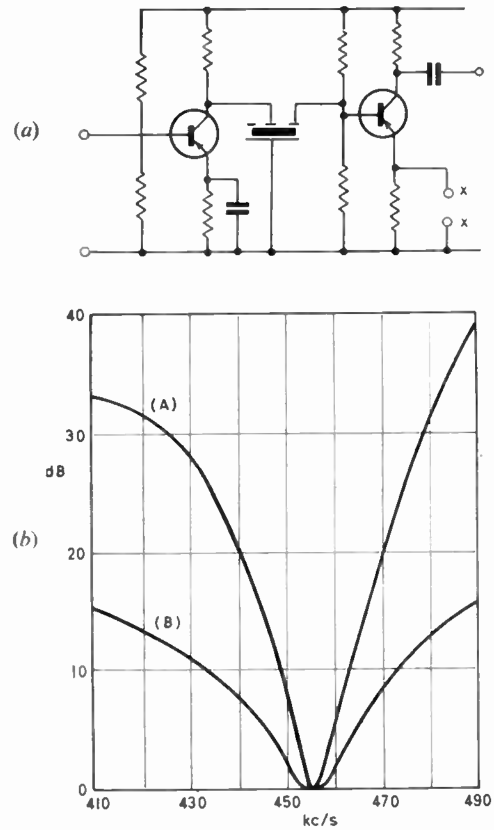



Fig. 15. I.f. stage and passband characteristics using Transfilters.

### 4.2. *Overtone Transfilter*

Using a different electrode configuration a disc can be made to operate efficiently at its first overtone.[7] Operating in this mode the resonator becomes a three-terminal network with one terminal in the form of a centre dot, one in the form of a ring round the dot, and a common back electrode (Fig. 13).

The insertion loss versus frequency response of this type of element is shown in Fig. 14. The equivalent circuit resembles a $\pi$-type impedance transforming network. A difference is found in the capacitor $C$, in series with the inductance $L$. A maximum impedance transformation ratio of 10 to 1 is obtained between the ring and dot electrodes, being governed by the minimum size of the dot for reasonable excitation.

The input and output impedances of a ring and dot disc are directly proportional to the area of the electrodes and inversely proportional to the thickness. The frequency of operation is determined by the diameter of the disc and it will be seen that a different range of impedance will exist for each chosen frequency. For common i.f. values this has a maximum of 5000 ohms and a minimum of 200 ohms.

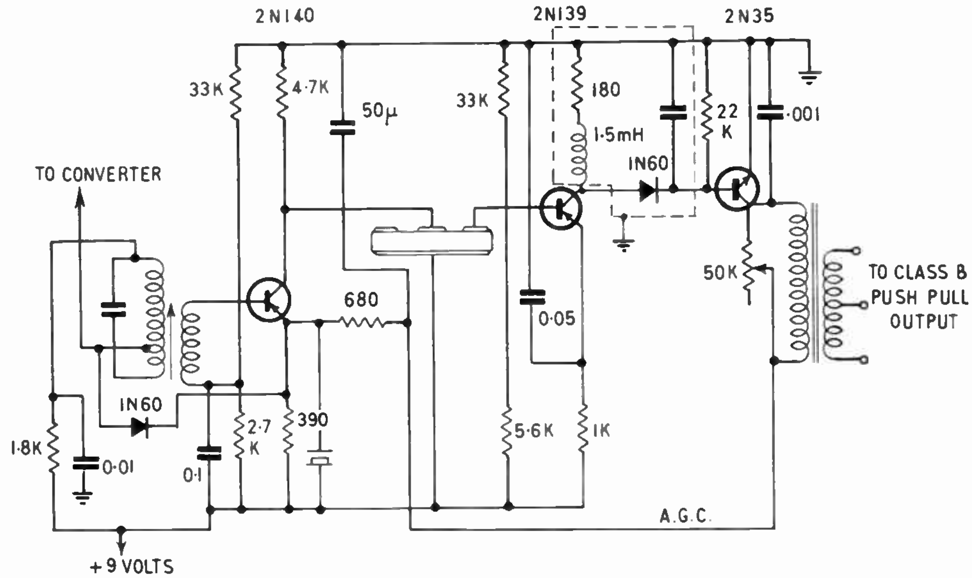



Fig. 16 (a). I.f. amplifier with partial replacement by Transfilters.

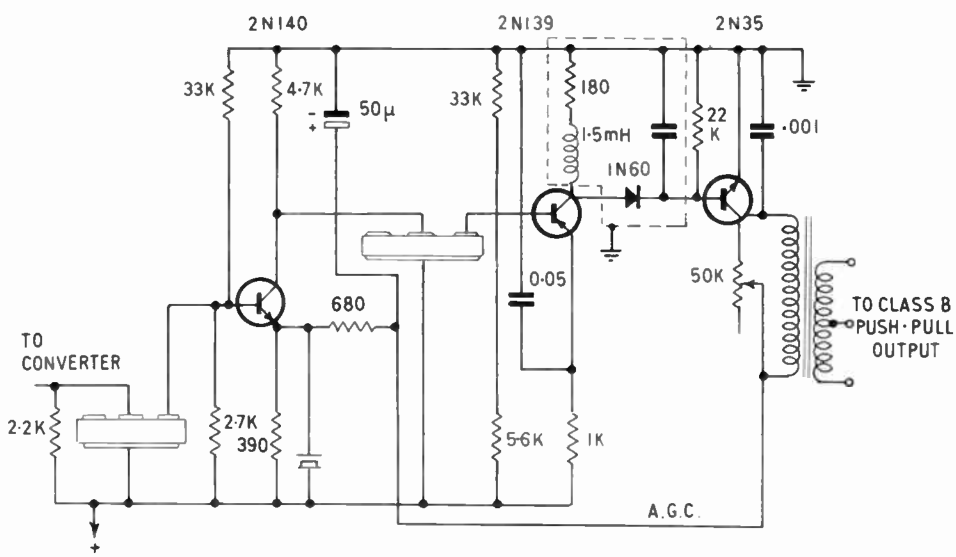



Fig. 16 (b). I.f. amplifier with Transfilters throughout.

The input and output impedances of the overtone resonator are equal to the capacitive reactances of the dot and ring when evaluated at $F_0$, the centre frequency of the pass-band. Therefore, knowing the required input and output values the capacitance of the ring and dot can be specified. The centre of the pass-band is determined by the series resonant frequency of the dot electrode open circuited. By controlling the series resonance of the dot electrode during manufacture the centre frequency $F_0$ can be held to close tolerances.

Bandwidth is related directly to the frequency difference $\Delta F$ between the resonant and anti-resonant frequencies and this is a function of the electromechanical coupling of the disc. Therefore by knowing $\Delta F$, the cut-off frequencies $F_1$ and $F_2$ of the pass-band can be easily determined from the following expressions.

$$F_1 = F_0 - \Delta F$$
$$F_2 = F_0 + \Delta F$$

The high electromechanical coupling of LZ-6 results in the ring and dot disc being a moderately wide-band tuned circuit with bandwidths of 3·5% to 5·5% of the centre frequency. With this in mind the element should be used as an impedance transformer rather than a filter. An i.f. stage circuit is given in Fig. 15(a) and the passband characteristics are shown in Fig. 15(b) with either a capacitor or a fundamental Transfilter in the circuit.

Transistor i.f. amplifier circuits using transfilters are shown in Figs. 16 (a) and (b). The first has a conventional transformer stage from the converter, while the second employs a ring and dot element for this input stage.[8]

These devices offer many advantages to the circuit designer. They are extremely rugged and their small size permits miniaturization with no loss of selectivity. There is no initial alignment of circuits or periodic adjustment and they have been physically designed for use in automatic assembly procedures involving printed circuits.

## 5. Conclusions

The piezoelectric devices which have been described give an idea of the scope of these new solid-state techniques. In line with current semi-conductor trends they offer the electronic engineer a range of components possessing similar characteristics of small size, robust construction and the minimum of maintenance. It can be considered that they are a forerunner of a number of future devices employing similar principles.

## 6. Acknowledgments

The author fully acknowledges the assistance of the research staff of Brush Crystal Co. Ltd., particularly Mr. R. F. J. Orwell, in the investigation of the piezoelectric transformer. The Transfilter was developed by his colleagues in the Clevite Corporation, U.S.A.

Thanks are also due to the Directors of Brush Crystal Co. Ltd. for permission to publish this paper.

## 7. References

1. R. B. Grey, U.S. Patent No. 2,488,560.
2. G. Goodman, "Ferroelectric properties of lead metaniobate", *J. Amer. Ceram. Soc.*, **36**, pp. 368–72, November 1953.
3. A. E. Crawford, "Lead zirconate piezoelectric ceramics", *Brit. Commun. and Electronics*, **6**, pp. 516–9, July 1959.
   See also: B. Jaffe, R. S. Roth, and S. Marzullo, "Properties of piezoelectric ceramics in the solid solution series lead titanate–lead zirconate–lead oxide: tin oxide and lead titanate–lead hafnate", *J. Res. Bur. Stand.*, **55**, pp. 239–54, November 1955.
4. C. M. Van der Burgt, "Piezomagnetic ferrites", *Electronic Technology*, **37**, pp. 330–41, September 1960.
5. A. E. Crawford, "Piezoelectric voltage transformers", *Wireless World*, **66**, pp. 510–4, October 1960.
6. C. A. Rosen, "Solid State Magnetic and Dielectric Devices", Chapter 5. (Wiley, New York, 1959.)
7. A. Lungo and K. W. Henderson, "Application of piezoelectric resonators to modern band-pass amplifiers". I.R.E. National Convention Record (1958), p. 235.
8. Data provided by Clevite Corp., Cleveland, Ohio.

*Manuscript first received by the Institution on 13th October 1960 and in final form on 14th February 1961 (Paper No. 629).*

# DISCUSSION

**Dr. R. C. V. Macario:** I should like to ask Mr. Crawford a question regarding the suppression of overtones when using radial mode discs in i.f. amplifiers. Could he please quote a figure for the suppression in decibels when:

(i) using discs in conjunction with a conventional i.f. transformer.

(ii) using overtone discs in conjunction with fundamental resonance discs.

Also, what figure would he suggest for the tolerable variation in centre frequency, that is $f_R$, in kc/s, when Transfilters are to be employed in broadcast receiver i.f. amplifiers.

**Mr. A. E. Crawford** (*in reply*): No figures of suppression are available for the use of fundamental disc resonators with conventional i.f. transformers, assuming that this refers to the use of the disc as a cathode or emitter bypass device. We would expect this to be a minor effect in any case.

The use of the overtone disc in conjunction with a fundamental resonance disc should give a maximum overtone response — 20 dB below that of the pass band. The centre frequency of the overtone disc can be kept within $\pm$ 2 kc/s of the nominal figure. Assuming the use of a 7% bandwidth figure for $\pm$ 6 dB this should be adequate for broadcast receiver i.f. strips without causing excessive side band clipping.

U.D.C. 621.373.421.112 : 621.374.32

# High Frequency Oscillator Stabilization by Pulse Counting Techniques

By

R. P. THATTE, M.Sc., Ph.D.†

*Presented at a Symposium on Stable Frequency Generation held in London on 25th May 1960*

**Summary:** A crystal oscillator source at 5 Mc/s provides reference frequencies at 1 Mc/s and 1 kc/s, the latter providing a pulse chain output. Two LC oscillators are controlled by a ganged two decade switch. One oscillator is pulse locked to the 1 Mc/s reference, producing frequencies at 1 Mc/s intervals from 3–31 Mc/s. The second oscillator, on the same switch position, can be continuously tuned over 1 Mc/s range, the frequency at the high end of the range being equal to that of the pulse-locked oscillator. These two frequencies are mixed, the output appearing in a band-pass filter of range 1 kc/s–1 Mc/s. This output frequency is fed to a Trochotron divider which can be set to divide by an integer between 1 and 1000 by operating a set of three decade switches.

The output will be 1 kc/s if the oscillator is on frequency. The 1 kc/s output pulse train from the divider is compared with that from the standard in a sine wave phase comparator circuit and the error voltage used to lock the variable oscillator via a reactance valve.

## 1. Introduction

The design of an oscillator drive for transmitters and receivers in the h.f. band of 2–30 Mc/s is governed largely by the requirement of stability and the number of channels. With the introduction of single-sideband systems, and particularly the suppressed carrier s.s.b. transmission, the stability requirements demand a frequency tolerance of 5 parts in $10^7$ or better. At present, various systems are being developed where such a degree of stability is achieved by referring the oscillator control to a single frequency source of high stability.

These methods may be roughly divided into two categories. In the first one, a single oscillator of the desired stability generates a frequency which is usually a whole number of megacycles. This frequency is divided to generate a frequency of 1 Mc/s and is further divided by decades to generate frequencies of 100 kc/s, 10 kc/s, and 1 kc/s, assuming a wanted channel separation of 1 kc/s. All these frequencies have the same percentage stability as the master crystal oscillator. By choosing the correct harmonics (up to the 9th) of these four decade reference frequencies and combining them in a mixer, the wanted frequency can be obtained if the unwanted cross-modulation-frequency products due to this mixing be reduced to a very low level by adequate filtering.

The second method consists of controlling the frequency of an L-C oscillator by a servo system which is fed by an error voltage derived from phase-comparison of the wanted frequency with the same frequency generated by synthesis of frequencies derived from the master crystal oscillator as in the first method. This phase-comparison can be made either once or in successive stages in which the frequency error is successively reduced. The advantage of this method is the freedom from unwanted frequencies which are inherently present in the first method even though reduced to an acceptable level.

The aim of this paper is to describe a system based on the second method of a servo-corrected oscillator but using pulse-counting technique.

## 2. The Proposed Scheme

Suppose it is desired to generate frequencies in the 2–30 Mc/s range at 1 kc/s intervals, with a stability of 5 parts in $10^7$. Referring to the block diagram of Fig. 1, (1) is the crystal oscillator of the desired stability, shown here as a 5 Mc/s frequency source. A regenerative divider (2) divides this frequency to a frequency of 1 Mc/s, which is further divided by 1000 in a three decade divider (3), to give an output frequency of 1 kc/s. The L-C tuned oscillator (4) is locked to the standard crystal frequency at the megacycle points between 3 and 31 Mc/s, and is operated by two decade switches calibrated in tens of megacycles and unit megacycles. A second L-C tuned

† Marconi's Wireless Telegraph Company Limited, Chelmsford.

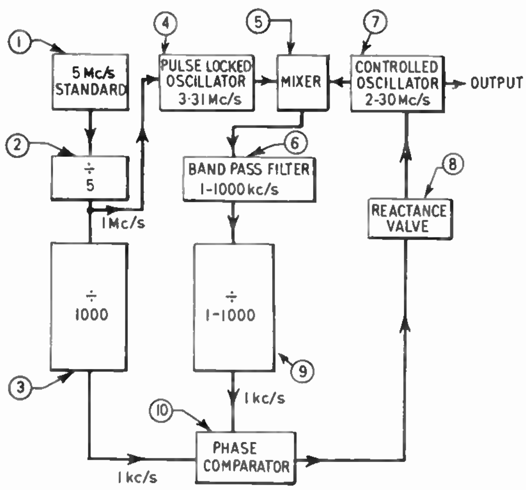



Fig. 1. Block diagram of the pulse counting frequency generator.

oscillator (7) furnishes the desired frequency of the required stability, and is controlled by the reactance valve (8) to cover the range of 2 to 30 Mc/s; its switched ranges are ganged to the two decade switches of oscillator (4). Oscillator (7) can be continuously tuned in each switch position over a 1 Mc/s range in such a way that its highest frequency is equal to the locked megacycle frequency of the oscillator (4). These two oscillator frequencies are mixed in a mixer (5) and the difference frequency appears in a band-pass filter (6), covering the range of 1 kc/s to 1 Mc/s. This frequency is passed to a pulse-counting type of divider (9), where the division ratio can be set to any whole number between 1 and 1000 by a set of 3 decade switches.

Suppose the wanted frequency from the oscillator (7) is 17·253 Mc/s. Setting the two decade dials of the ganged switches of oscillators (4) and (7) to number 17, sets the oscillator (4) to 18 Mc/s and the range of the oscillator (7) from 17 to 18 Mc/s. Suppose this oscillator is now tuned to generate the wanted frequency of 17·253 Mc/s. The difference frequency appearing in the band-pass filter output will be $18-17{\cdot}253 = 0{\cdot}747$ Mc/s. The variable ratio divider decade switches are so wired that their combined dial readings correspond to a division of $1000 - n$, where $n$ is the dial reading. Thus, when the three dials indicate a number of 253, the division ratio has been set to $1000 - 253 = 747$. Since the band-pass filter is feeding to this divider an input frequency of 0·747 Mc/s = 747 kc/s, the divider output will be 747 kc/s ÷ 747 = 1 kc/s.

Thus setting the 5 decade switch dials to the desired frequency (17 253 in above example), when the oscillator (7) is tuned to this frequency, the output of divider (9) will always be 1 kc/s. The phase of this 1 kc/s output from divider (9) is compared in a phase comparator circuit (10) with the 1 kc/s standard frequency derived from the master crystal, through the fixed ratio dividers (2) and (3). The error voltage derived from (10) is fed to the reactance valve (8) which then locks the frequency of oscillator (7) to the standard frequency crystal source of the desired stability.

## 3. Circuit Details

### 3.1. *Crystal Reference*

A crystal oscillator using crystal temperature control by a change-of-state oven as described by Fewings,† can give a day to day stability of 1 part in $10^7$ and a long term stability of 5 parts in $10^7$ or better for the 5 Mc/s reference standard. The scheme outlined would then furnish any frequency in the 2–30 Mc/s band at 1 kc/s spacing having the same degree of stability.

### 3.2. *The Regenerative Divider*

The regenerative divider of ratio 5 [(2) in Fig. 1] reduces the 5 Mc/s frequency from the standard oscillator to a frequency of 1 Mc/s. A regenerative divider has been chosen since it has the virtue of not producing any output in case of the failure of the input voltage. Figure 2, the circuit diagram of this divider, is self-explanatory.

### 3.3. *The Pulse-Locked Oscillator*

The block (4) in Fig. 1 is an oscillator which furnishes frequencies at exact megacycle intervals in the range of 3–31 Mc/s, locked to the standard reference crystal frequency.

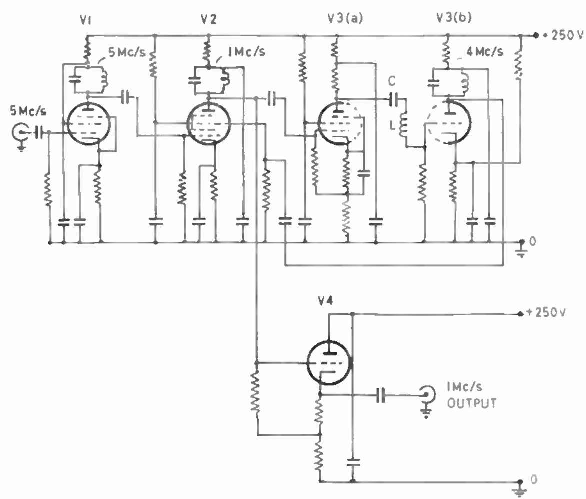



Fig. 2. Regenerative divider.

† D. J. Fewings, "The change-of-state crystal oven", *J.Brit. I.R.E.*, **21**, pp. 137–42, February 1961.

This method of generating spot frequencies at megacycle intervals from the crystal standard is preferred to the straight method of generating harmonics from the 1 Mc/s crystal reference frequency obtained from the divider described in Section 3.2, because of the ease with which a large output voltage, even at 31 times the 1 Mc/s reference, can be obtained. The circuit has the disadvantage that if the locking action fails, the output from this oscillator will not have the stability of the crystal reference. However, this oscillator is locked to the crystal reference frequency by watching a current meter, reading the reactance valve current. A relay operated by this current can inhibit the oscillator output, in case the locking circuit fails. This method of locking an L-C oscillator to a sub-harmonically related crystal frequency, has been described by Bell.†

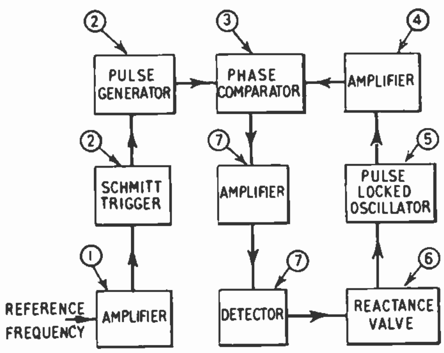



Fig. 3. Pulse locked oscillator schematic.

Figure 3 is a block schematic of this pulse locked oscillator. Amplifier (1) amplifies the 1 Mc/s reference voltage from the regenerative divider. A Schmitt trigger circuit of fast rise time (2) generates a sharp negative going square pulse at 1 Mc/s repetition rate, which discharges through a small inductance on a 5 mm square-loop ferrite core. It causes core switching, and negative- and positive-going pulses of a few volts amplitude and a duration of 15 to 16 millimicroseconds at half amplitude are generated. The L-C tuned oscillator (5) generates spot frequencies spaced one megacycle apart in the range of 3–31 Mc/s. This oscillator output, amplified by stage (4) is fed together with 15 mμs pulses at 1 Mc/s repetition rate from (2) to a phase comparator (3). The output of this phase comparator is amplified and rectified in (7), giving an error voltage proportional to the phase difference between the 1 Mc/s pulse and half of the sine wave of the oscillator frequency. This voltage controls a reactance valve (6) which corrects the frequency of the oscillator (5) until it is exactly phase-locked to the 1 Mc/s reference frequency.

Figure 4 is the complete circuit diagram of this pulse-locked oscillator. Valve V1 amplifies the 1 Mc/s reference frequency. Valves V2 and V3 form a short rise-time Schmitt trigger circuit. The negative-going square pulse from V3 anode discharges through the small inductance L1 (wound on a 5 mm core of square-loop hysteresis curve) generating positive and negative going pulses of about 15–16 mμs duration at 1 Mc/s repetition rate. V8 is the Hartley oscillator, and V9 the reactance valve in parallel with the oscillator tuned circuit. L2–C2 form the resonant circuit of the oscillator at one of the megacycle frequencies. Thirty such L-C combinations are switched in by two gangs of the switch having two 10-position decades. The trimming capacitor C3 is of a small value and at the highest frequency can change the oscillator frequency by less than half a megacycle on either side of its midcapacitance position. The switch banks bringing the desired L-C combinations in the oscillator circuits have shorting contacts on the reverse so that the unused coils are short circuited to avoid dead-spot effects due to stray resonant circuits. A further gang of this switch introduces four capacitances of 47, 22, 10 and 5 pF in the phase-shift network of the reactance valve grid circuit. These four capacitances are switched in only at four intervals of the range of the oscillator, i.e., 47 pF up to 3 Mc/s, 22 pF from 3 to 7 Mc/s, 10 pF from 8 to 15 Mc/s and 5 pF from 16 to 31 Mc/s. As mentioned before, the switch dials are so calibrated that the dial readings correspond to the pulse-locked oscillator frequency 1 Mc/s higher than the dial reading. The reason for this will be explained later. The oscillator output, taken through a 4·7 pF capacitor, feeds an isolating amplifier V5, having a very low anode load, which feeds its output voltage to the phase comparator valve V4. The low coupling capacitance also ensures a constancy of the oscillator input voltage to V4 over the entire range.

The phase comparator valve V4, is a 6AS6 dual control grid valve having equal grid base characteristics. Either grid can cut off the anode current by having large negative voltage applied to it. But, if it is kept in a cut-off condition, applying a positive voltage to one grid does not make it conduct; both grids must have positive voltages applied to them to make the valve conduct, the anode current being proportional to the positive voltage difference between the two input voltages. Potentiometer RV1 is preset so that the anode current is cut-off. The 15 mμs positive pulse is applied to one grid and the sine wave voltage from the oscillator is applied to the second grid. The anode current is then proportional to the phase difference between the narrow pulse and the positive half-cycle of the oscillator sine wave. A

† N. R. Bell, "Locked L-C oscillators", *Marconi Instrumentation*, 5, p. 66, No. 3, September 1955.

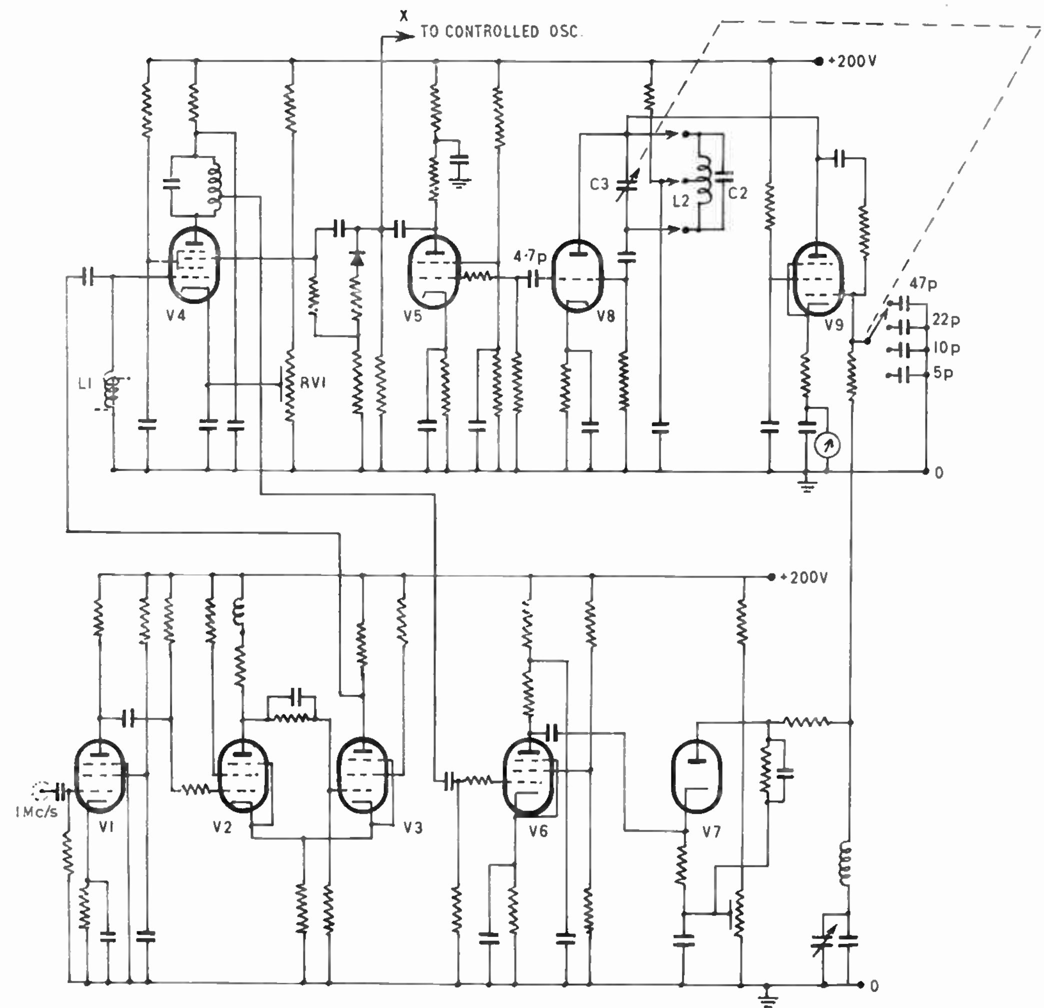



Fig. 4. Pulse locked oscillator.

resonant circuit tuned to 1 Mc/s is in the anode circuit of this valve. The result is that the 1 Mc/s repetition rate current pulse, whose amplitude is proportional to the phase difference, generates a 1 Mc/s sine wave voltage across the tuned circuit. The envelope of this sine wave is amplitude-modulated by a low frequency sine wave which is proportional to the phase difference at successive 1 Mc/s pulses sampled against the oscillator sine wave. This voltage is extracted from a tap on the 1 Mc/s tuned circuit and amplified in valve V6. It is rectified in diode V7 and the negative voltage developed in its anode circuit is proportional to the beat frequency modulating the 1 Mc/s frequency envelope, and thus proportional to the phase difference between the oscillator voltage and the 1 Mc/s reference frequency; this voltage becomes zero when the phase difference is zero. This negative voltage is applied as bias voltage to the reactance valve V9, which has a 0–10 mA meter in its cathode circuit, and the dip of this current meter is an indication of a phase-lock. A series tuned circuit of 1 Mc/s is connected between the reactance valve bias rail and earth to remove the 1 Mc/s frequency component and prevent it from modulating the oscillator frequency through the reactance valve.

### 3.4. *The Controlled Oscillator*

The design of the controlled oscillator and its reactance valve is very similar to the corresponding part of the pulse-locked oscillator. Figure 5 is a circuit diagram showing the controlled oscillator, the reactance valve, the mixer, the bandpass filter and amplifier represented by blocks (7), (8), (5) and (6) of Fig. 1.

Valve V7 is the Hartley oscillator and V6 the reactance valve. The inductance L1 and combination of capacitors C1, in parallel with the variable capacitor C2, which is in series with a padder C3, form the resonant circuit. The variable capacitor C2 can be

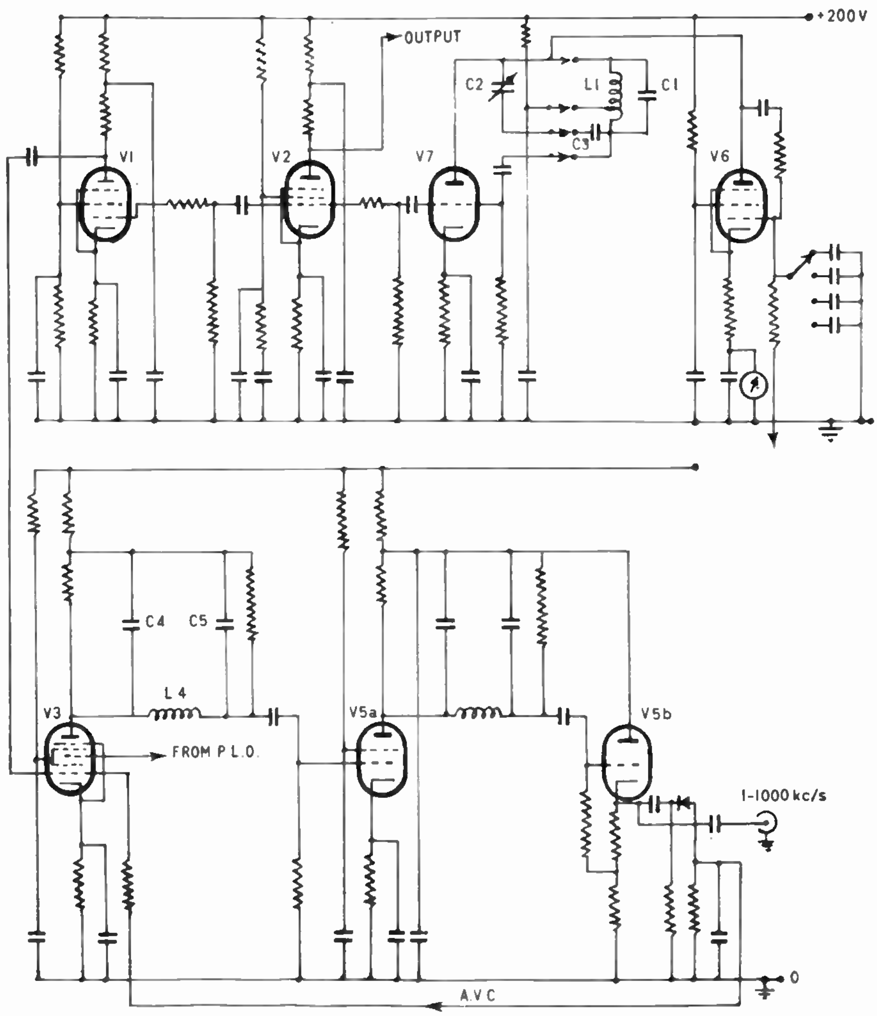



Fig. 5. Controlled oscillator and bandpass filter amplifier.

operated from the front of the panel for manual control or can be controlled by the servo-controlled motor mentioned later. It enables the oscillator frequency to be set over 1 Mc/s range and is common to all range positions of this oscillator over 2–30 Mc/s band. As in the pulse-locked oscillator, L1 with its range setting capacitor C1 and the series padder C3 is switched across C2 by a three wafer switch which is again in two decades and is ganged to the range controlling switch of the pulse-locked oscillator. There are thus 30 combinations of L1, C1, C3 for each megacycle range in the 2–30 Mc/s band. The choice of a suitable value of C3 as a capacitor in series with the main tuning capacitor C2 enables it to cover a range of only 1 Mc/s at all of the 30 ranges in the band. A further bank of switch wafers brings four different capacitors in the phase-shift network of the reactance valve as in the pulse-locked oscillator. The output from this oscillator is fed through a low capacitance to identical amplifier stages V1 and V2. The output of V2 gives the controlled frequency for the external use and is the output end of the system. Output of V1 is fed to a mixer valve V3 (shown by block (5) of Fig. 1), the other input grid of V1 being fed from the output of amplifier V5 in the pulse-locked oscillator of Fig. 4. The ranges of the coils of oscillator V7 of Fig. 5 are arranged to be the same as the dial readings of the controlling switch.

The reason for arranging the ranges of the controlled oscillator and the pulse-locked oscillator in such a way that the dial reading of the first two decade switches indicates directly the lower end of the range of the controlled oscillator when the pulse-locked oscillator frequency is 1 Mc/s higher than the dial reading, can be seen from an examination of Tables 1 and 2.

If we keep the pulse-locked oscillator frequency 1 Mc/s below the dial readings we get the conditions

**Table 1**

| Wanted Frequency Mc/s | 5 Decade Switch Dial Reading | Pulse-locked Oscillator Frequency Mc/s | Controlled Oscillator Mc/s | Beat Frequency Bandpass Filter Output kc/s | Division Ratio of Variable Divider |
|---|---|---|---|---|---|
| 2·000 | 02 000 | 1 | 2·000 | 1000 | 1000 |
| 2·001 | 02 001 | 2 | 2·001 | 0001 | 1 |
| 2·999 | 02 999 | 2 | 2·999 | 0999 | 999 |
| 3·000 | 03 000 | 2 | 3·000 | 1000 | 1000 |
| 3·001 | 03 001 | 3 | 3·001 | 0001 | 1 |
| 3·999 | 03 999 | 3 | 3·999 | 0999 | 999 |
| 4·000 | 04 000 | 3 | 4·000 | 1000 | 1000 |
| 4·001 | 04 001 | 4 | 4·001 | 0001 | 1 |

shown in Table 1. Column 2 indicates the dial readings of the 5 decade switch; the last three figures of these dials which control the division ratio of the variable divider 9, are the same as the division ratio, shown in the sixth column. However, inspection of columns 1, 2 and 3 shows that the No. 2 dial which reads megacycles, requires the pulse-locked oscillator to change its frequency setting depending upon whether the last three dials read 000 or any other number. This complicates the high-frequency range switching of the pulse-locked oscillator, requiring more wafers on the switch.

This difficulty can be overcome if we keep the pulse-locked oscillator frequency 1 Mc/s higher than that of the controlled oscillator, as shown in Table 2.

**Table 2**

| Wanted Frequency Mc/s | 5 Decade Switch Dial Reading | Pulse-locked Oscillator Frequency Mc/s | Controlled Oscillator Frequency | Beat Frequency Bandpass Filter Output kc/s | Division Ratio of Variable Divider |
|---|---|---|---|---|---|
| 2·000 | 02 000 | 3 | 2·000 | 1000 | 1000 |
| 2·001 | 02 001 | 3 | 2·001 | 0999 | 999 |
| 2·999 | 02 999 | 3 | 2·999 | 0001 | 1 |
| 3·000 | 03 000 | 4 | 3·000 | 1000 | 1000 |
| 3·001 | 03 001 | 4 | 3·001 | 0999 | 999 |
| 3·999 | 03 999 | 4 | 3·999 | 0001 | 1 |
| 4·000 | 04 000 | 5 | 4·000 | 1000 | 1000 |

With this scheme, it will be seen from columns 2 and 3 that the pulse-locked oscillator frequency remains the same for each setting of the megacycle dials. However, the last three decade dials, reading the kilocycle part of the wanted frequency, require a division ratio given by last column which is a number complementary to 1000; for example, for a dial reading of 001 in the last three dials, the division ratio is $1000-1 = 999$ and for 999 it is $1000-999 = 1$. Therefore the switch controlling the division ratio of the variable divider is wired to give a ratio which is the complement of 1000 to the dial reading. This is dealt with in Section 6.3, where a further modification is also shown.

### 4. Band-Pass Filter and Amplifier

The output from the pulse-locked oscillator is taken from point X in Fig. 4 and fed to one grid of the mixer valve V3, in Fig. 5. The output from the controlled oscillator is taken from the anode of the amplifier V1 and fed to the second grid of the mixer valve V3. The anode circuit of V3 consists of a low-pass filter consisting of C4, L4, C5, feeding into the grid circuit of the pentode section of amplifier V5a. The anode to earth output capacitance of V3 in parallel with C4 and the wiring capacitance and the input capacitance of V5a in parallel with C5, form the net input and output capacitance of the low-pass filter connected between the matched resistance loads of 4·7K. The filter is designed to give 0·1 dB ripple in passband, the response of the filter being $-3$ dB at 1·098 Mc/s and $-40$ dB at 4·12 Mc/s. A low-pass filter of identical response characteristics connects the output of V5a to the grid of the cathode follower V5b, furnishing the beat frequency output in the range of 1 kc/s–1000 kc/s from oscillators (4) and (7) of Fig. 1 of sufficient amplitude to operate the variable ratio divider (9). A part of the output from this cathode follower is rectified and applied as an a.g.c. bias to mixer valve V3 to achieve a reasonably constant output over the band of 1 kc/s–1000 kc/s.

### 5. The Frequency Dividers

From the above outline it will be seen that the operation of the system depends upon the two dividers, the fixed ratio (1000) divider, and the variable ratio divider (1–1000) of Fig. 1. The design of these dividers will be dealt with in some detail, as these embody a comparatively unfamiliar device, the Trochotron.

#### 5.1. *The Trochotron (Fig. 6)*

In recent years, a number of special tubes have been developed for counting purposes. For high counting rates, hot-cathode vacuum tubes are necessary because of the effect of gas deionization times in cold-cathode tubes. Among such tubes is one originally developed by Professor Alfven at the Royal Institute of Technology in Stockholm, in collaboration with the Swedish Ericsson Co., and called Trochotron or beam-switching tube. It was improved and is now manufactured by Burroughs Corporation in U.S.A., and recently by the British Ericsson Telephones Ltd.

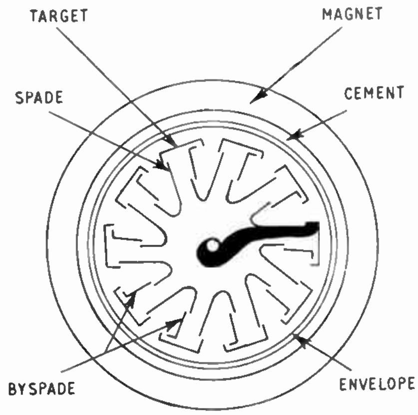



Fig. 6. Sectional view of Trochotron.

and Mullard Ltd., in this country. One such tube, in conjunction with a single-stage, hard-valve binary, can count at rates over 2 Mc/s, replacing four double triodes plus ten diodes or eight to sixteen transistors plus diodes in conventional scale-of-ten circuits. The Trochotron has also a minimum life in excess of 50 000 hours, and can work in high ambient temperatures. A full description of the design and operation of the Trochotron manufactured by Burroughs Corporation has been given by Fan.†

### 5.2. *The Beam Forming, Setting to Zero or Any Position*

As described by Fan, the Trochotron remains in the cut-off condition when the power is first applied. The beam may be formed in any one of the ten "on" positions, by lowering the potential of the respective spade by a d.c. voltage or a high speed pulse. Once the beam is formed on any one spade, it is necessary to clear the tube first to form it on any other spade. This can be conveniently done by lowering the common spade supply voltage to a cut-off value and if the chosen spade has a larger time-constant load, it will remain at this lowered potential longer than the other spades and the beam will now form in this position. This is the method employed in the design of the variable ratio divider. Figure 7 shows the spade characteristics.

## 6. The Variable Ratio Divider

The variable ratio divider will be considered first, though it is more complicated than the fixed ratio divider. For the economy of manufacturing design, a unit around a Trochotron is designed which forms a building block, with which the variable ratio divider is built. The same unit is used in the fixed ratio divider, though a simplified version could be used for this purpose.

† S. P. Fan, "The magnetron beam switching tube: its operation and circuit design criteria", *J.Brit.I.R.E.*, 15, pp. 335–54, 1955.

### 6.1. *A Trochotron Divider Unit*

Figure 8 is a detailed circuit of such a unit using a Trochotron. With appropriate connections it can be used as a divider capable of giving any division ratio between 1 and 10, up to a maximum input pulse rate of a little over 1 Mc/s.

V3 is a Trochotron. A cathode resistor is included to introduce current feedback to linearize the target output current. The spade supply voltage is determined by the potential divider R1, R2 and R4; the target supply is +300 V applied at pin D. The target load resistors R10 in series with R110 etc., are connected to each of the ten targets, their values satisfying the condition $V_T - R_T \times I_T \geq \frac{1}{2}V_s$. The junction points of the two target load resistors are brought to pins P0 to P9 and enable the output pulse to be taken from any desired numbered target. The spade load resistors in the number 1 to 9 spade circuits are identical. Spade 0 load circuit consists of three resistors in series, the total value being equal to the other spade load resistors. The two capacitors C1 and C2 introduce the delay for high and low speed resetting of the beam at the 0 position. Valve V4 is a special neon tube, called a numerical indicator tube, which has a common anode and ten cathodes, each shaped to a numeral from 0 to 9. Each of these numbered cathodes are connected to the appropriately numbered target. When the beam is at rest on a particular target, the potential drop across its load resistor is sufficient to strike the glow on that particular numbered cathode and the position of the beam of the Trochotron is thereby indicated. At high operating speeds all these numbers glow simultaneously but this indicator is included in the design as it helps to check the zero setting in the variable ratio divider at low speeds, and it also indicates that the tube is not in a cut-off condition.

Valve V1 is working in a high-speed Eccles-Jordan type binary circuit, capable of operating up to 1·5

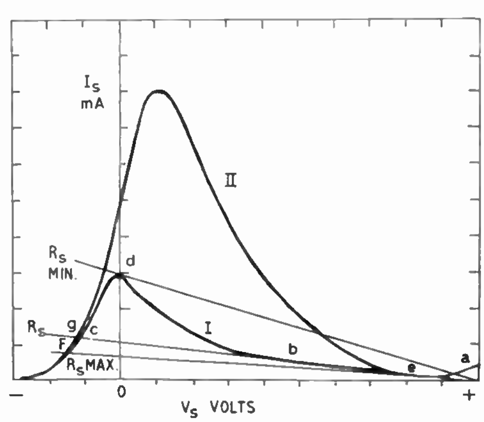



Fig. 7. Trochotron spade characteristic.

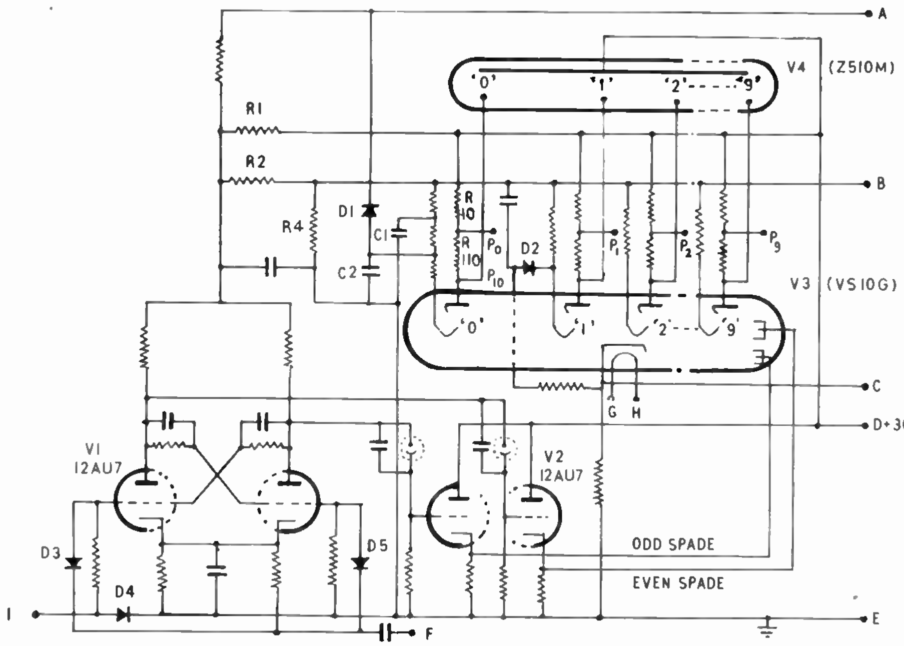



Fig. 8. Trochotron divider unit.

Mc/s on the application of negative input pulses of proper duration to pin F. These pulses trigger the grids of the binary through diodes D3 and D5. The left hand grid of the binary is returned to earth through another diode D4, so that by applying a negative pulse to pin 1, the binary can be reset, corresponding to the 0 ready condition. The two flip-flop outputs are d.c. connected through neons to the grids of the two cathode followers V2, and their cathodes connected directly to the odd and even byspades. The cathode load resistors are chosen to supply the appropriate d.c. bias to the two byspades. The zero target is also directly connected to pin P10 and, for normal cascade connection of a number of Trochotrons, this pin is connected to pin F of the succeeding Trochotron. The pulse output at P10 and also at the ten pins P0 to P9 is of sufficient amplitude to trigger the succeeding Trochotron binary. Application of a negative pulse of sufficient amplitude simultaneously to pins A and B lowers the potential of the common spade supply point B, cuts off the Trochotron and because of the time constant of the R-C circuit in the 0 spade, after the termination of the reset pulse, it recovers to the spade supply potential, at a slower rate than the other spades and the beam is formed on the 0 target.

### 6.2. *Division Ratios less than Ten*

If we examine the waveforms on the different targets as the beam rotates around through ten stable positions on the application of a train of pulses to the flip-flop input, we note the sequence of events shown in Fig. 9. On the line (a) are represented the negative-going input pulses, equally spaced, as applied to the pin F of the Trochotron unit. Lines (b) and (c) represent the push-pull waveforms as derived by the binary from these pulses and as applied to the byspades of the Trochotrons. Lines (d) to (m) show the d.c. waveforms on the ten targets, 0 to 9 respectively. When the beam is initially set on the 0 target, its voltage, as represented by line (d) is at the lowered potential corresponding to the *IR* drop in the target

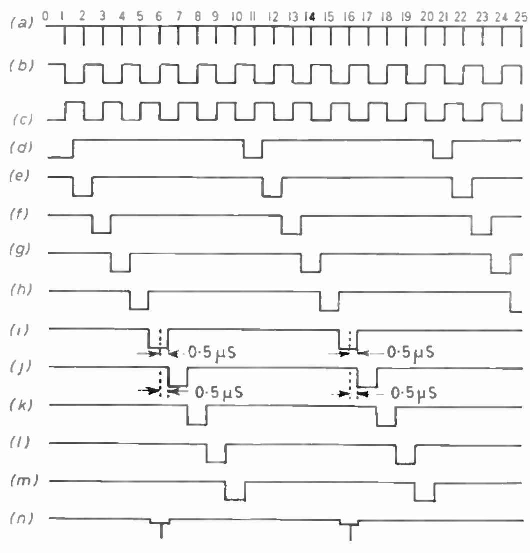



Fig. 9. Trochotron waveforms.

load. On the application of the first pulse, the beam switches on to the target 1, the potential of the 0 target rises to the h.t. supply potential, and remains there till the application of the 10th pulse when it is again lowered, rising again when the beam leaves the 0 target on the application of the 11th pulse. Similarly, the voltage on 1 target, shown by line (e), is at the h.t. potential till the application of the 1st pulse, when it is lowered and raised again after the 2nd pulse. This sequence is repeated at the 11th and the 12th pulse. Similarly, lines (f) to (m) show the sequence of events on targets 2 to 9, on the application of sequential pulses. It will be seen that the potential on any target exhibits similar waveforms whose interval corresponds to the interval between ten consecutive pulses. Therefore, if we take the waveform from any target, the output is at a rate of 1/10th of the input pulse rate, i.e., the Trochotron is acting as a divider of ratio 10. It should also be noted that there is an inherent delay between the application of the input pulse and the time when the beam comes on, or leaves the corresponding target. This delay is always 0·5 μs, whatever be the input pulse rate and whichever be the target chosen. This delay is a characteristic feature of the Trochotron. In sequential counting by division ratio of 10, this delay being repetitive is automatically cancelled, as the interval between successive waveforms is the same as the interval between 10 pulses. This delay is serious when two or more Trochotrons are cascaded and a forced resetting is used to secure a division ratio less than multiples of 10.

Suppose in a single Trochotron unit it is desired to have a ratio other than 10, say 5. Also suppose that the beam is on target 0. The potential, on the 5th target will go down on the application of the 5th pulse to the Trochotron unit input. If the negative-going edge of this waveform is used to generate a negative-going pulse of amplitude sufficient to reduce the spade supply voltage to the cathode potential, and this pulse is applied to points A and B simultaneously, as soon as this pulse appears it will clear the tube and because of the R-C time-constant introduced in the 0 spade load circuit, the beam will reform on the 0 target again. This resetting pulse should be of such duration that the resetting is complete before the 6th pulse arrives at the input, when again the Trochotron will count 5 pulses and reset to zero, repeating this cycle. The Trochotron unit is then acting a divider of ratio 5. Thus by selecting any target from 0 to 9 to generate the reset pulse, any ratio from 10 to 1 can be obtained.

Let us now suppose that a division ratio of 456 is wanted. Starting with the beams on the 0 targets in a three-stage Trochotron divider, if the output from the 1st Trochotron unit be taken from the 6th target, the output will first be obtained on the application of the 6th pulse to the 1st unit input, and will be repeated at the 16th, 26th, 36th input pulse and so on. If these outputs are connected to the input pin F of the second Trochotron, its beam will move to successive target positions at these 6th, 16th, 26th etc., input pulses. If the output is taken from the 5th target of the second Trochotron, an output will be obtained after the 56th, 156th, 256th input pulse to the first unit; if this output triggers the third unit its beam will move 1 target position at these pulses. If the output from the 4th target of this third unit is taken, an output will be obtained from it for the first time after the 456th input pulse. If this output is used to generate the reset pulse to reset the beams on all the three units to the '0'th target, the unit will count 456 input pulses every time. However it cannot be said that the input pulse rate has been divided by a ratio of 456. If the input pulse rate is 1 Mc/s, the output from the third unit will be obtained after 456+3 times the 0·5 μs (inherent delay of the three Trochotrons). That is, the reset is made after 457·5 μs after the '0'th pulse, when one more pulse (457th), will have operated the first Trochotron. The interval between successive resetting pulses will be 457·5 μs and the input rate is being divided by a factor of 457·5 instead of required 456. This means that, referring to Fig. 1, to secure the 1 kc/s output from the variable ratio divider, the beat frequency feeding it will be 457·5 kc/s, locked to the crystal frequency, when the desired frequency was 456 kc/s.

This accumulated delay of 1·5 μs in a three stage divider can be eliminated in the following manner. On examining Fig. 9 again, though the output from the 6th target appears 0·5 μs after the application of the 6th pulse, on the preceding 5th target there is still a waveform present when the 6th pulse has arrived at the input. If the input pulse train be gated by the waveform on the previous target, by a diode AND coincidence gate, the output of this gate will be coincident with the 6th input pulse as shown on line (n) in Fig. 9. The delay in one stage is thus eliminated. This coincident negative pulse is connected to the input of the second Trochotron stage. A coincidence between this 2nd stage input pulse and the output from $5-1 = 4$th target of this stage will give an output pulse, coincident with the 56th input pulse to the first stage. However, the efficiency of a diode gate is low and the loss in amplitude can be serious. Therefore, a triple coincidence is secured between the input pulses, the 5th target output from the first stage and the fourth target output from the 2nd Trochotron. The output gives a pulse coincident with the 56th input pulse which then triggers the third stage. Similarly to avoid the delay in the third stage, a quadruple coincident gate circuit between (1) the input pulses, (2) the 5th target output of 1st stage, (3) the

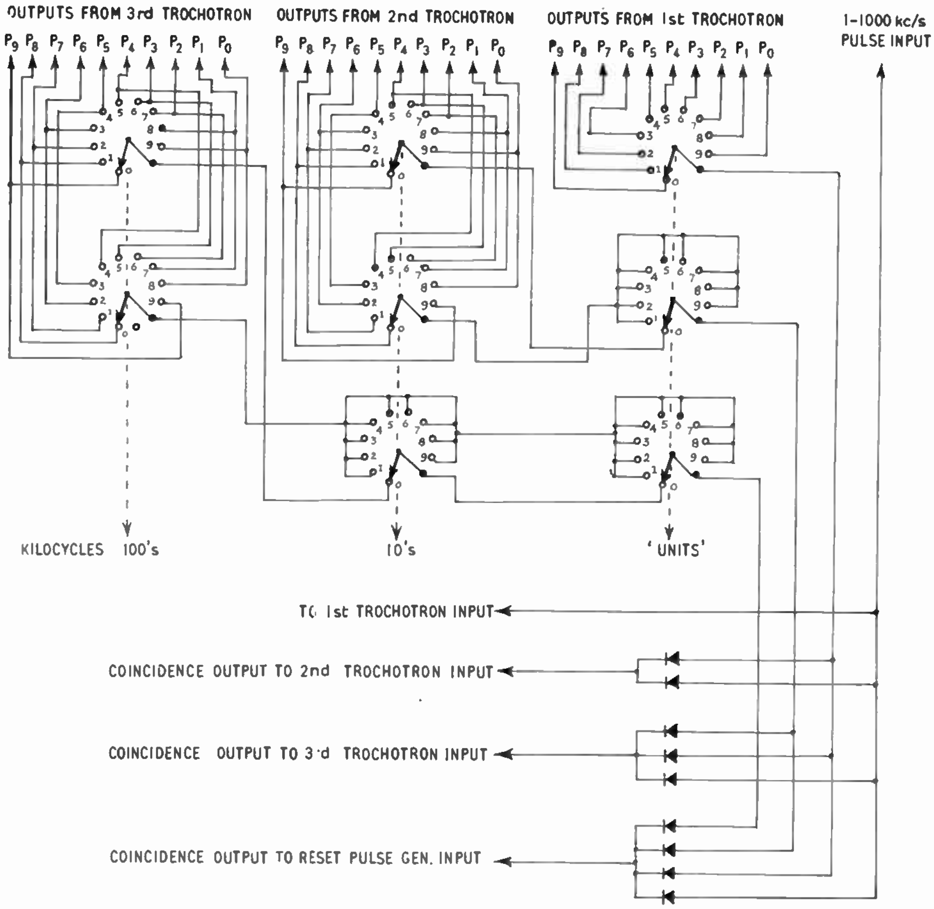



**Fig. 10.** Variable ratio divider switch wiring and coincidence coupling.

4th target output of the 2nd stage and $4-1 = 3$rd target output of the third stage, gives a pulse which is coincident with the 456th input pulse to the three-stage divider. This coincidence pulse is used to generate the narrow reset pulse which can thus reset the beams to 0, 0, 0 in all the three Trochotrons before the 457th pulse arrives at the divider input and thus the successive reset pulses will be exactly 456 μs apart, giving a division factor 456 as desired.

Thus, to secure a division ratio, $pqr$ ($p$, $q$ and $r \leqq 9$), the output is taken from the ($r$–1)th target of the 1st Trochotron, ($q$–1)th target from the second, and the ($p$–1)th target from the 3rd stage, to the three coincidence circuits.

### 6.3. *The Design of the Ratio-selection Switch*

As explained in Section 3.4, the division ratio of this three-stage divider is controlled by a three-decade switch, the switches being wired in such a way that the three dial readings correspond to a division ratio of (1000 – dial reading). This complication adds to the need mentioned above of connecting to a target number one less than each digit in the ratio number. Thus, a dial reading of 544 correspond a division ratio of $1000-544 = 456$ when the targets selected from the 3rd, 2nd and the 1st Trochotrons are $4-1$, $5-1$, $6-1$, i.e. 3rd, 4th and 5th. Figure 10 shows the wiring arrangement of this three decade switch; it also shows the inter-decade coincidence coupling.

### 6.4. *Reset Pulse Generator*

When resetting at high input rates of 1 Mc/s the whole resetting operation has to be performed in a time less than a microsecond if the next pulse is not to be missed. The last coincidence output pulse generates this pulse of large negative amplitude and duration of about 0·2μs. When doing such a fast resetting, if any Trochotron beam is being reset from target No. 1, the reset pulse, which also has a positive swing of nearly the same amplitude and duration, and which applied to the anode of the diode D2, pulses spade No. 1 positive, creating a low impedance path between spade No. 1 and the spade supply voltage for the duration of the backswing. This ensures that during the reset operation, the beam always forms and becomes stable in the zero position. Figure 11 shows the reset pulse generating circuit which produces a negative and positive swing, short duration pulse. The use of a video amplifier valve ensures sufficient power to reset three circuits with a low impedance output. This blocking oscillator circuit can work with maximum repetition rate up to 100 kc/s; that means, if the input frequency to the variable-ratio divider be 1 Mc/s, the minimum ratio obtainable is 1000 ÷ 100 = 10. However, at the lowest ratio of 1, required in our system, the input frequency will be about 1 kc/s only, therefore, the divider can give any ratio between 1 and 1000 necessary in our problem.

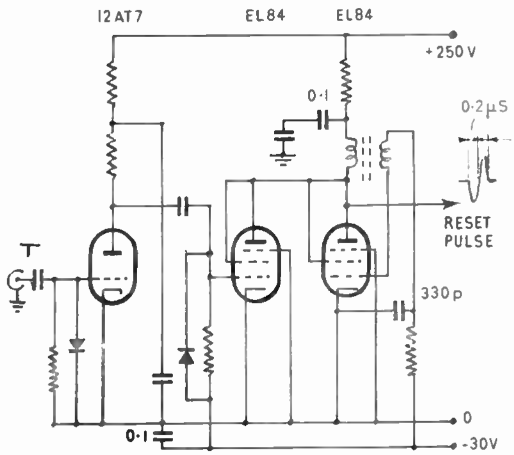



Fig. 11. Reset pulse generator.

The band-pass filter output of 1 kc/s–1000 kc/s is a sine wave output of about 20V peak-to-peak amplitude. This is fed to a Schmitt trigger circuit similar to the one used in the pulse-locked oscillator shown in Fig. 3. The output of the Schmitt trigger is differentiated and amplified, giving a negative pulse of about 50V amplitude and about 0·1 μs duration. This pulse is used as the input pulse to the first Trochotron stage and the interdecade coincidence coupling circuits.

## 7. The Fixed Ratio Divider

The 1 Mc/s frequency from the regenerative divider is to be divided by 1000 to secure the 1 kc/s reference frequency. This is most conveniently obtained by cascading three Trochotron units, similar to the ones used in the variable-ratio divider. As explained in Section 6.2, in a sequential operation in a cascaded divider without resetting, there is no inherent delay; therefore the three Trochotron units are simply cascaded by connecting P10 (the 0 target) to pin F of the following Trochotron, the final Trochotron output being taken from pin P0. To form the beam initially, the points B from these three Trochotron units are momentarily earthed through diodes by a press-to-close switch, another gang of which applies a positive trigger to the reset circuit of the variable ratio divider causing the beams to be formed in position 0 in both the dividers. The input to this fixed ratio divider is obtained from a Schmitt-trigger circuit similar to the one used in the other divider.

## 8. The Final Phase Comparison

Assuming the oscillator (7) in Fig. 1 is tuned to the desired frequency, set up by the five decade dials, the outputs from the variable and the fixed dividers will always be a train of pulses of 1 kc/s rate. A phase-comparison between a pulse from each channel should produce an error voltage, which, when applied to the reactance valve, phase-locks the oscillator (7) to the crystal source. An estimate of the accuracy of the required phase comparison, is necessary. Having set the target of the required stability as ± 5 parts in $10^7$ it means that on the highest range of 30 Mc/s, the error allowed will be ± 15 c/s. Suppose the wanted frequency is 30 Mc/s (30 000 000 c/s) and the oscillator is actually 10 cycles above, i.e. 30 000 010 c/s, the beat frequency as fed to the variable ratio divider (set to division ratio of 1000) will be 999 990 c/s and the output frequency from the divider will be 999·99 c/s. The interval between successive pulses from the variable divider will be $10^6 \div 999{\cdot}99 =$ 1000·01 microseconds. The reference channel gives pulses separated by 1000 μs (1 kc/s repetition rate). Thus a phase displacement of 0·01 μs should produce the minimum significant change in the error output. This minimum discrimination of detectable phase depends upon the division ratio and it is much larger for a division ratio of unity (it is 10 μs for ratio of 1, as can be easily calculated). Therefore, we should aim at this minimum sensitivity of 0·01 μs. It was first thought that a radar type time-demodulator circuit with gate apertures of 0·1 microseconds might be required. However, it was found experimentally that converting the pulse trains at 1 kc/s rate from the two channels into sine wave packets of 1 kc/s frequency and comparing the phases of these two 1 kc/s frequencies in a conventional phase discriminator circuit, enabled a phase-error detection, adequate to lock the final oscillator to the reference crystal oscillator within ± 1 c/s. Obviously, with a phase discriminator of sine wave inputs, the two frequencies have to be very nearly the same before the system control secures a lock. A means is therefore needed to set the controlled oscillator manually, to such a position that the divided frequency is very near to the 1 kc/s reference frequency. A very reliable and un-

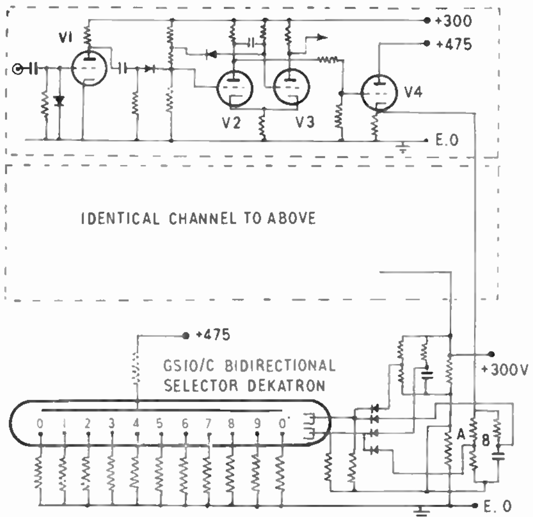



Fig. 12. Bi-directional Dekatron circuit.

ambiguous solution has been found in a bi-directional Dekatron circuit, shown in Fig. 12. Pulses from each of the 1 kc/s channels generate a negative-going pulse, which are fed to two identical circuits shown in Fig. 13 and generate a negative-going pulse of 140V amplitude and 80 μs duration at the output of cathode followers V4. The circuit also gives a positive pulse of about 200V amplitude and of the same duration from the anode of V3. This negative pulse from cathode follower V4 is divided by a potentiometer R1 and R2 at point A and is fed to one guide of a bi-directional Dekatron GS10/C. The pulse is also integrated by R3, C3 at B and is applied to the other guide. Thus the pulses from one divider make the beam go round in one direction. The negative pulses from the other divider also generate negative pulses of identical duration and amplitude in a similar circuit and the resistance potential divider pulse and the integrated pulse are also connected to the opposite guides, which make the glow go round in reverse direction. These connections to the two guides are made through isolating diodes. Thus, when the output pulses from the two dividers are exactly at the same rate, the glow remains stationary and goes round clockwise or anti-clockwise depending upon the difference between the two pulse rates. The oscillator (7) of Fig. 1 can therefore be set exactly to the required frequency when the glow becomes stationary. The sine wave phase lock circuit then takes over the final phase lock.

### 8.1. *The 1 kc/s Sine Wave Phase Lock*

When the two dividers are thus set to generate identical pulse rates of 1 kc/s, as mentioned in Section 8, positive pulses of 80 μs duration are also available from anodes of valves V3 in each drive channel of the Dekatron in Fig. 12. These two positive pulses are fed to cathode follower valves V1 and V2 in Fig. 13. Bandpass filters of 800–1200 c/s bandwidths are connected across the cathode follower loads. The outputs of the two bandpass filters are therefore sine waves of the same frequencies as the input pulse repetition frequency. These sine wave voltages are amplified by valves V3 and V4 and feed a conventional phase detector circuit. The output of the phase detector is filtered by a simple R-C filter in addition to a shunt-connected, series tuned circuit resonating to 1 kc/s, and is fed to the reactance valve grid of Fig. 5.

It has been found that, with this simple phase comparison system, the oscillator of Fig. 5 locks within ± 1 c/s of the crystal frequency and needs to be set within about ± 4 kc/s of the required frequency before locking. The output of the controlled oscillator, when it is phase-locked, is a good sine wave and shows no phase or amplitude modulation of 1 kc/s or any other frequency.

## 9. Automatic Servo-Controlled Tuning

It is shown in Section 8 how the controlled oscillator can be manually tuned to the desired frequency by adjusting it till the glow in the bi-directional Dekatron becomes stationary, when the phase-locking circuit locks the oscillator to the crystal reference frequency. This tuning may also be effected by a simple servo-control system.

The 80 μs pulses from the two drive circuit of the bi-directional Dekatron can each feed a diode-pump type frequency counter circuit whose d.c. voltage output is proportional to the pulse input frequency. These voltages will be exactly equal when the two pulse rates are equal. If these voltages are applied to a difference amplifier operating a servo-control motor

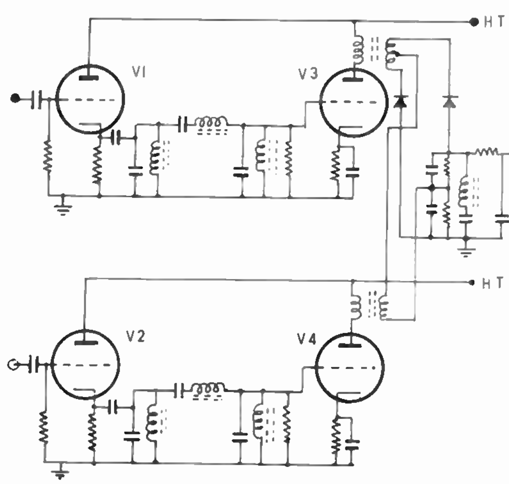



Fig. 13. 1 kc/s sine wave phase detector.

and if this motor turns the shaft of the tuning capacitor of the controlled oscillator, it will be tuned by the motor to lock on to the crystal reference, when the required frequency is produced. Such a simple system will also correct any drifts due to temperature or supply voltage variations which are likely to put the controlled oscillator out of lock and give an incorrect frequency.

## 10. Conclusion

It has been shown how, by designing a variable ratio divider using new devices called Trochotrons, pulse counting technique can be successfully applied to control the frequency stability of an L-C tuned oscillator, referring it to a single crystal frequency source of the required high stability. All the circuitry, including the dividers, could be transistorized as a Trochotron has now been developed which is capable of being driven by transistors.

## 11. Acknowledgments

The author is grateful to Messrs. C. G. Kemp and Dr. G. L. Grisdale for the encouragement to pursue this new approach to the frequency-control problem, and to Mr. D. J. Fewings for many useful discussions, and to the Engineer-in-Chief for permission to publish this paper.

*Manuscript first received by the Institution on 17th May 1960 and in final form on 8th February 1961 (Paper No. 630)*

# DISCUSSION

**Mr. A. J. H. Oxford:** Has Dr. Thatte weighed carefully the advantages and disadvantages of the Trochotron compared with more conventional counting circuitry. Could he give us reasons for choosing special valves rather than, say, transistorized binary counting circuits?

**Dr. R. P. Thatte** (*in reply*): The Trochotron requires only two additional twin triodes in the auxiliary driving circuit and therefore is more economical as a divide-by-ten stage than a feed-back cascaded binary decade. It has the additional virtue of extremely long life of 50 000 hours, and the ten target outputs are free from cross-talk; any mutual stray coupling introduced by the ratio-selection switch has negligible effect.

The choice of a Trochotron was also governed by the h.t. supply being chosen as greater than + 200 volts as required by the 15 millimicrosecond pulse generating circuit of the pulse-locked oscillator; the high-voltage Trochotron was therefore a more suitable choice than a hard valve binary.

With the introduction of the avalanche transistors for generating short duration pulses the whole circuitry could now be transistorized. One has to choose either transistor binary chains to secure a variable divider circuit capable of operating up to 1 Mc/s or use the low voltage Trochotron (Ericsson VS10K). Both types of circuit will have to be developed and compared before making a final choice.

# of current interest . . .

### Radio Trades Examination Board

The results of the Radio Servicing Examination for 1960 have recently been announced by the Radio Trades Examination Board who draw attention to the rising number of candidates taking the Board's examination in Radio and Television servicing.

Of the 2038 candidates who entered, 1965 sat the examination, 911 passed, 471 were referred in the practical test but passed the written papers, and 583 failed. The pass percentage for the year—46·4%—is a considerable improvement on 1959 (29·6%).

Details for the Television Servicing Examination for which the results were announced last Summer were as follows: 665 entered, 642 sat, 298 passed, 96 were referred in the practical test and 248 failed. The pass percentage of 46·5% was comparable with the previous year's results.

The 1961 examinations are being run under a new syllabus, for which the titles have been changed to Intermediate Radio and Television Servicing Certificate Examination and Final Radio and Television Servicing Certificate Examination. Entries are likely to be at least 25% over the 1960 figure. In addition, the first Electronics Servicing Examination is being held at the Intermediate level.

These increasing numbers of candidates require the services of additional examiners, both for written papers and practical tests, and there has been a good response to the request for examiners for practical tests in 1961. In addition six members of the Institution have been appointed examiners for the written papers in Radio and Television Servicing at Intermediate and Final levels, and the examiner in Electronic Servicing (Intermediate level) is also a member of the Institution.

### Anglo-French Space Group Formed

Details of what was described as "the first major move into space by British Industry" have been made known by Hawker Siddeley and S.E.R.E.B. (Société pour l'Etude et la Réalisation d'Engins Ballistiques), who have been working together for many months on feasibility studies of three space projects. These are: (1) Civil communications satellites, (2) Civil navigation satellites, and (3) Military anti-orbital missiles.

Sir Roy Dobson, Managing Director of the Hawker Siddeley Group, described these studies as an essential first step to any government-sponsored aero-space programme involving Europe and the Commonwealth. The projects are not allied to any particular rocket combination.

The two groups say that if their Governments authorized and financed a communication or a navigation satellite this year, the first experimental version could be launched by the end of 1964. The cost of such a programme would be about £140 million with the United Kingdom contributing £21 million and France £14½ million a year. Both groups emphasized that the finance must be provided by their Governments, since neither Hawker Siddeley nor S.E.R.E.B. could afford the cost of such a programme.

The Hawker Siddeley–S.E.R.E.B. initiative is in no way linked with the Anglo-European space proposals discussed at the European Space Research conference at Strasbourg recently, nor with the newly-formed British Space Development Company.

### 4th International Conference on Medical Electronics

As announced in the March *Journal* the 4th International Conference on Medical Electronics is to be held at the Waldorf Astoria, New York, on 16th-21st July. The British National Committee, concerned with contributions from this country, has been formed by the Bio-Engineering Society and the Institution is represented on the Committee by Dr. C. A. F. Joslin (Associate Member).

The International Conference is being combined with the 14th Annual Conference on Electrical Techniques in Medicine and Biology.

The Conference Programme is being organized along the following lines:

PLENARY SESSIONS

Mathematical and Biophysical Models for Simulation and Analyses

Biological Effects of Radio Frequency and Microwave Fields

A General Plenary Session

ROUND-TABLE DISCUSSIONS

Computers for Medicine and Biology

Perspectives in Bio-Medical Engineering Science

Electrocardiography

Acoustics and Ultrasonics

SPECIAL SESSIONS

Monitoring of Physiological Variables in Obstetrics

Control Systems in the Human Body

Advances in Recording Storage and Retrieval Techniques

Gastrointestinal Measurements

Electrodes and Amplifiers

Physiological Instrumentation

Education in Bio-Medical Engineering

Simultaneous translation to four languages will be provided and tours will be included to Scientific, Educational and Industrial Laboratories.

# 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".*

### VALVE MICROPHONY

Microphony is frequently a troublesome effect in valves and a Dutch paper describes a number of methods for its investigation. Certain direct methods requiring no special circuit arrangement can serve for comparing one valve with another, but these give no information on the cause of the microphony. A vibrator has therefore been designed by means of which a valve can be subjected to a vibration of constant peak acceleration and variable frequency. With the aid of a microscope and a stroboscope the components responsible for microphony can then be traced by directly observing their vibration. Some results achieved are illustrated by spectrograms. Finally, a method of using a noise generator is described, where the spectrogram is displayed on the screen of an oscilloscope.

"Microphony in electron tubes", S. S. Dagpunar, E. G. Meerburg and A. Stecker. *Philips Technical Review*, **22**, No. 3, pp. 71–88, 1960–61.

### QUARTER-WAVE PLATES

The quarter-wave plate is widely used as the conversion device between linearly polarized waves and circularly polarized waves. A broad-band characteristic is required of quarter-wave plates used in the circularly polarized aerials employed in microwave communication systems. In a paper from a Japanese telecommunications company, theories of the techniques of broad-banding are described for three types of quarter-wave plates which are obtained by utilizing two perpendicular modes in circular waveguide; they are dielectric slab type, differential type and additional type. The lumped susceptance type phase shifter which is combined with the hybrid junction to obtain the broad-band quarter-wave plate is also discussed.

"Broad-band microwave quarter-wave plate", T. Kitsuregawa, S. Nakahara and S. Tachikawa. *Mitsubishi Denki Laboratory Reports*, **1**, No. 4, pp. 21–52, October 1960. (In English.)

### RECTIFIER MODULATORS

An inter-connection of two filters via a ring-type or double push-pull modulator gives rise to attenuation distortion, since the filters are terminated at one end by the complex input impedance of the modulator which in turn is terminated by the other filter. This modulator input impedance can be calculated relatively simply if either the modulator input voltage or the modulator input current is sinusoidal or the source impedance is resistive. The input voltage is approximately sinusoidal, if the impedance of the preceding filter has a Π-characteristic at the modulator end, and the input current is sinusoidal if this filter input impedance has a T-characteristic. It is shown in a recent German paper that, to avoid attenuation distortion between modulator input and output of the second filter, the turns ratio of the output transformer of the modulator must be so proportioned that, with a sinusoidal input voltage, the second filter is matched to the short-circuit output impedance of the modulator, and that, with a sinusoidal input current, it is matched to the open-circuit output impedance of the modulator. If the output transformer is proportioned according to this rule and a Π-(T-)filter at the input of the modulator is combined with a Π-(T-)filter at the output, the modulator input impedance is approximately resistive and frequency-independent within the pass band of the first filter so that practically no attenuation distortion is introduced by imperfect termination of the first filter. A combination of a T-filter and a Π-filter is less favourable in this respect, but offers the advantage of a smaller frequency-independent portion of the modulator attenuation. This impedance-matching method is evaluated by reference to an example. The agreement between measured and calculated impedance and attenuation values is good.

"The impedance matching problem of ring and double push-pull modulators with filter terminations", H. Bauch. *Archiv der Elektrischen Ubertragung*, **15**, pp. 1–17, January 1961.

### GENERATOR FOR DELAYED PULSES

The Research Group on high temperature techniques at the Technical High School, Stuttgart, describes a device which produces at a number of outputs short rectangular pulses with very steep edges whose onset in time can be shifted from 0·1 μs to 1 ms with respect to each other with high relative accuracy. The delay times are varied by the coarse setting of four decadic controls while a fine setting of intermediate values is effected on a rotary potentiometer. The pulses generated are 0·4 μs long and have an amplitude of 120 V, while the leading edge has an amplitude of 150 V. The rise time of the leading edge, measured between 10% and 90% of the pulse amplitude, is 0·02 μs. For setting the oscilloscope the device can also be run periodically with a variable pulse repetition rate of 0·5 to 50 c/s. An extra trigger input provides for optional external triggering. The time jitter of the pulse onset is $< 0{\cdot}1\%$. An application of the instrument to high temperature plasma investigations is briefly described.

"A multi-channel impulse time delay generator", H. H. Maier, H. G. Hartner and E. Pfender. *Archiv der Electrischen Ubertragung*, **14**, pp. 515–9, November 1960.

## MICROWAVE NOISE GENERATOR

The determination of the sensitivity of a receiver by measuring its noise factor is one of the more important measurements which have to be made in connection with microwave equipment: it is usual to employ a noise generator and numerous designs have been put forward. Two engineers at the Institute of Wireless Telecommunication in the Polytechnical University of Budapest have recently described a microwave generator which makes use of a gas discharge tube as part of the inner conductor of a coaxial line using the TEM mode. The basic principles of this device are similar to those described in a paper in the Brit.I.R.E. *Journal* by Kollanyi (published in September 1938), but it is claimed that by avoiding the use of a helical line, the mechanical construction is simpler and no special additional matching elements are required. Although the noise generator does not match the connecting transmission line when the gas discharge tube is inactive, the authors state that this is an inconvenience rather than a disadvantage and they further state the generator can be used in a very broad band with the use of proper matching elements. In a theoretical section of the paper it is proved that the noise power of the generator is a function of the noise temperature of the electron in the gas discharge tube alone.

"New microwave noise generator for the 2000 Mc/s band", G. Almassy and I. Frigyes. *Periodica Polytechnica, Budapest*, **4**, No. 4, pp. 293–303, 1960. (In English.)

## AERIAL RADIATION MEASUREMENTS

Some tests to establish the field patterns of various aerials used by the Australian Department of Civil Aviation have been described in a paper presented at the last Australian I.R.E. Convention. A feature of these tests was the use of aircraft. The types of array tested were an aerial of horizontally arrayed dipoles, a rhombic aerial, an inclined vee and Franklin aerials. The test assisted investigation of aerial interaction effects and provided information of the off-frequency performance of the aerial of horizontally arrayed dipoles. The paper gives details of the techniques employed to determine aircraft position relative to the aerials.

"Field pattern measurements of various h.f. directional aerials using aircraft", R. T. Rye. *Proceedings of the Institution of Radio Engineers Australia*, **21**, pp. 879–85, December 1960.

## COMPUTATION OF ELECTRON GUN CHARACTERISTICS

A recent French paper gives an account of the electron optics methods used at the Centre National d'Etudes des Telecommunications for the study of guns for producing rotating electron fields such as those of klystrons and travelling-wave tubes. After recalling the basic optical principles for intense beams in rotating electric and magnetic fields, a description is given of the rhegraphic cavity which makes it possible to measure the potential and the analogue computer with which the automatic tracing of the path has been studied.

"Study of electron guns with the help of rhegraphic and analogue computers", Mrs. J. Henaff and J. Le Mezec. *Onde Electrique*, **41**, pp. 36–53, January 1961.

## F.M. SIGNAL/NOISE CALCULATION

By comparing the information contents in the signal at the input and output of an f.m. demodulator a Czech engineer has derived the improvement of the signal/noise ratio attainable theoretically by frequency modulation as compared with amplitude modulation. The results compare well with data gained in the classical way. The analysis is carried out for discriminators of the normal type as well as of the counting type; they are shown to be equivalent from the viewpoint of signal/noise ratio.

"Signal/noise ratio with frequency modulation", V. Pollak. *Slaborproudy Obzor (Prague)*, **22**, pp. 36–40, January 1961.

## IMPROVEMENT OF MICROWAVE MIXER PERFORMANCE

In a microwave repeater for frequency modulated multichannel telephone and television relay links, the reduction of delay distortion, the flatness of amplitude-frequency response, and the stabilization of these characteristics are essential requirements. These characteristics in conventional crystal mixers fluctuate considerably at each adjustment of the matching device. Observing that the cause of such deterioration and unstability come chiefly from undesired sideband components, a Japanese engineer has developed a very simple but effective method. For the receiving mixer, he examined the variation of the image frequency impedance due to the adjustment of the various matching devices and found the most suitable matching devices. A very simple phase shifter which keeps the image frequency impedance to a desired value was therefore developed. The transmitting mixer which suffers from many undesired sideband components, is improved by the application of a special balanced mixer, which is substantially free from undesired sideband components by locating one of the crystal detectors in a position shifted one-fourth of the wavelength. The result of experiments showed an improvement of about 15–20 dB compared with the conventional balanced mixer.

It is claimed that the quality of a 4000 Mc/s heterodyne repeater has been considerably improved and stabilized.

"Microwave mixer without the influence of undesired sideband components", T. Kawahashi. NEC Research and Development (Tokyo), No. 1, pp. 36–45, October 1960. (In English.)

## COMPONENT RELIABILITY

In the November issue of *NTZ* (the organ of the Radio Engineering Section of the German Association of Engineers), four papers are published which were read at a meeting on Component Reliability. The titles are as follows:—

" Technological Measures for an Improvement of the Reliability of Components " by K. H. J. Rottgardt.

" Life-time Investigations on Capacitors " by W. Ackmann.

" The Effect of Humidity on the Electrical Characteristics of Capacitors " by H. Veith.

" Poor Reliability of Electronic Equipment and its causes " by H. J. Frundt.

*Nachrichtentechnische Zeitschrift*, **13**, No. 11, pp. 505–12, 513–8, 519–23, 524–8, November 1960.