JOURNAL OF

The British Institution of Radio Engineers

FOUNDED IN 1925 - INCORPORATED IN 1932)

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

Vol. XV No. 11 NOVEMBER, 1955

REPORT OF THE THIRTIETH ANNUAL GENERAL MEETING

THE INSTITUTION'S THIRTIETH ANNUAL GENERAL MEETING (the twenty-second since incorporation) was held at the London School of Hygiene and Tropical Medicine, Gower Street, London, W.C.1, on Wednesday, October 26th, 1955.

Mr. G. A. Marriott (Vice President) stated that the President, Rear-Admiral Sir Philip Clarke, was unable to attend the meeting as he had suffered the very sad loss of his mother on the previous morning.

Mr. Marriott therefore took the Chair and was supported by other Officers of the Institution and members of the General Council. Forty-four corporate members had signed the minute book when the meeting opened at 6.5 p.m.

The Secretary read the notice convening the meeting which was published on page 377 of the August *Journal*.

1. To confirm the Minutes of the 29th Annual General Meeting held on October 27th, 1954

The Secretary stated that a report of the last Annual General Meeting was published on pages 509-512 of the November 1954 *Journal*. Mr. Marriott's proposal that these Minutes be signed as a correct record of the proceedings was approved unanimously.

2. To receive the Annual Report of the General Council

In calling upon Professor Emrys Williams, Chairman of Council, to present the Annual Report, Mr. Marriott referred to the fact that Professor Williams had attended every Council meeting during the year as well as having served on the various Institution Committees.

Professor Emrys Williams stated that this was the first occasion on which he had formally presented the Annual Report of the Council. He expressed appreciation for having had the opportunity of serving as Chairman of Council for the past twelve months, and had been much inspired by the painstaking care which the President and his Officers gave to the Institution's problems.

The keynote to the whole of the Report, and indeed to the work of the Institution, said Professor Williams, was to be found in the penultimate paragraph of the introduction which stated "The Council welcomes these further opportunities to be of service to the profession." Apart from developing the services to the membership and to the profession during the past twelve months, the Council had been particularly concerned with making the objects of the Institution more widely known and understood outside the membership.

Referring to the problems of the training, supply and employment of radio and electronics engineers, Professor Williams stated that through its educational work the Institution could make an important contribution to the well-being and future of industrial development, including Government and Service establishments. In this connection, the Professional Purposes Committee had been active in continuing negotiations with the Ministry of Labour on the matter of more adequate recognition being given to the work of the qualified radio engineer. Their efforts had been much strengthened by the care which the Membership Committee exercised in examining proposals for election to membership. Whilst such caution inevitably brought disappointment to many applicants, it was far better that the Institution should exercise caution, for the

Institution could then truthfully claim that election to membership conferred a status at least equivalent to membership of other comparable bodies. Professor Williams continued:

"If, however, the Membership Committee is to do its work properly, it is of even greater importance that we should move with the times and yet retain our high standard of examination. It has never been the Council's policy to make the Institution primarily an examining body. If other equivalent examinations were available, the Brit.I.R.E. examination would be necessary only as an interpretation of the Institution's educational policy.

"As the applications for admission to the Institution have increased, the demand for training and courses of study to meet the Institution's requirements for membership has also grown. An indication of the importance of this demand is reflected in the fact that although the Institution now recognises a number of university degrees, diplomas and other examination for the purpose of exemption from its Graduateship Examination, 1,198 candidates entered for the Institution's examination during 1954. I think it is worth recording the fact that since the cessation of hostilities in 1945, the Institution has accepted no less than 8,910 applications for entry to the Graduateship Examination, which indicates that a growing number of people realize the necessity for membership of a professional body adequately representing their particular activity, work or profession.

"The Education Committee has, therefore, not only the responsibility to the Institution of setting minimum standards for admission to the Institution, but also has the great responsibility of indicating the requirement for membership and the course of study which so many candidates must pursue in their efforts to become responsible members of a great and growing industry and profession."

Paying tribute to those members who had given so much time to studying education problems and who had been able to overcome even the cautious criticisms of the Council in framing revisions to the syllabus of the Graduateship Examination, Professor Williams said that they now had their reward in seeing the Institution so much better able to serve the profession.

Professor Williams also referred to the great burden of work which also fell upon the Papers Committee in ensuring that the standard of the Journal and of papers read and discussed at meetings and Conventions was maintained. He felt that the progress which had been made in publishing such an invaluable Journal, in providing meetings throughout Great Britain and overseas, and in the holding of Conventions, for which the Institution now had a great reputation, was a tribute to the careful discrimination and wise judgment of the members who served on that Committee.

Commenting on the Institution's Library work, Professor Williams particularly paid tribute to the efforts of the Committee in maintaining and extending the reference and lending library services notwithstanding the handicap of insufficient accommodation. Expansion of this important service was necessary and it was one of the main reasons why he personally hoped that it would soon be possible to acquire suitable premises.

In making a brief reference to the financial affairs of the Institution, Professor Williams affirmed that the Council would continue to consider how best to provide services to members, without the necessity of calling upon them to make greater financial contributions. He concluded, "I hope that my remarks have given some impression of the amount of work which the Council and its Committees undertake on behalf of the members and the profession. As more and more is accomplished, one sees always pressing tasks ahead. From my own observations as Chairman of Council during the past 12 months, I am in no doubt that there will always be members ready to give of their best in tackling such problems as the future holds."

Professor Williams then formally moved the adoption of the 29th Annual Report of the Council. The motion was supported by Mr. W. E. Miller (Immediate Past President), who said that as a member of long association with the Institution he was particularly pleased to see the progress which was being made. He considered that the Report gave a very true and impressive account of the Institution's work during the year.

The proposal that the Annual Report be adopted was carried unanimously.

3. To elect the President

Professor Williams proposed that Rear-Admiral Sir Philip Clarke, K.B.E., C.B., D.S.O., be re-elected President of the Institution for a second year. Supporting the motion, Mr. G. A. Marriott spoke of the work which Admiral Clarke had done on behalf of the Institution. The Council was particularly appreciative of the fact that Admiral Clarke had found time to attend every Council meeting and Mr. Marriott was sure that members would convey the warmth of their feelings toward Admiral Clarke in unanimously re-electing him as President for a second year. The motion was carried with acclamation.

4. To elect the Vice-Presidents

Mr. W. E. Miller (Immediate Past President) proposed the re-election of Messrs. G. A. Marriott, L. H. Paddle, and J. L. Thompson, and Professor E. E. Zepler as Vice-Presidents. Since they were all very well known to members for their work on behalf of the Institution, Mr. Miller did not think that any further introduction was necessary from him.

The proposal was approved unanimously.

5. To elect the General Council

The Agenda showed the Council's nominations for the vacancies on the General Council and Mr. Marriott felt that it was a measure of the confidence reflected in the Council that there had not been any opposing nominations. He therefore formally declared that the following members were elected to fill the five vacancies:—

Members: Air Vice-Marshal C. P. Brown, C.B., C.B.E., D.F.C.
J. W. Ridgeway, O.B.E.
Professor Emrys Williams, PH.D., B.ENG.

Associate Member: Mr. R. N. Lord, M.A.

Companion: Mr. A. H. Whiteley, M.B.E.

Mr. Marriott thanked the retiring members of Council for their support and for the time they had given to the work of the Institution.

6. To elect the Honorary Treasurer

In moving the re-election of Mr. G. A. Taylor as Honorary Treasurer, Mr. Marriott commented that Mr. Taylor had given convincing proof of his ability to act in the interests of the membership and on behalf of the General Council during the last two years. The motion was supported by Mr. W. E. Miller and the proposal was carried unanimously.

7. To receive the Auditor's Report, Accounts and Balance Sheet for the year ended March 31st, 1955.

Mr. Marriott called upon Mr. G. A. Taylor to present the Accounts.

After thanking the members for re-electing him as Treasurer of the Institution for the third consecutive year, Mr. Taylor said that he was again pleased to report that the Institution continued to make progress; members would have seen from the Accounts that revenue from normal sources of income had increased by over £1,000.

A comparison of the Income and Expenditure Account with previous years would indicate that whilst income showed steady progress, costs were very far from steady. One example of steeply rising costs was the greatly increasing cost of paper and printing involved in the publication of the Institution's *Journal*. The increased expenditure on this item also included the cost of publishing the List of Members.

During the year it had been possible to save a little on the premises account, but this could not be done every year and future accounts would show a greater cost for repairs and maintenance to the building. Mr. Taylor reminded members that expenditure on this item did not include the charges involved in renting outside accommodation for holding Institution meetings and examinations in London and those costs would be another form of saving when the Institution acquired a suitable building.

Mr. Taylor felt that the detailed way in which the Accountants showed the expenses left very little for him to explain, except perhaps two items of non-recurring expenditure—the preliminary expenses for the Institution's Coat of Arms, and the 1954 Convention. It was always the Council's hope that Conventions might become self-supporting, thus avoiding the need for subsidising such activities from the General Fund. Conventions were, however, a very desirable and necessary feature of the Institution's life and the Council was anxious to facilitate the participation of as many members as possible without imposing upon them charges which might make their attendance difficult.

Mr. Taylor then referred to the progress of the Building Appeal. He recalled that at the previous Annual General Meeting he had stated that the Institution could not expect industry to give greater support to the Building Appeal if members were not prepared to help as much as possible. He was pleased to report that during the year there had been a good response from members who had so far contributed £2,940. The Council was, however, anxious to see this figure increased to £5,000 and with that evidence of the members' support Mr. Taylor considered that the Institution's efforts to secure further support from industry would not be unavailing.

Finally, the Balance Sheet indicated on the one side the efforts which had been made to build up financial stability, and on the other the liabilities which had accrued over the years and which could only be reduced by producing an annual surplus of revenue. Mr. Taylor concluded: "In doing this, we can anticipate reasonable progress in our income, but we face ever-rising costs, much of which is beyond our control. I can only give you my assurance that the Finance Committee is well aware of the battle they have and are doing all they can to help the Institution to rise above the handicap of increasing costs."

Mr. Taylor then moved the adoption of the Accounts and the report of the Auditors.

Air Vice-Marshal C. P. Brown, C.B., C.B.E., D.F.C. (Member), seconded the proposal that the Accounts of the Institution be adopted and the motion was carried unanimously.

8. To appoint Auditors

Mr. Taylor proposed the re-appointment of Gladstone, Jenkins & Co., who had for many years acted as the Institution's Auditors, and the proposal met with unanimous approval.

9. To appoint Solicitors

Mr. Marriott expressed the Institution's appreciation for the advice and assistance given by Mr. Charles Hill. The proposal that Messrs. Braund and Hill be re-elected as Solicitors to the Institution was approved unanimously.

10. Awards to Premium and Prize Winners

Mr. Marriott stated that during the year a number of notable papers were presented to the Institution and his pleasure at presenting awards was tinged with regret that it was not possible to make an award to all who had contributed to the Proceedings of the Institution. He assured those authors who had not been successful in obtaining a Premium that their contributions were very much appreciated.

Mr. Marriott congratulated the authors whose papers had been selected as specially worthy of commendation and presented the following awards:—

The Clerk Maxwell Premium to Mr. F. N. H. Robinson, M.A.;

The Heinrich Hertz Premium to Dr. T. B. Tomlinson (Associate Member);

The Louis Sterling Premium to Mr. L. H. Bedford, O.B.E., M.A., B.SC. (Member):

The Brabazon Premium to Messrs. J. W. Jenkins, J. H. Evans (Associate Member), D. Chambers, B.SC., and G. A. G. Wallace;

The Marconi Premium to Mr. W. T. Brown (Associate Member);

The Students' Premium to Mr. R. W. Walker.

The J. C. Bose Premium and the Leslie McMichael Premium had been awarded to Dr. S. Deb, and Dr. G. N. Patchett (Member) respectively, who were unable to be present at the meeting.

The following received special awards for their outstanding contributions to the 1954 Convention:—

Mr. P. Huggins (Associate Member);
Mr. C. W. Miller, M.SC. (Associate Member);
Mr. K. Kandiah, M.A., B.SC., and Mr. D. W. Chambers;
Mr. J. L. Thompson (Member);
Mr. M. T. Elvy (Associate Member).

Mr. Marriott regretted that he would not have the pleasure of presenting the Examination Prizes for 1954 as the winners were unable to be present.

11. Any other business

The Secretary stated that he had not received notification of any other business, but before closing the meeting Mr. Marriott particularly referred to the work accomplished by Mr. Clifford and his staff and asked that the members should express their appreciation in the usual way.

Mr. Marriott then declared the 30th Annual General Meeting closed.

The Annual General Meeting of Subscribers to the Benevolent Fund then took place and a report will be given in the December Journal.

NOTICES

Fund for the Advancement of Scientific Education

Seventeen major industrial organizations have established an industrial trust "Industrial Fund for the Advancement of Scientific Education in Schools." More than £1,500,000 has already been guaranteed.

Alarmed by the growing shortage of scientists, mathematicians and technologists, the sponsoring companies plan to help independent and direct-grant schools where facilities for teaching science subjects are seriously inadequate through lack of capital resources. The Trust has pointed out that public funds are available for capital works at maintained schools.

The Fund is to be administered by an executive committee of thirteen prominent industrialists and educationalists, and the sponsoring companies include: Associated Electrical Industries Ltd., British Insulated Callenders Cables Co. Ltd., the English Electric Co. Ltd., and the General Electric Co. Ltd.

Norman W. V. Hayes Memorial Medal

The Norman W. V. Hayes Memorial Medal of the I.R.E. Australia for 1955 has been awarded to J. Anderson and W. D. Meewezen for their paper entitled "A New Mobile Phase-Modulated Transmitter-Receiver Unit." The paper was published in the February 1954 issue of the Proceedings of the Institution of Radio Engineers, Australia, and was read at the I.R.E. Convention held in Sydney during August 1952. The Hayes Medal is awarded annually for the most outstanding paper published in the Proc. I.R.E. Aust. during the year. The Brit. I.R.E. and the I.R.E. (America) make the selection in alternate years, and the American body was the adjudicator for this year's award.

Whitworth Foundation Awards 1956

It was announced in the June 1955 issue of the *Journal* that it had been decided to increase the number of Whitworth Senior Scholarships from two to three, to raise their value from £325 to £500 per annum, and to discontinue the award of ordinary Whitworth Scholarships.

Applications are now invited from candidates for the first awards to be offered under the new arrangements. These awards, now known as Whitworth Fellowships, will be available to holders of engineering degrees, or of Higher National Diplomas or Certificates in engineering with at least two distinctions, who have had at least two years' practical engineering experience. The Fellowships may be used to follow, over a period of at least two years, a course of further training in industry or postgraduate study or research in an approved establishment.

Full particulars of the conditions of entry and method of application may be obtained from the Ministry of Education, F.E.I.(D), Curzon Street, London, W.1.

Importance of Further Education

Mr. F. Bray, Under Secretary, Ministry of Education, recently referred to the long-term aspect of education in relation to manpower, industrial productivity and the approach to automation. At present only about a third of the boys and girls who left school at 15 and 16 years were attracted to Further Education. This was a grave weakness which could be wholly remedied only by the introduction of the County College.

On the training of technicians, he said that the responsibility of the technologist will tend to increase as automation and electronic controls develop, and we might well find that the technician of tomorrow must be as well trained as the technologist of to-day.

The Ministry has given much thought to the needs of the future technologist, for it was on his quality that the progress of industry so largely depended. More time is required to provide breadth of education and a greater knowledge of fundamentals. This was one reason for the introduction of the "sandwich" course with alternate works and college teaching.

Exhibition Dates, 1956

The Radio Industry Council has announced that the 23rd annual National Radio and Television Exhibition ("the Radio Show") will be held at Earls Court, London, from August 22 to September 1, 1956, with a pre-view for overseas and other special guests on August 21.

The Radio and Electronic Component Manufacturers Federation states that its annual private exhibition will be at Grosvenor House, London, W.1, from April 10 to 12, with a pre-view for overseas and other specially invited guests on the afternoon of April 9. Application for admission will have to be made in advance to the Radio and Electronic Component Manufacturers Federation, 21, Tothill Street, London, S.W.1.

APPLICANTS FOR MEMBERSHIP

New proposals were considered by the Membership Committee at a meeting held on 20th October, 1955, as follows:—35 proposals for direct election to Graduateship or higher grade of membership, and 37 proposals for transfer to Graduateship or higher grade of membership. In addition, 84 applications for Studentship registration were considered. This list also contains the name of one applicant who has subsequently agreed to accept a lower grade than that for which he originally applied.

The following are the names of those who have been properly proposed and appear qualified. In accordance with a resolution of Council and in the absence of any objections being lodged, these elections will be confirmed 14 days from the date of the circulation of this list. Any objections received will be submitted to the next meeting of the Council with whom the final decision rests.

Direct Election to Member

SEELEY, Wg. Cdr. Eric Charles, M.B.E. *Fontainebleau, France.*

Transfer from Associate Member to Member

DRISCOLL, Louis, B.Sc. *Welwyn Garden City.*

Direct Election to Associate Member

ATKINSON, Alan. *Marlow.*
BALL, Maj. Lawrence Frederick. *London, W.5.*
COOPER, Lieut.-Com. George William. *Titchfield Common, Hants.*
GILLARD, Jim Henry. *Stanmore.*
HEAD, Leonard Russell. *Chelmsford.*
INIGO-JONES, Lt.-Col. Charles Meredith. *London, S.W.1.*
JAMIESON, Lieut.-Com. Jack Revol, D.S.C. *Farnham.*
JOSHI, Madhusudan Vasudeo, B.E. *London, W.2.*
LITTLER, Flt. Lt. James, R.A.F. *Malta G.C.*
NORTHROP, Edgar Waite, M.Eng. *Chorleywood.*
ROBERTS, Frederick Alexander. *London, W.5.*
WAREHAM, Eric Maynard. *Letchworth.*

Transfer from Associate to Associate Member

FLACK, Maurice. *Harrow.*
RIMMER, John Barlow. *Glasgow.*

Transfer from Graduate to Associate Member

FISHER, Bernard, B.Sc. *Sidcup.*
HOGG, Douglas. *Tunbridge Wells.*
JOSLIN, Charles Albert. *London, W.2.*
NORTHALL, Bertram Victor. *Middlesbrough.*
PITTILO, Robert Dawson, B.Sc. *Glasgow.*
POULTER, Arthur Frederick. *Wembley.*
STANBROOK, Donald. *Stevenage.*

Transfer from Student to Associate Member

GOULDSTONE, Paul. *Carshalton Beeches.*

Direct Election to Associate

GRAHAM, William M'Gregor. *North Borneo.*
HEYS, Harry. *Heston.*
PERERA, Seneviratnage Lawrence Valarian Anthony. *Wattala, Ceylon.*
SHEARS, Rowland George. *Wilmington.*

Transfer from Student to Associate

BEAVEN, Horace Mawson. *Bournemouth.*
GRANT, Peter. *Wembley Park.*
HARE, Ronald Edwin. *London, N.W.7.*
HOLGERSON, Tor. *Bodo, Norway.*
MCNAIR, Archibald, B.Sc. *Barnet.*
ROBERTS, William Bentley. *Liverpool.*
SIMPSON, Frederick Thomas. *London, S.E.18.*
UPCOTT, Gilbert Spencer. *London, N.17.*

Direct Election to Graduate

ARDEN, William Francis. *London, S.E.5.*
ASARE, Jacob Kwame, B.Sc.(Eng.) *Manchester.*
BHATNAGAR, Brijnandan Saroop. *Allahabad.*
CHAPMAN, Roy Kenneth. *Canterbury.*
CLARKE, Christopher Dowse. *Loughborough.*
HUMPHREY, Ralph. *East Bedfont.*
KETTLETY, Arthur. *Derby.*
LLOYD-CLAMP, David Warwick. *London, S.W.19.*
MULLARD, John Eric. *Basingstoke.*
SARKAR, Ananta Bikram, M.Sc. *London, S.E.18.*
TAYLOR, Denis Frederick. *Hockley.*
TAYLOR, Dennis Philip. *Harlington.*
WALKER, Kenneth. *Southampton.*
WOODWARD, Grieg Roger. *Stourbridge.*
ZENIOS, Andreas Charalambous. *Cyprus.*

Transfer from Student to Graduate

BENEDETTI, Rev. Edmund Joseph. *Southampton.*
JACKSON, Eric. *London, E.11.*
JOHARI, Maharaj Bahadur, B.Sc. *Amalner, Bombay State.*
KANDASAMY, Manicavasagam, B.A. *Palam Cottah, Madras State.*
MARSHALL, Gerald William Arthur. *Singapore.*
MURCH, Leslie John. *Southend-on-Sea.*
OAKES, Mervyn Brian. *Bristol.*
RAMACHANDRAN, Gopalaswamy. *Secunderabad.*
RUBIN, Moshe. *Haifa.*
SRINIVASAGOPALAN, C., B.Sc.(Hons.). *Madras.*
THAKER, Kantilal Jethalal, B.Sc. *London, N.19.*
TRIBBLE, Kenneth Alan. *Waihi, New Zealand.*
WALDRON, Richard Arthur, B.A. *Chelmsford.*

STUDENTSHIP REGISTRATIONS

ADLLAGUN, Samuel Adebayo. *London, S.W.9.*
ANTONPOULOS, Angelos. *Athens.*
ASHMAN, Roy John. *St. Albans.*
AWOMOLO, George Oludaisi. *London, N.W.5.*
BANSAL, Vijai Kant, B.Sc. *Rajkot.*
BHALLA, Lt. Parkash Chander. *Lucknow.*
BHANU RAJA, K.C., B.Sc. *Madras.*
CHAKRABORTY, Gour Prasad. *Bangalore.*
CHMIELOWSKI, Jerzy Czeslaw. *London, W.11.*
CORNICK, Milton L. *Wanganui, New Zealand.*
DEVAN, Deva Das. *Brighton.*
DIGGINS, James Edward. *Zomba.*
EXARCHOS, Vladimir. *Athens.*
FLORIDIS, Georges. *Athens.*
GREY, John. *Belfast.*
GULATI, Madan Lal. *New Delhi.*
HADJATHANASIOU, Demetrios. *Athens.*
HANS RAJ SINGH SIWACH. *Delhi.*
HARI SINGH SAINI. *Ahmednagar.*
HARIKANOS, Constantinos. *Athens.*
HARRIS, Aubrey. *Chelmsford.*
ILIESCU CONSTANTINE, Rodu Mircea Rodric. *Hounslow.*
JOGINDER SINGH MUDEN. *Delhi.*
JONES, William Byron. *Folkestone.*
JOYCE, Arthur. *Watchet.*
KHAN, Mohammad Salim. *London, N.W.2.*
KITAJEWSKI, Ryszard. *London, N.5.*
KONTOPIADIS, Demetrios. *Athens.*
KORMAZOPOULOS, Anastasios. *Athens.*
KOTWAL, Jogdish Prashad. *Cochin.*
KOUROUKLIS, Paulos Gerasimos. *Athens.*
KRECIOCH, Henryk Teodor. *London, W.12.*
KRIEGER, Moshe. *Tel-Aviv.*
KUCHELA, K. S., M.Sc. *Bangalore.*
KUMAR, Rajindra. *New Delhi.*
KUMARASWAMI, Thiruvalam S., M.A. *Madurai.*
KURUVILLA, Poorakulam M., B.E. *S. Malabar.*
LANG, Robert Charles. *London, S.W.1.*
LAWSON, Bernard John. *Henlow.*
LEONG VIE YING. *Singapore.*
LESNIAK, Janusz. *London, N.W.10.*
MCCONNELL, Geoffrey James. *London, S.E.18.*
MACMILLAN, Hector. *Glasgow, W.1.*
MAH, Seck Wah. *Brighton.*
MASSELOS, Christos. *Athens.*
MAVROMMATIS, John. *Athens.*
MEHTA, Dinshaw M. *Bombay.*
MEMON, Haji M. Haji Ramtoola. *Rajkot.*
MENKAL, Raymond. *London, E.5.*
MICHALOPOULOS, Antonios. *Athens.*
MILLS, Samuel John. *Southsea.*
MOHAMED SHAFI SHAIKH, B.Sc. *Karachi.*
MUKHERJEE, Asoke, M.Sc. *Kyaukpyu, Burma.*
NICHOLSON, Horace Leonard. *Glasgow.*
NIRBHAI SINGH, B.A. *Kuala Lumpur.*
PAPANTONOPOULOS, Panayotis. *Athens.*
PIATEK, Tadeusz-Ludwik. *London, N.4.*
POTTER, Reginald, B.Sc. *Swansea.*
PRAHARAJ, Arakhita. *Bombay.*
RAMDEV, Hardev Sahai. *Delhi.*
RAYBONE, Norman Edward. *Birmingham.*
RICKERS, Denis. *Wrexham.*
ROWAN, David. *Glasgow.*
RYNKIEWICZ, Arthur Edward. *London, N.8.*
SAHARAN, Capt. Manphul Singh, B.A. *Mhow.*
SANTOKH SINGH CHAVEN, B.A. *Punjab.*
SANYAL, Somenath. *Calcutta.*
SASTRY, Kuruganti Venkateswara. *Madras.**
SEN GUPTA, Rabindra Nath. *Bangalore.*
SHAH, Sureshchandra R., B.Sc. *Ahmedabad.*
SHAHI, Brij Lal, B.A. *Amritsar.*
SMITH, Alex Michael. *Weybridge.*
SOOD, Omkar Nath. *Dehra Dun.*
STAVRIDIS, Orpheus. *Athens.*
STRATOS, Therapon. *Athens.*
SUBBARAMAN, M. S., M.A. *Madras.*
SUMMERBELL, Kenneth Wandsworth Browning. *Salisbury North, South Australia.*
TRUONG VAN PHU. *London, N.7.*
TSANG HIN-WA, Albert. *Hongkong.*
TSONIS, Fotios. *Messenian.*
TURNER, Dennis John. *London, S.W.19.*
TZORSZOPOULOS, Constantin. *Athens.*
WAI-CHI WONG. *Southampton.*
WOOD, Leslie Gilbert. *Arborfield.*
WORTH, Ernest William. *Leicester.*
YUNAS, Shiek Muhammad. *London, S.E.27.**

* Reinstatement.

SLOTTED SECTION STANDING WAVE METER*

Techniques of Measurement and Analysis of Results

by

E. M. Wareham (Associate Member) †

SUMMARY

The formal theory of a single discontinuity is briefly stated and developed in terms of the complex reflection coefficient. The concept of multiple reflections is used to extend the theory to cover a transmission line having two discontinuities. Expressions are derived which are then used to analyse the effect of the reflection from the travelling detector in a slotted section on the measurements of standing waves. A convenient method of analysing errors in measurement due to the probe reflection factor is described and charts are given which enable the errors to be calculated.

The techniques for making accurate measurements of v.s.w.r. are discussed for both small and large ratios. Attention is then turned to the problem of accurately determining the position of the standing wave in the line and different techniques are given for small ratios and for large ratios. Methods are suggested for eliminating the inherent errors of the measuring instrument. The determination of normalized impedance or admittance is considered in some detail. Causes of inaccuracy in the results are discussed and techniques for avoiding them are given.

The electrical and mechanical problems involved in the design of waveguide slotted section standing wave meters are discussed and some relations are evaluated. An instrument design by the author is analysed critically in order to demonstrate some of the ways in which the problems may be solved. The methods of test used are described, illustrating the application of the theory outlined.

LIST OF SYMBOLS AND DEFINITIONS

λ Wavelength of radiation in free space.

λ_c Critical or "cut off" wavelength in waveguide. In a rectangular waveguide propagating the dominant (TE_{01}) mode $\lambda_c = 2a$ where a is the broad dimension of the guide.

λ_g Wavelength of radiation in a guide. This is related to the cut-off wavelength and the free-space wavelength as follows:

$$\frac{1}{\lambda_g^2} = \frac{1}{\lambda^2} - \frac{1}{\lambda_c^2}$$

E The maximum value of an alternating electric field.

H The maximum value of an alternating magnetic field.

$\mathbf{H}_z$ The axial component of magnetic field in a waveguide.

$\mathbf{H}_x$ The transverse component of magnetic field in a wave guide.

Field Equations:

The equations relating the field intensities and the physical dimensions of transmission lines. In the case of a coaxial line these are, for the TEM mode:

$$\mathbf{E}_x = \frac{r}{x}\mathbf{E}_r$$ where $\mathbf{E}_r$ and $\mathbf{H}_r$ are the field strengths at the inner conductor.

$$\mathbf{H}_x = \frac{r}{x}\mathbf{H}_r$$

r = radius of inner conductor.

x = radius at which these are defined.

* Manuscript first received 20th December, 1954, and in final form on 9th July, 1955. (Paper No. 330.)
† W. H. Sanders (Electronics) Ltd., Gunnels Wood Road, Stevenage, Hertfordshire
U.D.C. No. 621.317.742:621.372.8.

For a rectangular waveguide TE_{01} mode they are as shown in Fig. 1.

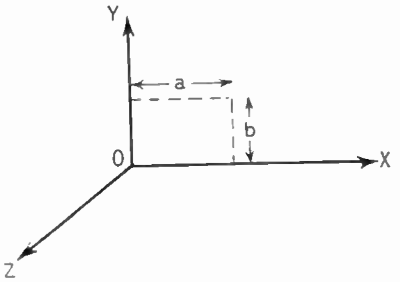


Fig. 1.—Diagram for wave notation in waveguide.

Taking $\mathbf{H}_z$ to be unity:—

$$\mathbf{H}_z = \cos \frac{\pi x}{a}$$

$$\mathbf{E}_y = j \frac{2a}{\lambda_g} . \mathbf{Z}_w . \sin \frac{\pi x}{a}$$

$$\mathbf{H}_x = j \frac{2a}{\lambda_g} . \sin \frac{\pi x}{a}$$

$\mathbf{Z}_w$ Wave impedance. In a coaxial line this is the same as in free space. In a waveguide it is the ratio $\mathbf{E}_y/\mathbf{H}_x$.

$\mathbf{Y}_w$ Wave admittance. $(1/\mathbf{Z}_w.)$

$\mathbf{Z}_0$ Characteristic impedance of a transmission line. In a coaxial line this is:

$$\mathbf{Z}_0 = 138 \frac{\log R/r}{\sqrt{\varepsilon}} .$$

where

R is radius of outer conductor

r is radius of inner conductor

ε is permittivity of the medium

In a waveguide it is:

$$\mathbf{Z}_0 = \sqrt{\frac{\mu}{\varepsilon}} \cdot \frac{\lambda_g}{\lambda} \cdot \frac{b}{a} \text{ for the } TE_{01} \text{ mode}$$

where a, and b, are the broad and narrow dimensions respectively, and μ is permeability of the medium.

$\mathbf{Y}_0$ Characteristic admittance. $(1/\mathbf{Z}_0)$

ρ The voltage (or current) reflection coefficient. A vector quantity defining the ratio of field strength in two waves having the same frequency but travelling in opposite directions in a transmission line.

$$\rho = Re^{j\varphi}$$

where R is a scalar, the modulus, and φ is an angle in radians, the argument.

S The voltage (or current) standing wave ratio. It may be expressed as a number either smaller than or greater than unity. In this paper the latter convention will be used. It is the ratio of the maximum field strength to the minimum field strength in a standing wave pattern.

$$S = \frac{E\ max}{E\ min} = \frac{H\ max}{H\ min} = \frac{E\ forward + E\ reflected}{E\ forward - E\ reflected}$$

$$= \frac{H\ forward + H\ reflected}{H\ forward - H\ reflected} = \frac{1+R}{1-R}$$

γ The propagation constant of a transmission line. This has two component parts, α the attenuation constant (nepers/cm) and β the phase constant (radians/cm). Consider the field strengths at two points in a transmission line separated by a distance l cm. If the line is propagating a pure travelling wave in the dominant mode then:—

$$\frac{\mathbf{H}_1}{\mathbf{H}_2} = \frac{\mathbf{E}_1}{\mathbf{E}_2} = e^{\gamma l} \quad \text{and} \quad \gamma = \alpha + j\beta.$$

PART I. THE ANALYSIS OF DISCONTINUITIES IN A TRANSMISSION LINE

1. Introduction

In a transmission line it can be shown that excitation at one end results in a flow of energy along its length at a finite velocity. Furthermore it is possible for energy to be simultaneously flowing in opposite directions without mutual interference between the two waves. There is an interaction, however, for at each point in the line the electric and magnetic components of the two waves combine to form a single electric and a single magnetic field. If the two waves are of the same frequency, it is found that whilst these fields are alternating in intensity at this frequency, their maximum amplitudes vary from point to point along the line. The pattern of this variation is stationary in space and cyclic in form and is called a standing wave. It is found that the maxima in the electric field pattern correspond to the minima in the magnetic field pattern and conversely. The ratio of the electric and magnetic field strengths defines a constant of the particular line known as the wave impedance. (In the case of a waveguide it is the ratio of the transverse components of the fields.) This constant is related to the characteristic impedance of the line by the physical dimensions. If a line is terminated in its characteristic impedance, all the energy incident upon the termination will be absorbed therein. For all other terminations there will only be a partial absorption of the incident energy and a reflected wave will be excited. Thus a standing wave exists in a line for all values of terminating impedance other than the one special case where it is terminated in its own characteristic impedance. Analysis of the standing wave pattern enables the properties of the termination to be evaluated, and the first section of this paper will formally develop the theory in terms of the complex reflection coefficient :

$$\frac{\text{E } \textit{reflected}}{\text{E } \textit{incident}} = -\frac{\text{H } \textit{reflected}}{\text{H } \textit{incident}}$$

$$= \rho \equiv Re^{j\varphi} \equiv R\cos\varphi + jR\sin\varphi$$

where ρ is a vector quantity whose modulus is R and argument φ.

Fundamentally, the measurement of a standing wave consists of two parts, (*a*) the ratio of maximum to minimum field strength, and (*b*) the position of the standing wave—in effect the measurement of an electrical phase angle. This is most conveniently done by examining either the electric or the magnetic field, but not both. The usual device for this purpose is a slotted section standing wave meter. This, as the name implies, consists of a section of totally enclosed transmission line (e.g. coaxial line or waveguide) in the outer conductor of which an axial slot has been made in such a way as to cause negligible disturbance to the wave in the tube. Projecting through the slot into the tube, a small probe is usually arranged so that it couples to the electric field only. The coupling is made small in order to avoid excessive disturbance to the wave and some suitable external arrangement is provided to enable the energy extracted by the probe to be detected. A facility is also provided for the probe to be moved along the slot thus enabling it to sample the electric field strength over a section of the length of the line. It is the object of the first part of this paper to demonstrate the existence of factors affecting the accuracy of measurement with instruments of this type. As far as possible, the error factors so revealed will be evaluated and, in the second section, which deals with techniques of measurement, means will be considered whereby such errors can be avoided. In the third section, the design and construction of a waveguide standing wave meter will be discussed in some detail so as to reveal the causes and probable magnitudes of the error factors. The paper concludes with some notes on typical techniques whereby standing wave meters can be used to analyse themselves.

2. General Theory

The analysis is based on three premises which are strictly in accordance with the general transmission line theory and are therefore applicable to all types of uniform line, in which only the dominant mode is propagated.

1st Premise.—In a travelling wave the ratio of the magnetic and electric field strengths shall be taken as defining a constant with the dimension of an admittance which shall be a characteristic of the particular type of transmission line that is being considered, e.g. coaxial, waveguide, etc.

For convenience a unit line will be analysed, that is, its characteristic admittance will be taken as unity and it will be considered to be excited with unit power.

Thus in a travelling wave

$$E = H = \sqrt{P} = 1 \qquad (1)$$

$$\text{Y}_0 = H/E = 1 \qquad (2)$$

where P = power delivered to the line. (Incident wave).

2nd Premise.—In a wave travelling away from the generator the electric and magnetic fields will be considered to be in time phase, whilst in a wave travelling towards the generator they will be considered to be in antiphase.

3rd Premise.—The field strength at a discontinuity is the vector difference between that of the travelling wave incident upon it and that of the travelling wave reflected from it. The ratio of the reflected and incident electric fields shall be defined in phase and magnitude as $\rho = Re^{j\varphi}$ and the ratio of reflected and incident magnetic fields as $\rho = Re^{j\varphi+\pi}$.

Consider now a transmission line excited by a matched source and terminated in a matched load ($\mathbf{Y} = \mathbf{Y}_0$) and having a discontinuity defined by the reflection co-efficient ρ. (Fig. 2)

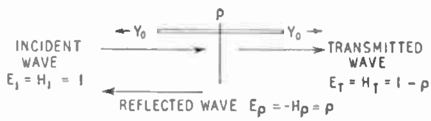


Fig. 2.

Power travelling in either direction from the discontinuity will be totally absorbed in the source and termination respectively. As a plane of reference situated at a distance l cm towards the generator from the discontinuity the incident wave will be defined as unity and the reflected wave will therefore be $\rho e^{2\gamma l}$ where γ is the propagation constant of the line. At this plane the ratio of reflected to incident field strengths will be

$$\left.\begin{aligned}\frac{\mathbf{E}_\rho}{\mathbf{E}_1} &= Re^{2\gamma l}\,.\,e^{j\varphi}\\ \frac{\mathbf{H}_\rho}{\mathbf{H}_1} &= -Re^{2\gamma l}\,e^{j\varphi}\end{aligned}\right\} \tag{3}$$

where $\gamma = \alpha + j\beta$ and $\beta = 2\pi/\lambda_g$ radians.

From this it can be seen that as l is varied, the phase of the reflected wave changes with respect to that of the incident wave in such a way that a complete cycle is made after l has changed by $\lambda_g/2$.

At the plane of reference the field strengths are, by definition

$$\left.\begin{aligned}\mathbf{E} &= \mathbf{E}_1 - \mathbf{E}_\rho = 1 - \rho e^{2\gamma l}\\ &= 1 - Re^{2\alpha l}\cos(\varphi + 2\beta l)\\ &\quad - jRe^{2\alpha l}\sin(\varphi + 2\beta l)\\ \mathbf{H} &= \mathbf{H}_1 + \mathbf{H}_\rho = 1 + \rho e^{2\gamma l}\\ &= 1 + Re^{2\alpha l}\cos(\varphi + 2\beta l)\\ &\quad + jRe^{2\alpha l}\sin(\varphi + 2\beta l)\end{aligned}\right\} \tag{4}$$

From this it can be seen that for $\varphi + 2\beta l = 0$, 2π etc. $\mathbf{E}$ is a minimum and $\mathbf{H}$ is a maximum. Also that for $\varphi + 2\beta l = \pi$, 3π etc. the converse is the case. It follows that the maximum field strength varies in a stationary pattern along the length of the line. This is the standing wave and the standing wave ratio is the ratio of maximum to minimum values (or conversely depending on the convention used) of the expression. Thus we have

$$S = \frac{1 + R}{1 - R} \quad \text{and hence} \quad R = \frac{S - 1}{S + 1} \tag{5}$$

The ratio of $\mathbf{H}$ to $\mathbf{E}$ at the plane of reference defines an admittance having value

$$\mathbf{Y} = \frac{1 + Re^{2\alpha l}\cos(\varphi + 2\beta l) + jRe^{2\alpha l}\sin(\varphi + 2\beta l)}{1 - jRe^{2\alpha l}\cos(\varphi + 2\beta l) - jRe^{2\alpha l}\sin(\varphi + 2\beta l)} \tag{6}$$

Now unless the particular line has a very high attenuation constant it is usual to ignore the term $e^{2\alpha l}$ and this will be done henceforth unless its retention is necessary. Resolving into real and imaginary parts to give an expression in the form of $\mathbf{Y} = G + jB$ and ignoring the term $e^{2\alpha l}$ we have :—

$$\mathbf{Y} = \frac{1 - R^2}{1 - 2R\cos(\varphi + 2\beta l) + R^2} + \frac{j\,2R\sin(\varphi + 2\beta l)}{1 - 2R\cos(\varphi + 2\beta l) + R^2} \tag{7}$$

This shows that $\mathbf{Y}$ varies cyclically, becoming real and minimum at values of $\varphi + 2\beta l = \pi$, 3π, 5π etc., and real and maximum at values of 0, 2π, 4π, etc. The maximum and minimum values of $\mathbf{Y}$ are respectively :

$$\left.\begin{aligned}\mathbf{Y}_{max} &= \frac{1 + R}{1 - R}\\ \mathbf{Y}_{min} &= \frac{1 - R}{1 + R}\end{aligned}\right\} \tag{8}$$

By setting $\beta l = 0$, the plane of reference can be transferred to the plane of the discontinuity, which is usually the most convenient place to define its admittance. Knowing the wavelength λ_g in the transmission line and the physical distance between a minimum in the pattern (where $\varphi + 2\beta l = 0$), and the discontinuity, the value of φ can be found. Then from a measurement of the standing wave ratio, R can be evaluated and the admittance at the plane of the discontinuity can be determined. This is made up of two parts :

$$\mathbf{Y} = \mathbf{Y}_0 + \mathbf{Y}_D$$

from which the admittance of the discontinuity alone can be derived, and as we are considering the unit line ($\mathbf{Y}_0 = 1$) :

$$\frac{\mathbf{Y}_D}{\mathbf{Y}_0} = \frac{\mathbf{Y}}{\mathbf{Y}_0} - 1 \tag{9}$$

Substituting in eqn. (7) this gives

$$\frac{\mathbf{Y}_D}{\mathbf{Y}_0} = \frac{2R(\cos\varphi - R)}{1 - 2R\cos\varphi + R^2} + j\,.\,\frac{2R\sin\varphi}{1 - 2R\cos\varphi + R^2} \tag{10}$$

Expressed in these terms the admittances are said to be normalized to the characteristic admittance of the line and have the dimensions of complex numbers. The absolute values are obtained by multiplying by Y_0 which is generally calculated from the physical dimensions of the particular line being used. It is usual, however, to consider normalized admittances only, and this practice will be used throughout this paper, the symbols for admittance and impedance being retained only for clarity.

It is possible to maintain the concept of two waves travelling in opposite directions in a line without loss of factual reality. Two waves travelling in the same direction with the same frequency however are no more than a mathematical concept. The field components add together vectorially to form a single wave from which the original components cannot be separated by any artifice or device save a mathematical manipulation. This fact leads to a powerful method of analysis described by L. G. H. Huxley.*

3. The Case of Two Discontinuities

Consider a transmission line with propagation constant γ, and having two discontinuities ρ_1 and ρ_2 separated by a distance l cm. Let the incident electric wave be unity, then there will be a reflected wave ρ_1 at the first discontinuity and a travelling wave $(1 - \rho_1)$ beyond it. This will give rise to a second reflected wave $\rho_2 (1 - \rho_1) e^{\gamma l}$ at the second discontinuity and a final travelling wave $(1 - \rho_1) . (1 - \rho_2) . e^{\gamma l}$ beyond it. Multiple reflections will take place between the two discontinuities as shown in Fig. 3.

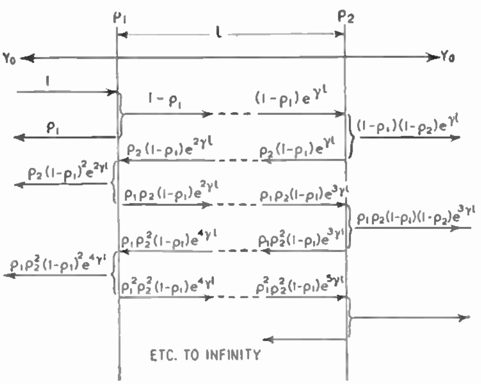


Fig. 3.—Diagrammatic representation of multiple reflections between two discontinuities in a transmission line.

* L. G. H. Huxley, "A Survey of the Principles and Practices of Waveguides", Section 5.4 (Cambridge University Press, 1947).

After an infinite time the total reflected wave to the generator will be the sum of the series

$$\rho = \rho_1 + \rho_2 (1 - \rho_1)^2 e^{2\gamma l} + \rho_1 \rho_2^2 (1 - \rho_1)^2 e^{4\gamma l} + \ldots\ldots \text{ to } \infty$$

$$= \rho_1 + \rho_2 (1 - \rho_1)^2 e^{2\gamma l} \left[\begin{array}{l} 1 + \rho_1\rho_2 e^{2\gamma l} + \\ \rho_1^2 \rho_2^2 e^{4\gamma l} + \text{ to } \infty \end{array} \right]$$

This is a geometric progression with the solution:

$$\rho = \rho_1 + \frac{\rho_2 (1 - \rho_1)^2 e^{2\gamma l}}{1 - \rho_1 \rho_2 e^{2\gamma l}} \tag{11}$$

Hence

$$\rho = \frac{\rho_1 + \rho_2 e^{2\gamma l} (1 - 2\rho_1)}{1 - \rho_1 \rho_2 e^{2\gamma l}} \tag{12}$$

A similar analysis of the travelling wave from ρ_1 to ρ_2 gives, at the plane of ρ_1

$$E_T = \frac{1 - \rho_1}{1 - \rho_1 \rho_2 e^{2\gamma l}} \tag{13}$$

and for the wave travelling from ρ_2 to ρ_1 at the plane of ρ_1

$$E_\rho = \frac{\rho_2 (1 - \rho_1) e^{2\gamma l}}{1 - \rho_1 \rho_2 e^{2\gamma l}} \tag{14}$$

Taking the ratio of these two waves gives the value of the apparent reflection coefficient of ρ_2 and hence the standing wave in this part of the line :—

$$\frac{E_\rho}{E_T} = \rho_2 e^{2\gamma l}.$$

This illustrates an important basic principle. Suppose the transmission line has a reflection coefficient of ρ_1 at its source and is terminated by an admittance giving a reflection coefficient ρ_2, then "the standing wave ratio on the line is independent in phase and magnitude of the source admittance."

4. The Accuracy of Admittance Measurements

Having established a general theory we may now turn to the specific effect of discontinuities on the accuracy of admittance measurements. It is clear from the foregoing analysis that in a line with a matched source having a discontinuity ρ_1 at a distance l from termination ρ_2 being measured, there will be an error in the measurement due to the presence of ρ_1. Equation (11) defines the reflected wave which would result at the plane of ρ_1. From this it can be seen that $\rho_2 e^{2\gamma l}$ is modified by a factor $(1 - \rho_1)^2/(1 - \rho_1\rho_2 e^{2\gamma l})$ and added to ρ_1. Thus if ρ_1 is larger than ρ_2 it can dominate the result. Even if ρ_1 is smaller than ρ_2, it can still cause serious uncertainty in a measurement for it may have any phase relationship to $\rho_2 e^{2\gamma l}$.

It will be shown later that such discontinuities may arise from the end of the slot in a standing wave meter or a change of characteristic admittance at the terminal coupling.

Standing wave meters in general have a travelling detector which is sensitive to only one component of the fields in the line, usually the electric field. (In this case a small aerial probe projects into the line to form the coupling element.) As the coupling consists of only one element it can never be reflectionless and can always be resolved into an admittance, across the line at that point. By designating the probe reflection co-efficient as ρ_1 and the termination as ρ_2, eqns. (12), (13) and (14) can be used to analyse the conditions which will obtain. The field strength at the plane of ρ_1 is, by definition, the vector difference between the incident and reflected waves at that plane, so that

$$\mathbf{E}_P = 1 - \rho = \frac{1 - \rho_1 - \rho_2 e^{2\gamma l} + \rho_1\rho_2 e^{2\gamma l}}{1 - \rho_1\rho_2 e^{2\gamma l}}$$

from eqn. (12)

and with a little manipulation :

$$\mathbf{E}_P = \left[1 - \rho_1\right] \left[\frac{1 - \rho^2 e^{2\gamma l}}{1 + \rho_1\rho_2 e^{2\gamma l}}\right] \quad (15)$$

Considering an ideal case in which the source is matched and the probe reflectionless, i.e. $\rho_1 = 0$ we have :

$$\mathbf{E}_P = 1 - \rho_2 e^{2\gamma l}$$

By moving the travelling detector along the line, maxima in the detected signal will occur when $(\varphi_2 + 2\beta l) = \pi$, 3π etc. and minima when $(\varphi_2 + 2\beta l) = 0$, 2π, etc., the maximum and minimum values being $1 + R_2 e^{2\alpha l}$ and $1 - R_2 e^{2\alpha l}$ respectively. Reference to eqn. (5) shows that a perfect measurement is thus obtained apart from the effect due to the attenuation of the line. In practice ρ_1 is never zero and an attempt is usually made to tune out the reactive component of the probe admittance by means of suitable adjustments in the probe circuit of the travelling detector. The second case to consider is the one in which this is achieved, in which case $\varphi_1 = 0$. This gives

$$\mathbf{E}_P = \left[1 - \rho_1\right] \left[\frac{1 - R_2 e^{2\alpha l} e^{j(\varphi + 2\beta l)}}{1 - R_1 R_2 e^{2\alpha l} e^{j(\varphi + 2\beta l)}}\right].$$

There is no phase error because this expression is real, and maximum and minimum when $(\varphi_2 + 2\beta l) = \pi$, 3π etc. and 0, 2π etc. respectively and the measured v.s.w.r. is :—

$$S_m = \left[\frac{1 + R_2 e^{2\alpha l}}{1 - R_2 e^{2\alpha l}}\right] \left[\frac{1 - R_1 R_2 e^{2\alpha l}}{1 + R_1 R_2 e^{2\alpha l}}\right] \quad (16)$$

This shows that the measured standing wave ratio is smaller than the true value.

The value of the v.s.w.r. is $S' = \dfrac{1 + R_2 e^{2\alpha l}}{1 - R_2 e^{2\alpha l}}$

and the error term is therefore

$$\frac{S_m}{S'} = \frac{1 - R_1 R_2 e^{2\alpha l}}{1 + R_1 R_2 e^{2\alpha l}} \quad (17)$$

When S' is large it is necessary to allow for the attenuation of the line in order to find the true v.s.w.r. at the reference plane

$$S = \frac{S'(1 + e^{2\alpha l}) - (1 - e^{2\alpha l})}{(1 + e^{2\alpha l}) - S'(1 - e^{2\alpha l})} \quad (17a)$$

Here l = length of line between the minimum and the reference plane.

It often happens in practice that the tuning adjustments in the probe circuit are insufficient to reduce φ_1 to zero. The error term then becomes complex giving rise to both phase and magnitude errors in the measurement. A rigorous analysis of the standing wave pattern would be extremely laborious under these conditions, but by making the restriction that R_1 is small (less than 0·1) equations (4) and (15) can be used to define the pattern traced by the probe. (See appendix.)

$$\mathbf{E}_P = \frac{(1 - 2R_2\cos\theta + R_2^2)^{\frac{1}{2}}}{1 - R_1 R_2 \cos(\theta + \varphi_1)} \quad (18)$$

where $\theta = \varphi_2 + 2\beta l$.

If we now introduce to this expression a term δ to define the phase error at the maximum and minimum, then, by differentiating, the maxima and minima are found to occur when $\theta = \pi + \delta_{max}$ and $\theta = \delta_{min}$ respectively where

$$\begin{aligned}\delta_{max} &\simeq \sin\delta_{max} \simeq -(1+R_2)^2 R_1 \sin\varphi_1 \text{ rad}\\ \delta_{min} &\simeq \sin\delta_{min} \simeq (1-R_2)^2 R_1 \sin\varphi_1 \text{ rad}\end{aligned} \quad (19)$$

In this, $R_1 \sin\varphi_1$ is the reactive component of the probe reflection coefficient and it is of interest to note that the magnitude of the phase error also depends on the terminating reflection. The maximum of the pattern is shifted more than the minimum, where the phase error becomes zero when $R_2 = 1$. Also the error changes sign with the result that the pattern becomes lopsided and asymmetrical about both maximum and minimum. The magnitude error factor in the v.s.w.r. is approximately,

$$\frac{S_m}{S} \simeq \frac{1 - R_1 R_2 \cos\varphi_1}{1 + R_1 R_2 \cos\varphi_1} \quad (20)$$

From eqns. (19) it is clear that phase measurements should always be made at a minimum. Errors of the order of 0·1 radian are the greatest that may occur, but extremely adverse conditions are required to produce errors of this order, and more usual figures would be 0·01 radians. In a good modern measuring instrument, however, an accuracy of 0·001 radian may be required and it is then necessary to minimize $R_1 \sin \varphi_1$ by tuning the detector accurately to resonance.

In the previous paragraphs we have discussed the effect of interaction between the probe and the termination. If the source is not matched, however, a second interaction term appears. Suppose that the source has a reflection coefficient ρ_3, then the field strength at the probe becomes modified due to the interaction between its reflection and the source reflection. This gives rise to a factor

$$\Delta_P = \frac{1}{1 - R_1 R_3 e^{j(\varepsilon - \theta)}} \tag{21}$$

where $\varepsilon = \varphi_1 + \varphi_2 + \varphi_3 + \beta l_1$

and $\theta = \varphi_2 + 2\beta l$

and l_1 = length of line between source and termination. This is equivalent to a change in the probe coupling factor as the travelling detector moves along the line and the pattern traced by the probe now becomes

$$E'_P = \left[1 - \rho_1\right] \left[\frac{1 - R_2 e^{j\theta}}{1 - R_1 R_3 e^{j(\varepsilon - \theta)}}\right] \tag{22}$$

(Note that the previously derived interaction term between load and probe has been omitted for simplicity.) This expression can have maximum and minimum values such that

$$\frac{1 - R_1 R_3}{1 + R_1 R_3} \leq \frac{S_m}{S} \leq \frac{1 + R_1 R_3}{1 - R_1 R_3} \tag{23}$$

It follows that it is never possible to make a perfect measurement and even in the roughest work it is always necessary to make a compromise between the power coupled to the travelling detector and the resulting errors in measurement. The best that can be done is to tune the probe to resonance and to reduce the coupling as far as possible. A coupling coefficient of −26 db gives a very reasonable compromise for the most accurate instruments whilst for lower grade measurements a figure of perhaps −12 db could be tolerated with a resonant probe. With the same restriction as before ($R_1 < 0{\cdot}1$) the value for θ can be found at maxima and minima in the pattern, where the following relationship must obtain:—

$$R_2 \sin \theta + R_1 R_2 R_3 \sin (\varepsilon - 2\theta) - R_1 R_3 \sin(\varepsilon - \theta) = 0 \quad \ldots\ldots\ldots(24)$$

it is worth noting that if $R_2 = 1$ or $R_1 R_3 = 0$ or $\varepsilon = n\pi$, z rosoccur when $\theta = 0, \pi, 2\pi$, etc.

When R_2 is quite small the shape of the standing wave pattern is given approximately by

$$E'_P = 1 - R_2 \cos \theta - R_1 R_3 \cos (\varepsilon - \theta) \quad \ldots(25)$$

In this expression $R_2 \cos \theta$ is the standing wave and when it is small it will finally be swamped by the third term. Under these conditions an unwary operator attempting to measure a well-matched termination may merely be measuring the interaction between the probe and the source. This fact is also shown clearly in equation (23) by putting S near to unity.

The probe coupling coefficient to a travelling wave is given by

$$C \cong 10 \log_{10} \left(\frac{2R_1 \cos \varphi_1}{1 - 2R_1 \cos \varphi_1 + R_1^2} \right) \text{db} \quad \ldots(26)$$

so that for a given coupling the probe reflection coefficient is a minimum when it is tuned to resonance and $\varphi = 0$.

5. Analysis of Errors

A convenient method of analysing the errors in measurement which will arise due to the reflection from the travelling detector is as follows:

A conventional waveguide measuring circuit is shown in Fig. 4 with the addition of an auxiliary slotted section or fixed probe situated between the calibrated attenuator and the slotted section to be analysed. The auxiliary probe is connected to a crystal detector and the meter M2, and provision is made to measure the length of line, l, between this and the probe of the slotted section.

Sufficient padding should be introduced by A1 and A2 for the source as seen from the auxiliary probe to be well matched (say v.s.w.r. 1·1 : 1). If the travelling detector of the section under test is moved along, the reflection from it will change in phase relative to the forward wave at the auxiliary probe causing a variation of M2. This variation will be substantially independent of any other reflections in the system not associated with the travelling detector provided they are fairly small. Denoting the reflection from the travelling detector as $R_1 e^{j\varphi_1}$ as before, it follows that the maximum value of M2 will be proportional to $1 + R_1$

and the minimum to $1 - R_1$. The ratio of these two values will be the v.s.w.r., S_1, due to the travelling detector. From this R_1 may be calculated :

$$R_1 = \frac{S_1 - 1}{S_1 + 1}$$

By a measurement of l at the minimum of M2, it is then possible to calculate φ_1 by use of eqn. 4. Fig. 5 gives a diagram in which l/λ_g and S_1 can be used to calculate $R_1 \cos \varphi_1$ and $R_1 \sin \varphi_1$. An additional scale is shown against $R_1 \cos \varphi_1$ which indicates the coupling loss of the travelling detector in db (assuming that it has no internal attenuation). Fig. 6 shows a family of error curves giving S_m/S as a function of $1/S$ for various values of $R_1 \cos \varphi_1$, and Fig. 7 shows a family of error curves giving δ_{min} and δ_{max} as a function of $1/S$ for various values of $R_1 \sin \varphi_1$.

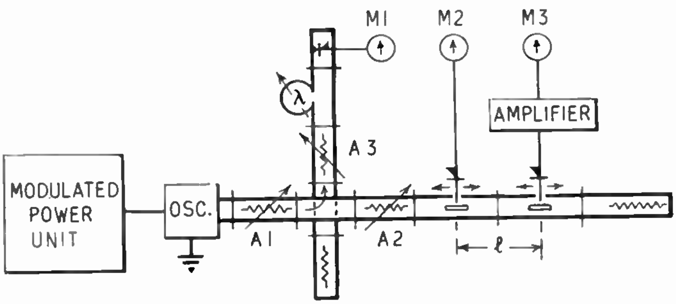


Fig. 4.—Conventional waveguide measuring circuit.

6. Conclusions to Part 1

The principal sources of error have been shown to arise from

(*a*) A discontinuity between the travelling detector and the termination being measured.

(*b*) Interaction between the detector and the termination, the real component of the detector admittance contributing a magnitude error and the imaginary component a phase error.

(*c*) Interaction between the detector and the source. In this case a periodic variation is added to the standing wave, the magnitude of which is, to a first order, independent of the standing wave itself. This could in extreme cases lead to totally false readings both of phase and magnitude of v.s.w.r. near to unity.

7. Appendix to Part 1

From eqn. (15)

$$E_P = (1 - \rho_1)\,\frac{1 - R_2\,e^{j(\varphi - 2\beta l)}}{1 - R_1 R_2\,e^{j(\varphi_1 + \varphi_2 + 2\beta l)}}$$

Now $(1 - \rho_1)$ is independent of θ and so does not effect the shape of the standing wave. In the second bracket the numerator defines the standing wave and the denominator is the error term. If we make the restriction that $R_1 < 0{\cdot}1$ then the fraction can be treated, with fair accuracy, as a scalar quantity instead of a vector. Thus the pattern recorded by the detector will be:—

$$E_P = K\,.\,\frac{(1 - 2R_2 \cos\theta + R_2^2)^{\frac{1}{2}}}{(1 - 2R_1R_2\cos(\theta + \varphi_1) + R_1^2R_2^2)^{\frac{1}{2}}}$$

$$\simeq K\,.\,\frac{(1 - 2R_2\cos\theta + R_2^2)^{\frac{1}{2}}}{1 - R_1R_2\cos(\theta + \varphi_1)}$$

where $\theta = \varphi_2 + 2\beta l$

Differentiating with respect to θ and equating to zero,

$$[1 - R_1R_2\cos(\theta + \varphi_1)] + [1 - 2R_2\cos\theta + R_2^2]^{\frac{1}{2}} R_2 \sin\theta$$
$$= [1 - 2R_2\cos\theta + R_2^2]^{\frac{1}{2}}\,[R_1R_2\cos\varphi_1\sin\theta - R_1R_2\sin\varphi_1\cos\theta]$$
$$R_2\sin\theta - R_1R_2^2\cos(\theta + \varphi_1)\sin\theta = [1 - 2R_2\cos\theta + R_2^2]\,[R_1R_2\cos\varphi_1\sin\theta - R_1R_2\sin\varphi_1\cos\theta].$$

Now put $\theta = n\pi + \delta$ with the restriction that $0{\cdot}1 > \delta > -0{\cdot}1$ and eliminate second order terms.

$$\sin\delta = [1 - 2R_2\cos(\delta + n\pi) + R_2^2] \times [R_1\sin\varphi_1\cos(\delta + n\pi)]\,.$$

If n is even, E_P will be near a minimum and

$$\sin\delta_{min} = (1 - R_2)^2\,R_1\sin\varphi_1$$

If n is odd, E_P will be near a maximum and

$$\sin\delta_{max} = -(1 + R_2)^2 R_1\sin\varphi_1.$$

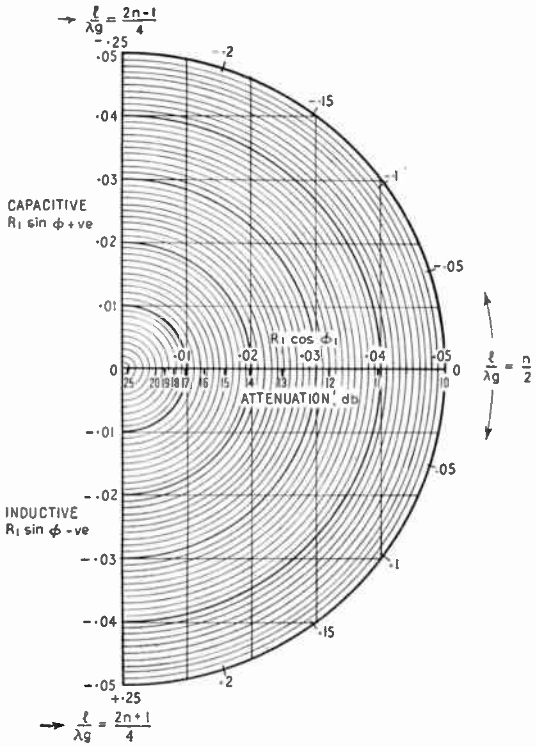


Fig. 5.—Analysis of probe reflection in standing wave meter chart for determining R_1 sin φ_1 and R_1 cos φ_1 from R_1 and the distance l between the s.w.m. probe and the auxiliary measuring probe (see Fig. 4).

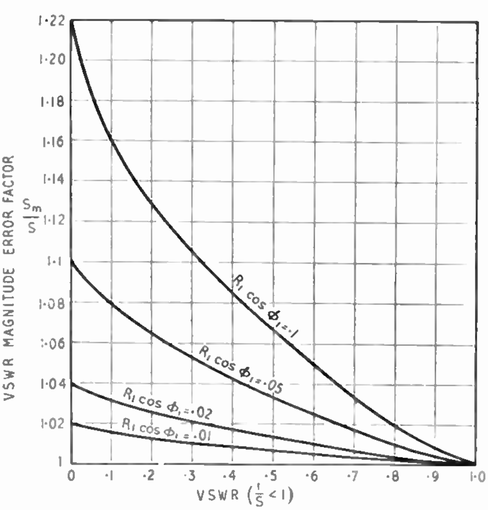


Fig. 6.—Chart showing v.s.w.r. magnitude error factor as a function of the real part of the probe reflection coefficient and the v.s.w.r. (1/S<1).

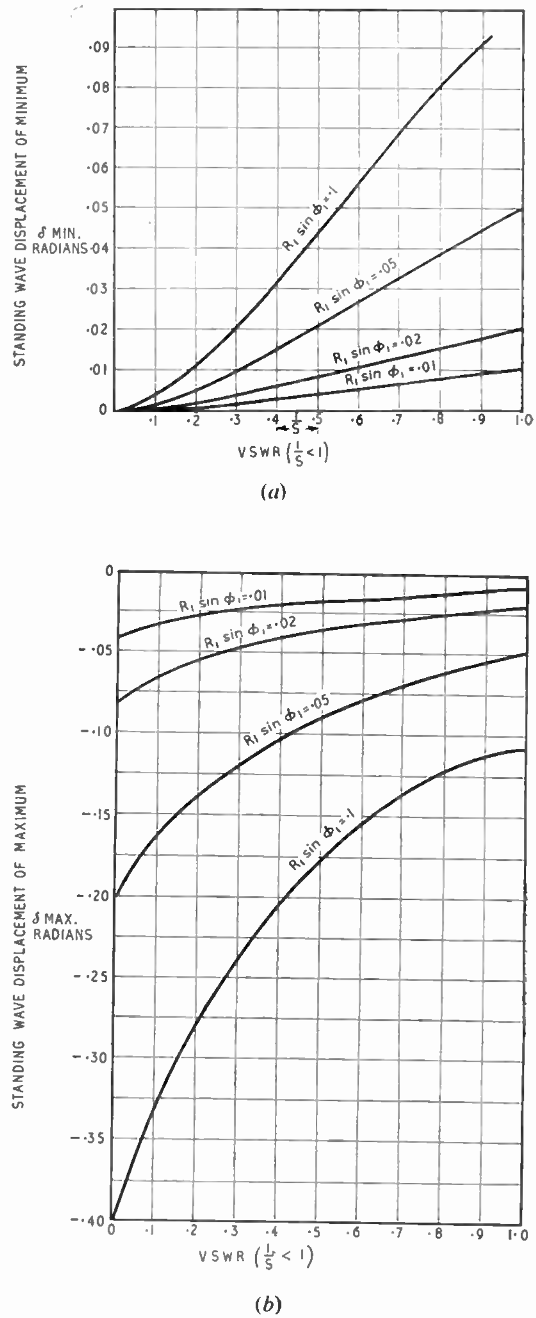


Fig. 7.—Charts showing displacement of (a) the minima, (b) the maxima, as a function of the reactive component of the probe reflection coefficient at the v.s.w.r. 1/S. (S>1).

Part 2—TECHNIQUES OF MEASUREMENT AND ANALYSIS OF RESULTS

8. Measurement of Small Standing Wave Ratios

The two principal causes of error in the measurement of small v.s.w.r. are (1) variation of the power from the signal source during the measurement, and (2) variation of coupling of the probe to the line as the travelling detector is moved along.

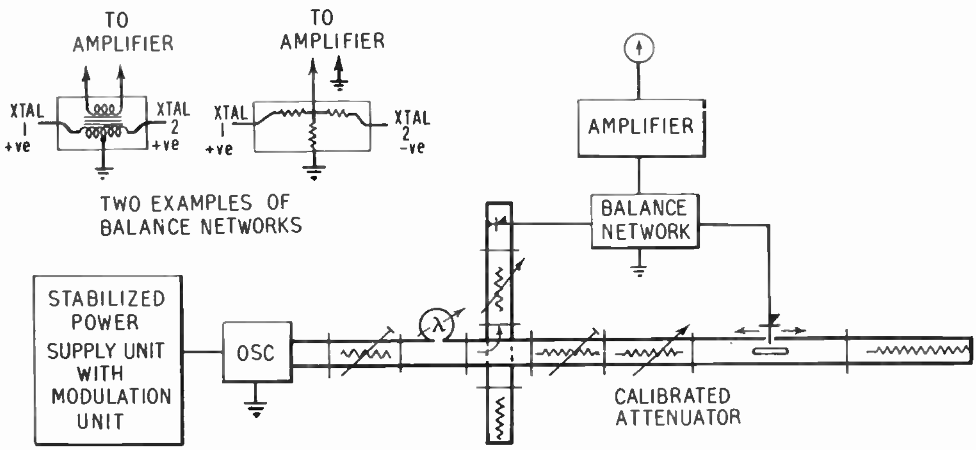


Fig. 8.—A balanced bridge arrangement in which two arms are at r.f. in waveguide and the other two are at d.c. in wire. Crystal detectors are used for the conversion from one form to the other.

The first of these may be dealt with fairly readily by setting up a bridge system as shown in Fig. 8. By balancing the output from the fixed detector against the travelling detector in a transformer or other network, and amplifying the difference, the sensitivity of indication can be very considerably increased. The bridge can be operated in the balanced condition and the measurements of v.s.w.r. made with a calibrated r.f. attenuator in one arm, or, alternatively, it can be operated unbalanced with the output meter calibrated from an r.f. attenuator. There is little to choose between the two methods, both depend upon the characteristics of the detectors being matched to one another. The latter, however, has an additional error in that a fractional change in the output from the signal source results in the same fractional change of the meter reading. In practice the signal source is usually stable to better than 1 per cent. and the crystals are rarely matched to better than 10 per cent. As a result it calls for very careful adjustment to obtain better than 10:1 improvement in sensitivity with such a bridge, that is to say a reading accuracy of 0·005 db r.f., and stability of the same order. Under these conditions the unbalanced bridge has a slight advantage because it is rapid in operation and gives direct visual indication of the variations being measured. A few practical hints may be worth while. In general, it will be found that the silicon crystal rectifiers used at these frequencies are sensitive to temperature and signal level. Quite small variations of temperature can produce significant changes in the rectification efficiency, e.g. in one particular rectifier the change is as high as 1 per cent. per ° C. It is also found that the crystal impedance as seen from the r.f. side and from the output side varies considerably with signal level. In addition the crystals show random fluctuations of output at low level. In short the operation of the bridge depends upon a steady ambient temperature and a high power level at the crystals. Most silicon crystals will require about 100 microwatts r.f. drive for a stability of indication of 0·005 db. r.f.

Variations in coupling as the travelling detector moves along are more difficult to deal with. These fall into two groups: (*a*) random variations such as those due to variable contact, r.f. leakage and looseness in the moving mechanism, etc.; and (*b*) systematic variations which are repeatable and are directly related to the position of the carriage. In a good instrument the first group should be small in magnitude compared with the second. They are impossible to deal with apart from the usual technique of making several measurements and averaging. Those in the second group, being repeatable, can be measured and, after calibration, measurements taken with the instrument can be corrected. The technique recommended for this will be discussed in a later section dealing with calibration of standing wave meters. In general the operator will prefer to make a visual assessment of the v.s.w.r. if it is very near to unity rather than make a series of measurements and corrections. In this case he may move the detector over its full travel and average the successive ratios observed. Alternatively he may prefer to average all the maxima and all the minima and calculate the ratio. This gives a slightly different result (see Fig. 9), but the one method will suit the temperament of one operator whilst the other will please another. Broadly speaking, however, it may be suggested that according to current standard practice there is little point in measuring the v.s.w.r. nearer to unity than 1·005 with any accuracy, except in the calibration of the measuring instrument.

9. Measurement of Large Standing Wave Ratios

The most obvious and certainly the best method is simply to measure the ratio of minimum to maximum as one operation with a calibrated r.f. attenuator. In this the travelling detector is set on a minimum and a convenient reading obtained on the output meter. The detector is then moved to a maximum and the change in insertion of the calibrated attenuator required to bring the meter back to its original reading is noted. By this method the rectifier law does not require to be known. It is important to monitor the output of the oscillator and to ensure that there is adequate padding between the source and the calibrated attenuator which itself should be well matched at all settings, otherwise there is a risk of the oscillator output changing during the measurement. This is due to a change in the impedance presented to the oscillator due to the change in the degree of padding provided by the calibrated attenuator during adjustment. It has also been shown in Part 1 that a change in the source impedance as seen by the probe can also change the magnitude of the forward wave.

It sometimes happens that the sensitivity of the output indicator is not sufficient to enable the minimum of the pattern to be clearly indicated. Under these conditions, and when efforts to obtain more r.f. power have failed, there is a technique which does not require the actual value of the minimum to be measured directly. This makes use of a calibrated r.f. attenuator as before, which establishes the ratio between the maximum and some suitable level which is near the minimum but clearly visible on the output meter. The phase gear is then used to measure the electrical width of the pattern at that level and from this information the v.s.w.r. is calculated. The method is probably the most accurate of the many variations on the same theme but even so it can only extend the measurement of the attenuator by some 10 db with certainty. An extension of 20 db is the limit, beyond which results are meaningless. This is due to the fact that the measurement of the electrical angle becomes so critical beyond this point that it is beyond the accuracy of even the best of instruments without excessive and laborious precautions. Fig. 10 illustrates the method. By inspection this equation can be made up:

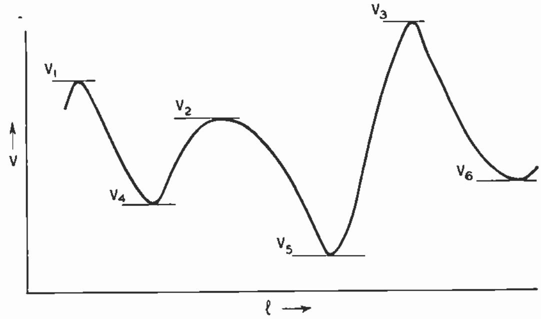


Fig. 9.—Methods of averaging v.s.w.r. near to unity when distorted by variation of probe coupling.

Method 1 : $$\frac{\frac{V_1}{V_4}+\frac{V_2}{V_4}+\frac{V_2}{V_5}+\frac{V_3}{V_5}+\frac{V_3}{V_6}}{5} = \textit{average v.s.w.r.}$$

Method 2 : $$\frac{V_1+V_2+V_3}{V_4+V_5+V_6} = \textit{average v.s.w.r.}$$

$$\left[\frac{E_3^2-E_1^2}{2}\right](1-\cos\theta)+E_1^2=E_2^2$$

Rearranging this we have

$$E_3^2\sin^2\tfrac{1}{2}\theta+E_1^2\cos^2\tfrac{1}{2}\theta=E_2^2$$

Whence $$\frac{E_3^2}{E_1^2} = S^2 = \frac{\cos^2 \frac{1}{2}\theta}{(E_2/E_3)^2 - \sin^2 \frac{1}{2}\theta} \qquad ..(27)$$

From the form of the expression it is now clear why the measurement of θ must be very accurate. Two examples will suffice to illustrate the order of accuracy needed.

(*a*) v.s.w.r. $= 100:1$ $(E_3/E_2)^2 = 1{\cdot}6 \times 10^3$ ($E_3/E_2 = 32$db). θ is found to be 0·032 radian and an error of 0·001 radian gives an error of 1·5 db in the answer. $-0{\cdot}001$ radian gives an error of $-1{\cdot}5$ db in the answer.

(*b*) v.s.w.r. $= 100:1$ $(E_3/E_2)^2 = 10^2$ ($E_3/E_2 = 20$db). θ is found to be 0·141 radian and an error of 0·001 radian makes the answer infinity. $-0{\cdot}001$ radian gives an error of -3db in the answer.

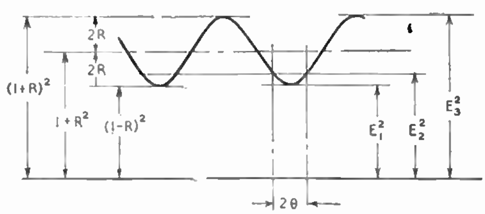


Fig. 10.—Measurement of large v.s.w.r. when minimum cannot be measured.

By definition:
$$E^2 = 1 - 2R\cos\varphi + R^2$$
$$S^2 = \frac{(1 + R^2)}{(1 - R)^2} = \frac{E_3^2}{E_1^2}$$

By measurement: $\left(\frac{E_3^2}{E_2}\right)$ *and* $2\,\theta$.

From the above considerations it is to be expected that a good result requires the physical measurement of length to be equivalent to an accuracy of better than 0·001 radian, the stability of frequency to be of the order of one part in 10^4 and the wavelength to be known with an accuracy of the order of one part in 10^3. Care is also necessary to ensure that the probe reflection does not distort the shape of the pattern and lead to a wrong result. On the whole, this is not a very encouraging picture as the technique demands an advanced knowledge of the capabilities of the measuring instrument. Yet the author believes that it is more accurate than any of the other ones in which only two measurements are made.

In view of the high accuracy required in the measurement of θ, it would certainly seem advisable to take particular pains with this part of the measurement. One way of improving the results considerably is to take several measurements at different output levels and plot them on a graph. The best curve can then be fitted to the points.

10. Measurement of Position of Standing Wave

Quite a lot of useful information can be derived from a knowledge of the position of the standing wave with respect to some convenient plane of reference. As in the measurement of v.s.w.r., the smaller the standing wave the greater is the effect of the errors of the instrument upon the answer. In current practice it may be assumed that when the v.s.w.r. is less than 1·005 : 1 position measurements have but slight meaning or usefulness. This is due to the fact that there are always small discontinuities and variations of coupling in a standing wave meter and these displace and distort the measured pattern. For v.s.w.r. above 1·5 : 1 these effects become insignificant and the error due to the reflection from the probe assumes increasing importance. It has been shown that the reactive component of the probe reflection causes a shift of the minimum of the standing wave in one direction and a shift of the maximum of the pattern in the opposite direction. This results in the pattern becoming lopsided and assymetrical about the minima and the maxima. The distortion is always worst at the maxima. For ratios in excess of 10 : 1 the shift in the minima becomes less and less until, when the ratio is very large, the error becomes negligible. As a result there are three quite different techniques possible for position measurement.

10.1. *Technique for ratios up to 1·5 : 1*

In this case a very accurate definition of the position of the standing wave is rarely required and it is usually sufficient to average two or more points near the axis of symmetry of the curve. At this part of the pattern the slope is steepest and therefore the discrimination on a meter is greatest. In addition, as the ratio is small, the effect of the probe reactance will be to shift the maxima in the pattern by nearly as much in one direction as it shifts the minima in the other direction. Thus the points of intersection of the curve with its axis will be near to those for the true curve. (See Figs. 11 and 12.)

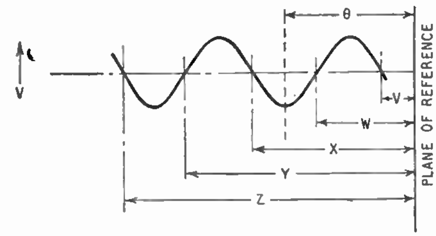


Fig. 11.—*Measurement technique for finding position of minimum from measurements near the mean of the curve.*

$$\theta = \frac{X+W}{2} = \frac{V+W+X+Y}{4}$$
$$=(Z-5\pi/4)=(Y-3\pi/4)=(X-\pi/4), \text{ etc.}$$

All dimensions in radians.

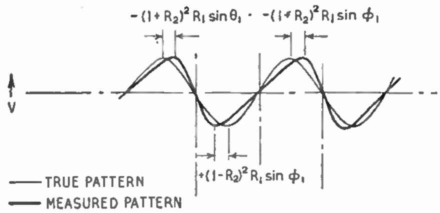


Fig. 12.—*Distortion of standing wave pattern due to probe reactance. If R_2 is small, the points of intersection with the axis will nearly coincide with those of the true curve.*

10.2. *Technique for ratios between 1·5 : 1 and 10 : 1*

When making position measurements on standing waves of moderate size, the position of the minimum in the pattern is fairly clearly defined and with care a good accuracy is possible. At the same time the field strength at the minimum is not small enough, compared with the forward wave in the guide, for the effect of the probe reactance to be ignored. The ratio is too large for the approximate correction used with the smaller ratios to be valid and it

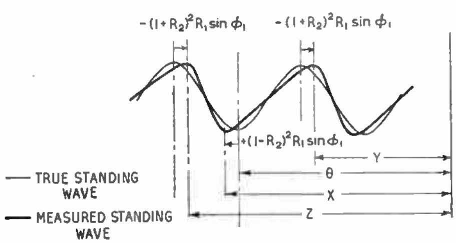


Fig. 13.—*Distortion of standing wave pattern due to probe reactance. If R_2 is between 0·2 and 0·8 it is necessary to use both the maxima and minima to find the true phase of and standing wave.*

is necessary to make a more precise correction. Fig. 13 shows a standing wave pattern with the true positions of the maxima and minima indicated by chain dotted lines. θ is the true electrical angle required between the plane of reference and a minimum, and x, y and z are measurements, converted to electrical angles, from the plane of reference to the indicated minimum and its adjacent maxima respectively. Now

$$\theta = x - (1-R_2)^2 R_1 \sin\varphi_1$$

$$\theta = \tfrac{1}{2}(y+z) + (1+R_2)^2 R_1 \sin\varphi_1 \qquad \text{(i)}$$

$$S = (1+R^2)/(1-R^2)$$

$$S^2\theta = S^2 x - (1+R_2)^2 R_1 \sin\varphi_1 \qquad \text{(ii)}$$

Adding (i) and (ii) and transposing terms

$$\theta = \frac{S^2}{1+S^2}\left[x + \frac{y+z}{2S^2}\right] \qquad \text{.....(28)}$$

The maxima, being less clearly defined than the minima, the process of averaging the two adjacent maxima is desirable as it tends to compensate for the greater errors which occur in establishing their positions. In fact, for the larger ratios, it is advisable to make an average of several measurements of y, and z, for each measurement of x.

10.3. *Technique for ratios over 10 : 1*

Providing that there is sufficient sensitivity in the detector for the minimum of the pattern to be seen on the meter, this is the easiest measurement to make as the definition is good and no corrections are required. If the indicator is not sufficiently sensitive so that the minimum is lost in the noise level, its position is found by the technique of "bracketing" that is, by making observations at a fixed level on either side of the minimum. This may be done by measurement or simply visually, by moving the travelling detector back and forth and finding the mean by feel and observation of the meter.

11. Measurement of Normalized Impedance

There are two points in each cycle of a standing wave which are easily measured and which precisely define an impedance without the need for calculation. These are the maximum and the minimum respectively. Their ratio defines the modulus and their position the phase or argument.

At the minimum the impedance is defined as:—

$$\mathbf{Z}_{min} = \frac{1}{S} e^{j0}$$

$$\text{where } S = \frac{V_{max}}{V_{min}}$$

and at the maximum as: $\mathbf{Z}_{max} = Se^{j0}$. At all other points in the pattern a complex impedance is defined from the above relationships by the equation:—

$$\mathbf{Z} = \frac{\mathbf{Z}_{min} + \tanh \gamma l}{1 + \mathbf{Z}_{min} \tanh \gamma l}$$

where l is the length of line between a minimum of the pattern and the plane of reference.

This equation is somewhat cumbersome to manipulate as would be usually necessary. As a result the by now familiar Smith Chart or Circle Diagram was devised,* see Fig. 14. This can be used as a calculator by the addition of a radial cursor, when it will quickly transform an impedance through any desired length of line. Alternatively the chart may be used as a graph on which may be plotted the variation of impedance against some other parameter (e.g. frequency). The use of the circle diagram as a calculator in this application is as follows. Knowing the impedance in the line at a minimum of the pattern, a point is defined on the Circle Diagram. Then, knowing the electrical length of line between the minimum and the desired plane of reference the diagram can be used to transform the known impedance at the minimum to the unknown impedance at the plane of reference. The electrical length is calculated from the measurement of physical distance between the minimum in the pattern and the reference plane, by dividing into it the wavelength in the line.

It is at this point that a certain amount of care becomes necessary, due to the fact that the guide wavelength in the slotted section is not the same as that in the rest of the guide. It is quite common to choose for reference the plane of the output coupling face of the standing wave meter, as this is readily accessible both in the instrument and the component being measured. The slotted section does not usually extend the full length of the instrument for reasons of physical strength, so that there is a change in guide wavelength between the probe and the reference plane. In a waveguide standing wave meter the wavelength in the slotted section may well differ from that in the rest of the guide by as much as 1 per cent. or as little as 0·1 per cent. depending upon the width of the slot. In either case it is advisable to take the effect of the slot into account. This is done by considering each section separately. The electrical length between the probe and the end of the slot is its length l_1 divided by the wavelength in the slotted section, and similarly the length l_2 of the rest of the guide divided by the wavelength in that section gives the remainder.

$$\text{Effective electrical length} = \frac{l_1}{\lambda_{g1}} + \frac{l_2}{\lambda_{g2}}$$

where $l = l_1 + l_2$ and l_1 = length of line having λ_{g1}
and l_2 = length of line having λ_{g2}

As λ_{g1} differs from λ_{g2} by quite a small amount it is clear that l_1 and l_2 do not themselves require to be known with high accuracy providing that their sum l is accurate. It is also necessary to know the frequency and the velocity ratio in the line in order to find the wavelength λ_{g2} and, as this technique requires measurements over at least one wavelength as a rule, it follows that both of these parameters must be known with an accuracy of the order of one part in 10^4 if the electrical length is required to 10^{-3} radian as may happen. It is rarely that measuring instruments in normal use in microwave laboratories have an absolute accuracy of this order and this technique is generally only used when an accuracy of about 1 per cent. is required in the measurement of impedance.

It is not uncommon to find instruments which are able to discriminate to one part in 10^4 and signal sources which are stable to this order and this fact leads to a much more accurate technique of measurement. In the first place the wavelength in the slotted section is measured by placing a short circuit on the reference plane and using the phase gear of the instrument to measure the physical distance between minima in the standing wave pattern. Secondly, the position of one of these minima is noted and used as a new reference plane. Thus whatever the actual wavelength in the line beyond the slotted section, the electrical distance between the new reference plane and the short circuit is an integral number of half wavelengths. As the pattern repeats itself every half wavelength it is necessary now only to find how many half

* P. H. Smith, "An improved transmission line calculator," *Electronics*, **17**, pp. 130-133, January 1944.

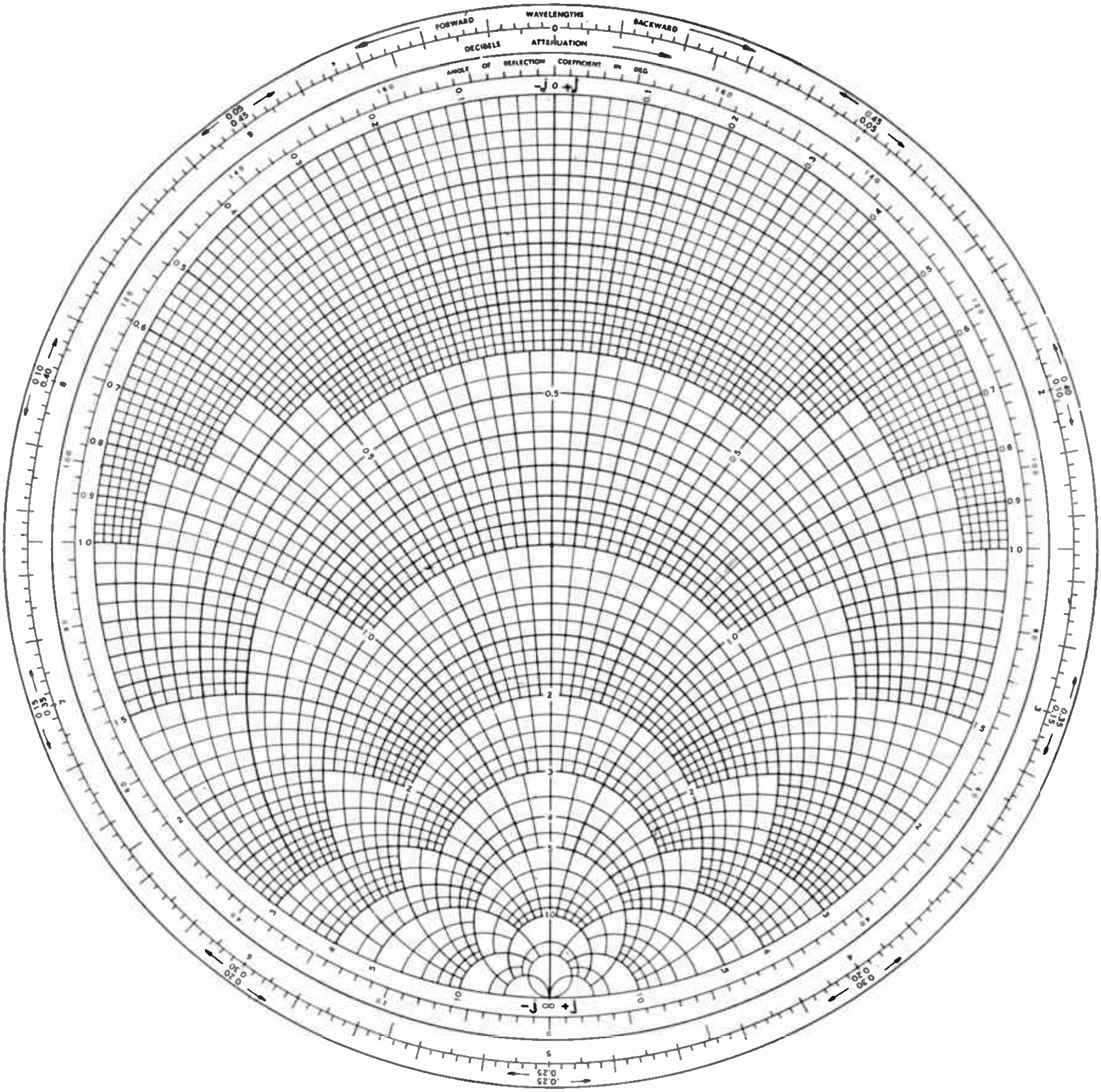


Fig. 14.—The "Smith Chart." The circles centred on the vertical axis represent constant values of R/Z_0. Those centred on the horizontal axis represent values of jX/Z_0. An impedance is defined by the intersection point of two circles. Points of constant v.s.w.r. are defined by radii about the centre of the diagram. The magnitude of the v.s.w.r. is defined by the point of intersection of these circles with the vertical axis.

wavelengths, a relatively easy matter. Having established a reference plane within the slotted section and being able to make an accurate measurement of the wavelength therein, the measurement technique is as before, i.e. the impedance at a minimum in the pattern is determined and then transformed to the new reference plane.

12. Measurement of Voltage Reflection Coefficient

One way of describing the effect of a termination upon a transmission line, as we have seen, is to define the ratio of its impedance to the characteristic impedance of the line. This method of description has certain advantages when design work is being carried out using the analogies of low frequency lumped circuits and filter theory, and is quite widely used. However, the algebra based on the use of the complex voltage reflection coefficient, v.r.c., is quite compact as has been shown in the general treatment of discontinuities in the first part of this paper, and offers powerful methods of analysis. The use of the v.r.c. has other advantages, for example, the fact that as its modulus does not change with the electrical length of line between the discontinuity and the reference plane, the use of a special graph paper and calculator becomes unnecessary and polar coordinates can be used for plotting results. The modulus, R of the v.r.c. is found from the v.s.w.r.:

$$R=\frac{S-1}{S+1}$$

and the argument, φ, from the physical distance l between the probe and the reference plane:—

$$\varphi = 2\pi\left[n-\frac{2l}{\lambda_g}\right]$$

when the probe is at a minimum. In this expression, n is an integer indicating the number of half wavelengths between the probe and the reference plane.

13. Analysis of Two Terminal Pair Networks

In the above considerations we have studied all the tests which can be made on a single ended termination, that is, the equivalent of a single terminal pair filter network. The only further information to be derived is a repetition of these tests at various frequencies. Consider now, a length of line which need not necessarily be uniform and which may contain one or more discontinuities. This has two ends and is equivalent to a two terminal pair filter network. It is clear that more information concerning its properties is required if it is to be defined completely. As before it can be treated as a filter, in which case it is defined by its image impedances and its transfer constant. Denoting the two ends of the network by the suffixes 1, and 2 and the image impedances as $\mathbf{Z}_1$ and $\mathbf{Z}_2$ respectively it is possible to show that

$$\mathbf{Z}_1 = \sqrt{\mathbf{Z}_{oc}\,.\,\mathbf{Z}_{sc}} \qquad\qquad \mathbf{Z}_2 = \sqrt{\mathbf{Z}_{oc}\,.\,\mathbf{Z}_{sc}}$$

where $\mathbf{Z}_{oc}$ and $\mathbf{Z}_{sc}$ are the impedances measured at terminals 1 with terminals 2 open circuited and short circuited respectively, and $\mathbf{Z}'_{oc}$ and $\mathbf{Z}'_{sc}$ are the impedances measured at terminals 2 with terminals 1 alternatively open and short circuited.

The image transfer constant, θ, defines the ratios of input and output voltages and currents:—

$$\frac{\mathbf{E}_2}{\mathbf{E}_1} = \sqrt{(\mathbf{Z}_2/\mathbf{Z}_1)}e^{-\theta} \qquad\qquad \frac{\mathbf{I}_2}{\mathbf{I}_1} = \sqrt{(\mathbf{Z}_1/\mathbf{Z}_2)}e^{-\theta}$$

and

$$\tanh\theta = \sqrt{\frac{\mathbf{Z}'_{sc}}{\mathbf{Z}'_{oc}}} = \sqrt{\frac{\mathbf{Z}'_{sc}}{\mathbf{Z}'_{oc}}}$$

The transfer constant has a real part which is the attenuation constant in nepers and an imaginary part which is the phase constant in radians. θ is the equivalent of γ the propagation constant of a uniform line.

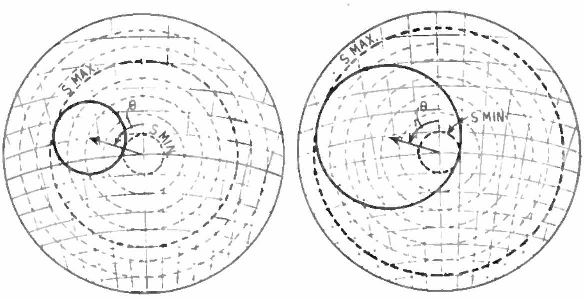
The foregoing definitions immediately point to the measurement technique to be used with a standing wave meter. The section to be analysed is connected to the instrument and terminated with a short circuit. A measurement of impedance as already described gives $\mathbf{Z}_{sc}$ at the plane of reference. The short circuit is then moved through a quarter wavelength, either by inserting a quarter wavelength of line or by moving the short circuit in the line, and the measurement is repeated. This gives $\mathbf{Z}_{oc}$ whence $\mathbf{Z}_1$ and θ can be found. If the component is intended to be worked between identical lines, it may then be reversed and the analysis repeated to find $\mathbf{Z}_2$. Bearing in mind the fact that the impedances measured with a standing wave meter are normalized to the characteristic impedance of the line, $\mathbf{Z}_1$ and $\mathbf{Z}_2$ define the impedance transformations necessary to match the network impedances to that of the line.

The above technique is convenient when some device which cannot be altered is to be matched to the line. In design, however, it frequently happens that some or all of the parameters affecting the properties of the network are controllable and require to be analysed in order to produce a device having optimum performance. In this case it may sometimes be preferable to terminate with the impedance into which it will finally have to work and to examine the effects of the controllable parameters upon the input impedance. In this kind of analysis the network is frequently nearly lossless, in which case the characteristics of the termination can profoundly modify the measurements. In modern equipments it is often necessary to design component parts having input and output impedances differing by less than one per cent. from that of the lines into which they are connected. In the measurement of such components it is therefore necessary to eliminate the effects of any mismatches which may exist in the dummy loads used as terminations. This may be done by fitting a variable matching transformer to the load and setting it up by means of the standing wave meter to be as nearly perfect as possible at each frequency. However, this is a rather laborious procedure and a more accurate and more simple one is as follows:—The load used as a termination is made to consist of a uniform length of transmission line which is dimensionally accurate; within this, a suitably shaped piece of dissipative material is arranged to closely fit the metal walls but still be free to move axially. Considering the properties of this device, it may be stated that if the impedance which it presents is measured at various positions of the dissipative load along its guide, the locus of the points will be a circle centred on the origin of the circle diagram. The radius of this circle will, of course, be the v.s.w.r. It is a principle that a circle always transforms through a linear network to another circle in this type of diagram. Hence, if the component to be measured is inserted between the standing wave meter and the termination and the procedure repeated, a new circle will result. Providing both the match of the load and the network is better than about 1·2:1, the centre of this new circle will be the impedance which would have been measured with a perfect termination to a good first order approximation. It must be stressed that the technique is only approximate but becomes more nearly exact as the centre of the circle approaches the origin. In practice it is only necessary to measure three or four points in order to define the circle but even this is somewhat laborious and a simplification is often employed. According to this method the load is slowly moved along and at the same time the travelling detector is made to follow the minimum in the standing wave pattern. The value of this will vary cyclically and its minimum value is determined. With the load in this position, the impedance is measured giving S_{max} and an angle φ. The load is then moved through a quarter wavelength and the impedance measured again, giving S_{min}, when it should be found that the position of the minimum is either unchanged or differs by exactly a quarter wavelength from the previous measurement. This corresponds to the two conditions shown in Fig. 15. Here the measured impedances are shown transformed to the plane of reference and from this it is clear that in either case the phase is found from S_{max} and in case (*a*) the standing wave ratio is:—

$$S = \frac{S_{max} + S_{min}}{2}$$

and in case (*b*)

$$S = \frac{S_{max} - S_{min}}{2}$$



(a) (b)

Fig. 15.—Circle diagram showing the separation of two small discontinuities.

Having considered both the case of the single ended circuit and that of the double ended circuit there is little more to be said. The detailed analysis of more complex circuits can normally be built up from the basic measurements which have been described combined with the application of the circuit theory outlined in the first part of this paper. The measurement techniques are the same and there is nothing new to be added from the viewpoint of impedance measurement.

It is worth while to mention that attenuation in a network has the effect of reducing the apparent magnitude of the reflection coefficient of the termination. In fact, the attenuation factor is always found in the equations, together with the terminal reflection coefficient. This leads to a method of measuring the attenuation of a system whereby the magnitude of the reflection coefficient of a short circuit termination is first measured by itself and then as seen through the section of line which is to be measured. The difference between the two reflection factors when they have been converted to decibels then gives twice the attenuation of the section. This method is very accurate in a uniform line of low loss, but a word of warning is necessary. If there are any discontinuities in the line, or between the line and the standing wave meter or short circuit termination they can lead to large and serious errors which are difficult to detect and much more difficult to correct. The application of equation (12) to this case will show the presence of terms which can not be eliminated even when the discontinuities are small. There is not even an approximate technique of measurement which will give an adequate idea of the way in which a discontinuous line will behave when the short circuits are replaced with matched or nearly matched terminations. In such a case the line must be completely analysed in terms of its image impedances and the image transfer constant and its behaviour, when operated in anything but the matched case, then has to be derived from this data, or measured under the particular conditions which apply.

14. Conclusions to Part 2

The correct analysis of a standing wave depends upon an understanding of the ways in which the measuring instrument can distort the results. Approximate corrections are valuable so long as their limitations are realised. Having analysed the standing wave the results can be converted into normalised impedances and four terminal networks can also be resolved into the equivalents of low frequency filters. The analyses can also be used for the derivation of voltage reflection coefficients.

Part 3—DESIGN AND ANALYSIS OF STANDING WAVE METERS

15. Introduction

The type of standing wave meter to be described consists of a length of rectangular section waveguide propagating the dominant (TE_{01}) mode, slotted along the centre of the broad face parallel to the axis. The electric fields within the guide are examined by means of a short aerial probe held in a carriage which is free to move in a direction parallel to the guide axis. This probe is an extension of the inner conductor of a coaxial network which incorporates an impedance matching system.

In the design of such an instrument there are three distinct sets of problems to solve once the electrical parameters are known. Firstly it is necessary to establish accurate relationships between the electrical and mechanical parameters. This is capable, in general, of quantitative analysis in exact mathematical terms. Secondly it is necessary to devise a physical structure which will be a practical manufacturing proposition and which will conform as closely as possible to the mechanical requirements. The design of such a structure, whilst making use of scientific principles and methods, is not in itself scientific—it is an art and as such is highly individualistic. No two designers will solve the problems in quite the same ways. In fact the form of their solutions should be, to a large extent, dependent upon the nature of the manufacturing organisation to be employed in the production of the instrument. Thirdly, there is a set of aesthetic problems concerned with placing of controls, degree of control, display of information, general appearance, ease and speed of operation, etc. etc. (See Fig. 16).

The form of the solutions in the second and third cases must be largely influenced by the important considerations of cost of manufacture. The whole design is a compromise in which the best possible results must be achieved without rendering the instrument uneconomic from the purchaser's point of view. The instrument which will be referred to in this paper should not, therefore, be regarded in any way as a suggested ideal solution—it is merely the author's approach under the specific set of conditions obtaining in this particular case. The skilled user of such instruments will no doubt find much to criticise.

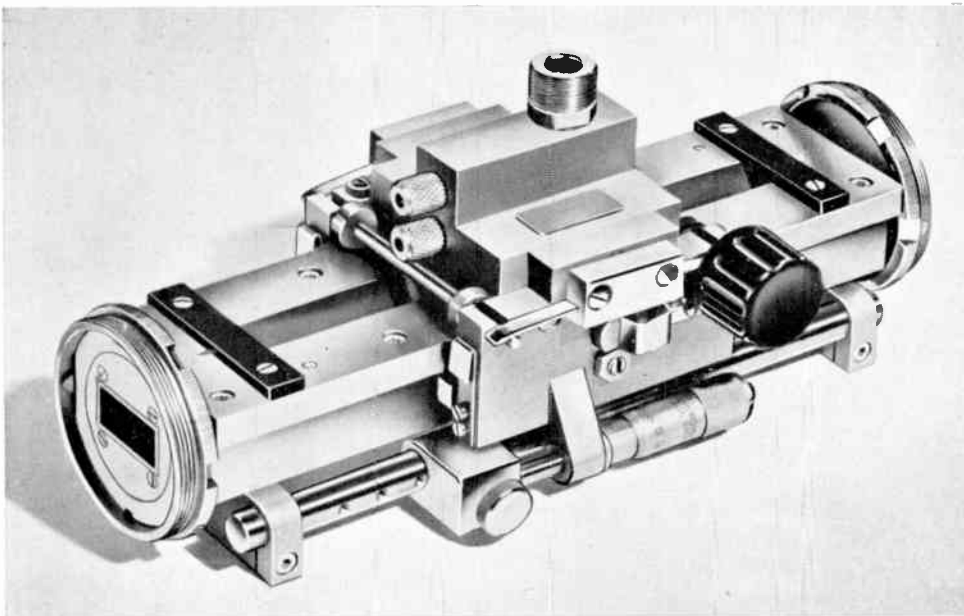


Fig. 16.—The slotted section of the standing wave ratio meter described in the paper.

16. Assessment of Mechanical Tolerances

16.1. *Characteristic impedance of guide*

16.1.1. Effect of guide dimensions

When making a measurement of impedance the result is normalised to the characteristic impedance of the slotted section. It is therefore necessary to ensure that this is close to the nominal value for the waveguide being used. The characteristic impedance of a waveguide is a function of both its dimensions and of the wavelength being propagated. It is defined by the expression:—

$$\mathbf{Z}_0 = \sqrt{\frac{\mu}{\varepsilon}} \cdot \frac{\lambda_g}{\lambda} \cdot \frac{b}{a} \qquad \text{.........(29)}$$

where

$$\frac{1}{\lambda_g^2} = \frac{1}{\lambda^2} - \frac{1}{\lambda_c^2}$$

substituting for λ_g and taking partial derivatives we have, after rearranging the terms:—

$$\frac{\delta \mathbf{Z}_0}{\mathbf{Z}_0} = -\frac{\delta a}{a} \cdot \left(\frac{\lambda_g}{\lambda}\right)^2 + \frac{\delta b}{b} \qquad (30)$$

where λ_g is the guide wavelength which would have existed in a waveguide of dimensions "a" and "b."

This expression defines the fractional change in characteristic impedance $\frac{\delta \mathbf{Z}_0}{\mathbf{Z}_0}$ which would occur as a result of small fractional changes $\frac{\delta a}{a}$ and $\frac{\delta b}{b}$ in the dimensions "a" and "b" of a rectangular waveguide.

Having decided on a permitted maximum departure of $\mathbf{Z}_0$ from the nominal value, it is then possible to allocate mechanical tolerances δa and δb for the dimensions of the tube. The term $(\frac{\lambda_g}{\lambda})^2$ normally has values between 1·5 and 3·0 within the recommended operating range of wavelengths for a given size of guide. With most British standard waveguides the aspect ratio $\frac{a}{b}$ is about 2:1 so that there is a strong temptation to allocate equal mechanical tolerances to both dimensions. That is to say, at mid-band where $(\frac{\lambda_g}{\lambda})^2$ is about the same as the aspect ratio, an error of 0·001 in. in the broad dimension will produce the same effect on $\mathbf{Z}_0$ as an error of $-0{\cdot}001$ in. in the narrow dimension.

16.1.2. Effect of Slot

It is now generally accepted that the presence of a slot modifies the $\mathbf{Z}_0$ in that section of the guide by a relationship suggested by Montgomery:*

$$\frac{\delta Z_0}{Z_0} = \frac{1}{2\pi} \cdot \left(\frac{w}{a}\right)^2 \cdot \frac{a}{b} \cdot \left(\frac{\lambda_g}{\lambda_c}\right) \quad .... \quad (31)$$

where w is the width of the slot and λ_g and λ_c are as calculated for unslotted guide. This means that unless steps are taken to prevent it, there will be a discontinuous change of $\mathbf{Z}_0$ at the end of the slot which will give rise to a reflection combining with that from the termination being examined and result in a false measurement as described in Part 1 of this paper.

There are two practical methods of overcoming this difficulty, both involving impedance matching transformations. The first consists of tapering the slot from its full width to zero width. In effect the waveguide, over the length of the tapered slot, is a section of non uniform transmission line whose impedance and propagation constant are slowly varying. By forming the wave equations for a non uniform line and setting the restriction that there must be no reflected wave we find that the length of the taper must be a half guide wavelength or multiples thereof. If the slot is to be corrected at both ends, however, this means that one full wavelength of the slot at the design centre is unusable and the instrument becomes rather unwieldy in length. For this reason an alternative method finds favour with the author.

According to this method the principle of the quarter wave matching transformer is used. Here the slot is reduced in width over the last quarter guide wavelength in such a way as to satisfy the equation:—

$$\mathbf{Z}_1 . \mathbf{Z}_2 = \mathbf{Z}_0^2$$

where $\mathbf{Z}_1$ is the characteristic impedance of the slotted section of guide.

$\mathbf{Z}_2$ is the characteristic impedance of the unslotted section of guide.

$\mathbf{Z}_0$ is the characteristic impedance of the matching section of guide.

This can be expressed in terms of reflection coefficients by considering the matching section to have reduced the total reflection to two equal components separated by a quarter wavelength. The quarter wavelength separation results in these two components mutually cancelling. (See Fig. 17.) Reference to eqn. (31) shows that the desired result is achieved by reducing the slot width by a factor of approximately $1/\sqrt{2}$ in the matching section.

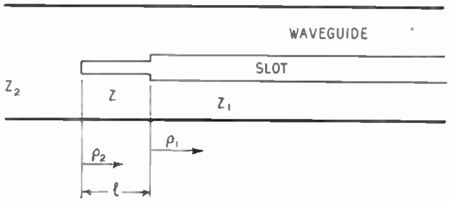


Fig. 17.—Matching the end of slot discontinuity.

$\rho_1 = \rho_2 \qquad \beta l = \pi/2$

total reflection $\simeq \rho_1 + \rho_2 e^{j2\beta l}$

16.2. *Excitation of the Slot*

Reference to the field equations shows that a narrow slot in the broad wall whose axis coincides with that of the guide does not couple to the wave within the guide. If this condition does not obtain, energy will be coupled into the slot which will radiate and also couple into the probe circuit of the travelling detector. This latter effect is serious as the coupling will vary with the distance between the travelling detector and the ends of the slot giving rise to sharp changes in output corresponding to the various resonant modes. The precise relationship which exists must depend to a very large extent upon the physical arrangement and no general treatment seems to have been made.

Some general principles have been established however. Firstly, if the slot is accurately positioned on the guide axis, only evanescent modes will be excited within it and the fields will die away exponentially towards the outer surface. The rate of decay will depend upon its width and a good general practice is to make the depth of the slot at least twice its width in order to ensure negligible radiation. Secondly, displacement of the slot parallel to the guide axis results in coupling which is proportional to

*C. G. Montgomery, "Technique of Microwave Measurements." M.I.T. Radiation Laboratory Series, Volume 11. (McGraw-Hill, New York, 1947.)

the square of the displacement. A similar result is obtained if the slot is inclined to the guide axis. Thirdly, as it cannot be expected that perfect alignment will be achieved, the effect of coupling to the slot can be reduced by arranging for a tongue on the carriage to shroud the probe until it is within the guide. This tongue acts as a shield which serves the dual function of reducing unwanted coupling between the probe and any fields within the slot and reducing the impedance of the return path from the probe circuit to the guide. The carriage and tongue may make electrical contact with the guide outer surface and the slot walls, or it may be held clear in the vicinity of the slot in which case an electrical choking system will be used. The first method finds little favour at the present time due to difficulties in ensuring adequate electrical contact as the carriage is moved along. Variations in point of contact and contact resistance give rise to fluctuations in output and sometimes quite severe radiation. As a result the choked tongue is most frequently used in high precision instruments. To obtain the low impedance required, it is usual to reduce the gaps between the tongue and the slot walls to a minimum and to maintain this spacing over a distance of a quarter wavelength in a plane at right angles to the guide axis. In effect, the tongue and the slot walls form two parallel strip transmission lines a quarter wavelength long which are of low $\mathbf{Z}_0$. If these are terminated with a high impedance then a transformation will take place such that a very low impedance is presented to the guide for transverse current flow between the tongue and guide wall, so that the gaps are electrically closed or choked. Axial current flow at the inner face of the tongue only excite evanescent modes in the slot but it is highly desirable to prevent this as magnetic leakage results which can give rise to coupling between the probe and external objects near the ends of the tongue. The author therefore makes a practice of adding further transverse quarter wave chokes, to present high impedances to any axial current flow at the ends of the tongue.

To summarize, it should be arranged that the slot width is less than twice its depth, its axis coincides with that of the guide within 0·001 a throughout its length, and that there is a tongue, occupying at least 80 per cent. of the width of the slot and about half a wavelength long which has a low impedance for transverse current flow to the guide and a high impedance for axial current flow at its ends. Under these conditions the effects of slot excitation which have been discussed above should not give rise to variations in output of more than 0·005 db.

16.3. *Effect of variation of penetration of tongue*

It is to be expected that if the carriage moves up and down with respect to the top wall of the guide there will be a variation of coupling between the probe and the field in the guide. There seems to be no theoretical analysis of this effect but the author has determined experimentally that a variation of about one part in five thousand of the b dimension results in a change of output of 0·01 db. This effect seems to be linear when the tongue is protruding into the guide but if it is withdrawn into the slot it rapidly becomes much greater. Even when the tongue protrudes into the guide, however, the variation of probe coupling seems to be rather surprisingly fierce. The effect has been measured on several different designs of choked tongue and there can be no doubt of the validity of the observations so that an acute mechanical problem results. For example, in a first grade instrument on No. 16 waveguide (0·9 in. × 0·4 in.) it is desirable to ensure that the tongue moves parallel to the upper inside surface of the waveguide within 0·0001 in. over its entire travel.

17. Mechanical Design

From the assessment of tolerances it is clear that the three primary mechanical problems to be solved are: (*a*) Maintenance of guide dimensions, (*b*) Accurate alignment of axis of slot to axis of guide, and (*c*) Ensuring that the carriage moves parallel to the upper inside wall of the guide. The author has chosen the following solutions.

The guide is made to consist of a channel section and a flat top plate which projects beyond the edges of the channel. The channel is first milled out to within a few thousandths of an inch of the desired waveguide dimensions, suitable heat treatment schedules being applied between machining operations to minimise internal stresses in the metal. In a similar way the top plate is slotted and machined to near finished size. The under surface of the top plate is then surface ground flat and checked against an optical proof plane. It may also be lapped a little until the operator is satisfied that it is flat overall within 0·0001 in. This surface now forms a reference plane. Similarly one edge is ground flat and parallel to the axis of the slot to form a reference edge. The remaining machining operations are carried out on the channel and it is then set up and surface ground on all internal faces until the width ("a" dimension) of the channel is within limits plus or minus 0·0005 in. and all faces are flat and parallel. The operator then carefully measures the actual width obtained and grinds the top of the channel in such a way to compensate in the "b" dimension for the errors in the "a" dimension. That is the error in "a" less the error in "b" has to lie within the limits $\pm$ 0·0003 in.

$$\left[[\delta a]^{+.0005''}_{-.0005''} - [\delta b]^{+.0005''}_{-.0005''} \right]^{+.0003''}_{-.0003''}$$

The reasons for this procedure are chiefly associated with economizing in surface grinder's time which is expensive, and minimizing the risk of spoiled work. For example, the mechanical measurement of the width of a channel is a tedious and time consuming process involving special measuring instruments. On the other hand the measurement of the depth of a channel is easy precise and rapid with the aid of a depth micrometer. Also, if a mistake is made in grinding the top of the channel it can be corrected by grinding a few more ten thousandths of an inch from the bottom without resetting the work, whereas a mistake in the width is irremediable.

As a result, both $\mathbf{Z}_0$ and λ_g are within 7 parts in 10^4 of nominal at the design centre. The channel and top plate are then electroplated, first with a film of nickel about 100 microinches thick and secondly with a film of rhodium about 20-30 microinches thick. As a result of the plating the waveguide attenuation is about 0·005 db per cm and, at the worst, both the $\mathbf{Z}_0$ and λ_g are within one part in a thousand of nominal.

Consideration of the wave equations will show that the junctions between the top plate and the walls of the guide lie across current flow lines of maximum intensity. On first sight this might seem to be unwise. In fact, as the surfaces are accurately flat and coated with rhodium, a very intimate and excellent electrical contact is achieved and care is taken to ensure that the top plate is pulled down evenly and strongly to the channel. As both the top plate and channel were flat to start with, they remain flat when bolted together so that the overhanging ledges formed by the projecting top plate represent an extension of the internal upper surface of the waveguide wall to the outside. This surface and the reference edge are used to locate the carriage on kinematical principles.

A great deal could be written on the application of the theory of kinematics to the design of instrument mechanisms. The subject is, however, somewhat outside the scope of the present paper. It is sufficient to say that the use of kinematical theory makes it possible to devise precise physical structures which are practical manufacturing propositions. By the same principles it is also possible to devise moving mechanisms which are free from mechanical variance and whose construction makes a minimum demand for high precision machining. It is worth while, however, to make reference to two most important points in connection with the kinematical design of moving mechanisms. First, it is axiomatic that such mechanisms are sprung together. The springs which hold the moving parts in place must therefore be sufficiently strong to cater for the largest forces which are likely to be met in operation. The author makes a point of ensuring that all such devices are capable of withstanding accelerations of at least $20g$ in any direction without either damaging the tracks or overcoming the springs. This might seem to be rather stringent but it should be borne in mind that if, due to accidental mishandling, the locating surfaces are damaged, the instrument is ruined. Similarly, if the springs are overcome the moving parts of the mechanism are no longer constrained and may acquire sufficient energy to do considerable damage both to themselves and other parts of the structure. Secondly, arising from the first point, as contact between the moving parts is limited to precisely defined and extremely small areas, the surface loading at the points of contact becomes very large. This means that the metallurgical properties of the materials used in these locations require to be studied in considerable detail and careful calculations should be made to ensure that the deformation at the points of contact is entirely elastic.

In this instrument the carriage is located by means of rollers. Three preselected ball races less than 50 microinches eccentric, disposed in the form of an equilateral triangle on the reference surface, and two other races running on the reference edge, provide five constraints. These are held to their tracks by five independently sprung rollers, two of which, fitted with rubber tyres, provide a coarse drive for the carriage. The one degree of freedom remaining to the carriage is motion parallel to the guide axis. This may be controlled by a knob on the coarse drive shaft, or alternatively by means of a micrometer. In this instrument it was considered desirable to be able to set and indicate the position of the carriage with an accuracy of 0·001 radian. This has been achieved as follows. A hard steel bar, having a number of jig bored conical holes spaced by exactly 1cm. apart, is attached to the body running parallel to the guide axis and below the carriage. Sliding on the bar is a stop which can be located on any of the holes with an accuracy of better than 0·001 cm. A 1-cm micrometer attached to the carriage is sprung against a ball anvil on the stop. By releasing the locating mechanism, the stop is free to slide along the bar with the carriage as it is traversed by the coarse drive. When an accurate phase measurement is required the stop is located at an appropriate hole and the micrometer is used as a fine drive to the carriage and an accurate interpolating scale. The scale bar is adjustable in position and is set up during calibration to indicate the electrical distance between the probe and the terminal coupling face.

There are many other ways of achieving the same results but this method appealed to the author, as the only precision work involved was the jig boring of seven small conical holes in a length of silver steel ground stock.

18. Probe Circuit

So far we have considered only the probe by itself. There exists the problem of arranging a suitable transformation between the probe and some external indicator. This may be a meter or amplifier fed via a crystal detector, a bridge fed via a bolometer or thermistor, or a superheterodyne receiver fed via a flexible r.f. cable. In general the impedance of the probe relative to that of the coaxial system will be very high and the matching transformation will be of high Q and hence very frequency sensitive. It is also clearly desirable that the matching transformer should be as close to the probe as possible. The most common arrangement consists of a single variable reactance situated at a chosen point in the line between probe and output. This arrangement will always show a maximum output at one setting but can only give optimum matching at one frequency. As a result it is not only high Q but also narrow band in operation. The considerations outlined in Part 1 show that this leads to an undesirable situation because the reactive component of the probe reflection coefficient in the guide becomes large and cannot be eliminated. The use of two variable reactances spaced apart in the coaxial line by about $\lambda/8$ gives a very much more versatile transforming network. It has the disadvantage, however, that the real and imaginary parts of the impedances to be matched together are not independent of either control, hence progressive adjustments are required to achieve the optimum condition.

The ideal arrangement is one which provides adjustment of the real parts of the impedances with one control and the imaginary parts with a second control. There are several excellent circuits which can closely approach this ideal but in any case two controls are essential for adequate matching over a wide frequency range. For reasons of economy the author chose the two stub reactive matching unit rather than one of the more elegant circuits as it has no serious drawbacks and can be tuned up in about 15 seconds.

19. Methods of Test

As the slotted section standing wave meter is the final impedance reference in the microwave laboratory against which all other instruments and components are judged, the instrument must be made to test itself as far as possible. In such respects as require some external standard, the device must be such that its properties can be completely defined in terms which can be measured separately. The three parameters which require to be known are respectively, frequency, attenuation, and impedance. Both frequency and attenuation can be referred to physical standards, and the characteristic impedance of a waveguide can, as we have shown earlier, be defined in terms of its linear dimensions. A discussion on the standards of frequency and attenuation is outside the scope of this paper and it will be assumed that reliable means are available for making these measurements. By making use of the information in the second part and in this part of the paper it is possible to devise a standard of characteristic impedance of the waveguide in use. This is done as follows:

A standard waveguide is machined to the closest possible limits of accuracy. This is best done by milling (possibly in a jig boring machine) two channels which, when put together, form a waveguide split down the centre of its broad faces. Although laborious, it is possible in this way to define a waveguide whose $\mathbf{Z}_0$ is known to better than one part in five thousand. It is also possible by conventional techniques to design a dissipative wedge which when inserted in the guide will have a voltage reflection coefficient of the order of one part in a thousand. This wedge should be accurately located in the standard guide but be free to be moved along axially within it by known amounts. As the guide is accurate, the reflection from the wedge is therefore determined in both phase and amplitude by the wedge itself.

If this standard termination is now fixed to the standing wave meter under test and the output of the instrument measured at intervals over the full movement of the carriage, the resulting pattern will contain the following components:

(*a*) a standing wave due to the reflection from the dissipative wedge, combined with any reflections within the instrument and at its junction with the standard termination.

(*b*) an exponential variation of output due to the attenuation of the slotted section of guide.

(*c*) a variation of output due to systematic changes in probe coupling. That is to say, repeatable changes which are a function of the position of the probe along the guide. This may be due to inaccuracies in the guide and the locating surfaces of the moving mechanism. As rolling contacts are generally used in the location of the moving parts there may also be detectable sinusoidal variation due to eccentricity in the rollers.

(*d*) random variations of output which may be due to instability of the r.f. oscillator, crystals, amplifiers etc., or to bad contacts, mechanical variance in the mechanism, or to radiation effects, slot resonances and to other causes.

The method of treatment to be described will determine (*a*) and eliminate from it the effect of the reflection from the dissipative wedge; (*b*) and (*c*) will also be found and (*d*) will be qualitatively analysed.

In this analysis the waveguide circuit of Fig. 8 is suitable and the measurement technique is that described in section 2 of Part 2 of this paper. The author finds from experience that it is more convenient to operate the bridge in the unbalanced condition with a calibration of the meter sensitivity by means of the precision r.f. attenuator. It will be found, with conventional apparatus of good quality that the sensitivity of the bridge can be increased to the point where the meter will just resolve 0·001 db without short term random variations becoming too obtrusive. To achieve this result, some care is needed to ensure that good electrical contacts are made throughout the circuit, that the oscillator stability is at its best and that all crystals and the oscillator valve are well shielded from draughts and other causes of random variation of temperature. The power level at the two crystals should also be as high as possible and should not be less than about 100 microwatts.

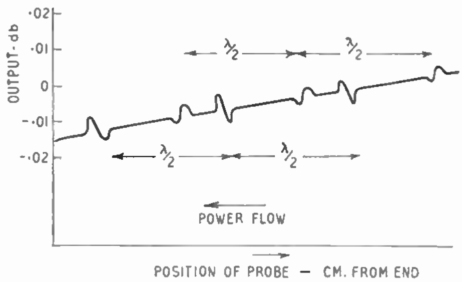


Fig. 18.—A typical display of slot resonances. Two sets of resonances are observed, between the tongue and the two ends of the slot. The standing wave and variations of coupling have been omitted for clarity, but the attenuation slope has been left in.

The first step is a general examination of the instrument. The travelling detector is slowly traversed from end to end of its travel and the meter is observed. It should be found that the output varies cyclically with no sharp changes in excess of about 0·005 db for a first grade instrument. If such changes are noted it may be found that they are clearly repeatable and occur at regular intervals. This will probably be due to slot resonances and it can be checked by measuring the distance between them, when it should be found that they occur at intervals of $\lambda/2$. (See Fig. 18.) If the changes cannot be repeated on resetting or if they do not occur at half wavelength intervals they may be attributed to mechanical variance. Once the operator is satisfied that the system is stable and that random changes in the instrument are negligible it is safe to proceed to the detailed analysis.

The method suggested is to measure the output from the travelling detector at intervals of displacement of, for instance, exactly 1/24th wavelength from one end of its travel to the other. The process should then be repeated from the end back to the beginning. Comparison of the results at each position will show whether the test is likely to be satisfactory. For example, a steadily increasing deviation would indicate oscillator drift, a random variation in excess of that observed with the travelling detector stationary would indicate mechanical or electrical variance, etc. If this is satisfactory, the dissipative wedge is moved through exactly a quarter guide wavelength and the two measurement runs are repeated. The four sets of figures are then averaged for each position of the travelling detector and the results plotted.

The resulting curve is now analysed to extract a sinusoidal function with a period equal to a half guide wavelength. This may, with a little practice, be done quite rapidly and with good accuracy by inspection and a little trial and error. If this should prove difficult, a precise treatment may be made by the method of Fourier Analysis. (Hence the intervals of 1/24th wavelength). It is worth while to note that for variations of the order of a few tenths a db the logarithmic scale is linearly proportional to both volts and power and that the standing wave is sinusoidal with sufficient accuracy for this purpose.

The measured curve has now been analysed into two component parts, and we will now consider the sinusoidal function. This represents the standing wave which the instrument measures when it is terminated with a perfectly matched standard waveguide, see Fig. 19. By moving the dissipative wedge through a quarter wavelength for the second runs and averaging the results, its reflection has been eliminated from the measurement. At the same time the two pairs of measurements afford comparison checks on the reliability of the results and the process of averaging gives a statistical improvement of a factor of two times in the accuracy of each point. To summarise, this standing wave represents the sum of the end of slot discontinuity, any other discontinuities in the instrument, failure of the terminal coupling to mate properly with that on the standard termination and the discontinuity between the $\mathbf{Z}_0$ of the instrument and that of the standard termination.

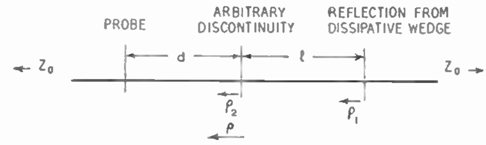


Fig 18.—Separation of two small discontinuities.

Let ρ_1 *and* $\rho_2 << 1$

from eqn. 12, *eliminating second order terms:—*

$$\rho = \rho_2 + \rho_1 e^{2\gamma l}.$$

Let the voltage at the probe $V_p = 1 + Pe^{2\gamma d}$,

Then, changing l *by* $\lambda_g/4$

$$\rho' = \rho_2 + \rho_1 e^{2\gamma l + \pi} = \rho_2 - \rho_1 e^{2\gamma l}.$$

in the first case $V_p = 1 + \rho_2 e^{2\gamma d} + \rho_1 e^{2\gamma(d+l)}$

in the second case $V_p' = 1 + \rho_2 e^{2\gamma d} - \rho_1 e^{\gamma 2(l+d)}$

and the average is $1 + \rho_2 e^{2\gamma d}$.

The residual variation can now be analysed further. By replacing the standard termination with a fixed short circuit termination, the large standing wave ratio can be analysed by the method suggested in Part 2. The travelling detector must, of course, cover at least a half guide wavelength, so two minima can be obtained and hence two v.s.w.r's. From these, two values of reflection factor can be found and their ratio gives the voltage attenuation of twice the length of line between the two minima. The attenuation rate so derived can now be extracted from the residual variation graph to give the actual variation of coupling between the travelling detector and the guide. At the same time, if the frequency is known accurately, a measurement of the distance between the minima will give the guide wavelength in the slotted section which can be compared with that calculated for an unslotted guide of nominal dimensions.

As a result of these tests it is thus possible to derive the attenuation constant of the guide, the phase constant of the slotted section, the variation of coupling between the travelling detector and the guide, the residual discontinuities within the instrument, and any random effects can also be shown up. The instrument has effectively been used to measure itself by the analysis of its results and with the aid of a standard guide which can be defined entirely in terms of its linear dimensions. Other tests, such as the measurement and analysis of the probe reflection coefficient, have already been described in the earlier parts of the paper.

Use can be made of the information, should the need arise, to enable impedances to be defined in relation to the standard termination with an absolute accuracy of better than one part in five thousand when operating an instrument of first grade performance. This points out an important distinction. For normal measurements the standing wave meter is the reference standard, but this method of analysis converts the instrument to a substandard which is calibrated against a standard waveguide. The residual v.s.w.r. obtained from the analysis can be converted to a reflection coefficient and this information can be put into a measurement as a correction for the inaccuracies of the instrument, thus making the standard waveguide the reference and the standing wave meter a comparator. It is obvious that such a precise measurement would be extremely laborious as it would be necessary to go through the whole procedure outlined above, twice at each frequency, once for the standard and once for the termination to be measured. On the other hand, it may well be that in research, certain problems can be more readily studied by impedance measurement when this degree of accuracy is achievable, than by other means.

20. Conclusions

There is nothing particularly new here but there is a representative selection of measurement techniques and some practical hints which the author has found useful. The microwave field is expanding rapidly and there are many newcomers both in the laboratory and in the factory test room who are having to acquaint themselves with these instruments. It is admitted that learning the hard way produces the most lasting results and it is certain that the technique of microwave measurement can never be acquired from a book. At the same time mistakes in measurements can lead to much frustration in the laboratory and may be very expensive in the factory. It is hoped that this paper may help to ease the pains of acquiring skill in the art without spoiling the fun.

A SHUTTER TESTER USING A PHOTO-ELECTRIC INTEGRATOR *

by

R. J. Hercock, B.Sc.† and D. M. Neale, B.Sc.‡

SUMMARY

The requirements for testing between-lens camera shutters are discussed and it is concluded that an ideal instrument should combine a photoelectric integrator with a single-sweep oscilloscope, the two units operating simultaneously. The instrument described uses a valve voltmeter to measure the voltage across the integrating capacitor. Methods of calibrating the circuit are given. The cathode-ray oscilloscope employs a phantastron time base; time marking facilities are provided. The integrator measures the effective exposure time and the oscilloscope shows the total open time of the shutter besides indicating faults such as sticking or bouncing of the shutter blades.

1. Introduction

Because of the limited exposure latitude of reversal materials, in particular colour film, in recent years greater emphasis has been placed on the accuracy of camera shutters. The effective time during which a shutter is open should be within a few per cent of the nominal value and its operation should be consistent.

The shutter on a camera is used for two purposes—to arrest motion and to control the exposure. In the majority of cases it is better to use as short a time of exposure as possible, as this not only arrests subject motion but helps to minimize effects of camera vibration. However desirable this may be, it is often in conflict with the necessity of allowing sufficient light to pass through the lens to produce a satisfactory negative.

The function of a shutter as an exposure determining mechanism probably calls for a much greater accuracy of timing than would be necessary if its ability to arrest motion were the only consideration. The ideal shutter would open to its full extent instantaneously, remain open for the specified time and then close equally rapidly (Fig. 1*a*). It will be seen from this figure that during the whole time that the shutter is open, the transmission is 100 per cent and there can be no ambiguity in defining the open time.

Practical between-lens shutters depart, more or less, from this ideal in that they take a finite time to open and to close. Hence the transmission of the shutter at the beginning and end of the exposure is reduced and a plot of the energy passing through the lens has a form similar to that shown in Fig. 1*b*. Peculiarities of the blade shape, and the acceleration and deceleration of the blades, may produce asymmetry and curvature as shown in Fig. 1*c*.

For a shutter with a characteristic as shown in Fig. 1*b* or Fig. 1*c*, it is obviously more difficult to assign a time value to a particular setting.

The *total open time*, t_o, may be defined as the time which elapses between the shutter just opening and just closing. However, the quantity of light transmitted during the time t_o is proportional to the area under the curve and this will, as can be seen, vary with the slope of the rise and fall of the curve.

The *efficiency*, η of a shutter can be defined as the actual energy transmitted divided by the energy which would be transmitted if an ideal shutter were open for the total open time t_o of the actual shutter.

The *effective exposure time*, t_e, can be defined as the time for which an ideal shutter would have to be open to pass the same amount of energy as the actual shutter. Expressed otherwise $t_e = \eta t_o$. When $\eta = 1$, then the effective exposure time equals the total open time.

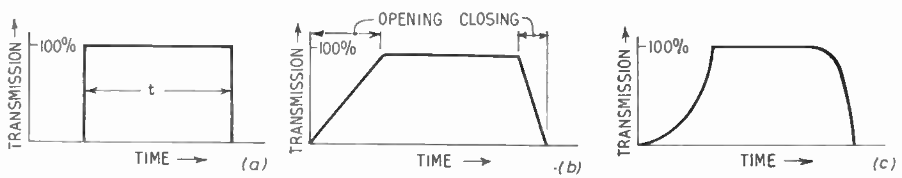
As the values of the terms defined above depend on the energy transmitted, they will obviously be affected by the lens components,

* Manuscript first received 14th March, 1955, and in final form on 13th August, 1955. (Paper No. 331.)

† X-ray Laboratory, Ilford Limited, Watford, Hertfordshire.

‡ Physics Research Laboratory, Ilford Limited, Brentwood, Middlesex.

U.D.C. No. 621.317.75 + 621.383:771.36.



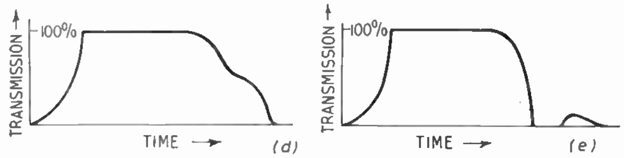


Fig. 1.—Shutter characteristics.

(a) Ideal shutter.
(b) Practical shutter.
(c) Showing effect of acceleration of shutter blades during opening and closing.
(d) Effect of sticking blades.
(e) Effect of shutter bounce.

diaphragm settings, etc. Hence these factors should be specified and care taken to see that they are the same as those used in practice.

It is evident that, although oscillographic methods of shutter testing can be used to determine the total open time for the shutter, they cannot readily indicate the effective exposure time. From the photographic point of view, it is the area under the energy-time curve which is of importance. This factor can be measured by using a photoelectric integrator,[1] which will give a meter reading proportional to the area under the energy-time curve, and hence, by suitable calibration, a value in terms of the effective exposure time.

Irregularities in the shutter performance will, of course, produce variations in the integrator reading. Nevertheless sticking (Fig. 1*d*) or bouncing (Fig. 1*e*) of the shutter blades may occur without producing observable variations in the effective exposure time. Such faults must be detected, however, because not only do they reduce the shutter efficiency, but they may become better or worse with use and so change the shutter timing. An oscilloscope will, of course, indicate such defects immediately.[2]

2. Requirements

2.1 *Functional*

For the rapid testing of camera shutters before and after assembly into camera bodies, an instrument is required which will provide the maximum information with the minimum effort. It seems, therefore, that an ideal shutter tester for this purpose combines a photoelectric integrator with a single sweep oscilloscope. Both parts are operative simultaneously so that when faulty performance is observed on the oscilloscope the corresponding integrator readings may be noted.

For production testing, the most important measurement is that of effective exposure time. The integrator gives a direct indication of this quantity and the oscilloscope is used primarily as a qualitative check on shutter performance.

If some form of time scale is provided on the oscilloscope, this display may also be used for quantitative measurements of total open time and shutter efficiency. Provision for doing this may also be desirable as a cross-check on the integrator performance.

To allow for the effects of gravity on the moving parts of the shutter, the camera should be tested with the shutter in a substantially vertical plane. A precisely vertical arrangement is, however, a special case and therefore it is preferable to incline the camera at a few degrees to the vertical. Provision is required also for rotation about the lens axis so that the camera may be tested in both the " upright " and " horizontal " positions. The camera should face the operator so that the lens and shutter scales may be clearly visible.

The various controls may be operated several thousand times a day and must therefore be positioned on a basis of time and motion study rather than electrical or mechanical convenience. It is an obvious step to gang together the integrator range switch and the oscilloscope time-base speed selector. In the present instrument, the ranges were chosen with a particular camera (the Ilford " Advocate ") in mind and accordingly were as follows :—1/25, 1/50, 1/100, 1/200 and 1/500 sec.

2.2. *Reliability*

Failure of laboratory test equipment is seldom more than irksome and there is usually someone at hand who can get the apparatus working again within a short while. On the other hand, failure of test gear used in an assembly line can be an expensive matter. The more valuable the instrument while active, the greater its expense while inactive. Thus even minor faults become serious in the absence of a skilled serviceman.

With these considerations in mind, the present instrument was designed to give a high degree of reliability. Remembering that valve failures account for the majority of breakdowns in electronic equipment, the number of valves employed has been restricted so far as consistent with good performance.

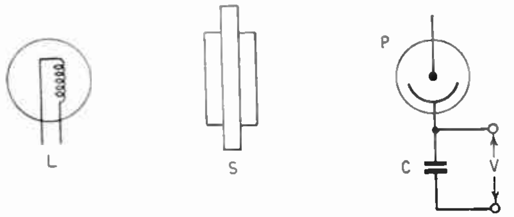


Fig. 2.—Basis of simple photoelectric integrator.

3. The Integrator

3.1 *Principle of Operation*

In principle, the photoelectric integrator is very simple. Light from a lamp, L, (Fig. 2) passes through the open shutter to fall on the photocell, P. A vacuum cell is used so that the photo-current is strictly proportional to the light intensity. With an anode voltage above about 20 V, the vacuum cell has a constant-current characteristic, and so the rate of charging of the integrating capacitor, C, is substantially unaffected by a reduction in anode voltage on the photocell. If the capacitor is initially discharged, it follows that the voltage, v, appearing across the capacitor after an operation of the shutter will be proportional to the energy-time integral of the shutter.

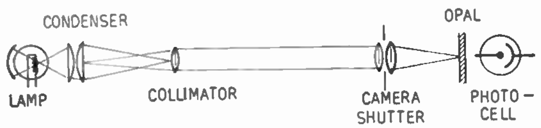


Fig. 3.—Optical system used for testing between-lens shutters.

The capacitor voltage may be read with a valve voltmeter. Provided the valve voltmeter grid current, the capacitor leakage and the photocell dark current are all extremely low, the effective shutter timing will remain displayed on the valve voltmeter until the circuit is reset by discharging the capacitor once more.

3.2. *Optical System*

In order that the characteristic of the shutter shall not be falsified, it is important to test it under conditions simulating as nearly as possible typical operating conditions. This requires that the lens and shutter shall be uniformly illuminated by a substantially parallel light beam. Light passing through the shutter must subsequently be diffused so that results are not affected by variations in sensitivity from one part to another of the photocathode. It is advisable also to preserve the normal direction of light through the lens and shutter assembly. Reflections between the shutter blades and the glass-to-air surfaces of the lens may introduce small errors if the direction of the light is reversed.

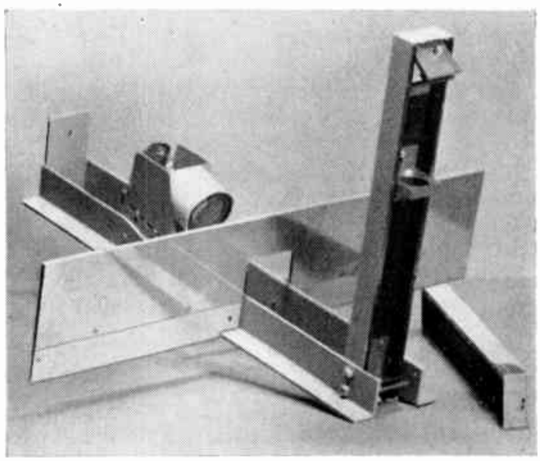


Fig. 4.—The folded optical system removed from the shutter tester.

The optical system shown in Fig. 3 is folded by the addition of two plane mirrors so that the lamp can be accommodated inside the body of the instrument. In this way, maximum accessibility is retained and the hot lamphouse is removed from the vicinity of the operator's hands. For cleaning and adjustment, the

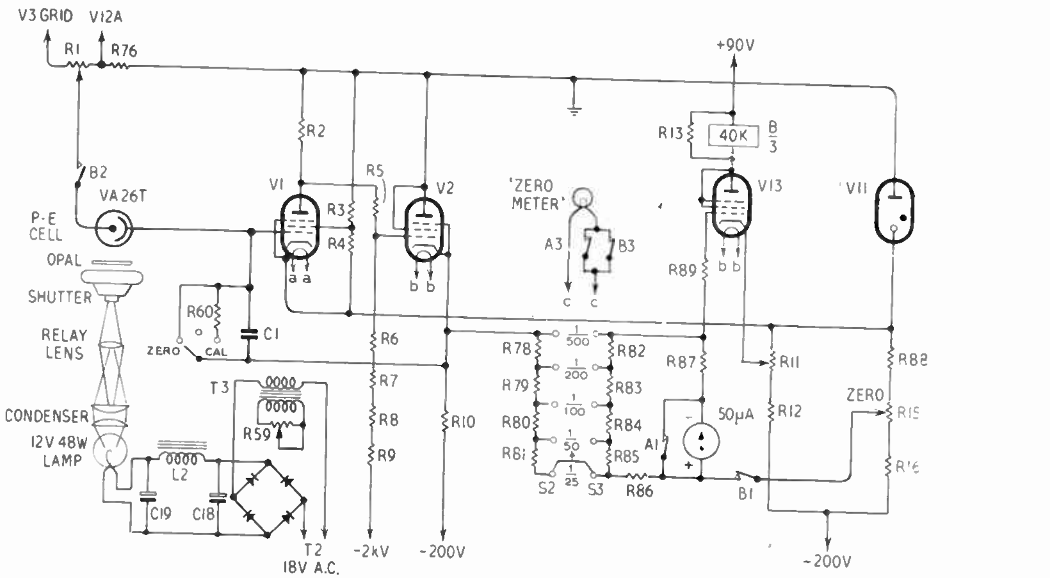


Fig. 5.—Photoelectric integrator circuit used to measure effective exposure time as a percentage of nominal value.

whole optical system may readily be removed as shown in Fig. 4.

When a camera is being tested, the camera lens forms on the opal an image of the condenser lens next to the lamp. When a shutter alone is being tested, a shadow of the shutter is thrown on the opal by the collimated light beam. In either case the opal is of ample size to collect all the light transmitted by the shutter, also the photocell is spaced sufficiently from the opal to ensure substantially uniform sensitivity to all active parts of the opal.

3.3. *The Photocell*

The photocell must be chosen to give the smallest possible dark current and the largest possible photocurrent when illuminated. The rate of drift of the integrator will depend upon the ratio of these two currents. For example, if the dark current is 0·1 per cent of the photocurrent, then in 1 second the dark current will produce the same integrator deflection as an exposure of 1 millisecond.

Leakage across the photocell has an effect similar to that of the dark current. In the photocell employed, the anode and cathode connections have been brought out at opposite ends of the envelope to minimize such leakage effects.

In order to obtain a high photocurrent while the shutter is open, a photocell is used having a photocathode which is highly sensitive to tungsten light. A high intensity of shutter illumination is required also because much of the transmitted light is lost by scattering at the opal.

3.4. *Valve Voltmeter*

To provide accurate indications, free from drift, the valve voltmeter must have the highest possible input impedance. The integrating capacitor should have a low-leakage dielectric free from after-effects.[3] Valve leakage and grid current are minimized by using an electrometer valve operating under optimum conditions. It is desirable also that the relationship between capacitor voltage and meter reading shall be linear and that the sensitivity shall be independent of supply voltage variations and valve ageing.

These requirements are satisfied by the two-valve cathode follower circuit shown in Fig. 5. The electrometer pentode, V1, operates as a d.c. voltage amplifier to drive the cathode follower, V2. The negative electrode of the integrating capacitor, C1, is connected directly to the output, so that the circuit operates with a high degree of negative feedback.[4] As the photocell charges C1, therefore, the fall in

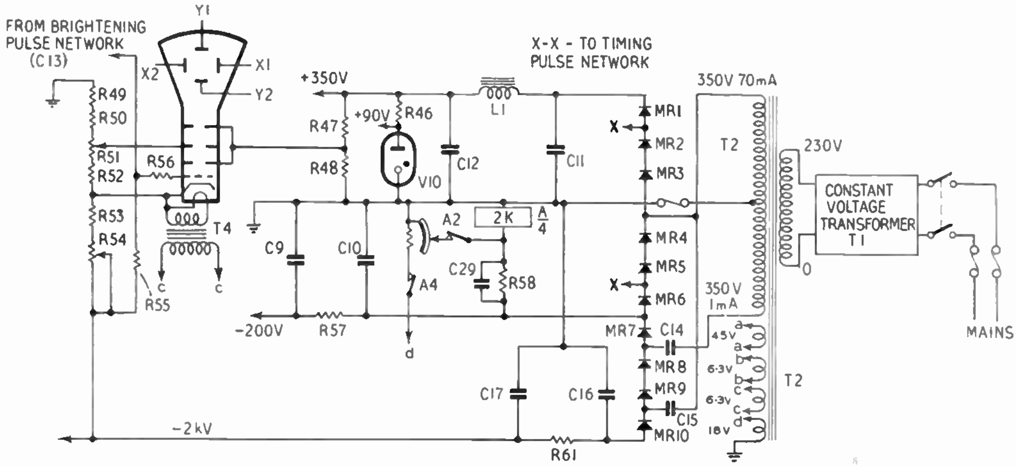


Fig. 6.—Power supply unit for shutter tester. The relay A/4 protects the integrator meter when the instrument is switched on or off.

KEY TO COMPONENTS SHOWN IN FIGS. 5, 6 and 7

R1	100 kΩ	R25	100 kΩ	R49	2·2 MΩ	R73	100 kΩ	C4	0·1 μF	C28	300 pF
R2	220 kΩ	R26	100 kΩ	R50	2·2 MΩ	R74	1 MΩ	C5	0·05 μF	C29	2 μF
R3	22 kΩ	R27	100 kΩ	R51	200 kΩ	R75	1 MΩ	C6	0·02 μF	C30	0·01 μF
R4	22 kΩ	R28	33 kΩ	R52	1 MΩ	R76	10 kΩ	C7	0·01 μF	C31	0·001 μF
R5	1·5 MΩ	R29	22 kΩ	R53	150 kΩ	R77	100 kΩ	C8	0·25 μF	C32	0·001 μF
R6	5 MΩ	R30	100 kΩ	R54	10 kΩ	R78	3·6 kΩ	C9	8 μF		
R7	5 MΩ	R31	250 kΩ	R55	1 MΩ	R79	2·8 kΩ	C10	8 μF		
R8	5 MΩ	R32	1 MΩ	R56	220 kΩ	R80	1·7 kΩ	C11	8 μF	V1	ME1400
R9	5 MΩ	R33	330 Ω	R57	10 kΩ	R81	925 Ω	C12	8 μF	V2	EF40
R10	33 kΩ	R34	15 kΩ	R58	50 kΩ	R82	12 kΩ	C13	0·1 μF	V3	EF40
R11	500 Ω	R35	10 kΩ	R59	10 Ω	R83	4 kΩ	C14	1 μF	V4	EF40
R12	22 kΩ	R36	470 kΩ	R60	800 kΩ	R84	2 kΩ	C15	1 μF	V5	EF40
R13	68 kΩ	R37	470 kΩ	R61	220 kΩ	R85	1 kΩ	C16	0·5 μF	V6	EB34
R14	22 kΩ	R38	1 MΩ	R62	470 kΩ	R86	1 kΩ	C17	0·5 μF	V7	VR116
R15	100 Ω	R39	12 kΩ	R63	1 MΩ	R87	18·5 kΩ	C18	1,000 μF	V8	EB34
R16	22 kΩ	R40	2·2 MΩ	R64	470 kΩ	R88	100 Ω	C19	1,000 μF	V9	EF40
R17	10 kΩ	R41	2·2 MΩ	R65	1 MΩ	R89	1 MΩ	C20	0·25 μF	V10	7475
R18	470 kΩ	R42	330 kΩ	R66	33 kΩ	R90	10 kΩ	C21	50 pF	V11	7475
R19	22 kΩ	R43	2·2 MΩ	R67	47 kΩ			C22	8 μF	V12	EB34
R20	56 kΩ	R44	5 MΩ	R68	47 kΩ			C23	0·25 μF	V13	EF40
R21	100 kΩ	R45	5 MΩ	R69	33 kΩ			C24	0·25 μF		
R22	100 kΩ	R46	33 kΩ	R70	1·2 MΩ	C1	0·05 μF	C25	0·001 μF		
R23	1 MΩ	R47	100 kΩ	R71	1 MΩ	C2	0·003 μF	C26	0·001 μF	MR1- 6	RM1
R24	1 MΩ	R48	100 kΩ	R72	100 kΩ	C3	0·25 μF	C27	300 pF	MR7-10	K2/25

cathode potential of V2 almost eliminates the rise in grid potential of V1. This has the advantage of keeping almost constant the voltage applied to the photocell, even though the integrating capacitor may be charged to over 25 volts. Even with exposures of 1 or 2 milliseconds, the capacitor is charged to several hundred millivolts and so the exposurė can readily be measured to within 1 or 2 per cent.

To make best use of this circuit, it was necessary to display the meter readings as a percentage of the nominal exposure, the meter being calibrated from 0–200 per cent. This meant changing the integrator sensitivity over a total range of 20 : 1. It would have been possible to have changed the value of the integrator capacitor, C1, but this would have increased the current leakage and so the drift. So far as possible, it was considered preferable to perform all switching at the low impedance output. Accordingly the meter was provided with a universal shunt which was made to present a constant impedance by the addition of the resistors R78-R81.

Some means of protecting the microammeter was required as excessive exposures could otherwise result in serious overloading. V13 is provided for this purpose. When the potential across the meter and R87 becomes excessive, V13 is cut off, the relay B/3 releases and opens the contacts B1 to break the meter circuit. Simultaneously the contacts B2 break the photocell circuit so that the photocathode shall not be damaged by excessive emission. The protection circuit is reset automatically when the integrating capacitor is discharged through the switch marked "zero". This is a telephone key switch modified by the substitution of polythene for the standard insulation.

A second circuit is provided to protect the meter from transients when the instrument is switched on or off. A thermal delay switch initially short-circuits the relay A/4 (Fig. 6), but when the thermal switch opens, A/4 is energized through R58, the contacts A2 open and so A/4 remains energized even when the thermal switch cools due to the opening of the contacts A4. When the instrument is switched off, the collapse of the negative h.t. voltage is transmitted through C29 to make A4 release before V11 extinguishes. The contacts A1 then close to short-circuit the microammeter once more and protect it from the severe overloads which would otherwise occur.

When either protection circuit is in operation, contacts A3 or B3 cause a warming lamp to be illuminated. This lamp normally extinguishes automatically about two minutes after switching on. Thereafter it lights again only if an overload on the integrator causes B/3 to release as described already.

3.5. *Calibration*

The voltage, v, across C1 after an operation of the shutter is given by

$$v = \frac{1}{C}\int_0^t i \,.\, dt \qquad \ldots \quad \ldots \quad (1)$$

where C is the capacitance of C1 in μF and i is the photocurrent which, it will be remembered, is proportional to light intensity.

With an ideal shutter having the performance indicated by Fig. 1*a*, the photocurrent has a constant value, i_0, between the limits 0 and t.

$$\text{Thus } v = \frac{1}{C} \,.\, i_0 t \qquad \ldots \quad \ldots \quad \ldots \quad (2)$$

where t = open period of the shutter.

Now suppose the integrating capacitor is replaced by a resistor of such a value, R, that while the shutter is open,

$$v = i_0 R$$

Then, substituting in (2),

$$t = CR \qquad \ldots \quad \ldots \quad \ldots \quad (3)$$

where t is in seconds, C in microfarads, R in megohms.

In the integrator, the values of C1 and R60 are chosen so that $CR = 1/25$ sec. With the range selector set to 1/25 sec. and R60 in circuit, therefore, the shutter is opened and the lamp intensity is adjusted to bring the meter deflection to a reading of 100 per cent. The shutter is then closed and R60 is removed from circuit. Subsequent operations of the shutter will produce meter deflections indicating the exposures as a percentage of 1/25 sec. No further calibration procedure is required when the selector switch is set to another nominal exposure.

3.6. *Lamp Supply*

A low-voltage, heavy current lamp was used to ensure a high optical efficiency. To minimize errors due to ripple in the light intensity, it was necessary to operate it either on d.c. or on a.c. of a frequency high enough to be smoothed by the thermal inertia of the lamp filament. A valve oscillator to supply the lamp at a high audio frequency would require very careful screening in order to prevent interaction with the oscilloscope display. A stabilized oscillator capable of supplying a 48-W lamp would add considerably to the h.t. drain and by using additional valves would add to the potential causes of breakdown. It was considered more satisfactory to supply the lamp with d.c., rectified and smoothed from the a.c. mains supply. The lamp supply is stabilized by the 150-W constant voltage transformer through which the mains supply to the whole instrument is drawn.

Smooth control of lamp brilliance on a low-voltage, heavy current d.c. supply can present difficulties. A low resistance rheostat employs usually relatively few turns of heavy gauge wire or strip and consequently the series resistance can be adjusted only in rather coarse increments. Variations in contact resistance can also cause trouble. These difficulties were

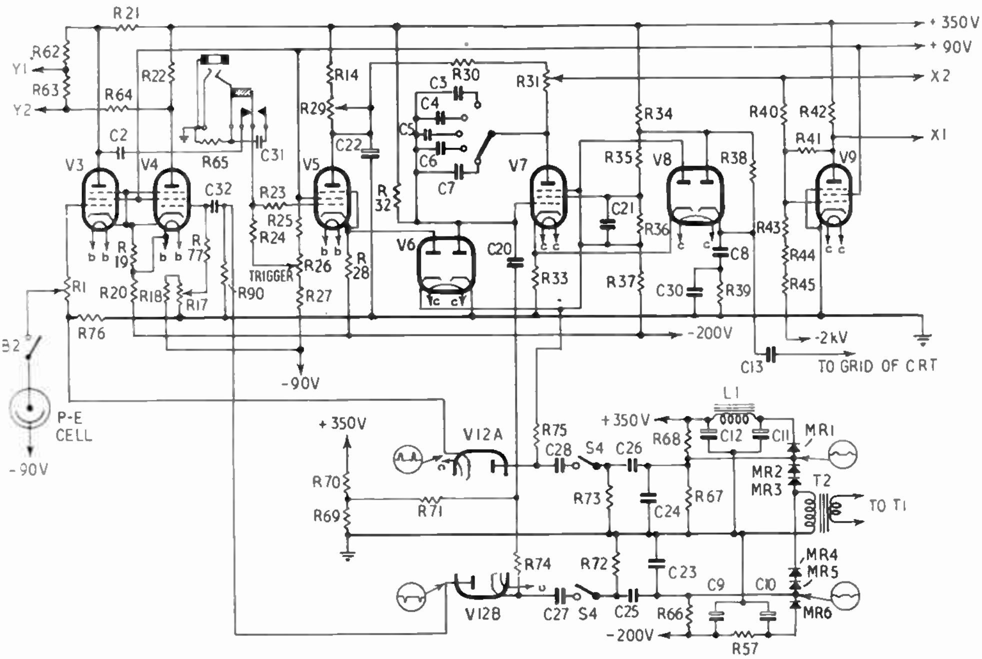


Fig. 7.—Y-amplifier, phantastron time base and pulse generating network of oscilloscope unit.

overcome by using a 10-ohm rheostat, R59, coupled by a 4 : 1 step-down transformer, T3, into the a.c. supply to the bridge rectifier. This arrangement in effect provides a smoothly adjustable series resistance of 0·6 ohm.

4. The Oscilloscope

The oscilloscope used in the instrument was provided with a long persistence tube so that the trace could be examined in detail after each shutter operation. A single sweep time-base was required which was to be triggered automatically on operation of the shutter and the sweep velocity was to be selected automatically by the integrator range switch. It was required also that the oscilloscope should draw its own base line to facilitate detection of shutter bounce. As a refinement, provision was made for superimposing 10-millisecond timing pips on the trace.

4.1. *The Y Amplifier*

The long-tailed pair used to provide a symmetrical Y deflection is represented by V3 and V4 in Fig. 7. Gain control is provided with particular ease since, by adjustment of the slider of R1, the load resistance of the photocell can be varied. In order to obtain time-base triggering at an early stage in the opening of the shutter, high values of anode load resistors R21 and R22 are selected so that a large triggering pulse can be obtained. In consequence, it is found desirable to attenuate the Y signal by the network R62-64. The potentiometer R17 acts as a Y shift control by adjustment of the bias of V4.

4.2. *The Time Base*

As the shutter opens, the rising wave front at the anode of V3 is differentiated by the combination of C2 and R24 to apply to the grid of the cathode follower V5 a positive-going pulse. This positive pulse is transmitted by the diode V6 to trigger the screen-coupled phantastron time-base V7.[5] R29 and R31 serve respectively as X shift and X amplitude controls and an anode follower, V9, provides a symmetrical deflection on the horizontal axis.

If it is desired to study particularly the first part of the shutter operation, a mechanical device may be used to close contacts triggering the time base before the shutter begins to open. By using the jack indicated in Fig. 7, a triggering pulse is applied to V5 through C31, the normal coupling from V3 through C2 being broken at the same time. At the optimum setting of R26, a positive-going step function of about 1 volt will provide reliable triggering. This facility may thus be used, in conjunction with a time calibration of the X-sweep, to test the synchronization of flash contacts. In this case, the jack is connected to an adaptor replacing the flash bulb and the pulse is provided by the usual 1½ volt cell. For routine shutter testing, it is more convenient to use the automatically triggered time base. The high amplifier gain ensures that very little of the shutter characteristic is lost at the beginning of a sweep.

The testing of flash contact synchronization is normally effected separately, using a specialized device.[6]

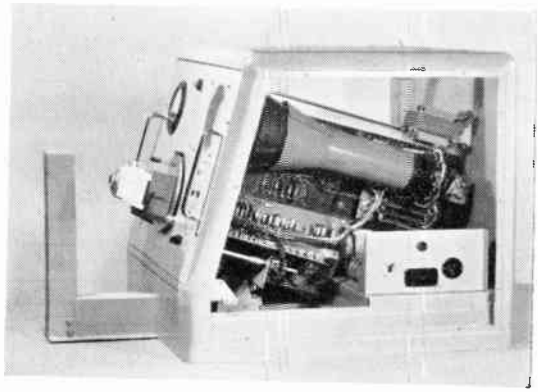
4.3. *Trace Brightening*

In most single-sweep oscilloscopes, it is desirable to suppress the stationary beam so that glare and phosphor burn may be eliminated. A positive-going brightening pulse must therefore be applied to the c.r.t. grid during the active cycle.

During the forward sweep of the time base, a positive-going rectangular pulse appears at the screen of V7. This affords a convenient point from which to obtain a brightening pulse. In order to obtain the base line to the trace, however, it is necessary to extend the duration of this brightening pulse so that the beam current is held on during the re-trace period. For this purpose a pulse stretching circuit has been included comprising V8, R38, R39 and C8. At the lower sweep speeds the re-trace is far from linear but a substantially uniform brightness is maintained on the oscilloscope by choosing the component values of the pulse-stretching network to produce a gradual reduction in beam current as the writing speed decreases.

The capacitor C30 has been added to delay slightly the initial brightening of the c.r. tube and so conceal the jump in the X deflection which is characteristic of the phantastron time base. The time taken to effect this initial jump is determined by the stray capacitance of the grid circuit and hence is substantially the same at all sweep speeds. Consequently the same delay of brightening pulse proves satisfactory at all speeds.

With a high efficiency shutter, having an almost rectangular characteristic, the c.r.t.



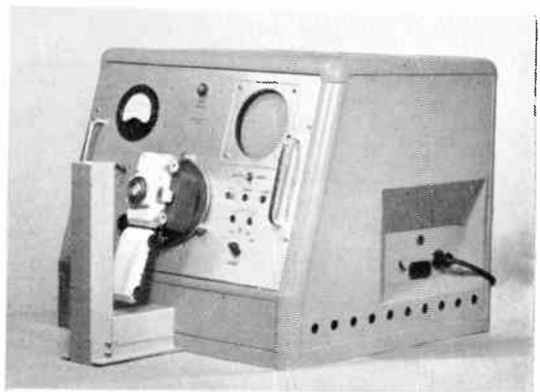


Fig. 8.—(a) Shutter tester with side panel removed. The valves are mounted in a central ventilating duct over the lamp. The power unit, on the right of the picture, may be withdrawn from the back of the instrument.

(b) Complete shutter tester with camera on turntable. Oscilloscope and range switch are on right, integrator meter and controls on left of front panel.

writing speed varies over a wide range. Fortunately this range is restricted somewhat because, owing to the inertia of the shutter blades, the efficiency is limited at the shortest exposures. Nevertheless, it is worth considering whether some form of brightness modulation is not desirable to equalize the brightness of the various parts of the oscilloscope display.

Circuits have been described [2, 7] providing automatic intensification of the c.r. beam in those parts of the display at which the writing speed is high. Such arrangements tacitly assume a restriction of the beam intensity used for the greater part of the display. This may not be a desirable condition, however.

In the single-sweep oscilloscope under discussion, it is necessary to examine briefly a succession of displays produced, perhaps, at $1\frac{1}{2}$-second intervals. If there is to be no confusion between successive displays, the afterglow must decay in $1\frac{1}{2}$ sec to less than half the initial brightness. On the other hand, occasional displays may require more detailed examination. It is then necessary to have a useful amount of afterglow remaining after 10 seconds.

It might appear that these apparently conflicting requirements could be satisfied by using a mixed phosphor screen for the oscilloscope. One component would have a fairly rapid decay whilst the other would provide the long afterglow necessary for detailed examination. In this event, however, the long afterglow component would distinguish most indifferently between the required display and one produced $1\frac{1}{2}$ sec previously.

With a single phosphor, decaying to 50 per cent in $1\frac{1}{2}$ seconds, a much clearer distinction is maintained. A trace which is $11\frac{1}{2}$ sec old remains only half as bright as another which is 10 sec old. On the other hand, the 10-sec trace will have decayed to about 1 per cent of its initial brightness. It is thus important that the initial brightness shall be as high as possible, in order to provide a trace which is clearly visible after 10 seconds.

To this end, the c.r.t. of the present instrument is driven into grid current during the active cycle. This necessarily means that the beam cannot be intensified selectively to compensate for variations in writing speed. Despite this, no practical difficulties are encountered due to variations in trace intensity. It is concluded therefore that phosphor saturation is providing a high degree of compensation for the differences in writing speed. During several years' use, no practical objections have been found to this method of working. On the credit side, the reliability has almost certainly benefited by the elimination of some 6 or 8 valves otherwise required for trace brightening.

4.4. *The Time Marker*

Although, as indicated in Section 2.1, a time marker is scarcely necessary on an instrument incorporating an integrator, it was decided to provide one in order to increase the flexibility of the instrument.

Four methods of providing a time scale were considered :—

- (*a*) By using a double-beam tube,[1] the shutter characteristic being indicated by one beam, the time scale by the other.
- (*b*) By using a double-stroke time base,[2] the shutter characteristic being indicated during the first forward sweep, the time scale during the second.
- (*c*) By using the timing oscillator to modulate the c.r. beam intensity.[8]
- (*d*) By using a ruled graticule, the accuracy of which may be checked by superimposing timing pips on the shutter characteristic.

The relative merits of the various methods may be summarised as follows :—

Method (*a*) provides a time scale simultaneously with the tracing of the shutter characteristic. Consequently the origin of the time scale cannot be made to coincide with a particular part of the characteristic.

Method (*b*) seeks to overcome this limitation by arranging for the second sweep to be triggered by the timing oscillator. This means that the delay between the first and second sweep is variable. Consequently h.t. regulation effects will recover to a variable extent between the two sweeps. The origin and speed of the second sweep are therefore likely to differ from those of the first sweep. These defects could be eliminated by using a highly stabilized h.t. supply, but this would involve the use of more valves and so a reduction in overall reliability.

Method (*c*) dispenses with the need for a double beam tube or a complex time base circuit but, as in (*a*), the time scale cannot be synchronized with the shutter action.

Methods (*a*), (*b*) and (*c*) suffer from the common objection of requiring a 10-kc/s timing oscillator and counting circuits providing additional pulses at 1 kc/s and 100 c/s. The additional valves required may again prove a liability.

Method (*d*) involves least complication and affords the greatest reliability. When timing pips are added to the shutter characteristic, they are not synchronized to it. The graticule may be slid into registration, however, and the X amplitude control may be adjusted until the spacing of the timing marks fits the graticule. The origin of the graticule is subsequently set opposite that of the X sweep and the timing pips need not be used.

Since interpolation is provided by the rulings of the graticule, there is no need for the higher frequency timing pips. The sweep speed used for 1/200 sec exposures may be checked quite readily using 100-c/s pulses and, of course, these pulses may be used for the longer exposures as well. Method (*d*) thus dispenses with the need for the timing oscillator and counting circuits, for the 100-c/s pulses may be derived from the 50-c/s mains with sufficient accuracy.

Method (*d*) was therefore adopted as it promised to be simpler and more reliable than any of the alternatives whilst yielding results at least as satisfactory.

Two similar networks, shown in the lower half of Fig. 7, are used to generate complementary trains of 50-c/s pulses. One pulse train consists of positive-going pulses which are 10 milliseconds in advance of the complementary, negative-going pulses. A potential divider is connected across the positive h.t. supply so that about 200 V is developed across R67. During that fraction of the cycle in which the rectifier voltage exceeds 200 volts, however, the voltage across R67 increases. A train of part-sine waves is therefore applied to C26 which provides a differentiated signal across R73. The waveform is further differentiated by C28 and R75 so that a positive-going pulse is produced when the voltage on R67 increases and a negative going pulse when it returns once more to 200 volts. The diode V12 passes the

(*a*) (*b*) (*c*) (*d*) (*e*)

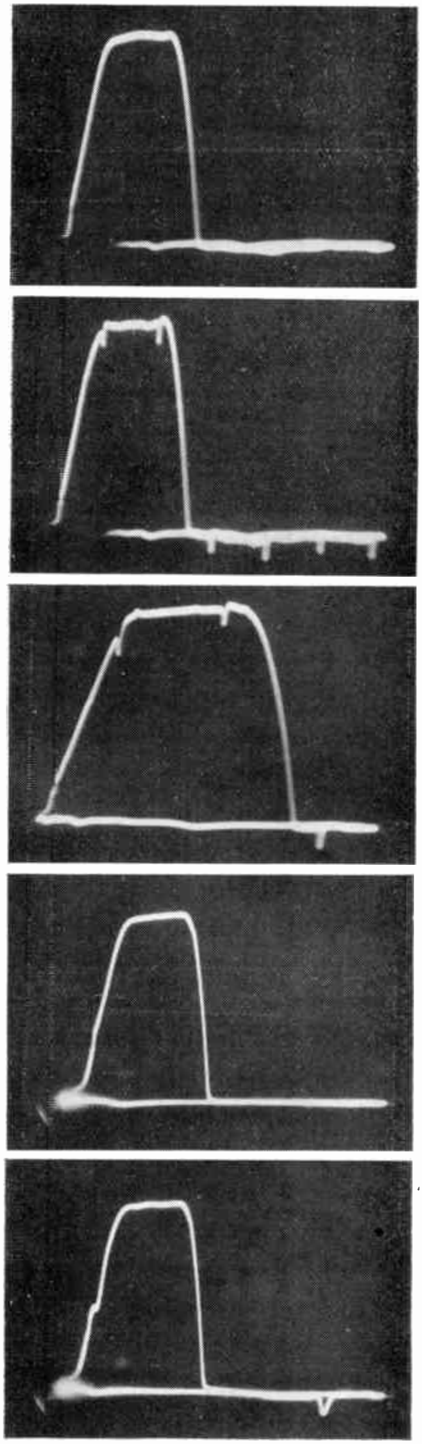


Fig. 9.—Typical oscillograms of "Advocate" shutter.

(*a*) *Nominal exposure 1/50 sec, tested with range switch set at 1/50 sec. The relatively slow time base facilitates detection of any shutter bounce.*

(*b*) *As (a), but with 10 msec timing pips added.*

(*c*) *As (b) but range switch set to 1/100 sec. The faster time base provides greater accuracy of time measurements.* $t_e = 20$ *msec,* $t_o = 23$ *msec,* $\eta = 83$ *per cent.*

(*d*) *Nominal exposure 1/200 sec, range switch set to 1/200 sec.*

(*e*) *As (d) but with 10 msec timing pips added.* $t_e = 4{\cdot}7$ *msec,* $t_o = 5{\cdot}6$ *msec,* $\eta = 84$ *per cent.*

positive-going pulses to the grid of V3 in which they are amplified to produce negative-going pulses in the trace. In a similar way, the complementary train of negative pulses is applied to the grid of V4 so that the final trace shows equally spaced 100-c/s pulses.

It is, of course, essential that the timing pulses shall appear on the oscilloscope in a negative-going direction in order that they shall not produce spurious triggering of the time-base circuit.

A complication arose due to the necessity for eliminating the pulses during the return stroke of the time base. If this is not done, the pulses appear unequally spaced during the non-linear re-trace. Fortunately, the grid and suppressor of V7 provide the necessary rectangular waveforms with which to bias the two halves of the diode V12 so that the pulses are transmitted to V3 and V4 only during the forward sweep.

It is important that the origin and velocity of the X sweep shall remain the same whether or not the timing pulses are added to the Y deflection. For this reason, the application of the timing pulses must not introduce h.t. regulation effects. Accordingly the pulse-generating network remains operative at all times, only the high impedance signal to the amplifier grids being switched. The push-pull arrangement of the Y amplifier ensures a substantially constant current drain whether or not the pulses are used.

5. Conclusion

The testing of between-lens camera shutters is greatly simplified by an instrument embodying a photoelectric integrator as well as an oscilloscope display. The integrator measures the effective exposure time and the oscilloscope can be used to determine the total open time of the shutter, besides indicating faults such as sticking or bouncing of the shutter blades.

6. Acknowledgments

The authors thank the directors of Ilford Limited for permission to publish this paper.

7. References

1. K. Onwood, "Apparatus for the investigation of inter-lens shutters," *Electronic Engineering*, **20**, p. 9, 1948.
2. D. T. R. Dighton and H. McG. Ross, "Cathode-ray tube shutter-testing instrument," *Journal of Scientific Instruments*, **24,** p. 128, 1947.
3. A. L. Biermasz and A. J. Michels, "An electronic d.c. millivoltmeter," *Philips Technical Review*, **16**, p. 117, October 1954.
4. M. J. Tucker, "A high-sensitivity photo-electric cell circuit," *Electronic Engineering*, **20,** p. 401, 1948.
5. F. C. Williams and N. F. Moody, "Ranging circuits, linear time-base generators and associated circuit," *J. Instn Elect. Engrs*, **93,** Part IIIA, p. 1188, 1946.
6. R. J. Hercock and D. M. Neale, "A device for testing the synchronization of shutters for use with incandescent flash bulbs." *Photographic Journal*, **92B**, p. 106, July-August 1952.
7. J. de Klerk, "Automatic c.r.t. trace brightening for varying amplitude r.f. signals," *Electronic Engineering*, **25**, p. 388, 1953.
8. J. D. Kelley, "Intensity curves of light pulses of short duration," *J. Optical Soc. America*, **28**, p. 27, January 1938.

COLOUR TELEVISION - B.B.C. EXPERIMENTS

It was announced in the October *Journal* that the B.B.C. has started a series of experimental colour television transmissions from the London station at Alexandra Palace and a demonstration was given recently of the current state of the investigations to which a representative of the Institution was invited. The Director of Technical Services of the B.B.C., Sir Harold Bishop, gave the background to the tests and pointed out that the trials did not indicate the imminence of a public service but were purely for obtaining information on various possible systems. This information would be considered by the Television Advisory Committee which, at the request of the Postmaster-General, was reporting on colour television.

The tests at present are concerned with investigating the American N.T.S.C. system as modified to British standards. The colour standards evolved by the National Television Systems Committee were approved by the F.C.C. in December 1953 and a service of a few hours per week is being radiated in the U.S.A. at present. The principal features of the N.T.S.C. signal are:

1. The colour signal is transmitted in the same radio frequency channel and by the same transmitters as carry the established monochrome service.
2. It is claimed that the system is "compatible," i.e. that existing monochrome receivers can produce a monochrome version of the colour picture which is as good as if the picture had originated from a normal monochrome camera.
3. It is further claimed that the standards are such as to allow for considerable future development in the quality of the colour picture, in the same way as the original specification for the monochrome television service has allowed a continuous improvement in quality over the course of the years.

Since the scanning and transmission standards of the U.S.A. and this country differ in important ways the question as to whether the N.T.S.C. system would show the same advantages when modified to suit British television standards has been investigated by the B.B.C. Research Laboratories and certain industrial organizations.

The equipment which has been installed at Alexandra Palace is intended:

1. To explore the degree of compatibility of the system by making observations on some thousands of monochrome receivers.
2. To see whether the system is capable of producing a consistently good quality colour picture.

The tests in connection with the first question are already proceeding and it is hoped to provide a statistical answer in due course. Naturally, since colour pictures are being transmitted, some experience and knowledge is being obtained on the second point, but no wide-scale observations are yet taking place because sufficient colour receivers are not yet available.

The N.T.S.C. Colour Signal.—It is well-known that the human eye's perception of practically any colour can be stimulated by additive mixtures of red, green and blue light.[1] The N.T.S.C. system therefore reproduces these primary colours at the receiver at intensities controlled by three separate signals from the transmitter. Three signals are transmitted as: (*a*) a luminance (brightness) component, and (*b*) a chrominance (colour) component, having two separate parts.

The luminance component is the same as that which would be produced by a panchromatic monochrome television camera looking at the same scene, and this signal therefore produces a normal monochrome representation of the coloured scene on a monochrome receiver.

The chrominance component consists of two colour-difference signals which, in the simplest terms, may be said to convey the hue and degree of saturation of the colour information. In the colour receiver, these three signals representing brightness, hue and saturation are combined to produce the required intensity from each of the red, green and blue lights. The fact that a monochrome receiver and a colour receiver can each produce simultaneously its own version of the scene from the same signal gives the N.T.S.C. system its valuable feature of "compatibility."

It would be possible to transmit the chrominance signal quite independently of the luminance signal and in this case the compatibility would be virtually perfect. However, the second unique feature of the N.T.S.C. signal is that the two components have been combined in such a way that they occupy the same total bandwidth as that used by the equivalent monochrome signal. Due to the manner in which the human eye perceives colour, the separation of luminance and chrominance enables the bandwidth of the chrominance signal to be reduced to about one third of that of the luminance. Further saving of bandwidth is achieved by placing this reduced bandwidth information at the upper end of the luminance band in such a way that the inevitable interference (cross-talk) between the two signals has a minimum effect on the compatible picture on the monochrome receiver. The actual mechanism, by which this band sharing takes place, employs a colour sub-carrier (in the British version 2·66 Mc/s) which is simultaneously modulated in amplitude and phase by the two colour difference signals, the carrier itself being suppressed so that the chrominance signal exists only when colour is present in the scene being transmitted. The colour sub-carrier is an odd multiple of half the line scanning frequency, and, under these circumstances the visibility of the best pattern produced between it and the scanning lines is a minimum.

This combination of band saving, band sharing, suppressed carrier modulation and "frequency interleaving" is claimed in the U.S.A. to produce an adequately compatible signal. Whether or not such is the case in the British version applied to typical domestic receivers in this country is the chief matter under investigation at the present time.

DESCRIPTION OF THE PRESENT EQUIPMENT

The main items of equipment installed at Alexandra Palace are: (1) colour slide and film scanner; (2) colour camera; (3) signal coding equipment; (4) colour picture monitors; and (5) colour test equipment.

Colour Slide and Film Scanner.—The colour slide and film scanner is the source of the pictures which are being transmitted for the present series of tests of the compatability of the N.T.S.C. signal. It produces pictures from slides either $3\frac{1}{4}'' \times 3\frac{1}{4}''$ or $2'' \times 2''$ or from 16-mm film, by selection of the appropriate optical system.

The scanner employs the flying spot principle and the source of light is therefore a cathode ray tube of which the phosphor emits light as evenly as can be achieved over the whole of the visible spectrum. The light from the raster on the face of the scanning tube is passed either through the slide or the film as desired and the coloured image so produced is then split into three separate parts, which represent respectively the red, green and blue information in the picture. This colour analysis process is performed by a combination of dichroic mirrors, colour filters, plane mirrors and lenses. The three colour separation pictures, which emerge from the analyser as three physically separate rays of light, are then focused each on to a photo-multiplier tube which turns the intensity of the light, which is varying in accordance with the scene being scanned, into corresponding electric voltages. The three voltages are then passed through three separate and identical chains of electronic equipment which supply gamma correction, correction for the distortion introduced by the finite decay time of the light from the scanning tube phosphor, and equalization for aperture loss, exactly as in the case of a monochrome flying spot scanner.

The film transport mechanism is a standard intermittent motion 16-mm projector with a "pull-down" time of about 4 milliseconds. Since the time available for "pull-down" is only 1·4 milliseconds if all the lines of the television picture are to contain information, some picture information is inevitably lost. This loss occurs at the top and bottom of the picture where about 15 lines are presented as black. In order to preserve the usual aspect ratio of 4 : 3 an equivalent area at the sides of the picture is also black. The picture therefore appears as in a black frame, but this disadvantage is accepted because the arrangement permits of a simple and efficient optical system. Synchronism between the film motion and the television picture repetition rate is achieved in a simple way by supplying power to the synchronous motor of the film transport mechanism by amplifying the 50 c/s component of the frame pulses.

The photograph of the scanner in Fig. 1 shows the principal mechanical features. At the bottom right of the right-hand cubicle is a large rectangular box containing the scanning tube. The recessed three-spoked handle moves the scanning tube physically for focusing the image. The optical system contained in the drum in the middle left deflects the light from the raster either into the slide holder immediately above or into the film scanning mechanism at the right. The rectangular box at the top left, which is on pivots, contains the colour analysing filters and the three photo-multiplier tubes. It is shown in the slide scanning position: for film it is swung over to the right so that it accepts light emerging from the film scanner through the funnel-shaped outlet above the transport mechanism. The rack on the left contains power supplies, a control panel and a monochrome picture monitor. A further cubicle (not shown) contains the electronic equipment associated with the red, green and blue separation signals.

The Colour Camera.—Coloured light entering the lens of the camera is split into three colour separation images by a colour analyser similar in principle to that used in the slide and film scanner. In place of the three photo-multiplier cells are three image orthicon camera tubes of a type developed specifically for colour work. These tubes produce the three colour separation signals in electrical form. Each of the tubes is supplied with the necessary scanning waveforms and electrode potentials just as in the case of the single-tube monochrome camera. It will be realized that the output of each tube is a separate picture of which not only the transfer-characteristic between light input and voltage output must be maintained in a precise manner for the three signals, but the geometry of the three pictures must be the same within very close limits so that any particular detail of the picture occurs at the same point in the scanning cycle of all three.

The signals from the tubes are amplified in the camera and transmitted to the control room over three identical cables. In the control room, each signal is gamma corrected and equalized in a manner very similar to that used in monochrome equipments employing the same type of camera tube, and finally emerges as a colour separation signal of the same form as that produced by the slide and film scanner.

The control desk of the camera is seen in the foreground of Fig. 3, which shows the control room. The three sets of controls, one for each camera tube, can be clearly seen. The electronic equipment for the camera is mounted in the cubicle nearest to the control desk.

Signal Coding Equipment.—The signal coding equipment includes the special colour waveform generating equipment and the "encoder" in which the luminance and chrominance signals are formed from the incoming three-colour information.

Fig. 1.—Colour television scanning equipment for showing 3¼" × 3¼", and 2" × 2" slides, and 16mm film.

The "master" frequency, from which all the other scanning and pulse waveforms are derived, is obtained from a temperature controlled crystal oscillator whose frequency is 2·6578125 Mc/s ± 8 c/s. This frequency is multiplied and divided to produce the usual double line frequency of 20,250 cycles/second (i.e. 4/525 times sub-carrier) from which the standard 405-line interlaced waveform is generated. (It will be noted that the frame repetition rate is asynchronous with respect to mains frequency, in contrast to the existing monochrome service in which synchronous working is almost always employed. Multiple outputs of line and frame trigger pulses, mixed synchronizing pulses and mixed suppression pulses are available.

in the suppression period following every line synchronizing pulse. This "burst" is used at the receiver to synchronize a sub-carrier generator which is needed for detection of the quadrature modulated chrominance signal.

The waveform generator and the encoder are mounted in the two cubicles adjacent to the camera control equipment. The three other

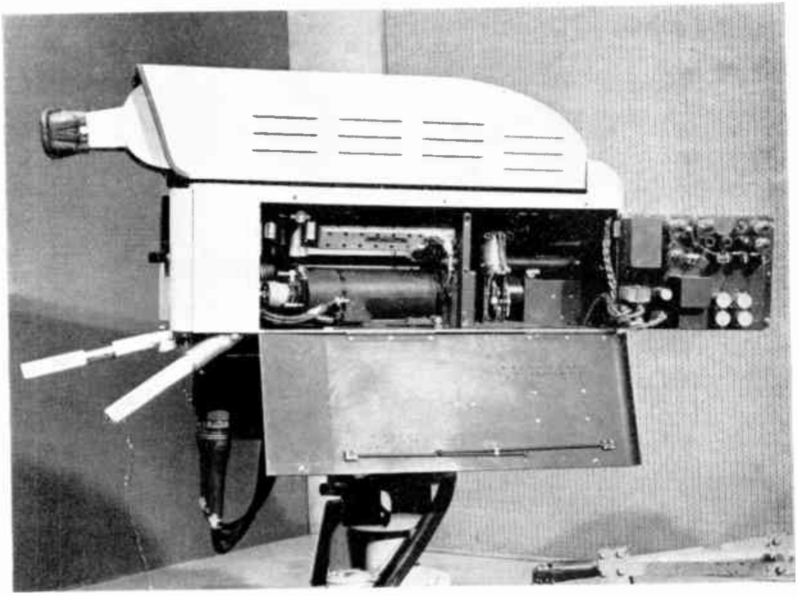


Fig. 2.—Three-tube colour camera, side view, showing one of the camera tubes in its yoke and the associated amplifiers.

(Note.—The four-position lens turret is obscured by the small hinged chassis on the right which has been swung back to show the focusing and lens changing devices.)

The input to the encoder consists of the three gamma corrected colour separation signals (red, green and blue) which are produced by either the slide and film scanner or by the camera. The encoder may be considered as performing a single linear transformation of the three incoming signals, red, green and blue, to the other three quantities, Y, I and Q, of which Y is the luminance signal. The colour sub-carrier is then modulated by the I and Q signals in such a way that the amplitude of the resultant signal conveys the saturation information and the phase conveys the hue. In the absence of colour information the sub-carrier is suppressed. The complete chrominance signal is added to the luminance which is, of course, in video form. Finally, the synchronizing waveform is added to produce the complete waveform. The synchronizing waveform is of the normal type except that a "burst" of nine cycles of the colour sub-carrier is added

cubicles in the background at the right supply power for the whole of the equipment, with the exception of the slide and film scanner.

Colour Picture Monitors.—There are two colour picture monitors. One employs three separate tubes, the phosphors of which emit respectively red, blue and green light. The application of the colour separation signals to the grids of these tubes produces three colour separation images which are combined optically by dichroic mirrors to produce a direct viewed colour picture.[2] This method brings with it the attendant difficulty of superimposing the three separate images accurately, just as in the colour cameras. However, up to the present, this method produces the best pictures and its complication is worthwhile in a monitor intended for technical purposes. This monitor is seen in the centre of the photograph of the control room.

The other monitor uses a 15-in. R.C.A. shadow-mask tri-colour tube. Since the monitor incorporates its own decoder, the input signal is of the N.T.S.C. type and the unit is therefore used for general checking and monitoring of the transmitted signal. It can be seen on the extreme right of Fig. 3.

Colour Test Equipment.—The complicated nature of the N.T.S.C. signal requires special test signals and measuring apparatus to ensure that its specification is met. The main signal for this purpose, "colour bars," is generated electronically and produces on the picture monitor seven vertical strips which, from left to right, are white, yellow, cyan (blue-green), green, magenta (purple), red and blue. These signals represent saturated colours for which the amplitude and phase of the colour sub-carrier are known. The amplitude is measured in the usual way with a waveform monitor; the phase is measured by a special piece of test equipment known as a Colour Signal Analyser. Distortion occurring in the transmission of the signal after it has left the encoder can, of course, be measured similarly.

Other electronically generated signals such as "dots" and a grid pattern of lines covering the whole picture are provided for the purpose of adjusting the picture monitors. The camera and slide and film scanner have a series of special test cards for the alignment of the apparatus.

Note on Compatibility

During the B.B.C. demonstration, colour slides, films and studio shots were shown on colour receivers and also on monochrome receivers of the studio monitor type. The latter showed that the compatibility was satisfactory; the slight imperfections, e.g., crawling dot pattern, were only visible at ex-

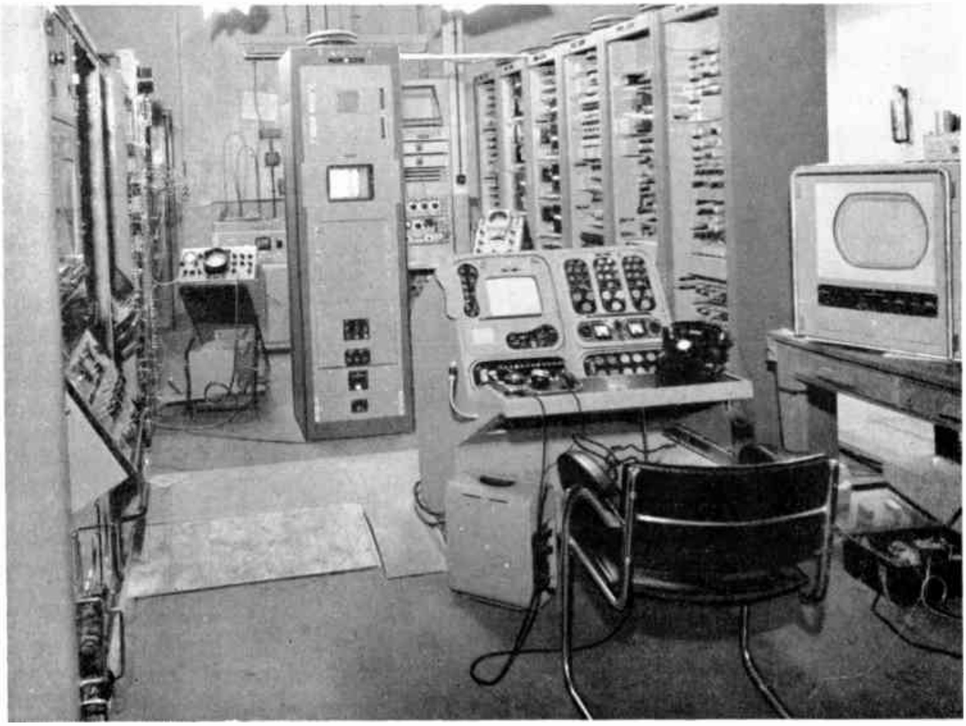


Fig. 3.—General view of colour studio control room—left, film scanner—right, power supply and pulse generator equipment—centre, control console and three-tube colour picture monitor—on extreme right, radio check colour receiver.

tremely close range and would almost certainly not be noticeable on a standard commercial receiver. There was no evidence of any interaction of the colour information signals on vision or sound. The reproduction on the colour receivers was good for bold colours but the more delicate pastel and flesh tints, as well as white, were not reproduced very faithfully; there was also a certain amount of colour fringeing. Possibly these faults may have been due to the three-gun masked-dot tubes used which, it is understood, were not the latest type.

References

The interest in the theory and practice of colour television is, of course, very considerable and a number of local section meetings during the present session are being devoted to it. These include papers by Mr. D. W. Heightman, (*Member*), Dr. G. N. Patchett (*Member*), and Mr. B. V. Somes-Charlton. The Papers Committee will be considering these papers for possible publication in the *Journal* in due course.

Meanwhile members may like to be reminded of the following papers on colour television which have been published in the *Journal*.

1. J. E. Benson, "A survey of the methods and colorimetric principles of colour television," January 1953.
2. A description of a dichroic mirror system with three tubes, similar to that used in one of the B.B.C. monitors, is given in "A 15-by-20-in. projection receiver for the R.C.A. colour television system," April 1950.

Mention is made of certain aspects of colour television in the following papers:—

3. D. A. Bell, "Economy of bandwidth in television," September 1953.
4. L. H. Bedford, "Problems of television cameras and camera tubes," October 1954.

Other colour systems developed in the U.S.A. prior to the N.T.S.C. system are described in reference 2 above and in "Synchronization for colour dot interlace in the R.C.A. colour television system," (April 1950), and "Description of C.B.S. colour television system," (April 1950).

SPECIFICATION OF THE MODIFIED N.T.S.C. SIGNAL

Time delay between luminance and chrominance

E_Y* E_Q* E_I* and the components of these signals match each other in time within 0.07 μ sec. (* Gamma corrected signals)

Delay specification

The overall delay characteristics of the transmitted signal conform to the shape set out in the following table to within ± 0.07 μ sec.

f	Envelope Delay
0.1 Mc/s	+0.22 μ sec.
1.0	+0.22
1.5	+0.22
1.8	+0.19
2.0	+0.17
2.2	+0.133
2.4	+0.085
2.6	+0.015
2.8	−0.12
3.0	−0.35

Colour sub-carrier frequency

2.6578125 Mc/s ± 8 c/s with a maximum rate of change of 0.1 c/s/s.

Q Channel bandwidth

at 300 kc/s less than 2 db down
at 340 kc/s „ „ 6 db „
at 450 kc/s „ „ 6 db „

I Channel bandwidth

at 1.0 Mc/s less than 2 db down
at 2.5 Mc/s „ „ 20 db „

Picture carrier/sound carrier frequency relation

The difference in frequency between the picture signal carrier and the sound signal carrier is a multiple of the line frequency. The exact relation is:

$$f_{c(vision)} = f_{c(sound)} + 350\, f_{(line)}$$

All but the last of these items are direct modifications of the American specification to suit British standards. The last is an additional feature of the British signal and its purpose is to make the beat frequency between the sound carrier and colour sub-carrier an odd multiple of half line frequency, so that the pattern produced on the screen is both stationary and frequency-interleaved, thus giving minimum visibility. Only receivers which have an appreciable amount of the sound carrier present in the input to the vision detector will produce this beat frequency, but since its frequency is low (approx. 880 kc/s), every precaution to minimize its effects is worthwhile. In the U.S.A. frequency modulation, as opposed to the British amplitude modulation, is employed for the sound signal. This obviously makes rigid frequency locking an impossibility.

EXAMINATION OF THE "NEGATIVE FREQUENCY" CONCEPT †

by

ir. A. P. Bolle‡ and ir. J. L. Bordewijk‡

SUMMARY

The advantage of the combined use of positive and negative angular frequencies over the current rise of positive angular frequencies only is pointed out. A number of examples are given to illustrate this, e.g., complex symbolism for linear fixed and variable networks, the Fourier series and integral, and the calculation of intermodulation noise.

1. Introduction

In many fields of engineering and physics the simple harmonic variable quantities, or compositions of such quantities, play an important part. By way of example we refer to the well-known phenomenon that the projection on any diameter of a point moving with uniform speed along the circumference of a circle describes a simple harmonic motion. Of this projection, the distance to the centre, the path covered, the speed and the acceleration are, *inter alia*, simple harmonic variable quantities and consequently they can be mathematically described with the aid of the well-known trigonometric functions, " sine " and " cosine ".

2. Geometric Interpretation of Simple Harmonic Variable Quantities

The distance of the above-mentioned projection to the centre or, which amounts to the same thing, the length of the projection of a line of constant length rotating around a fixed point with uniform angular velocity (vector) is considered as a very useful geometric representation of simple harmonic variable physical quantities in the various technical applications.

Figure 1 gives an elucidation. A line (vector) with length E rotates counter-clockwise with constant angular velocity ω around the point P. The length of the projection of this line on any arbitrary line through P is a simple harmonic variable quantity. For convenience's sake it has been assumed in Fig. 1 that at the time $t = 0$ the rotating line coincides with the line onto which the projection is made. For the length of the projection e is then given by:

$$e = E \cos \omega t$$

It is to be observed that in Fig. 1 preference is given to a counter-clockwise direction of rotation which actually is not justified by any physical or mathematical consideration.

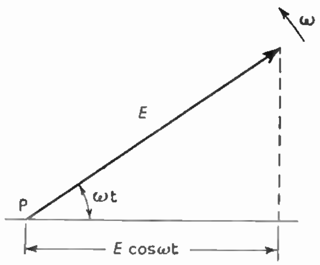


Fig. 1

Besides the representation contained in Fig. 1, with a single rotating line of length E, another representation is possible which gives e as a vectorial sum of two lines (vectors), each with a length $E/2$, and with equal but opposite angular velocities. For convenience's sake it is assumed that the two rotating lines coincide at the time $t = 0$ (cf. Fig. 2). Through the latter manner of representation the concept of negative angular frequency has now been introduced.

3. Phasor Concept

As follows from the above, simple harmonic quantities are characterized by two indices. Quantities with two indices can mathematically

† Reprinted by permission of the Editors, from *Het PTT—Bedrijf*, Vol. 6, No. 2, October 1954. (Paper No. 332.)

‡ Central PTT Laboratory of the Postal and Telecommunications Service of the Netherlands, The Hague.
U.D.C. No. 517.512.2 : 621.37.

be represented in an elegant manner by complex numbers, since a complex number can also be described by two indices, for example a real and an imaginary part.

Now if two complex numbers with modulus $E/2$ move around the circumference of a circle with equal but opposite velocities, both starting from the real axis at the moment $t = 0$, then the sum of these complex numbers is a real number :

$$e = \tfrac{1}{2} E \left[\exp (j\omega t) + \exp (-j\omega t)\right]$$

This sum can be constructed in the well-known graphical manner (cf. Fig. 3).

Due to the similarity of the rules for adding vectors and complex numbers, the latter are sometimes called vectors too. This may cause confusion, because the further rules are fundamentally different. Therefore some mathematicians indicate the arrow connecting the point which is the geometrical representation of the complex number with the origin by the name of " phasor ".

Thus a simple harmonic quantity can also be represented as the sum of two phasors

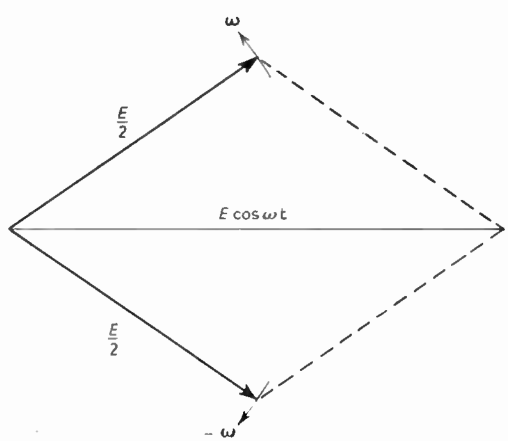


Fig. 2

rotating in opposite directions. The vector and phasor representations of a simple harmonic quantity are obviously identical because :

$$2 \cos \omega t = \exp (j\omega t) + \exp (-j\omega t).$$

Through this phasor notion the simple variable harmonic quantity is represented as the sum of two phasors, one of which has a positive angular velocity or angular frequency, and the other a negative one. Here the negative angular frequency concept has therefore been likewise introduced.

The positive angular frequency alone can also be used in phasor representations. Then the following may be written :

$$e = E \cos \omega t = \mathrm{Re}\left[E \exp (j\omega t)\right]$$

Apart from the disadvantage that one direction of rotation is given preference, the fact is

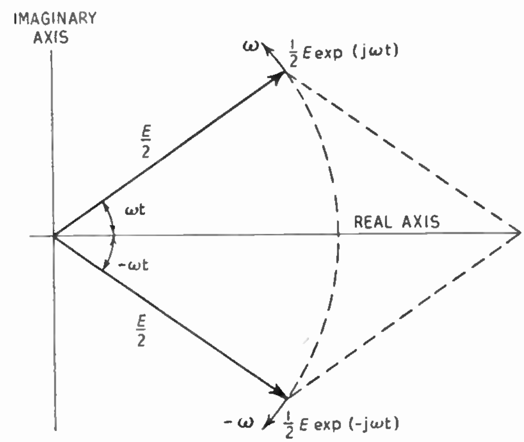


Fig. 3

to be observed that use is made of a non-analytical function Re [. . .], with all the relative objections.

Instead of negative angular frequencies it is better to speak of negative argument, because in order to represent $\cos (\omega t + \varphi)$ two phasors are required :

$$\tfrac{1}{2} \exp j (\omega t + \varphi) \text{ and } \tfrac{1}{2} \exp -j (\omega t + \varphi)$$

From this it will be clear that now there are two starting positions for the phasor, image-symmetrical with regard to the real axis, and that there is no longer any preference as to either starting position (cf. Fig. 4).

The use of the cosine function instead of the sine function in the above is not essential. The relation between the cosine function and the exponential is slightly simpler, however, and therefore it was preferred here.

4. Applications

4.1. In electrical engineering the concept of negative angular frequency or negative argument can be advantageously used in analysing linear

electrical circuits. For example if the applied voltage is: $e = E \cos \omega t$,
first the response I due to the complex voltage: $E \exp(j\omega t)$, can be calculated. The response II due to the complex voltage $E \exp(-j\omega t)$, is then known at once, so that the response due to the applied voltage: $E \cos \omega t$, is found to be half the sum of I and II, on the basis of the superposition principle in linear networks.

For a series arrangement L and R the following holds:

$$L\frac{\mathrm{d}i}{\mathrm{d}t} + iR = e$$

For the inserted voltage:

$$e = E\cos\omega t = \frac{E}{2}\Big[\exp(j\omega t) + \exp(-j\omega t)\Big]$$

The current $\bar{i}_{\mathrm{I}}$ due to the complex voltage:

$$\bar{e} = E\exp(j\omega t),$$

is determined by assuming:

$$i_{\mathrm{I}} = I\exp j(\omega t + \varphi)$$

The result is:

$$i_{\mathrm{I}} = \frac{E}{(R + j\omega L)}\exp(j\omega t)\ I\exp j(\omega t + \varphi)$$

The current $\bar{i}_{\mathrm{II}}$ due to the complex voltage:

$$E\exp(-j\omega t)$$

then becomes:

$$\bar{i}_{\mathrm{II}} = \frac{E}{(R - j\omega L)}\exp(-j\omega t) = I\exp -j(\omega t + \varphi)$$

The current i due to the inserted voltage: $E \cos \omega t$, is now found from:

$$i = \frac{\bar{i}_{\mathrm{I}} + \bar{i}_{\mathrm{II}}}{2}.$$

It is further found that:

$$i = \frac{E}{\sqrt{R^2 + \omega^2 L^2}}\cos(\omega t + \varphi) = I\cos(\omega t + \varphi)$$

from which:

$$\tan\varphi = -\omega L/R$$

It is easier to determine the values of I and φ with the aid of:

$$I = \sqrt{i_{\mathrm{I}}\,.\,i_{\mathrm{II}}} = \frac{E}{\sqrt{R^2 + \omega^2 L^2}}$$

and

$$\exp j(\omega t + \varphi) = \sqrt{\frac{\bar{i}_{\mathrm{I}}}{\bar{i}_{\mathrm{II}}}} = \sqrt{\frac{R - j\omega L}{R + j\omega L}}\exp(j\omega t)$$

so that

$$\tan\varphi = -\omega L/R$$

4.2. In trigonometry the operations with exponentials having positive and negative exponents are very simple. This is illustrated, for example, in calculations of products with cosine and sine functions. Multiplications of exponentials are more easily carried out and are easier to remember than multiplications of trigonometric functions.

For example:

$$\begin{aligned}
\cos^2\omega t &= \tfrac{1}{2}[\exp(j\omega t) + \exp(-j\omega t)]^2 \\
&= \tfrac{1}{4}[\exp(2j\omega t) + \exp(-2j\omega t)] + \tfrac{1}{2} \\
&= \tfrac{1}{2}\cos 2\omega t + \tfrac{1}{2}
\end{aligned}$$

$$\begin{aligned}
\cos^3\omega t &= \tfrac{1}{2}[\exp(j\omega t) + \exp(-j\omega t)]^3 \\
&= \tfrac{1}{8}[\exp(3j\omega t) + \exp(-3j\omega t)] \\
&\quad + \tfrac{3}{8}[\exp(j\omega t) + \exp(-j\omega t)] \\
&= \tfrac{1}{4}\cos 3\omega t + \tfrac{3}{4}\cos\omega t
\end{aligned}$$

$$\begin{aligned}
&\cos\omega_1 t\cos\omega_2 t \\
&= \tfrac{1}{2}[\exp(j\omega_1 t) + \exp(-j\omega_1 t)]\,.\\
&\qquad .\,\tfrac{1}{2}[\exp(j\omega_2 t) + \exp(-j\omega_2 t)] \\
&= \tfrac{1}{4}[\exp j(\omega_1 + \omega_2)t + \exp -j(\omega_1 + \omega_2)t \\
&\qquad + \exp j(\omega_1 - \omega_2)t + \exp -j(\omega_1 - \omega_2)t \\
&= \tfrac{1}{2}\cos(\omega_1 + \omega_2)t + \tfrac{1}{2}\cos(\omega_1 - \omega_2)t.
\end{aligned}$$

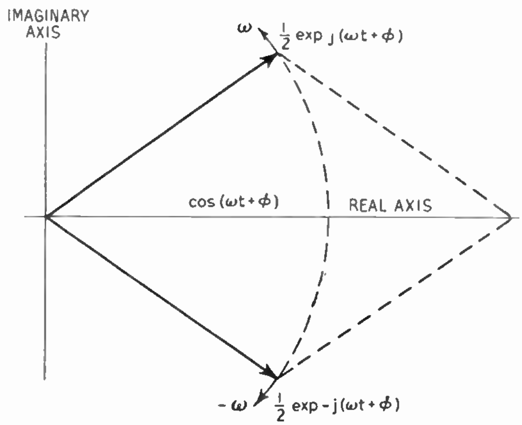


Fig. 4

4.3. In calculations of intermodulation noise which are *inter alia* important in the design of multiplex transmission systems, the introduction of negative angular frequencies often produces much simplification.

This is particularly apparent in the graphical method for the determination of intermodulation noise, because in that case there is no difference between sum and difference frequencies.

Since this would be too involved a matter to deal with here, reference should be made to a paper by one of the present authors.[2]

5. Complex Frequencies

If the inserted voltage, again for the series connection of L and R, is the following:

$$e = E\,[\exp(at)]\cos\omega t$$

e can be given as the sum of two complex terms:

$$e = \tfrac{1}{2}E\,[\exp(\lambda t) + \exp(\lambda^* t)]$$

in which $\lambda = a + j\omega$

and $\lambda^* = a - j\omega$

By analogy with the example of Section 4, the solution for i is given as:

$$i = (\tfrac{1}{2}i_{\mathrm{I}} + i_{\mathrm{II}}) = \tfrac{1}{2}I\exp(j\varphi)\exp(\lambda t) + \tfrac{1}{2}I\exp(-j\varphi)\,.\,\exp(\lambda^* t)$$

After substitution in the differential equation of Section 4 the following solution is found:

$$i = \tfrac{1}{2}\left[\frac{E}{R+\lambda L}\exp(\lambda t) + \frac{E}{R+\lambda^* L}\exp(\lambda^* t)\right]$$

From this it follows that:

$$i\exp(at) = \sqrt{i_{\mathrm{I}}\,.\,i_{\mathrm{II}}} = \frac{E}{\sqrt{(R+aL)^2+\omega^2L^2}}\exp(at)$$

$$\exp(j\varphi)\exp(j\omega t) = \sqrt{\frac{i_{\mathrm{I}}}{i_{\mathrm{II}}}} = \sqrt{\frac{R+\lambda^* L}{R+\lambda L}}\exp(j\omega t)$$

so that:

$$I = \frac{E}{\sqrt{(R+aL)^2+\omega^2L^2}}$$

and

$$\exp(j\varphi) = \sqrt{\frac{R+\lambda^* L}{R+\lambda L}}$$

from which it follows that:

$$\tan\varphi = -\frac{\omega L}{R+aL}$$

6. Application to the Fourier series and the Fourier integral

The system of writing in the Fourier series is simplified by application of negative angular frequencies, when the analogy with the Fourier integral is more clearly shown.

If $\xi(t)$ be a periodical function with a period T (see Fig. 5) the following can be written:

$$\xi(t) = p_0 + \sum_{n=1}^{n=\infty}(p_n\cos n\omega_0 t + q_n\sin n\omega_0 t)$$

in which:

$$p_0 = \frac{1}{T}\int_{-\frac{1}{2}T}^{+\frac{1}{2}T}\xi(t)\,dt\,; \quad p_n = \frac{2}{T}\int_{-\frac{1}{2}T}^{+\frac{1}{2}T}\xi(t)\cos n\omega_0 t\,dt$$

and

$$q_n = \frac{2}{T}\int_{-\frac{1}{2}T}^{+\frac{1}{2}T}\xi(t)\sin n\omega_0 t\,dt$$

$$\left(\omega_0 = \frac{2\pi}{T} = 2\pi f_0\right)$$

However, when negative angular frequencies are used the expression for $\xi(t)$ becomes:

$$\xi(t) = \sum_{n=-\infty}^{n=+\infty}(p'_n\cos n\omega_0 t + q'_n\sin n\omega_0 t)$$

in which

$$p'_n = \frac{1}{T}\int_{-\frac{1}{2}T}^{+\frac{1}{2}T}\xi(t)\cos n\omega_0 t\,dt$$

and $q'_n = \dfrac{1}{T}\displaystyle\int_{-\frac{1}{2}T}^{+\frac{1}{2}T}\xi(t)\sin n\omega_0 t\,dt$

The varying formula for p'_0 has now disappeared.

If it is assumed that $\bar{r}_n = p'_n - jq'_n$ and $r_n = r^*_{-n}$ the following is obtained:

$$\xi(t) = \sum_{n=-\infty}^{n=+\infty}\bar{r}_n\exp(jn\omega_0 t)$$

For $\bar{r}_n$ the following is found from the formulae for p'_n and q'_n:

$$\bar{r}_n = \frac{1}{T}\int_{-\frac{1}{2}T}^{+\frac{1}{2}T}\xi(t)\exp(-jn\omega_0 t)\,dt$$

The amplitude of a component with frequency

$$nf_0 = \frac{n\omega_0}{2\pi} = \frac{n}{T}$$

then amounts to:

$$A_n = 2r_n$$

It is apparent that only one integral is sufficient to find the coefficients of the series, the constant term being found at the same time without any difficulty.

Now $\bar{r}_n$, which only differs from zero for whole values of n, can be made equal to :

$$\bar{r}_n = \frac{\mathbf{S}\left(\frac{n}{T}\right)}{T} = \mathbf{S}(f)\,.\,\Delta f$$

The character **S** used indicates the discontinuous character of **S** (f).

With its aid it is found that :

$$\left.\begin{aligned} \xi(t) &= \sum_{n=+\infty}^{n=-\infty} \mathbf{S}(f)\Delta f \exp(j\omega t) \\ \mathbf{S}(f) &= \int_{-\frac{1}{2}T}^{+\frac{1}{2}T} \xi(t)\exp(-j\omega t)\,\mathrm{d}t \end{aligned}\right\} \begin{aligned} &\mathbf{S}(f) = \overset{*}{\mathbf{S}}(-f) \\ &\omega = n\omega_0 = n.2\pi f_0 = \frac{2\pi n}{T} \end{aligned} \qquad \text{(I)}$$

For purposes of comparison the Fourier integral transformations, applied to time functions possessing a non-periodical character, are given (cf. Fig. 6).

$$\begin{aligned} \xi(t) &= \int_{-\infty}^{+\infty} S(f)\exp(j\omega t)\,\mathrm{d}f \\ S(f) &= \int_{-\infty}^{+\infty} \xi(t)\exp(-j\omega t)\,\mathrm{d}t \end{aligned} \qquad \text{(II)}$$

According to the conventional definition the following holds for the spectral intensity :

$$I(f) = S(f)\,.\,S^*(f) = S(f)\,.\,S(-f).$$

$2I(f)\,\Delta f$ is then the energy dissipated in a part of the frequency spectrum with a width of Δf by $\xi(t)$ in a resistance of 1 ohm, if $\xi(t)$ represents a current or a voltage.

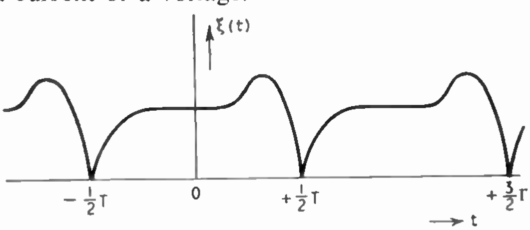


Fig. 5

From the fact that equations I and II show considerable analogies it can be concluded : $S(f)$ and $\mathbf{S}(f)$ are equal for whole values of n if, for the time function in Fig. 6, is chosen the time function of Fig. 5, inasmuch as the latter lies between $\pm\frac{1}{2}T$. From this it follows immediately that the course of the continuous function $S(f)$ is entirely determined by the values of $S(f)$ at n/T (n being a whole number). This property is part of the well-known sampling theorem.

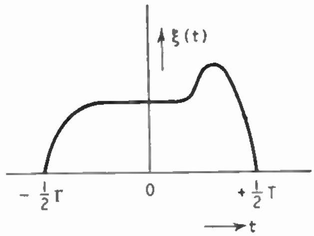


Fig. 6

From the spectrum $S(f)$ the single curve is contained. With a spectrum **S** (f) the single curve repeats itself with a period T. The shape of the curve is exactly the same for series and integral. (See e.g. Fig. 5 and 6.)

7. Linear Variable Networks

In the preceding examples the use of negative angular frequencies resulted in simpler expressions and a mathematically simpler discussion of the problems. The advantages of using negative angular frequencies is even more marked with linear variable networks.

Linear variable networks are networks in which the currents and voltages are controlled by homogeneous linear differential equations with coefficients that are dependent on time, but independent of current or voltage.

In these networks the principle of superposition holds and thus complex symbolism can be used. The remarkable thing is that with these networks the complex symbolism can only be applied if use is made of the concept of negative angular frequencies.

This will be shown for the series arrangement of a constant R and a linear variable L.

For L the following holds :

$$\begin{aligned} L &= L_0 + 2\,\hat{L}\cos\omega t \\ &= L_0 + \hat{L}\,[(\exp(j\omega t) + \exp(-j\omega t))] \end{aligned}$$

If voltage is applied of frequency p ($< \omega$) and if it is assumed that in the circuit only currents can flow of frequencies p and $\omega - p$, then only voltages will arise across the coil with frequencies p, $\omega - p$, $\omega + p$ and $2\omega - p$, which can be easily verified.

To obtain a complex expression for the circuit it must be assumed that :

$$i = I_0 \cos(pt + \varphi_0) + I_1 \cos(\omega t - pt - \varphi_1)$$
$$= \frac{I_0}{2}[\exp(jpt + j\varphi_0) + \exp(-jpt - j\varphi_0)] +$$
$$+ \frac{I_0}{2}[\exp(jpt - j\omega t + j\varphi_1) +$$
$$+ \exp(-jpt + j\omega t - j\varphi_1)].$$

If the expressions for L and i are inserted in

$$L\frac{di}{dt} + i\left(R + \frac{dL}{dt}\right) = e$$

in which

$$e = \tfrac{1}{2}E\,[\exp(jpt + j\psi) + \exp(-jpt - j\psi)]$$

then the following 4 equations are obtained :

$$I_0(R + jpL_0)\exp(j\varphi_0) + jpL\, I_1 \exp(j\varphi_1) -$$
$$- E\exp(j\psi) = 0$$

(coefficient of $\exp(jpt) = 0$)

$$I_1[R + j(p-\omega)L_0]\exp(j\varphi_1) +$$
$$+ j(p-\omega)\hat{L} I_0 \exp(j\varphi_0) = 0$$

(coefficient of $\exp j(p-\omega)\,t = 0$)

$$I_0(R - jpL_0)\exp(-j\varphi_0) - jp\hat{L}\, I_1 \exp(-j\varphi_1) -$$
$$- E\exp(-j\psi) = 0$$

(coefficient of $\exp(-jpt) = 0$)

$$I_1\ [R - j(p-\omega)L_0]\exp(-j\varphi_1) -$$
$$- j(p-\omega)\hat{L}\, I_0 \exp(-j\varphi_0) = 0$$

(coefficient of $\exp - j(p-\omega)t = 0$)

Obviously the equations that arise due to the coefficients of the other exponentials being assumed to be equal to 0 do not play any part.

In the first two equations only $I_0 \exp(j\varphi_0)$ and $I_1 \exp(j\varphi_1)$ appear as unknown factors and in the other two only the conjugate values of these two quantities appear as unknowns, the coefficients of the two pairs of equations being conjugate. From these four equations I_0, I_1, φ_0 and φ_1 can be determined and the problem thus solved.

The first two equations can also be represented as follows :—

$$\bar{I}_0(R + jpL_0) + jp\hat{L}\bar{I}_1 = \bar{E}$$
$$\bar{I}_0 j(p-\omega)\hat{L} + \bar{I}_1[R + j(p-\omega)L_0] = 0$$

if $\bar{I}_0 = A\ I_0 \exp(j\varphi_0)$, $\quad \bar{I}_1 = AI_1 \exp(j\varphi_1)$

and $\qquad E = A\ E\exp(j\psi)$

Thus a complex expression has been found for the circuit in question. In it the negative angular frequency is also used because $(p-\omega) < 0$.

An expression comparable to the latter would not have been obtained without the use of negative angular frequencies.

Thanks to the possibilities of applying the complex system of calculation, the insight into the manner of operation of many linear variable networks, which in practice very often occur as modulators, can be greatly extended.[1]

8. References

1. A. P. Bolle, " Application of complex symbolism to linear variable networks," *Tijdschrift van het Ned. Radio Genootschap* (Journal of the Netherlands Radio Society), **18,** No. 4, pp. 231-246, July 1953.
2. J. L. Bordewijk, " Intermodulation noise ", *Ibid.*, **17,** No. 5 and 6, pp. 261-279, Sept./Nov. 1952.

SOME FACTORS AFFECTING TRANSMITTING VALVE LIFE*

by

T. N. Basnett, B.Sc., B.E.†

SUMMARY

A knowledge of the principles underlying the operation of vacuum tubes enables a user to appreciate what performance may be expected and to understand to what extent factors under his control may influence the life obtained. The object of the paper is to discuss briefly the principal factors which affect valve life, and to do this it is convenient to study the various modes of failure. A number of good operating policies will become evident as the reasons for failure are studied, and these will be stated as the occasion arises. Emission and filament failures in valves with pure tungsten filaments, thoriated tungsten filaments and oxide-coated cathodes are dealt with first. A few remarks are then made about glass failures and operating temperatures. Mechanical failures are mentioned very briefly, and finally certain important factors relating to storage and handling of valves are dealt with.

1. Emission and Filament Failures

1.1. *Valves with Pure Tungsten Filaments*

In this case the emission is simply that from a pure metal. The operating temperature is so high that contaminants are quickly removed, and the phenomenon of poisoning so frequent with other types of emitter is rarely observed. Gas has practically no effect on the emission from pure tungsten, but on the other hand pure tungsten at very high temperatures has the ability to take up very large volumes of gas as it is released in the tube or as it enters the tube by way of infinitesimal leaks. Gas failures, therefore, are relatively uncommon and life is more usually terminated by burn-out of the filament.

The temperature at which the filament is operated is limited by evaporation of the tungsten. As the operating temperature is increased, so also is the rate of evaporation of the metal. As the metal is evaporated, the cross-sectional area of the filament is reduced. At those parts where the temperature is highest, the reduction in cross-sectional area is obviously greatest, and thus there is a tendency for a further increase in temperature and evaporation rate. This effect is counteracted by conductivity along the wire which tends to keep the temperature uniform. It is found, in general, that by the time the cross-sectional area of the filament has been reduced to about 90 per cent of its original value, the irregularities in diameter are such that temperature irregularities become serious. At the thinnest point the temperature rise rises, the evaporation rate increases, the wire " necks down," the tungsten at the centre of the reduced area melts and surface tension causes the tungsten to draw up into a droplet on one or other of the limbs of the filament, thus producing an open circuit.

The emission available from a filament is a function of surface temperature and surface area. Life is determined only by temperature and thus very long life could be obtained by producing a valve with a filament of large surface area operating at low temperature. Such a filament requires high heating power. A reduction of heating power requires a smaller, higher temperature filament, which obviously will give a shorter life. The size of a filament and thus the life it will give at full ratings has to be determined on economic grounds.

Figure 1 shows curves of the percentage life and emission against filament voltage expressed as a percentage of the normal value. These curves are based on normally chosen operating temperatures. It will be seen that life is greatly increased as filament voltage is reduced. In particular, at 90 per cent of normal voltage a four-fold increase in life is obtained, but of course the available emission is reduced to 45 per cent of normal.

It frequently happens that a valve chosen for a particular job has a greater reserve of

* Reprinted from *The Proceedings of the Institution of Radio Engineers, Australia*, Vol. 15, No. 7, July 1954. (Paper No. 333.)

† Assistant Manager, Valve Division, Standard Telephones & Cables Pty. Ltd., Sydney N.S.W.

U.D.C. 621.385.1, 621.396.61.

emission than necessary for satisfactory operation. If the valve has a pure tungsten filament it should be used at as low a filament voltage as is consistent with satisfactory operation. In Fig. 2 is shown the burnt-out filament structure of a type 4220Z water-cooled, pure

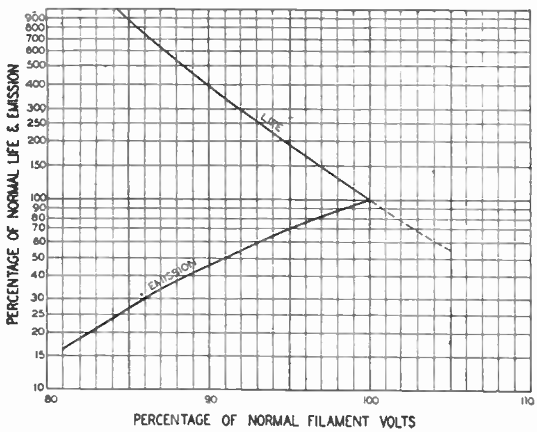


Fig. 1.—The percentage of normal life and emission plotted as a function of the percentage of normal filament voltage.

tungsten filament valve. It was operated in an equipment where there was a considerable reserve of emission so that the filament voltage could be reduced to 90 per cent of the normal rated value. The life obtained was 27,590 hours, which compares with 6,000 to 7,000 hours, usually obtained with valves of this type when operated at maximum ratings. Fig. 2*b* shows a greatly enlarged view of the burn-out. The necking-down of the filament and the formation of a droplet of tungsten on one of the limbs can be seen clearly.

The resistance of a hot tungsten filament is approximately fifteen times that of a cold filament and as a result care must be taken to limit the starting current when lighting up a filament, as otherwise the mechanical forces resulting from the high magnetic field associated with the filament current may distort or, in extreme cases, even shatter the filament. Normally a limit of about 70 per cent above the normal operating current is imposed as a maximum starting current. In addition, in order to reduce the effects of thermal shock on the filament and its supports, it is desirable to reduce to a minimum the number of times the filament is subjected to switching stresses. For the same reason, it is usual to recommend that during stand-by periods of up to two hours, the filament voltage be maintained at 80 per cent of normal. At this figure the evaporation rate is negligible, so that without appreciably affecting the life of the valve, the filament is protected from the effects of thermal shock.

As an example of the serious damage which can be caused by excessive filament starting current, Fig. 3 shows an x-ray photograph of the elements of a 4220Z valve which failed in an industrial equipment after 1500 hours. In the first place it was found that the starting current was over 100 per cent above normal and even more important, the filament was switched on and off as many as thirty times per day. Failure was caused by contact between filament and grid.

1.2. *Valves with Thoriated Tungsten Filaments*

In this case, emission depends on the maintenance on the surface of the filament of a layer of thorium atoms. The operating temperature is considerably above the melting point of thorium, but a unimolecular layer can be maintained provided the operating temperature is held within very rigid limits.

One condition necessary to the maintenance of this layer is that the diffusion rate of thorium from the body of the wire to the surface must be sufficiently high. The diffusion rate increases with temperature, but of course so does evaporation from the surface. In order to achieve the required diffusion rate at a suitable temperature, the outer layers of the wire are converted to tungsten carbide through which thorium diffuses more rapidly than through tungsten. Unfortunately tungsten carbide is extremely brittle so the layer converted to tungsten carbide contributes little to the mechanical strength of the filament. Usually about 20 to 30 per cent of the cross-sectional area of the wire is converted to carbide. The operating temperature of a thoriated tungsten filament is usually about 2000°K as compared with about 2500°K for a pure tungsten filament. Evaporation of tungsten is negligible at this temperature so that the mechanical life of a thoriated tungsten filament is potentially very great. Also the amount of thorium consumed during the life is a negligible part of that contained in the wire. Permanent failure of emission in this case usually occurs as a result of decarbonization, which takes place at a rate depending on both temperature and gas pressure.

If a thoriated tungsten filament is operated at too high a temperature, decarbonization takes place at a greatly accelerated rate and permanent, early loss of emission results. If the temperature is too low, the thorium layer may not be maintained and loss of emission result. In this case, however, reactivation may be possible, provided there is no secondary damage.

The widest tolerable limits of filament voltage for this type of valve are $\pm$ 5 per cent but improved life is usually obtained if closer control can be maintained.

1.3. *Valves with Oxide-Coated Cathodes*

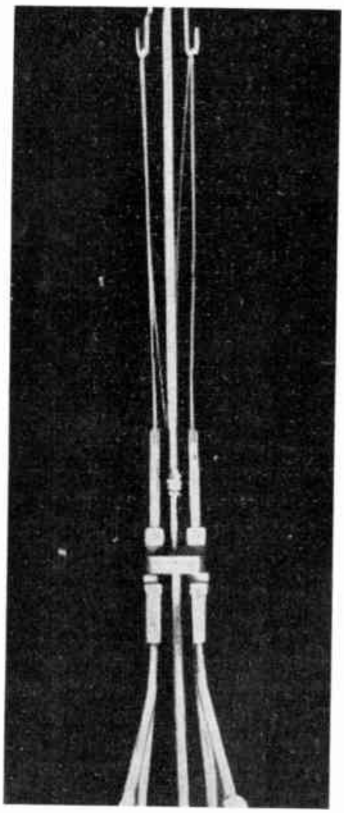
The mechanism of emission in this case is extremely complex and to obtain successful results a very delicate balance has to be maintained between various temperature sensitive processes. The permissible operating temperature range is quite narrow.

In general, the life of an oxide-coated cathode if operated within the correct temperature range is so great that very rarely indeed is it the primary cause of failure. Loss of emission frequently occurs, but almost invariably this is the result of an increase of gas pressure within the valve, of mechanical damage to the coating, or of poisoning by some contaminant.

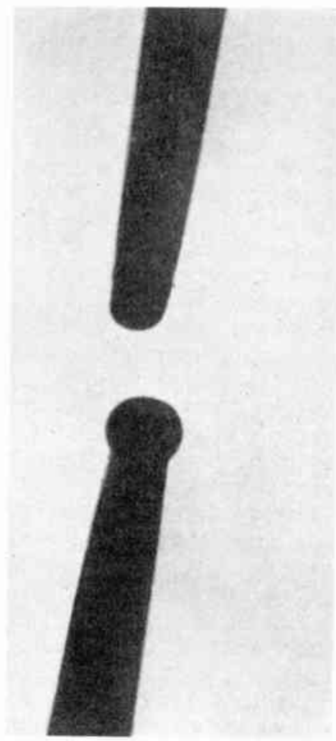
1.4. *General*

Filament and emission failures may be called fundamental limitations to the life of a valve. We have seen that valves with pure tungsten filaments frequently reach their full potential life, which is determined only by filament operating temperature and is terminated by mechanical failure. In cases where there is more reserve of emission available than necessary for satisfactory performance, great improvement in life may be obtained by reducing the filament voltage.

In the case of thoriated tungsten filaments,



(*a*)



(*b*)

Fig. 2 (a)—The burnt-out tungsten filament of a 4220Z water-cooled valve. (b) An enlarged view of the burnt-out filament.

potential lives are generally much longer, but to achieve the greater potential life the operating temperature must be maintained strictly within the correct range and gas pressure must remain low. It is frequently found that faults other than emission are the primary cause of failure with these valves.

The life of a valve with an oxide-coated cathode may be terminated by any one of a large number of faults. Provided the cathode is operated within the narrow range of temperature set by the manufacturer and provided gas pressure remains very low, cathode life is so long that it can be regarded as almost indefinite.

In the case of valves with thoriated tungsten filaments and oxide-coated cathodes, life is determined not only by filament temperature but also by operating gas pressure and anode voltage. Anode dissipation and glass temperature, therefore, as they affect gas pressure, are also important factors in determining valve life.

2. Glass Failures

All the valves referred to in this paper use glass envelopes with connections made to electrodes by way of metal to glass seals. Inevitably, the walls of any irregularly shaped, evacuated vessel are stressed, and the stresses are accentuated by the presence of the metal to glass seals. In operation, the leads sealed through the glass are heated by the currents flowing in them and the glass is heated by the radiation from the internal electrodes. Temperature gradients are thus set up when the valve is in operation, so that stress patterns are different in a hot valve from those in a cold valve.

Generally speaking it is not possible to produce a metal to glass seal that is completely stress free. What has to be done by the manufacturer is to produce a valve in which stresses in the various parts of the envelope are safe at room temperature and under operating conditions. Photo-elastic methods are used to reveal the position and magnitude of stresses and rigid inspection procedures are necessary to control production.

The bond between the metal and glass of a vacuum tight seal is a layer of oxide of the metal, which adheres very strongly to the metal and part of which dissolves in the glass. This bond is quite stable over very long periods provided the temperature is not excessive and provided it is not subject to electrolysis or attack by moisture or other active agents.

The most common modes of failure of the glass work in a valve are the development of a crack, or the destruction of seals by prolonged overheating, frequently in combination with electrolytic action. The conductivity of glass is normally very low indeed, but it increases rapidly with temperature. In operation, therefore, the glass acts as an electrolyte and the various leads, where they are sealed through the glass, act as electrodes. The products of electrolysis are deposited at the metal surface of the seals. It is, of course, to the most negative

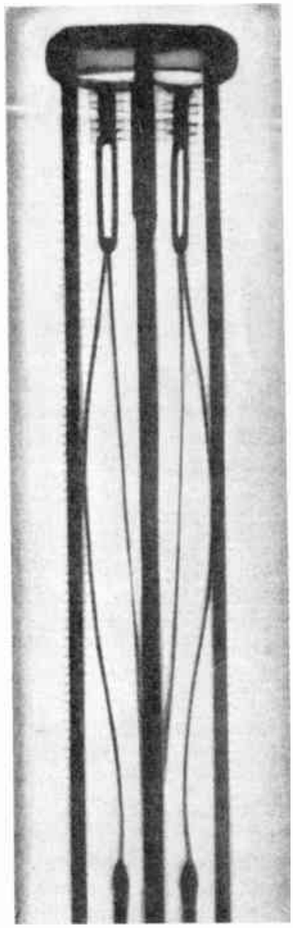


Fig. 3.—An X-ray photograph of the filament of a 4220Z valve distorted by frequent, excessive starting current.

electrode that any hydrogen or metal ions migrate and this is the seal usually most affected. If the process continues long enough (or fast enough) the metal to glass bond is attacked and high stresses may be built up. Sudden failure may result either by destruction of the vacuum tight seal or by development of a crack. In extreme cases failure may occur from this cause in a few hundred hours, but it is not uncommon to find serious attack on seals after 3,000–4,000 hours operation with barely adequate cooling.

3. Temperature of Operation

It cannot be too strongly stressed that the longest life will be obtained from a valve only if the utmost care is taken to keep the operating temperatures of the glass and leads to a minimum. The requirement for care in the establishment of operating conditions is obvious, as this determines working voltages and internal electrode temperatures. Having established suitable operating conditions, considerable variation of life is possible depending on the efficiency of the cooling systems employed. This is true for all types of valve, whether they be radiation cooled, forced air or water cooled.

An increase in glass temperature frequently involves an increase in stresses which may become dangerous, invariably accelerates electrolytic action, and in spite of the rigorous pumping to which valves are subjected, results in the liberation of gas from the envelope and hence raises the operating gas pressure. This last, as mentioned above, has a very marked influence on the life of valves with thoriated tungsten or oxide-coated cathodes. The maximum operating temperatures which might be considered safe for transmitting valves using hard glass envelopes are 250°C for those parts of the bulb away from any seals and 150°C for the glass adjacent to any seal.

We have found that crayons made with temperature sensitive pigments give reliable indications of temperature to an accuracy of about ± 10 per cent. They are certainly much simpler to use in the presence of strong r.f. fields than other methods.

4. Mechanical Failures

Space does not permit more than a passing reference to failures of this nature. A good deal of the manufacturer's effort is directed towards strengthening and improving mechanical construction of valves. Rigid control of materials and methods, together with careful inspection are essential, but the state of the art is still such that a valve is a fragile article which requires careful handling. The user must take every reasonable precaution to avoid undue shock and vibration.

5. Storage and Handling of Valves

Valves should be stored in a dry place, away from exposure to sunlight, which as well as having a weathering action on glass can cause unnecessary and even dangerous temperature cycling.

If a valve is warm it should never be subjected to thermal shock by being placed on a cool heat conducting surface.

Great care should always be exercised to avoid scratching the glass, as a scratch in a highly stressed area of the envelope may initiate a crack or even cause an implosion.

Valves tend to deteriorate if left on the shelf too long. The reason for this is two-fold. The attack on seals by moisture has been mentioned already, but infinitesimal leaks and liberation of gas internally, also result in an increase of gas pressure. When the valve is in operation these minute quantities of gas are continuously being taken up by the gettering action of electrodes and the operating pressure remains very low. On the other hand, if the valve is left for very long periods without use, when at last it is put into service, the amount of gas which has collected in the envelope may be sufficient to damage the filament or cathode permanently before it can be taken up. This danger may be countered by operating the valve at regular intervals no longer apart than say three months. Suitable conditions for taking up residual gas may be obtained by interchanging spare and working valves, so that the spare is given at least one hour of operation—preferably more. If this is not possible, simple ageing equipment is sometimes constructed so that suitable conditions can be applied.

Unfortunately there is no way in which damage caused by the effect of moisture on a seal can be repaired, A good deal of work is being done however on the use of protective coatings and moisture barrier packages to overcome this trouble and to make storage of valves, particularly in sub-tropical conditions less hazardous.