FOUNDED 1925
INCORPORATED BY
ROYAL CHARTER 1961

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

THE RADIO AND ELECTRONIC ENGINEER

The Journal of the Institution of Electronic and Radio Engineers

VOLUME 33 No. 5 MAY 1967

Integrating Design and Production

IT is vitally important in any industry for the inventions of the research and development engineer to be translated efficiently and economically into production models. In no branch of technology is the pace of invention and innovation more rapid than in electronics—and the need for the processes of design and production to be linked smoothly together is correspondingly greater. One problem of increasing specialization in technology is the difficulty of ensuring adequate communication between different branches of engineering where, as in so many fields of endeavour, advances depend on human understanding and co-operation. Certainly the 'integration' of design and production relies on the creation of a full and sympathetic understanding of the problems of development, production and 'user' engineer.

The purpose of the three-day Conference which is being organized by the Institution at Nottingham University in July is to enable design engineers and production engineers to explore the *processes* of design and production and the ways in which new discoveries can make their integration more effective. Accordingly the Institution's Council invited the Institution of Production Engineers and the Institution of Electrical Engineers to participate in planning the Convention. Authoritative speakers from a wide range of organizations were invited to contribute papers which would enable the theme to be developed in a logical and complete fashion: how far this has succeeded may now be judged from the outline programme on page 282 of this issue. The final list of papers, with their synopses, will be published in the June issue of *The Radio and Electronic Engineer*.

After an introductory address which will point out likely developments for the future, the opening paper of the first session on 'Design for Production' will discuss the decisions which determine *what* shall be made. Subsequent papers will describe some of the considerations affecting *how* a product shall be made, including the great importance of designing for ease of maintainability in the finished product.

The theme for the second session, 'Economical Production', will be set by a paper discussing the interface between design and production, and other papers will deal with cost and quantity considerations, including the application of computers to control of production. A symposium of short papers will present some recent developments in materials, components and techniques.

The final day, on 'Future Developments', includes a group of papers summarizing the possibilities in automatic assembling and testing, followed by a discussion of the impact of integrated electronic circuits and the future of computer-aided design.

The earnest hope of the Organizing Committee is that the Conference will provide a really effective opportunity for design engineers and production engineers to meet, discuss their successes and failures, and go some way along the road to drawing into an integrated whole what are all too often separate, and perhaps sometimes even conflicting, disciplines. We are all engineers, and achieving common ground for the discussion of development and manufacturing techniques is basically a communication problem which should be well within the grasp of all members.

J. A. SARGROVE

INSTITUTION NOTICES

Election as Fellows of the Royal Society

Council has learned with pleasure of the election, on 16th March 1967, of two members of the Institution as Fellows of the Royal Society:

Mr. Ieuan Maddock, O.B.E., B.Sc., C.Eng. (Member), who has served on the Council for a number of years and was from 1963 to 1966 a Vice-President, has recently been appointed Controller, Industrial Technology, at the Ministry of Technology. He was previously Head of the Applied Physics Division, Atomic Weapons Research Establishment, U.K. Atomic Energy Authority, and the Royal Society's citation reads: 'Distinguished for his contributions to the design, engineering and operation of recording systems, including work on the recording of nuclear explosions'.

Dr. Francis Edgar Jones, M.B.E., B.Sc., Ph.D., C.Eng. (Member) is Managing Director of Mullard Ltd. During the war he was the leader of a group at T.R.E., Malvern, working on the blind bombing system known as 'Oboe'; before joining the Board of Mullard Ltd. in 1956, he was Deputy Director of the Royal Aircraft Establishment. The Royal Society states that he is 'Distinguished for his contributions to radar and infra-red technology, and for his outstanding technological leadership in an advanced industry'.

Magnetic Materials and their Applications

A conference on Magnetic Materials and their Applications will be held at the Institution of Electrical Engineers, London, from 26th to 28th September, 1967. The conference is sponsored jointly by the I.E.E., the Institute of Physics and the Physical Society, and the U.K. and Eire Section of the Institute of Electrical and Electronics Engineers.

Topics to be discussed will include the preparation, properties and applications of all forms and types of hard and soft magnetic materials. The applications concerned will be those in which the choice of the particular magnetic material is important. The physical theory of magnetism will not be included.

Registration forms and further details are available from the Joint Conference Secretariat, Institution of Electrical Engineers, Savoy Place, London, W.C.2.

Course on Microwave Techniques

A Summer Course on Microwave Techniques will be offered from 4th to 15th September, 1967, at the Electronics Laboratories of the University of Louvain, Belgium. The course is intended for graduate engineers who have not so far specialized in microwave work. The topics to be covered are: Guided waves, obstacles in waveguides, cavities, junctions and passive non-reciprocal devices, periodic structures and filters.

Several seminars will be held, covering measurement techniques, microwave tubes, ferrites, low-noise devices and their use in spatial communications, microwave solid-state and industrial applications.

The tuition fee is $200 and further information can be obtained from the course co-ordinator: Professor A. S. Vander Vorst, Sc.D., Electronics Laboratories, 94 Kardinaal Mercierlaan, Heverles-Leuven, Belgium.

U.S. National Medal of Science Awarded to Dr. V. K. Zworykin

Dr. Vladimir K. Zworykin (Honorary Member) received the National Medal of Science from President Lyndon B. Johnson at a ceremony at the White House, Washington, on 6th February last. The citation for Dr. Zworykin's award read: 'for major contributions to the instrumentation of science, engineering and television and for his stimulation of the application of engineering to medicine'.

Dr. Zworykin is now Honorary Vice-President of the Radio Corporation of America, and in recent years he has been Chairman of the Board of Governors of the International Institute for Medical Electronics and Biological Engineering.

Fuller details of his career appeared in the April and September 1959 issues of the *Journal* in connection with his presentation of the Fourth Clerk Maxwell Memorial Lecture and his election as an Honorary Member of the Institution.

Aerospace Instrumentation Symposium

The Fifth International Aerospace Instrumentation Symposium will be held at the College of Aeronautics, Cranfield, from 25th to 28th March 1968. The Symposium is sponsored by the College of Aeronautics and the Instrument Society of America, Aerospace Industry Division.

Papers are invited for presentation at this symposium and summaries, of about 300 to 500 words, should be submitted to the Symposium Organizer for consideration by 1st June 1967. Appropriate subjects in the applied science and technology of instrumentation of aerospace vehicles may be chosen, but papers concerned with the following subjects will be particularly welcome: instrumentation for helicopters and/or VTOL aircraft; development in transducing techniques; on-line data processing for aerospace flight testing; analysis of non-stationary data.

Further information is available from the Symposium Organizer, Department of Flight, The College of Aeronautics, Cranfield, Bedford, England.

U.D.C. 621.372 : 656.25

High Frequency Guided Electromagnetic Waves in Application to Railway Signalling and Control

By

Professor H. M. BARLOW,
F.R.S., C.Eng., F.I.E.E., M.I.Mech.E.†

Presented at a meeting of the Institution's East Anglian Section in Chelmsford on 16th February 1967.

Summary: The use of guided high frequency electromagnetic waves transmitted along the side of a railway track for the purpose of providing radar location of trains as well as continuous telephonic communication with drivers and guards on the trains is discussed. Attention is drawn to the advantages and shortcomings of different systems already investigated experimentally, and the direction in which further work now being pursued is likely to develop.

1. Introduction

The long-distance railway systems of the world are today facing a serious challenge from the rapidly developing air services, and if trains are to maintain their place as a vital means of transport for both people and freight, high speed operation is essential. Some excellent surveys of the requirements, particularly in regard to British Railways, and of the present limitations imposed by equipment available, have been given by various authors.[1-4] In these papers the history of railway development towards more and more electrical signalling and control is outlined, emphasizing the enormous capital outlay that necessarily goes with innovations and the consequent importance of the time factor in future progress. As the pioneer nation in railway working we have inherited a network of tracks all over the country that do not always lend themselves readily to the needs of the present day and rationalization of the system clearly presents many difficulties. However, the task is not insoluble and, as everyone knows, much time and effort is being given in many countries of the world to ensuring that their long-distance railway systems will continue as viable enterprises side-by-side with their competitors in the air and, to a lesser extent, on the roads.

For some time past British Railways have had in mind the possibility of using high frequency electric waves guided along a main line railway track as a means of locating trains, signalling, control and communication with those in charge on the trains (Fig. 1). Similar ideas have been very much to the fore in Japan where the new high-speed Tokaido line between Tokyo and Osaka providing for passenger trains running regularly at 150 miles/h has emphasized the importance of a new approach to the problem. There seems little doubt that continuous information, to drivers and guards about the progress of the train ahead and to the various control points about the disposition of all the traffic in the area, is essential at the very high speeds envisaged for the future. A number of possible solutions have been under consideration for some time, including the ingenious zigzag conductor coupling circuit designed by Ogilvy[3] and a variety of high frequency guided-wave systems. It is the object of the present paper to discuss particularly the possible application of a form of surface waveguide as a means of meeting the problem and to consider the advantages and disadvantages of this arrangement side-by-side with other guided-wave systems that have been proposed for operation at high frequencies. In doing this it must be made clear that the work is still in its infancy and results so far obtained are only on an experimental basis, not directly applicable to practical railway operations.

† Department of Electrical Engineering, University College London.

2. The Particular Features of a High-frequency Guided Wave System that Appear to Offer the Most Attractive Possibilities

Continuous information, which is now regarded as so essential to satisfactory running of trains at high speeds, can only be made available by uninterrupted coupling between trains and a circuit running along the track. Block working cannot fully satisfy the requirements nor can the use of track circuits to provide multiple-aspect signalling. There are many different ways of guiding high-frequency electromagnetic waves along a pre-determined route but the need to couple continuously to an adjacent object, namely the train, implies incomplete screening of the wave. Thus an ordinary waveguide can only be applied when arrangements are made for some radiation along the length (by, for example, slots in the outer wall). If such radiation takes place all the time

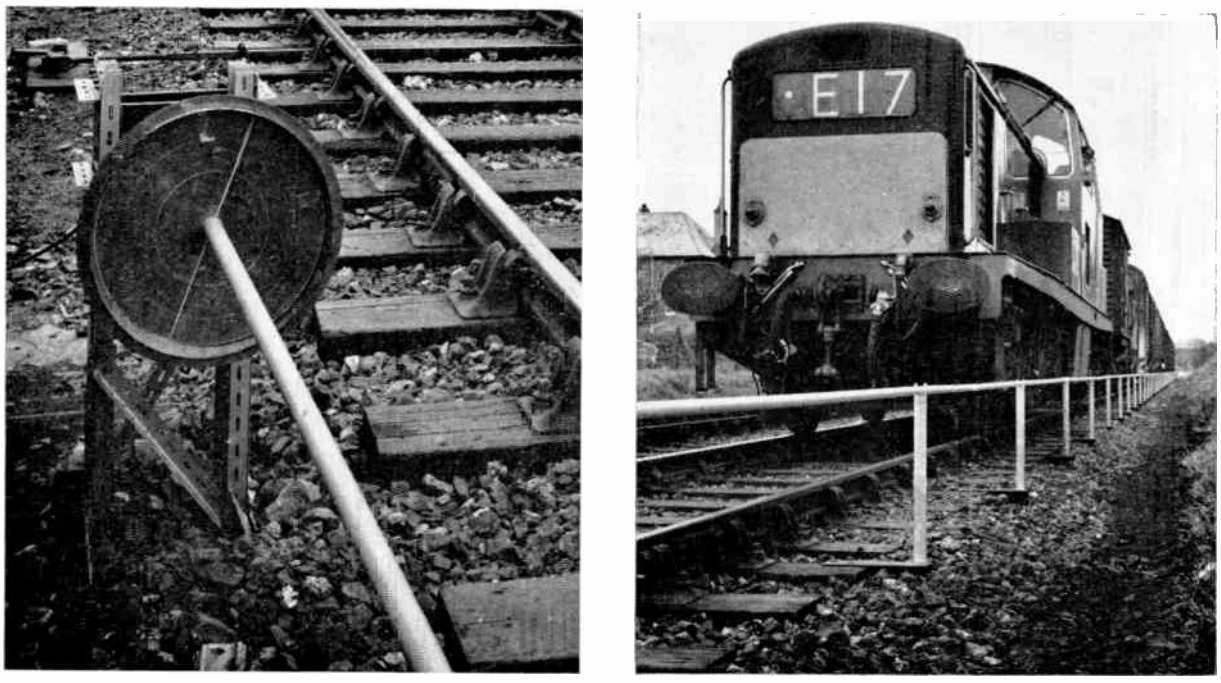


Fig. 1. Experimental trackside axial cylindrical surface waveguide designed by British Railways and Ferranti Ltd.

(Courtesy of the Editor of Electronics & Power)

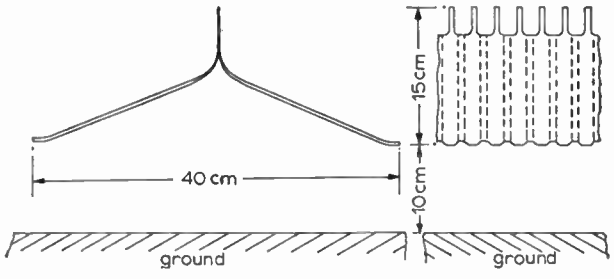
and is distributed over the whole waveguide run, the losses build up, requiring quite a large power supply even although the waveguide itself may be of a low dissipation kind. It happens that the single-wire transmission line as a surface waveguide has particularly attractive features in this situation.[3] It is well known that such a guide when perfectly straight does not normally radiate while carrying 80% of the power along its surface with a field spread in the air outside of only a few wavelengths. Nevertheless any obstacle in the path of the wave sets up radiation with power flow to the obstacle and if the obstacle moves along the side of the waveguide, the radiation moves with it. Thus, when the obstacle is represented by a train, the power flow is exactly as required, giving maximum conservation of usable energy, while at the same time a partial reflection of the surface wave provides for radar location of the train along the track.

If the surface waveguide has a fairly high surface reactance (e.g. dielectric coated conductor) the guide can be bent to any radius of curvature likely to be met on a railway track, without setting up any significant radiation due to the bend.[5] There is, however, a serious difficulty with the ordinary single-wire transmission line as a waveguide, arising from the fact that it is normally unprotected and a bird perching on the wire, ice or snow, causes a discontinuity with consequent unwanted radiation. For this reason a screened form of surface waveguide has been proposed[6] and is now under investigation. The particular features of this arrangement will be discussed as well as those of other waveguides of interest in connection with the railway problem.

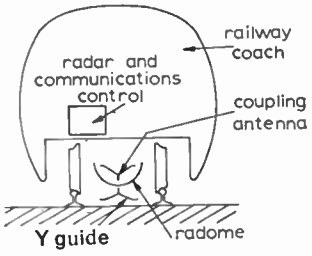
3. Guided Wave Systems Developed in Japan for Application to Railway Signalling and Control

Outside Great Britain perhaps the most significant work on the use of very high frequency guided waves has been done in Japan. The attractive features of a surface wave were appreciated at an early stage, with the result that the inverted Y-section corrugated aluminium guide (Fig. 2), with an attenuation of about 15 dB per km, at 900 MHz, was developed.[7] In this arrangement the surface wave sits on top of the guide, hugging the corrugated ridge, and coupling to it is provided by a short length of the same guide inverted and attached to the underside of the train. The configuration avoids serious disturbance of the wave due to the proximity of the earth but it gives no protection against interference with smooth propagation of the wave by extraneous objects in the vicinity of the guide.

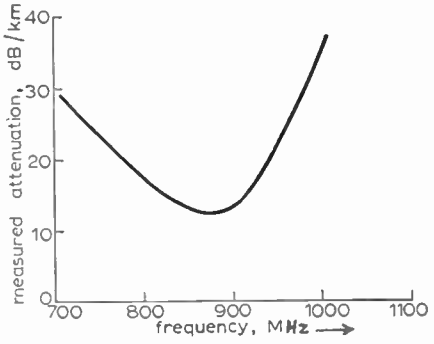
The Japanese Railway Authority, in collaboration with industry in that country, evidently came to the



(a) corrugated aluminium Y-guide for support of surface wave;



(b) arrangement of Y-guide and coupling to it;



(c) surface-waveguide attenuation at different frequencies.

Fig. 2. Application of surface wave to control of trains. (After Nakahara[7]).

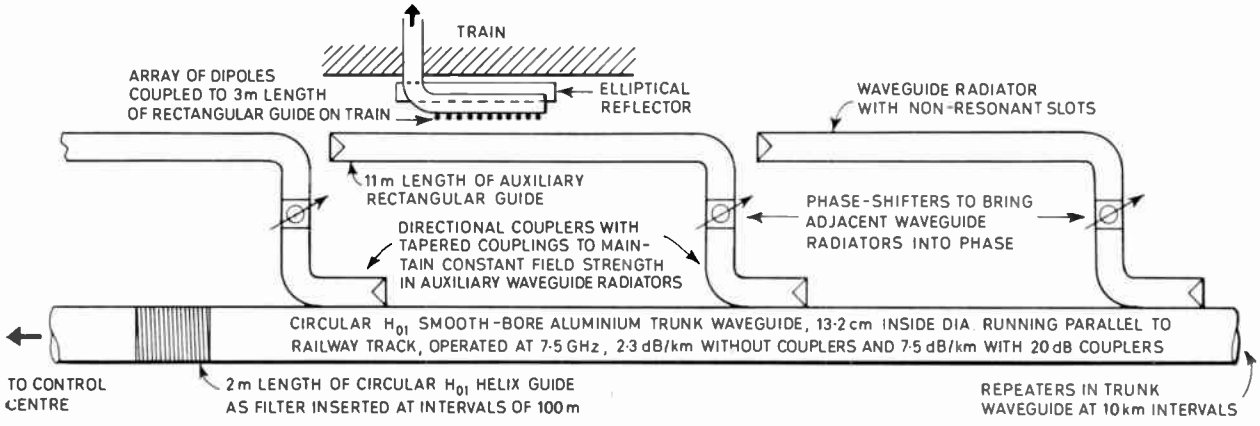
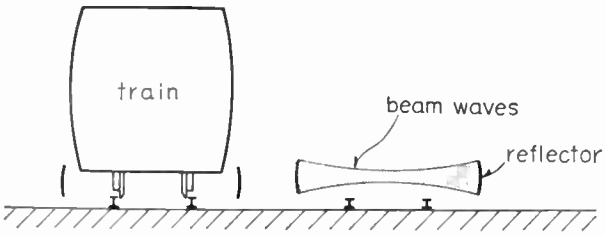


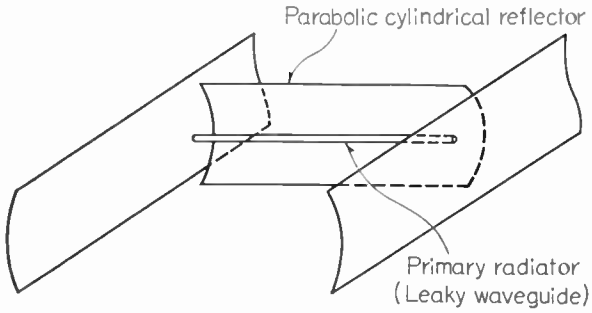
Fig. 3. Waveguide communication system for centralized railway traffic control developed for Japanese National Railways. (After Kawakami *et al.*[8])

conclusion as a result of their surface wave experiments that a leaky low-loss waveguide might be more appropriate, and the system shown in Fig. 3 was set up for investigation.[8] Here a circular H_{01} waveguide is employed at 7·5 GHz as a means of feeding auxiliary slotted waveguide radiators coupling to an array of dipoles on the train. It is claimed that the overall loss of the system, including the power coupled out, which goes on independently of any train on the track, is only 7·5 dB/km, about half the loss in the surface waveguide. Thus the very low transmission loss of the circular H_{01} waveguide more than compensates for the continuous radiation.

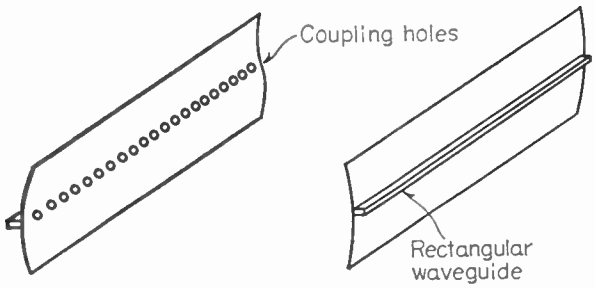
Still more recently a beam wave transmission arrangement (Fig. 4) has been described.[9] Running along each side of the track is a reflector so shaped as to form a beam of high frequency energy flowing in the space between them. A confocal setting of the reflectors proved to be most suitable and they were excited at about 10 GHz, either by a transverse leaky waveguide radiator of the insertion type or through coupling slots distributed along the reflectors. A train in the path of the beam not only forms an obstacle giving a radar picture of its location, but is also provided with means to pick up energy from the beam for telecommunication purposes. Unfortunately



(a) Railway application of beam wave transmission.



(i) Insertion type.



(ii) Side coupling type.

(b) Beam mode exciter.

Fig. 4. Guided beam waves between parallel concave reflectors in application to railways (after Nakahara and Kurauchi[9]).

the overall attenuation of the beam is very high, amounting to about 200 dB/km even when great care is taken in the alignment of the reflectors.

Ingenious as all these Japanese developments are, they appear to present many engineering difficulties in application and to be very costly.

4. The Screened Surface Waveguide Investigation for British Railways

In 1953 Professor Cullen and the author[10] drew attention to a form of hybrid wave mode that can be set up between two parallel reactive supporting surfaces. Further investigation[6] has shown that it is possible to produce this wave in asymmetrical form (Fig. 5) and that in such circumstances it can have very interesting properties. In the first place, the wave is not subject to cut-off, being a combination of TEM and surface waves, one of which is associated with each of the two surfaces. When supported by a coaxial guide the outer conductor can be removed and the wave then transforms itself into that of the single wire transmission line, described as an axial cylindrical surface waveguide. Provided that the smooth progression of the wave is not disturbed there is no radiation and a gradual transition from the screened line to the open wire substantially ensures this. Moreover this hybrid TEM/dual surface wave may be partially screened using a concentric outer conductor with a sector opening running along its length. Along a line midway between the sides of the opening the field spreads out theoretically to an infinite radius and follows the usual Hankel function distribution of the unscreened surface wave, while within the part covered by the coaxial cylinder the hybrid mode is required to satisfy the boundary conditions. Thus in terms of the circumferential magnetic field component H_θ we can write, referring to Fig. 5:

(i) inside the screen

$$H_\theta = \frac{j\omega\varepsilon_0}{h} e^{-\gamma x}[C' H_1^{(1)}(hr) + C'' H_1^{(2)}(hr)] e^{j\omega t} \quad \text{.....(1)}$$

where

$$C''/C' = -\frac{H_0^{(1)}(hr_m)}{H_0^{(2)}(hr_m)} \quad \text{.....(2)}$$

with $E_x = 0$ at $r = r_m$

$$h = ju = j(a - jb) \quad \text{.....(3)}$$

$$\gamma^2 + u^2 + \omega^2\mu_0\varepsilon_0 = 0 \quad \text{.....(4)}$$

$$\gamma = \alpha + j\beta \quad \text{.....(5)}$$

(ii) in the centre of the sector opening

$$H_\theta = \frac{j\omega\varepsilon_0}{h} e^{-\gamma x}[C' H_1^{(1)}(hr)] e^{j\omega t} \quad \text{.....(6)}$$

with eqns. (3), (4) and (5) still applicable.

If the sector opening in the outer wall of the guide is flared there will be a gradual transition from the distribution of magnetic field given by eqn. (1) to that of eqn. (6) while the current in the supporting surface will remain longitudinal throughout. The requirements can in principle be satisfied if $r_m = \infty$ at the centre of the aperture, so that $C''/C' = 0$ along the line $\theta = 0$. The corresponding wave impedance, looking radially outwards, transforms from

$$Z_{r_3} = -\frac{h}{j\omega\varepsilon_0}\left[\frac{H_0^{(1)}(hr_3) + C''/C' H_0^{(2)}(hr_3)}{H_1^{(1)}(hr_3) + C''/C' H_1^{(2)}(hr_3)}\right] \quad \text{.....(7)}$$

at the inner surface (radius r_3) of the screen to:

$$Z_r = -\frac{h}{j\omega\varepsilon_0}\left[\frac{H_0^{(1)}(hr)}{H_1^{(1)}(hr)}\right] \quad \text{.....(8)}$$

in the absence of the screen.

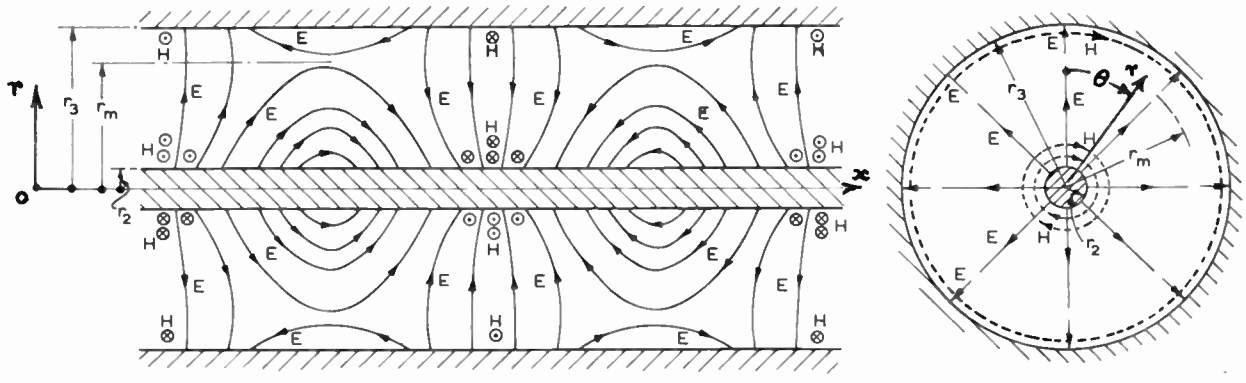


Fig. 5. Hybrid TEM/dual surface wave in coaxial guide (after Barlow[6]).

Assuming that the radius r_m, which defines the range at which the axial component of electric field is zero and has a constant value within the screen, increases progressively to infinity at $\theta = 0$ in accordance with the relation:

$$r'_m = \frac{r_m}{\sqrt{\pm \sin 2\theta}} \qquad \text{.....(9)}$$

For angles between $\theta = 0$ and $\theta = \pm \pi/4$ we get an r'_m contour (Fig. 6(a)) which follows a rectangular hyperbola in cartesian co-ordinates based on the centre-line of the aperture. If the screen is a bare metal surface, as seems convenient in this application, r_m within it is very nearly equal to r_3 and the surface impedance is

$$Z_{r_3} = R_{r_3} + jX_{r_3} \qquad \text{.....(10)}$$

where

$$R_{r_3} = X_{r_3} = \sqrt{\frac{\omega\mu}{2\sigma}} \qquad \text{.....(11)}$$

and σ is the conductivity of the metal.

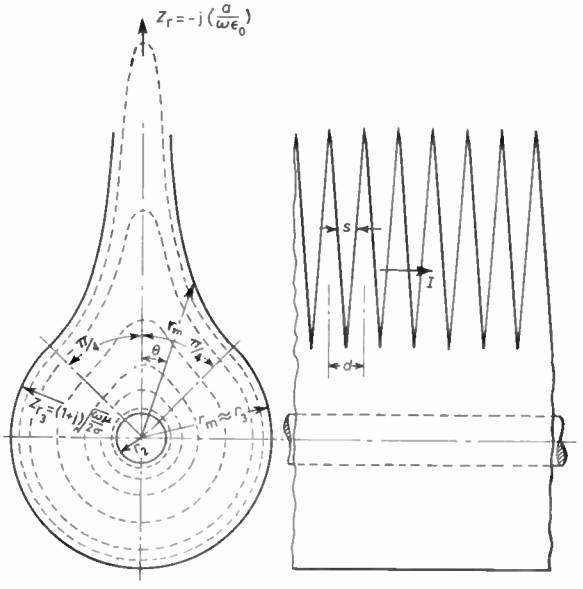
On the other hand eqn. (8) gives a value for Z_r which is substantially a capacitive reactance and the transition from Z_{r_3} to Z_r which must be made gradual can be suitably provided for by using tapered teeth cut in the screen between $\theta = \pm 45°$, as shown in Fig. 6(a).

For given values of s and d the equivalent circuit for this toothed structure is approximately as in Fig. 6(b). Following Macfarlane[11]

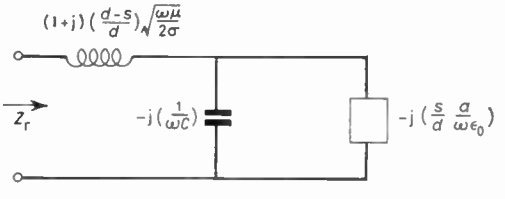
$$C = \frac{2\varepsilon_0 d}{\pi} \log_e \left(\frac{1}{\sin(\pi s/2d)} \right) \qquad \text{.....(12)}$$

and using the large argument approximations for the Hankel functions in eqn. (8)

$$Z_r \simeq -j\left(\frac{a}{\omega\varepsilon_0}\right) \qquad \text{.....(13)}$$



(a) Coaxial waveguide with flared longitudinal slot in outer conductor.



(b) Equivalent circuit from which surface impedance of flared metal teeth is calculated.

Fig. 6. Partially screened waveguide for application to railway signalling and control.

As more and more of the metal is cut away in approaching the sharp ends of the teeth, the proportion of surface impedance presented by the metal falls while that of the air rises. This is reflected in the values of the components of the equivalent circuit. Thus the input impedance Z_r of this circuit is given by:

$$Z_r = \left[\left(\frac{d-s}{d}\right)\sqrt{\frac{\omega\mu}{2\sigma}}\right] - \mathrm{j}\left[\frac{1}{\omega\left(C+\frac{\varepsilon_0 d}{sa}\right)} - \left(\frac{d-s}{d}\right)\sqrt{\frac{\omega\mu}{2\sigma}}\right] \quad \ldots\ldots(14)$$

When $s = d$ then $C = 0$ and

$$Z_r = -\mathrm{j}\left(\frac{a}{\omega\varepsilon_0}\right)$$

while for $s = 0$ we have $C = \infty$ and

$$Z_r = Z_{r_3} = (1+\mathrm{j})\sqrt{\frac{\omega\mu}{2\sigma}}$$

as required.

It is, of course, important that the wavelength should cover about five teeth to give an average surface impedance as in eqn. (14). Taking as an example the case depicted in Fig. 6(a) for which $f = 1$ GHz, $r_3 = 7{\cdot}5$ cm, $r_2 = 1{\cdot}5$ cm with $a = 6\ \mathrm{m}^{-1}$ we find that the most significant term in eqn. (14) represents the impedance of the air space between the teeth. The profile of the metal screen in the vicinity of the aperture may be expected to follow closely the r'_m contour.

Very recent investigations[12] have shown too that by suitably arranging the surface reactances of the walls of a coaxial guide, supporting the hybrid TEM/dual surface wave, the power density distribution over the cross-section can be so adjusted as to make a significant reduction in the overall losses.

The partially screened guide with the sector opening in the outer wall not only gives the much sought after measure of protection for the wave against extraneous objects, but it also provides for access to the wave by a deliberate obstruction. Thus this form of guide offers attractive possibilities in application to railways. Coupling to the wave through the sector opening can be on the under side, giving protection from weather and at the same time providing for a part of the train itself or a pick-up dipole attached to the train to probe the field. The possibilities are being investigated at present working at 1 GHz with an inner metal conductor of 3 cm diameter coated with dielectric to enhance its surface reactance and a coaxial outer cylinder 15 cm inside diameter which may or may not be coated with a thin layer of dielectric.

Sideways movement of the field probe attached to the train relative to the guide presents something of a problem, but it is expected that this can be overcome by a differential coupling arrangement whose net output does not vary much due to such movements.

5. Proposals for Locating and Identifying a Number of Trains on a Given Track

Figure 7 shows a schematic of a simple surface waveguide system, developed by the author's colleagues H. G. and M. Effemey, for locating, identifying and providing communication with a number of separate trains along a railway track. Continuous wave transmitters operating at discrete frequencies f_1, f_2, f_3, etc., with adjacent ones separated by about 200 MHz, are used in parallel to feed one end of the waveguide, while on each train is a resonant cavity coupled to the waveguide and having a centre frequency corresponding to that of one of the transmitters. The other end of the waveguide is connected to a number of receivers, one for each frequency f_1, f_2, f_3, etc. Under resonant conditions the input impedance of the probe coupling to a given cavity is small and consequently in such circumstances the probe makes a significant disturbance to the smooth propagation of the surface wave transmitted at the relevant frequency. Within each cavity is mounted a varactor diode which, when subject to a modulation voltage, detunes the cavity and makes the associated coupling probe of lesser consequence as an obstacle in the path of the surface wave. Thus not only is a radar picture locating each train on the track obtainable at the receiving end of the guide, but the trains may be separately identified and contacted by telephone at all times by using

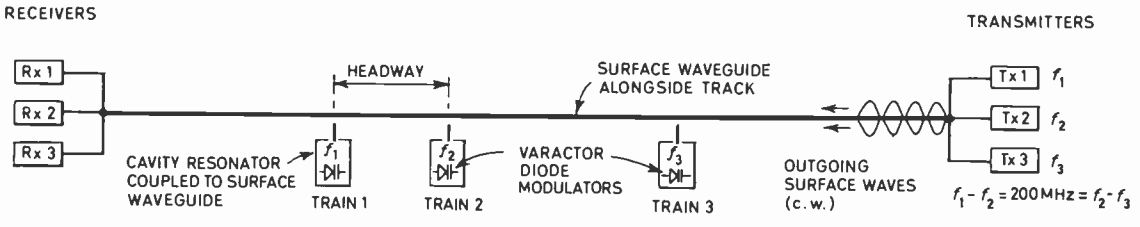


Fig. 7. Application of high frequency surface wave to location and identification of trains on a railway track.

suitable modulation of the varactor diodes. The 200 MHz frequency separation between adjacent surface surface waves is found to be quite sufficient to avoid interference one with the other. Moreover the equipment on any one of the trains, with this arrangement, is exceedingly simple.

In a practical application it would be necessary to present each train driver with continuous information about the position on the track of the next train ahead, thus enabling the requisite headway to be maintained inviolate. This would, of course, add to the equipment requirements, but the system described offers a sound basis on which to build.

6. Conclusions

The problem of railway signalling and control in application to trains operated at high speeds has been discussed from the point of view of providing continuous information derived from high frequency electromagnetic waves guided along the side of the track. It is concluded that techniques now available offer an encouraging prospect in this kind of development, but further experimental work is required to establish the most suitable arrangement, including particularly the waveguide system.

7. Acknowledgments

The author wishes to thank British Railways (Dr. S. Jones, Dr. L. L. Alston, Mr. H. H. Ogilvy) for continuing support of research on this project, and to acknowledge his indebtedness to his colleagues Mr. H. G. Effemey, Mr. M. Effemey, Miss M. Sen and Mr. E. M. A. Eid for their contributions to the work.

8. References

1. S. Jones, 'Automatic control in railway systems', *Electronics & Power*, **12**, p. 180, June 1966.
2. F. T. Barwell, 'Research for railway electrification', *Electronics & Power*, **12**, p. 311, September 1966.
3. H. H. Ogilvy, 'Radar on the railways', *Electronics & Power*, **10**, p. 146, May 1964.
4. F. T. Barwell and H. H. Ogilvy, 'Communications and their effect on railway operations', *Proc. Instn Railway Signal Engineers*, p. 135, 5th January 1966.
5. H. M. Barlow and J. Brown, 'Radio Surface Waves'. (Clarendon Press, Oxford, 1962.)
6. H. M. Barlow, 'Screened surface waves and some possible applications', *Proc. Instn Elect. Engrs*, **112**, No. 3, p. 477, March 1965.
7. T. Nakahara, 'Corrugated Y-guide for surface wave transmission', *Sumitomo Electric Technical Review*, April 1963.
8. T. Kawakami, T. Maruhama, T. Takeya and S. Kohno, 'Waveguide communication system for centralized railway traffic control', *Trans. Inst. Elect. Electronics Engrs on Vehicular Communications*, VC–13, No. 1, p. 1, September 1964.
9. T. Nakahara and N. Kurauchi, 'Guided beam waves between parallel concave reflectors', *Sumitomo Electric Technical Review*, No. 7, p. 44, February 1966.
10. H. M. Barlow and A. L. Cullen, 'Surface waves', *Proc. I.E.E.*, **100**, Pt. III, p. 329, 1953.
11. G. G. Macfarlane, 'Quasi-stationary field theory and its applications to diaphragms and junctions in transmission lines and wave guides', *J.I.E.E.*, 93, Part IIIA, p. 703,1946.
12. H. M. Barlow, 'New features of wave propagation not subject to cut-off between two parallel guiding surfaces', *Proc. I.E.E.*, **114**, No. 4, p. 421, April 1967.

Manuscript received by the Institution on 25th January 1967 (Paper No. 1114.)

Joint I.E.R.E.-I.Prod.E.-I.E.E. Conference on

"The Integration of Design and Production in the Electronics Industry"

10th to 13th JULY 1967, UNIVERSITY OF NOTTINGHAM

OUTLINE PROGRAMME

Monday Evening, 10th July (7 for 7.30 p.m.)

Conference Dinner

Tuesday, 11th July (9.30–12: 2–5.30: 8–9.30)

Session I: DESIGN FOR PRODUCTION

Opening Address: 'Looking Ahead with Hindsight', Dr. E. EASTWOOD

Papers to be presented will include:

'The Relationship between Design and the Market in the Electronics Industry'
'Design Team Management'
'Design for Production'
'Value Engineering in the Electronics Industry'
'Quality Failure Cost Analysis'
'Unit Product Cost Control'
'The Implications of Applying Basic A.G.R.E.E. Principles to the Design and Production of Commercial Electronic Equipment'
'System Engineering for Reliability and Ease of Maintenance'
'Design for Maintainability with particular reference to Domestic Radio and Television Receivers'

Wednesday, 12th July (9.30–12: 2–5)

Session II: ECONOMICAL PRODUCTION

Papers to be presented will include:

'Handing over the Design to the Works'
'Batch and Flow Production'
'The Use of Computers in Production and Management'

A Symposium of short papers on 'Modern Materials, Components and Techniques' including

'Modern Electronic Materials'
'Integrated Circuit Packages'
'Plastics'
'Resin Finishes'
'Electrophoretic Finishes'

Thursday, 13th July (9.30–12 noon: 2–5.30)

Session III: FUTURE DEVELOPMENTS

A Symposium of papers on 'Automatic Assembling and Testing' including:

'Automatic Assembly on Printed Wiring Boards'
'Mechanical Assembling'
'Automatic Testing of Electronic Equipment for Aircraft'
'The Application of Automation to Assembly and Testing in Computer Manufacture'

Other papers presented will include:

'The Changing Pattern of Equipment Design'
'The Impact of Integrated Electronic Circuits on the Company'
'Some Applications of Computers in Electronics Design'
'Computer Techniques in the Drawing Office'
'New Vistas in Design'

Closing Address—Mr. J. LANGHAM THOMPSON

U.D.C. 621.372.832.8

A Simplified Theory of the Three-port Junction Ferrite Circulator

By

J. HELSZAJN,
M.S.E.E., C.Eng., A.M.I.E.R.E.†

Summary: A simplified theory of the three-port junction circulator is given. This theory is applicable to all three-port junctions which rely on the splitting of a pair of degenerate modes and gives the two circulation conditions of the three-port junction circulator in terms of the elements of the equivalent shunt resonator. This is done by first developing the degenerate normal-mode description of the resonator and then modifying it to allow for the known splitting of the resonant frequencies. The loaded-Q of the junction is also given. All the results obtained are consistent with the classical junction behaviour.

1. Introduction

The most general form of the junction circulator[1, 2] consists of a symmetrical distribution of ferrite material at the junction of three transmission lines as shown in Fig. 1. The ferrite material is magnetized perpendicularly to the plane of the disks by a static magnetic field. In this device, power entering port 1 emerges from port 2, and so on, in a cyclic manner. Important contributions to the understanding of the symmetrical junction circulator were made by Auld[3] and Bosma.[4] Auld considered the theory in terms of the scattering matrix of the device. An important property of the device is that a perfect circulator is obtained when the junction is matched. For a three-port junction this requires two independent variables. Bosma analysed the junction under certain simplifying conditions. An essential feature of the circulator adjustment involves a standing wave of the electric field pattern within the disks due to the interference of a pair of field patterns rotating in opposite directions. When the junction is unmagnetized, the resonant frequencies of the two field patterns are the same. When the junction is magnetized, the degeneracy is removed, and the standing wave pattern within the disks is rotated. One circulation condition is obtained by operating between the two split frequencies. The second circulation condition is met by adjusting the splitting until the standing wave pattern is rotated through 30 deg.[5, 6] From symmetry, port 3 is then situated at a null of the standing wave pattern and is therefore isolated. The junction then behaves as a transmission line resonant cavity between ports 1 and 2.

It is the purpose of this paper to present a new simplified theory of the junction circulator which applies to all three-port junctions which rely on the splitting of a pair of degenerate modes. This will be done by first developing the degenerate normal mode description of the equivalent resonator network and then modifying it to allow for the known splitting of the two modes. The two circulation conditions are obtained by matching the admittance of the split resonator at the plane of the junction.

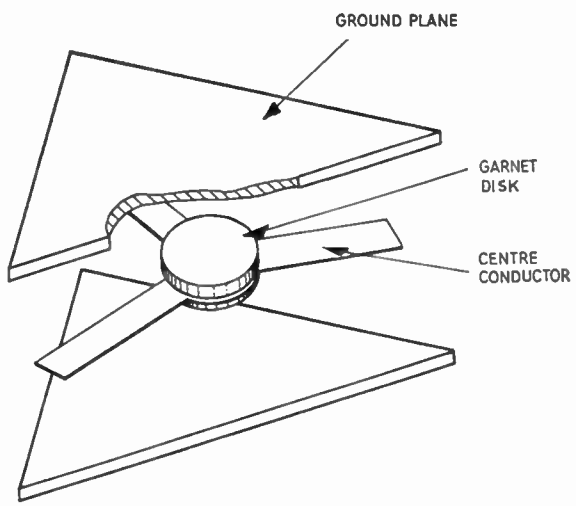


Fig. 1. Three-port ferrite junction circulator.

2. Characterization of the Junction Circulator

The classical junction has been represented by a parallel shunt resonant circuit.[4–8] The notion of the loaded-Q of a junction circulator has also recently been introduced in connection with this equivalent circuit. The bandwidth is defined once the loaded-Q is known.

It is well known that any junction geometry can be experimentally adjusted for parallel resonance at a pair of reference terminals taken at the edge of the

† Formerly with Microwave Associates Inc., Burlington, Massachusetts, U.S.A.; now at the Department of Electrical and Electronic Engineering, University of Leeds.

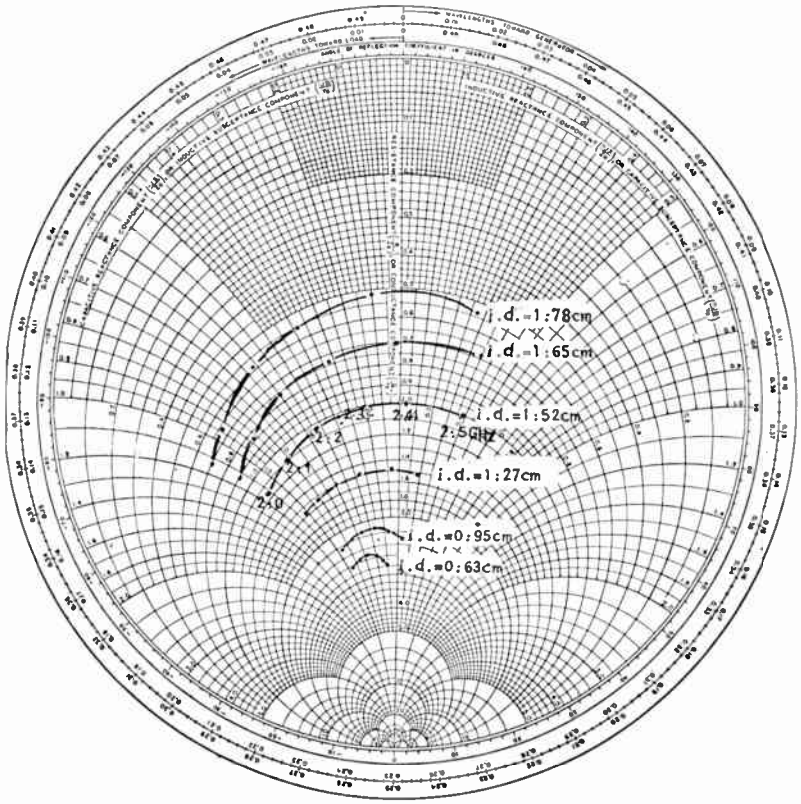


Fig. 2. Admittance of ring circulator for fixed outer diameter (2·35 cm). (After J. Helszajn[10].)

junction. The junction can then always be characterized by a parallel resonant circuit from the admittance function.

A useful formula for the loaded-Q, of a resonant circuit in terms of the admittance function is[9]:

$$Q_L = \frac{\omega_0}{2G}\frac{\partial B}{\partial \omega}\bigg|_{\omega_0} \qquad \text{.....(1)}$$

where B is the shunt susceptance and G is the shunt conductance. This representation is particularly useful when working with a Smith chart. It is worthwhile to notice that Q_L is determined by the susceptance slope parameter and the shunt conductance of the circuit.

The equivalent shunt elements can also be related to the susceptance slope parameter by:

$$b = \frac{\omega_0}{2}\frac{\partial B}{\partial \omega}\bigg|_{\omega_0} = \omega_0 C = \frac{1}{\omega_0 L} \qquad \text{.....(2)}$$

It has been noted experimentally in a number of devices that for a given geometry the susceptance slope remains essentially constant over fairly large variations in shunt conductance. The shunt conductance can be adjusted by varying the ferrite parameters or the magnetic field. The admittance plot of a ring circulator[10] is shown in Fig. 2. The equivalent circuit is shown in Fig. 3.

3. Normal Mode Description of Shunt Resonator

In this section, we will develop the normal mode description of a shunt resonator. According to Louisell, we can do this by taking a linear combination of the voltage current relations of the LC circuit that will decouple them.[11] From Fig. 4 we have

$$\frac{di}{dt} = \frac{-1}{L}\,V \qquad \text{.....(3)}$$

$$\frac{dV}{dt} = \frac{1}{C}\,i \qquad \text{.....(4)}$$

If we multiply eqn. (4) by an arbitrary constant Y and add it to eqn. (3), we have

$$\frac{d}{dt}(i + YV) = \frac{Y}{C}\left(i - \frac{C}{LY}\,V\right) \qquad \text{.....(5)}$$

This equation is satisfied by making $Y = \pm j\sqrt{C/L}$.

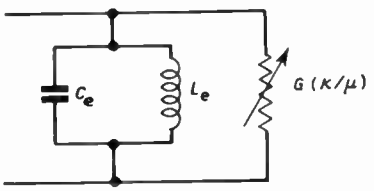


Fig. 3. Equivalent circuit of junction circulator at plane of ferrite disks.

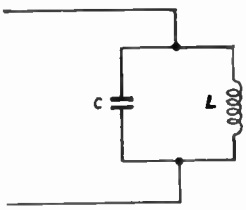


Fig. 4. Shunt resonator.

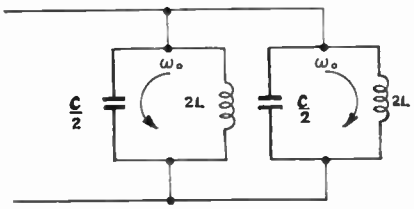


Fig. 5. Normal-mode form of shunt resonator.

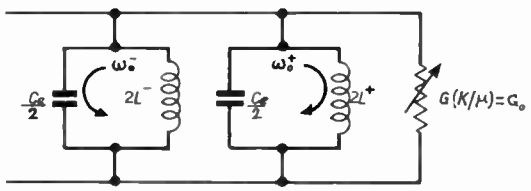


Fig. 6. Normal-mode form of junction circulator at plane of ferrite disks.

Equation (5) now becomes

$$\left(\frac{d}{dt} - j\omega\right)\alpha = 0 \qquad \ldots\ldots(6)$$

$$\left(\frac{d}{dt} + j\omega\right)\alpha^* = 0 \qquad \ldots\ldots(7)$$

where

$$\alpha = \sqrt{\tfrac{1}{2}}L(i + j\omega CV) \qquad \ldots\ldots(8)$$

and

$$\alpha^* = \sqrt{\tfrac{1}{2}}L(i - j\omega CV) \qquad \ldots\ldots(9)$$

The solutions of the normal mode equations are

$$\alpha(t) = \alpha(0)\, e^{j\omega t} \qquad \ldots\ldots(10)$$

$$\alpha^*(t) = \alpha^*(0)\, e^{-j\omega t} \qquad \ldots\ldots(11)$$

The sum of the normal modes gives the total energy in the system

$$E = \tfrac{1}{2}[CV^2(t) + Li^2(t)] = |\alpha(t)|^2 + |\alpha^*(t)|^2 \qquad \ldots\ldots(12)$$

From eqns. (10) and (11) we note that α and α^* are a pair of normal degenerate modes rotating in opposite directions. We also note that each of the normal modes is circularly polarized and has half of the total energy of the system. This result is satisfying because it phenomenologically describes the uncoupled disks of the classical junction. Figure 4 can now be redrawn in terms of the normal mode representation, as shown in Fig. 5.

4. Equivalent Circuit of Junction Circulator

When the junction is magnetized, the degeneracy between the normal modes is removed and the resonant frequencies of the normal modes split. We can represent this situation by replacing the shunt inductances in Fig. 5 by new inductances $L^{\pm}$. We will further write $L^{\pm} = L(\mu \pm K)$ because the normal modes are circularly polarized. We will find that this representation will give answers consistent with the known results. In the above μ and K are the diagonal and off-diagonal components of the Polder tensor permeability $[\mu]$.[12]

$$[\mu] = \begin{bmatrix} \mu & -jK & 0 \\ jK & \mu & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \ldots\ldots(13)$$

where

$$\mu = 1 - \frac{p\sigma}{1-\sigma^2} \qquad \ldots\ldots(14)$$

$$K = \frac{p}{1-\sigma^2} \qquad \ldots\ldots(15)$$

and

$$\sigma = \frac{|\gamma| H_i}{\omega} \qquad \ldots\ldots(16)$$

$$p = \frac{|\gamma| M_s}{\mu_0 \omega} \qquad \ldots\ldots(17)$$

In the above, $|\gamma|$ is $2{\cdot}21 \times 10^5$ rad/s/At/metre, M_s is the saturation magnetization of the material in weber/metre^2, H_i is the internal biasing magnetic field in At/metre within the ferrite material, μ_0 is the free-space permeability and ω is the microwave angular frequency in rad/s. The material is saturated when $\sigma > 0$.

The equivalent circuit of the circulator is now shown in Fig. 6, where the shunt conductance is shown to be a function of K/μ, the splitting factor.

5. Circulation Adjustment

The total admittance of the split modes at the plane of the ferrite disks can now be obtained from Fig. 6.

$$Y = Y^+ + Y^- = G(K/\mu) + \\ +j\left\{\left(\frac{\omega C_e}{2} - \frac{1}{2\omega L^+}\right) + \left(\frac{\omega C_e}{2} - \frac{1}{2\omega L^-}\right)\right\} \quad \text{.....(18)}$$

The two circulation conditions will now be obtained. It is well known that two independent variables are required to match a three-port junction circulator.

One of the two circulation conditions is satisfied by adjusting the operating frequency to lie between the two split resonant frequencies ω_0^+ and ω_0^-. This condition is met by equating the imaginary parts of Y to zero.

$$\left(\frac{\omega C_e}{2} - \frac{1}{2\omega L^+}\right) + \left(\frac{\omega C_e}{2} - \frac{1}{2\omega L^-}\right) = 0 \quad \text{.....(19)}$$

The circulation frequency is then given by

$$\omega_0^2 L_e C_e = 1 \quad \text{.....(20)}$$

where

$$L_e = \frac{2L^+L^-}{L^+ + L^-} \quad \text{.....(21)}$$

If we write $L^\pm = L(\mu \pm K)$ we have

$$L_e = \left(\frac{\mu^2 - K^2}{\mu}\right)L = \mu_{eff} L \quad \text{.....(22)}$$

We also have $C_e = \varepsilon_r C$, where ε_r is the relative dielectric constant of the ferrite material.

The variation of μ_{eff} as a function of p and σ is shown in Fig. 7. The junction can in general be biased either above or below ferrimagnetic resonance.

We further note that

$$\omega_0 = \frac{\omega_0^- + \omega_0^+}{2} \quad \text{.....(23)}$$

Hence the operating frequency lies between the two split resonant frequencies.

The second circulation condition requires that the standing wave pattern be rotated until port 3 is decoupled. From the symmetry, this condition is met by making the phase angles between the split admittances $\pm 30°$. This gives

$$\frac{\omega_0 C_e}{2} - \frac{1}{2\omega_0 L^-} = -\frac{G(K/\mu)}{2\sqrt{3}} \quad \text{.....(24)}$$

$$\frac{\omega_0 C_e}{2} - \frac{1}{2\omega_0 L^+} = +\frac{G(K/\mu)}{2\sqrt{3}} \quad \text{.....(25)}$$

Subtracting eqn. (24) from eqn. (25) and rearranging gives

$$\frac{1}{\omega_0}\left(\frac{L^+ - L^-}{2L^-L^+}\right) = \frac{G(K/\mu)}{\sqrt{3}} \quad \text{.....(26)}$$

If we write $L^\pm = L(\mu \pm K)$ eqn. (26) becomes

$$G(K/\mu) = \sqrt{3}\sqrt{\frac{C_e K}{L_e \mu}} \quad \text{.....(27)}$$

The second circulation condition relates the junction shunt conductance to the intrinsic admittance of the junction and K/μ. The variation of K/μ as a function of p and σ is shown in Fig. 8. A perfect circulator is obtained by adjusting the splitting to give $G(K/\mu) = G_0$, where G_0 is the conductance of the external transmission line.

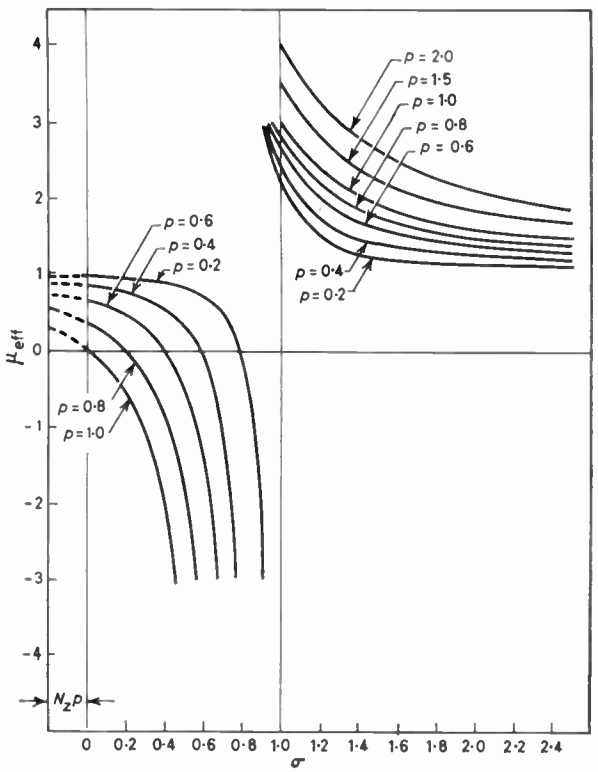


Fig. 7. Variation of μ_{eff} as a function of p and σ. (After Fay and Comstock.[5])

The splitting between the two resonant frequencies is also related to K/μ by

$$\frac{\omega_0^- - \omega_0^+}{\omega_0} = \frac{K}{\mu} \quad \text{.....(28)}$$

We can compare the above relation with the more exact one which is for the stripline circulator given by

$$\frac{\omega_0^- - \omega_0^+}{\omega_0} = 0{\cdot}81\,\frac{K}{\mu} \quad \text{.....(29)}$$

Hence the assumed form for $L^\pm$ gives results in close agreement with the more exact theory.

6. Bandwidth of Direct Coupled Circulator

The bandwidth of the junction can be obtained by forming the loaded-Q from the admittance function using the definition for Q given previously.

The susceptance slope parameter for the admittance function given in eqn. (18) is

$$b = \frac{\omega_0}{2}\left.\frac{dB}{d\omega}\right|_{\omega_0} = \sqrt{\frac{C_e}{L_e}} \qquad \text{.....(30)}$$

We note from eqn. (30) that the susceptance slope parameter is essentially independent of the splitting for K/μ small, which is the usual case. This is in agreement with Fig. 2.

The loaded-Q is in general given by

$$Q_L = \frac{1}{G(K/\mu)}\sqrt{\frac{C_e}{L_e}} \qquad \text{.....(31)}$$

When the circulator is matched directly to the connector admittance we have

$$G(K/\mu) = G_0$$

hence

$$Q_L = \frac{1}{G_0}\sqrt{\frac{C_e}{L_e}} \qquad \text{.....(32)}$$

The loaded-Q can also be given in terms of the splitting with the help of eqn. (27). This gives

$$Q_L = \frac{1}{\sqrt{3}}\frac{\mu}{K} \qquad \text{.....(33)}$$

This last equation has previously been given for the waveguide junction.

Equation (33) is useful in that it allows K/μ to be determined from measurement of Q_L.

The maximum v.s.w.r. and the bandwidth can also be related to the loaded-Q by

$$Q_L = \frac{r-1}{4\delta\sqrt{r}} \qquad \text{.....(34)}$$

where

$$\delta = \frac{\omega-\omega_0}{\omega_0} \qquad \text{.....(35)}$$

The bandwidth of a directly coupled junction is completely determined by the splitting of the two normal modes. The value of the splitting corresponding to a directly-coupled circulator can be obtained from the circulation condition given by eqn. (27). The maximum splitting possible can be determined from Fig. 8.

7. Quarter-wave Coupled Junction

The loaded-Q of a directly-coupled junction is usually relatively large, typically about 3·1 (see, for instance, Fig. 2). Equation (33) indicates that this last number corresponds to a splitting of about 0·2. On the other hand, Fig. 8 indicates that the splitting for a saturated material is about 0·65 for a typical value of p equal to 0·65. The loaded-Q can therefore be reduced by increasing the splitting. Hence, directly-coupled junctions are usually restricted to large average power devices where the splitting is necessarily low as a consequence of the techniques used. Low average power devices are normally quarter-wave coupled. For instance, the loaded-Q associated with the bottom curve of Fig. 2 is about 1·3 and the corresponding splitting is approximately 0·52. However, some kind of admittance transfer is now needed because the conductance at resonance is no longer matched to the external transmission line. In practice quarter-wave transformers are used to do the matching as their use results in further bandwidth improvement. An example of a quarter-wave coupled junction is described by the author in another paper.[10]

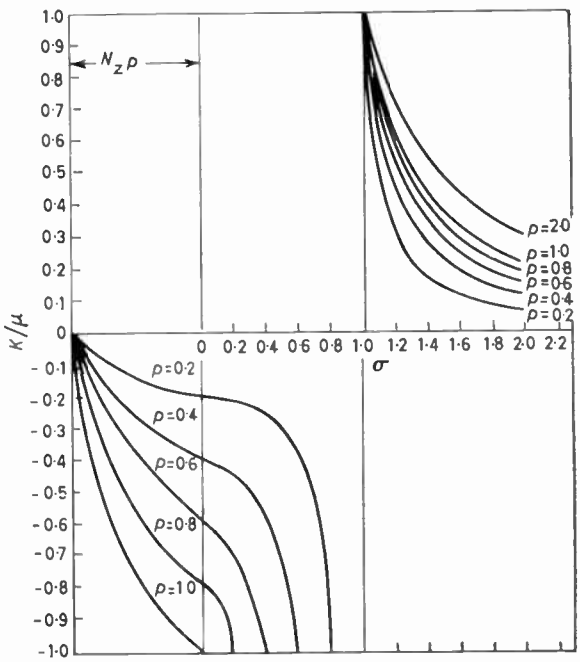


Fig. 8. Variation of K/μ as a function of p and σ (after Fay and Comstock.[5])

8. Conclusions

A general characterization of the three-port junction circulator has been described which allows the two circulation conditions of the junction to be derived in the most general terms. The results obtained apply to all three-port junctions which rely on the splitting of a degenerate pair of normal modes. The approach is particularly satisfying because it gives the two circulation conditions in easily understood terms. All the results obtained are consistent with the classical theory.

9. Acknowledgment

The author wishes to thank Microwave Associates Inc. for permission to publish this paper.

10. References

1. H. N. Chait and T. R. Curry, 'A new type Y-circulator', *J. Appl. Phys.*, Suppl., **30**, pp. 152S–153S, April 1959.
2. U. Milano, J. H. Saunders and L. Davis, Jr., 'A Y-junction stripline circulator', *Trans. Inst. Radio Engrs on Microwave Theory and Techniques*, **MTT-8**, pp. 346–51, May 1960.
3. B. A. Auld, 'The synthesis of symmetrical waveguide circulators', *Trans. I.R.E.*, **MTT-7**, pp. 238–46, April 1959.
4. H. Bosma, 'On stripline Y-circulation at u.h.f.', *Trans. Inst. Elect. Electronics Engrs*, **MTT-12**, pp. 61–72, January 1964.
5. C. E. Fay and R. L. Comstock, 'Operation of the ferrite junction circulator', *Trans. I.E.E.E.*, **MTT-13**, pp. 15–27, January 1965.
6. H. J. Butterweck, 'Der Y-zirkulator', *Arch. Elekt. Übertragung*, **17**, pp. 163–76, March 1963.
7. C. E. Barnes, 'Integrated circulator design for parametric amplifier design', I.E.E.E. G-MTT International Symposium Technical Program Digest, pp. 170–5, 1964.
8. J. Helszajn, 'A high-speed TEM junction modulator using a wire loop', *The Radio and Electronic Engineer*, (To be published.)
9. C. G. Montgomery, R. H. Dicke and E. M. Purcell (Ed.), 'Principles of Microwave Circuits'. (McGraw-Hill, New York, 1948.)
10. J. Helszajn, 'A ring stripline junction circulator', *The Radio and Electronic Engineer*, **32**, pp. 55–60, July 1966.
11. W. H. Louisell, 'Coupled Mode and Parametric Electronics', pp. 4–7. (John Wiley, New York, 1960.)
12. D. Polder, 'On the theory of ferrimagnetic resonance', *Phil. Mag.*, **40**, pp. 99–115, February 1949.

Manuscript first received by the Institution on 30th August 1966 and in final form on 23rd January 1967. (Paper No. 1115.)

U.D.C.621.374: 639.2

A Digital Echo-counting System for use in Fisheries Research

By

B. R. CARPENTER
(*Graduate*)†

Reprinted from the Proceedings of the I.E.R.E. Conference on 'Electronic Engineering in Oceanography' held in Southampton on 12th–15th September 1966.

Summary: The echo-counting system to be described was designed for use in conjunction with a high-frequency (100 kHz) echo-sounder and was primarily intended for use in surveying pelagic fish stocks. The counter incorporates a variable gating system, synchronized to the transmission pulse repetition frequency, enabling signals occurring between any two selected depth limits to be processed.

Time-swept gain amplifiers are utilized to compensate for fall-off of signal strength with depth, although the counter differs from previously published systems in as much as echo *duration* rather than echo *amplitude* is the parameter under consideration.

The processed signal information is displayed as an in-line digital read-out from a conventional scaler. A permanent record is obtained from a printer coupled to the scaler.

1. Introduction

Echo survey has long been accepted as a fairly reliable method of estimating fish abundance based on the number of echo signals received per unit distance steamed.[1] In 1959, Richardson *et al.* indicated that the abundance of Barents Sea cod could be determined by the visual counting of echo signals displayed on a cathode-ray tube which was coupled to an echo-sounder.[2] Two years later Mitson and Wood described a three-channel automatic counter which performed a similar task at differing amplitude thresholds.[3] In each case the principle employed was a summation of total echo-signal amplitude multiplied by depth squared during a trawling period. A relationship was derived between fish caught and total echo signal received.

More recently a fish-echo integrator system has been described which was developed for studies on O-group fish.[4]

Attempts have recently been made to process, independently, signals obtained from single targets as distinct from shoals, but again tedious calculations have been necessary.[5]

The present paper describes a system, developed at the Lowestoft Fisheries Laboratory, which utilizes pulse duration as the parameter to be measured and, above a given amplitude threshold, is independent of amplitude fluctuations. Development of the system is continuing and an experimental shoal or single target recognition circuit is outlined.

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

2. Factors Influencing Choice of Echo-sounder Frequency, Pulse-width and Target Amplitudes

Many of the early estimates of the relative abundance of fish stocks were made by the visual addition of target echoes received and recorded on a paper chart. Great difficulty was experienced in estimating numbers of fish present whenever dense shoals occurred so that target echoes merged to give a continuous black trace. The problem was aggravated by the fact that the echo-sounders being used were commercial instruments of relatively low frequency, long pulse-width and wide beam-angle—being intended as much for navigational use as for fish finding. The use of sonic or low ultrasonic frequencies, coupled with the impracticability of producing transducers of high sensitivity and low Q (to limit pulse-width), inevitably resulted in poor resolution in both horizontal and vertical planes.

The depth anomaly, inherent in all echo-sounders, was also magnified by the large beam-width. Any advantage gained from the low attenuation of a low-frequency acoustic wave by the sea-water (resulting in a greater depth range for a given output power) was almost immediately nullified by the low amplitude echo returns from fish targets.[6–8] For example, the amplitude of the echo returned from a standard target at a given depth (an air-filled sphere of radius 2 metres) is, theoretically, shown to be greater than that from a 50 cm cod, at the same depth, by +58 dB at 1 kHz, +38 dB at 10 kHz, and +28 dB at 30 kHz.[6]

Obviously, so far, it would appear that a high-frequency sounder would be superior in every respect

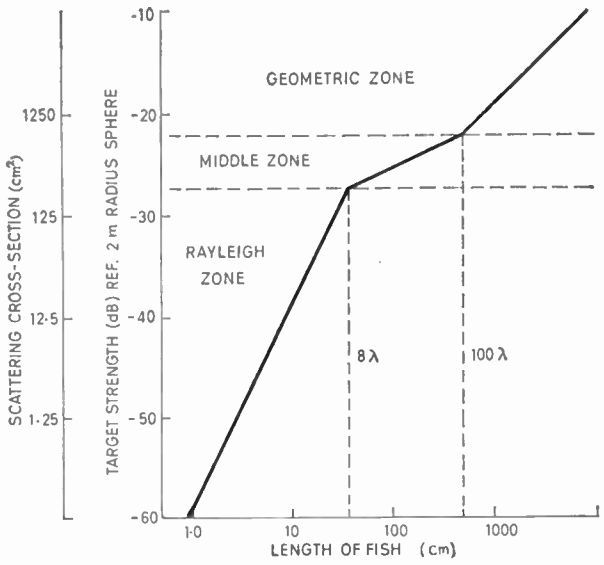


Fig. 1. The relation between scattering cross-section, target strength and the length of fish measured at 30 kHz (after Cushing *et al.*[6]).

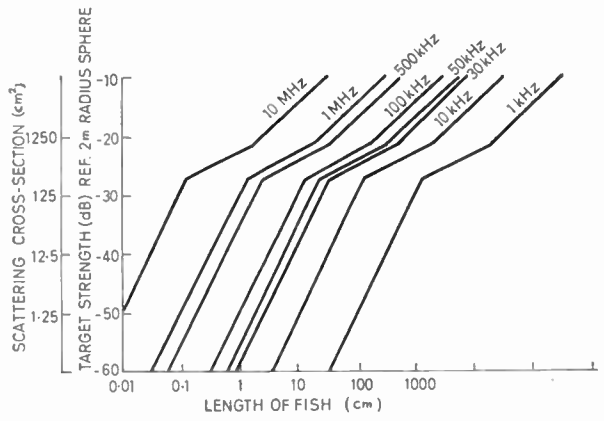


Fig. 2. The relation between scattering cross-section, target strength and the length of fish at different frequencies (after Cushing *et al.*[6]).

for fish echo-survey, but there remains one important point to be considered.

Haslett, in his adaptation of the back-scattering theory of electromagnetic waves to acoustic waves, has shown that there are three distinct regions when considering target-strength signal returns from fish.[9-12] In the first region, where the target is small compared with the incident wavelength of the sound, the back-scattering follows Rayleigh's law (back-scattering cross-sectional area proportional to $1/\lambda^4$). This region extends to the point where the fish-length approaches 8λ; within it predictable results for fish-echo amplitudes are obtained and also in the region where the fish-length exceeds 100λ, at which point the target can be considered plane and the acoustic reflection follows a geometric law (scattering cross-section proportional to $1/\lambda^2$). The middle region, however, for fish-lengths between 8 and 100λ, has been shown to yield target strengths of a complex nature differing from maxima to minima by a factor of 10 (Fig. 1).[6]

Thus, the low-frequency echo-sounder, whilst providing operation in the Rayleigh zone, does not have the necessary resolution for accurate fish-counting and an alternative must be sought. Referring to Fig. 2, it may be readily ascertained that if the middle zone, with its unpredictable target amplitudes, is to be avoided, a sounder frequency of the order of 2 MHz or more is required. This would give geometric zone scattering of all fish from 10 cm upwards in length. Unfortunately, the attenuation by the sea of acoustic waves at this frequency is very great and so an enormous output power would be required from the transmitter. This in turn would necessitate a large transducer assembly, and as a piece of sea-going equipment suitable for deep-water working the 2 MHz echo-sounder is no more than a theoretical ideal. Adoption of a lower frequency of between, say, 15 kHz and 2 MHz inevitably means that many of the fish encountered will have dimensions which place them in the middle zone, and therefore *amplitude* must immediately be rejected as a suitable parameter for summation in fish-counting systems.[14]

Of the remaining parameters (phase or pulse width) pulse width was adopted as being a suitable basis for development.

3. The Fish-echo Counting System

3.1 *The Echo-sounder and Paper Recorder*

Since target resolution was to be the main criterion, it was evident that an acoustic wavelength of not greater than 2 cm should be employed. In practice, the echo-sounder used was a 100 kHz machine built at Lowestoft by R. B. Mitson in 1962. Taking the velocity of sound in sea-water to be 15×10^4 cm/s, it follows that the acoustic wavelength was 1·5 cm. Reference to Fig. 2 shows that all fish-lengths between 10 and 180 cm lie in the middle zone when using a 100 kHz sounder.

Three values for transmission pulse widths, namely 100, 200 or 300 μs, were chosen. Normally the shortest pulse is used, but in a choppy sea where aeration becomes a problem the longer pulses can be used with some advantage. In sea-water, the 100 μs pulse (i.e. ten complete cycles at 100 kHz) is contained in a vertical column of 15 cm. For an infinitely narrow beam-width a theoretical vertical resolution of slightly better than 3 in is obtained In practice, with finite beam-width of the order of 10 degrees, a vertical resolution of better than the pulse column-length is maintained to a depth of slightly greater than 20 fathoms. Typical commercial echo-sounders operate

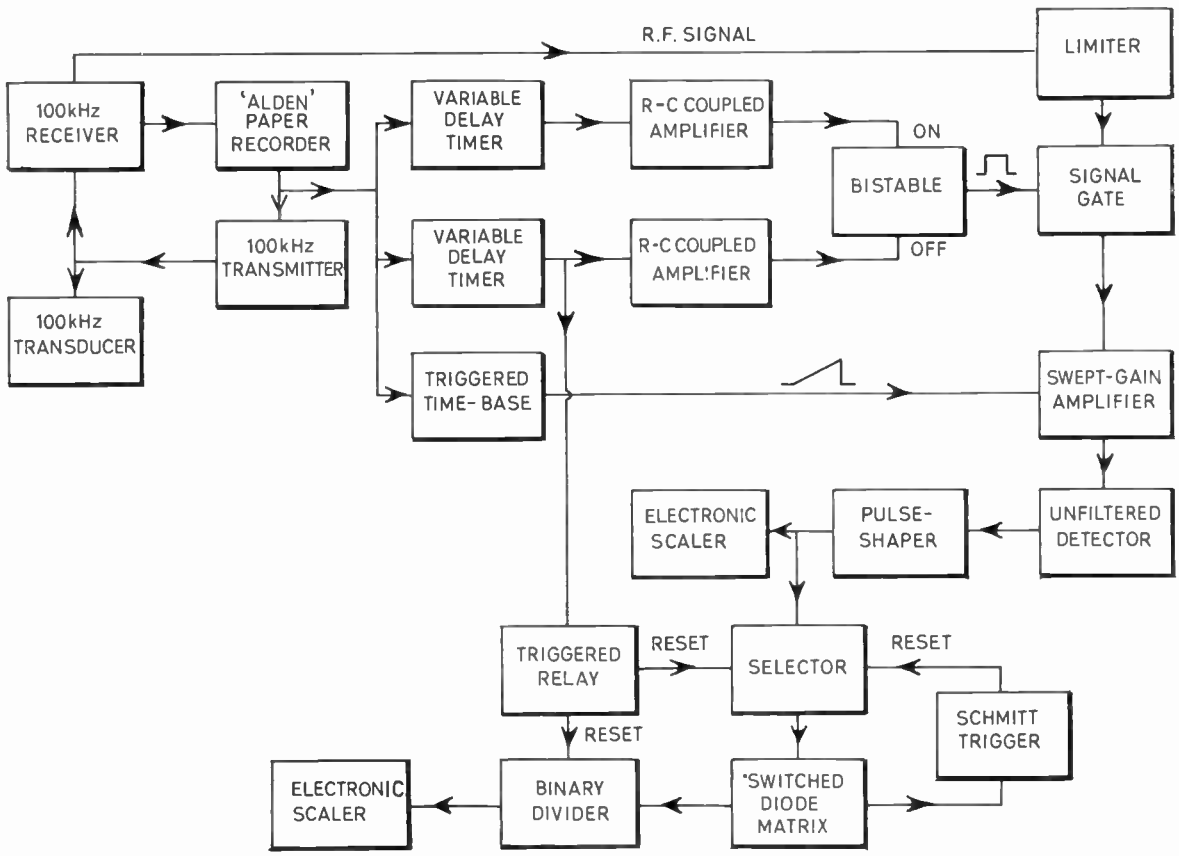


Fig. 3. Block diagram of the complete digital fish-echo counting system.

at only one-third of this frequency, with pulse widths of the order of $\frac{1}{2}$, 1 and 2 ms and approximately twice the beam-width; thus it is evident that the echo-sounder developed and built at Lowestoft has a vertical resolution of at least one order of magnitude better than any fish detection equipment fitted on any British commercial trawler.

It was felt desirable that the permanent paper record of echo surveys should, itself, show far more detail than had been obtained in the past from a 30 kHz commercial echo-sounder. The commercial recorder utilized a 6-in wide paper to display, at maximum scale expansion, a 35-fathom slice of the water column. Some 'optical distortion' was apparent on the record obtained, due to the fact that the echo trace was a series of consecutive arcs, of radius 4 in, rather than a parallel straight-line recording.

The recorder chosen for use with the 100 kHz equipment was a modified Alden 319DA paper recorder employing a 19-in wide paper. The recorder uses a helical writing wire giving a straight line 'sweep' and is similar to that used for facsimile reception.

A synchronous motor was used to drive the recorder through a continuously variable gear-box—the fitting of this gear-box being the principal modification to the existing instrument. This facility enabled the operator to make full use of the 19-in paper for any sea-bed depth between 15 and 180 fathoms. This was accomplished by increasing the writing speed and the transmission rate, since the recorder was employed as the master timing element—it being far easier to synchronize the transmitter to the recorder than vice versa. Transmitter triggering was effected by means of a perforated timing-disk fitted to the extended shaft of the writing helix of the recorder. A Lilliput Edison screw lamp and phototransistor provided the actual trigger pulse, which was shaped by a Schmitt trigger circuit. The differentiated pulse, now only 2 μs wide, was used to fire the transmitter.

3.2 *The Gating and Signal Processing Circuits*

A functional diagram of the entire fish-echo counting equipment appears in Fig. 3.

Simultaneous with the firing of the echo-sounding transmitter, a pulse was applied to two variable delay circuits. The time-delayed pulses obtained from the two timers determined the counting range of the equipment: the pulse from timer 1 started the count whilst that from timer 2 stopped it. The variable

elements in the timers (potentiometers) were calibrated directly in fathoms. The two pulses, after amplification, were fed to the separate inputs of a bistable. The bistable output could be considered as a positive-going square wave of the same pulse rate frequency as the echo-sounding transmitter and with a variable mark to space ratio.

The bistable output was applied to a gate circuit which passed signals only during the positive excursion of the square wave. The entire water-column signal from transmission pulse (i.e. surface) to sea-bed echo was fed to the gate, but in order to minimize undesirable signal break-through, amplitude limiting of transmission pulse and sea-bed echo was applied. Typically, the amplitudes of the transmission pulse and sea-bed echoes are greater than fish-echo signals by three orders, or more. The gain of the system was adjusted so that there was never any danger of fish-echoes reaching clipping level.

The gated fish-echo signals were then fed to a two-stage swept-gain amplifier, each stage having a linear variation of gain with respect to time, measured from the start of each transmission.

The cascaded stages exhibited a gain characteristic which varied as the square of the depth. Thus compensation of signal amplitude fall-off with increased depth was automatic. Gain variation with respect to time was achieved by means of a sawtooth bias voltage derived from a time-base triggered at the instant of transmission. Preset sweep widths accommodated all depth ranges likely to be encountered.

On first inspection, it appeared that it was unnecessary to apply any form of amplitude correction since it will be recalled that amplitude was regarded as an unsuitable parameter for measurement. There were, however, two reasons for applying this correction. Firstly, alternative methods of signal processing were envisaged, including integration, where correction is essential; secondly, it is important to realize that as depth increases and the signal levels fall, so also does the mean noise level. Evidently, increased gain is desirable to bring the noise level back to its shallow-water value, at which level the risk of missing deep-water signal echoes is minimized, since later in the system an arbitrary amplitude threshold must be chosen. For convenience this might typically be two or three times the mean noise level and with the amplitude correction it does mean that there is a constant value of minimum recognizable signal, irrespective of depth.

After passing through the swept-gain amplifiers, the fish-echo signals were fed to a detector stage which removed the negative half-cycles but *no subsequent high frequency filtering was applied.* Thus, the output from the detector stage consisted of a series of positive-going, variable amplitude 100 kHz half-sinewaves, the number of which gave an indirect measurement of signal-echo duration.

The half-cycles pulse train was fed to a pulse shaper (Schmitt trigger plus differentiation) to yield a series of constant-amplitude, constant-width (2 μs) pulses which could be readily totalized on an electronic scaler.

A printer was coupled to the scaler, and a crystal chronometer on board the research vessel gave out print commands each minute. Resetting of the electronic scaler was related to the distance steamed by the use of a pair of contacts on the ship's log.

3.3 *The Experimental Fish Shoal Recognition Circuit*

It is frequently desirable to know whether the total fish-echo count is built up from predominantly single targets or discrete fish shoals, and with this problem in mind the experimental shoal recognition circuit was devised.

The pulse train fed to the electronic scaler was also fed to four cascaded binary stages, and the eight collectors of the transistors used in the binary circuits were compared for d.c. level in a diode matrix. This matrix could be switched to give an output for every 6 or 7 or 8 or 9 or 10 . . . (up to 16) input pulses to the selector. The d.c. state of the diode matrix switched a Schmitt trigger stage to give an output pulse which was used both to reset the selector and as a 'carry-forward' signal to yet another bistable. The output of this bistable was fed to a second, extremely simple, electronic scaler. This meant that, in practice, a count of one would be registered for every 12 or 14 or 16 . . . 32 pulses received at the input, depending on the position of the diode matrix switch. Thus the operator could decide how many pulses were to constitute a shoal, and adjust the selector switch accordingly. A triggered relay stage, initiated by the 'stop count' pulse from timer 2, reset both the bistable and the selector stage before each subsequent transmission.

4. Analysis of Results

The reason for not filtering the detector output only becomes apparent if one considers the case of two fish swimming sufficiently close together as to be unresolved by the echo-sounder. In the case of a small single fish target (where the vertical dimension of the swim-bladder can be considered as being much less than λ) the echo returned is similar in form to the transmitted pulse, i.e. ten cycles of a 100 kHz sine-wave. Thus, after detection, the ten half-cycles will eventually register as a count of ten on the scaler. (In practice, the first and last half-cycle are frequently too low in amplitude to trigger the Schmitt circuit and a count of eight per target is more usual.) In the case of two overlapping targets the echo returned will be complex in form, exhibiting phase and amplitude changes,[13] but will still show between ten and twenty

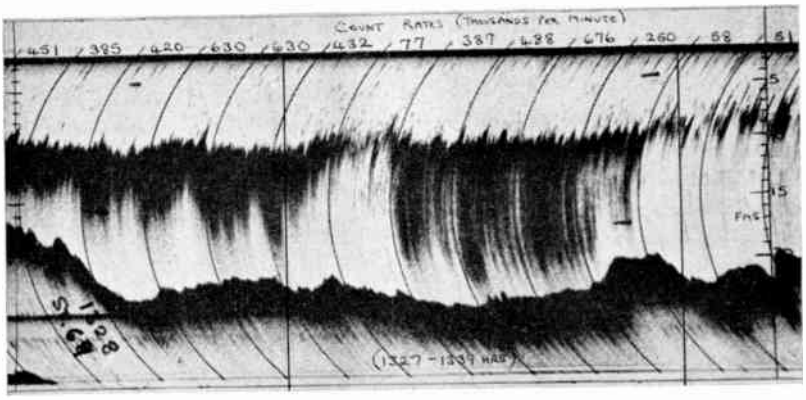


Fig. 4. 30 kHz echogram of sprat shoals encountered in the Wash, November 1965. (One-third reduction of original record.)

cycles and will register as such on the scaler. Obviously if high frequency filtering had been applied at the detector the voltage envelope fed to the Schmitt stage would have appeared as a pulse of 100 μs duration for the single target, or between 100 and 200 μs for the overlapping pair. In either case, only *one* output pulse would have resulted from the Schmitt trigger. The author makes no apology for results obtained in this manner, being aware that they are not absolutely valid, but undoubtedly a much closer approximation to the true result can be obtained by the 'half-cycle count' method. Trials have so far indicated that the system is superior to methods previously exploited, and the equipment has proved to be of particular value when working in sprat or herring fisheries.

To date, four cruises have been made using the digital system—the first in November 1964 using an early prototype counter with no amplitude/depth correction. Results obtained from the two most recent echo surveys and shown in Figs. 4 and 5, (Wash sprat survey, October/November 1965) have been quite encouraging, yielding theoretical catch rates not far removed from those actually achieved by nearby fishing vessels.

Admittedly in this case the vertical orientation of the fish was ideal, practically the entire distribution

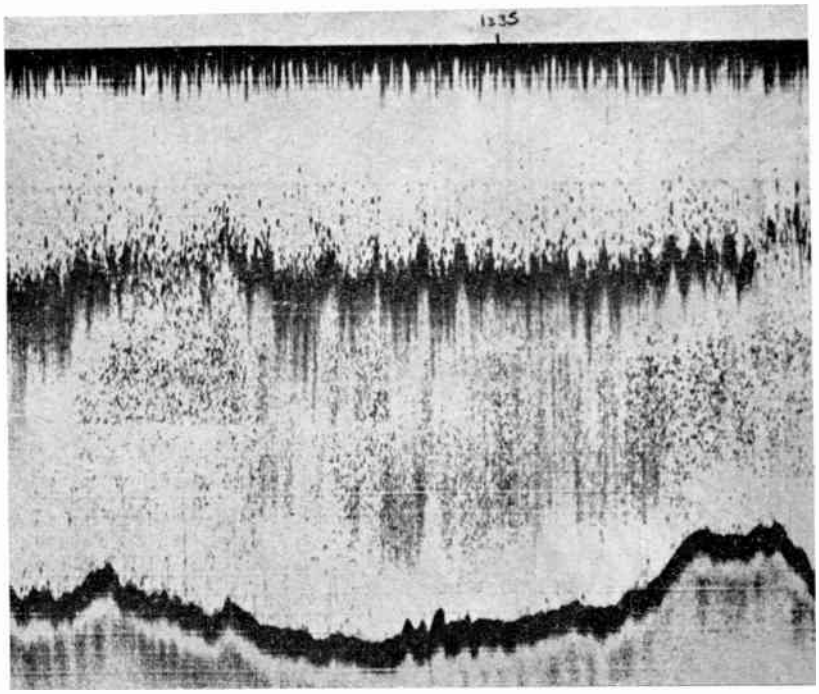


Fig. 5. 100 kHz echogram (same photographic reduction) taken at the same time as the record in Fig. 4. The portion shown lies between the vertical lines on the 30 kHz echogram.

being confined to a three-fathom slice. Upper and lower gating limits were set at 2 and 7 fathoms respectively and, in conjunction with transducer beam width, the projected area of the insonified frustrum was calculated. This area represented only a fraction of the cross-sectional area of the fishing net aperture and so a multiplying factor was derived. Likewise the survey speed of the research vessel (10 knots) was approximately five times the speed of the fishing vessels and thus a five times reduction factor was applied. The results were then corrected for the number of times each target would be recorded. In the calculation it was assumed that the fish were stationary and that current effects were minimal. The number of times each fish target was recorded was, of course, dependent on depth, transducer beam-width and directivity, transmission rate and survey speed. Another correction was then applied to eliminate the cyclical error inherent in the half-sinewave count, the dividing factor being that obtained from tank tests or by allowing the research vessel to drift over individual targets and taking camera records of oscillograph traces.

The final corrected figures were the theoretical fish 'encounters' per minute at the net aperture, i.e. the potential catch rate, assuming that the net retains all fish caught. Corrected figures were compared with the catch rate statistics obtained from the fishing vessels.

In spite of the remarkable agreement between predicted and actual catches, the need for echo-survey continues. Fishing vessels tend to work only in the areas of greatest fish concentration, whilst the research vessel needs to survey the entire fishery in order to establish not only the magnitude, but also the distribution of the fish stock.

Relative abundance estimates are still being made by the use of a trawl, but for absolute determinations the digital system described represents a real advance in the technique of acoustic fish-counting and, as far as is known to the author, is the only system so far described in which an attempt has been made to separate multiple fish echoes from individuals.

5. Acknowledgments

The author wishes to thank Mr. R. B. Mitson for the initial suggestion of cyclical counting within the pulse period and for other helpful comments during the preparation of the paper. Thanks are also due to Dr. D. H. Cushing for useful discussions on earlier echo-sounding/-counting techniques and to Dr. P. O. Johnson for information concerning the analysis of the results.

The permission of the Director, Fisheries Laboratory, Ministry of Agriculture, Fisheries and Food, Lowestoft, to publish this paper is acknowledged. Crown copyright reserved.

6. References

1. D. H. Cushing, 'Echo-surveys of fish', *J. Cons. Perm. Intl Explor. Mer*, **18**, pp. 45–60, 1951.
2. I. D. Richardson, D. H. Cushing, F. R. Harden-Jones, R. J. H. Beverton and R. W. Blacker, 'Echo sounding experiments in the Barents Sea', *Fishery Investigations*, Ser. 2, **22**, No. 9, 1959.
3. R. B. Mitson and R. J. Wood, 'An automatic method of counting fish echoes', *J. Cons. Perm. Intl Explor. Mer*, **26**, pp. 281–91, 1961.
4. O. Dragesund, S. Olsen and I. Hoff, 'On the possibility of estimating year class strength by measuring echo abundance of O-group fish', *Fisk Dir. Skr. Ser. Havunders.*, **13**, No. 8, pp. 47–75, Appendix 1, 1964.
5. M. D. Truskanov and M. N. Scherbino, 'Methods of Direct Calculation of Fish Concentrations by Means of Hydroacoustic Apparatus', F.A.O. Etap Report No. 1937, II, pp. 101–13, 1965.
6. D. H. Cushing, F. R. Harden-Jones, R. B. Mitson, G. H. Ellis and G. Pearce, 'Measurements of the target strength of fish', *The Radio and Electronic Engineer*, **25**, pp. 299–303, April 1963.
7. D. H. Cushing and I. D. Richardson, 'Echo sounding experiments on fish', *Fishery Investigations*, Ser. 2, **18**, No. 4, 1955.
8. D. H. Cushing and I. D. Richardson, 'A triple frequency echo sounder', *Fishery Investigations*, Ser. 2, **20**, No. 1, 1955.
9. R. W. G. Haslett, 'The Application of a Scale-model Technique to the Analysis of the Principles Underlying the Formation of Ultrasonic Echoes, with Special Reference to Fish in Water', Ph.D. Thesis, University of London, 1960.
10. R. W. G. Haslett, 'Determination of acoustic backscattering patterns and cross sections of fish', *Brit. J. Appl. Phys.*, **13**, No. 7, pp. 349–57, July 1962.
11. R. W. G. Haslett, 'Measurement of the dimensions of fish to facilitate calculation of echo strength in acoustic fish detection', *J. Cons. Perm. Intl Explor. Mer*, **27**, pp. 261–9, 1962.
12. R. W. G. Haslett, 'Physics applied to echo sounding for fish', *Ultrasonics*, **2**, pp. 11–22, January–March 1964.
13. A. Freedman, 'A mechanism of acoustic echo formation', *Acustica*, **12**, No. 1, pp. 11–21, January 1962.
14. D. H. Cushing, 'The acoustic estimation of fish abundance', in 'Marine Bioacoustics' (ed. Tavolga). (Pergamon Press, London, 1966.)

Manuscript received by the Institution on 11th August 1966. (Paper No. 1116/RNA68/EA34.)

U.D.C.621.398:621.39

Communications Aspects of Underwater Telemetry

By

H. O. BERKTAY,
Ph.D.†

AND

B. K. GAZEY,
Ph.D.†

Reprinted from the Proceedings of the I.E.R.E. Conference on 'Electronic Engineering in Oceanography' held in Southampton on 12th–15th September 1966.

Summary: Problems associated with reliable underwater telemetry are studied in three categories dictated by the degree to which multi-path propagation effects are experienced. The maximum range obtainable under ideal conditions of a typical telemetry system is investigated as a function of the required information bandwidth.

The relevant acoustic properties of the sea are summarized and systems, originally used in electromagnetic communication to combat 'flat' and 'frequency selective' fading are reviewed and their adaptability to the present application is discussed. Inter-symbol distortion arising from large differential delays during multi-path propagation is discussed and some proposals are made for the realization of a practical system particularly suitable for underwater telemetry where large path differentials are encountered.

1. Introduction

At the present time data transmission (telemetry) systems are commonplace, their chief function being to monitor and supervise the operation of remote equipment which may be anything from an electrical distribution network to an aircraft, a rocket, a satellite or an interplanetary probe. At first sight it is rather surprising therefore that so far, apart from a few simple systems,[1–5] telemeters have found little application underwater where obviously they would be of great benefit. Vast economy of ship-time would result from the use of expendable instrument probes either capable of relaying their information in real time over distances of several miles or which store information gathered over long periods and relay this in a few minutes upon the reception of a command signal.

It is the purpose of this paper to investigate the possibilities of realizing such telemetry systems and to comment upon the techniques which may have to be employed.

The essentials of a telemetry system consist of a transmitter, a transmission channel and a receiver. Data gathered by instruments are first encoded in some way at the transmitter and then used to modulate a carrier signal which is sent via the transmission channel to the receiver where it is decoded to yield the original data. The complexity of the transmitting and receiving equipment largely depends upon the transmission properties of the transmission channel which for the present application is the sea.

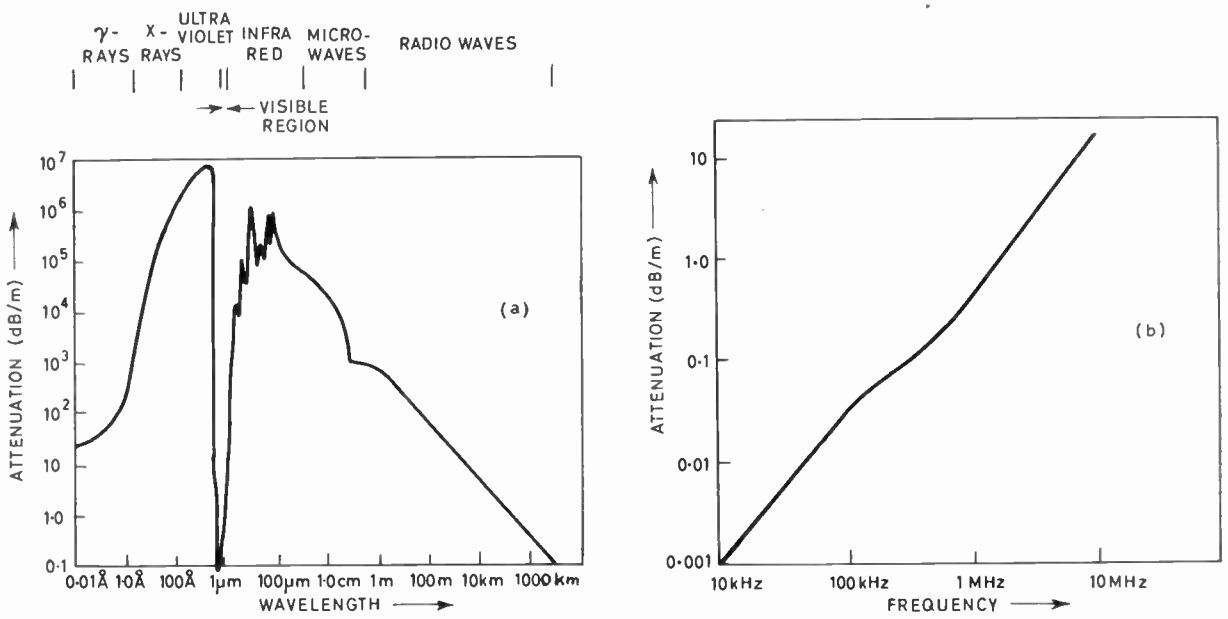
Conventional telemeters employ either electrical signals transmitted over closed wire systems or, more commonly, some form of electromagnetic radiation (radio waves, microwaves, etc.) for the carrier signal. The former method is inconvenient at sea and the latter is unfortunately impractical owing to the high attenuation imposed upon the entire electromagnetic spectrum by sea water. Figure 1(a) shows that only the visible region holds any promise for underwater telemetry and even then in most practical environments the measured values of the absorption coefficients far exceed the clean water values shown.

It is clear therefore that the carrier signal must be an acoustic wave for the attenuation imposed by sea water is then very much less severe provided the frequency is restricted to below 100 kHz (see Fig. 1(b)).

2. The Characteristics of the Transmission Channel

The sea is far from the ideal medium for sound propagation. Its vertical velocity gradient causes the transmitted 'beam' to be refracted, sometimes in such a way as to render impossible any communication between two specific points owing to the formation of shadow zones (see Fig. 2). Also the carrier signal frequency is subject to a randomly fluctuating Doppler shift, as a result of surface and internal waves.

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



(a) Attenuation of electromagnetic radiation in clear sea-water.

(b) Attenuation of acoustic radiation in sea-water at 5°C.

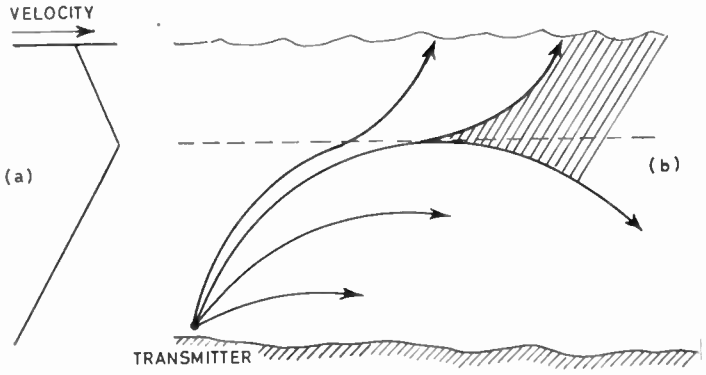
Fig. 1.

The signal is, however, most seriously distorted by multiple path (multi-path) propagation. This term describes a condition which causes the signals from a given transmitter to be received via a number of different paths, each path involving a different transit time delay. Two major factors contribute to this multi-path propagation.

Firstly, the sea is inhomogeneous; its own wave motion breaks up the surface layer (which is usually at a different temperature to the deeper layers) and produces a turbulent and thermal microstructure. This, being a random distribution of small patches of water with slightly differing temperatures to their surroundings which are continually in motion, produces variations of the acoustic refractive index. It is thus possible for several signals each radiated in different initial directions to be refracted in such a way that they all arrive at the receiver where interference occurs owing to the multiplicity of signal transit times and hence relative phases involved.

This phenomenon, often referred to as forward scattering, is illustrated by Fig. 3(a) and produces random fluctuations in both the amplitude and phase of the carrier signal which become progressively more troublesome as range increases.

The signal transit-time between transmitter and receiver is also subject to random fluctuations, although fortunately the r.m.s. transit time differential is small enough to be ignored[7] at all but the longest ranges.



(a) A typical velocity gradient (greatly simplified).

(b) The corresponding ray path diagram. (Horizontal scale is compressed). (The shaded area defines a shadow zone and it would be impossible to communicate with any receiver sited here.)

Fig. 2.

Owing to the random nature of these fluctuations their magnitudes can only be described using statistical methods, the most convenient parameters for this being the coefficient of variation in the case of amplitude fluctuations and the mean-square value in the case of phase fluctuations. The range dependencies of the amplitude and phase fluctuations of a carrier signal at a frequency of 50 kHz are shown in Fig. 4 which is based upon theoretical work performed by Obukhov.[8] At very short ranges, the range dependence of the amplitude fluctuations obeys a three-halves power law but this reduces first to a one-half power law and then becomes constant as the range is increased. At long ranges the signal amplitudes follow a Rayleigh distribution. Throughout, the phase fluctuations are directly proportional to range, and their probability distribution becomes constant only at very long ranges.

Multi-path propagation is also produced by reflections from the boundaries of the medium. Apart from the direct path it is possible for sound to travel between transmitter and receiver after being reflected from a large number of favourably orientated surfaces on the sea floor and in the sea surface (see Fig. 3(b)). Once again interference, and consequently, signal fluctuation at the receiver, results. In general this factor has a far more deleterious effect upon long range acoustic telemetry than forward scattering owing to the much larger transit-time differentials that can occur.

Multi-path propagation produces two additional effects which are worthy of mention. If the transmitted signal is received simultaneously by two closely-spaced hydrophones a comparison of the amplitude, phase or transit-time differential of the

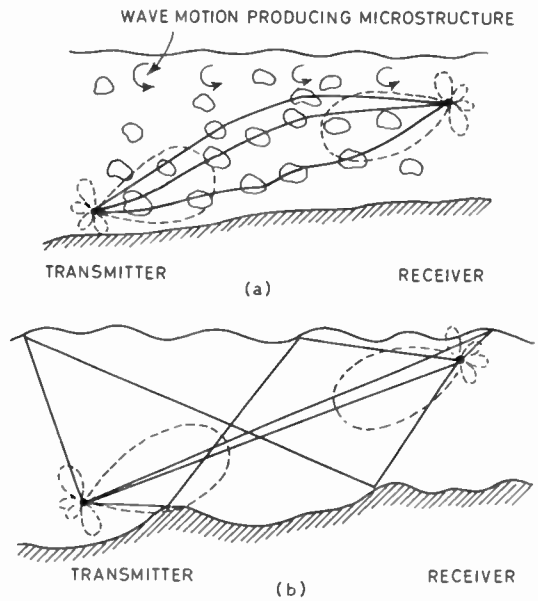


Fig. 3. (a) Illustrating forward scattering.
(b) Illustrating multiple path propagation.
(Both situations are highly exaggerated.)

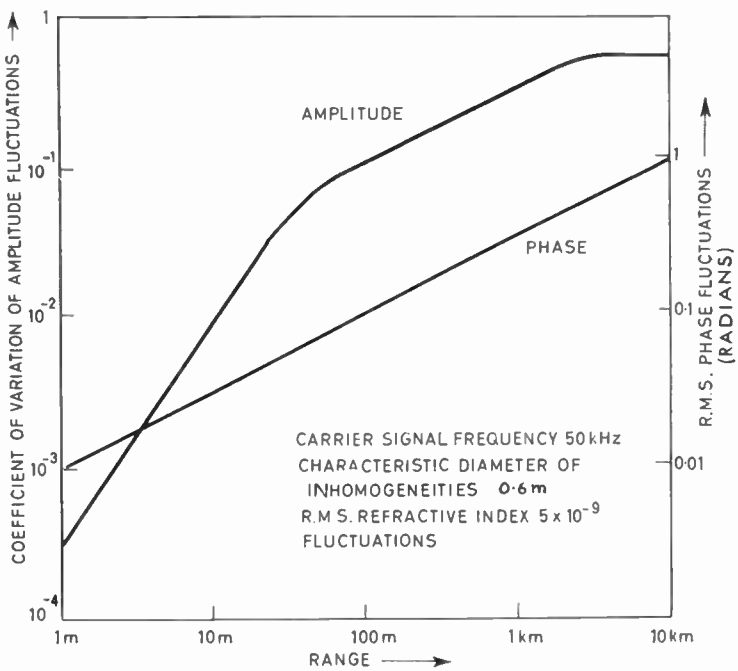


Fig. 4. Carrier signal amplitude and phase fluctuations as a function of range.

two received signals shows that they rapidly become uncorrelated as the hydrophone separation is increased along the transverse direction normal to the direction of sound transmission. Furthermore, if two signals of differing frequency are simultaneously transmitted from a common source and received on a common hydrophone the amplitude, phase, etc., of the two signals become uncorrelated as the frequency separation is increased from zero. As will be shown in Section 3, both of these effects can be made use of to overcome to some extent the errors produced by the channel.

The fluctuations described in this section cause the received signal to fade in a similar manner to the way radio signals reflected from the ionosphere fade and it is proposed in the remainder of this paper to investigate the well-established techniques available to combat the fading of radio signals to see if any are adaptable for our present needs. Before doing this, however, it is useful to examine the range performance to be expected from a typical telemeter operating under ideal conditions so that the full potential of underwater telemetry may be clarified.

3. Communication Under Ideal Conditions

If propagation in the sea is assumed to be free from the effects of refraction, forward scattering and multi-path propagation, the major factors influencing the useful range of a telemeter are as follows:

(*a*) acoustic power radiated by the transmitter,

(*b*) transmitter array gain,

(*c*) acoustic signal attenuation due to 'spreading' of the acoustic beam as it propagates,

(*d*) acoustic signal attenuation due to absorption in the sea,

(*e*) gain of the receiving array,

and

(*f*) noise experienced at the receiver.

Looking at these in more detail enables an optimum specification for a typical telemeter to be evolved.

3.1. *Transmitted Power*

This is limited by the rating and capacity of the battery supplying the transmitter. If the underwater unit is to be reasonably compact and have a satisfactory working life an upper limit to transmitted power of 10 acoustic watts seems reasonable.

3.2. *Transmitter Array Gain*

By confining the transmitted energy along a particular direction thereby producing a 'beam' the effective transmitted power can be greatly increased. The amount of directionality obtainable is a function of the size, operating frequency and method of connection of the transmitting array.[9,10]

A beam which is too narrow would be extremely difficult to align with the receiver, however, and for most situations a beamwidth of 15° between the half-power directions is considered as the lower limit. This increases the effective transmitted intensity by a factor of nearly 200.

3.3. *Spreading Loss*

Within the volume of the sea the intensity at any point obeys an inverse square law with distance. (In shallow water where the sound is bounded by the surface and the sea-floor the spreading loss is often considerably smaller than the value in open water.)

3.4. *Attenuation due to Absorption*

This is a function of frequency as has already been shown by Fig. 1(b). The total attenuation imposed is the product of the range and the attenuation coefficient. Spreading loss is usually combined with the attenuation due to absorption to give the total loss in intensity during transmission, the so-called transmission loss.

3.5. *The Gain of the Receiving Array*

The receiving array can be made more sensitive along a particular direction also using the same techniques employed for increasing the transmitting array gain. The problem of positioning the receiver accurately applies once again, however, so a receiver beamwidth of 15° will be taken as the convenient minimum.

If additional system complexity can be justified, it is possible to use a much narrower receiving beam provided, of course, some tracking facility is available to compensate for motion of the receiver relative to the transmitter. The use of a narrow receiving beam (1°) considerably enhances the operational range of the telemeter.

3.6. *Noise at the Receiver*

Apart from noise in the receiving transducer and the subsequent electronic circuitry, receiver noise comprises three components in general,

(1) noise induced by motion of the sea past the receiving transducer,

(2) thermal noise due to thermal agitation of the water molecules,

and

(3) noise due to the ceaseless activity of the sea, particularly the surface layer, known as sea-state noise.

For simplicity it can be assumed that the receiving array will be stationary to obviate motion-induced noise.

The frequency dependence of the sea-state and thermal noise expressed as the intensity in watts/square metre in a 1 Hz band is shown in Fig. 5. This noise

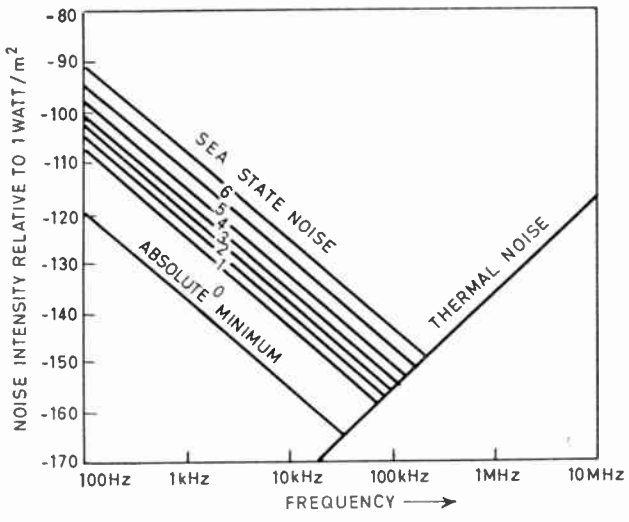


Fig. 5. Sea-noise experienced at the receiver (after Horton[11]).

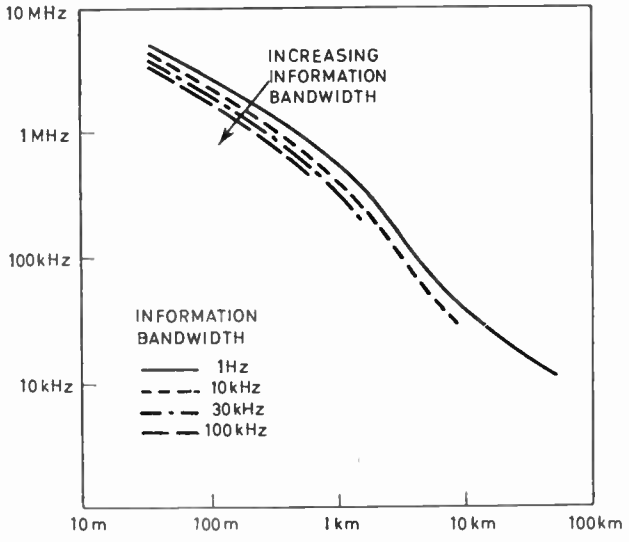


Fig. 6. The range of a typical telemeter transmitter as a function of carrier frequency for various information bandwidths.

intensity is increased proportionally as the receiver bandwidth is increased.

For the purpose of range computations it will be assumed that sea-state 4 will be taken as the general operational limit owing to the difficulty of orientating the receiving gear in rough seas.

This noise is isotropic and the receiving array gain can be used to discriminate against it in favour of the signal. Figure 6 shows the theoretical range, as a function of carrier signal frequency with required information bandwidth as parameter, of a telemeter based on the outline specification just developed. Several points emerge from a study of this.

(*a*) As the required bandwidth is increased the carrier frequency must be increased to accommodate it (owing to the difficulty of obtaining transducers with relative bandwidths (Q-factors) of less than about 5) and the range decreases in consequence.

(*b*) For any given carrier frequency because of the large loss incurred during transmission an increase in the information requirement has relatively little effect upon the range expectation (e.g. at a carrier frequency of 1 MHz increasing the information bandwidth from 1 Hz to 10 kHz changes the range by only 20%).

(*c*) Long ranges (in excess of 10 km) can only be obtained by the use of carrier signals below about 20 kHz.

4. Communication in Presence of Multi-path Propagation

The overall effect of multi-path propagation on a communication link is to reduce the effective bandwidth of, and thus the rate of transmission of information in, the communication channel. Detailed studies of the effects of multi-path propagation have been made in the literature,[12] with particular reference to communication channels using electromagnetic waves. However, the underwater (acoustic) channel has certain characteristics which justify a new and separate analysis of its communications potential.

4.1. *A Brief Outline of the Effects due to Multi-path Conditions*

The effect of multi-path propagation on a single pulse is to spread it in time, the pulse-spreading being determined by the maximum difference in the transit delays associated with the various paths. Relative movement in the parts of the medium which cause the multi-path propagation to come about results in a difference in the frequencies of the signals coming via different paths (Doppler-shift). Thus, a modulated signal suffers a time spreading due to the differential time delays and a frequency spreading due to different Doppler shifts arising on the various paths.

As mentioned previously if a single sine-wave is transmitted through such a channel, the received signal becomes the resultant of a number of sine-waves of slightly different frequencies, which have been subjected to different time delays. As a result, the amplitude and the phase of the received signal fluctuate in a random manner. The amplitude fluctuations ('fading') of two sine-waves of different frequencies transmitted through such a channel are correlated if the frequency difference is not greater than the inverse of the pulse-spreading in the channel. Hence, if a band-limited signal has a bandwidth

which does not exceed the inverse of the pulse-spreading in a multi-path channel, then it retains its frequency-spectrum (apart from an ambiguity arising from the Doppler shift and the Doppler spreading), when transmitted through such a channel. Therefore, the inverse of the pulse-spreading can be said to be the coherent bandwidth of the channel. If the signal bandwidth is increased beyond this value, the various frequency components of the signal fade in an uncorrelated manner, causing the signal to be distorted.

If all the spectral components of the signal fluctuate in a correlated manner 'flat-fading' of the signal occurs, while lack of correlation of the fluctuation characteristics of different frequency components signifies 'frequency-selective fading' of the signal.

Another effect which occurs in an underwater communication channel is 'inter-symbol interference'. If, for example, two distinct signal paths exist and the difference between the time-delays associated with the two paths is at least of the order of the duration of a binary symbol (i.e. of a mark or a space), then the demodulated signal will be distorted because of interference with its delayed replica.

4.2. *Means of Combating Effects due to Multi-path Propagation in a Communication Channel*

4.2.1. The case of a flat-fading signal

If a received modulated-carrier is subject to 'flat-fading', the carrier/noise ratio (referred to the input port of the receiver) becomes a time dependent quantity. Then the performance of a receiver, being dependent upon the carrier/noise ratio, becomes also time dependent. This leads to the performance requirements for a given system to be specified as those obtaining for a prescribed fraction of time.† Thus, for a given receiver, the performance requirements can be realized if the carrier level at the input port can be maintained above a limiting value for the prescribed fraction of time.

For given fading statistics the system can be designed either to realize such a requirement for the received carrier level or by altering the effective statistics of the fading through the use of diversity techniques. The first solution generally leads to the over-engineering of a communication system, while the use of diversity techniques normally necessitates the duplication of parts of the receiving equipment.

The fading statistics obtained in flat-fading conditions normally conform to a Rayleigh distribution; this is known to be the case in underwater propagation over long ranges. The characteristics of a carrier subject to Rayleigh-fading and the performance improvements obtainable under these conditions through the use of diversity techniques are well known. In underwater applications, space or frequency diversity can be used without reducing the rate at which information is being transmitted, provided the frequency separation or the receiving transducer separation is such that the signal fluctuations in the two channels are uncorrelated.

An example may help to illustrate the advantages of diversity combining. The received signal will be assumed to have Rayleigh type of fading characteristics. The method of diversity combining used will be a simple selection diversity system; i.e. one in which the highest signal in any one of the N diversity channels at any one time is selected. It will be assumed that the average signal powers in all the various channels are equal and the characteristics of the equipment in the diversity channels are all identical. If $p(V)$ is the probability that the signal amplitude in a channel is below a level V, then the probability that it is below the same level in N independent channels at the same time is $[p(V)]^N$.

Let the demodulating and decision making equipment be such that the error-rate requirement is met if the signal amplitude is not less than a value V_0 and the fraction of time this error rate may be exceeded is M. Then the requirement for the signal statistics at the output of the diversity combiner must be

$$[p(V_0)]^N = M$$

If $M = 1\%$ and $N = 2$,

$$p(V_0) = 10\%$$

To obtain a similar improvement (in the duration of time over which a given error rate is exceeded) without using diversity techniques, the average signal power would have to be raised. The required increase of the average signal power can be calculated by using the fading statistics of the signal. The Rayleigh distribution function is of the form

$$p(V) = 1 - \exp(-V^2/2v^2)$$

where v^2 is the mean-square value of the amplitude of the signal. Without diversity, the required system performance is met if

$$p(V_0) = M$$

If $$M \ll 1, \quad V_0^2/2v^2 \ll 1.$$

Hence,

$$p(V_0) \simeq V_0^2/2v^2$$

Therefore, to improve $p(V_0)$ from 10% to 1% the mean signal-power (that is, v^2) must be increased by 10dB.

In practical underwater communications flat-fading conditions will exist in particular if the multi-path propagation is due to forward scattering.

† For example, in a digital data system this could be in terms of the average error-rate which is not to be exceeded except for a small percentage of time.

As pointed out in an earlier section, the coherent bandwidth obtainable in such a case would be so large that frequency diversity cannot be used; space diversity techniques can, however, be realized.

4.2.2. Frequency-selective fading

A way of combating frequency-selective fading is the use of signals with large bandwidths to represent the binary symbols. If the signal bandwidth is large compared with the coherent bandwidth of the channel, a compression of the bandwidth associated with the demodulation process can provide effective frequency diversity.

A method which has been used successfully to communicate under severe frequency-selective fading conditions relies on a digital pulse-compression system, the pulse compression being used to separate and add coherently (thus eliminating the destructive interference between) all the signal components arriving via different paths.[13] The need to synchronize the pseudo-random sequence generators in the transmitting and the receiving equipment and, further, to maintain this synchronism while one of the terminals (i.e. the ship-board equipment) is moving relative to the other, would cause serious problems in the use of such a system in underwater communications. A system based on a similar concept can be developed using an analogue type of pulse compression. For example, consider f_1 and f_2 to be two frequencies such that $f_1 < f_2$ and $f_2 - f_1$ is much greater than the rate at which binary signals are to be transmitted over the communication link. A binary signal can be devised in the form of a carrier whose frequency is being varied with a positive slope between the frequencies f_1 and f_2 (i.e. $f_1 \rightarrow f_2$) during the transmitted binary pulse. If the received signals are applied to a suitable dispersive delay line, the signal components arriving with different delays will form different pulses, each of duration $1/(f_2-f_1)$. Thus, if the delay difference between two paths is not less than $1/(f_2-f_1)$, the different contributions can be separated and, if necessary, added coherently by the use of an adaptive combining arrangement similar to that used in the Rake system.† If the pulse-spreading due to the multi-path effects in the communication channel is less than the duration of a binary pulse, such a system would provide reliable communications. However, if the pulse-spreading is greater than the duration of a binary pulse, some of the compressed pulses due to a binary signal will appear within the duration of subsequent binary symbols, causing an ambiguity in the determination of the identity of the received binary signals. This leads to what has been termed 'inter-symbol interference'.

In a binary communication system two separate identifiable signals need to be used for 'mark' and 'space'. In the pulse-compression system mentioned the second signal could occupy a different frequency channel or, more simply, could consist of a carrier the frequency of which is varied, this time, with a negative frequency gradient between f_2 and f_1 (i.e. $f_2 \rightarrow f_1$).

4.2.3. Inter-symbol interference

As pointed out in the previous sub-section, if the pulse-spreading in the channel exceeds the duration of a binary symbol, then errors may be caused because of interference between binary symbols in a message. A method used to improve the rate of transmission of information under these circumstances has been the coding of the information at the transmitting end on a multi-level rather than a binary basis by the use of a multi-tone system. In this way the number of symbols transmitted per second can be maintained within the limitation imposed by the coherent bandwidth of the channel, but the information content of each symbol is increased. The following simple example illustrated the point. The transmitter sends out three tones of different frequencies for the duration of each symbol, there being six possible frequencies to choose from. This is tantamount to sending one of twenty characters per symbol, and gives an information content of 4·3 bits per symbol instead of the 1 bit per symbol which would be obtained in a binary system transmitting one out of two possible frequencies.

To combat frequency-selective fading in such a system, either a wide-band signal can be used for each 'tone' in the way described in the previous sub-section or diversity reception techniques could be used for each of the receiving channels (i.e. for each tone).

Such a multi-level signalling scheme would be very useful for communication in channels in which the pulse-spreading is not unduly time-dependent. But paths with very large time delays may often occur in an underwater channel (e.g. due to specular reflection from the troughs of waves) which alter the overall limits of the pulse-spreading suddenly and by a large amount. If the symbol duration is fixed in such a way as to accommodate the delays associated with all such paths, the signalling rate would have to be reduced beyond limits desirable in practice.

These conditions may be overcome by adopting a scheme based on time-dependent coding. For example, let the first symbol (for the moment, a binary

† In underwater communications signals arriving via different paths usually have only a nuisance value, and in a simple system it may well be sufficient to detect the compressed signal by means of an envelope demodulator followed by a low-pass filter with a cut-off frequency corresponding to the speed of binary signalling.

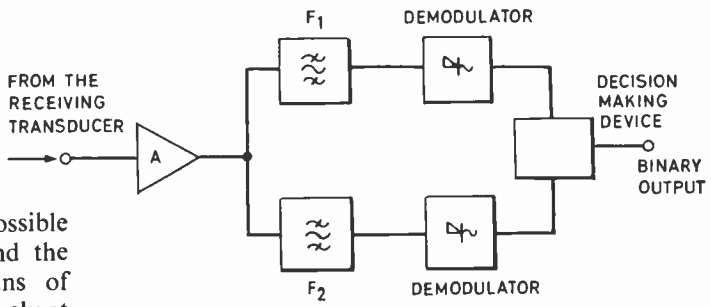


Fig. 7. Block schematic of a simple f.s.k. system.

system is assumed) consist of one of two possible signals in a frequency band centred at f_1, and the subsequent symbols be transmitted by means of similar signals but within bands centred successively at frequencies $f_2, f_3 \ldots f_k$. In this way inter-symbol interference can be avoided over a sequence of k symbols. Such a system would require more expensive transmitting (as well as receiving) equipment and its use could only be justified in extreme cases where a high rate of transmission of information is required and other means of collecting such information (e.g. cable, sono-buoys, etc.) are likely to lead to even more expensive solutions.

5. System Considerations

A simple binary communication technique suitable for underwater telemetry can be realized by the use of frequency-shift keying (f.s.k.). The receiver of an f.s.k. system is indicated in block-diagram form in Fig. 7. Signals at frequencies F_1 or F_2 are transmitted depending upon the binary symbol at the input to the transmitter. The receiver essentially consists of two matched filters (at the frequencies F_1 and F_2) and a decision-making device, which compares the outputs of the two matched filters and decides whether the transmitted binary symbol was a mark or a space. To prevent the overlapping of the sidebands, $|F_2-F_1|$ must be made large compared with the rate of signalling. Such a simple system provides satisfactory service under ideal propagation conditions, provided the signal/noise ratio in each channel is maintained at an acceptable level.

The problems encountered in a real channel have been summarized in Sections 2 and 3 and it is now proposed to study their effects on a practical telemeter, taking the simple f.s.k. system outlined above as the basis.

5.1. *Doppler Shift*

In the simple f.s.k. system each receiving channel will have a bandwidth which is approximately equal to the signalling rate, in order to obtain matched-filter reception. The possible existence of a Doppler shift introduces an ambiguity in the centre frequency of the channel which can be catered for in a system *either* by widening the pass-band of the receiving filters *or* by using 'tracking' filters. For example, a relative speed of 10 knots between the transmitter and the receiver will produce a Doppler shift corresponding to about 0·3% of the transmitted frequency. For a transmitted frequency of about 80 kHz, this means an uncertainty (as regards the frequency of the received signal) of about ± 260 Hz. For a signalling rate of about 500 bits/second, for example, the bandwidths of the receiving channels would have to be doubled in order to accommodate this frequency drift due to Doppler shift, thereby causing a deterioration of the system performance for a given received signal level.

The use of a self-tracking receiver prevents this deleterious effect on system performance, at the expense of increasing the complexity of the receiving equipment.

5.2. *Doppler Spreading*

This would be relatively small and can be accommodated by a slight increase in the bandwidths of the receiving channels, without affecting the system performance seriously.

5.3. *Flat-fading*

If the frequency difference $|F_2-F_1|$ is less than the coherent bandwidth of the propagation channel, B_c (see Sect. 4.1) then the mark and the space signals will fade simultaneously. Then, a space-diversity system can be used by duplicating the pre-amplifier (A in Fig. 7) and combining the signals from the two diversity paths before applying them to the mark and the space channels.

5.4. *Frequency-selective Fading*

In the context of the simple f.s.k. system, frequency-selective fading can be considered in two stages.

(*a*) If $|F_2-F_1|$ is greater than B_c, the coherent bandwidth of the channel, but the signalling rate is less than B_c, then the output of each filter will be subject to flat-fading, while the mark and the space signals will fade independently of one another. Hence, diversity techniques can be used on each receiving channel separately before each signal is demodulated. This would necessitate the duplication of almost the whole of the receiver, but would provide a signalling rate of the order of the coherent bandwidth of the propagation channel.

(*b*) If the signalling rate is greater than the coherent bandwidth, B_c, the inter-symbol interference will be

the most serious effect. For, although frequency-selective fading can be combated by the use of a band-spreading pulse-compression system (see Sect. 4.2), incorrect (binary) decisions may be made by the receiver because of signal overlap due to signals associated with different symbols arriving within the same symbol interval.

5.5. *Inter-symbol Interference*

In underwater data transmission the problem of interference between symbols is likely to arise in two situations. *Either* the pulse spreading obtaining in the propagation channel is very high causing inter-symbol interference even at low rates of data transmission, *or* the required information rate is very high. In the latter situation, in general, it would be justifiable to make the transmitting device more complex and expensive. In this case, a multi-frequency scheme as described in Sect. 4.2.3, may give acceptable results. For example, if each symbol transmitted consists of five of ten possible frequencies (transmitted simultaneously), each symbol has an information content of approximately 8 bits. A high-resolution sonar system developed for fisheries research can be shown to have a video information rate of about 60 000 bits per second if there are no more than about 300 individual targets per frame. Thus, if the propagation channel has a coherent bandwidth of not less than 10–15 kHz, such a multi-frequency acoustic data transmission system can be devised for the transmission of the video information to a surface vessel from a submerged sonar system.

If the data-rate requirement is low, then a simpler multi-tone system can be used to enhance the rate at which data can be transmitted for a given overall pulse spreading. For example, as mentioned in Sect. 4.2.3, by sending three out of six possible frequencies for each symbol the rate of transmission of information (in bits per second) is increased by a factor of 4·3.

6. Conclusions

The properties of the acoustic propagation channel consisting of the sea, together with its boundaries, dictate the degree of complexity of an underwater telemetry system for a given relative configuration of the transmitting and the receiving equipment. In this paper, the propagation characteristics of an underwater acoustic channel have been summarized and their effects upon the rapid and reliable transmission of information have been investigated. Some of the points arising from this investigation need to be underlined.

(i) The fluctuations produced by the non-ideal propagation properties of the medium itself (i.e. disregarding the effects due to the boundaries of the medium) do not limit seriously the rate of transmission of information. The use of space-diversity in such a channel is sufficient for improved reliability of communication even with moderate transmitted acoustic powers.

(ii) Fading, either frequency-selective or 'flat', forms the most likely cause of symbol errors at short and medium ranges involving transmission between an instrument in mid-water or on the sea floor and a ship on the surface. The use of diversity or pulse compression techniques or both together is promising as regards the improvement of the reliable transmission rate of information under such conditions.

(iii) Inter-symbol interference becomes a very serious limitation to the rate at which information may be transmitted over long ranges, particularly in shallow water. If the pulse spreading in the channel is known to be restricted to a certain upper limit, then a multi-tone system should permit reliable communication without reducing the data transmission rate substantially. If the degree of pulse-spreading is subject to large fluctuations a system in which successive symbols are transmitted on a different carrier signal frequency may be of use but the data rate would have to be reduced considerably.

(iv) When the amplitude (and phase) fluctuation rates are not excessive it is possible to modify the apparent properties of the channel to some extent. Thus some of the more offending multiple propagation paths may be eliminated (or their effect greatly reduced) by the use of adaptive (i.e. signal tracking) receiving arrays or ones designed to present their minimum sensitivity along the directions of the multi-path signals, e.g. along the sea surface.

Throughout, the emphasis has been placed on maintaining equipment simplicity at the transmitter so that it may be considered expendable. Complexity of the receiving equipment is considered of little consequence since this is situated on board ship where adequate space, power supplies, trained personnel, etc., are available.

In the opinion of the authors the preliminary work presented here shows that in spite of the fact that little attempt has so far been made to realize a highly reliable underwater telemeter every effort should be made to pursue this interesting topic, for although the problems encountered are vast they are not insurmountable and the rewards should make the labour involved worthwhile.

7. References

1. W. Dow, 'Underwater telemetering', *Deep Sea Res.*, **2**, p. 145, 1955.
2. W. Dow, 'A telemetering hydrophone', *Deep Sea Res.*, **7**, p. 142, 1960.

3. M. J. Tucker, R. Bowers, F. E. Pierce and B. J. Barrow, 'An acoustically telemetering depth gauge', *Deep Sea Res.*, **10**, p. 471, 1963.
4. A. N. Paramonov, 'The use of the principles of discrete counting for the transmission of data from deep-sea measurements', *Okeanologiya*, **1**, No. 4, p. 710, 1961.
5. D. E. Cambell, R. J. Cyr and C. Crozier, 'A f.m.-f.m. acoustic telemeter', *Electronics*, **35**, p. 141, 12th January 1962.
6. N. G. Jerlov, and B. Kullenberg, 'On radiant energy measurements in the sea,' *Svenska Hydrogr.-Biol. Komm. Skr., Series 3, Hydrogr.*, **1**, No. 1, 1946.
7. L. A. Chernov, 'Wave Propagation in a Random Medium', p. 29, (McGraw-Hill, New York, 1960).
8. A. M. Obukhov, 'On the influence of weak atmospheric inhomogeneities on the propagation of sound and light', *Akad. Nauk S.S.S.R., Ser. Geofiz.* **2**, p. 155, 1953.
9. S. A. Schelkunoff, 'A mathematical theory of linear arrays', *Bell System Tech. J.*, **22**, p. 80, January 1943.
10. D. G. Tucker and B. K. Gazey, 'Applied Underwater Acoustics', Chapter 6, (Pergamon, London 1966).
11. J. W. Horton, 'Fundamentals. of Sonar', (United States Naval Institute, Annapolis, 1957).
12. R. S. Kennedy, 'Signal design for dispersive channels', *I.E.E.E. Spectrum*, **1**, p. 231, March 1964.
13. R. Price and P. E. Green, Jr., 'A communication technique for multipath channels', *Proc. I.R.E.*, **46**, p. 555, March 1958.

Manuscript received by the Institution on 17th May 1966. (Paper No. 1117/EA35.)

U.D.C. 621.372.57

Active Filters with Low Sensitivity to Element Changes

By

Professor J. VLACH,
Dipl. Ing.†

AND

J. BENDÍK,
Dipl. Ing.‡

Summary: The active element has the greatest influence on the properties of active filters and may cause oscillations especially in more complicated networks. An active filter network is described which is almost independent of element changes. The analysis shows that the constant of the negative impedance converter may be changed by as much as $\pm 10\%$ without appreciably influencing the properties of the filter. Synthesis was carried out for a whole system of filters of fourth- and sixth-degree with various minimum attenuations in the stop-band. The changes of all other elements are briefly considered and it is shown that the influence of all elements not determining the position of the transmission zero is less than the same relative change of the impedance converter constant.

1. Introduction

The theory of active filters with lumped R-C elements is well established and realizations are theoretically possible for all kinds of transmission functions. From the viewpoint of practical use, there still remains, however, the question of the stability of these filters, because the changes of passive as well as active elements are to be considered. The sensitivity to such changes is usually investigated in the differential vicinity of the desired pole positions. From such considerations it follows that with a given type of approximation the sensitivity rapidly increases with the degree of the filter. For Butterworth and Chebyschev filters, such dependence was investigated by Holt[1] and it was shown that a 10% change of the elements and of the constant of the impedance converter already makes their use dubious at the fourth-degree. In another paper,[2] a filter with eight poles and four imaginary zeros was considered, in which the poles were chosen to fulfil the requirement for maximally flat group delay, but the realization was carried out using two negative impedance converters (n.i.c.), separated by a buffer amplifier. Probably, the most complicated filter with only one n.i.c. is described by Antreich and Gleissner.[3] Schüssler–Jess's filters for the best step-response were used and a filter of sixth-degree with four imaginary zeros was realized.[4-6] The stability of this filter is, however, very poor leading to oscillations with a 0·6% change of the n.i.c. constant. This paper describes the synthesis and analysis of a filter of the same degree, which is well suited to transmission of pulses and which is essentially insensitive to the element changes as well as to the change of the converter itself.

† Popov Research Institute of Radio Communications, Prague; now Visiting Professor, Department of Electrical Engineering, University of Illinois, U.S.A.

‡ Popov Research Institute of Radio Communications, Prague.

2. Principle of the Negative Impedance Converter and its Properties

The negative impedance converter is a four-pole network whose input impedance is negatively proportional to the load impedance; i.e.

$$Z_{in} = -k.Z_L \qquad \text{.....(1)}$$

If a general four-pole network with a load Z_L is considered, one can write

$$Z_{in} = \frac{A.Z_L + B}{C.Z_L + D} \qquad \text{.....(2)}$$

where the constants A, B, C and D are constants of the chain matrix. For a n.i.c. the following conditions must therefore be valid, namely,

$$B = C = 0 \quad \text{and} \quad A/D = -k \qquad \text{.....(3)}$$

The first condition is never fulfilled in practical circuits (i.e. $B \neq 0$, $C \neq 0$). This may be, however, taken into account by the synthesis of the filter. The second condition may be fulfilled using either a voltage inversion type ($A = -k$, $D = 1$), or a current inversion type of n.i.c. ($A = 1$, $D = -1/k$). There exist, therefore, two classes of n.i.c. Both may be used and they are in fact equivalent. In this paper, the current inversion type will be considered.

Several types of n.i.c.'s, employing two transistors, are described in literature. A classification of these converters using degenerate one-ports is described by Braun.[7] Ten basic circuits are described, half of them are of the current and half of the voltage inversion type. The realization of these ten basic circuits altogether, leads to twelve practical circuits.

Several authors have showed that the changes of the constant of n.i.c. strongly influence the properties of filters.[1-3] Experimental analysis of the properties of several n.i.c.'s employing two transistors was therefore carried out, the dependence of k being

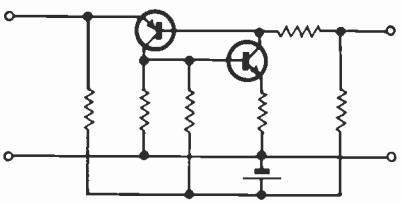


Fig. 1. Circuit for a negative impedance converter (after Yanagisawa[8]).

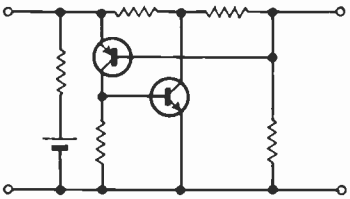


Fig. 2. Another form of n.i.c. circuit (after Nagata[9]).

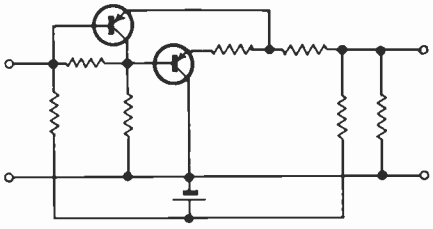


Fig. 3. Another stable n.i.c.

measured by changes of the power supply voltage and by changes in temperature. All these circuits used direct connections of the transistors, thus enabling the realization at low frequencies by some solid-state technology. Six of the possible twelve circuits were designed to fulfil the condition of minimum changes of k with the changes in temperature and in the supply voltage. Calculations and measurements showed that there are three most stable circuits. The best one, designed originally by Yanagisawa[8] is shown in Fig.1 (circuit III according to the classification given by Braun[7]). The second circuit, originated by Nagata,[9] is shown in Fig. 2 (circuit IV according to Braun's classification). The last (fifth according to the classification) is shown in Fig. 3.

All converters are of the current inversion type. The circuit in Fig 1 uses one p-n-p and one n-p-n germanium transistor. It has good characteristics with $k = 1$: a $\pm 20\%$ change in the supply voltage causes a $\pm 2\%$ change in k, whereas a temperature change from $+20°C$ to $+50°C$ causes a total change of 2% of k. The converter in Fig. 2 was realized with one p-n-p and one n-p-n transistor. To achieve good properties the value of k had to be chosen and made equal to 5 and silicon transistors had to be used. Such a high value of k leads, however, to a greater spread of the element values, which may be inconvenient especially in more complicated filters. A $\pm 20\%$ change of the supply voltage caused $\pm 5\%$ change of k; the change of k was not measurable in the temperature range from $+20°C$ to $+50°C$. Finally, the converter of Fig. 3 uses two similar transistors, either p-n-p or n-p-n. Satisfactory results may be achieved by making $k = 2$ and by the use of Ge transistors. A $\pm 20\%$ change of the supply voltage leads to a $\pm 5\%$ change of k. The temperature change from $+20°C$ to $+50°C$ causes an increase k by 5%. The use of silicon transistors increases the admissible temperature range, but to improve the stability of k, the supply voltage has to be considerably increased.

Without special means (i.e. stabilization of the power supply and stabilization of the temperature), smaller changes of k than approximately $\pm 5\%$ cannot be expected.

3. Consideration of Approximations

From the results of the preceding section it follows that a $\pm 10\%$ change of k has to be taken into consideration to meet all conditions encountered in practice. Such a change is, however, too great for more complicated filters, which have been described up to now. The problem would seem to have no solution if it were not realized that all present designs always used approximations useful for passive networks without any regard to the specific properties of active filters. If these filters are able, in principle, to oscillate, it is useful to find such approximations which would lead to oscillations after considerable changes of parameters. The oscillations can build up only when one of the poles is moved to the imaginary axis of the p-plane. If at the approximation an additional condition, namely, that the poles should be as far from the imaginary axis as possible, were considered, then an improvement in the properties of active filters could be expected.

From the approximations with poles remote from the imaginary axis, the Thomson approximation for maximally flat group delay[10] is very attractive. If imaginary zeros are added, the amplitude response will be improved without influencing the group-delay characteristic. Therefore such networks will be suitable for transmission of impulses. The choice of the transmission zeros can optimize the stop-band response in such a way that all the minima of attenuation will have the same value. This type of approximation is described by Feistel and Unbehauen;[11] a brief description is here given for even-order filters.

Let us start with an auxiliary complex plane

$$z = u + jv \qquad \ldots\ldots(4)$$

and a function

$$F_1(z) = \prod^{n} \frac{z+z_i}{z-z_i} \qquad \ldots\ldots(5)$$

where z_i's are either real or in complex conjugate pairs, $\mathrm{Re}\, z_i < 0$. The values of the function $F_1(jv)$ may be depicted by a radius vector of unit length, rotating around the origin, e.g. the locus of these points will be a unity circle with the centre at the origin.

If we create a new function

$$F_2(z) = \frac{1}{2}\left[F_1(z)+1\right] = \frac{1}{2}\,\frac{\prod^{n}(z+z_i)+\prod^{n}(z-z_i)}{\prod^{n}(z-z_i)}$$

then for $z = jv$ this function will be depicted by a circle of unit diameter, crossing the imaginary axis in the points 0 and 1. The absolute values of $|F_2(jv)|$ will be between one and zero. If we now create the inverse function of $F_2(z)$

$$G(z) = \frac{1}{F_2(z)} = \frac{\prod^{n}(z-z_i)}{\mathrm{Ev}\prod^{n}(z+z_i)} \qquad \ldots\ldots(6)$$

then $|G(jv)|$ will assume values between infinity and one, as shown in Fig. 4. (The symbol Ev means even part of the polynomial.) Approximations of amplitude functions are always carried out using the square of the absolute value, i.e. using the expression

$$G(z)G(-z) = \frac{\prod^{n}(-z^2+z_i^2)}{\left[\mathrm{Ev}\prod^{n}(z+z_i)\right]^2}$$

Should the function oscillate between e^{a_s} and infinity and not between the values one and infinity, we build a new auxiliary function

$$H(z)H(-z) = G(z)G(-z)\,e^{2a_s} = \frac{\prod^{n}(-z^2+z_i^2)}{\left[\mathrm{Ev}\prod^{n}(z+z_i)\right]^2}\cdot e^{2a_s} \qquad \ldots\ldots(7)$$

which will fulfil this requirement. This rational function has the desired response in the stop-band, it has, however, no pass-band and no pole in the infinity. This drawback can be removed by a suitable transformation which moves the point v_0 (the last attenuation pole) into infinity and the point $v = 0$ shifts to a certain frequency Ω_{a_s} of the p-plane, this being given by the condition that for $p = 0$ the function has to be

$$|H(p)|_{p=0} = 1 \qquad \ldots\ldots(8)$$

The transformation fulfilling this requirement is

$$z^2 = v_0^2\,\frac{\Omega_{a_s}^2+p^2}{v_0^2-p^2} \Leftrightarrow p^2 = v_0^2\,\frac{z^2-\Omega_{a_s}^2}{z^2+v_0^2} \qquad \ldots\ldots(9)$$

where up to now both v_0 and Ω_{a_s} are unknown. They can be computed using two conditions:

(i) The point v_0 is the zero of the denominator of eqn. (7),

$$\mathrm{Ev}\prod^{n}(z+z_i) = \prod^{n}\left(jv_0 + v_0\sqrt{\frac{\Omega_{a_s}^2+p_i^2}{v_0^2-p_i^2}}\right) + \prod^{n}\left(jv_0 - v_0\sqrt{\frac{\Omega_{a_s}^2+p_i^2}{v_0^2-p_i^2}}\right) = 0 \qquad \ldots\ldots(10)$$

(ii) If the transformation (9) is substituted in eqn. (7), then for $p = 0$ is $z = \Omega_{a_s}$, eqn. (7) acquires the form

$$|H(p)|_{p=0}^2 = \left.\frac{\prod^{n}(-z^2+z_i^2).e^{2a_s}}{\left[\mathrm{Ev}\prod^{n}(z+z_i)\right]^2}\right|_{z=\Omega a_s} = \frac{\prod^{n}\left(\Omega_{a_s}^2 - v_0^2\,\frac{\Omega_{a_s}^2+p_i^2}{v_0^2-p_i^2}\right)}{\left[\mathrm{Ev}\prod^{n}\left(\Omega_{a_s}^2 + v_0\sqrt{\frac{\Omega_{a_s}^2+p_i^2}{v_0^2-p_i^2}}\right)\right]^2}\cdot e^{2a_s} = 1 \qquad \ldots\ldots(11)$$

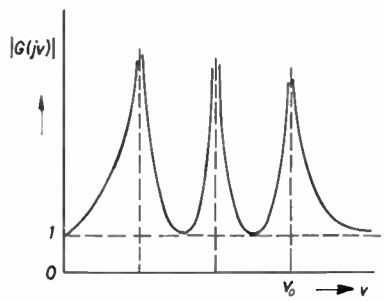


Fig. 4. Response of the function $|G(jv)|$.

Simultaneous solution of (10) and (11) determines the unknown v_0 and Ω_{a_s} so that all data of the filter may be found. In these equations the p_i's are the coordinates of poles of the Thomson filters, given in Table 1. All necessary data for the filter of fourth and sixth degree are rewritten in Tables 2 and 3. Here e^{a_s} is given in decibels in steps of 4 dB. The coordinates of the attenuation poles are Ω_1, Ω_2; $\Omega_{3\,\mathrm{dB}}$ is the frequency where the response is 3 dB down and Ω_{a_s} is the frequency at which the response acquires the minimal attenuation for the first time. The response of the function of eqn. (7) along the imaginary axis $j\omega$ is shown in Fig. 5.

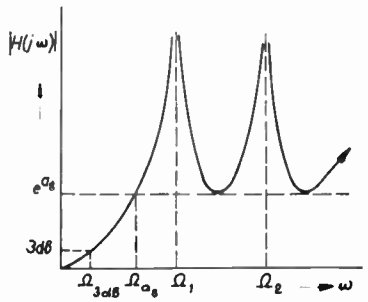


Fig. 5. Attenuation characteristic of the filters considered.

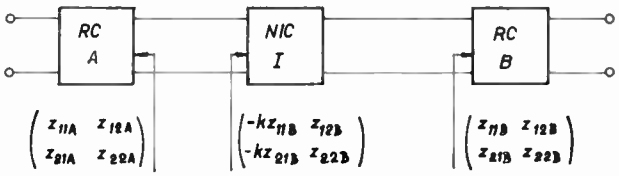


Fig. 6. Principle of an active filter with cascade connection and with a current type of n.i.c.

4. The Synthesis Chosen

Two methods are generally used for the synthesis of active R-C filters. One of them, described by Yanagisawa[8] connects one R-C circuit in cascade with the n.i.c. and the second R-C circuit is connected in parallel to this combination. Linvill[12] proposed another method where two passive circuits are connected in cascade by means of a n.i.c. The drawback of the Yanagisawa's synthesis is that not only the poles, but also the transmission zeros are influenced by changes of the converter and of all components used. On the contrary, Linvill's method determines the transmission zeros entirely by the R-C circuits, independently of the converter, thus leading to a decrease of sensitivity of the whole circuit. For this reason, Linvill's method was used.

As already mentioned, the current inversion type of n.i.c. with the cascade matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & -1/k \end{pmatrix}$$

will be used. The transfer impedance to be realized is

$$Z_{21} = K\frac{A(p)}{B(p)} = K\frac{\prod_{1}^{m}(p-p_i')}{\prod_{1}^{n}(p-p_i)} \qquad \text{.....(12)}$$

Every transfer impedance Z_{21} may be divided in a cascade connection of two circuits using the formula

$$Z_{21} = \frac{z_{21A}\cdot z_{21B}}{z_{22A}+z_{11B}}$$

where all the z_i's are seen from the point of mutual connection. If a current type of n.i.c. is inserted between these two circuits, the matrices change according to those given in Fig. 6 and the whole transfer

Table 1

Horowitz's decomposition and original poles for the maximally flat group delay.

n	α_1	B_0	β_1	p_1
2	−1·73205080	0·46410161	—	−1·5000000 ± j 0·8660254
4	−1·94058110	4·44186546	−3·11082937	−2·8962106 ± j 0·8672341
	−5·28035163			−2·1037894 ± j 2·657418
6	−1·87032119	16·0210479	−3·52278649	−4·2483594 ± j 0·8675097
	−4·48453287		−5·96190949	−3·7357084 ± j 2·6262723
	−12·1556697			−2·5159322 ± j 4·492673
8	−1·68106890	39·3134540	−3·69499079	−5·587886 ± j 0·8676144
	−4·99699323		−5·85540336	−5·2048408 ± j 2·6161751
	−7·0969041		−10·2590798	−4·3682892 ± j 4·4144425
	−23·8817608			−2·838984 ± j 6·3539113
10	−1·48640941	78·4321876	−3·66141681	−6·9220449 ± j 0·8676651
	−5·32043698		6·42946118	−6·6152916 ± j 2·6115683
	−7·22449552		−8·34415416	−5·9675282 ± j 4·3849471
	−10·6558969		−16·3363543	−4·8862195 ± j 6·2249855
	−42·0288552			

impedance is equal to

$$Z_{21} = \frac{kz_{21A} \cdot z_{21B}}{kz_{11B} - z_{22A}} \qquad \text{.....(13)}$$

This equation can be compared with eqn. (12) if $B(p)$ is decomposed into two polynomials with negative real roots, as is necessary for the realization of an R-C circuit. The question of the best decomposition was described by Horowitz;[13] the least influence of the element changes is secured if the decomposition has the form

$$B(p) = \prod^{n/2}(p-\alpha_i)^2 - B_0 p \prod^{n/2-1}(p-\beta_i)^2 \text{.....(14)}$$

A brief but clear description is also given by Antreich and Gleissner.[3] This decomposition was computed for even polynomials and for the maximally flat group-delay filter. The given poles and the results are shown in Table 1. Using this decomposition, the transfer impedance will be rewritten in the form

$$\frac{Z_{21}}{K} = \frac{A(p)}{\prod^{n/2}(p-\alpha_i) p \prod^{n/2-1}(p-\beta_i) \times} \times \left\{ \frac{\prod^{n/2}(p-\alpha_i)}{p \prod^{n/2-1}(p-\beta_i)} - \frac{B_0 \prod^{n/2-1}(p-\beta_i)}{\prod^{n/2}(p-\alpha_i)} \right\}$$

which may be further rearranged to the form

$$\frac{Z_{21}}{K} = \frac{\dfrac{A_1(p)}{p \prod^{n/2-1}(p-\beta_i)} \cdot \dfrac{A_2(p)}{\prod^{n/2}(p-\alpha_i)}}{\dfrac{\prod^{n/2}(p-\alpha_i)}{\prod^{n/2-1}(p-\beta_i)} - \dfrac{B_0 \prod^{n/2-1}(p-\beta_i)}{\prod^{n/2}(p-\alpha_i)}}$$

Comparison with (13) leads to

$$\left.\begin{aligned} kz_{11B} &= \frac{\prod^{n/2}(p-\alpha_i)}{p \prod^{n/2-1}(p-\beta_i)} & z_{22A} &= \frac{B_0 \prod^{n/2-1}(p-\beta_i)}{\prod^{n/2}(p-\alpha_i)} \\ z_{21B} &= \frac{A_1(p)}{p \prod^{n/2-1}(p-\beta_i)} & z_{21A} &= \frac{A_2(p)}{\prod^{n/2}(p-\alpha_i)} \end{aligned}\right\} \text{.....(15)}$$

The conversion constant, k, is set equal to one in all further calculations. Should another value be used it is sufficient to divide all the resistors and multiply all the capacitors of the B-circuit by k.

A substantial part of the whole synthesis is the realization of transmission zeros, in our case lying on the imaginary axis. Such zeros may be realized by a twin-T circuit, whose principal connection is shown in Fig. 7. The impedance matrix of this circuit, with transmission zeros in the points $p = j\Omega$, is

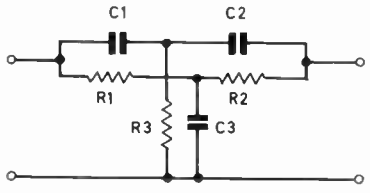


Fig. 7. Twin-T network.

$$[Z] = \begin{bmatrix} \dfrac{K_1 p^2 + \dfrac{p}{C_1} + K_1 \Omega^2}{p(p+\gamma)} & \dfrac{K_1(p^2+\Omega^2)}{p(p+\gamma)} \\ \dfrac{K_1(p^2+\Omega^2)}{p(p+\gamma)} & \dfrac{K_1 p^2 + \dfrac{a}{C_1} p + K_1 \Omega^2}{p(p+\gamma)} \end{bmatrix} \text{.....(16)}$$

the values of the resistors and capacitors being successively given by the equations

$$\left.\begin{aligned} R_1 &= \frac{1}{C_1 \gamma} & R_3 &= \frac{a}{\rho C_1 (1+a)} \\ \rho &= \Omega^2/\gamma & C_3 &= \frac{1+a}{a\rho R_1} \\ v &= C_1(\rho+\gamma) & C_2 &= C_1/a \\ a &= \frac{K_1 v}{1-K_1 v} & R_2 &= aR_1 \end{aligned}\right\} \text{.....(17)}$$

For the analysis of the influence of circuit elements on the whole transfer impedance it is necessary to know the whole unsimplified impedance matrix of the twin-T. Easy but rather tedious calculations give

$$z_{11} = \frac{p^3 M_1 + p^2 M_2 + p M_3 + 1}{p^3 N_1 + p^2 N_2 + p N_3} \qquad \text{.....(18)}$$

$$z_{21} = \frac{p^3 P_1 + p^2 P_2 + p P_3 + 1}{p^3 N_1 + p^2 N_2 + p N_3} \qquad \text{.....(19)}$$

where

$$\begin{aligned} M_1 &= R_1 R_2 R_3 C_1 C_2 C_3 \\ M_2 &= R_1 C_3 (C_1 R_3 + C_2 R_2 + C_2 R_3) + \\ &\quad + C_1 C_2 R_3 (R_1 + R_2) \\ M_3 &= R_1 C_2 + R_2 C_2 + R_1 C_3 + R_3 (C_1 + C_2) \\ N_1 &= C_1 C_2 C_3 (R_1 R_2 + R_1 R_3 + R_2 R_3) \\ N_2 &= R_1 C_1 (C_2 + C_3) + R_2 C_2 (C_1 + C_3) + \\ &\quad + R_3 C_3 (C_1 + C_2) \\ N_3 &= C_1 + C_2 + C_3 \\ P_1 &= M_1 \\ P_2 &= C_1 C_2 R_3 (R_1 + R_2) \\ P_3 &= R_3 (C_1 + C_2) \end{aligned}$$

Table 2

Data of the filter of sixth-degree. The significance of Ω_1, Ω_2 is shown in Fig. 5, the element values are valid for Fig. 8.

	dB	Ω_{3dB}	Ω_{a_s}	Ω_1	Ω_2	R_1	$C_1\times10^{-2}$	R_2	$C_2\times10^{-2}$	R_3	$C_3\times10^{-1}$	R_4
1	22	1·527	3·382	3·764	6·406	139·6	0·1606	2·916	1·801	6·299	0·3562	1·875
2	26	1·631	3·857	4·198	6·808	129·6	0·1730	2·486	1·657	6·073	0·3694	2·100
3	30	1·729	4·331	4·644	7·266	120·5	0·1861	2·102	1·510	5·861	0·3828	2·312
4	34	1·820	4·810	5·105	7·783	112·4	0·1996	1·768	1·364	5·665	0·3960	2·509
5	38	1·905	5·299	5·586	8·357	105·3	0·2130	1·483	1·223	5·488	0·4088	2·685
6	42	1·985	5·804	6·090	8·990	99·24	0·2261	1·243	1·089	5·331	0·4208	2·842
7	46	2·060	6·331	6·623	9·683	94·04	0·2386	1·042	0·9656	5·193	0·4320	2·980
8	50	2·129	6·888	7·191	10·44	89·62	0·2503	0·8749	0·8514	5·073	0·4422	3·100
9	54	2·193	7·480	7·797	11·26	85·89	0·2612	0·7358	0·7481	4·969	0·4515	3·204
10	58	2·251	8·111	8·447	12·15	82·72	0·2712	0·6199	0·6551	4·879	0·4598	3·294
11	62	2·304	8·787	9·144	13·11	80·05	0·2802	0·5235	0·5722	4·802	0·4672	3·371
12	66	2·352	9·513	9·894	14·15	77·79	0·2884	0·4428	0·4985	4·736	0·4737	3·437
13	70	2·395	10·29	10·70	15·27	75·87	0·2957	0·3754	0·4336	4·679	0·4795	3·494
14	74	2·433	11·13	11·57	16·49	74·23	0·3022	0·3183	0·3760	4·630	0·4846	3·544
15	78	2·467	12·04	12·51	17·80	72·85	0·3080	0·2705	0·3258	4·588	0·4890	3·586
16	82	2·497	13·01	13·52	19·22	71·67	0·3130	0·2301	0·2818	4·551	0·4929	3·622
17	86	2·524	14·06	14·61	20·75	70·66	0·3175	0·1959	0·2435	4·520	0·4963	3·653
18	90	2·547	15·20	15·78	22·41	69·80	0·3214	0·1669	0·2101	4·494	0·4993	3·680
19	94	2·588	16·42	17·05	24·20	69·07	0·3248	0·1424	0·1812	4·471	0·5018	3·702
20	98	2·585	17·74	18·42	26·12	68·45	0·3277	0·1216	0·1562	4·451	0·5040	3·722

Table 2 (contd.)

	dB	R_5	$C_5\times10^{-1}$	R_6	C_7	R_8	R_9	$C_9\times10^{-1}$	R_{10}	$C_{10}\times10^{-2}$	R_{11}	$C_{11}\times10^{-2}$
1	22	3·300	0·6241	1·000	0·1107	0·7872	5·871	0·4077	6·910	5·253	126·7	0·1888
2	26	,,	,,	,,	,,	1·394	5·264	0·4547	4·954	4·736	101·3	0·2362
3	30	,,	,,	,,	,,	1·849	4·808	0·4978	3·679	4·257	84·16	0·2844
4	34	,,	,,	,,	,,	2·200	4·458	0·5370	2·810	3·817	72·00	0·3324
5	38	,,	,,	,,	,,	2·477	4·180	0·5726	2·192	3·414	63·06	0·3796
6	42	,,	,,	,,	,,	2·701	3·957	0·6049	1·739	3·045	56·29	0·4252
7	46	,,	,,	,,	,,	2·883	3·774	0·6342	1·398	2·709	51·04	0·4690
8	50	,,	,,	,,	,,	3·035	3·623	0·6607	1·134	2·401	46·87	0·5107
9	54	,,	,,	,,	,,	3·162	3·496	0·6847	0·9287	2·122	43·53	0·5499
10	58	,,	,,	,,	,,	3·269	3·389	0·7063	0·7652	1·870	40·81	0·5866
11	62	,,	,,	,,	,,	3·359	3·299	0·7256	0·6342	1·643	38·58	0·6205
12	66	,,	,,	,,	,,	3·435	3·222	0·7428	0·5280	1·440	36·73	0·6516
13	70	,,	,,	,,	,,	3·500	3·157	0·7581	0·4416	1·258	35·20	0·6800
14	74	,,	,,	,,	,,	3·556	3·102	0·7716	0·3704	1·097	33·92	0·7056
15	78	,,	,,	,,	,,	3·603	3·055	0·7836	0·3115	0·9548	32·85	0·7287
16	82	,,	,,	,,	,,	3·643	3·014	0·7941	0·2629	0·8294	31·94	0·7493
17	86	,,	,,	,,	,,	3·678	2·980	0·8033	0·2223	0·7193	31·18	0·7676
18	90	,,	,,	,,	,,	3·707	2·950	0·8113	0·1885	0·6233	30·54	0·7837
19	94	,,	,,	,,	,,	3·733	2·925	0·8183	0·1599	0·5390	30·00	0·7979
20	98	,,	,,	,,	,,	3·754	2·903	0·8244	0·1359	0·4655	29·53	0·8104

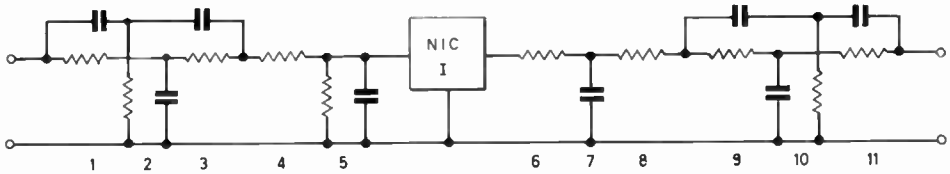


Fig. 8. Connection of the filter of sixth-degree. The element values are given in Table 2.

Note: In this Figure and in Fig. 9, the numerals at the foot of the filter circuit diagram indicate the component reference numbers of the resistors and/or capacitors immediately above.

Let us consider the synthesis of the filter of the sixth-degree. The desired transfer impedance has the form

$$Z_{21} = \frac{(p^2+\Omega_1^2)(p^2+\Omega_2^2)}{\prod_{i=1}^{6}(p-p_i)} \qquad \ldots\ldots(20)$$

where p_i are the poles from Table 1 and Ω_1, Ω_2 are transmission zeros, as given in Table 2. The synthesis of both circuits A and B—one pair of the zeros being considered in each—can be carried out by a simplified procedure. It consists of the separation of resistors and capacitors in such a way that the final impedance has the form of z_{11} in eqn. (16); this is then realized by a twin-T, thus realizing the transmission zeros at the same time.

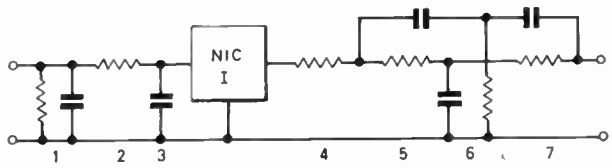


Fig. 9. Connection of the filter of fourth-degree. The element values to the right of the converter are given in Table 3. The element values to the left are the same for all attenuations:

$R_1 = 0{\cdot}2678$ $\quad R_2 = 1{\cdot}081$

$C_1 = 0{\cdot}6676$ $\quad C_3 = 0{\cdot}2251$

Table 3

Data of the filter of fourth-degree. The element values are valid for Fig. 9.

	dB	Ω_{3dB}	Ω_{a_s}	Ω_1	R_4	C_5	R_5	$C_6\times10^{-2}$	$R_6\times10^{-2}$	$C_7\times10^{-1}$	R_7
1	18	1·471	3·045	3·576	0·1986	0·1514	2·122	13·07	140·7	0·2136	15·04
2	22	1·573	3·558	4·075	0·3829	0·1658	1·938	11·17	97·66	0·2596	12·38
3	26	1·662	4·098	4·624	0·5207	0·1785	1·800	9·459	69·61	0·3045	10·55
4	30	1·740	4·677	5·231	0·6255	0·1895	1·695	7·931	50·69	0·3469	9·264
5	34	1·807	5·309	5·905	0·7061	0·1990	1·615	6·595	37·54	0·3860	8·327
6	38	1·863	6·005	6·655	0·7686	0·2070	1·552	5·443	28·19	0·4210	7·635
7	42	1·910	6·777	7·493	0·8174	0·2137	1·503	4·463	21·39	0·4517	7·115
8	46	1·949	7·636	8·429	0·8557	0·2193	1·465	3·639	16·38	0·4783	6·720
9	50	1·980	8·595	9·476	0·8858	0·2239	1·435	2·953	12·64	0·5009	6·417
10	54	2·006	9·667	10·65	0·9096	0·2277	1·411	2·386	9·805	0·5198	6·183
11	58	2·027	10·87	11·96	0·9283	0·2307	1·392	1·923	7·648	0·5355	6·001
12	62	2·044	12·21	13·44	0·9432	0·2332	1·377	1·543	5·976	0·5486	5·859
13	66	2·057	13·71	15·09	0·9549	0·2352	1·366	1·237	4·691	0·5592	5·748
14	70	2·068	15·40	16·94	0·9642	0·2369	1·356	0·9903	3·691	0·5678	5·661
15	74	2·077	17·29	19·01	0·9716	0·2381	1·349	0·7917	2·911	0·5747	5·592
16	78	2·084	19·41	21·34	0·9774	0·2392	1·343	0·6317	2·297	0·5804	5·538
17	82	2·089	21·79	23·95	0·9821	0·2400	1·339	0·5036	1·816	0·5849	5·495
18	86	2·093	24·46	26·88	0·9858	0·2407	1·335	0·4012	1·437	0·5885	5·461
19	90	2·097	27·45	30·16	0·9887	0·2412	1·332	0·3195	1·138	0·5914	5·435
20	94	2·100	30·80	33·85	0·9910	0·2416	1·330	0·2542	0·9018	0·5937	5·413
21	98	2·102	34·57	37·98	0·9928	0·2420	1·328	0·2023	0·7151	0·5956	5·397

Let us consider only the circuit A. Its input admittance, seen from the converter, is

$$y_1 = \frac{1}{z_{22A}} = \frac{p^3 + A_1 p^2 + A_2 p + A_3}{B_1 p^2 + B_2 p + B_3}$$

The admittance of a capacitance p/B_1 is realized first. Then, such a conductance G is subtracted from the remainder, which makes the absolute part of the numerator equal to zero. The inverse of the resulting admittance has the form

$$z_2 = \frac{A'_1 p^2 + A'_2 p + A'_3}{p(p+\gamma)}$$

which resembles the open circuit impedance z_{11} of the twin-T, only the constants A'_1 and A'_3 do not have the necessary values of eqn. (16). It is therefore necessary to subtract from this impedance such a constant (series resistor R), to make $A'_1 - R$ equal to A'_3/Ω^2. Then the whole remaining impedance may be realized by the twin-T according to eqns. (16) and (17). The circuit B would be synthesized similarly.

The above-described procedure is very simple and no residues have to be computed. If, however, the transmission zero is too close to the pass-band, it may happen that some or more elements are negative, so that such a case is unrealizable by the given procedure. This actually happened for e^{a_s} equal to 18 dB or less. Because these attenuations are very small and in practice unlikely to be used, no attempt has been made to realize them by another configuration of the circuit.

The results for the sixth-degree are given in Table 2 and the circuit is shown in Fig. 8. The fourth-degree, whose synthesis is simpler and not described in this paper, is given in Table 3 and Fig. 9. In the fourth-degree filter only one twin-T is used, so that the circuit A remains the same for all attenuations and its values are shown in Fig. 3.

5. Influence of Changes in the Converter

The influence of a change of the n.i.c. on the whole characteristic is usually considered by using the so-called sensitivity concept, which determines the properties in the differential vicinity of given values. The sensitivity of function F to changes of a certain parameter q is defined as a ratio

$$S = \frac{\mathrm{d}F}{F} \Big/ \frac{\mathrm{d}q}{q}$$

Here the parameter q may be either the constant of the n.i.c. or any other parameter, as for instance the position of a pole, transmission zero, some particular resistor and so on. Practically the differentials are substituted by differences

$$S = \frac{\Delta F}{F} \Big/ \frac{\Delta q}{q}$$

and the difference of the parameter is usually chosen, for instance 1%. This means, however, that we tacitly anticipate that such a change is so small, that we are still in the region of linear approximation—a presumption, which may be entirely wrong, especially at more complicated systems where such a change may even cause oscillations.

Because of this failure of the sensitivity concept

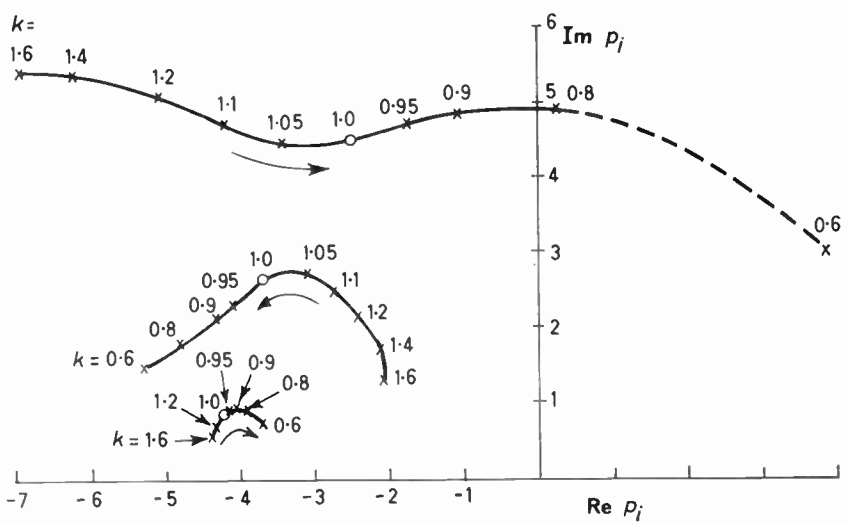


Fig 10. The movement of the poles in relation to the change of k.

more precise methods are used in this paper. As we are dealing with an active filter in which greater changes of the conversion constant may cause oscillations, it is advisable to find such a value of k, for which would make some pole of the transfer impedance fall on the imaginary axis. This condition will be fulfilled, if the denominator in eqn. (13) is put equal to zero for $p = j\omega$, that is

$$k\prod^{n/2}(p-\alpha_i)^2 - B_0 p\prod^{n/2-1}(p-\beta_i)^2 = 0\big|_{p=j\omega} \quad \text{.....(21)}$$

If the polynomial in eqn. (21) is rewritten for $n = 6$ and the coefficients are designated as follows

$$(p^6 + A_1 p^5 + A_2 p^4 + A_3 p^3 + A_4 p^2 + A_5 p + A_6) - \\ - B_0(p^5 + B_1 p^4 + B_2 p^3 + B_3 p^2 + B_4 p) = 0\big|_{p=j\omega}$$

then the equation may be divided into a real and an imaginary part, leading to two equations as follows:

$$\left.\begin{aligned} f_1(k,\omega) &= -k\omega^6 + (kA_2 - B_1)\omega^4 - \\ &\quad -(kA_4 - B_3)\omega^2 + kA_6 = 0 \\ f_2(k,\omega) &= (kA_1 - B_0)\omega^4 - (kA_3 - B_2)\omega^2 + \\ &\quad +(kA_5 - B_4) = 0 \end{aligned}\right\} \quad \text{.....(22)}$$

which must be simultaneously fulfilled. The second equation was divided by ω, so that the trivial solution $\omega = k = 0$ is removed. It may be shown that substitution of hp_i instead of p_i (h real constant) leads to such a change of the constants A_i, B_i in (22), that the solution for k remains unchanged. The change of k, which causes oscillations of the filter, is therefore independent of the width of the pass-band, so that eqn. (22) may be used for any filter of sixth degree and a comparison of various types of filters is much simplified.

For the filters described in this paper the solution is

$$k = 0{\cdot}81682$$
$$\Omega_{osc} = 4{\cdot}9422$$

so that an 18% change of k does not lead to oscillations yet. Using the same program, the stability of the filter described by Antreich and Gleissner[3] was checked, with the result

$$k = 0{\cdot}99377$$
$$\Omega_{osc} = 0{\cdot}8966$$

The frequency Ω_{osc} is here different, because of a different normalization of the pass-band. The oscillations start, however, when k is changed only by 0·6%.

The encouraging result of these first solutions led to the decision to study the influences of the converter changes more thoroughly. Several values for k were chosen and the trajectories of the pole movements were computed. The result is shown in Fig. 10. It is clear that the oscillations of this system of sixth degree may start only with the diminution of k and that the pole, which causes oscillations, is the one that is most remote from the origin.

Although the changes of the pole positions are quite considerable, they have, however, small influence on the characteristics of the filter. The characteristics were computed for k = 0·9, 1, 1·1 and for the filter 6 from Table 2 (42 dB minimum attenuation). The amplitude characteristic is drawn in Fig. 11 on a logarithmic scale. The step responses are shown in Fig. 12. In addition, a comparison with the filter described by Antreich and Gleissner[3] (Schüssler–Jess filters) was computed. The poles and zeros of this filter, transformed to the same bandwidth for a decrease of 3 dB, are

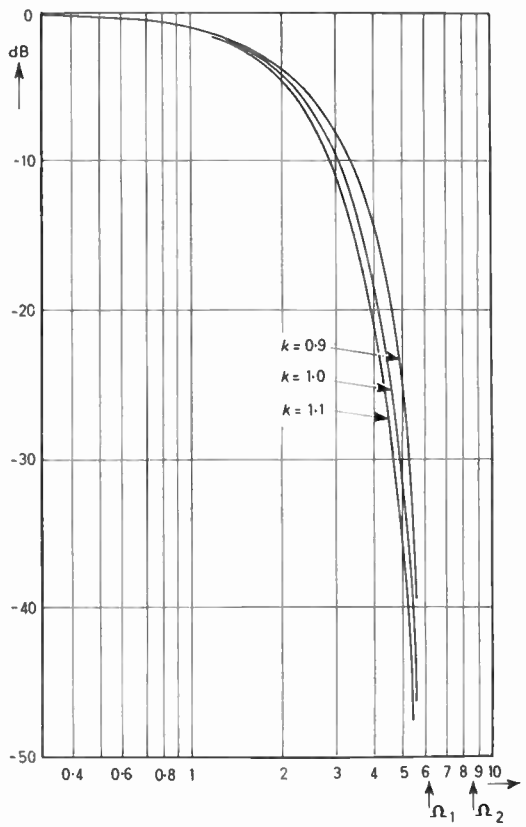


Fig. 11. The dependence of the amplitude characteristic on the change in the values of k.

poles	zeros
−2·335 555 88 ± j 0·990 889 20	0 ± j 4·854 343 2
−1·742 326 6 ± j 2·971 296 72	0 ± j 6·161 582
−0·217 208 32 ± j 4·360 483 68	

.....(23)

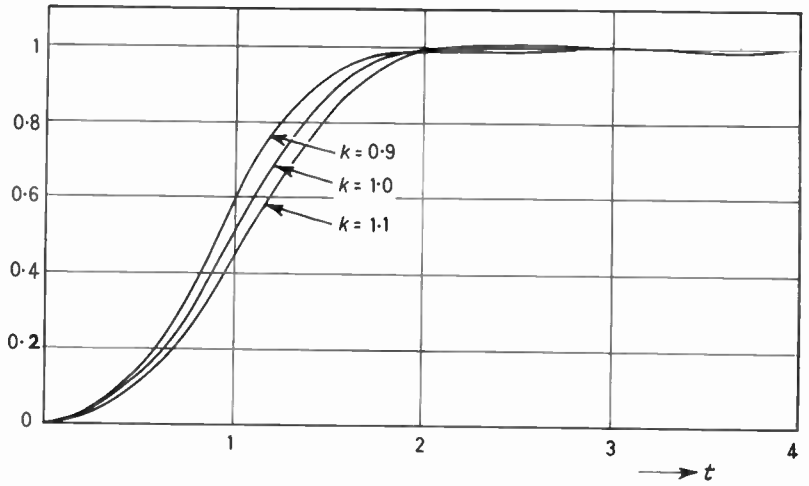


Fig. 12. The step-response of the filter for different values of k.

(The catalogue[4] of these filters was not available to the authors of this paper. The compared filters differ in the stop-band by 2 dB, the attenuation of eqn. (23) being 40 dB, whereas that of the filter considered is 42 dB.) The amplitude curves of both filters are drawn in Fig. 13. It is clear that they differ only negligibly down to −20 dB. From this point the selectivity of the filter (eqn. (23)) is greater and is given by the transmission zeros, shown in Fig. 13 by the arrows. The step-response of both filters is shown in Fig. 14. For passive filters, more thorough analysis of both classes of filters would be necessary; for active filters the design used by Antreich and Gleissner[3] is obviously useless. The filters proposed in this paper should cause no trouble because of the results described in Section 2.

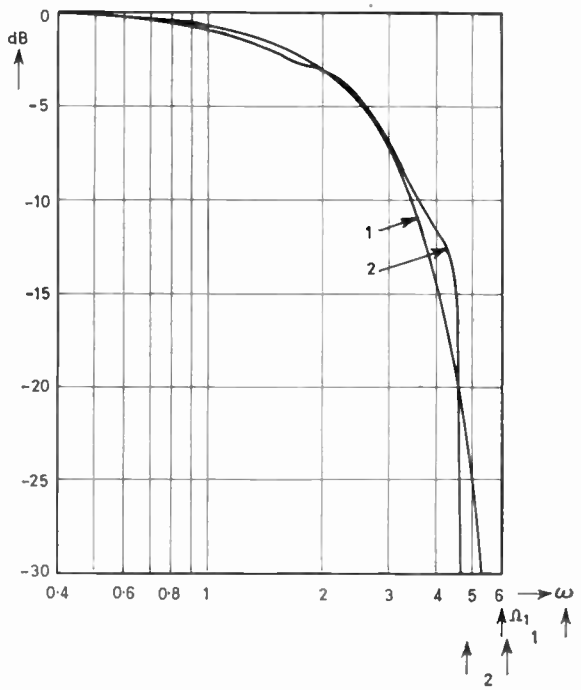


Fig. 13. Comparison of the amplitude characteristic of the proposed filter (curve 1) and of the filter described in Ref. 3 (curve 2), transformed to the same bandwidth for 3 dB.

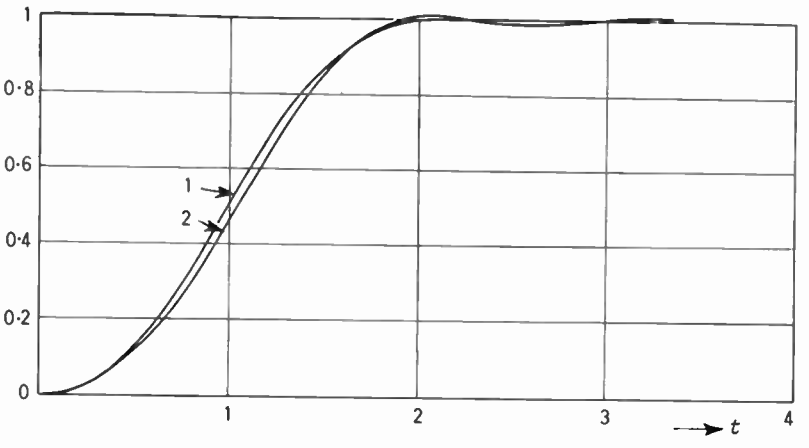


Fig. 14. Comparison of the step-response of the proposed filter (curve 1) with the filter described in Ref. 3 (curve 2), transformed to the same bandwidth for 3 dB.

6. Influence of Changes in the Passive Elements

The results of previous works show that the instability in active filters is mainly caused by the instability of the converter. Nevertheless it is useful to have some data of the influence of different component parts of the filter. This influence was considered by computing the amplitude characteristic. The input impedance of the unsimplified twin-T, as given in eqn. (18), was used. For the circuit A the impedance is

$$z_{22A} = \frac{1}{pC_5 + \dfrac{1}{R_5} + \dfrac{1}{R_4 + z_{11}}} = \frac{A_1(p)}{A_2(p)}$$

whereas for the circuit B

$$z_{11B} = R_6 + \frac{1}{pC_7 + \dfrac{1}{R_8 + z_{11}}} = \frac{B_1(p)}{B_2(p)}$$

holds. It is necessary that the indices of the twin-T's in the circuits A or B be changed in accordance with Fig. 8. The transfer impedance is then given by

$$Z_{21} = \frac{z_{21A} \cdot z_{21B}}{\dfrac{B_1(p)}{B_2(p)} - \dfrac{1}{k}\dfrac{A_1(p)}{A_2(p)}}$$

The denominator $B_2(p)A_2(p)$ may be cancelled by the denominator of the product $z_{21A} \cdot z_{21B}$. Only the two numerators of eqn. (19) (with appropriate indices in accordance with Fig. 8) remain in the numerator of Z_{21}. The whole transfer impedance is then

$$Z_{21} = \frac{[p^3P_1' + p^2P_2' + pP_3' + 1][p^3P_1'' + p^2P_2'' + pP_3'' + 1]}{B_1(p)A_2(p) - \dfrac{1}{k}A_1(p)B_2(p)}$$

The amplitude response of this impedance was computed. It was found, that a 10% change of those elements, which do not determine the position of the transmission zeros, has smaller influence on the characteristic than the same relative change of the converter constant. If some element of the twin-T is changed, then the imaginary zero changes to a complex zero, thus leading to the change of the characteristic. The influence of a ± 10% change in the value of the capacitor C_{10} is shown in Fig. 15. In both cases the amplitude characteristic is essentially the same (curve 2).

7. Conclusion

Active filters with poles chosen for maximally flat group delay and with zeros on the imaginary axis have been considered. The transmission zeros are chosen in such a way that the behaviour in the stop-band has Chebyschev character. The poles of such filters are comparatively remote from the imaginary axis. In agreement with the assumption, these filters are little sensitive to changes of filter elements. It was shown that filters of sixth degree would oscillate only if the n.i.c. changed its constant by more than 18%, whereas experimental results show that this constant would not change by more than ±5% in usual conditions. Comparison with another filter[3] of similar properties shows that those proposed in this paper are more suitable. The calculations, carried out by a computer, showed that a 10% change of various elements does not cause considerable changes of the characteristics, as long as no transmission zero is influenced by this change.

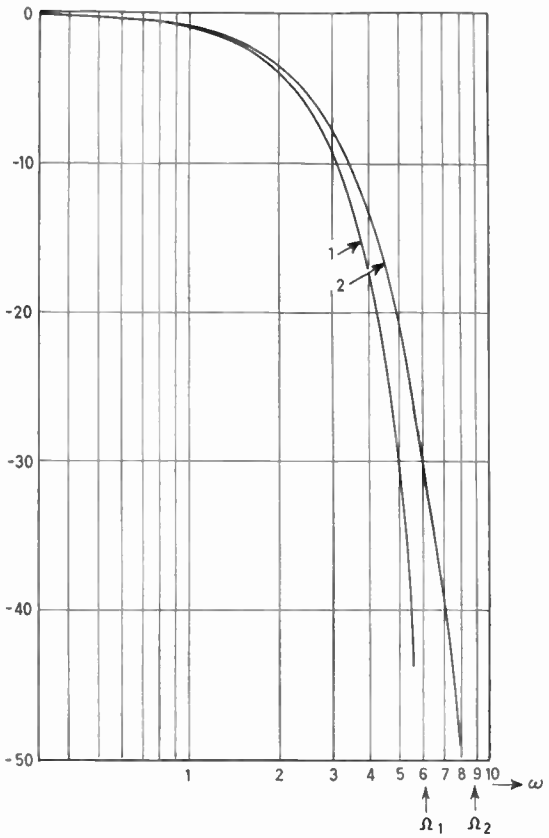


Fig. 15. Influence of a ±10% change of C_{10} on the amplitude characteristic of the filter.

Curve 1: the desired response;

Curve 2: the responses with C_{10} changed by ±10%.

8. References

1. A. G. J. Holt and F. W. Stephenson, 'The effects of error in the element values and the converters transfer characteristic on the response of active filters', *The Radio and Electronic Engineer*, **26**, No. 6, pp. 449–57, December 1963.
2. D. J. Storey and W. J. Cullyer, 'Active low-pass linear-phase filters for pulse transmission', *Proc. Instn Elect. Engrs*, **112**, No. 4, pp. 661–8, April 1965.

3. K. Antreich and E. Gleissner, 'Über die Realisierung von Impulsfiltern durch aktive RC-netzwerke', *Archiv der Elektrischen Übertragung*, **19**, pp. 309–316, June 1965.
4. J. Jess, 'Katalog normierter Tiefpassübertragungsfunktionen mit Tschebyscheffverhalten der Impulsantwort und der Dämpfung', Forschungsbericht Nr. 1329 des Landes Nordrhein-Westfalen. (Westdeutscher Verlag, Köln und Opladen, 1964.)
5. J. Jess and H. W. Schüssler, 'A class of pulse-forming networks', *Trans. Inst. Elect. Electronics Engrs on Circuit Theory*, **CT-12**, pp. 296–9, June 1965 (Letters).
6. J. Jess and H. W. Schüssler, 'On the design of pulse-forming networks', *Trans. I.E.E.E.*, **CT-12**, pp. 296–9, June 1965 (Letters).
7. J. Braun, 'Equivalent nic networks with nullators and norators', *Trans. I.E.E.E.*, **CT-12**, pp. 441–2, September 1965.
8. T. Yanagisawa, 'RC active networks using current inversion type negative impedance converters, *Trans. I.R.E.*, **CT-4**, pp. 140–4, September 1957.
9. M. Nagata, 'A simple negative-impedance circuit with no internal bias supplies and good linearity', *Trans. I.E.E.E.*, **CT-12**, p. 443, September 1965.
10. W. E. Thomson, 'Delay networks having maximally flat frequency characteristic', *Proc. I.E.E.*, **96**, pp. 487–90, August 1949.
11. K. H. Feistel and R. Unbehauen, 'Tiefpässe mit Tscheby-scheffcharakter der Betriebsdämpfung im Sperrbereich und maximal Geebneter Laufzeit', *Frequenz*, **19**, pp. 265–82, August 1965.
12. J. G. Linvill, 'RC active filters', *Proc. I.R.E.*, **42**, pp. 555–64, March 1954.
13. J. M. Horowitz, 'Optimization of negative-impedance conversion methods of active RC synthesis', *Trans. I.R.E.*, **CT-6**, pp. 296–303, September 1959.

Manuscript received by the Institution on 21st December 1966. (*Paper No. 1118.*)

STANDARD FREQUENCY TRANSMISSIONS

(Communication from the National Physical Laboratory)

Deviations, in parts in 10^{10}, from nominal frequency for **April 1967**

April 1967	24-hour mean centred on 0300 U.T.			April 1967	24-hour mean centred on 0300 U.T.		
	GBR 16 kHz	MSF 60 kHz	Droitwich 200 kHz		GBR 16 kHz	MSF 60 kHz	Droitwich 200 kHz
1	— 300·0	— 299·7	+ 0·7	16	— 300·0	+ 0·1	0
2	NM	NM	+ 0·1	17	— 299·7	+ 0·1	— 0·9
3	— 299·5	NM	+ 0·4	18	— 299·6	+ 0·4	— 0·1
4	— 299·5	+ 0·2	— 0·4	19	— 300·0	+ 0·1	0
5	— 299·6	+ 0·3	— 1·0	20	— 299·9	+ 0·2	+ 0·1
6	— 299·8	+ 0·3	— 0·8	21	— 299·7	+ 0·4	+ 0·2
7	— 299·8	+ 0·2	— 0·6	22	— 300·1	— 0·1	+ 0·4
8	— 299·8	NM	— 0·9	23	— 299·8	+ 0·2	+ 0·2
9	— 299·9	NM	— 0·9	24	— 299·9	+ 0·1	0
10	— 299·7	NM	— 0·5	25	— 299·9	+ 0·2	— 0·1
11	— 299·7	NM	— 0·3	26	— 299·9	+ 0·2	— 0·3
12	— 299·7	NM	— 0·1	27	— 299·7†	0	— 0·2
13	— 299·9	+ 0·2	— 0·5	28	— 299·9	+ 0·1	0
14	— 299·8	+ 0·2	*	29	— 300·0	0	— 0·4
15	— 300·0	— 0·1	0	30	— 299·9	0	— 0·4

Nominal frequency corresponds to a value of 9 192 631 770·0 Hz for the caesium F,m (4,0)–F,m (3,0) transition at zero field.
NM = not measured
* Droitwich oscillator changed at 11.45 U.T.
† 6μ second phase change at 23.15 U.T. subtracted

Special Note Concerning MSF and GBR Transmissions

Since 1300 UT on 2nd April 1967 the MSF and GBR transmissions have been controlled by a rubidium gas cell frequency standard. From this date the MSF carrier frequencies have been operated without frequency offset.

Anomalous phase excursions occurred on the MSF 60 kHz transmissions between 11th–14th April due to work in progress at the transmitter.

U.D.C.621.396.677

The Radiation Characteristics of the Backfire Helical and Zigzag Antennae

By

S. C. LOH,
B.Sc., Ph.D., C.Eng.,
M.I.E.E., A.M.I.E.R.E., A.Inst.P.†

AND

JOHANNES JACOBSEN‡

Summary: In this paper, a type of highly directional, surface-wave antenna is described. This consists of a conventional endfire structure which is terminated by a large surface-wave plane reflector and is thus operated under the backfire condition. Two commonly used endfire structures, namely helical and zigzag, are tested. The improvement is achieved in the half-power beam width.

1. Introduction

The backfire antenna discussed by Ehrenspeck[1-3] is a new type of surface-wave antenna. It consists of an ordinary endfire structure terminated by a large surface-wave plane reflector. (The term surface-wave reflector was suggested by Ehrenspeck to distinguish the ordinary plane reflector in antenna array). Ehrenspeck has used the Yagi antenna as the endfire structure and has shown that the 'backfire' principle greatly increases the gain of an endfire antenna without adding to its length. He first showed experimentally that the backfire (Yagi) antenna could achieve a gain of 5 to 6 dB above that of an equal-length ordinary endfire (Yagi) antenna with a plane circular reflector of radius 2λ, terminated at the open end opposite to the feed and later with a modified surface-wave plane reflector,[5] a higher gain of 8 dB or more could easily be made available. Recently he investigated a short-backfire antenna[6] (half-wave-length long) and has obtained desirable radiation properties.

The present paper describes an experimental study of the radiation characteristics of the well-known helical and zigzag antennae under the backfire conditions.

2. The Principle of Backfire Antenna

If a large surface-wave plane reflector R is placed at the open end of an ordinary endfire antenna structure, A, shown in Fig. 1, a surface-wave launched at the feed F travels along the antenna structure until it strikes on the surface-wave plane reflector; it then travels back towards the feed and radiates into the free space in a direction opposite to that of normal endfire direction. Because of the reversed direction of radiation, the antenna is called the 'backfire' antenna.

The function of the large surface-wave plane reflector R used is not generally considered to be the same as that of plane reflector G usually placed a little distance behind the feed of an antenna array for the purpose of producing a radiation pattern with great directivity in the forward direction. In the latter case, it is suggested that the plane reflector causes interference between direct and reflected wave in the antenna structure, a maximum is achieved in the forward direction. The function of the surface-wave plane reflector on the backfire antenna proposed by Ehrenspeck is that the surface-wave plane reflector images the original endfire structure, setting up a standing wave along the structure. Owing to the fact that the energy of the surface wave impinging on the surface-wave plane reflector, if properly placed, arrives concentrated in a narrow beam, only a relatively small conducting surface (instead of infinite in size) is needed for imaging.

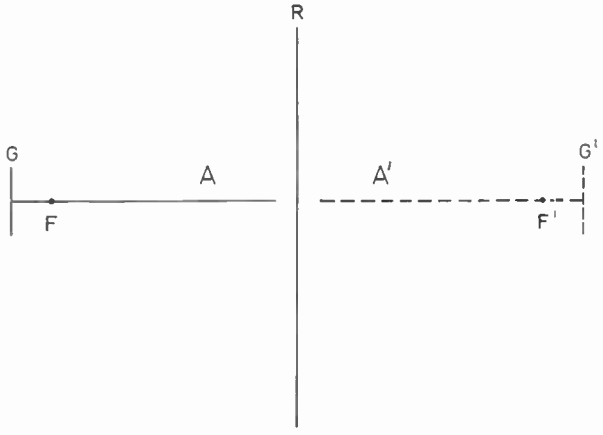


Fig. 1. Simplified version of backfire antenna.

† Formerly at the Laboratory of Electromagnetic Theory, The Technical University of Denmark, Lyngby, Denmark; now with the Physics Department, United College, The Chinese University of Hong Kong.

‡ Laboratory of Electromagnetic Theory, The Technical University of Denmark, Lyngby, Denmark.

The backfire antenna was first proposed by Ehrenspeck and the endfire antenna used is the Yagi. Though a detailed and thorough experimental study of the backfire (Yagi) antenna has been undertaken by Ehrenspeck with great success, and a qualitative approach to the design of the antenna suggested by Zucker[7] is available, a satisfactory theoretical solution to the problem is yet to be found.

3. Experimental Results

The backfire antennae were tested in the radio anechoic chamber in the Technical University of Denmark and the testing frequency was chosen to be in the 10 cm band (S-band) to allow a convenient antenna size.

3.1. *Backfire Helical Antenna*

An eight-turn backfire helical antenna shown in Fig. 2 was constructed, having a circular surface-wave plane reflector of diameter $D = 40$ cm with rims of three different widths, $W = 2{\cdot}5$ cm, 5 cm and 7·5 cm respectively. It has been found that the reflector with rim of 2·5 cm gave the best radiation patterns. The circular surface-wave plane reflector with rim was first used by Ehrenspeck[5] in his investigation of backfire antenna and it was found that an improved radiation pattern was possible. However in our study, we found that the rim has very little influence on the main lobe and only influences the side-lobe level slightly. The size of the plane reflector, which is relatively insensitive in the endfire condition, does influence the side-lobe level slightly and therefore circular plane reflector of various diameters, namely, $d_1 = 5$, 6, 7·5, 10, 12·5 cm were tested. The plane reflector of diameter 10 cm gave the best results. The distance L between the antenna open termination and the surface-wave plane reflector, was varied between 0·1 to 0·25 cm, and was found to have only slight effect on the operating frequency bandwidth and the half-power beam width, but it does influence the side-lobe level moderately.

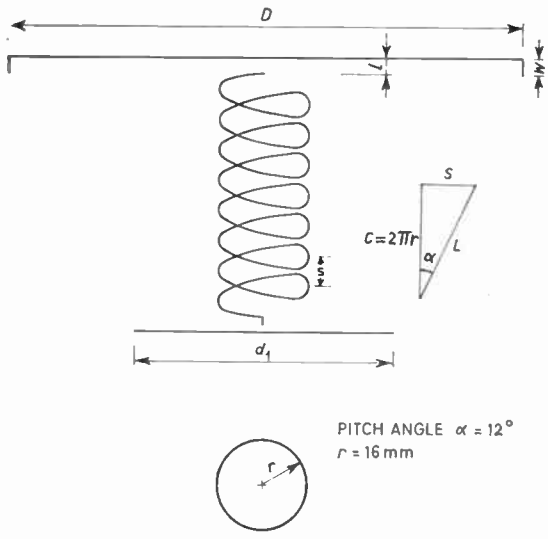


Fig. 2. Backfire helical antenna.

The radiation characteristics of the helical antennae operated under the backfire conditions shows marked improvements over those operated under the normal endfire conditions. The half-power beam width has improved from 40° minimum under the endfire condition to the maximum 17° for the backfire case. The variation of the half-power beam width against the frequency is shown in Fig. 3. As there is no

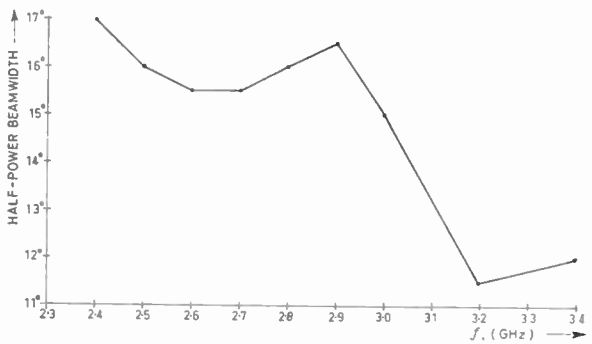


Fig. 3. Half-power beam width against frequency for backfire helical antenna.

significant difference between the E-plane and the H-plane patterns, for the backfire helical antenna, only the best radiation pattern of E-plane is shown in Fig. 4 at a frequency of 3·2 GHz. The new antenna can operate satisfactorily over the frequency range of 2·4 GHz to 3·4 GHz and the variation of worse side-lobe level over this frequency range is shown in Fig. 5. On the range of 2·4 GHz to 3·4 GHz (i.e. operating frequency range of 1·4 : 1) the following criteria are satisfied by the backfire helical antenna: (*a*) half-power beam width is less than 17°, (*b*) the side-lobe level is at least 5 dB down and (*c*) the directivity is more than 10 dB above the isotropic source. Figure 6 shows the variation of the back-lobe level in decibels against the frequency.

The power gain of the backfire helical antenna has been computed from the measured radiation pattern with the aid of an electronic computer. Figure 7 shows the variation of the gain over an isotropic source plotted against the frequency. This gain can readily be explained, as suggested by Zucker[7], on the basis that the surface-wave plane reflector, if assumed to be infinite, may be replaced by the image

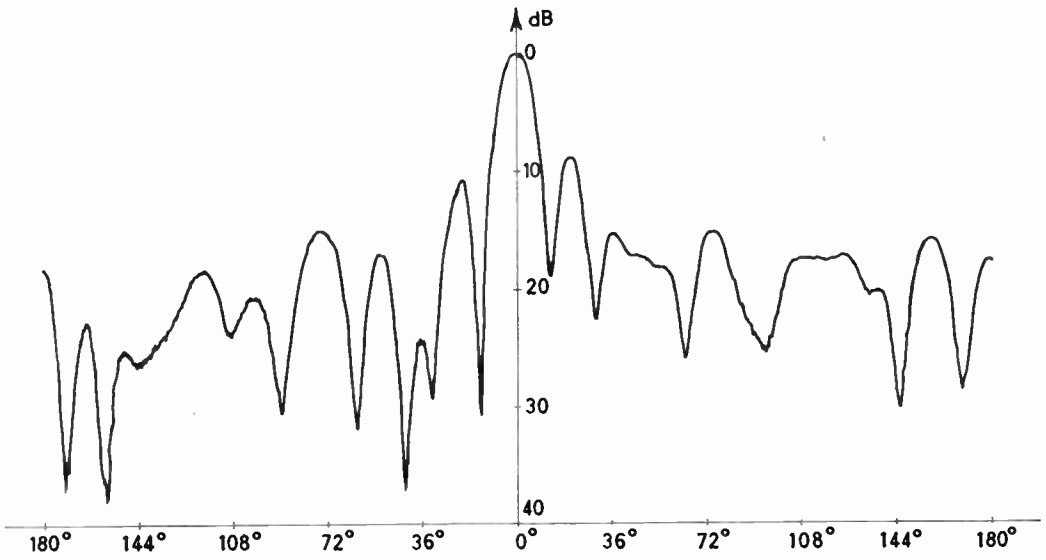


Fig. 4. Radiation pattern for backfire helical antenna.

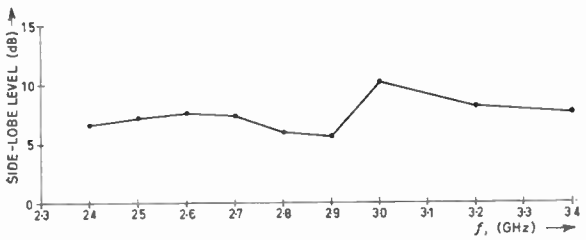


Fig. 5. Side-lobe level against frequency for backfire helical antenna.

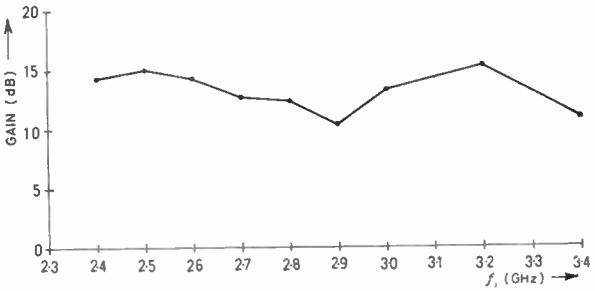


Fig. 7. Gain above isotropic source for backfire helical antenna.

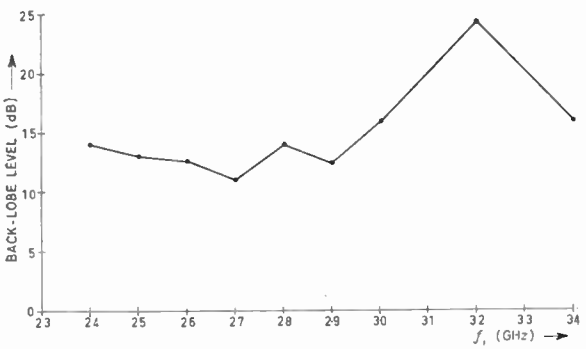


Fig. 6. Back-lobe level against frequency for backfire helical antenna.

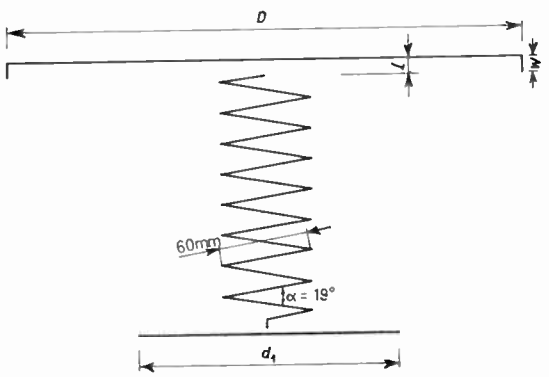


Fig. 8. Backfire zigzag antenna.

of A and G, with image feed F′ as shown in Fig. 1, thus the backfire antenna is equivalent to two endfire antennae, fed at F and F′ respectively.

3.2. *Backfire Zigzag Antenna*

An eight-turn zigzag antenna with dimensions as shown in Fig. 8 was constructed and tested. The best size for the surface-wave plane reflector is found to be the same as for the backfire helical antenna. The influence of the width of the rim on both the reflector and the distance L is insignificant in determining the half-power beam width and the operating frequency bandwidth except in the case of side-lobe level. Again it has been noted that the radiation characteristics of the backfire zigzag antenna compare favourably with those obtained when it is operated under the endfire condition. Both the operating frequency bandwidth and the half-power beam width are improved greatly and their properties are illustrated in Fig. 9. The best radiation patterns of a backfire zigzag antenna are shown in Fig. 10 for E- and H-polarizations respectively. The variation of the side-lobe level against the frequency is shown in Fig. 11 and Fig. 12 shows the variation of the back-lobe level against frequency. The operating frequency range is defined with the same criteria as for the case of backfire helical antenna.

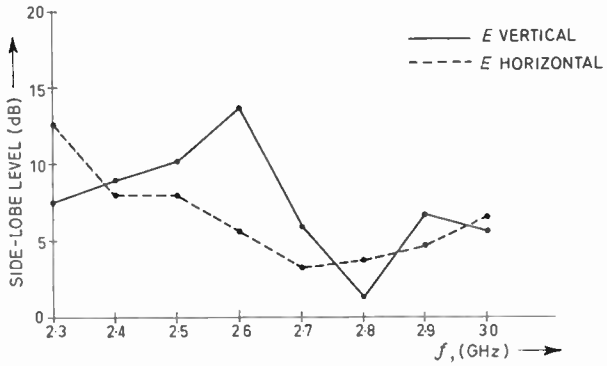


Fig. 11. Side-lobe level against frequency for backfire zigzag antenna.

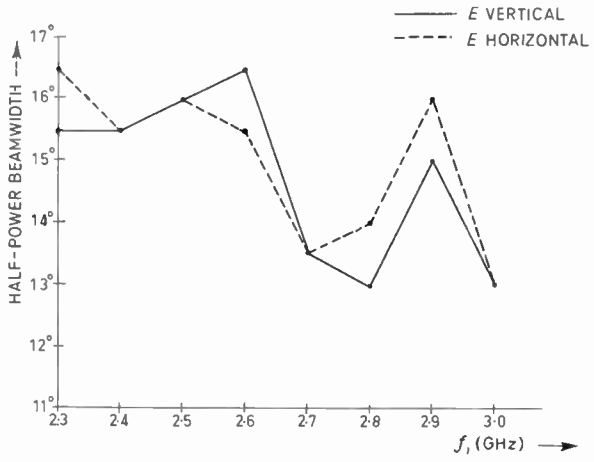


Fig. 9. Half-power beamwidth against frequency for backfire zigzag antenna.

As in the case of the helical antenna the power gain has been computed from the measured radiation patterns and Fig. 13 shows the variation of the power gain with frequency.

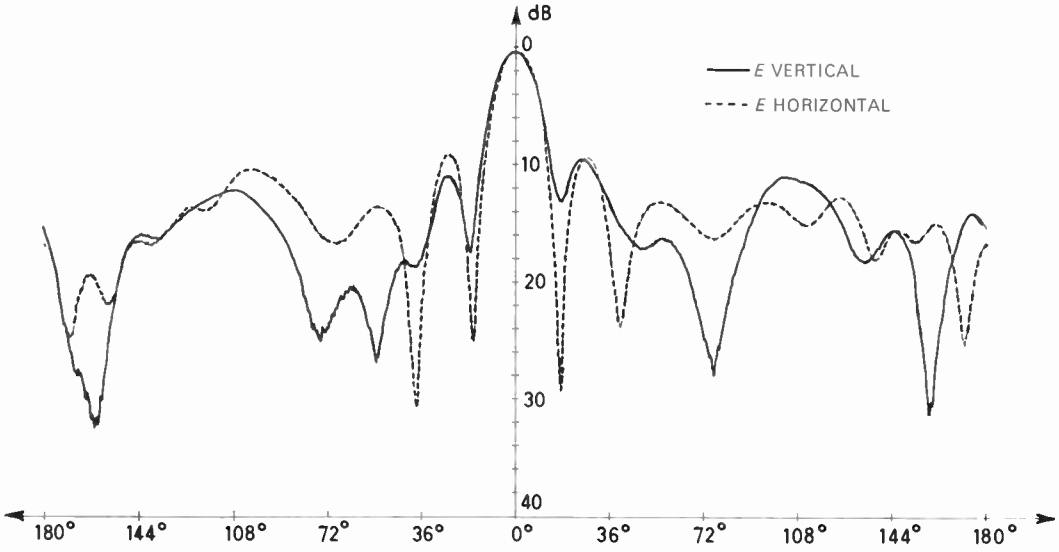


Fig. 10. Radiation pattern for backfire zigzag antenna.

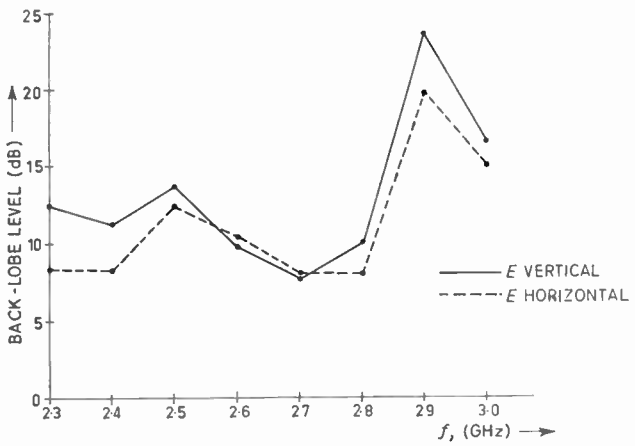


Fig. 12. Back-lobe level against frequency for backfire zigzag antenna.

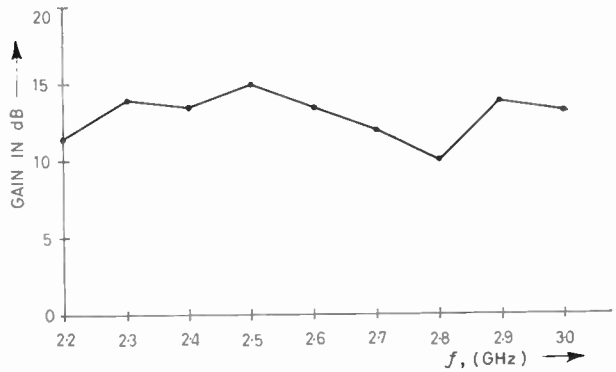


Fig. 13. Gain above isotropic source for backfire zigzag antenna.

4. Conclusions

The exact mechanism governing the radiation properties of the backfire antenna is not yet understood and no complete theory explaining the radiation from this antenna has yet been given; the theory of images suggested by Zucker may offer a first approximation to such a theory.

Experimentally, however, it has been found that the backfire antennae offer somewhat better performance than the similar endfire antennae and their radiation characteristics may be summarized as follows:

(1) The half-power beam width is greatly improved over the entire operating frequency range. This is true for all antennae tested.

(2) The operating frequency range for the backfire helical antenna is approximately 1·4 : 1 and for the backfire zigzag 1·13 : 1. The antenna may possibly be operated beyond the operating frequency range defined, but the endfire radiation is not well behaved, in fact multiple-lobed patterns often appeared.

(3) The side-lobe level for the backfire behaves roughly in the same way as in the case of endfire.

5. Acknowledgment

This work was carried out at the Laboratory of Electromagnetic Theory, the Technical University of Denmark and one of the authors (S.C.L.) is grateful to Prof. Lottrup Knudsen for the invitation to work in his laboratory. The authors wish to thank Prof. R. E. Kleinman for his interest and suggestions concerning the material in this paper.

6. References

1. H. W. Ehrenspeck, 'The backfire antenna, a new type of directional line source', *Proc. Inst. Radio Engrs*, 48, pp. 109–10, January 1960. (Letters.)
2. H. W. Ehrenspeck, 'The Backfire Antenna, a New Type of Directional Line Source', AFCRL Report No. 722, Air Force Cambridge Research Labs., Bedford, Mass., August 1961.
3. H. W. Ehrenspeck, 'Reflection antenna employing multiple director elements and multiple reflection of energy to effect increased gain', U.S. Patent No. 3, 122, 745, February 1964.
4. J. A. Strom and H. W. Ehrenspeck, 'Backfire Antennas for SHF, UHF and VHF Bands', AFCRL Report No. 63–114, Air Force Cambridge Research Labs., April 1963.
5. H. W. Ehrenspeck, 'The backfire antenna: new results' *Proc. I.E.E.E.*, 53, pp. 639–41, June 1965. (Letters.)
6. H. W. Ehrenspeck, 'The short backfire antenna', *Proc. I.E.E.E.*, 53, pp. 1138–40, August 1965. (Letters.)
7. F. J. Zucker, 'The backfire antenna: A qualitative approach to its design', *Proc. I.E.E.E.*, 53, pp. 746–7, July 1965. (Letters.)

Manuscript first received by the Institution on 1st June 1966 and in final form on 6th December 1966. (Paper No. 1119.)

U.D.C. 621.373

Calculation of Resonance Frequencies of Microwave Cavities Containing Electro-optic Crystals

By

M. J. P. MUSGRAVE, B.Sc., A.R.C.S.†

AND

H. PURSEY, B.Sc.(Eng.)†

Summary: Calculations appropriate to the design of microwave cavities containing uniaxial electro-optic crystals used in the modulation of optical beams are reported. Results for a series of crystals (KDP and isomorphs) embedded in various filler dielectrics are presented as graphs of a frequency parameter against ratio of crystal radius/cavity radius.

1. Introduction

Successful experiments in the modulation of light by microwave frequencies depend upon the attainment of a high electric field within a non-centrosymmetric crystal and the spatial matching of the phases of the microwave field and the optical beam.

Chen and Lee[1] have recently published some data, for a wave guide containing a crystal surrounded by another concentric dielectric material, in the form of graphs of the propagation frequency *versus* the effective microwave refractive index. They do not, however specify the phase-matching points on their plots and, furthermore, they treat KDP (KH_2PO_4) as an isotropic dielectric when, in fact, the ratio of the principal dielectric constants at microwave frequencies $\varepsilon_T/\varepsilon_L > 2$. Measurements of the electro-optic coefficient r_{63} for KDP have recently been described by Pursey, Merran and Musgrave[2] in which a resonant microwave cavity containing the crystal was phase-matched with the optical beam. The design of such a cavity is facilitated by the results presented in this note wherein the effect of the approximations invoked by Chen and Lee are considered in relation to a more correct form of the theory.

2. TM Modes for a Modulating Cavity

The form of the cavity is shown in Fig. 1. It consists of an optically uniaxial crystal, prepared in the shape of a circular cylinder with the optic axis parallel to the generators; the crystal is mounted concentrically with an isotropic dielectric material within the outer casing.

Thus we have a system specified by

$$\left.\begin{matrix}\varepsilon_L \text{ (axially)}\\ \varepsilon_T \text{ (radially)}\end{matrix}\right\} \text{ for } 0 \leqslant r \leqslant a$$

$$\varepsilon_D \text{ (isotropic) for } a \leqslant r \leqslant b$$

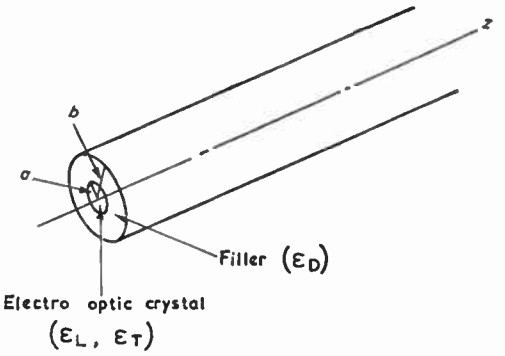


Fig. 1. Form of guide or cavity.
Crystal: radius a, dielectric constants, ε_L (axial), ε_T (radial).

We are here concerned with modes whose components of electric and magnetic field have the form

$$\left.\begin{matrix}[E_r, 0, E_z]\\ [0, H_\theta, 0]\end{matrix}\right\} \times \exp j(kz-\omega t):$$

the suffixes r, θ and z indicate radial, circumferential and axial components respectively.

Within $a \leqslant r \leqslant b$ (isotropic dielectric)

$$E_r = jk[AJ_1(\beta_0 r)+BY_1(\beta_0 r)]/\beta_0, \quad \text{.....(1)}$$

$$E_z = AJ_0(\beta_0 r)+BY_0(\beta_0 r), \quad \text{.....(2)}$$

where

$$\beta_0^2 = \frac{\varepsilon_D}{\varepsilon_0}\frac{\omega^2}{c^2} - k^2,$$

A and B are constants to be determined, ε_0 is the free-space permittivity.

Within $0 \leqslant r \leqslant a$ (crystal)

$$E_r' = \frac{jk\beta_L}{\beta_T^2} CJ_1(\beta_L r) \quad \text{.....(3)}$$

$$E_z' = CJ_0(\beta_L r) \quad \text{.....(4)}$$

where

$$\beta_L^2 = \frac{\varepsilon_L}{\varepsilon_0}\frac{\omega^2}{c^2} - \frac{\varepsilon_L}{\varepsilon_T}k^2, \qquad \beta_T^2 = \frac{\varepsilon_T}{\varepsilon_0}\frac{\omega^2}{c^2} - k^2$$

† Division of Molecular Science, National Physical Laboratory, Teddington, Middlesex.

and C is a further constant. The components of magnetic field have not been itemized since the boundary conditions in the radial components of electric field and the circumferential components of magnetic field are equivalent.

The ratios $A : B : C$ are found by substituting solutions (1) to (4) into the appropriate boundary conditions:

$$E_z(b) = 0 \qquad \text{at } r = b \qquad \text{.....(5)}$$

$$\left.\begin{aligned} \varepsilon_T E_r'(a) &= \varepsilon_D E_r(a) \\ E_z'(a) &= E_z(a) \end{aligned}\right\} \text{ at } r = a \qquad \begin{aligned} &\text{.....(6)} \\ &\text{.....(7)} \end{aligned}$$

We obtain

$$AJ_0(\beta_0 b) + BY_0(\beta_0 b) = 0 \qquad \text{.....(8)}$$

$$\frac{\varepsilon_D}{\beta_0}\left[AJ_1(\beta_0 a) + BY_1(\beta_0 a)\right] = C\frac{\varepsilon_T \beta_L}{\beta_T^2} J_1(\beta_L a) \qquad \text{.....(9)}$$

$$AJ_0(\beta_0 a) + BY_0(\beta_0 a) = CJ_0(\beta_L a) \qquad \text{.....(10)}$$

Elimination of C between eqns. (9) and (10) followed by elimination of $A : B$ between the resulting equation and eqn. (8) yields the dispersion equation in $\omega(k)$ which may be written as:

$$\begin{vmatrix} J_0(\beta_0 b) & Y_0(\beta_0 b) \\ \Delta_J & \Delta_Y \end{vmatrix} = 0 \qquad \text{.....(11)}$$

where

$$\Delta_J = \begin{vmatrix} \rho J_0(\beta_L a) & \sigma J_1(\beta_L a) \\ J_0(\beta_0 a) & J_1(\beta_0 a) \end{vmatrix}$$

$$\Delta_Y = \begin{vmatrix} \rho J_0(\beta_L a) & \sigma Y_1(\beta_L a) \\ J_0(\beta_0 a) & Y_1(\beta_0 a) \end{vmatrix}$$

with $\rho = \beta_L/\beta_0$ and $\sigma = \varepsilon_L/\varepsilon_D$.

The dependence of the argument β_L, defined after eqn. (4), upon both ε_L and ε_T should here be noted. It is clear that the assumption $\varepsilon_L = \varepsilon_T$ will be progressively less accurate as k increases.

The first zero of eqn. (11) specifies the simplest radial mode and for a modulation experiment we are concerned to match $\omega/k = c/n$, where n is refractive index for light travelling along the optic axis. Hence, for a given electro-optic crystal, we wish the zero of eqn. (11) to imply a root ω/k with the phase velocity of the light wave. For given physical properties and dimensions of the crystal, the necessary condition may be met by adjustment of either of two parameters—the outer radius of the cavity, b, or the dielectric constant of the filler, ε_D.

It is convenient to plot a dimensionless frequency parameter $\Omega = a\omega/c$ against the ratio $\gamma = b/a$ for different values of ε_D. A series of such curves for KDP and its isomorphs has been prepared and examples are shown.

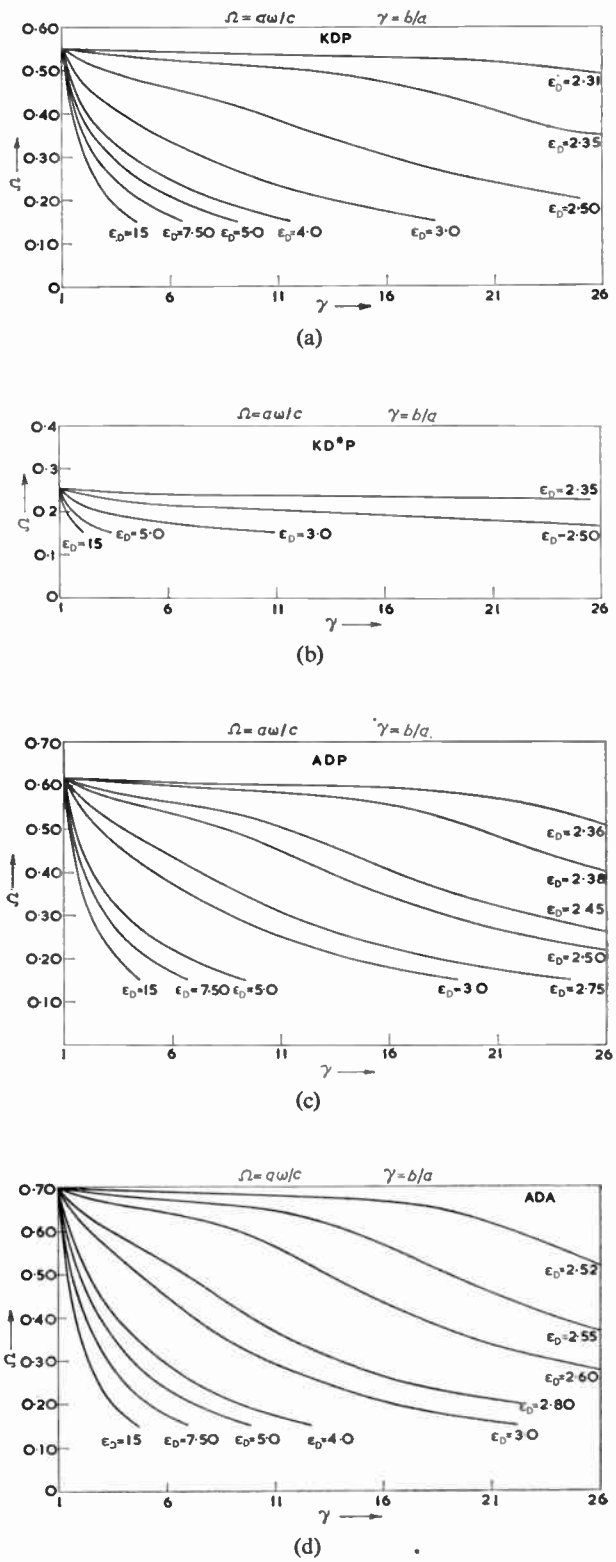


Fig. 2. Curves of Ω ($= a\omega/c$, frequency parameter) vs. γ ($= b/a$) under the matching condition $\omega/k = c/n$.
(a) KDP; (b) KD*P; (c) ADP; (d) ADA.

Table 1

Microwave dielectric constants and optical refractive indices

	$\varepsilon_L/\varepsilon_0$	$\varepsilon_T/\varepsilon_0$	n
KDP	19·7	42·0	1·510
KD*P	90	88	1·510
ADP	15·5	56	1·525
ADA	12	75	1·577

3. Results and Discussion

Figure 2(a) shows curves of $\Omega(\gamma)$ for different ε_D appropriate to KDP (KH_2PO_4); Figs. 2(b), (c) and (d) show corresponding curves for deuterated KD*P (KD_2PO_4), ADP ($NH_4H_2PO_4$) and ADA ($NH_4H_2AsO_4$). The values of dielectric constants and optical refractive index used are shown in Table 1.

Given a cylindrical specimen of electro-optic crystal of radius a and an operating microwave angular frequency ω, the radial dimension b required, when a filler of dielectric constant ε_D is to be used, is immediately obtainable.

A comparison of results for KDP when ε_L and ε_T are given their correct values and when $\varepsilon_T = \varepsilon_L$ (Chen and Lee approximation) shows that under the matching condition, which is not far removed from the cut-off frequency, the error in Ω for a given γ is about 2%; however, owing to the small slopes of the curves for low ε_D, the error in γ for a given Ω could be 50%.

At higher frequencies, away from the matching condition, the propagation (or resonance) characteristics of the guide (or cavity) are in error by a factor $\varepsilon_T : \varepsilon_L$. This may be seen immediately by considering the simplified dispersion equation when $\gamma = 1$, which may be written

$$\frac{1}{\varepsilon_0} - \frac{1}{\varepsilon_L}\left[\frac{2{\cdot}4048}{\Omega}\right]^2 = \frac{k^2a^2}{\varepsilon_T} \qquad \text{.....(12)}$$

Clearly, for large Ω,

$$k \to \sqrt{\varepsilon_T}/(a\sqrt{\varepsilon_0}) \quad \text{or} \quad \sqrt{\varepsilon_L}/(a\sqrt{\varepsilon_0})$$

in the two cases and the high frequency portions of the curves given by Chen and Lee should be subject to such a correction in the case of anisotropic dielectrics.

4. Acknowledgment

We wish to acknowledge the able assistance of Miss A. Kemp in the computation and preparation of the graphs.

5. References

1. D. Chen and T. C. Lee, 'Propagation characteristics of a partially filled cylindrical waveguide for light beam modulation', *Trans. Inst. Elect. Electronics Engrs on Microwave Theory and Technique*, **MTT-14**, p. 482, October 1966.
2. H. Pursey, P. A. Merran and M. J. P. Musgrave, 'Measurement of the Pockels effect in KDP at 9 Gc/s,' *Brit. J. Appl. Phys.*, **18**, pp. 285–91, March 1967.

Manuscript first received by the Institution on 18th January 1967 and in final form on 9th March 1967. (Contribution No. 97.)

U.D.C. 620.168 : 681.33

A Special-purpose Computer for Particle-size Analysis

By

H. A. COLE,

C. Eng., A.M.I.E.R.E.†

Summary: An analogue computer is described which operates in conjunction with the Karl Zeiss particle-size analyser. The analyser is used to measure the diameter distribution of bubbles in a photograph of a column of foam, and the computer makes use of this information to provide simultaneous and separate indications proportional to total surface area and volume as the analysis proceeds. These indications enable an immediate determination of the mean surface area of foam per unit volume of contained air for the sample. The computer may also be used with the analyser for the total area analysis of particles of approximately circular shape, or for total area and volume analysis of particles of approximately spherical shape.

1. Introduction

In many branches of research, it is necessary to measure the size distribution of particles in matter. A particular requirement of this type is the determination of the size distribution of air bubbles in a column of foam so that the mean surface area of foam per unit volume of contained air may be ascertained.

A convenient method of measuring the diameter distribution of the bubbles is by means of the particle-size analyser described by Endter and Gebauer[1] (Karl Zeiss particle-size analyser, TGZ3) and shown in Fig. 1. A photograph of a representative area of foam, shown in Fig 2, is examined, a bubble at a time, on the analyser. The operator matches the diameter of a light spot with that of the bubble and then operates a foot switch causing the bubble diameter to be registered in one of 48 contiguous diameter groups. On completion of the series of operations the readings of 48 electromechanical registers directly provide the diameter distribution.

The normal procedure for the calculation of mean surface area for unit volume of air requires visual read-out of the diameter distribution followed by 49 calculations to obtain the total area, another 49 to obtain the total volume and finally the division of total area by total volume. It is obviously advantageous to perform this calculation by an 'on-line' device which provides the required information as the analysis proceeds. The analogue computer described in this paper achieves this object. Essentially the computer takes information directly from the analyser and, each time a bubble is analysed, produces signals proportional to the square and cube of the mean diameter of the group into which the bubble diameter falls. It separately stores this information to provide data proportional to total area and total volume.

For convenience, the computer has been designated PARSAC (PARticle Size Analogue Computer); it is the subject of British Patent Application 51510/65.

2. The Particle-size Analyser

Although the particle-size analyser is described in detail elsewhere,[2] it is necessary for the understanding of the computer design to describe it briefly here.

The photograph to be analysed is placed upon a transparent plate which is illuminated from the rear through an adjustable iris, manually controlled by the operator. The illumination, in shining through the photograph, causes a sharply defined light of adjustable diameter to appear on the viewing side of the photograph. The diaphragm of the iris is connected to a system of electrical commutators composed of 48 segments, each segment being connected to one of 48 electromechanical registers. The operator aligns the centre of each bubble with the centre of the light spot and then adjusts the iris until the two diameters are matched. The appropriate commutator segment is thus automatically selected and when a foot-switch is depressed the corresponding register operates. This process is repeated until a statistically adequate number of bubbles have been analysed.

The analyser may be set to operate in one of four modes, namely, linear distribution (Lin E), exponential distribution (Exp E), linear summation (Lin Σ) and exponential summation (Exp Σ). These correspond to linear or exponential, differential or integral curves of bubble diameter. The computer was specifically designed to operate on linear distributions and further description will therefore concentrate on this mode (Lin E).

† Electronics and Applied Physics Division, Atomic Energy Research Establishment, Harwell, Didcot, Berkshire.

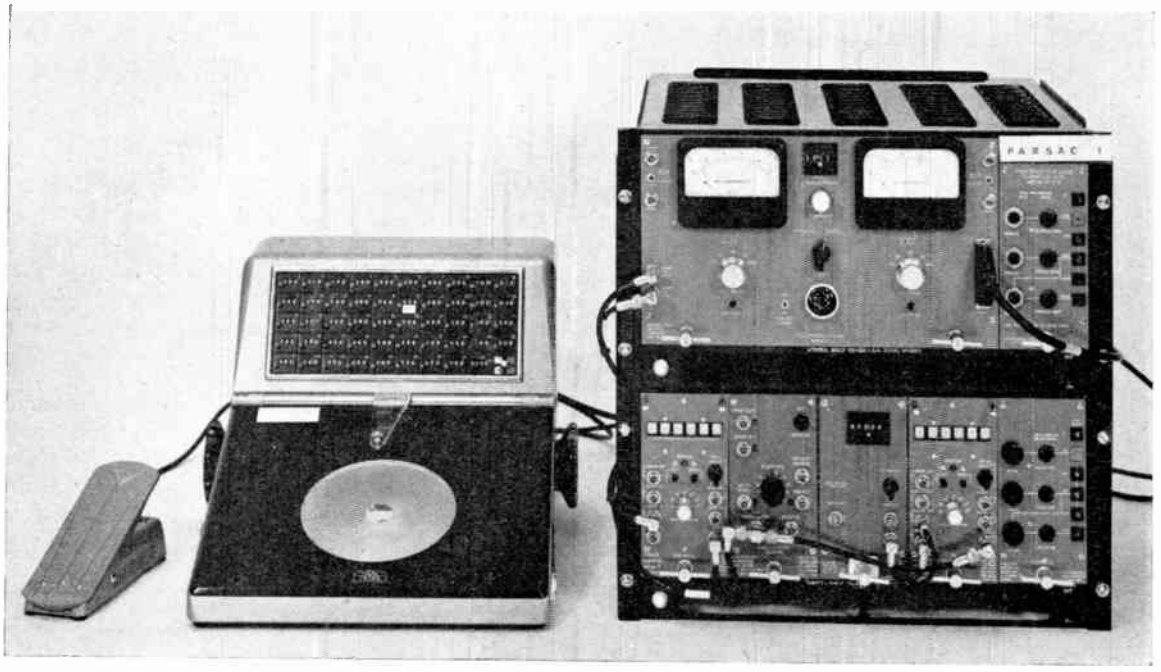


Fig. 1. PARSAC with particle-size analyser.

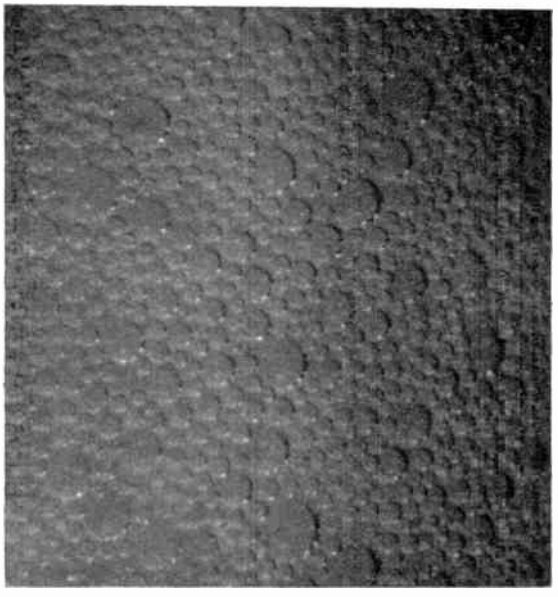


Fig. 2. Foam column photograph.

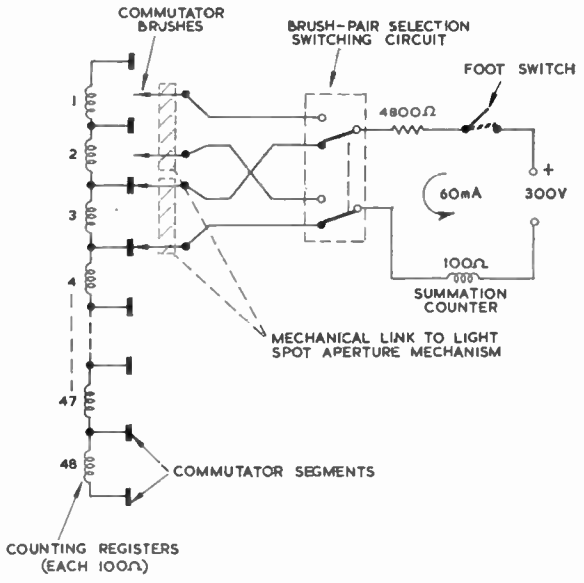


Fig. 3. Simplified circuit of particle-size analyser, Lin E mode.

Figure 3 shows the basic circuit for the Lin E mode of operation. The 48 registers are connected in series and each interconnection is taken to a commutator segment. Two pairs of brushes are mechanically linked to the iris mechanism, and these are arranged to make contact with four consecutive segments. When one pair of brushes is situated in the gaps between segments, the other pair is in contact with two adjacent segments. A second commutator (not shown in detail) senses the position of these brushes and selects the pair which are in contact with the segments. When all four brushes are in contact with segments only one pair is selected. In the linear distribution mode, connection is made to only one register at a time, the

actual register connected at any instant being determined by the position of the brushes on the commutator. When the foot switch is depressed, the solenoid of the selected register is connected through a 4800 Ω current-limiting resistor and a summation register to a 300 V d.c. supply. The summation register indicates the total number of bubbles analysed.

The analyser has two operating ranges which are achieved by a change in the optical magnification of the iris aperture. One range covers diameters from 0·5 mm to 9·44 mm and the second is a factor of about 3 greater, i.e. from 1·49 mm to 27·43 mm.

3. Access to Analyser

It was decided at the outset that any modifications to the particle-size analyser should be as simple as possible and should not interfere with its normal operation. Thus a 50-way socket was fitted to the rear panel of the analyser to provide parallel electrical access to each of the registers measuring diameter distribution. Each time a register operates, a signal of approximately 6 V (d.c.) amplitude appears across its solenoid and exists for as long as the foot-switch contact is made, which is generally about 0·5 seconds.

4. Analogue Signal Generation

To provide indications proportional to area and volume of the bubbles in the sample, each signal from the analyser has to initiate the generation of two separate analogue signals and one of these must be proportional to the square and the other to the cube of the diameter associated with a particular register.

The actual method chosen for producing the analogue signals is influenced by the operating range of the analyser. This is normally about 19·1 : 1 (for both optical ranges), but the squaring and cubing processes of the computer extend this to 365 : 1 and 7000 : 1 respectively. It is this large dynamic range which precludes the use of voltage and current levels to represent the squares and cubes of particle diameters. No such limitations exist if the time duration of a waveform is used as the analogue, and the slow operating rate of the particle size analyser allows the dynamic range to be accommodated without approaching the speed limitations of normal circuits.

If some form of monostable circuit is employed, the maximum pulse duration is governed by the maximum rate at which the foot-switch is likely to be operated during an analysis; about one operation every 2 seconds. If the longest pulse duration is limited to 1 second (corresponding to the cube of the diameter associated with register 48), then the shortest duration (corresponding to the cube of register 1 diameter) will be approximately 140 μs. A serious disadvantage of using active circuits of this type is that two circuits need to be associated with each of the 48 registers, i.e. one for d^2 and the other for d^3.

A method of analogue conversion which results in a considerable saving in circuit components, employs two capacitors associated with each register; the capacitance of one is made proportional to d^2 and the other to d^3. The analogue signals can be related to the quantities of charge appearing on the capacitors for each operation of the analyser, or to the time taken for the capacitors to accept the charges. This is the method adopted in the analogue computer. The advantage of this approach is that only 96 capacitors and a signal of defined voltage level is required, i.e. that used to charge the capacitors. Furthermore, the duration of this voltage pulse is not critical, provided it allows sufficient time fully to charge the largest value capacitor.

Unfortunately, the charging pulse cannot be derived directly from the analyser due to the absence of stabilization of its power supply and the unsuitable waveform developed across the registers. Contact bounce effects give rise to back-e.m.f.'s in excess of 100 V across a register each time it is energized. Figure 4 shows that these e.m.f.'s extend for a period of at

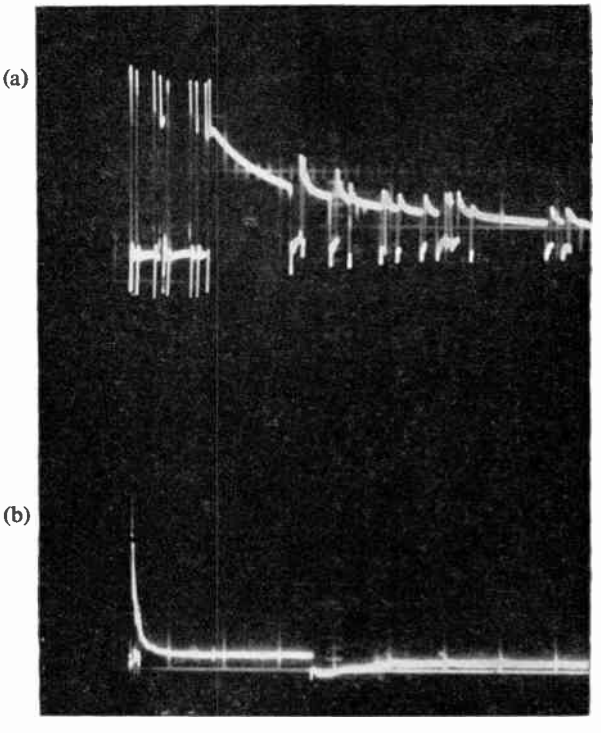


Fig. 4. Back e.m.f.'s generated by register solenoids.
(a) 0·5 ms/division
(b) 20 ms/division
Y axis = 40 V/cm

least 4 ms before the steady level of approximately 6 V is attained. Furthermore, the back e.m.f.'s from one register appear across neighbouring registers at sufficient amplitude to cause triggering of other analogue circuits. Complete isolation between the individual register signals and their associated analogue circuits is therefore necessary, and is readily achieved by connecting a subsidiary relay in parallel with each of the 48 register solenoids. High-sensitivity relays (3 mA, 1200 Ω) are used in this way so that they consume as little as possible of the 60 mA energizing current normally supplied to the register solenoids.

Although the capacitors defining the analogue signals can, in principle be charged from a stable voltage directly through the relay contacts, the closure time is indeterminate because it depends upon the time the foot-switch is depressed by the operator. Other means must therefore be employed to define the minimum duration of the charging voltage. In the computer, this condition is achieved by 48 monostable multivibrator circuits operated from the register signal relays.

The smallest acceptable capacitor value (corresponding to register 1) must be large compared with the stray capacitance. The largest value (corresponding to register 48) must meet the condition whereby its charging-time is less than the shortest period between successive operations of the foot-switch. In the computer, the smallest capacitor (register 1) chosen for both d^2 and d^3 outputs is 22 pF. The corresponding values for register 48 are then approximately 0·15 μF for d^3 and 7130 pF for d^2.

Assuming a 560 kΩ charging resistance (governed by subsequent circuits, to be explained), the charging voltage must be maintained for at least 0·42 s to ensure that the 0·15 μF capacitor is fully charged. Each monostable circuit is therefore arranged to produce a rectangular pulse of 15 V amplitude and approximately 0·6 second's duration when triggered. This period is sufficient for charging the largest analogue capacitor (0·42 s), but shorter than the shortest period between successive operations of the foot-switch (about 2 s).

5. Display Methods

The two analogue quantities of charge produced by the computer, for each bubble registration on the analyser, must be separately integrated and made to produce indications proportional to total surface area

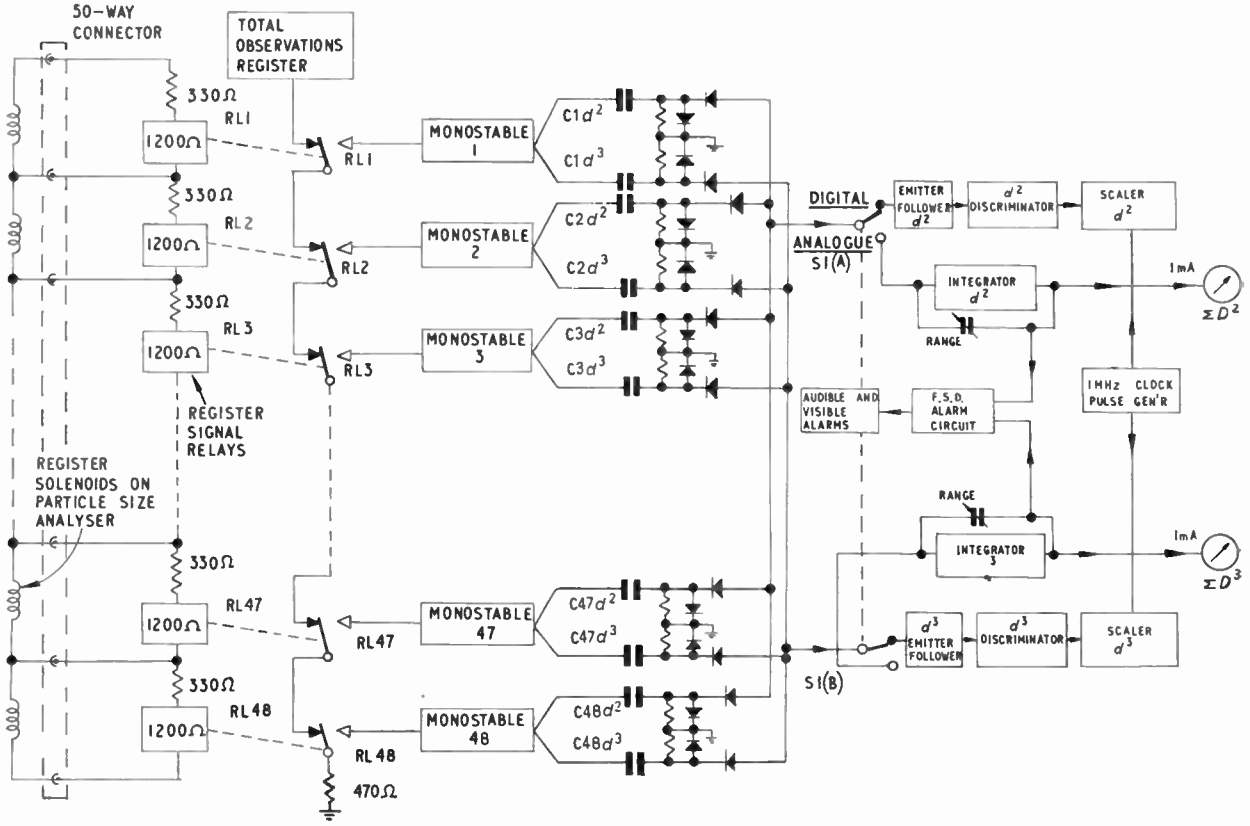


Fig. 5. Block schematic of PARSAC.

and total volume. The actual indications may be in analogue or digital form, depending upon the method of integration. In the PARSAC, a choice of method of presentation is provided.

5.1 *Analogue Presentation*

In the analogue method of presentation (see Fig. 5), the analogue capacitors act as feed-capacitors for two diode pump circuits; the 48 d^2 capacitors supply one pump and the 48 d^3 capacitors supply the other. Associated with each diode pump there is a storage capacitor connected in the feedback loop of a d.c. amplifier to give a virtual earth input. Each capacitor stores the individual charges applied to it during the analysis of a sample and produces an output voltage proportional to the sum of their individual magnitudes,

$$\text{i.e. } V_0 = k \sum_1^n q.$$

The output voltages produced by the d^2 and d^3 amplifiers are displayed on separate wide-scale meters to give indications proportional to total surface area and total volume.

A disadvantage of the analogue presentation is the inherently poor dynamic range of the indicating meters. This can be extended in the normal way by providing the necessary numbers of switched ranges. It may then be possible for the operator to select an operating range suitable for the analysis of all the bubbles on a photograph. Usually, however, it will be necessary to change range during the analysis, especially if the photograph contains bubbles of widely varying size. Since it is not possible to change range on an integrator without losing some of the stored charge, the information displayed on the meters must be read-out before the range is changed. Provisions must also be made to warn the operator when a meter indication approaches full-scale deflection, otherwise information will be lost if this point is exceeded.

In the PARSAC, five ranges are provided for each integrator and audible and visible alarms are automatically initiated whenever either of the two meter indications reaches 80% of full scale.

5.2 *Digital Presentation*

Digital displays do not suffer from the problems of dynamic range associated with the analogue displays. The arrangement of capacitors already described can be used as part of an analogue-to-digital converter (a.–d.c.), in which the digital outputs from two a.–d.c. stages (d^2 and d^3) can be integrated and displayed by electronic scaling units.

The a.–d.c. process takes place in two stages. Firstly, the charging currents of any pair of analogue capacitors associated with a particular registration on the analyser, are arranged to flow through two separate 560 kΩ resistors. The voltage E developed across each resistor R follows the exponential variation in the charging current flowing through it, and therefore decays according to the expression:

$$E = E_0 . e^{-t/CR}$$

where E_0 is the magnitude of the voltage applied to the capacitor, from its associated monostable circuit.

Since E_0 and R are both constant, the charging time (t) of each capacitor is linearly related to C, and hence to the square (or cube) of the diameter of a particular bubble, i.e.

$$t_A = C_A R \log_e E_0/E_A$$

(proportional to the square of the diameter)

and

$$t_V = C_V R \log_e E_0/E_V$$

(proportional to the cube of the diameter)

where the subscripts 'A' and 'V' refer to area and volume respectively.

The two 560 kΩ charging resistors are connected across the input terminals of two d.c. discriminators via high impedance emitter-follower coupling circuits. One discriminator, and its associated input resistor, is common to all the d^2 capacitors; the other discriminator, and its input resistor is common to all the d^3 capacitors.

The thresholds of the two discriminators are set to approximately 1 V so that when voltages in excess of this value are developed across the resistors, during the charging of the analogue capacitors, the discriminators are triggered for periods proportional to the capacitor charging times t_A and t_V.

The finite rise-time of the monostable pulse (~ 10 μs) and the stray capacitance of the emitter-follower input circuits (~ 80 pF), modifies the predicted amplitude of the voltage appearing across the charging resistor, and hence the duration of the gating waveform produced by the discriminator circuits. The errors introduced in this way, however, are only significant when the lower-value analogue capacitors are being used and are easily compensated by padding during the initial calibration.

The discriminators produce steep-sided voltage waveforms for each registration on the analyser the durations of the waveforms being proportional to the square and cube of the diameter associated with the registration. In the absence of an input signal the output level of each discriminator is at (−)15 V. When a signal is received, this level rises to approximately 0 V, and remains there for the duration of the signal; this is shown in Fig. 6 for a capacitor value of 0·1 μF.

The second stage in the analogue-to-digital conversion process is the production of digital outputs by means of the discriminator output waveforms. Two electronic scalers, of the Harwell 2000 Series type (see Appendix), with decimal presentation and external

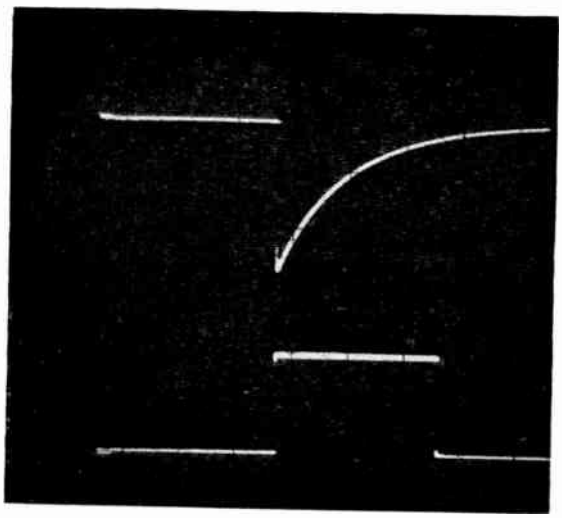


Fig. 6. Discriminator input and output waveforms for 0·1 μF analogue feed capacitor.

X axis 1 division = 20 ms
Y axis 1 division = 5 V (input)
= 10 V (output)

Upper trace: Input waveform
Lower trace: Output waveform

gating facilities, are each supplied with signals from a clock-pulse generator. Both scalers are normally held in the 'off' (inhibited) state by the (−)15 V levels obtained from the two discriminators in the absence of analogue signals. When d^2 and d^3 signals are received, the inhibit levels are removed from the scalers which are then free to count the clock pulses applied to their inputs. The number of pulses counted is determined by the durations of the 'enable' waveforms from the two discriminators. Since these periods are determined by the values of the analogue capacitors, the number of pulses accumulated in the scalers are related, respectively, to the square and cube of each bubble diameter; this is shown in Fig. 7.

The number of clock pulses accumulated by the scalers for each analyser registration must be large enough to permit adequate resolution of adjacent diameters. The inherent registration accuracy of the particle-size analyser varies with diameter (on the Lin E mode of operation), ranging from ±100% for the smallest diameter to a maximum of approximately ±2% for the largest diameter (i.e. ±1 diameter in 48). The subsequent squaring and cubing processes which take place during an analysis result in an effective maximum attainable accuracy of area and volume measurement of about ±4%, this corresponding to the square of diameter 48.

For the digital registration accuracy of the PARSAC to be not worse than ±4%, even for the smallest diameter, a minimum of 25 pulses must be registered for diameter 1, i.e. ±1 pulse in 25. Correspondingly, 175 000 pulses must be registered for the cube of diameter 48. The pulse-rate necessary to ensure this accuracy is determined by the duration of the gating waveform produced by the discriminator for a particular analogue capacitor, which is about 80 ms for the capacitor associated with diameter 48. The necessary pulse rate is therefore

$$\frac{175\,000}{80} \times 10^3 \text{ Hz} = 2{\cdot}2 \text{ MHz}$$

which falls in the 10 MHz operating range of the Harwell 2000 series counting equipment. To reduce the cost of the computer, it was therefore decided to restrict the pulse rate to 1 MHz, for which less expensive units in the series are readily available.

With a 1 MHz pulse rate the registration for diameter 1 (d^2 and d^3) is about 12 pulses, giving a 'worst' digital registration accuracy of ±8%. Due to the squaring and cubing processes, however, the accuracy improves steeply and the corresponding registration for the cube of diameter 48 is about 80 000 pulses, giving a 'best' digital registration accuracy of ±0·001%.

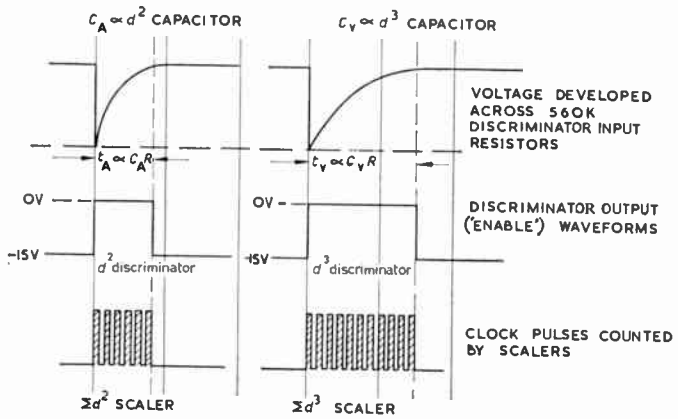


Fig. 7. Principles of the analogue-to-digital conversion process.

6. Computer Operation

6.1 *Analogue Output Mode*

When a single bubble is analysed by the particle-size analyser, a registration occurs in one of the 48 registers, the number of the selected register (R) being proportional to the diameter (d) of the bubble (within the limits of the diameter group represented by the register), i.e. $R \propto d_R$. The actual relationship between diameter and the selected register is given by the expression

$$d_R = f \cdot \frac{k_{p.s.a.}}{k_{mag.}} \cdot R$$

where f is a calibration factor and $k_{mag.}$ is the optical magnification factor of the photograph; $k_{p.s.a.}$ is the range factor of the analyser and has a value of 1 or 3.

The area of the bubble is πd_R^2 and, substituting for d_R,

$$\text{area} = \pi . f^2 . R^2 \cdot \frac{k_{p.s.a.}^2}{k_{mag.}^2} \qquad \ldots\ldots(1)$$

In the analogue mode, the increment in the indication of the 'area' meter due to the operation of register R, is m_R, where $m_R \propto R^2$, i.e.

$$m_R = \frac{ZR^2}{k_A}$$

where Z is a constant and k_A is the range factor of the integrator feeding the 'area' meter (this will have a value of 1, 2, 4, 8 or 16).

Substituting for R^2 in eqn. (1),

$$\text{bubble area} = \left(\frac{\pi . f^2}{Z}\right) k_A \cdot \frac{k_{p.s.a.}^2}{k_{mag.}^2} m_R$$

$$= F_A . k_A \cdot \frac{k_{p.s.a.}^2}{k_{mag.}^2} m_R \text{ cm}^2$$

where

$$F_A = \frac{\pi . f^2}{Z}$$

and is the analogue mode area calibration factor. This has a value of 73·079, determined from measurements.

The final indication (M_A) on the 'area' meter consists of the sum of the individual increments made during the analysis, i.e.

$$M_A = m_{21} + m_6 + m_{48} + \ldots m_R \ldots$$

where the subscripts 21, 6, 48, etc., represent arbitrarily chosen register numbers, to which the increments (m) are related.

The total surface area of foam is given by eqn. (2) as

$$\text{total surface area} = \left(F_A . k_A \cdot \frac{k_{p.s.a.}^2}{k_{mag.}^2}\right) M_A \text{ cm}^2 \qquad \ldots\ldots(2)$$

Similarly, since the volume of a bubble is $\pi . d_R^3/6$, the total volume of contained air is given by

$$\text{total volume} = \left(F_V . k_V \cdot \frac{k_{p.s.a.}^3}{k_{mag.}^3}\right) M_V \text{ cm}^3 \qquad \ldots\ldots(3)$$

where F_V is the analogue mode volume calibration factor and equals 4·513, determined from measurements. M_V is the final indication of the 'volume' meter and k_V is the range factor of the integrator feeding this meter.

Equations (2) and (3) represent the general case and enable measurements to be made with different operating parameters. Should it be necessary to change any of the range factors during the analysis of a sample, the corresponding equations must be evaluated for each range setting and the results subsequently summed.

6.2 *Digital Output Mode*

In the digital mode, the number of counts registered in the 'area' scaler, due to the operation of register R, is n_R, where $n_R \propto R$.

$$n_R = YR^2$$

where Y is a constant.

Substituting for R^2 in eqn. (1),

$$\text{bubble area} = \left(\frac{\pi . f^2}{Y}\right) \frac{k_{p.s.a.}^2}{k_{mag.}^2} \cdot n_R$$

$$= f_A \cdot \frac{k_{p.s.a.}^2}{k_{mag.}^2} \cdot n_R$$

where

$$f_A = \frac{\pi . f^2}{Y}$$

and is the digital mode area calibration factor. This has a value of $6{\cdot}72 \times 10^{-4}$, determined from measurements.

The final indication (N_A) on the 'area' scaler consists of the sum of the individual increments made during the analysis, i.e.

$$N_A = n_{21} + n_6 + n_{48} + \ldots n_R \ldots$$

The total surface area of foam is therefore given by

$$\text{total surface area} = \left(f_A \cdot \frac{k_{p.s.a.}^2}{k_{mag.}^2}\right) N_A \text{ cm}^2 \qquad \ldots\ldots(4)$$

Similarly,

$$\text{total volume} = \left(f_V \cdot \frac{k_{p.s.a.}^3}{k_{mag.}^3}\right) N_V \text{ cm}^3 \qquad \ldots\ldots(5)$$

where f_V is the digital mode volume calibration factor and has a value of $5{\cdot}47 \times 10^{-6}$, determined from measurements. N_V is the final indication on the 'volume' scaler.

In the digital mode there is no need for resetting to zero, or range changing operations on the 'area' and

'volume' scalers during an analysis. Consequently, eqns. (4) and (5) need only be modified to allow for changes which might be made in the value of $k_{p.s.a.}$. Substituting for $k_{p.s.a.} = 1$ and 3 in eqns. (4) and (5):

$$\text{total surface area} = \frac{f_A}{k^2_{mag.}} [N_{A(1)} + 9N_{A(2)}] \text{ cm}^2 \quad \ldots\ldots(6)$$

and

$$\text{total volume} = \frac{f_V}{k^3_{mag.}} [N_{V(1)} + 27N_{V(2)}] \text{ cm}^3 \quad \ldots\ldots(7)$$

The subscripts 1 and 2 associated with N_A and N_V indicate the p.s.a. range on which those indications were obtained.

When it is required to express the result of the analysis in the form 'mean surface area of foam per unit volume of contained air' (m.s.a./V), expressions (6) and (7) may be combined and the measured values of f_A and f_V substituted:

$$\text{m.s.a./V} = 122{\cdot}8\, k_{mag.} \left[\frac{N_{A(1)} + 9N_{A(2)}}{N_{V(1)} + 27N_{V(2)}} \right] \text{ cm}^2/\text{cm}^3 \quad \ldots\ldots(8)$$

7. Circuit Aspects

7.1 *Analogue Capacitor Feed Circuits*

The 48-monostable circuits are identical and of the conventional collector-base coupled type (Fig. 8). They are mounted on plug-in boards, with four circuits to each board. Each circuit is triggered by a negative pulse produced by the operation of contacts RL1–48 associated with the relay connected across the appropriate register in the analyser. The contacts discharge C1 and C2 through a 470Ω resistor causing a negative-going edge to be developed across R2; this edge is transmitted through MR2 and C3 to the collector of TR1 and triggers the circuit. Capacitors C1 and C2 remain discharged for the time the contacts are closed and then recharge on a 0·1 s time constant via R1. The slow recharge which is dictated by C1 and R1, is not readily transmitted by the short time constant of C2 R2 and the resulting positive overshoot is limited by the conduction of MR1; this prevents the monostable being triggered a second time at the end of the input pulse. The slow recharge of C1 and C2 also makes the circuit insensitive to contact bounce when the relay is de-energized. The duration, t, of the output waveform from the monostable is given approximately by $t = 0{\cdot}7\,C4\,R5$ and has been chosen to be 0·6 s. The emitter-follower TR3 feeds the output waveform to two parallel connected pre-set potentiometers RV1 and RV2. These provide calibration controls for adjusting the waveform amplitudes applied to the d^2 and d^3 capacitors.

Each of the analogue capacitors (C_{d^2} and C_{d^3}) feeds its charge into a diode pump circuit comprising diodes MR4 and MR3 for d^2 output and MR6 and MR5 for d^3 output. In the analogue mode of display these outputs go directly to the associated integrator circuits. The function of the diodes in the digital display mode is somewhat different. MR4 and MR6 provide isolation of the appropriate analogue capacitors when the outputs are commoned and MR3 and MR5, and their shunting resistors R10 and R11, allow the analogue capacitors to discharge. In this mode, the combined outputs of the analogue capacitors are fed via the diodes to two 560 kΩ resistors (one for d^2, the other for d^3) across the discriminator inputs. These resistors define the charging time constants of the analogue capacitors in use and hence the output signal durations for the discriminators. Since there are a large number of parallel feeds to each resistor, it is necessary to introduce further series diodes to reduce interaction between the input circuits. One low leakage diode is used to buffer groups of four capacitors (Fig. 5).

7.2 *Analogue Display Circuits*

The two integrators have identical circuits (Fig. 9) as described by Wilkins and Marsh.[3] The storage capacitors, C3–C7, are switched to provide five operating ranges differing by successive factors of two. Display is by means of 0–1 mA meters connected as 0–10 V voltmeters across the outputs of the d^2 and d^3 integrators.

If either meter indication is allowed to exceed 80% of its full scale value, interrupted visual (and audible)

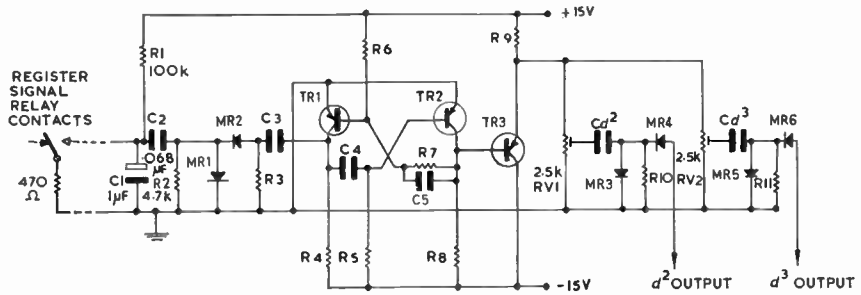


Fig. 8. Circuit diagram of a monostable multivibrator.

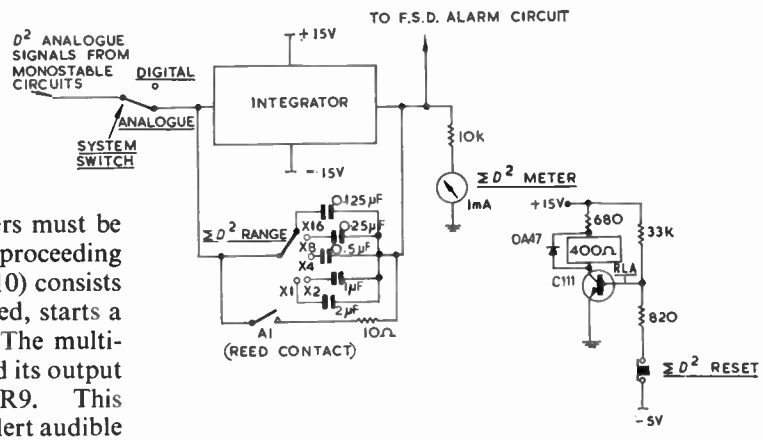


Fig. 9. Circuit diagram of the integrator.

alarms are set off to indicate that the meters must be read-out and the integrators reset before proceeding with the analysis. The alarm circuit (Fig. 10) consists of a d.c. discriminator which, when operated, starts a multivibrator comprising TR7 and TR8. The multivibrator operates at a frequency of 1 Hz and its output controls the conduction of transistor TR9. This transistor drives the relay RLA, and a Sonalert audible alarm; the relay operates the alarm lamp.

The discriminator is mounted on a standard plug-in circuit board and is identical with the discriminators described in the next sub-section.

7.3 *Digital Display Circuits*

The two discriminators[4] used in the digital display mode have identical circuits (Fig. 11) and are mounted on a single plug-in board (Harwell 5000 Series standard circuit board). The input impedance of the discriminator varies with pulse amplitude over the range 20–70 kΩ and it is therefore necessary to provide impedance transformation to avoid shunting the 560 kΩ resistor during the charging time of the analogue capacitors. Darlington-type emitter-followers are used for this purpose and increase the effective input impedance to more than 10 MΩ.

As previously discussed, each discriminator output pulse controls the number of 1 MHz clock pulses fed to a scaling unit. The scalers used with the computer are six decade units of 1 μs resolution,[5] and are controlled by the standard type 1 binary control levels of the Harwell 2000 Series.[5] In the quiescent condition the output of the discriminator is at −15 V which inhibits its associated scaler. When the discriminator is operated the inhibit is removed so that the scaler counts for the duration of the discriminator pulse.

7.4 *Total Observations Register*

A four-decade resettable register is provided on the computer to indicate the total number of bubbles analysed. The register is energized by a transistor

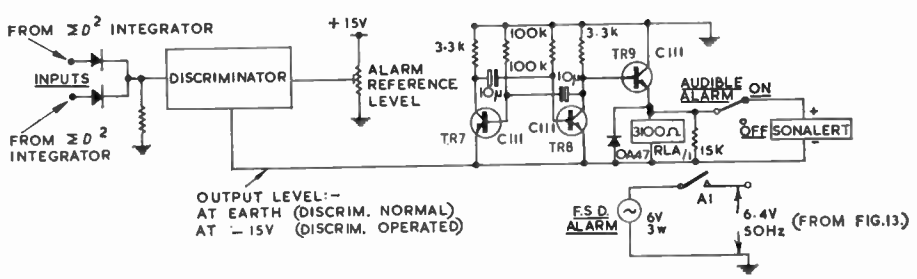


Fig. 10. Full-scale deflection alarm circuit.

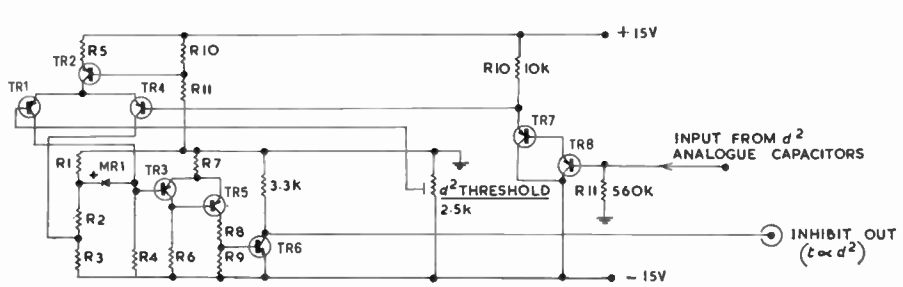


Fig. 11. Circuit diagram of the discriminator.

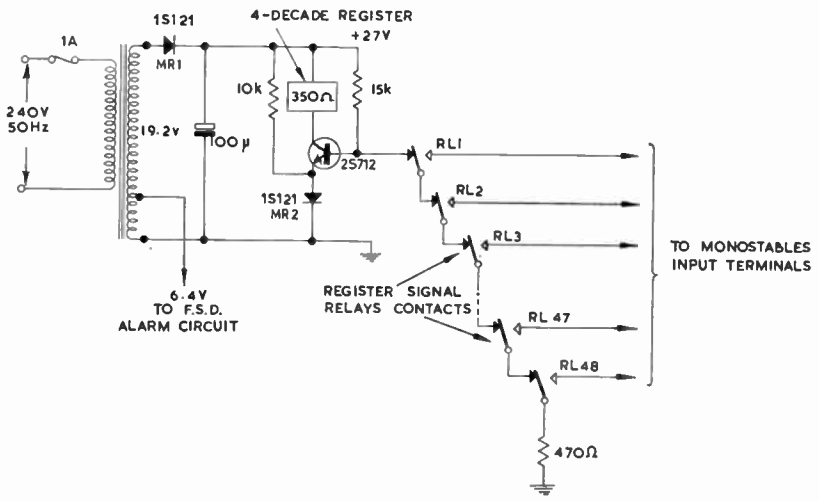


Fig. 12. Circuit diagram of total observations register.

(Fig. 12) which is made to conduct by breaking any pair of the series-connected contacts of relays RL1 to RL48. The silicon diode MR2 provides reverse bias for the transistor.

8. Conclusions

The results obtained with PARSAC are as accurate as the method of analysis requires. A direct comparison was made between PARSAC results and those calculated by the previous method using a central computer (Fig. 13). A hundred different foam photographs were analysed over a period of two months and 96% of the results from the two methods differed by no more than ±1%; 84% differed by no more than ±0·5%.

The time delay in the previous method of working, using manual read-out from the analyser and calculation by the central computer, was typically 1 day. With the present instrument the results are contemporary with the analysis except for a simple ratio calculation with a desk calculator, taking no more than half a minute.

The digital method of presentation has proved to be much more convenient than the analogue display and on further designs the latter might be omitted in the interests of simplicity. Detailed improvements might include the replacement of the existing relays by reed-types (which are cheaper and smaller) and the use of integrated circuits for the monostable multivibrators.

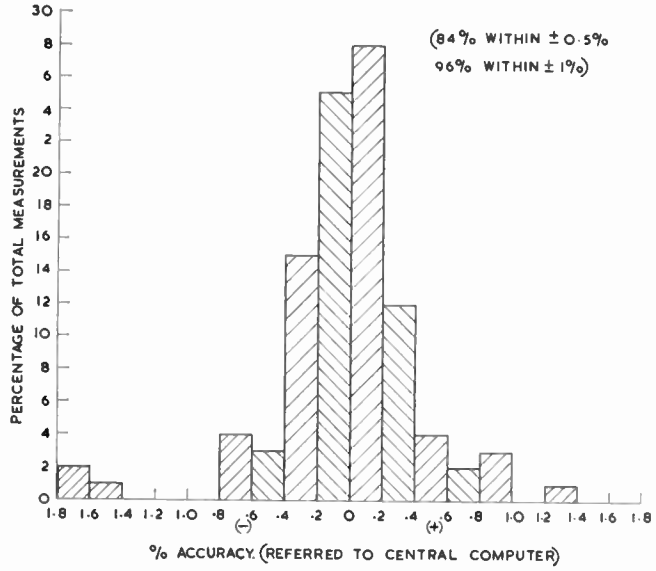


Fig. 13. Measurement accuracy of PARSAC.

9. Acknowledgments

The author wishes to acknowledge the advice and assistance given by Mr. D. Williams during the design of the analogue computer, and in the preparation of this paper. Acknowledgment is also due to Mr. L. Crook who constructed the entire computer and designed much of its mechanical layout.

10. References

1. F. Endter and H. Gebauer, 'Ein einfaches Gerätzur statistischen auswertung von Microscopischen BZW Elektronen Microscopischen ausnahmen', *Optik*, 13, pp. 97–101, 1956.
2. Operators' Handbook, Karl Zeiss Particle Size Analyser TGZ3.
3. J. Wilkins and J. Marsh, 'The Harwell 2000 Series Linear Ratemeter/Integrator', A.E.R.E. Harwell Report R.4676, 1964.
4. J. E. Maddock, A.E.R.E., Harwell, Unpublished work.
5. H. Bisby, 'The design principles and rôle of a comprehensive unit system of electronic equipment, with particular reference to the Harwell 2000 system', *The Radio and Electronic Engineer*, 29, No. 3, pp. 185–95, March 1965.
6. J. M. Richards *et al.*, 'Recording Systems for Use with Harwell 2000 Series Scalers', A.E.R.E. Harwell Report M.1322.

11. Appendix 1

The Harwell 2000 Series is a range of plug-in, modular equipment.[5] The scalers are fully transistorized modules of approximately $3\frac{1}{2}$ in width and are supplied with power from a low-voltage plug-in power unit of similar dimensions; this unit may also be shared by other modules.

There is a variety of scalers in the series which can be divided into four main categories, each category being designed to fulfil a particular requirement:

(i) 4 decade, 10 ns resolution, with and without read-out access,

(ii) 6 decade, 100 ns resolution, with and without read-out access,

(iii) 3 and 6 decade, 1 μs resolution, with and without read-out access,

(iv) 2 decade with electromechanical register, 5 μs resolution.

Each scaler accepts standard digital pulses and produces a similar output 'carry' pulse (after a preselected scaling factor) which may be accepted by other units in the series. The scalers may be started, stopped or reset either manually, or automatically by applying standard binary signals.

A binary output signal is available from a scaler (whether it is controlled manually or electrically) whenever it is inhibited, or whenever the accumulated number of pulses in the store exceeds a preselected number set by the scaling factor; exceeding this number therefore results in the generation of a digital output signal which is coincident with the 'carry' pulse.

If a binary signal is connected to the control input of a scaler, the count condition of the scaler will be inhibited when the signal is in its '1' state, and free to count when in its '0' ('enable') state. Groups of scalers may be controlled together in this way by a common signal and controlled from a remote source, e.g. from another scaler. Individual scalers may also be connected so as to stop themselves when a predetermined count has been accumulated. This is achieved by coupling the binary inhibit signal generated by the scaler into its own control socket. Connections like this enable groups of scalers to form automatic timing circuits or to perform ratioing, squaring and other computations on the input pulse trains.

The contents of the store of an access scaler can be obtained in electrically digital form when the scaler is suitably addressed. Thus a number of scalers can be connected into a common highway, which is controlled from a central point and the store contents read out on to punches, printers and tape recorders.[6]

Manuscript first received by the Institution on 15th June 1966 and in final form on 13th January 1967. (Paper No. 1120/C95.)

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.

TELEVISION FILM-RECORDING

There are two distinct operations in the conventional techniques used for television film-recording: recording a positive display on to a negative film, and printing on a positive film from this negative image. A theoretical and practical investigation of a negative display on the cathode-ray tube was undertaken in the Recording Laboratory of the Research Department of the Office de Radiodiffusion-Télévision Française (O.R.T.F.).

This investigation is developed systematically, utilizing the principle of the transfer function. After a brief recapitulation of that principle the several links constituting the recording chain are analysed, from the vision signal applied to the display tube to the electrical signal obtained from the film-scanner. The electronic reversal, for which the idea of 'gamma' is theoretically incompatible with the operation effected, is emphasized. Finally, the results of tests carried out by the O.R.T.F. on a number of test-patterns are described.

'Television film-recording from a negative image', J. Gregeois, *E.B.U. Review*, 101-A Technical, pp. 2–10, February 1967.

TESTING THE ENCAPSULATION OF SEMICONDUCTOR DEVICES

A method is described, in a Czech paper, for testing the leak-proof encapsulation of semiconductor devices with the aid of radioactive tracer gas—crypton 85. As compared with the methods used until now, the new method is much more sensitive and facilitates an accurate evaluation of the leakage rate. The laws of aero-mechanics are used as the basis for deriving the principal relations. Special attention is paid to measuring accurately the activity of the leaky component parts and to excluding the influence of tracer gas absorption on the surface of the component parts. A detailed description is given of the test equipment which enables a sensitivity of $1{\cdot}06\times10^{-10}$ torr-litre per second to be attained.

'Testing the leak-proof encapsulation of semiconductor devices with the aid of radioactive tracer gas', M. Tauer, *Slaboproudý Obzor*, **28**, No. 4, pp. 233–7, April 1967.

EFFECT OF NOISE ON PHASE-LOCKED A.F.C. SYSTEM

Phase-locked automatic frequency control is used for frequency stabilization in transmitters, receivers used for phase-shift keying and also for measuring Doppler shift in radar systems.

A particular feature of the phase-locked a.f.c. system is its noise immunity. However, in the presence of strong noise, locking fails and 'jumps' occur from one equilibrium region to another. This leads to 'reverse' working in the case of phase telegraphy and to serious frequency measurement errors in radar systems.

A Soviet paper experimentally investigates the effect of fluctuation noise on the automatic control process in a phase-locked a.f.c. system and determines the error signal probability density and correlation coefficient in such a system.

'Experimental study of the action of fluctuation noise on a phase-locked a.f.c. system', B. I. Shakhtarin and Yu. N. Shchepkin, *Telecommunications and Radio Engineering* (English language edition of *Elektrosvyaz* and *Radiotekhnika*), pp. 15–19, No. 9, September 1966.

RESONATORS WITH NON-UNIFORM LINES

Resonators using non-uniform lines offer a solution to the frequency-related limitations of uniform line resonators.

In a French paper, a variety of curves provide a means of easily determining the frequencies of series and parallel resonance, as well as the behaviour of the reactance and the input susceptance of the resonators in terms of frequency, without laborious calculations.

By the use of an analogue computer the results obtained have been checked by direct calculation and it has been possible to resolve problems relating to such lines where the equations do not have exact mathematical solutions. This is particularly true in the case of the conical line which is the easiest to construct. Resonators using non-uniform lines can be employed in harmonic filters, reactance equalizing circuits for negative-resistance amplifiers, as well as for wide-band applications.

'Resonators using non-uniform lines', J. C. Matheau, *L'Onde Electrique*, **47**, No. 478, pp. 91–104, January 1967.

STUDY OF MOS FIELD EFFECT TRANSISTORS UNDER PULSED CONDITIONS

A method for the analysis of the pulsed signal behaviour of the metal-oxide semiconductor field effect transistors below and within the 'pinch-off' range is proposed in a German paper.

A justification for the use of a model to illustrate the pulsed signal behaviour within the 'pinch-off' range is given on the basis of considerations of the static behaviour in the 'pinch-off' range. The results from such an analysis have been confirmed by measurements on an n-type MOSFET.

'On the static and dynamic behaviour of MOS field effect transistors under large signal conditions', A. Möschwitzer, *Nachrichtentechnische Zeitschrift*, **20**, No. 3, pp. 150–54, March 1967.