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. 2 FEBRUARY, 1955

THE ENGINEER IN MANAGEMENT

Increasing attention is now being given to the part which the engineer can play in management. A critical review of the subject was provided at the 1954 meeting of the British Association and one paper had the particularly provocative title of "The *Eclipse* of the Engineer in Management." In support of his thesis, the author compared the opportunities available to the engineer in various countries. The status of engineers, and the official recognition given to the profession, varies in almost every country, so that there is not really a basis for international comparison. This does not gainsay the fact that some countries give a greater measure of recognition to the engineer than others.

There is a similar difficulty in obtaining agreement on the interpretation of "manager." In Great Britain there are several bodies which aim at securing professional recognition for those engaged in management, but in its February 1955 issue, "The Times Review of Industry" poses the question, "Is Management a Profession?" The article suggests that the contemporary idea of a profession should have " . . . the following distinguishing requirements:—

(*a*) It should be founded upon a distinct body of knowledge requiring a measure of skill in its application.

(*b*) Members of the profession should have recognized responsibility for guarding and developing this body of knowledge and for the educational process associated with it.

(*c*) High standards of character, training, and proved competence, and the observance of an appropriate ethical code, should be required of members of the profession.

(*d*) Membership of the profession should be recognized by the public and the state as conferring a significant status based on integrity and technical competence."

The above requirements have long been well understood and practised by British professional engineering bodies; indeed, points (*a*) and (*b*) form the fundamental requirements for admission to membership of a recognized professional body.

The article further suggests that a convenient definition of management is the " . . . directing, organizing, and controlling the activities of a functional unit comprising more than one person."

Management is usually divided into various categories, e.g., sales managers, financial controllers, production engineers, etc.—whose ability is primarily based on specialized technical knowledge. Engineers wishing to enter the field of management usually require qualifications additional to their technological achievements. Thus, at the British Association meeting, there was some discussion on the desirability of including a paper on industrial administration as a compulsory subject in professional engineering examinations.

The importance of the subject is reflected in the approval given by the Department of Scientific and Industrial Research to a scheme of extra-mural courses for managers in industry which have been arranged by the University of Cambridge during the last three years. The third of these courses will be held in Madingley Hall from June 27th to July 22nd, 1955.

Such a course obviously has much to interest the radio and electronics engineer, whether he be engaged in manufacture in that field, or utilizing in some other sphere the products of the radio industry. In the latter connection, new techniques designed to improve industrial production frequently require management decisions based on experience and specialized technical knowledge.

NOTICES

Obituary

The Council of the Institution has expressed its sympathy with the relatives of the following members:—

Robert Hanford Banner (Associate Member) died, after a brief illness caused by a rheumatic heart, on January 3rd, aged 47 years. For the past 12 years he had been with Standard Telephones & Cables, Ltd., New Southgate, where he was in charge of a group of writers and technical library in the Radio Division. Mr. Banner joined the Institution as a Student in 1933 and transferred to Associate Membership in 1939. He had recently been invited to serve on the Library Committee.

* * *

Advice has only just been received of the death, on July 11th last, in a motor-car accident, of **Richard Gardiner Stuart** (Associate Member). Mr. Stuart was with the Department of Civil Aviation, Nyasaland and, was elected to membership of the Institution in 1948.

During the war he held appointments as Signals Officer with various R.A.F. Groups in the Far East, and was demobilized in 1946 holding the rank of Squadron Leader. From 1947 to 1948 he was Director of Communications in the Civil Aviation Department of the Government of India, and subsequently he joined the Department of Civil Aviation, Southern Rhodesia. He was 39 years of age.

* * *

Horace Swainsbury (Associate), who died suddenly on December 18th, was 56 years of age and at the time of his death was a civilian Technical Officer with Anti-aircraft Command, a position which he had held since 1948. Mr. Swainsbury served in the Royal Navy from 1912 to 1937, and also during the 1939-45 war when he was commissioned as Air Radio Officer and attained the rank of Lieutenant-commander; immediately before and after the war he was a Civilian Instructor in radio and radar with the R.A.F. He was elected an Associate of the Institution in 1953.

* * *

Howard Bardsley (Associate) died on February 5th, after a long and painful illness. He was 38 years of age and was elected to Associateship of the Institution in 1949.

From 1939 to 1946, Mr. Bardsley was an engineer with the British Broadcasting Corporation on the maintenance of high-powered transmitters. In 1949 he obtained an appointment as Communications Officer with the Iraq Petroleum Company, but had returned to his home near Rochdale in 1953 because of his illness.

Correspondence and Change of Address, etc.

In all correspondence with the Institution members are requested to include their grade of membership and reference number as other members may have similar initials and names. Changes of address should be sent, as soon as the change is made, to the Institution at 9 Bedford Square and such advice kept separate from other Institution correspondence.

It would be of great assistance in keeping the Institution's records up to date if members would advise the General Secretary of any decoration, award, or promotion which they receive, and of any change of employment made.

Jubilee of the Thermionic Valve

The 50th anniversary of the patenting, by Professor J. A. (later Sir Ambrose) Fleming, of a diode type of valve for the detection of radio waves has recently been celebrated in London.

An exhibition was held at University College, London, in the laboratories where Professor Fleming carried out his work during his 41 years' association with the College. As well as examples of the original valves made to Fleming's specification, the exhibition included very many documents relating to the invention and subsequent legal battles regarding its validity. The original equipment developed and used by Professor Fleming in his early experiments on this and many other aspects of light current engineering were shown side by side with research equipment used at present in the laboratory for work on microwaves.

A short account of Professor Fleming's life and work, particularly that relating to the "Fleming diode," has been written by Professor J. T. MacGregor-Morris,* who studied under Fleming and was later his personal assistant.

* *Inventor of the Valve—a Biography of Sir Ambrose Fleming.* The Television Society, London, 1954. 8½ in. × 5½ in., 134 pp., 10s. 0d.

INTERMODULATION NOISE IN V.H.F. MULTICHANNEL TELEPHONE SYSTEMS*

by

J. L. Slow, B.Sc.†

SUMMARY

This paper considers the noise which is introduced into the individual channels of a multichannel telephone system (frequency division multiplex) during transmission over a v.h.f. f.m. radio system due to various forms of distortion in the radio system. Formulae and curves are given for the computation of noise levels in systems of up to 60 channels and methods of measuring the parameters used in these formulae are discussed.

TABLE OF CONTENTS

1. Introduction

In recent years increasing use has been made of frequency-modulated radio systems in frequency bands above 30 Mc/s as alternatives to cable systems for the transmission of multichannel telephony signals, and a recent paper[1] discussed various engineering aspects of these radio systems for use in the v.h.f. bands. When a multichannel telephone signal consisting of a number of channels assembled in frequency divisions multiplex is transmitted over such a radio system, noise is introduced into the individual channels by the radio system. The main sources of this noise are:

(1) Thermal-type noise introduced mainly by the first stages of receivers at repeater and terminal stations. This noise varies with the r.f. input to the receiver, i.e. with propagation conditions, but is independent of the speech content of the transmitted multichannel signal.

(2) Intermodulation noise caused by non-linearity in the system, when speech signals in one or more channels give rise to harmonics or other distortion products which appear as unintelligible noise in other channels. Such noise is dependent on the frequency distribution and level of the speech signals in the transmitted multichannel signal.

The object of this paper is to analyse the various causes of the second type of noise, i.e. intermodulation noise and to describe means of determining the level of such noise to be expected in a given system.

* Manuscript first received June 2nd, 1954, and in final form October 19th, 1954. (Paper No. 299.)

† British Telecommunications Research Ltd., Taplow Court, Taplow, Bucks.

U.D.C. No. 621.396.822.1:621.396.43.029.62.

The paper deals in general with f.m. systems handling multichannel signals of 12 to 60 channels Larger numbers of channels are not usually transmitted by v.h.f. systems because of bandwidth limitations Much of the analysis is applicable to wideband u.h.f. and s.h.f. systems handling more than 60 channels, but in certain cases approximations are made which render the results invalid for these wideband systems. Furthermore, with the systems considered, it is shown that much of the information required to predict overall system performance is obtainable with test equipment which is relatively simple compared with transmission-measuring equipment necessary for system performance analysis in the case of, say, wideband microwave systems.

2. General

2.1. *Sources of Intermodulation Noise*

The main causes of intermodulation noise are:—

(1) Amplitude distortion in the terminal modulator, demodulator and amplifiers handling the baseband signals.

(2) Phase distortion, i.e. group delay distortion in circuits handling r.f. and i.f. frequency-modulated signals.

(3) Mismatches in aerial feeder systems or other long i.f. or r.f. connections, e.g. an i.f. cable connecting a receiver to a transmitter at a repeater. This produces a form of phase distortion.

These sources of intermodulation noise are the most important and are considered in detail in subsequent sections. There are, however, other forms of distortion which may produce intermodulation noise under certain conditions. These are:—

(4) Imperfect limiting at the demodulator. This form of distortion has been analysed by Libois[2] and by Albersheim and Schafer.[3] If any amplitude modulation at the baseband frequency which may be present is not effectively removed, distortion and consequent intermodulation noise is produced. This noise is similar in character to that produced by amplitude distortion (1 above). Normally amplitude modulation introduced by the modulator and transmission system should be small and should be adequately suppressed by well-designed limiters.

(5) Non-uniform amplitude/frequency characteristic in r.f. and i.f. circuits. This may, by affecting the relative sideband amplitudes, cause distortion of the modulating signal. Medhurst,[4] who has analysed the effect of phase and amplitude distortion for a single-tone modulating signal, points out that the amount of distortion produced by a particular amplitude characteristic depends on the phase distortion present, i.e. some amplitude characteristics are non-distorting in the absence of phase distortion. For multichannel signal modulation the intermodulation noise produced by this form of distortion has the same character and distribution as that produced by phase distortion. For the types of network normally used in v.h.f. systems it appears that this form of noise should be small compared with that due to phase distortion.

(6) Radio-frequency echo signals, i.e. signals arriving at the receiver via an indirect path. These are normally high-level short-delay echoes as distinct from feeder reflections which are low-level echoes having a relatively long delay. Albersheim and Schafer[3] give results of measurements on this type of noise together with a theoretical analysis.

2.2. *Multichannel Speech Signals*

Before considering the effect of the transmission characteristics on the multichannel signal the nature of the multichannel signal itself may be summarized. A frequency division multiplex signal in accordance with C.C.I.F. recommendations[5] takes the form of a spectrum of adjacent channels of 4 kc/s bandwidth, each containing a single audio channel translated and possibly inverted in frequency. Thus, in a given channel, signals only exist while speech is actually being transmitted, and the composite multichannel signal comprises a spectrum of signals whose frequency distribution and amplitude vary continuously with the number and volume of the speakers.

Throughout this paper, all levels are expressed in mW or db relative to 1 mW (dbm) at a point of zero relative level. At such a point in a system a single-channel test tone has a level of 1 mW. Speech power in a single channel varies from zero to a peak power of approximately + 7 dbm when peak limiting is used. The average speech power of a single subscriber is of the order of − 14 dbm. In a system with a large number of channels, at any given time during the busiest hour some channels are unoccupied, for example, while calls are being connected, while of those channels actually connected to subscribers engaged in a two-way conversation, on the average half only will be carrying speech. Thus the actual peak or r.m.s. level of the composite multichannel signal at any time is dependent on a large number of variables. From the point of view of circuit design sufficient information is given by a knowledge of two levels associated with the multichannel signal. These are:—

(*a*) The nominal peak level of the signal, i.e. the peak level which is exceeded for only a very small percentage of the time, say, 0·01 per cent. This enables the overload or "crash" level of line amplifiers, modulators and demodulators and the minimum bandwidth of r.f. and i.f. circuits to be determined.

(*b*) The r.m.s. level exceeded for a given percentage of the busiest hour. This enables the intermodulation noise power exceeded for this fraction of time to be determined.

It is evident that whatever method is used to determine intermodulation noise from the characteristics of the system, the accuracy of the result depends ultimately on the accuracy to which the latter level is known. Values have been given by a number of writers[6,7,8] determined by both statistical and analytical methods, but there is by no means complete agreement between all published results. Fig. 1 shows curves of peak and mean multichannel power which have been used by the writer. These are derived largely from the results of Holbrook and Dixon,[6] but using a slightly lower value of mean talker power, this lower value being the average of a number of published and unpublished results. Fig. 1a shows the mean multichannel power exceeded for 1 per cent. of the busy hour as a function of the number of channels in the system. The 1 per cent. value is given because C.C.I.F. recommendations are based on the noise exceeded for this fraction of time. Fig. 1b shows the corresponding peak values.

It should be noted that the term "channel deviation" is here used to denote the r.m.s. frequency deviation of the r.f. carrier produced by a single-channel test tone of 0 dbm at a zero level point. The r.m.s. and peak deviations due to the multichannel signals are then obtained from the voltage scale on the right of Fig. 1 which represents deviation relative to channel deviation.

2.3. *Weighting Factor*

Channel noise levels in this paper are given in terms of the actual or unweighted noise power in a telephone channel of 4 kc/s bandwidth. The psophometric noise level is approximately 3 db below this unweighted noise level, this factor relating the C.C.I.F. weighting curve to a flat noise distribution in a 4 kc/s band.

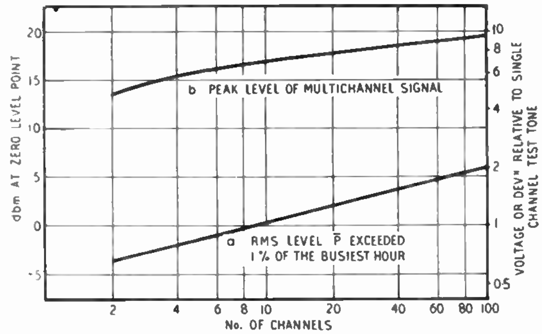


Fig. 1.—Multichannel signal levels.

3. Intermodulation Noise in Terminal Equipment

3.1. *Characteristics of Terminal Equipment*

The terminal equipment is here assumed to include the modulator, demodulator and all other circuits, e.g. line amplifiers, handling baseband frequencies.

For all types of modulators normally encountered the relation between frequency deviation f_d and input voltage v_1 may be represented adequately over the working range by a power series:

$$f_d = a_1 v_1 + b_1 v_1^2 + c_1 v_1^3 + - - - \ldots (1)$$

Similarly the output voltage v_2 from the demodulator may be represented:

$$v_2 = a_2 f_d + b_2 f_d^2 + c_2 f_d^3 + - - \ldots (2)$$

It follows that the output voltage from the demodulator is also related to the input voltage to the modulator by a power series:

$$v_2 = Av_1 + Bv_1^2 + Cv_1^3 + - - \qquad \ldots\ldots(3)$$

The relation between a_1, a_2, A, etc., may be found if required by approximating the series to a finite number of terms and substituting (1) in (2). In practice, the characteristics may usually be represented with sufficient accuracy by the first three terms of the series.

3.2. *Single-Tone Modulation*

If the modulator input is a single sinusoidal tone then v_1 may be represented by:

$$v_1 = V_1 \sin \omega_m t$$

Substituting in (3) and neglecting terms above the third order gives:

$$v_2 = AV_1 \sin \omega_m t + BV_1^2 \sin^2 \omega_m t + CV_1^3 \sin^3 \omega_m t \qquad \ldots\ldots(4)$$

It is well known that such a series may be expanded in terms of a fundamental frequency and harmonics. In this case the expression becomes:

$$v_2 = AV_1 \sin \omega_m t + \frac{BV_1^2}{2} \cos 2\omega_m t + \frac{CV_1^3}{4} \sin 3\omega_m t \qquad \ldots\ldots(5)$$

This neglects the d.c. component and very small fundamental frequency component contributed by 2nd- and 3rd-order distortion respectively.

It follows that:

$$\text{2nd-harmonic ratio (voltage ratio)} = \frac{BV_1}{2A} \qquad \ldots(6)$$

$$\text{3rd-harmonic ratio (voltage ratio)} = \frac{CV_1^2}{4A} \qquad \ldots(7)$$

3.3. *Two-Tone Modulation*

As shown in later sections a two-tone test has several advantages over a single-tone test in measuring linearity. If the two tones of frequency f_{m1} and f_{m2} have equal amplitudes the input v_1 may be expressed:

$$v_1 = V_1 (\sin \omega_{m1} t + \sin \omega_{m2} t)$$

where

$$\omega_{m1} = 2\pi f_{m1}, \text{ etc.}$$

Substituting in (3) gives:

$$v_2 = AV_1 (\sin \omega_{m1} t + \sin \omega_{m2} t) + BV_1^2 (\sin \omega_{m1} t + \sin \omega_{m2} t)^2 + V_1^3 (\sin \omega_{m1} t + \sin \omega_{m2} t)^3 \qquad \ldots\ldots(8)$$

Expansion shows that this expression is equivalent to a series of frequencies comprising the fundamental frequencies and harmonics and intermodulation products, having frequencies $(f_1 \pm f_2)$ in the case of 2nd-order distortion arising from the second term of (8) and frequencies $(2f_1 \pm f_2)$ and $(2f_2 \pm f_1)$ in the case of 3rd-order distortion arising from the third term of (8).

The frequencies and amplitudes of these intermodulation products may be summarized as follows:

Order of Dist.	Frequency of Product f_p	Amplitude
2nd	$f_{m1} \pm f_{m2}$	BV_1^2
3rd	$2f_{m1} \pm f_{m2}$, $2f_{m2} \pm f_{m1}$	$3CV_1^3$

The levels of these products relative to the fundamental tones may therefore be written down as:

$$\frac{\text{Voltage of 2nd-order product}}{\text{Voltage of single fundamental tone}} = \frac{BV_1}{A} \qquad \ldots(9)$$

$$\frac{\text{Voltage of 3rd-order product}}{\text{Voltage of single fundamental tone}} = \frac{3CV_1^2}{A} \qquad \ldots(10)$$

3.4. *Relation between Single- and Two-Tone Tests*

When distortion is measured by means of a two-tone test it is often an advantage to express the results in terms of the equivalent single-tone harmonic ratio.

From (6) and (9) it follows that the 2nd-harmonic ratio of one of the fundamental tones used in a two-tone test is one-half of, or 6 db less than, the ratio of the intermodulation product to one fundamental tone. Thus if two tones each of level 0 dbm result in a $(f_1 - f_2)$ product of $-$ 60 dbm, the harmonic ratio of a single 0 dbm test tone is $-$ 66 db.

Similarly from (7) and (10) the third harmonic ratios of one of the fundamental tones is one-twelfth of, or 21·5 db less than the ratio of the intermodulation product to one fundamental tone.

3.5. *Intermodulation Noise Due to Multichannel Signal*

When a composite signal containing a large number of different frequencies is applied to the modulator, as in the case of a multichannel telephone signal, intermodulation noise is produced, firstly, by harmonics of signals in one channel falling in another, and secondly, by intermodulation products of the form $(f_1 \pm f_2)$, $(2f_1 \pm f_2)$, $(f_1 + f_2 + f_3)$, etc. In any one channel the total intermodulation power is made up of many such products.

Expressions for the level of intermodulation noise power in a single channel of a multichannel system, when transmission takes place through a network having a power series amplitude response, have been derived by several writers.[7, 8, 9, 14] The results given by Brockbank and Wass[8] which are used in this paper may in general be expressed :

$$N_n = \frac{1}{M} y_n (K_n H_n P^n) \quad \ldots\ldots\ldots (11)$$

where

- N_n = nth-order intermodulation noise power (unweighted) in a single channel expressed in mW at a zero level point.
- M = number of channels.
- y_n = distribution factor dependent on the position of the channel in the baseband spectrum and on the order of distortion.
- K_n = constant for nth-order distortion.
- H_n = nth-harmonic ratio of a single-channel test tone expressed as a power ratio, i.e. $\dfrac{\text{harmonic power}}{\text{test tone power}}$
- P = total multichannel speech power in mW at zero level point, usually expressed in terms of mean power exceeded for a given percentage of the time (see section 2.2).

The bracketed part of (11) is the total nth-order intermodulation noise power and the factor $\frac{1}{M} y_n$ is the fraction falling in a particular channel.

Using decibel notation (11) may be rewritten as follows. Here $\overline{N}_n = 10 \log_{10} N_n$, etc.

$$\overline{N}_n = \overline{H}_n + n\overline{P} + C_n \quad \ldots\ldots\ldots\ldots (12)$$

where

- $\overline{N}_n$ = nth-order intermodulation noise (unweighted) in a single channel in dbm at a zero level point.
- $\overline{H}_n$ = nth-order harmonic ratio of single-channel test tone in db. (This is a negative quantity.)
- $\overline{P}$ = total multichannel speech power in dbm at zero level point (see section 2.2).
- C_n = constant for nth-order distortion $= 10 \log_{10} \dfrac{y_n K_n}{M}$.

Values of C_2 and C_3 (i.e. C_n for 2nd- and 3rd-order distortion) have been computed from the results of Brockbank and Wass for systems of 12 to 60 channels. These values are shown in Fig. 2. In the case of odd-order distortion two values of y_n are given by Brockbank and Wass corresponding to two groups of products which may add differently, i.e. on a voltage or a power basis when two or more networks (line repeaters, modulator/demodulator combinations, etc.) are cascaded. For the present, however, only systems with one modulator and demodulator are considered and since the two groups of products add together on a power basis the two values of y_n may be added to give a single value. This value has been used in the computation of the curves of 3rd-order distortion in Fig. 2.

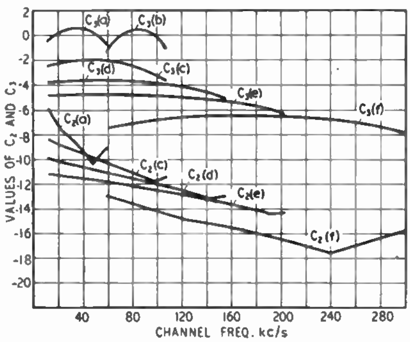


Fig. 2.—Values of C_2 and C_3 for typical multichannel systems.

(Derived from Brockbank and Wass,[8] Figs. 1 and 2.)

(a) 12 channels 12–60 kc/s
(b) 12 channels 60–108 kc/s
(c) 24 channels 12–108 kc/s
(d) 36 channels 12–156 kc/s
(e) 48 channels 12–204 kc/s
(f) 60 channels 60–300 kc/s

N.B. 2nd-order intermodulation noise is zero in 60–108 kc/s system.

In practice, therefore, the intermodulation noise due to modulator/demodulator distortion may be obtained by measuring the harmonic ratio H_n of a single-channel test tone, obtaining values of $\overline{P}$ and C_n from Figs. 1a and 2 respectively, and substituting in equation (12). This is dealt with further in section 7.

4. Intermodulation Noise due to Phase Distortion

4.1. *Effect of Phase Distortion on a F.M. Signal*

When a frequency-modulated signal is trans-transmitted through a network, for example an i.f. amplifier, in which the time taken for transmission is not equal for all frequencies within the transmitted bandwidth, i.e. which introduces group delay or phase distortion, the result is a distortion of the modulating signal. Thus, if the modulation is a multichannel telephone signal, intermodulation noise is caused.

Referring to Fig. 3, if the network introduces no distortion the phase/frequency characteristic is a straight line representing a constant group delay $d\phi/d\omega$. If phase distortion is present the phase/frequency characteristic departs from this linear function by an amount ϕ_d which is itself a function of frequency. Usually the phase/frequency characteristic of the network over the working bandwidth may be represented by a power series:

$$\phi = \phi_0 + A_p f_d + B_p f_d^2 + C_p f_d^3 + --- \quad (13)$$

where ϕ_0, A_p, B_p, etc., are constants and f_d is the deviation of the carrier frequency from the mean frequency f_0 so that:

$$f_d = f - f_0$$

As in the case of modulator/demodulator distortion terms above third order may usually be neglected and the phase distortion ϕ_d is then given by:

$$\phi_d = B_p f_d^2 + C_p f_d^3 \quad \ldots\ldots\ldots\ldots (14)$$

The curve of Fig. 3 is the steady-state phase/frequency characteristic of the network, and therefore a rigorous analysis of its effect on a frequency-modulated signal involves resolving the signal into carrier and sidebands, determining the phase distortion of each such component signal and thence re-combining to determine the distortion of the composite signal. A much simpler analysis may, however, be carried out using a "quasi-stationary"[10] treatment in which it is assumed that Fig. 3 not only represents the steady-state characteristic but also the phase at any instant of time when the frequency is varying rapidly, as in the case of a frequency-modulated signal. It appears that such an approximation should not result in any serious error when applied to the distortion of multichannel f.m. signals and this method is therefore used in this paper.

4.2. *Single-Tone Modulation*

In section 3.2 it was stated that in the case of a single tone, the 2nd- and higher-order terms of the amplitude characteristic of the modulator/demodulator combination give rise to amplitude variations at multiples of the fundamental frequency, and these appear as harmonics. Similarly, when the carrier frequency is modulated in a sinusoidal manner the phase distortion, as illustrated in the characteristic of Fig. 3, gives rise to phase variations at multiples of the fundamental frequency, and since such superimposed phase modulation may be represented by an equivalent frequency modulation, the carrier after passing through the distorting network is frequency modulated by harmonics of the fundamental tone in addition to the tone itself.

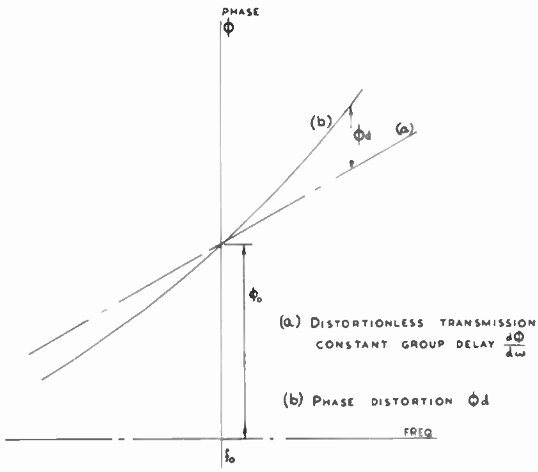


Fig. 3.—Phase distortion in r.f. and i.f. circuits.

If the frequency-modulated signal is represented by:

$$f_d = \Delta f \sin \omega_m t \quad \ldots\ldots\ldots\ldots (15)$$

where Δf = peak deviation from mean carrier frequency

$\omega_m = 2\pi$ times the modulating frequency f_m

Then substituting in (14) gives:

$$\phi_d = B_p \Delta f^2 \sin^2 \omega_m t + C_p \Delta f^3 \sin^3 \omega_m t \quad \ldots\ldots\ldots (16)$$

Expanding the powers of sin ωt in terms of harmonics gives:

$$\phi_d = \tfrac{1}{2} B_p \Delta f^2 (1 - \cos 2\omega_m t) + \\ + \tfrac{1}{4} C_p \Delta f^3 (\sin \omega_m t - \sin 3\omega_m t) \\ = \tfrac{1}{2} B_p \Delta f^2 + \tfrac{1}{4} C_p \Delta f^3 \sin \omega_m t - \\ - \tfrac{1}{2} B_p \Delta f^2 \cos 2\omega_m t - \tfrac{1}{4} C_p \Delta f^3 \sin 3\,\omega_m t \quad ..(17)$$

The first term represents merely a constant-phase shift of the carrier for a given value of Δf. The second, third and fourth terms represent phase modulation at the fundamental frequency, and the second and third harmonics respectively. Since $\omega = d\phi/dt$, by differentiating (17), the equivalent instantaneous frequency deviation f_d' produced by the distortion is obtained.

$$f_d' = \frac{1}{2\pi} \frac{d\phi_d}{dt} \\ = -f_m \,.\, \tfrac{1}{2} C_p \Delta f^3 \cos \omega_m t + \\ + 2f_m \,.\, \tfrac{1}{2} B_p \Delta f^2 \sin 2\omega_m t + \\ + 3f_m \,.\, \tfrac{1}{4} C_p \Delta f^3 \cos 3\omega_m t \quad ......(18)$$

The total instantaneous frequency deviation f_{d0} present after transmission through the network is therefore given by adding the initial frequency deviation f_d to the deviation f_d' arising from distortion:

$$f_{d0} = f_d + f_d' \\ = \Delta f \sin \omega_m t - f_m \,.\, \tfrac{1}{2} C_p \Delta f^3 \cos \omega_m t + \\ + 2f_m \,.\, \tfrac{1}{2} B_p \Delta f^2 \sin 2\omega_m t + \\ + 3f_m \,.\, \tfrac{1}{4} C_p \Delta f^3 \cos 3\omega_m t \quad ......(19)$$

The second term represents a small amount of modulation at the fundamental frequency in phase quadrature to the original modulation and for tolerably low distortion this term may be neglected. The harmonic voltage ratios of the modulating frequency are therefore given by:

$$\text{2nd-harmonic ratio} = 2f_m \,.\, \tfrac{1}{2} B_p \Delta f \quad ....(20)$$

$$\text{3rd-harmonic ratio} = 3f_m \,.\, \tfrac{1}{4} C_p \Delta f^2 \quad ..(21)$$

It will be noted that these are analogous to (6) and (7) where $A = 1$ but with the additional factor nf_m, i.e. the harmonic margin is proportional to the frequency of the harmonic.

4.3. *Two-Tone Modulation*

The case of modulation by two tones f_{m1} and f_{m2} of equal amplitude may be analysed in a similar manner to that used in 4.2. In this case:

$$f_d = \Delta f \sin \omega_{m1} t + \Delta f \sin \omega_{m2} t \quad ....(22)$$

Substituting in (14) gives

$$\phi_d = B_p \Delta f^2 (\sin \omega_{m1} t + \sin \omega_{m2} t)^2 \\ + C_p \Delta f^3 (\sin \omega_{m1} t + \sin \omega_{m2} t)^3 \quad ..(23)$$

This is analogous to the 2nd- and 3rd-order distortion terms of eqn. (8) and may be expanded in terms of harmonics and intermodulation products in the same way. In this case these intermodulation products represent phase modulation of the carrier at frequencies of $(f_{m1} \pm f_{m2})$ in the case of 2nd-order phase distortion and $(2f_m \pm f_2)$ or $(2f_2 \pm f_1)$ in the case of 3rd-order phase distortion. By analogy with the results of 3.3:

Phase deviation (i.e. amplitude of phase modulation) due to 2nd-order products

$$= B_p \Delta f^2 \quad ....(24)$$

Phase deviation (i.e. amplitude of phase modulation) due to 3rd-order products

$$= 3C_p \Delta f^3 \quad ....(25)$$

For a sinusoidal phase deviation ϕ at a modulating frequency f_m the equivalent frequency deviation is $f_m\phi$ c/s and therefore if the frequency of a particular intermodulation product is f_p the equivalent frequency modulation at f_p is given by:

Freq. deviation of 2nd-order products

$$= f_p B_p \Delta f^2 \quad ....(26)$$

Freq. deviation of 3rd-order products

$$= f_p 3C_p \Delta f^3 \quad ....(27)$$

The ratios of the intermodulation products to the fundamental tones are then given from (26) and (27):

$$\frac{\text{Amplitude of 2nd-order products}}{\text{Amplitude of single fundamental tone}} = f_p B_p \Delta f \quad ....(28)$$

$$\frac{\text{Amplitude of 3rd-order products}}{\text{Amplitude of single fundamental tone}} = f_p 3C_p \Delta f^2 \quad ....(29)$$

The amplitude of the intermodulation product is therefore proportional to f_p, the frequency of the product, but independent of the frequency of the fundamental tones, thus, for example, for two frequencies of 20 and 30 kc/s the $(f_1 + f_2)$ product has an amplitude five times or 14 db greater than the $(f_1 - f_2)$ product.

4.4. *Relation Between Single- and Two-Tone Tests*

By combining (20) and (28)

$$\frac{\text{Amplitude of 2nd harmonic of single tone}}{\text{Amplitude of 2nd-order product of two tones}} = \frac{f_m}{f_p} \quad \ldots\ldots\ldots\ldots\ldots(30)$$

If, therefore, the ratio of the 2nd-order intermodulation product f_p relative to a single tone is measured, the 2nd harmonic ratio of a single tone having the amplitude of one of the tones of the two-tone test is given by multiplying the ratio by $\frac{f_m}{f_p}$ where f_m is the single-tone frequency.

Similarly by combining (21) and (29)

$$\frac{\text{Amplitude of 3rd harmonic of single tone}}{\text{Amplitude of 3rd-order product of two tones}} = \frac{1}{4}\frac{f_m}{f_p} \quad \ldots\ldots\ldots\ldots\ldots(31)$$

Thus the third harmonic ratio of a single tone of frequency f_m is given by multiplying the two-tone $\frac{\text{intermodulation product}}{\text{single tone}}$ ratio by $\frac{1}{4}\frac{f_m}{f_p}$.

4.5. *Intermodulation Noise due to Multichannel Signal*

When a carrier, frequency modulated by a multichannel signal, is transmitted through a phase-distorting network, harmonics and intermodulation products are produced which, assuming a power series law, are analogous to those produced in modulator/demodulator distortion as described in section 3.5. However, by the same reasoning as used in sections 4.3 and 4.4, it is evident that these products take the form of unwanted phase modulation, and that the equivalent frequency modulation of each product is given by multiplying the phase deviation by the frequency of the product. Therefore the relative level of the products falling in the various speech channels may be obtained by multiplying the Brockbank and Wass power-distribution curves used in section 3.5 by the square of the frequency. This extension of the Brockbank and Wass analysis for the case of phase distortion of frequency-modulated signals was proposed by Gladwin.[11] The validity of this argument may perhaps be more evident by referring to the following comparative table of 2nd- and 3rd-order distortion products. In this table the bracketed figures refer to the appropriate equations. In the right-hand column f_p replaces $2f_m$ or $3f_m$ in (20) and (21).

From the examples shown it may be seen that the ratio between the level of a single-tone harmonic and a two-tone intermodulation product of the same order is the same in both the modulator/demodulator distortion and phase-distortion cases, provided that in the latter case the product frequency f_p is the same in both the single- and two-tone cases. It may be shown that this is also true in the case of multitone products other than the ones quoted, e.g. $(f_{m1} + f_{m2} + f_{m3})$ in the case of 3rd-order distortion. Generally speaking then, over a narrow band of frequencies in the baseband spectrum, e.g. a single telephone channel over which the change in the value of a particular product due to variation of f_p is negligible, the ratio between a single-tone harmonic and all intermodulation products of the same order falling in the channel is the same in both the modulator/demodulator and phase-distortion cases. Since, therefore, equation (11) in effect relates the single test tone harmonic ratios to the sum of intermodulation products due to distortion of the same order, in the case of modulator/demodulator distortion it may also be used to obtain total intermodulation noise in a given channel due to phase distortion, with the important provision that the harmonic ratio H_n is measured using a fundamental tone of such a frequency that the harmonic falls in the telephone channel being considered. For example, if the 2nd-order noise power falling in a 104-108 kc/s channel (106 kc/s mean) is required, the value of

		Amplitude of Distortion Product Relative to Single Tone	
Dist Order	Product Freq f_p	Mod/Dem Dist $v_2 = Av_1 + Bv_1^2 + Cv_1^3$	Phase Dist $\phi = A_p f_d + B_p f_d^2 + C_p f_d^3$
2nd	$2f_m$	$\frac{BV_1}{2A}$ (6)	$f_p \frac{B_p \Delta f}{2}$ (20)
2nd	$f_{m1} \pm f_{m2}$	$\frac{BV_1}{A}$ (9)	$f_p B_p \Delta f$ (28)
3rd	$3f_m$	$\frac{CV_1^2}{4A}$ (7)	$f_p \frac{C_p \Delta f^2}{4}$ (21)
3rd	$2f_{m1} \pm f_{m2}$ or $2f_{m2} \pm f_{m1}$	$\frac{3CV_1^2}{A}$ (10)	$f_p 3C_p \Delta f^2$ (29)

N.B. V_1 = peak amplitude of single tone
Δf = peak deviation produced by single tone.

H_2 to be used in equation (11) is the 2nd harmonic ratio of a single-channel test tone of frequency $106/2 = 53$ kc/s.

The harmonic ratio H_n of a test tone is proportional to the square of the frequency of the tone (from equation (20) and (21)) when H_n is expressed as a power ratio. If the harmonic ratio is measured at some convenient test frequency f_{mt}, which will normally be as high as possible to give the maximum level of harmonic, the harmonic ratio of a test tone whose nth harmonic falls in the telephone channel of centre frequency f_{mc} is $H_n\left(\frac{f_{mc}}{nf_{mt}}\right)^2$

and it is this term which replaces H_n in equation (11) when dealing with phase distortion.

In the phase-distortion case, then, equation (11) may be rewritten:

$$N_n = \frac{1}{M} y_n \, (K_n H_n P^n) \frac{f_{mc}^2}{n^2 f_{mt}^2} \quad \ldots\ldots(32)$$

where notation is the same as in (11) with the additional terms:

f_{mc} = mid-frequency of telephone channel considered

f_{mt} = frequency of test tone used for measurement of H_n

n = order of distortion.

The values of y_n given by Brockbank and Wass are still valid but they no longer represent the distribution of intermodulation noise through the baseband. As stated earlier this relative power distribution is obtained by multiplying the original curves by a factor proportional to f_{mc}^2; in this case the factor $\frac{f_{mc}^2}{n^2 f_{mt}^2}$.

Broadly speaking, the variation of intermodulation power over the baseband in the case of modulator/demodulator distortion is not normally very great, and therefore the addition of the f_{mc}^2 factor in the phase-distortion case results in a power-distribution curve approaching a square law, i.e. a triangular voltage distribution. This being similar to the basic thermal-noise distribution in a frequency-modulation system, it means that provided modulator/demodulator intermodulation noise is kept low compared with phase distortion intermodulation noise, a degree of pre-emphasis may often be used with advantage in improving the overall signal/noise ratio of the system.

Equation (32) may conveniently be rewritten in decibel form:

$$\overline{N_n} = \overline{H_n} + n\overline{P} + C_n + D_n \quad \ldots\ldots\ldots(33)$$

Where notation is the same as in (12) with the additional term:

$$D_n = 20 \log_{10} \frac{f_{mc}}{nf_{mt}}$$

To summarize, the intermodulation noise power in a given channel due to phase distortion may be obtained by measuring the nth harmonic H_n of a single-channel test tone, obtaining values of $\overline{P}$ and C_n from Figs. 1a and 2 respectively, computing the factor D_n to take into account the relation between channel frequency and test-tone frequency and substituting in (33). This procedure is dealt with in more detail in section 7.

5. Intermodulation Noise due to Feeder Mismatch

5.1. *Phase Characteristics of Mismatched Feeder*

If a long length of transmission line carrying a frequency-modulated multichannel signal is mismatched at the ends, intermodulation noise is produced. Such long connections commonly exist in the form of feeders connecting aerials to transmitters or receivers but may also be present in some types of equipment in the form of an i.f. link between receiver and transmitter at a repeater. The means whereby the distortion and consequent intermodulation takes place may be described briefly as follows.

If in Fig. 4a Z_1 and Z_2 represent the terminations of reflection coefficients ρ_1 and ρ_2 respectively, a signal of amplitude V_t transmitted from Z_1 to Z_2, after reflection from Z_2 and Z_1, gives rise to an echo signal V_e at Z_2 of amplitude $\rho_1\rho_2 \; 10^{-\alpha/10}$ where α is the attenuation of the single length of cable in db. The phase of V_e is dependent on the real and imaginary coefficients of ρ_1 and ρ_2 and also on the time taken for the signal to traverse the double length of feeder. If the frequency of the signal is varied, the electrical length of the cable varies and the phase of V_e relative to V_t changes accordingly. With a long cable a relatively small change of frequency will cause V_e to describe the locus shown in Fig. 4b and the phase of the resultant signal (for tolerably small reflections) will vary in a sinusoidal manner as shown in Fig. 4c. Corresponding group delay and amplitude characteristics are given by Friis.[12] This variation may be considered as a phase distortion superimposed on the linear phase characteristic representing the constant group delay of the feeder.

For the echo-signal vector V_e to describe one rotation relative to V_t the electrical length of the echo-path length $2l$ must change by λ and the frequency must therefore change by $f_1 . \frac{\lambda}{2l}$ where f_1 is the original frequency.

Now $f_1 . \frac{\lambda}{2l} = \frac{v}{2l}$ where v is velocity of propagation.

Therefore the frequency spacing f_s of the maxima of the phase-distortion characteristic is given by:

$$f_s = \frac{v}{2l} \text{ c/s} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (34)$$

Also from Fig. 4b the maximum phase distortion $\hat{\phi}_d$ is given by:

$$\hat{\phi}_d = \frac{V_e}{V_t}$$
$$= \rho_1 \, \rho_2 \, 10^{-\alpha/10} \qquad \ldots\ldots\ldots\ldots\ldots\ldots (35)$$

It should be noted that the foregoing assumes that the reflection coefficients remain constant over the band of frequencies considered. In practice this is not normally the case, and further distortion of the phase/frequency characteristic may be present. In order to simplify the analysis, however, it is assumed that the reflection coefficient is constant and equal to the mean value over the band of frequencies considered.

A numerical example will serve to illustrate the order of magnitude of f_s. In the case of a typical semi-air-spaced coaxial cable, v has a value of $2{\cdot}9 \times 10^8$ m/sec and f_s would then be 3·2 Mc/s for a 150-ft feeder. When the deviation of a frequency-modulated signal is several Mc/s, as may be the case in a wide-band microwave system handling several hundred telephone channels, the carrier frequency sweeps through several cycles of the sinusoidal phase-distortion characteristic and it is evident that in this case high-order distortion products will be produced and that the distortion cannot be analysed in terms of a power series law as is the case in section 4. If, however, the deviation or feeder length is small enough so that the carrier frequency sweeps over less than ½ cycle of the phase-distortion characteristic the distortion may be approximated to 2nd-order distortion when the frequency is centred at, say, f_A on Fig. 4c, and 3rd-order distortion when the frequency is centred at f_B, or, of course, a combination of 2nd- and 3rd-order for intermediate points. The distortion of single-tone and multichannel signals due to feeder reflections has been analysed by a number of writers[2, 3, 9, 13, 15, 16] for both the general case and the case approximating to a power series law.

In the case of v.h.f. multichannel systems handling up to approximately 60 channels and with feeders not exceeding 250 ft in length the deviation is normally such that this approximation to a power series law is valid and in fact it may be shown that the intermodulation noise obtained by this method differs from that obtained by the rigorous analysis by less than 1 db.

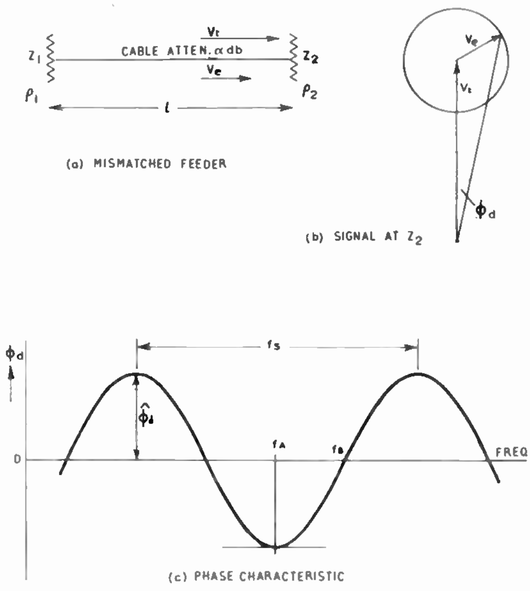


Fig. 4.—Phase distortion due to feeder mismatch.

The approximate analysis is here carried out for the cases of maximum 2nd- and maximum 3rd-order distortion, i.e. centre frequencies at f_A and f_B.

In the case of a carrier frequency centred at f_A

$$\phi_d = -\hat{\phi}_d \cos \frac{2\pi f_d}{f_s}$$

For small values of $\frac{f_d}{f_s}$, as is assumed here,

$$\phi_d \simeq -\hat{\phi}_d\left(1-\frac{1}{2}\frac{4\pi^2 f_d^2}{f_s^2}\right)$$

$$= -\phi_d + \hat{\phi}_d\frac{2\pi^2 f_d^2}{f_s^2}$$

The first term may be neglected as it represents only a constant phase shift and ϕ_d may be written:

$$\phi_d = \hat{\phi}_d\frac{2\pi^2 f_d^2}{f_s^2}$$

Substituting for f_s and $\hat{\phi}_d$ from (34) and (35)

$$\phi_d \text{ (Max 2nd order)} = \frac{8\pi^2\rho_1\rho_2\,10^{-\alpha/10}\,l^2}{v^2} f_d^2 \quad (36)$$

Similarly when the frequency is centred at f_B

$$\phi_d = \hat{\phi}_d \sin\frac{2\pi f_d}{f_s}$$

$$\simeq \hat{\phi}_d\left(\frac{2\pi f_d}{f_s} - \frac{8\pi^3 f_d^3}{6 f_s^3}\right)$$

Here the term $\hat{\phi}_d\frac{2\pi f_d}{f_s}$ represents a linear phase characteristic, i.e. a small additional constant group delay, and from the point of view of distortion may be neglected. ϕ_d may then be written:

$$\phi_d = -\hat{\phi}_d\frac{4\pi^3 f_d^3}{3 f_s^3}$$

Substituting for f_s and $\hat{\phi}_d$

$$\phi_d \text{ (Max 3rd order)} = \frac{32\pi^3\rho_1\rho_2\,10^{-\alpha/10}l^3}{3v^3} f_d^3 \quad (37)$$

The coefficients $\frac{8\pi^2\rho_1\rho_2\,10^{-\alpha/10}\,l^2}{v^2}$ and $\frac{32\pi^3\rho_1\rho_2\,10^{-\alpha/10}\,l^3}{3v^3}$ of (36) and (37) are therefore equivalent to the constants B_p and C_p respectively in equation (13). The distortion products due to single-tone, two-tone and multichannel modulation may then be obtained by substituting these coefficients in the appropriate equations of section 4.

5.2. *Single-Tone Modulation*

From equations (36) and (20)

Maximum 2nd harmonic ratio of single tone

$$= \frac{8\pi^2\rho_1\rho_2\,10^{-\alpha/10}\,l^2}{v^2} f_m\Delta f \quad \ldots\ldots\ldots(38)$$

From equations (37) and (21)

Maximum 3rd harmonic ratio of single tone

$$= \frac{8\pi^3\rho_1\rho_2\,10^{-\alpha/10}\,l^3}{v^3} f_m\Delta f^2 \ldots\ldots\ldots(39)$$

These are the maximum values. The actual amplitude of a particular harmonic varies sinusoidally with the length of the feeder and small variations of feeder length about a nominal value may result in a considerable change in the harmonic level. It is not therefore practical to attempt to "tune" the feeder for minimum 2nd or 3rd harmonic distortion. The average value of distortion for a random variation of feeder length about a nominal value is given by multiplying the above values by $1/\sqrt{2}$

5.3. *Two-Tone Modulation*

From equations (36) and (28)

$$\frac{\text{Maximum amplitude of 2nd-order products}}{\text{Amplitude of single fundamental tone}}$$

$$= \frac{8\pi^2\rho_1\rho_2\,10^{-\alpha/10}\,l^2}{v^2} f_p\Delta f \quad \ldots\ldots\ldots(40)$$

From equations (37) and (29)

$$\frac{\text{Maximum amplitude of 3rd-order products}}{\text{Amplitude of single fundamental tone}}$$

$$= \frac{32\pi^3\rho_1\rho_2\,10^{-\alpha/10}\,l^3}{v^3} f_p\Delta f^2 \quad \ldots\ldots(41)$$

As in 5.2 the average value is given by multiplying the above by $1/\sqrt{2}$

5.4. *Intermodulation Noise Due to Multichannel Signal*

If the "channel deviation" as defined in section 2 is denoted by f_{d0} then the maximum 2nd and 3rd harmonic ratios of a single-channel test tone due to feeder mismatch are given by substituting $\sqrt{2}f_{d0}$ for Δf in (38) and (39), f_{d0} being an r.m.s. deviation, and writing $f_m = f_{mt}$, the test-tone frequency. By converting these equations to decibel form the appropriate values of $\overline{H}_n$ for use in equation (33) are obtained and are:

$$\overline{H}_2 = 20\log_{10}\left(\frac{8\pi^2\rho_1\rho_2 10^{-\alpha/10}l^2}{v^2} f_{mt}\sqrt{2}f_{d0}\right) \ldots\ldots\ldots\ldots(42)$$

$$\overline{H}_3 = 20\log_{10}\left(\frac{8\pi^3\rho_1\rho_2 10^{-\alpha/10}l^3}{v^3} f_{mt}\,2f_{d0}^2\right) \ldots\ldots\ldots\ldots(43)$$

By substitution of these values in equation (33) the maximum 2nd- and 3rd-order intermodulation noise due to feeder mismatch is obtained. The actual intermodulation noise voltage varies sinusoidally with small changes of feeder length as shown in Fig. 5. In practice the maximum 3rd-order noise is small compared with the maximum 2nd order and from the point of view of system performance only the 2nd-order noise need be considered, thus the maximum intermodulation noise is given by substituting the value of $\overline{H}_2$ from (42) in equation (33). By putting $nf_{mt} = f_{mc}$ where $n = 2$ the term D_n becomes zero and the noise $\overline{N}_n$ in dbm in a channel of mid-frequency f_{mc} is given by:

$$\overline{N}_{n\,(feeder\ mismatch)} = \overline{H}_2 + 2\overline{P} + C_n \ldots\ldots (44)$$

where

$$\overline{H}_2 = 20 \log_{10}\left(\frac{4\pi^2\rho_1\rho_2 10^{\ \alpha/10\ l^2}}{v^2} f_{mc}\sqrt{2}f_{d0}\right)$$

$\overline{P}$ and C_n are as in (12).

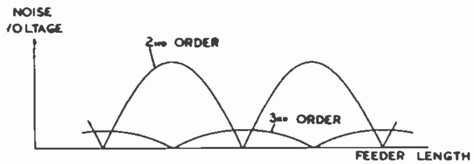


Fig. 5.—Variation of intermodulation noise with small changes of feeder length.

The sinusoidal variation of noise voltage with feeder length results in an average intermodulation noise power for random variations of length about a nominal value which is one-half of, i.e. 3 db less than, the value given by (44).

It should be re-emphasized that the analysis used to obtain these results is approximate and is valid only for systems with peak deviations up to the order of $\frac{f_s}{4}$, which is equivalent to about ± 500 kc/s for a 250-ft feeder length or correspondingly more for a shorter feeder.

6. Total Intermodulation Noise in System

6.1. *Addition of Noise due to Modulator/Demodulator and Phase Distortion*

Intermodulation products due to phase distortion are in phase quadrature to corresponding products, i.e. products of the same order and frequency, due to modulator/demodulator distortion. This is illustrated by comparing equations (5) and (18) where the same harmonics are represented by a sine term in one equation and a cosine term in the other. Corresponding products due to the two types of distortion therefore add on a power basis and the total intermodulation power at the terminal receiver output is the sum of the noise power due to modulator/demodulator distortion and that due to phase distortion, the latter being the total phase distortion over all links of a multi-repeater system and including the phase distortion due to feeder mismatch.

6.2. *Addition of Noise due to Phase Distortion over Several Links*

When it is required to determine the total intermodulation noise due to phase distortion which will be present in a system with a number of non-demodulating repeaters from a knowledge of the distortion produced by a single link or by a single feeder of given length the problem is not a straightforward one. For example, in hypothetical extreme cases the phase distortion over a number of links might cancel completely resulting in a zero intermodulation noise, or it might add due to identical phase characteristics over successive links causing in effect a voltage addition of intermodulation products. In this case the intermodulation noise power in the system would be proportional to the square of the number of links. In practice, of course, the total noise power lies somewhere between these extremes.

The distortion due to i.f. and r.f. circuits alone and neglecting feeder mismatch will be considered first. In general, the overall phase or group delay characteristic of the i.f. and r.f. circuits in a single link will have some average characteristic corresponding to the design characteristic, and there will be variations about that average value in different links due to misalignment, component tolerances, frequency drift, etc. For example in the case of 2nd-order phase distortion, i.e. a linear group delay/frequency characteristic, the group delay distortion expressed as $A_t f_d$, where A_t is a constant and f_d the deviation from the mid-frequency, might vary between $(A_t + \delta A_t)f_d$ and $(A_t - \delta A_t)f_d$ due to the random effects mentioned. Where these random variations are large compared with the average or design characteristic, i.e. if δA_t is large compared with A_t in the above example, the tendency will be for the intermodulation noise to add on a power basis over a number of links. If on the other hand the average distortion is large compared with the random variations the intermodulation noise will tend to add on a voltage basis. An exception to this latter case occurs when the modulation polarity is reversed at successive repeater stations as in the case, for example, when receiver local oscillator frequencies are above and transmitter beating oscillator frequencies are below the intermediate frequency. Under these conditions the average 2nd-order distortion tends to cancel in each pair of links, and the total noise power due to 2nd-order distortion in n links will be rather less than n times the noise power due to 2nd-order distortion in one link.

It is obvious that no inflexible rules can be laid down for the determination of total intermodulation noise in a system from simple distortion measurements over one link. In practice the 2nd-order noise power due to phase distortion on n links should normally be between n and $n^{1\cdot5}$ times that for a single link. In the case of 3rd-order distortion the factor might vary between n and n^2. In the absence of precise knowledge of the mean and extreme phase characteristics of a single link it appears reasonable to use the higher factors, i.e. $n^{1\cdot5}$ for 2nd-order distortion and n^2 for 3rd-order distortion in predicting overall system performance.

In the case of distortion due to feeder mismatch, the intermodulation noise voltage due to a single feeder varies sinusoidally with small variations of feeder length. Since the exact length of each feeder may be regarded as a random quantity the amplitude and polarity of the distortion is also random and the total intermodulation noise due to feeder mismatch tends to add on a power basis. The total intermodulation noise power may therefore be determined by the addition of the average noise power due to each feeder; this value being 3 db less than that obtained by substitution in equation (44). The total noise power thus obtained may be added on a power basis to that due to modulator/demodulator distortion and phase distortion in r.f. and i.f. circuits.

6.4. *Addition of Intermodulation Noise in Two or More Systems*

If two or more v.h.f. systems each containing a single modulator/demodulator combination are connected in tandem, the interconnection being at baseband frequency, the manner in which intermodulation products combine in the two systems depends largely on the phase characteristic of the baseband frequency circuits.[8] A suitable amount of phase shift of the intermodulation products relative to the multichannel signals producing them will prevent. voltage addition of the products in the two or more systems. Such phase distortion at baseband frequency may be present in the existing circuits or it may be introduced artificially, but in either case the result is to make the intermodulation noise in the two systems add on a power basis. Where breakdown to individual audio channels or 12-channel groups is provided at the interconnection between systems the overall performance of the "through" circuits may be improved by transposing channels or groups in the baseband spectrum so that the noisier higher frequency channels of one system become the lower frequency channels of the next.

7. Measuring Techniques for the Determination of Intermodulation Noise

7.1. *Basic Types of Measurement*

In general, intermodulation noise may be determined:

(*a*) By direct measurement of the power series constants B, C or B_p, C_p in the case of modulator/demodulator or phase distortion respectively.

(*b*) By measurement of harmonic or intermodulation products of a single- or two-tone test.[16] This is in effect an indirect method of measuring the power series constants.

(*c*) By direct measurement of intermodulation noise using a suitable test signal, conveniently a band of random noise, to simulate an actual multichannel signal. Intermodulation noise introduced into "stopped" bands is filtered off and measured.

In addition there are measurements which may be made on individual circuits which enable the phase characteristic to be determined. A particular case is the measurement of reflection coefficients which enables distortion due to feeder mismatch to be determined.

Each method has advantages under particular conditions. Method (*a*), which involves plotting and analysing the modulator/demodulator or group delay characteristics, enables the fundamental causes of intermodulation to be investigated. The measuring equipment required is, however, fairly complex. Method (*c*) enables the actual intermodulation noise due to all or part of a system to be measured under conditions as nearly as possible those of an actual multichannel signal. The filter requirements are, however, very severe and also there is the possible disadvantage when measuring a working system that thermal noise may mask intermodulation noise in the "stopped" bands. Method (*b*) may be regarded as something of a compromise between (*a*) and (*c*). The necessary equipment in this case is relatively simple and it is this method of determining intermodulation noise using single- and two-tone measurements which is discussed in the subsequent sections.

7.2. *Single- and Two-Tone Distortion Measurements*

Formulae have been derived (equations 12 and 33) for the intermodulation noise in terms of the harmonic ratio of a single-channel test tone. The use of a single tone for obtaining the harmonic ratio, however, has the disadvantages that, firstly the injected tone must be extremely pure so that harmonics are well below those generated in the circuit under test, and secondly, when it is required to use test tones of a high frequency the generated harmonics may fall outside the transmission bandwidth of the circuit and be lost. It is therefore often preferable to measure the harmonic ratios by measuring the intermodulation products of two tones and obtaining the equivalent single-tone harmonic ratios using the relations given in 3.4 and 4.4.

Figure 6 shows the schematic arrangement of a two-tone test set for measuring the performance of v.h.f. f.m. systems handling 24 or 36 channels. The equipment is in two self-contained units. The first contains the two tone generators, each tunable over the range 20-150 kc/s. The second unit, which is used for measuring the level of the intermodulation products, contains an input attenuator and amplifier used for level setting. The output from this preamplifier is switched either directly to a metering circuit for the measurement of the levels of the fundamental tones or to a second amplifier with highly selective band pass filters for the measurement of intermodulation products at 10, 50 and 90 kc/s. Amplifiers 1 and 2, which are used to increase the range of levels over which the equipment will operate, are very linear feedback amplifiers whose distortion is small compared with that of circuits under test.

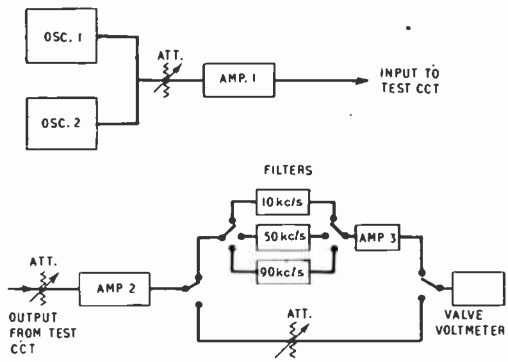


Fig. 6.—Two-tone test set.

As an example of the use of this test set, Fig. 7 shows typical results of measurements of 2nd-order distortion over a single link of a system in which the modulator/demodulator distortion was small and phase distortion predominated. At each of the three filter frequencies various tone-frequency combinations were used to check that the level of the intermodulation product of a given frequency was, in fact, independent of the frequency of the fundamental tones as should be the case from theoretical considerations. The graph shows the voltage of the intermodulation product as a function of the product frequency. In the absence of any modulator/demodulator distortion this function should be a linear one passing through the origin, but in practice the level tends to a constant value at low product frequencies due to modulator/demodulator distortion.

7.3. *Determination of Intermodulation Noise due to Modulator/Demodulator Distortion*

If a single- or two-tone test is carried out using test frequencies such that the harmonic or intermodulation product is at a low frequency in the baseband spectrum, any distortion due to the phase characteristic may be assumed to be negligible and even if the test is carried out over a complete test link, including r.f. and i.f. circuits, the resulting distortion is that due to the modulator/demodulator characteristics only. For measuring this distortion a pure single test tone of, say, 1 kc/s is convenient, the 2nd and 3rd harmonics being measured using a commercial wave analyser. The first step is to check that the distortion is a power series law by varying the deviation produced by the test tone from zero to the peak deviation determined from Fig. 1b. The 2nd harmonic ratio should be proportional to the deviation and the 3rd harmonic ratio to the square of the deviation. The 2nd and 3rd harmonic ratios may then be measured with the test tone at a level corresponding to channel deviation. These ratios expressed in db are the values of $\overline{H}_2$ and $\overline{H}_3$ which may be substituted in equation 12 to obtain intermodulation noise. The values of $\overline{P}$ and C_n are obtained from Figs. 1a and 2 respectively.

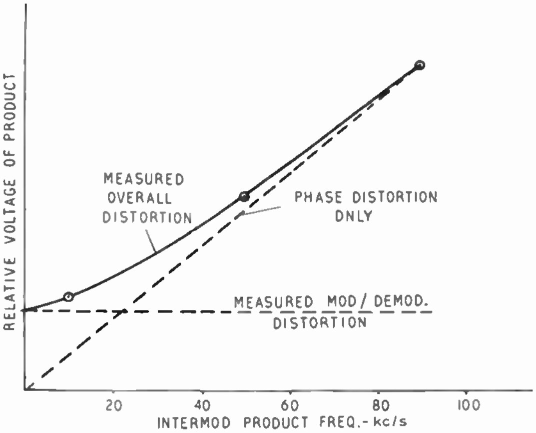


Fig. 7.—Measurement of 2nd order distortion using two-tone method.

7.4. *Determination of Intermodulation Noise due to Phase Distortion*

In measuring the level of intermodulation products due to phase distortion using a two-tone method, it is necessary to ensure that the products due to modulator/demodulator distortion are small compared with those due to phase distortion at least for the higher product frequencies. If the standard equipment modulator and demodulator does not meet this requirement it is necessary to use a test modulator and demodulator aligned for minimum overall distortion.

When measuring the distortion due to a single link using a localized test, the transmitter may be connected directly to the receiver via an attenuating cable and pads, which provide a good match and thus eliminate any distortion due to feeder mismatch. The procedure then is to measure the level of 2nd- and 3rd-order products relative to the level of a single tone when each tone is producing single-channel deviation. A high product frequency should be used to give maximum intermodulation product level for ease of measurement. Using the test set described previously, the 90-kc/s product frequency would be employed, produced by, for example, (35 + 55) kc/s in the case of 2nd-order distortion and (2 × 75 — 60) kc/s in the case of 3rd-order distortion. It is also advisable to check the power series character of the distortion in a manner similar to that mentioned in section 7.3.

If the levels of the intermodulation product of two single-channel tones measured in db relative to one tone are denoted by $\overline{I}_2$ and $\overline{I}_3$ for 2nd- and 3rd-order distortion respectively, then by putting $f_m = \dfrac{f_p}{n}$ in (30) and (31) the equivalent values of $\overline{H}_2$ and $\overline{H}_3$ for substitution in equation (33) are given by:

$$\overline{H}_2 = \overline{I}_2 - 6 \quad \ldots\ldots (45)$$

$$\overline{H}_3 = \overline{I}_3 - 21{\cdot}5 \quad \ldots\ldots (46)$$

By putting $f_{mt} = \dfrac{f_p}{n}$, D_n in equation (33) becomes $20 \log_{10} \dfrac{f_{mc}}{f_p}$.

As an example, if two-tone measurements on a particular link give values of $\overline{I}_2$ and $\overline{I}_3$ equal to — 50 and — 55 respectively at a product frequency $f_p = 90$ kc/s, the intermodulation noise in the highest frequency channel of a 24-channel system ($f_{mc} = 106$ kc/s) is given by substitution in equation (33) as follows:

$H_2 = -56$ $\qquad$ $H_3 = -76{\cdot}5$

$\overline{P} = +2{\cdot}5$ from Fig. 1a

$C_2 = -11{\cdot}5$ $\qquad$ $C_3 = -3{\cdot}5$ from Fig. 2

$$D_n = 20 \log_{10} \frac{106}{90} = 1{\cdot}4$$

C

Substitution of these values in (33) gives

2nd-order noise $\overline{N}_2 =$

$$-56 + 2(2{\cdot}5) - 11{\cdot}5 + 1{\cdot}4 \simeq -65 \text{ dbm}$$

3rd-order noise $\overline{N}_3 =$

$$-76{\cdot}5 + 3(2{\cdot}5) - 3{\cdot}5 + 1{\cdot}4 \simeq -71 \text{ dbm.}$$

7.5. *Determination of Intermodulation Noise due to Feeder Mismatch*

In order to obtain the value of intermodulation noise due to feeder mismatch by substitution in equation (44) it is necessary to know the reflection coefficients ρ_1 and ρ_2 of the aerial and transmitter output or receiver input. The techniques for such measurements using a slotted or slab line or a reflectometer are well known and will not be described here.

The modulus of the reflection coefficient should be plotted as a function of frequency. As stated in section 5.1 the theory assumes a constant value of reflection coefficient over the band of frequencies transmitted. If this is not the case, a general rule would be to use either the mean value over the band occupied by the peak-to-peak deviation, or the value at the centre frequency, whichever is the greater.

From a knowledge of these reflection coefficients and the length and characteristics of the feeder, the intermodulation noise due to each feeder may be obtained by substitution in (44) and the total noise due to all systems is then given by summation on a power basis.

7.6. *Two-Tone Measurements on a Complete System*

As a check of overall systems performance when the total intermodulation noise in a given channel is required, rather than the relative contribution of modulator/demodulator noise and phase distortion noise, a fairly simple two-tone measurement is sufficient. The procedure is to measure the 2nd- and 3rd-order intermodulation products of two single-channel test tones, the product frequency being the mid-frequency of the channel whose intermodulation noise is required. Thus in the case of the worst channel of a 12-108 kc/s 24-channel system the intermodulation products should be measured at 106 kc/s, say (66 + 40) kc/s for 2nd-order distortion and (2 × 75 − 44) kc/s for 3rd-order.

Under these conditions D_n of equation (33) becomes zero and the formulae for modulator/demodulator distortion noise (12) and phase distortion noise (33) become the same; substitution in either will therefore give the total intermodulation noise in the channel regardless of the relative contributions of each type.

Therefore using the notation of sections 7.4 and 3.5 the total 2nd and 3rd order intermodulation noise in the system is given by:

$$\overline{N}_2\,(\text{total}) = \overline{I}_2 - 6 + 2\overline{P} + C_2 \quad \ldots\ldots(47)$$

$$\overline{N}_3\,(\text{total}) = \overline{I}_3 - 21{\cdot}5 + 3\overline{P} + C_3 \quad \ldots\ldots(48)$$

In obtaining the values of $\overline{I}_2$ and $\overline{I}_3$ a series of measurements may often be necessary to average out variations due to circuit drift, slight oscillator frequency variations, etc.

8. Conclusions

In this paper expressions for different types of intermodulation noise have been derived. In general these expressions are of the form:

$$\overline{N}_n = \overline{H}_n + \overline{P}_n + \text{Constant}$$

where $\overline{N}_n$ is the nth-order noise power in dbm, $\overline{H}_n$ is the nth harmonic ratio of a test tone in db, and $\overline{P}$ is the multichannel speech power in the system in dbm under the conditions for which the intermodulation noise is required, e.g. 1 per cent of the busy hour. It is the accuracy to which $\overline{P}$ is known which usually limits the accuracy of the calculated intermodulation noise level.

In general, it may be concluded that with a knowledge of this multichannel signal level, the results of relatively simple measurements using single- or two-tone tests are sufficient to predict the level of intermodulation noise in a v.h.f. system. Some approximations are used in the determination of the formulae for intermodulation noise, but these should cause no serious errors in the results for v.h.f. systems handling up to 60 channels.

9. Acknowledgments

The author wishes to acknowledge the assistance given by colleagues at British Telecommunications Research Ltd., particularly in

connection with the measuring techniques described, and to thank Mr. F. O. Morrell, the Executive Director, for permission to publish this paper.

10. References

1. W. T. Brown. "Some factors in the engineering design of v.h.f. multichannel telephone equipment." *J. Brit. I.R.E.*, **14**, No. 2, p. 51, February, 1954.
2. L. J. Libois. "Distorsions non-linéaires en modulation de fréquence." *Câbles et Transmission*, **4**, p. 297, 1950.
3. W. J. Albersheim and J. P. Schafer. "Echo distortion in the f.m. transmission of frequency-division multiplex." *Proc. Inst. Radio Engrs*, **40**, No. 3, p. 316, March, 1952.
4. R. G. Medhurst. "Harmonic distortion of frequency-modulated waves by linear networks." *Proc. Instn Elect. Engrs*, **101**, Part III, p. 171, 1954.
5. Comité Consultatif International Téléphonique, Tome III Bis.
6. B. D. Holbrook and J. T. Dixon. "Load rating theory for multichannel amplifiers." *Bell Syst. Tech. J.*, **18**, p. 624, October, 1939.
7. B. B. Jacobsen. "The effect of non-linear distortion in multichannel amplifiers." *Electrical Communication*, **19**, p. 29, 1940.
8. R. A. Brockbank and C. A. A. Wass. "Non-linear distortion in transmission systems." *J. Instn Elect. Engrs*, **92**, Part III, No. 17, p. 45, 1945.
9. L. Lewin. "Interference in multichannel circuits." *Wireless Engr*, **27**, No. 12, p. 294, December, 1950.
10. Balth. Van der Pol. "The fundamental principles of frequency modulation." *J. Instn Elect. Engrs*, **93**, Part III, No. 23, p. 153, 1946.
11. A. S. Gladwin. Discussion on "The application of frequency modulation to v.h.f. multichannel radio-telephony." *J. Instn Elect. Engrs*, **95**, Part III, No. 36, p. 278, 1948.
12. H. T. Friis. "Microwave repeater research." *Bell Syst. Tech. J.*, **27**, p. 183, 1948.
13. L. Lewin, J. J. Muller and R. Basard. "Phase distortion in feeders." *Wireless* Engr, **27**, No. 5, p. 143, May, 1950.
14. P. J. M. Clavier. "Distorsions non-linéaires d'une signal aléatoire; application à la modulation en fréquence." *Câbles et Transmission*, **7**, No. 4, p. 293, April, 1953.
15. L. J. Libois. "Distorsion d'une onde modulée en fréquence à la traversée d'un quadripole désadapté." *Câbles et Transmission*, **8**, No. 2, p. 195, February, 1954.
16. A. T. Starr and T. H. Walker. "Microwave radio links." *Proc. Instn Elect. Engrs*, **99**, Part III, No. 61, p. 241, 1952.

GRADUATESHIP EXAMINATION—NOVEMBER 1954

FIRST PASS LIST

This list contains the results of all candidates in the British Isles and those of the oversea candidates which were available on January 7th, 1955

The following candidates have now completed the requirements of the Graduateship Examination and are eligible for transfer or election to graduateship or higher grade of membership:—

BANNOCK, Keith. (S) *Hounslow.*
BEEFTINK, Johan. (S) *Welwyn Garden City.*
BOWN, Geoffrey Charles Stanley. (S) *Penarth, Glamorgan.*
BROWN, Alfred Dixon. (S) *Sevenoaks.*

CHOPRA, Krishna Singh. (S) *Coventry.*
CLEARY, Alan. (S) *Twickenham.*
COFFEE, Ronald Alan. (S) *Thornton Heath.*

HANDS, Raymond Kenneth. (S) *Blyth.*

JARVIS, John Walter. (S) *Stevenage.*
JHA, Chandra Shekhar. (S) *Edinburgh.*

KLIMEK, Andre Mathieu. (S) *London, S.W.*

LONGLAND, David Arthur. (S) *Northampton.*

MILES, Ronald Boyce. (S) *Harrow.*
MORDEKHAI, Ephraim Hay. (S) *Ramat Gan, Israel.*

NIGHTINGALE, Daniel Edgar. (S) *London, S.W.*

PIDDINGTON, Vivian Henry. *Ashford, Middlesex.*

ROW, Edward Francis. (S) *Chessington.*

SEALE, Edward Gilbert. (S) *Dublin.*
SMITH, John Hugh. *Send, Surrey.*
STANBROOK, Donald. (S) *Stevenage.*
STEWART, Edward Samuel. (S) *Christchurch, New Zealand.*

TUCK, Leslie. (S) *Bloemfontein, South Africa.*

VINNELL, Lionel Frederick. (S) *Wells, Somerset.*

WOLFE, Brian Sinclair. (S) *London, N.*
WOOD, Jas. Gladstone Stewart. (S) *Wellington, New Zealand.*

The following candidates passed the parts indicated against their names:—

AHMAD, Saiyed Amin Uddin. (S) *Rawalpindi.* (I).
ALLEN, George Edward Elphick. (S) *Chelmsford.* (II).
ARDITTI, Joseph. (S) *Tel-Aviv.* (I).
AYLWARD, Patrick William. (S) *Dublin.* (I).

BAKHSHI, Manohar Singh. (S) *Surbiton.* (II).
BARRETT, Brendan. *Coventry.* (IV).
BOULTER, Sidney Nelson. *Manchester.* (IIIa).
BRIGGS, Peter George. *London, S.E* (I).
BURGESS, Thomas George. (S) *Wellington, New Zealand.* (I).

CARPENTER, Dennis Arthur. (S) *London, S.W.* (I).
CHALHON, Moise Oriel. (S) *Beirut.* (I).
CHOLMONDELEY-SMITH, Douglas Reginald. *Auckland, New Zealand.* (IIIa).
COLLINS, Gerald. (S) *Exeter.* (I, IIIa).
CONROY, Richard Peter. (S) *Hayes, Middlesex.* (I).
CROLE-REES, David George. (S) *London, W.C.* (II).

DAMERELL, Antony George. (S) *London, W.* (II, IIIa).
DARSHAN LAL. (S) *Dehra Dun.* (I, II, IIIa).
DE JONG, Leo Kingsley. (S) *Dehiwala, Ceylon.* (I).
DICKIN, Frank Douglas. (S) *Hereford.* (IV).
DURGOTRA, Diwan Chand. (S) *Amritsar.* (IIIa).

FITCH, Lawrence William. (S) *Shepparton, Australia.* (II, IIIa).

GARDINER, Alan. (S) *Hounslow.* (I).

HAND, Alan John. *Dalkey, Co. Dublin.* (I, II).
HAYDEN, Edward Clifford. (S) *Gt. Yarmouth.* (I).
HOLDEN, Dennis George. (S) *Kings Lynn.* (II).
HOLTZHAUSEN, Petrus Johannes. (S). *Pretoria, South Africa.* (I, II).
HURST, Sydney. *Blackpool.* (IIIa).
HUTCHINGS, Giles Samuel. (S) *London, N.W.* (I).

INDER, James Haviland. (S) *Waimauku, New Zealand.* (II).

JAMES, James Roderick. (S) *Swindon.* (II, IV).
JOHNSON-BROWN, Arthur. *Tonyrefail, Glamorgan.* (I).

LEVY, Yermiyahu. (S) *London, N.* (I, II, IIIa).

MARTIN, Arthur William. (S) *London, N.* (IIIb).
MCDONALD, Brendan Antony. (S) *Dublin.* (II).
MILLER, John Robert. *South Shields.* (IIIa).
MITCHELL, Roy Herbert. (S) *Wellington, New Zealand.* (I, II, IIIa).
MOSS, Geoffrey. (S) *Halifax.* (II).
MURPHY, Edward Colman. (S) *Blackrock, Co. Dublin.* (II).

NEWELL, Allen Frederick. (S) *Sheerness.* (I).
NOWICKI, Verzy Ryszard. (S) *London, S.W.* (I).

OVERDYKING, Raoul Jacques. (S) *Torquay.* (II).

PANCHOLI, Chandrashanker Girdharlal. (S) *London, W.* (II).
PIAK, Chan Joon. (S) *Singapore.* (I).

RUBIN, Moshe. (S) *Haifa.* (II, IIIa, IV).

SMIT, Frederik Nicolaas Johannes. (S) *Maitland, South Africa.* (I).
SMITH, Edward James. (S) *Ilford.* (IV).
SMITH, Maxwell Bendall. (S) *Bristol.* (II).
SRINIVASAN, Kalyanarama. (S) *Jamnagar, India.* (II).

TRIBBLE, Kenneth Alan. (S) *Waihi, New Zealand.* (IV).

VAN ZYL, Nicolaas. (S) *Transvaal, South Africa.* (I).

WAITE, William James. (S) *Sydney, N.S.W.* (I).
WARIN, John William. (S) *Lydd.* (II).
WOOLFORD, Alan John. (S) *Harrow.* (II).

ZERAY, Pir Mohammad. (S) *London, N.* (II).

(S) denotes a Registered Student

The second pass list, giving the results of the remainder of the successful oversea candidates, will be published in the March issue of the *Journal.*

SOME TYPICAL CIRCUITS FOR INDUSTRIAL X-RAY APPARATUS*

by

J. Jeremy Bliss, B.Sc. (Eng.) †

A paper presented during the Industrial Electronics Convention held in Oxford in July 1954

SUMMARY

The nature of X-rays and their use in non-destructive testing is briefly outlined. Suitable X-ray tubes and the associated h.t. generators are described and problems of exposure control and timing referred to. A typical 250-kV industrial X-ray generator is described in some detail, special attention being paid to the control equipment.

1. Introduction

The applications of X-rays in industry and research[1] have shown a rapid increase in recent years. Metallurgical investigation of crystal structure by X-ray diffraction methods has become standard practice, while the more conventional application of X-rays in the field of non-destructive testing now embraces a range of materials from foodstuffs to ferrous metals. Progress in the design of X-ray equipment has made possible such diverse uses as the examination of wooden structures to detect woodworm or flaws, automatic thickness control in sheet metal mills and detection of weevils in grain or foreign bodies in packed foodstuffs. X-rays also show flaws in fibrous and plastic materials and are particularly valuable in checking the setting of metal insertions in plastic components.

The more familiar aspects of X-rays in industry involve the examination of welded or cast metal components for detection of faults. The more common faults thus discovered may be divided into two groups, those formed in the liquid state and those formed in the solid state. The former include gas holes, porosity and piping; slag and sand inclusions and cold shuts. The latter are most commonly shrinkage cracks or cracks occurring during service.

To say that X-ray testing merely increases the output of scrap is an unjust and short-sighted statement. The cost of machining a casting may represent many times its initial value and it is this cost of machining which will be saved by previous X-ray examination. In a large casting the presence of blowholes or slag inclusions may not necessitate rejection. Radiographic examination from several directions may indicate that the flaws are not of a nature to cause failure in service and will not interfere with the subsequent machining.

Radiography of pilot castings may prove most valuable in indicating the improvements necessary in casting technique to eliminate repeated faults. X-ray examination may replace destructive testing in some cases, while in others it will indicate which parts of a workpiece should be selected for micro-examination.

2. The Nature of X-rays[2]

The comparative transparency of materials to X-rays depends on their atomic numbers and upon their density. Aluminium, for instance, is about as transparent as glass but ten times more opaque than water, whereas lead is almost entirely opaque.

Clearly the human body is composed of materials which are highly transparent when compared with metal workpieces. The boldest contrasts occur between bone and tissue or between the lungs filled with air and the heart filled with blood. The entire human body, in fact, is not much more opaque than water. Obviously industrial radiography calls for apparatus and techniques differing considerably from those used in the hospital.

X-rays are generated by conversion of energy produced by the bombardment of a metal target with a stream of electrons having a velocity of one-third to four-fifths of that of light. They are believed to be manifestations of electro-magnetic wave motion occurring at a frequency higher than that of ultra violet

* Manuscript first received April 27th, 1954, and in revised form June 4th, 1954. (Paper No. 300.)

† Marconi Instruments Limited, St. Albans, Hertfordshire.

U.D.C. No. 621.386.1 : 620.179.

light but lower than that of cosmic rays, their wavelengths varying from 0·5 to 0·01 Angstrom units.

The shorter their wavelength, the higher is their penetrating power. The wavelength, in turn, depends upon the energy dissipated in the bombardment. The energy converted as each electron strikes the target is proportional to the square of its velocity. The velocity of the electrons is a function of the anode voltage of the X-ray tube which may vary from some thirty kilovolts to several megavolts.

An unsmoothed voltage applied to the tube will produce X-rays varying in wavelength throughout the cycle. This produces a contrast on the film when dealing with a subject composed of parts varying only slightly in opacity and is thus more suitable for clinical than industrial work.

Constant anode potential, while not producing homogeneous radiation, provides a considerably higher proportion of X-rays of the minimum wavelength obtainable at the applied kilovoltage. It is thus suitable for industrial radiography of metal workpieces. In addition to the penetrating power of such a beam, its intensity is constant over the entire cycle as opposed to that produced by a sinusoidal voltage where X-rays are only generated for a part of each cycle, and the beam is accordingly pulsating. Thus the constant potential set, though unsuitable for clinical radiography, since the scarcity of soft radiation results in a lack of contrast, is ideal for radiography of metals since exposure times are considerably reduced.

3. The X-ray Tube

The generation of X-rays involves, as has already been stated, the bombardment of a target of heavy metal by a stream of high-velocity electrons. In the conversion of energy which takes place some 99 per cent. is released as heat. Taking some typical figures for an industrial X-ray set, a tube current of 15mA at a voltage of 250 kV will result in a dissipation of over three-and-a-half kilowatts as heat at the focal spot of the anode. Obviously, the smaller the area of the focal spot, the clearer will be the radiograph, since the ideal condition would be a point source of X-rays. Reduction of the focal spot size, however, is limited by the problem of heat dissipation, since too small a spot size for a given current and voltage in the tube will cause the anode to melt.

The problem may be approached in numerous ways. In clinical work a rotating anode is used, so that, while the effective focal spot is small, the rotation of the anode increases the area in which the heat is dissipated, thus the actual area under bombardment is in the form of a ring many times the area of the effective focal spot. Tubes of this type will withstand tube currents of about 500 mA at about 100 kVp, or 100 mA at 300 kVp, but only for exposures of a few hundredths of a second. This time limitation renders them unsuitable for most industrial work, where exposures are usually of several seconds or even minutes duration.

The reason for this limitation lies in the conductivity of the anode assembly. While the rotation of the anode prevents melting of the anode at the point under bombardment, the heat storage capacity of the anode assembly is quickly expended. Since the entire rotating assembly must be within the vacuum, heat can only leave it by radiation or by conduction through the anode itself. Bearings on which the anode may rotate at some 3,000 r.p.m. in vacuo offer many design problems and present a path of very low thermal conductivity. Thus the temperature of the entire anode assembly may rise excessively if the current, kilovoltage and time values are not carefully restricted within the limits imposed by the tube rating.

The logical step in tube design to permit continuous operation at higher ratings is to employ some means of conducting the heat away from the anode. This is usually done by employing a hollow anode into which cooling oil may be introduced. The circulation of the oil may be confined to the space between the glass envelope and the earthed metal shield, but for higher ratings a pump is employed to circulate the oil. In such cases the oil circulator, which is connected to the tubehead by high-pressure hoses, incorporates a substantial reservoir and a heat exchanger employing either water cooling, air-draught cooling or both. This solution to the problem of heat conduction from the anode unfortunately rules out the use of a rotating anode and thus involves the initial difficulty of heat dissipation in the area of the small focal spot. Thus, even with the most efficient cooling methods available, choice must be made between small focal spot size, and greater power with longer exposure times.

For certain specialized applications continuously evacuated tubes may be used. In crystallography interchangeable anodes of different materials are employed, and in other special techniques anodes having different anode angles may be used. The continuously evacuated tube has many advantages. For instance, damaged electrodes may be replaced, thus saving the cost of a new tube. Against this there are the added complications associated with the operation of evacuating pumps.

4. Timers

An X-ray exposure may be initiated and terminated either by interruption of the h.t. primary circuit or by the interposition of lead shutters between the tube and the film. The first method is suitable for tubes where heat dissipation is the primary problem or where short exposures are required. The second method is often used where the tube kilovoltage is so high that it is applied to the tube in steps at automatically determined intervals. Where this is done, more accurate exposure determination is possible if the X-rays generated during the starting period are not used and the exposure occurs while radiation is constant.

For accurate timing of short exposures ignitron switching[3] may be employed and the timer phased to the frequency of the supply mains so that the make and break occur at the zero point of the cycle. For lower power sets, phased timers employing thyratrons and mechanical contactors[4] are used. These may be entirely non-mechanical, with resistance-capacitance timing circuits.

For industrial work in general, exposures are relatively long and mechanical or electro-mechanical timers are used. Clockwork timers are generally confined to the smaller portable sets and the most common timer for higher power equipment is the synchronous motor type, which may control either the h.t. primary circuit or the radiation shutters.

One circuit in common use provides interchangeable links to connect the timer in either circuit as required. Interlock circuits ensure that X-rays cannot be switched on with the shutters open, thus eliminating radiation during the automatic delay period while the tube kilovoltage is being stepped up to its selected value. The synchronous motor driving the timer will only operate when radiation is in progress, that is, when X-rays are on and the shutters have been subsequently opened. If X-rays are switched off, or the shutters closed, the timer stops. On completion of the selected exposure time, either the shutters are automatically closed, while X-rays remain on, or the h.t. primary circuit is interrupted, terminating radiation from the tube while leaving the shutters open.

5. Types of H.T. Generator[5]

Since the generation of X-rays depends upon the conversion of the kinetic energy of a stream of electrons, an essential part of any X-ray equipment is a generator of the high tension which imparts the necessary velocity to the electrons.

For many years when only direct current was available, first from batteries and later from d.c. mains supplies, high tension was generated by means of an induction coil with primary interruption. So efficient did these interrupters become, that they were even used with a.c. mains supplies.

However, as a.c. mains became more generally available, the coil and interrupter was gradually replaced by the high tension transformer. The change was slow, since the h.t. transformer was more expensive than the coil and interrupter and, in addition, the disposal of the unwanted half cycle had to be faced. Interrupters, however, required much laborious maintenance if they were to remain efficient, and it was on this score that they were finally ousted.

The X-ray tube, being a form of diode, may be used as a rectifier. In sets of low output up to about 90 kVp and 100 mA, the X-ray tube is connected directly across the secondary winding of the h.t. transformer. Apart from the general disadvantage of half-wave rectification, this system means that the X-ray tube is limited to a working kilovoltage considerably below its actual rating. Suppose, for instance, that a tube can withstand a certain voltage without flash-over occurring between its electrodes. If the anode is raised to that potential when a current is passing through the tube, the inverse voltage which occurs during the alternate half-cycles will be considerably higher since, during these non-working half-cycles, no current passes and, as a result, there is no voltage drop in the transformer. Thus, if the tube is to withstand the inverse voltage, the working voltage which serves to generate the X-rays must be considerably reduced.

There is also a serious restriction of tube current since, if the anode becomes too hot, it will emit electrons. If this happens it will act as a cathode during the inverse half-cycle and the resulting electron stream may bombard the filament or even the glass wall of the tube, which will melt, so that the tube is destroyed. Thus the current which a tube will stand under self-rectified conditions is considerably below that permissible with rectified high tension, where anode temperatures may be higher without damage. This system has, however, the great advantage in saving of cost and reduction of weight and volume.

5.1. *Simple Rectifier Circuits*

The logical remedy to relieve the X-ray tube of the inverse voltage is to connect a rectifier valve in series between the tube and transformer (see Fig. 1). This gives only half-wave rectification, but relieves the tube of the strain of high inverse voltages. However, the rectifier itself is now subject to a high inverse voltage.

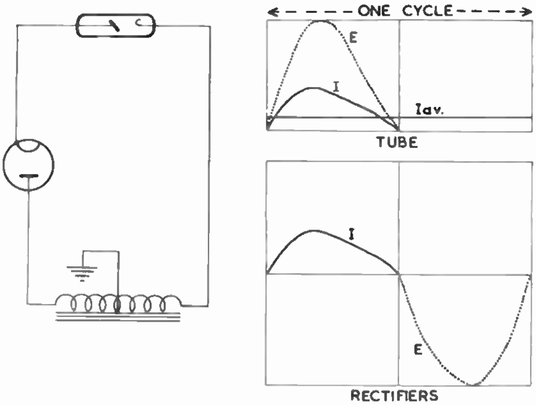


Fig. 1.—Single-rectifier half-wave.
$E_p(Trans) = E_p(X\text{-}RT) = E_p(Rect)$
$I_p = 2.8\ I_{av}$.

A further disadvantage of this arrangement is that the system is not completely balanced about the earthed centre-point of the transformer secondary, since there will be a voltage drop of about a kilovolt across the rectifier. This may lead to breakdown of the h.t. cable on the other side owing to the alternating tension between its central conductor and earthed sheath.

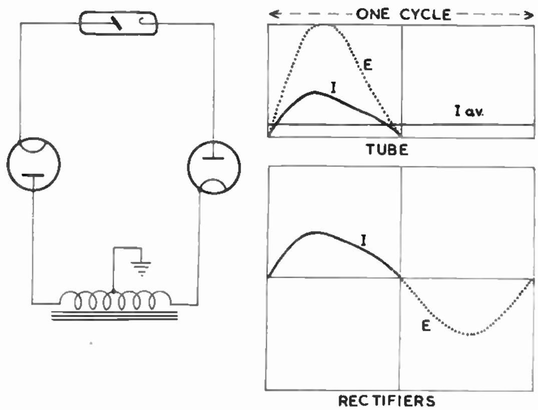


Fig. 2.—Two-rectifier half-wave.
$E_p(Trans) = E_p(X\text{-}RT) = 2E_p(Rect)$
$I_p = 2.8\ I_{av}$.

The essential difference between the operation of the rectifier and that of the tube is that, while the former has an excess of electron emission from its filament, with the result that it is space-charge limited and the voltage drop across it is comparatively low, the X-ray tube, on the other hand, operates at saturation, so that virtually all the electrons emitted from the filament reach the anode. Control of the tube current is thus by adjustment of filament current and almost all the voltage drop in the circuit is across the tube.

By adding another rectifier in series with the tube the inverse voltage across each rectifier is reduced to half the working voltage of the tube and the system is properly balanced about

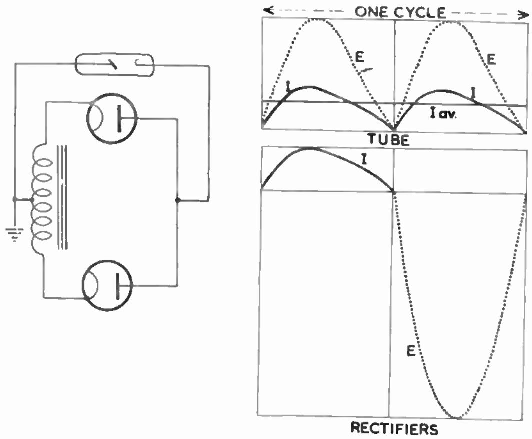


Fig. 3.—Two-rectifier full-wave.
$E_p(Trans) = 2\ E_p(X\text{-}RT) = E_p(Rect)$
$I_p = 1.4\ I_{av}$.

the earthed centre-point of the transformer secondary (see Fig. 2). This relieves the strain on the cable insulation since the potential to earth is uni-directional.

It is, of course, possible to obtain a full-wave rectified output, thus doubling the average tube current, using only two rectifiers (see Fig. 3); but this is seldom done since the inverse potential across the rectifiers is twice the working voltage of the tube, while the tube voltage is only half the voltage developed across the transformer secondary. A further disadvantage is that the system is unbalanced, since one electrode of the tube is permanently at earth potential. This leads to troublesome insulation problems, although tubes with extended anodes at earth potential are used in certain specialized applications.

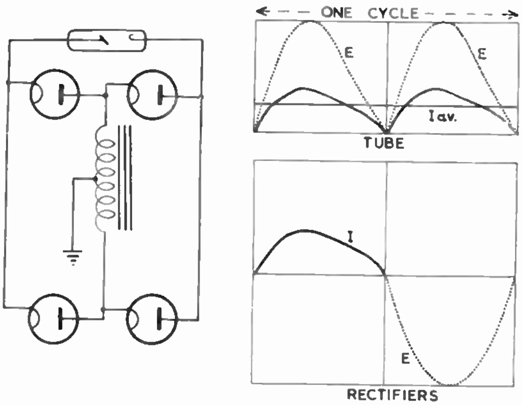


Fig. 4.—Four-rectifier full-wave.
$E_p(Trans) = E_p(X\text{-}RT) = E_p(Rect)$
$I_p = 1.4\ I_{av.}$

Probably the most common of the pulsating output generators is the four-rectifier full-wave (see Fig. 4). Here the rectifiers are connected in a bridge circuit and the system is balanced about the earthed centre-point of the transformer secondary. In this circuit the transformer output voltage, the tube voltage and the inverse voltage across the rectifiers are all equal.

5.2. *Voltage Doubling Circuits*

The higher power industrial sets frequently employ voltage doubling circuits to supply h.t. These utilize the principle of charging capacitors in parallel and discharging them in series, as seen in the appropriate diagrams. One common type is the Villard circuit which is shown in Fig. 5.

During the first half cycle connection T1 of the transformer secondary is negative relative to the earthed centre-point, while T2 is positive. C1 is thus charged as shown by the current passing through V1, and C2 by the current through V2. As the output voltage of the transformer falls to zero, so the voltage across

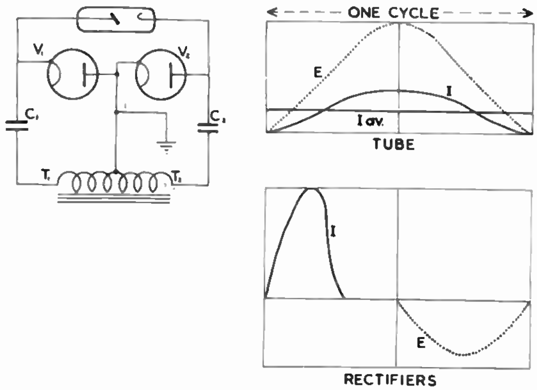


Fig. 5.—Villard voltage doubler.
$E_p(Trans) = 0.5\ E_p(X\text{-}RT) = E_p(Rect)$
$I_p(X\text{-}RT) = 2.0\ I_{av.},\ I_p(Rect) = 2.5\ I_p\ (X\text{-}RT)$

the tube becomes equal to the voltage across C1 plus that across C2, since they are effectively in series. This voltage is approximately equal to the peak voltage developed across the secondary. As connection T1 becomes positive and T2 negative, so the voltage across the tube rises to twice the peak voltage of the transformer since the transformer secondary is effectively in series with C1 and C2 with the two rectifiers blocking.

For satisfactory operation of this circuit the capacitance of C1 and C2 must be sufficiently high for their voltage to remain almost constant. Also the charging current through V1 and V2 must be considerably higher than the tube current.

The resultant output will be pulsating, but will have a frequency of half that of the transformer. This makes the circuit unsuitable for sets where very short exposure times are required since the output during the initial cycle will be distorted while the capacitors are being charged from zero. The circuit is, however, quite suitable for normal industrial work, where exposures are of a length which makes this irregularity negligible.

Another voltage doubling circuit which has wide application in industrial sets of the order of 200 to 300 kilovolts is the Greinacher circuit (see Fig. 6). This provides a constant potential output and thus reduces exposure times in industrial work.

On consideration of the circuit diagram it can be seen that during the first half-cycle, if the free-end of the winding becomes positive, C1 will be charged through V1 in the direction indicated, while V2 blocks. During the next half-cycle, it becomes negative so that C2 is charged as indicated while V1 blocks. The only discharge path for C1 and C2 is through the tube, and since they are effectively in series their potentials combine to give a tube voltage of double that of the transformer output.

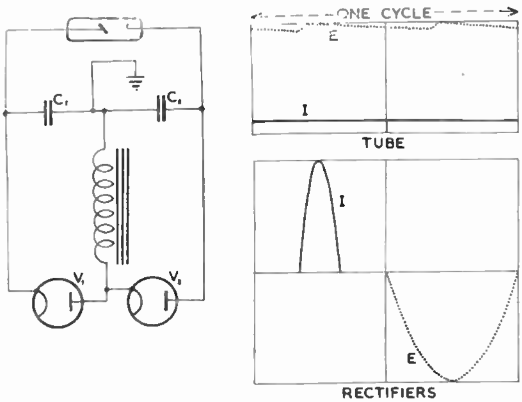


Fig. 6.—Greinacher constant potential voltage doubler.
$E_p(Trans)=0.5\ E_p(X\text{-}RT)=0.5\ E_p(Rect)$
$Ip(X\text{-}RT)=I_{av.},\ I_p\ (Rect)=10\ I(X\text{-}RT)$

A double Greinacher circuit is often employed in which the secondary of the transformer is split into two and a single Greinacher circuit connected to each winding, as shown in Fig. 9. This enables metering and current stabilizing circuits to be connected between the two secondaries so that they are maintained at earth potential. It also reduces the ratings of the components employed.

Another similar circuit which is sometimes used is the voltage quadrupler (see Fig. 7). Initially C1 is charged through V1 to the voltage developed across the transformer secondary and having the polarity indicated. During the subsequent half-cycle C2 is charged through V2 to the same voltage. At the same time, while C1 is being charged through V1, the transformer secondary is effectively in series with

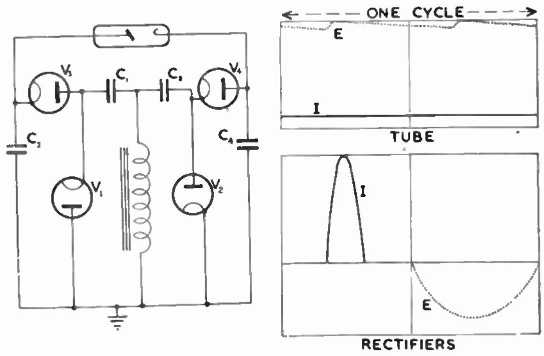


Fig. 7.—Constant potential voltage quadrupler.
$E_p(Trans)=0.25\ E(_pX\text{-}RT)=0.5\ E_p\ (Rect)$
$Ip(X\text{-}RT)=I_{av.},\ I_p(Rect)=10\ I_p(X\text{-}RT)$

C2 and C4 through V4, so that C4 is charged to the combined voltage of C2 plus the transformer secondary, that is to twice the transformer output voltage. Similarly, while C2 is being charged through V2, C3 is charged through V3 by C1 in series with the transformer secondary. C3 and C4 are effectively in series with the X-ray tube and apply across it a voltage four times that of the ouput of the transformer. At the same time, the circuit is symmetrically connected about the earth point.

Three-phase h.t. generators for X-ray equipment have never been very popular in this country, except for very high voltage equipment. In the three-phase full-wave circuit shown in Fig. 8, six rectifiers are required in addition to the three-phase transformer and the necessary filament transformers. Thus the cost, weight and volume are all high. Costs of manufacture are increased by the limited demand for equipment requiring a three-phase supply of low resistance.

In such a circuit the voltage never drops to zero, but varies in a complex manner over an appreciable range. The magnitude of the change from minimum to maximum is a function of many variables including the transformer leakage reactance and the tube current. In most cases its magnitude will be of the order of 20 to 40 per cent. of the maximum voltage, though some improvement in the wave-form may be achieved by use of compensating reactors. It is sometimes believed that three-phase rectification gives a far higher X-ray output than single-phase. This is probably due to the popular misconception regarding the output voltage wave-form. Actual experiments show little difference in radiographic speed between three-phase and single-phase rectified equipment operating at the same peak kilovoltage and the same milliamperage.

Another method, which was formerly used where the mains regulation was extremely poor, was that of charging a large capacitor and then allowing it to discharge through the tube. Among the disadvantages of such a system is the fact that the exposure time and current are interdependent and can be changed only by substituting a different capacitor. The very short exposure times produced are sometimes of value in medical practice, but would be of little use in industrial work.

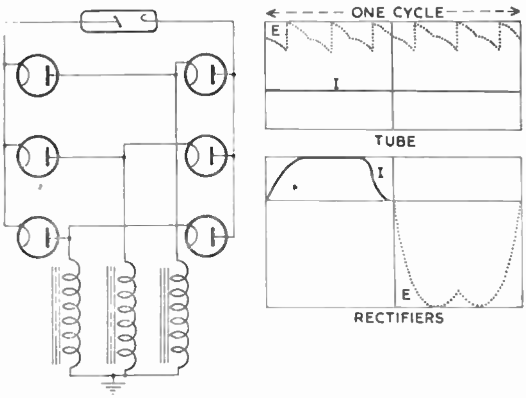


Fig. 8.—Three-phase full-wave.
$E_p(Trans)=0.56\ E_p(X\text{-}RT)=0.56\ E_p(Rect)$
$I_p(X\text{-}RT)=I_{av.}=I_p(Rect)$

There are, of course, a number of generators, such as the betatron, which come under the heading of electron accelerators. The X-rays generated by such methods are extremely hard and of a similar nature to the emission from radioactive substances. The nature of the problems involved and the circuits used with such generators are so different from the more conventional types that they have been omitted from this paper since discussion of them would afford scope for another paper at least as long as this one. The future of such X-ray generators is controversial since, while experimental research in this field is still in progress, it would appear that such expensive equipment may soon be ousted from certain industrial fields by some of the new radio-active isotopes.

At one time, before shockproof cables and X-ray tubes were produced, X-ray equipment was built into a suitable room with nothing but the air as an insulator for the high tension. While such an arrangement was, perhaps, ideal from the service engineer's point of view, the space required and the danger of electrocution soon rendered it obsolete. In certain cases, where a building is specially designed for the purpose, it might still be used, but such cases are rare, and, from a manufacturer's point of view, such apparatus, being virtually custom-built, is expensive to produce. Nowadays, the oil- or gas-insulated h.t. generator contained in a shock-proof tank and connected to the X-ray tube by shock-proof cables is almost universal, the only common variation being the use of an h.t. generator oil-immersed within the tube-shield to form a single shock-proof unit. Oil insulation permits the use of smaller clearances and more compact rectifying valves and enables the entire h.t. generator, with its rectifiers, capacitors and filament transformers, to be contained in a tank of reasonable dimensions.

Much experimental work has recently been carried out on gas-insulated h.t. generators.[6] The comparison is not within the scope of this paper but some salient points are worth mention. Freon has now been fairly generally discarded in favour of sulphur hexafluoride but in spite of its better insulating properties the clearances necessary are still several times as great as those required in oil. Even then, to reduce the size of the equipment to practicable dimensions, the gas must be subjected to considerable pressure. This necessitates an increase in thickness of the tank, which goes far towards nullifying any saving in weight.

A further disadvantage is that, whereas leakage of oil is readily detected before serious loss occurs, a leakage of gas may not be apparent until some major breakdown has occurred. In addition to this, the servicing of gas-insulated equipment presents difficulties not encountered with the oil-immersed type.

6. A Typical 250-kV Industrial X-ray Generator

To appreciate the application of the circuits employed it is necessary to consider some typical example of a complete X-ray generator. The equipment chosen is a 250-kV constant potential generator[7] of wide application and

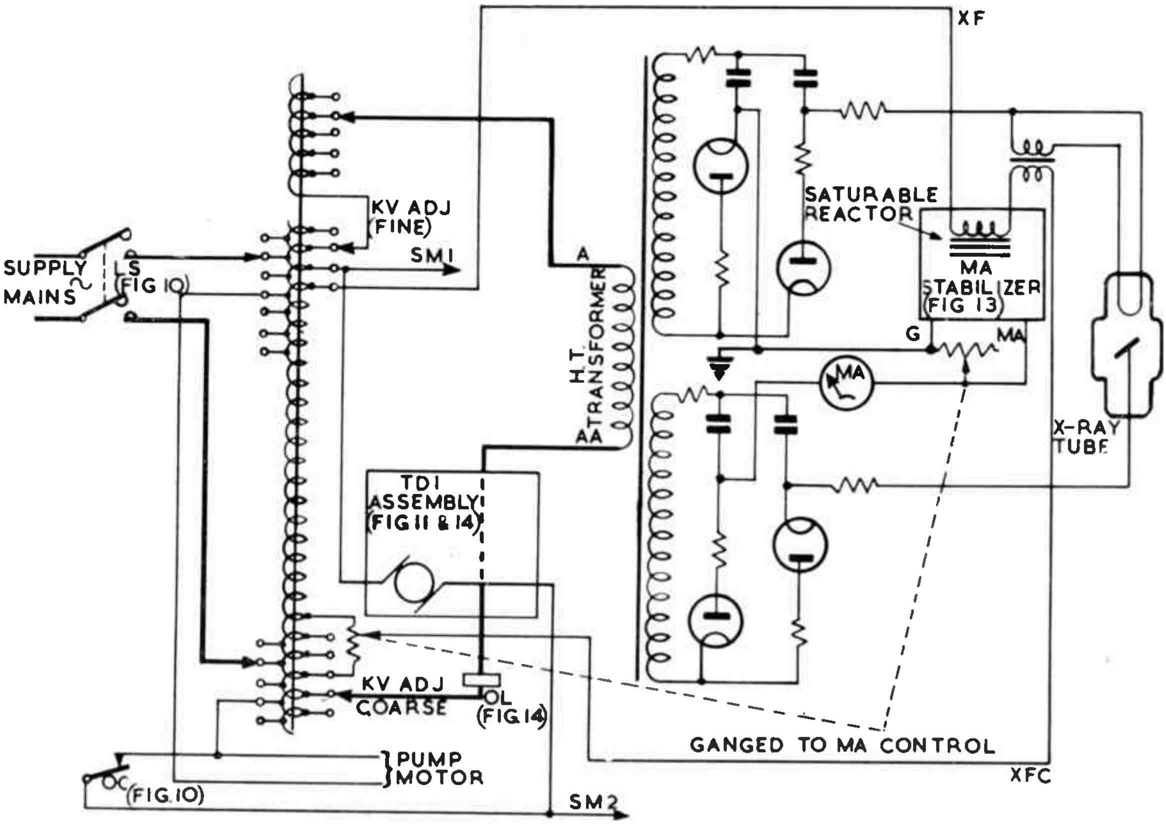


Fig. 9.—Overall functional diagram of typical 250-kV constant-potential X-ray equipment.

suitable for examination of any material from plastics and timber to bronzes and ferrous metals of thicknesses up to about three inches.

The first functional diagram (Fig. 9) shows the basic circuit of the equipment from the mains input terminals to the X-ray tube. The mains supply is taken to an auto-transformer feeding the h.t. transformer primary, the pump motor and the internal supply lines (SM1 and SM2), which are used to energize all relays and auxiliary circuits.

The h.t. generator is a double Greinacher voltage doubler, with earthed centre-point. This is normally oil-immersed and provides a constant potential output which will be almost ripple-free if the capacitors are sufficiently large. The four separate filament transformers for the rectifier valves are incorporated in the h.t. generator, as is the transformer for the filament of the tube itself. Separate transformers for each rectifier are essential, since their directly heated cathodes are all at different potentials. One, for instance, is at earth potential, while another is connected to the anode of the tube.

The primary winding of the X-ray tube filament transformer is fed from XF and XFC via the saturable reactor of the milliamperage stabilizer, the impedance of which controls the filament current and hence the tube current. The method by which this impedance is varied is discussed later.

The tube itself is contained in a shockproof, rayproof shield and cooled by oil which is pumped into the hollow copper anode whence it flows between the glass envelope and tube-shield to the return pipe. The tubehead is mounted on a stand which allows freedom of

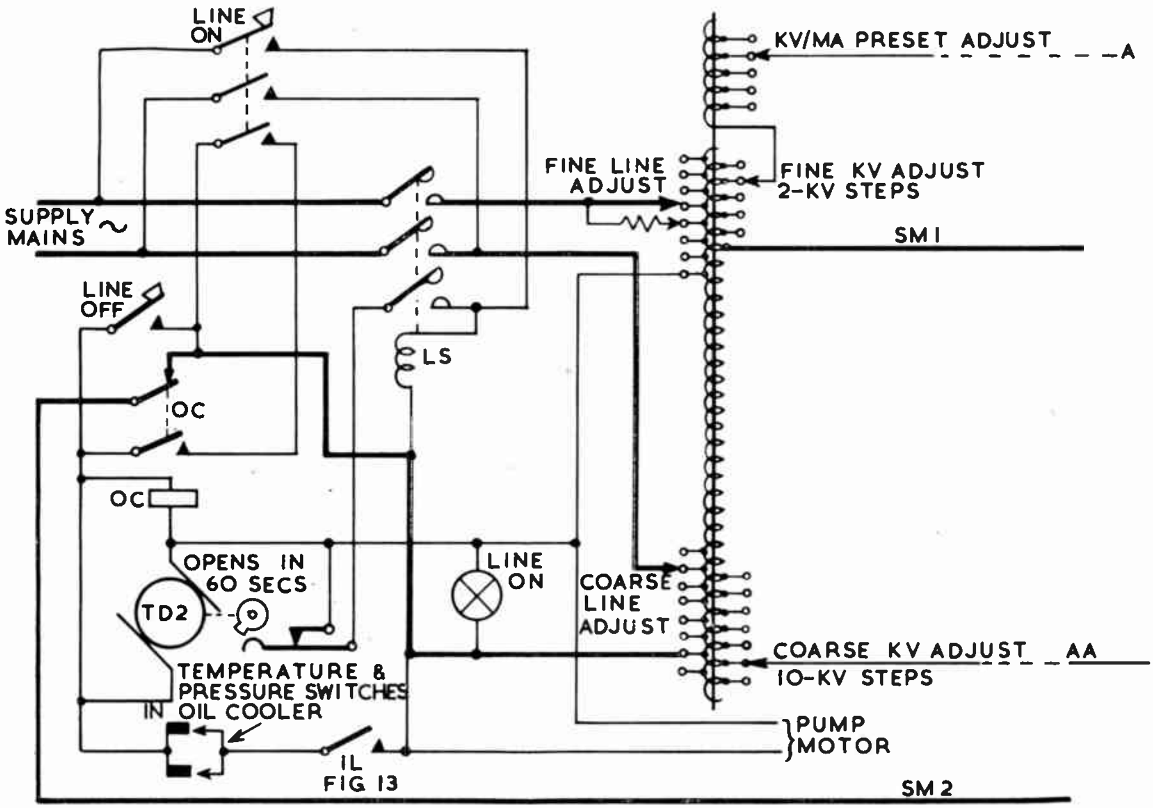


Fig. 10.—Functional diagram of Line On/Off circuits.

movement in all directions and connected to the h.t. generator by shockproof h.t. cables.

The oil cooler comprises a reservoir, with immersed cooling water pipes, and an electrically driven pump. The unit has an automatic water valve, to turn on the cooling water as soon as the equipment is energized, and safety relays which shut down the equipment in the event of rises in oil temperature or drops in oil pressure in excess of the prescribed amount.

The control unit is normally installed in a protected room adjacent to the X-ray inspection room and bears on its front panel all meters and controls necessary for remote regulation of the exposure.

The second functional diagram (Fig. 10) shows how the set is switched on. This circuit is described in some detail since a time delay mechanism is incorporated which ensures that cooling oil continues to be pumped through the tubehead for at least 60 seconds after X-rays have been switched off in order to prevent carbonization of the oil in the anode.

The diagram shows that pressing "Line On" button will energize the Line Switch contactor (LS). The hold-on interlock of this contactor will keep it closed so long as the Time Delay Contacts (TD2) remain closed. This maintains the supply to the auto-transformer from which are taken the tappings for the SM1 and SM2 lines (internal supply leads), the A and AA connections to the h.t. transformer primary and the feed to the pump motor.

The main supply is taken to the auto-transformer via coarse and fine adjustment tappings. These give compensation for changes in mains voltage and thus ensure a constant output voltage from the auto-transformer. The "Line On" light on the control panel indicates that the auto-transformer is energized.

Pressing the "Line Off" button does not immediately disconnect the auto-transformer from the supply mains but energizes the OC relay and starts the Time Delay (TD2) motor. The OC relay breaks the SM2 line, thus disconnecting the supply to all the relays other than OC and LS. This serves to break the AA line, shutting of the h.t., and renders the "X-rays On" and "portal" buttons inoperative. The pump motor, however, continues to run, the X-ray tube filament remains on and, since the auto-transformer is still energized, the "Line On" pilot light is not extinguished.

The TD2 motor drives a cam which, after 60 seconds, opens the TD2 contacts, disconnecting the supply from the LS contactor and thus breaking the connection between the auto-transformer and the mains input.

The second pair of contacts on the OC relay forms a "hold on" interlock, but is in series with a pair of normally closed contacts on the "Line On" button. Thus, should the "Line Off" button be pressed and then the "Line On" button be pressed again before 60 seconds have elapsed, the OC relay will open, the TD2 motor will stop and the cam, which is spring loaded, return to zero. This will restore normal operating conditions and avoid the necessity of waiting until the set has been switched off, at the end of the time delay period, before it can be switched on again.

It is clear from the diagram that there is a way in which the OC relay may be energized other than by operation of the "Line Off" button. The temperature and pressure safety switches in the cooler are effectively in parallel with the "Line Off" button contacts, but are in series with the IL relay in the milliamperage stabilizer. The IL relay, which is discussed more fully later and shown in Fig. 13, provides an adequate time delay after the initial switching on of the equipment for the oil pump to build up sufficient pressure. If this relay did not nullify the action of the pressure safety switch for a few seconds, the set would automatically switch itself off every time it was switched on.

The third functional diagram (Fig. 11) shows the circuit which must be completed before h.t. can be applied to the tube. This "X-rays On" circuit is drawn boldly in the diagram and consists of a series of interlocks, all of which must be closed before the AA line to the h.t. transformer primary can be completed.

If the circuit is traced from the SM2 line shown boldly at the top of the diagram, the first interlock is the Overcurrent Breaker. This protects the X-ray tube against excessive current and the details of its circuit are discussed later.

The Filter Interlock provides a means of checking, from the control room, that the correct filter holder is in position in the tube-head, since, only when the filter acknowledgment switch on the control panel is set to the corresponding filter value will the filter pilot lamp light and the relay close. Although the use of filters in industrial radiography is a controversial subject, there are many industrial radiographers who favour them. To make a number of exposures with the wrong filter in position is to waste time and film and thus the filter acknowledgment system provides a valuable safeguard by enabling the operator to check the filter value from the control panel.

Next in the circuit is the door safety interlock, which ensures that X-rays cannot be switched on unless the X-ray room doors are closed. This interlock may be wired in an alternative position to ensure that the radiation shutters cannot be opened unless the doors are closed. In either case this precaution is an essential one for the safety of all who work near to the X-ray room.

The electronic milliamperage stabilizer is discussed later. It employs a number of thermionic valves and it is important that these have warmed up so that the stabilizer is operating properly before h.t. can be applied to the tube. The output from a diode rectifier in the stabilizer operates the next relay (IL), which cannot close until sufficient time has elapsed after switching on the equipment for the cathode of the diode to reach its operating temperature. This delay not only allows a warming-up period for the other valves, but provides sufficient delay period for the oil pump to raise the oil pressure to the prescribed value and for the X-ray tube filament to reach its operating temperature.

Operation of the "X-rays Off" button opens a pair of normally closed contacts which form part of the circuit.

Next in the circuit is the Test/Operate Switch. This is normally kept in the "Operate" position but serves to keep the "X-rays On" circuit open when set to the Test position, in

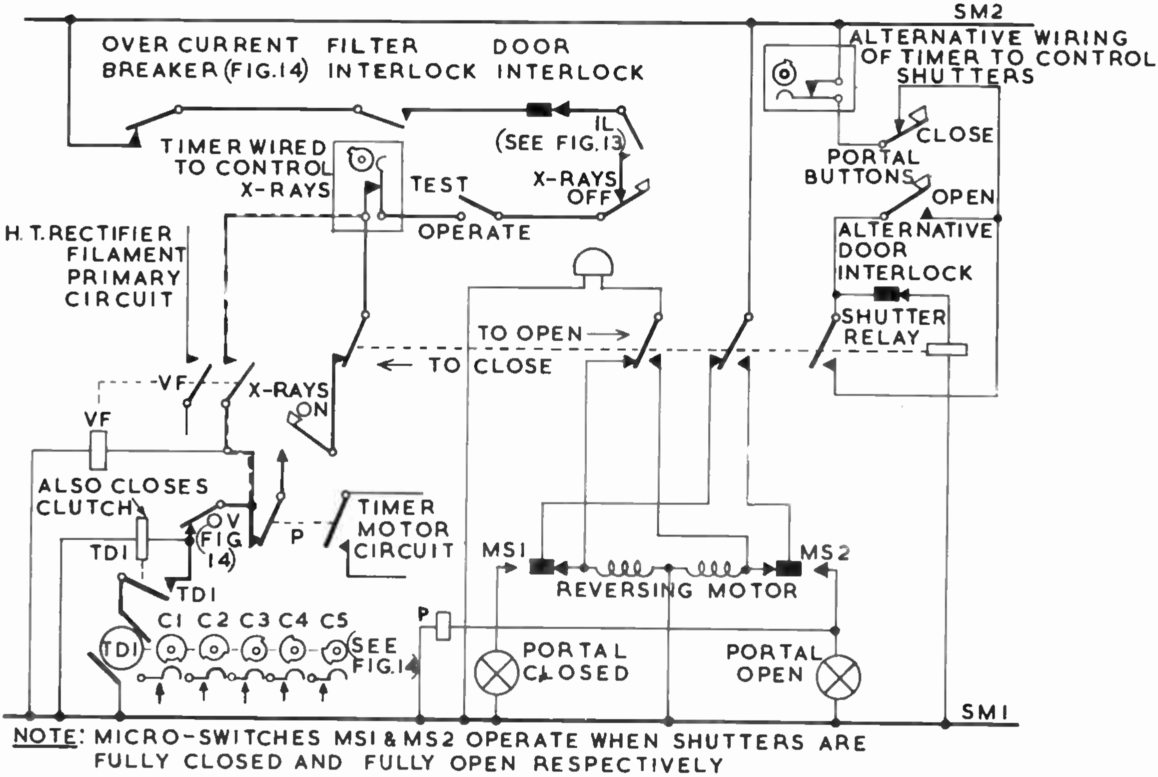


Fig. 11.—Functional diagram of " X-rays on " circuit and shutter control circuit.

order to facilitate tests which are carried out during servicing and setting up.

The timer, which in this equipment, is of the synchronous motor type with a range from a hundredth of a minute to 60 minutes, may be wired—like the door switches—either to control the X-rays or the radiation shutters. When it is wired to control X-rays, it must be set to some value other than zero before the " X-rays On " circuit can be complete.

The next part of the circuit is so closely associated with the shutter control system that the two circuits have been shown in the same diagram. The operation of the shutters is discussed in greater detail in a later section. For the moment it is only necessary to appreciate that when the main shutter relay is energized the shutters will open, when it is not, they will close. Also, when the shutters are fully open a micro-switch operates, opening the contacts of the relay P. Thus the first interlock ensures that X-rays cannot be switched on when the main shutter relay is energized. The P relay co-ordinates the opening of the shutters with the starting of the timer and, at the same time, prevents X-rays from being switched on with the shutters open. This is an added safeguard since, if the shutters were mechanically held open, the main shutter relay might not be energized, thus closing that part of the " X-rays On " circuit without the shutters actually closing.

These two shutter-controlled interlocks, together with the " X-rays On " button contacts, which are connected between them, form a part of the circuit which is by-passed by the VF relay as soon as this is closed. Thus, once the " X-rays On " circuit is complete, the

" X-rays On " button may be released and the shutters opened without the circuit being interrupted. The VF relay also closes the circuit to the h.t. rectifier filament primary windings.

The final relay in the circuit is the normally closed Over-Voltage breaker. The circuit details of this protective device are discussed later, but, provided no excessive voltage occurs across the h.t. transformer primary, the " X-rays On " circuit will be completed to the Time Delay Assembly (TD1). The TD1 solenoid will thus be energized, simultaneously completing the supply to the motor and closing the clutch through which the motor drives its camshaft.

6.1. *The Time Delay Contactor Assembly*

The h.t. transformer primary is energized over a period of ten seconds by five contactors. As the TD1 motor drives its camshaft the cams close switches which energize the contactors closing the AA line through a resistance which is reduced in five steps from an initial value of about ten ohms, thus gradually increasing the tube anode voltage to its selected value. This system greatly increases the life of the X-ray tube and is essential where voltages in excess of about 200 kV are employed. The series/parallel connection of the contactors is shown in Fig. 14.

The first of these contactors connects one side of the timer motor to the SM2 line. The other side is connected to the SM1 line, starting the timer, when the shutters are fully open and the P relay operates.

These cam-operated switches also energize the exposure counter and the exposure time totalising counter, both of which give valuable information concerning the tube life in the event of tube failure. In addition, the first switch of the series (C1) puts on the " X-rays On " pilot light.

When any of the interlocks in the " X-rays On " circuit open, or the " X-rays Off " button is pressed or the timer reaches zero (if wired in this circuit), the supply to the TD1 motor is broken and the clutch opened. As soon as the camshaft is thus released, it is automatically reset to its initial position by a spring-loaded return mechanism. When this occurs, the primary circuit of the h.t. transformer is broken, the timer stops and X-rays can only be initiated by pressing the " X-rays On " button again with the shutters closed.

6.2. *Operation of the Timer*

The timer may be set manually like a clock. The synchronous motor drives the two hands and a micro-switch opens as each hand reaches zero. These micro-switches are wired in parallel, so that the circuit is only broken when both hands are at zero simultaneously. This system provides accuracy to about a hundredth part of a minute. When it reaches zero it breaks either the " X-rays On " circuit or the shutter control circuit according to the system of connections employed. Its driving motor will run only when both the P relay and the C1 contactor are closed as already explained. Thus it may be stopped either by switching off the X-rays or by closing the shutters.

6.3. *Shutter Operation*

In Fig. 11, to facilitate explanation, the shutter circuit is shown with the shutters at rest and deliberately left in some intermediate position. Supposing that the control is then switched on, SM1 and SM2 become alive. The shutter motor and the warning bell (to indicate that the shutter is neither open nor closed) will be energized through the first micro-switch (MS1 in Fig. 11), closing the shutters while the bell rings. When the shutters are fully closed this micro-switch will operate, cutting off the supply to the bell and lighting the " Closed " pilot lamp on the control panel.

The main shutter relay may be energized by pressing the " Portal Open " button. Four sets of contacts will then operate as shown in the functional circuit diagram (Fig. 11). Considering these in order from the right (as shown in the diagram), the first is a hold-on interlock to keep the relay energized when the button is released. The second energizes the reversing winding of the motor which opens the shutters. The third connects the warning bell in parallel with the motor, so that it rings while the shutters are moving. The supply to both the motor and the bell is cut off by the operation of the micro-switch MS2, which operates when the shutters are fully open. This micro-switch operates the P relay and lights the " Portal Open " pilot lamp on the control panel.

The shutters may be closed again by breaking the circuit supplying the solenoid of the main shutter relay. This will occur when the " Portal Closed " button is pressed or when the door interlock opens or when the timer reaches zero, if the last two are connected in this circuit.

6.4. *Stabilization*

Since an X-ray tube operates at saturation and its anode current is thus directly proportional to the electron emission of the filament, the filament voltage is extremely critical, a 15 per cent. voltage drop reducing the anode current to zero while a 5 per cent. drop in X-ray tube filament voltage might reduce the tube current by 50 per cent. This alone would decrease the photographic effect by 50 per cent., but if accompanied by a reduction of anode voltage, as it will be if the drop is due to mains fluctuation, there will be an additional loss. Since the penetrating power of the X-rays decreases with anode voltage, the film blackening will decrease also. The photographic blackening is a complex function of the anode voltage and several other factors and frequently varies as the third or fourth power of the voltage. Thus, a 15 per cent. decrease in tube voltage may produce a 50 per cent. loss in film blackening. The two effects combined give a loss of blackening of some 60 per cent. for a 5 per cent. reduction in supply voltage.

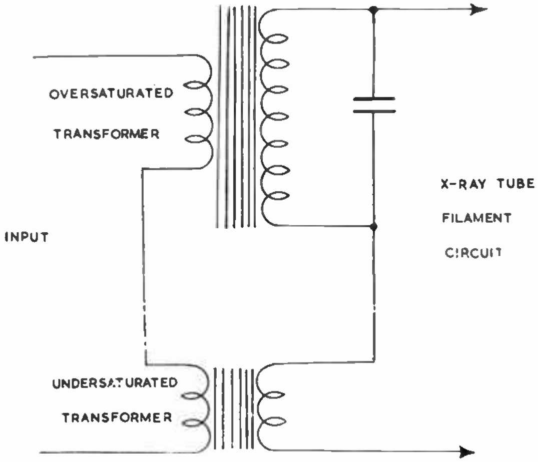


Fig. 12.—A typical filament voltage stabilizer.

To compensate for voltage drops resulting from the sudden load taken by the X-ray set is not difficult. The selection of boosting turns on the auto-transformer is quite simply achieved. Selection may be automatic, but since it depends on the regulation of the mains supply, manual compensation is sometimes more convenient.

A typical automatic filament voltage stabilization circuit is shown in Fig. 12. Two transformer primaries are connected in series with the input. The one that supplies most of the output voltage operates with a high flux density. It is also partly resonated by means of a capacitor. The other transformer operates with a low flux density and in phase opposition to the first transformer.

A change in line voltage affects the oversaturated transformer less than the other and whatever change in output does occur in the secondary of the first is compensated by the bucking action of the second. This can be adjusted to provide very satisfactory filament voltage stabilization.

This system, which is simple and relatively cheap, has one grave disadvantage. Since it depends partly on resonance it is highly frequency sensitive. While, by change of component values, it may be adapted to either 50 or 60 c/s mains supplies, any fluctuation in mains frequency will result in most serious changes in filament voltage.

A more elaborate stabilization system, and one which is not frequency conscious within reasonable limits, is the tube current stabilizer incorporated in the typical industrial set under discussion. This circuit offers complete filament stabilization not only against the voltage drop resulting from mains and auto-transformer regulation, but also against independent fluctuations in supply voltage.

This stabilizer, which is shown in its functional essentials in Fig. 13, depends on a feedback circuit in which the filament voltage is automatically decreased by any increase in X-ray tube current. This holds the X-ray tube current within 1 per cent. against mains voltage fluctuations of plus or minus 10 per cent., although there is a time lag of a few cycles between the fluctuation and the correction. This is unavoidable since the tube current is dependent on the filament heat, so that some hysteresis will occur between a drop in filament voltage and the resultant decrease in tube current.

The XF and XFC tappings on the auto-transformer are shown in Fig. 9. These are connected to the primary winding of a transformer within the h.t. generator tank, the secondary of which supplies the X-ray tube filament. All control of the filament temperature is effected in this primary circuit. There are two control items in the circuit: one a potentiometer by which one of the tappings may be adjusted

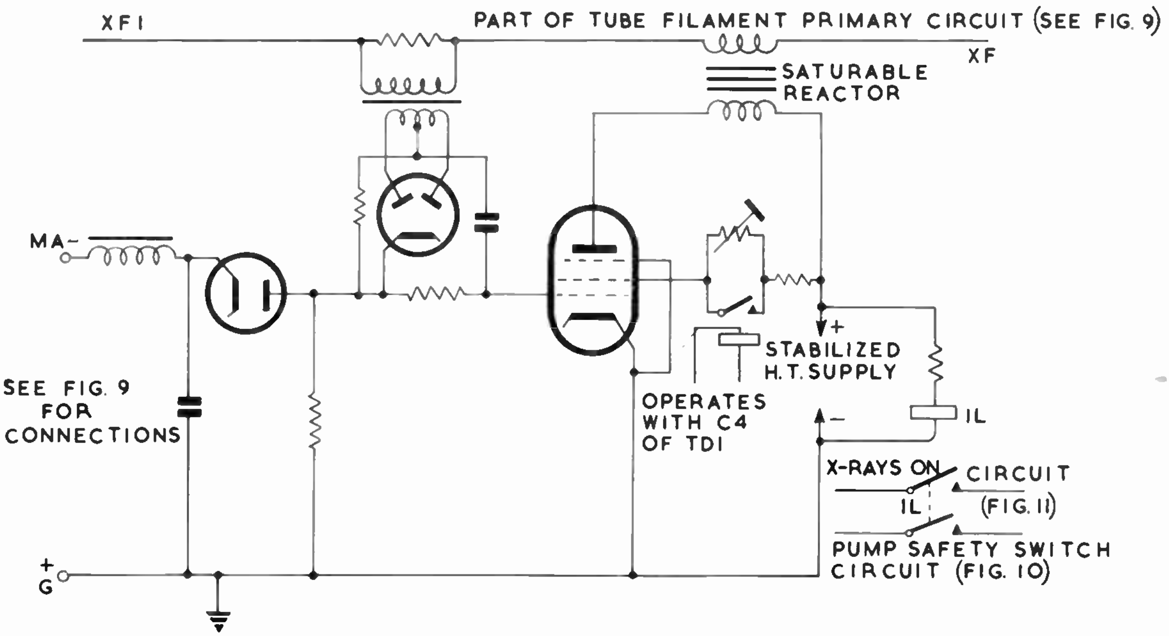


Fig. 13.—Functional diagram of tube current stabilizer.

(AT) and the other a variable reactor. Variation of the reactance is achieved by varying the d.c. in another winding and thus varying the extent to which the core of the reactor is saturated.

Obviously an increase in d.c. will tend to increase the saturation of the core, lowering the impedance to a.c. in the other winding and thus causing an increase in filament current resulting in higher filament temperature and higher tube current. The d.c. which saturates the reactor is the anode current in a pentode control valve (see Fig. 13). Control of the tube current is thus achieved by regulating the grid potential of this valve.

A stabilized h.t. supply is derived from a conventional power pack circuit deriving its primary current from the SM1 and SM2 lines. While the " X-rays On " circuit is incomplete the anode current in the valve is maintained at a suitable level by the connection of resistance in the screen grid circuit. As the fourth contactor in the TD1 assembly closes this resistance is substantially reduced and the screen potential raised to approach that of the anode. The control grid then takes full control of the anode current.

The potential of the control grid is derived from the terminals MA and G. Between these is applied the voltage developed across a variable resistance network (see Fig. 9) through which the tube current flows. Thus the voltage applied between MA and G depends upon two variables: the resistance of the network—which is determined by the setting of the tube milliamperage control—and the instantaneous tube current. The higher the tube current, the higher will be the voltage between MA and G and the more negative will the grid of the control valve become. This will reduce the anode current in the valve and tend to reduce the current in the tube.

The potential between MA and G is not applied directly to the grid of the control valve. After a conventional LC smoothing circuit it is applied to a resistor with a neon stabilizer in series. The effect of this circuit is to amplify the percentage voltage change across MA and G.

Consider, for example, 100 volts applied to the circuit. If there is a voltage drop of 70 volts across the neon stabilizer, there will be 30 volts across the resistor, that is between the

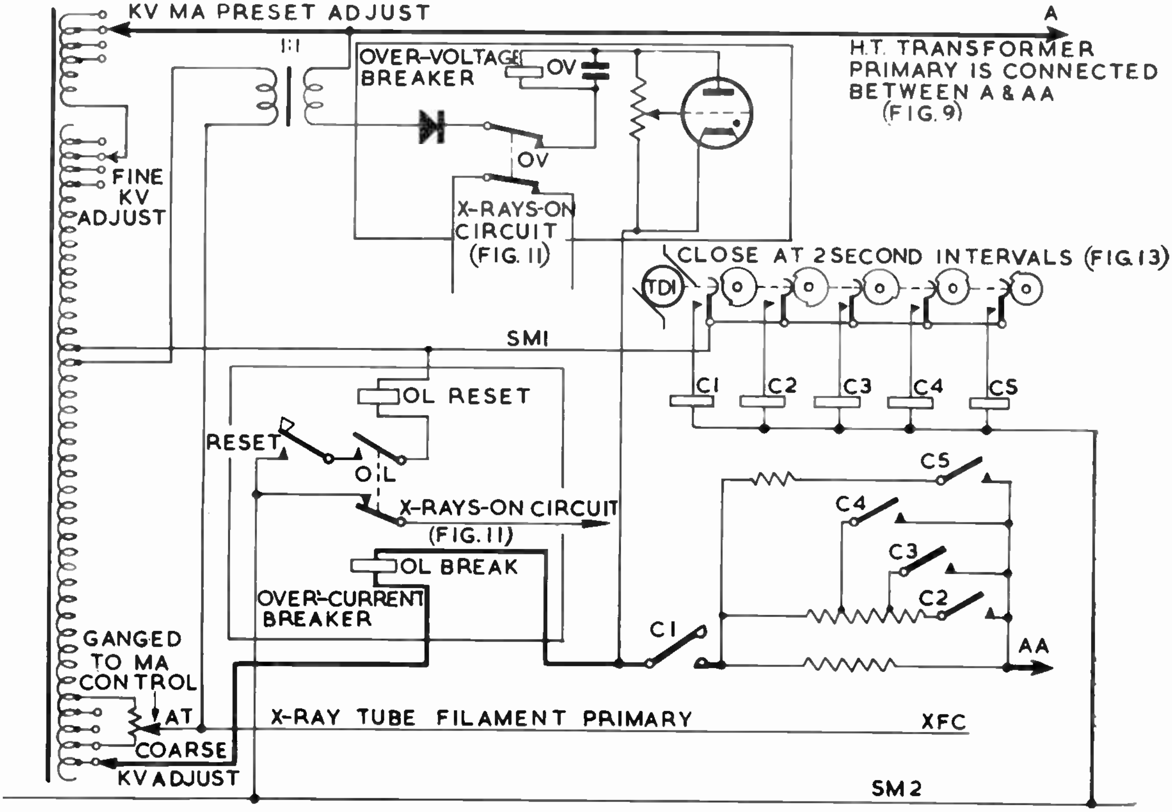


Fig. 14.—Functional diagram of over-current and over-voltage breakers and of TD1 assembly.

grid and cathode of the control valve. If the applied voltage now drops from 100 to 85 volts, a drop of 15 per cent., the neon stabilizer will still dispose of 70 volts, leaving 15 volts across the resistor to be applied to the grid. This gives a reduction of 50 per cent. in the grid voltage for a 15 per cent. change in applied voltage.

The sensitivity of such a system provides a very fine control of the tube current, but under certain conditions oscillation might occur if some form of damping were not provided. The damping circuit provides negative feedback into the control grid circuit whenever a surge occurs in the tube filament primary. The circuit is shown functionally in Fig. 13 and consists of a transformer, a double diode, a capacitor and two resistors. Under conditions of steady current in the tube filament circuit no appreciable current flows through the diode since the only return path for d.c. is through a resistor of high value, thus there is no voltage drop across the resistor in the grid circuit of the control valve. Any sudden surge, however, will produce a current through the diode and thus increase the charge on the capacitor. This charging current flows through the resistor in the control grid circuit, producing a volt drop which lowers the grid potential and counteracts the surge in the tube filament. Once the filament current resumes a steady value, the extra charge on the capacitor leaks away through the two resistors in series and stable conditions are restored. Since tube emission is dependent on filament temperature, the damping is improved by the thermal hysteresis and hunting is eliminated.

A further and virtually incidental function of the stabilizer is the provision of a time delay afforded by the IL relay. This relay operates as soon as the cathode of the double diode rectifier in the power-pack has reached its operating temperature. Two pairs of contacts are closed, one in the " X-rays On " circuit and the other in the pump safety switch circuit. The first ensures that X-rays cannot be switched on until the stabilizer is operating correctly—that is, until all its valves have warmed up and h.t. is applied to the control pentode. It also allows time for the X-ray tube filament to heat up before h.t. can be applied to the anode.

The second pair of contacts is connected in series with the oil pump safety switches, so that time is allowed for the oil to reach the necessary pressure before the switches become effective.

6.5. *Over-Current and Over-Voltage Breakers*

Protection of the h.t. circuits against excessive currents and voltages is of the highest importance, since damage to them can be costly both in replacement of components and in time during which the equipment is out of use.

The over-current and over-voltage protection circuits are shown in Fig. 14. Both of them break the " X-rays On " circuit when they operate, thus safeguarding the X-ray tube and the h.t. circuits against damage.

A solenoid in series with the primary winding of the h.t. transformer actuates contacts in the " X-rays On " circuit, breaking the circuit should the current rise above the prescribed maximum value. These contacts are normally open and held closed by a mechanical latch which the solenoid releases. To reset the breaker a button on the control panel must be pressed, closing the contacts, which are then held closed by the latch.

The over-voltage relay employs a thyratron, which fires if excessive voltage occurs across the h.t. transformer primary, operating a relay which breaks the " X-rays On " circuit.

As the tube current is increased it is necessary to apply boost to the h.t. transformer primary by adding more auto-transformer turns and thus preventing a fall in tube voltage. In order that an increase in this preset kV/mA compensating voltage will not cause the over-voltage breaker to trip, a bucking transformer with a 1 : 1 ratio is connected as shown in Fig. 14. The bucking voltage, in phase opposition to the A-AA voltage and proportional to the supply to the tube filament primary is introduced between the A line and the rectifier of the over-voltage protection circuit. This obviates the effect of the increased preset adjustment on the thyratron and ensures that the preset value is not excessive relative to the selected tube current.

7. Conclusion

The recent growth in the use of X-rays in industry and research has necessitated a vast increase in techniques and corresponding ramifications in circuit design. A paper of this length can only hope to embrace a few of those circuits which are in most general use and thus little mention has been made of apparatus having more specialized applications.

X-ray inspection has now been generally accepted as a standard method of testing castings and welds in light alloys and ferrous metals and it is mainly with apparatus suitable for these uses that this paper has been concerned. The design problems involved in such apparatus differ considerably from those met with in clinical radiography. The differences are largely concerned with the penetrating power and length of exposure time necessary to produce the required contrast on the film in radiographs of the comparatively opaque metal objects involved in industrial work, as opposed to the relatively transparent tissues of the human body.

The choice of X-ray generator depends upon the output required and the economic cost for the job involved. Many of the circuits in more general use have been described.

8. Acknowledgments

Figs. 9, 10, 11, 13 and 14 are reproduced by courtesy of Marconi Instrumentation.

9. References and Bibliography

1. H. Hirst. " X-rays in Research and Industry."
2. W. E. Schall. " X-rays." (Schall, 1951).
3. C. W. Wickens. " Ignitron switching for radiographic timers." *Marconi Instrumentation*, 3, 1951, No. 3, p. 34.
4. C. W. Wickens. " Electronics in X-ray control," *Marconi Instrumentation*, **3**, 1952 No. 4, p. 56
5. J. J. Bliss. " Generation of power for X-ray equipment." *Journal of the Nottingham University Engineering Society*, 1953.
6. M. J. Gross. " Gas Insulation—A new trend in Roentgen-Ray apparatus." Presented at the Sixth International Congress of Radiology, London, 1950.
7. J. J. Bliss. " Control circuits of the industrial X-ray apparatus type TF 1555A," *Marconi Instrumentation*, 4, 1953, No. 3, p. 55

INDUSTRIAL X-RAY FLUOROSCOPY WITH PARTICULAR REFERENCE TO ULTRA-FINE FOCUS X-RAY EQUIPMENT*

by

E. W. Kowol, M.A.†

A Paper presented during the Industrial Electronics Convention held in Oxford during July 1954

SUMMARY

The development of fluoroscopy as an inspection method and as an aid to radiography is discussed and the economics of both applications, their advantages and limitations, considered. A 150-kV mobile fluoroscopic x-ray unit and a 150-kV fine focus unit, which permits an enlargement of the image, are described. Typical users include light alloy foundries, electrical engineering and the food industry. Special problems of fluoroscopy are discussed, such as the protection and fatigue of the operator, and the subjective element which is introduced by the method.

1. Introduction

In spite of the fact that x-rays were discovered by Roentgen through the fluorescence they caused on a screen coated with barium platino-cyanide, their industrial application over nearly 60 years has been concentrated mainly on film radiography. In medicine fluoroscopy has been more widely applied, partly because of its value in the observation of movement and, more recently, fluorography (the photography of fluoroscopic images) has been used in the campaign against pulmonary tuberculosis.

The restriction of fluoroscopy to a comparatively small number of industrial uses results partly from the rather slow development of suitable equipment and partly from certain limitations which have hitherto been characteristic of the technique but are now being gradually overcome.

2. Fluoroscopic Screens

One of the chief problems has been to find a suitable substance for coating fluorescent screens. Roentgen's original screen was coated with barium platino-cyanide and this substance was in use until about 1910. When subjected to x-rays it fluoresced with a brilliant green light but it had the disadvantages of being costly and deteriorating rather rapidly in use. The next coating was zinc orthosilicate and this was in general use for only a few years. These screens also gave a green light and also had two grave faults: they produced afterglow and there was a distinct lack of fluorescence with hard, i.e., higher voltage x-rays. In 1914 C. V. S. Patterson introduced cadmium tungstate screens and these were in use for almost 20 years.

Cadmium tungstate gives no afterglow and can be used satisfactorily for long periods without marked deterioration. It fluoresces with a bluish white light but is wanting in brilliance. In 1933 L. Levy and D. W. West developed zinc-cadmium screens which are incorporated in most present-day fluoroscopic units. A tendency to afterglow was eliminated by the addition of one part in two million of nickel. These screens fluoresce with a yellowish-green light ten times as bright as that given by the older cadmium tungstate screens.

Research and development on fluoroscopic screens is going on all the time, and in 1951 D. T. O'Connor and D. Polansky, of the U.S. Naval Ordnance Laboratory, Spring Field, Maryland, published results of research into screen brightness.[1] Twenty different commercially available screens were tested and four selected for closer study. The brightness differences ranged from 2·1 to 4·0 foot-candles at 80 kVp (120 roentgens per minute) and from 2·6 to 6·8 foot-candles at 100 kVp (170 roentgens per minute). They also published data on the points at which maximum brightness was reached. Zinc sulphide, zinc silicate and cadmium tungstate screens reach maximum brightness at 120 kVp, whereas barium platino-cyanide and calcium tungstate screens reach it at 150 kVp.

*Manuscript received May 26th, 1954. (Paper No. 301.)

†Formerly with Solus-Schall Ltd., London.

U.D.C. No. 621.386.832 : 620.179.

A likely future development is that a variety of fluoroscopic screens, each suitable for its own range of applications, will be available for fluoroscopy, in the same way as different types of x-ray film are available for radiography.

3. Fluoroscopic X-ray Units

In essence the requirements are simple, an x-ray tube and a fluorescent screen on which x-rays fall after passing through the specimen being examined.

X-ray tubes and high-tension generators in fluoroscopic units are of the standard types commonly used in radiographic units. However, as a fluoroscopic unit must be able to work continuously at its maximum rating for long periods, an efficient cooling system is essential.

The most marked differences in the design of commercially available fluoroscopic sets are found in the viewing arrangements. These vary considerably with the particular application for which the unit is intended.

As the method requires the presence of the operator near the equipment whenever it is in operation, special attention has to be given to protection from radiation and, in order to speed up inspection, mechanical handling and rejection devices have in many cases been incorporated in the units. To provide for various uses under different conditions a variety of arrangements of shutters, mirrors, jigs and turntables are available. The use of mirrors is not generally recommended, however, because to view the image in a mirror causes a loss in brightness of 20 to 40 per cent. or more.[1] In any case, if a mirror is used, its edges must be hermetically sealed to prevent oxidation and further loss of light.

A unit which was developed during the last war for the fluoroscopic scanning of shell fuses is now being produced, with certain modifications, for use in many other industries where large quantities of specimens have to be examined but where finest detail is not needed. The set operates up to a maximum voltage of 150 kV, and the tube with a 2·8-mm line focus is oil cooled for continuous running. The set is mounted on a trolley so that it can be wheeled to any point of the production line where

Fig. 1.—A 150-kV mobile fluoroscopic x-ray unit.

fluoroscopic viewing is required. Usually, if the equipment is to be fixed at one particular point on the production line, conveyor belt feeding and an automatic rejection mechanism are incorporated. For hand feeding there is a door at the left-hand side of the cabinet. This door is coupled to a lead diaphragm which cuts off x-rays as soon as the door is opened. This prevents the operator from accidentally exposing his hand to radiation. The cabinet is lead lined and the window is made of lead glass 25 mm thick. Inside the cabinet is a turntable upon which the specimen is placed. It can be raised, lowered and rotated during inspection, its movement being controlled by handles below the viewing cabinet. The specimen can also be moved closer to and away from the x-ray tube. The operator sits in front of the screen which is provided with an eyepiece. A lead-leaf diaphragm fitted in front of the x-ray tube can be adjusted by means of levers outside the cabinet to provide masking and prevent glare from the area surrounding a specimen. The control panel is mounted on a vertical column and can be adjusted to any height convenient to the operator (Fig. 1).

Most commercially available fluoroscopic units are not at present suitable for the detection of very fine defects, but an interesting development is the ultra-fine focus unit designed by R. Seifert, Hamburg (Fig. 2). This unit has an x-ray tube with a focal spot only 0·2 mm in diameter and operates at 150 kV constant potential. Because the focal spot is so fine, sensitivity is materially improved and the image can be magnified up to seven or eight times its normal size without any visible loss of sharpness.

Both the units described have slots in front of the screen where a film cassette can be fixed. Thus, if the observer is doubtful about a particular specimen, he can take a radiograph of it.

The selection of a suitable unit depends upon the application for which it is intended. For instance, the fine focus unit would not normally be used in a food factory but would be a good choice where small assemblies are to be screened. In both cases fluoroscopy can be used an an independent inspection method, but in a light alloy foundry fluoroscopy should only be regarded as an extremely valuable method of bringing down the cost of radiography. In this case, it is not used as an independent but as an auxiliary inspection technique.

4. Advantages and Disadvantages

Fluoroscopy has undoubtedly three distinct advantages which make it an attractive commercial proposition as an inspection method for industry.

In the first place, the rate of inspection is high and compares very favourably with that of radiography. L. R. Curtis[2] reports an average speed of 189·6 light alloy castings per hour while testing 20,000 castings of various shapes. In some cases several views of each casting were required, but even so the lowest rate was 28·2 per hour. With simple shapes the rate of inspection was 371. The castings were loaded on a conveyor belt by one man and viewed by another. When more than one view was needed

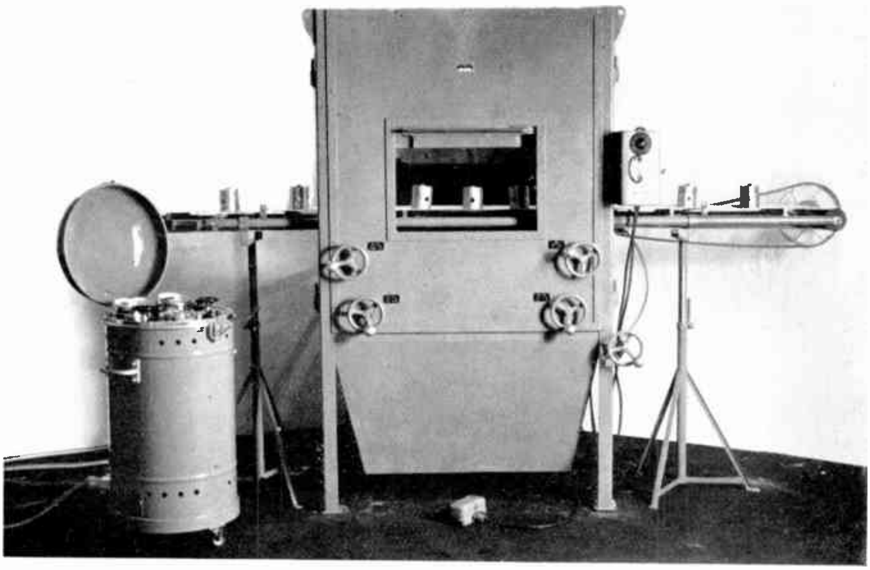


Fig. 2.—Fine focus fluoroscopic x-ray unit.

all the castings were viewed from one angle first, reloaded in a different position and viewed again.

The second advantage is that fluoroscopy is cheap and saves labour, materials and space, all of which are necessary to standard radiography. Outstanding savings are made in cases where the specimen to be inspected would have to be radiographed from more than one angle. The rotation of the specimen in the viewing cabinet allows the operator to see it from every angle and gives him the opportunity to find faults which might be missed if an insufficient number of radiographs were taken. Field tests were carried out by B. Cassen and D. S. Clark,[3] of the California Institute of Technology, with 10,000 aluminium alloy castings. These were radiographed and fluoroscoped and a total of about 14,000 defects was found. Five thousand of these were very small and were found by radiography only. On the other hand, about 1,700 larger defects spotted fluoroscopically were not picked up by radiography. If the castings had been radiographed from more angles, no doubt all of these and other defects as well would have been discovered. The cost would, however, have been augmented by at least the price of 10,000 films and additional processing.

Thirdly, the operator does not require prolonged or difficult training and, as the specimens are light and are generally handled mechanically, a woman is as suitable as a man. Good eyesight and a reasonable standard of intelligence are, of course, necessary.

Unfortunately, fluoroscopy has inherent limitations which have hitherto confined its application to a fairly narrow field.

First and foremost, there is no permanent record of the x-ray image seen by the observer. (If the image is photographed, the technique can no longer fairly be described as fluoroscopy; it becomes fluorography.) A good radiograph is a valuable record of the internal soundness of the article. It can be shown to the buyer, inspected by the insurer's surveyor and filed for reference.

The second serious weakness of the method is that sensitivity, definition and contrast of the image are below the standard obtained by conventional radiography. Recent developments in the design of screens and fine focus tubes have reduced the discrepancy between the two techniques, but because a higher tube voltage must always be used in fluoroscopy than would be needed in radiography, the technical inferiority of the fluoroscopic image will persist.

Thirdly, because the fluoroscopic screen does not accumulate the effect of x-rays as a film emulsion does, fluoroscopy can be applied only to comparatively thin metallic specimens or thicker objects of lower density. Moreover, fluoroscopy could not be used on site work such as the inspection of large welded structures.

Finally, fluoroscopy depends more than radiography on the judgment of one individual. This is felt to be an important limitation in certain applications and is discussed more fully later.

In many cases the limitations of fluoroscopy are of such importance that industrial users have been obliged to forgo the economic advantages, concentrating instead on radiography which is in comparison costly and time consuming. There are, however, a number of applications, particularly in light alloy engineering where the use of the two inspection methods in combination results in the user profiting from the advantages of both and at the same time discounting their disadvantages.

A number of industrial users screen parts for which 100 per cent. radiography is stipulated and are able to eliminate all those specimens which have gross defects. The speed of fluoroscopic testing is high, because the operator is not looking for fine defects, and in turn the radiographer is not wasting skilled man-hours and material on castings which contain gross defects. It does not normally matter that there is no record of castings with gross defects because, in any case, they would have to be scrapped. If, however, the incidence of gross defects is high and it is thought that the foundry technique should be altered, it will be useful to have radiographic records of castings rejected fluoroscopically for study in the foundry. In this case, fluoroscopy can be suspended temporarily and radiography of all castings carried out until a satisfactory foundry technique has been established.

5. Fluoroscopic Examination of Light Alloy Castings

More x-ray sets are used in this country and more film consumed on the inspection of light alloy aircraft parts than on any other single industrial application. The development of safer aeroplanes, which at the same time are light in weight and consequently can carry heavier loads and larger fuel supplies, is due in no small part to the use of radiography, indeed it is doubtful whether it could have been achieved without it.

The weight of an aircraft has been substantially reduced by the introduction of strong light alloys, but even so, much more metal would have to be used to ensure operational safety if radiography were not available to guarantee the soundness of each part. Moreover, there is a direct saving in metal and in man- and machine-hours which would otherwise be wasted on faulty castings, and an indirect saving due to the help derived from radiography in improving foundry techniques. Obviously, radiography is a sound commercial proposition, but the fact remains that, compared with other inspection methods, radiography is expensive. A film 10 in × 8 in costs two shillings, and several films, sometimes of a larger size, may be needed for a single part. To this must be added the darkroom and processing cost, the radiographer's time and the running of the x-ray set. Clearly, in the same way as machine time must not be wasted on unsound castings, radiography should not be wasted on detecting defects which could be found more quickly and cheaply by fluoroscopy.

Fluoroscopy has not been widely used, chiefly because of its reputation for poor sensitivity. This reputation is, however, founded on very early experience when sensitivity of the order of 15 per cent. was regarded as normal.*

Nevertheless as long ago as 1930 in Germany, Dr. R. Berthold[4] produced figures for defect sensitivity for aluminium and copper, using zinc silicate screens.

In calculating sensitivity it is usual to use a penetrameter of which various types are available. Those are made of the same metal as the specimen examined and are normally graduated in thickness. It should, however, be noted that the sensitivity obtainable with a penetrameter does not necessarily correspond to the defect sensitivity. This is largely due to the regular shape of the penetrameter.

The results in the Tables were obtained under the following conditions:—

1. Constant potential
2. Size of illuminated field: 5 cm. dia.
3. Focus—screen distance: 50 cm.
4. Tube current: 10 mA.

Dr. Berthold's conclusions are that for general application the following sensitivities were valid:—

5% for aluminium up to 70 mm thick.
7% for copper up to 8 mm thick.
7% for steel up to 15 mm thick.

*% Sensitivity
$$= \frac{\text{Smallest detectable thickness difference}}{\text{Thicknesses of specimen being examined}} \times 100.$$

Aluminium

Thickness (mm) ..	5	10	20	30	40	50	60	70
Tube voltage (kV) ..	65	75	95	105	115	125	140	165
Smallest visible thickness difference	0·3+	0·5—	1·0+ 2·0×	1·5+ 1·0—	1·5— 2·0+	2·0+	2·0— 3·0×	3·0+

Copper

Thickness (mm) ..	2	4	6	8
Tube voltage (kV) ..	105	130	150	170
Smallest visible thickness difference	0·1— 0·2×	0·25+	0·4×	0·5×

— Hardly perceptible
+ Quite easily visible
× Very clear.

Ten years later Dr. Berthold produced a penetrameter sensitivity curve for aluminium up to 50 mm thick,[5] shown in Fig. 3.

It is interesting to compare these results with those obtained in America. In 1944 M. Iona reported[6] that he obtained sensitivity better than 5 per cent. at 150 kVp with aluminium 2 in thick, magnesium 3 in thick, and steel $\frac{3}{8}$ in thick. With 170 kVp the sensitivity was about 6 per cent. In the same year R. W. Mayer[6] stated that "a sensitivity of 5 per cent. or better can be counted on as regular day-in day-out production standard."

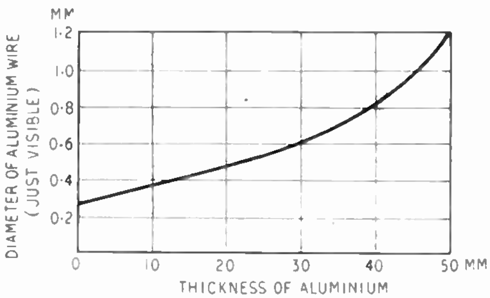


Fig. 3.—Penetrameter sensitivity curve for aluminium. (After Berthold.[5])

L. R. Curtis[2] conducted some interesting experiments in 1948 into the problems of fluoroscopic inspection of aluminium and magnesium castings. He found that spherical cavities are the most difficult to detect, cylindrical cavities somewhat easier, and washer-type cavities the easiest of all. Ten observers were used, only one of whom was experienced. The average sensitivity was 6·667 per cent., while the experienced observer obtained 4·25 per cent. The best sensitivity recorded was 1·5 per cent. in the detection of washer defects by the experienced observer, but an inexperienced observer was able to detect them with a sensitivity of 2 per cent.

D. T. O'Connor and D. Polansky[1] carried out experiments on fluoroscopic penetrameter sensitivities and published their results in 1951. At 96 kVp, 24 mA, 2 per cent. was obtained with $1\frac{3}{4}$ in of aluminium and 1·5 per cent. with 2 in of magnesium at 90 kVp, 27 mA.

It will be seen from these figures that there has been a steady improvement in sensitivity over the last 20 years and, although fluoroscopy is not generally comparable with radiography, where a penetrameter sensitivity of 2 per cent. or better is required, it is sufficiently advanced to merit very serious consideration as a preliminary inspection tool for light alloy castings.

The differences in results are somewhat disconcerting but they can be explained by the wide variations in the equipment used, the conditions under which experiments took place and the lack of uniformity in penetrameters used for assessing sensitivity.

The same reasons would account for the equally diverse data available on the thicknesses which can be successfully penetrated.

6. Fluoroscopic Inspection of Electrical Components

In the electrical industry, where the lack of a permanent record is not a serious disadvantage, the applications of fluoroscopy are extremely varied, although it is mainly used to examine the position and condition of wires embedded in insulating material. As the wires are of higher density than the surrounding material and therefore absorb more radiation, they appear on the screen as darker lines and are easy to observe. Breaks, misplaced or bunched wires can be seen without difficulty. In the case of very small and delicate assemblies where accurate positioning is essential, the enlarged image given by a fine focus unit is particularly useful.

In the electrical industry, moreover, it is frequently necessary to view an assembly from several angles so that the use of fluoroscopy instead of radiography considerably reduces inspection costs.

In the United States a special fluoroscopic unit employing two x-ray tubes, two fluoroscopic screens and an arrangement of mirrors, has been designed for the inspection of electric cables for concentricity of core and uniformity of insulation. The x-ray beams penetrate the cable at right angles to each other and the operator sees a double image. Any variations in either dimension are easily observed.

The following applications give some idea of the range of the usefulness of fluoroscopy. Assemblies such as metal-jacketed radio and transmitter valves are screened for alignment of the parts before evacuation. The cathode-grid distance is, of course, important and this can be measured with sufficient accuracy on the screen image. Switchgear is inspected for broken parts and loose connections; heater elements for position and continuity of the resistance wire; carbon electrodes for concentricity of the core; dry batteries for correct spacing of parts. Telephone equalizer units and capacitors in metal containers are also suitable for fluoroscopic inspection. The faults which can be picked up include incorrect alignment, voids in the sealing compound and broken soldered joints.

Another field of application is the inspection of insulating materials. Ceramic, vulcanite, bakelite, polythene and mica insulators are viewed for the presence of foreign inclusions or cavities which might detract from their efficiency.

In this country some 20 electrical engineering firms regularly use fluoroscopy for routine inspection.

7. Fluoroscopic Inspection of Foodstuffs

The field in which fluoroscopy has its widest application in the United States is the food industry, and it seems likely, particularly in view of increasing parliamentary and public interest in the purity and cleanliness of food, that a similar situation will soon be established in Britain. There is no doubt that food packers have a great many problems which can be solved by fluoroscopic inspection of prepared foods in tins, foil, bottles, paper and other packages.

One of the commonest complaints which the suppliers receive from the general public is that some kind of inclusion is present in a tin or other package of food. These include an enormous variety of likely and unlikely objects, but unfortunately it is only rarely possible to detect such foreign bodies as cigarette ends, pieces of string and paper and splinters of wood. However, almost all the other common inclusions can be seen on the fluorescent screen by an operator with normal vision and experience in viewing large batches of packed foods.

In addition to the detection of inclusions, food manufacturers can use their fluoroscopic equipment to check that containers are properly filled. This is particularly useful in cases where each packet contains a number of separate ingredients.

Some manufacturers inspect their entire output as a matter of routine and use a conveyor belt to channel the goods through the viewing cabinet. In other cases only suspect batches or a percentage of the output is examined and in such circumstances, where numbers are small, hand feeding is usual. Where fluoroscopy is only required occasionally manufacturers can make use of hired equipment and the services of an experienced operator.

Apart from these three main fields of application fluoroscopy is used on a wide variety of other products. The range extends from packages of razor blades, which are checked to make sure that no chips have been caused by mechanical packing, to golf balls for location and roundness of the core.

8. Protection of the Operator

As fluoroscopy is mainly used for mass inspection and the observer is close to and directly in front of a powerful x-ray beam for long periods every day, it is absolutely essential that he should be completely protected from radiation. Even a very small leakage may have serious effects on the operator if he is subjected to it for a long time and it is important to check the efficiency of protection as a matter of everyday routine. A well-designed unit, however, gives the operator complete protection and prevents him from carelessly exposing any part of his body to radiation. The measures which are taken to prevent accidental exposure of the hands to x-rays are mentioned in Section 3. Where conveyor belt feeding is used, loading takes place at a safe distance and a lead-lined tunnel covers the conveyor belt at both sides of the equipment. The observer's head is protected by a sheet of lead glass equivalent to at least 2·5 mm of lead for a 150-kV unit. In the United States protective screens made of a plastic material filled with a strong solution of lead perchlorate have been successfully used. These screens are 3 in thick for a 150 kV set and give the same protection as would be afforded by $\frac{1}{4}$ in of metallic lead. They have the advantage over lead glass that they do not become discoloured by the exposure to x-rays.

Discoloration is a serious problem. According to M. Iona[6] as much as 30 per cent. loss in light transmission can be caused by the discoloration due to the exposure of glass to high-voltage x-rays during one day. The glass can be cleared by heating it to 200°-250°F or by exposing it to long wave x-rays.

9. Operator Fatigue

As the operator has the sole responsibility for detecting flaws where fluoroscopy is being used as an independent inspection method, he must be given conditions of work which do not in any way tend to make his judgment unreliable. The fluoroscopic unit should be in a well-ventilated, dimly lit room. The operator should be allowed 5 to 10 minutes to accustom his eyes to a low level of illumination.

It is advisable to change observers frequently. A spell of one hour at a time is generally recommended and during breaks the observer should be in the daylight but wearing dark glasses. In many cases two observers take it in turns to load. Comfortable seating accommodation at the screen should be provided and the screen itself should be set at an angle which the observer can view from a sitting position. The lead glass should be cleared frequently as the discoloration causes eye strain. Observers often find it less tiring to view slowly moving objects rather than stationary ones. Enlargement of the image is also a help, but only if it can be achieved without too much loss of sharpness. For this reason the fine focus unit will be found less tiring.

10. The Subjective Element in Fluoroscopy

With fluoroscopy, in the same way as with many other methods of non-destructive testing, the judgment of the observer is all important. Having selected and trained a suitable person and provided him with good equipment and viewing conditions, the employer must then delegate to him the responsibility for passing or rejecting the goods. The case is similar to that of magnetic or ultrasonic testing and differs from radiography in that a snap decision by one person is sufficient to reject a specimen. There is no permanent record against which the observer's accuracy can be checked but it is possible to make periodic spot checks by means of prepared samples containing flaws of various types and sizes. Where fluoroscopy is being used as a preliminary to radiography there is, of course, a constant check on the reliability of the operator, and in any case his responsibility is limited to deciding which castings are to be radiographed.

The obvious way of reducing the dependence of fluoroscopy on the judgment of the operator is to obtain a permanent record of the pictures which appear on the screen.

11. Industrial Fluorography

During the past 20 years a considerable amount of research has been carried out into the practicability of photographing fluoroscopic images on small-size films. In the medical field the technique is now highly developed and is widely used in mass radiography, but for industrial use it is still in the development stage.

Fluorography has obvious economic advantages over radiography. Far less film is used and there is a substantial reduction in the amount of space needed for storing records. Moreover, a high rate of inspection can be maintained over long periods.

During the war when speed of examination was vital and x-ray film and skilled labour were in short supply, industrial fluorography became the subject of serious research. It was clear that if fluorography could provide a comparable standard of sensitivity, it would be an attractive alternative to industrial radiography.

The basic difference between fluoroscopy and any method which uses a film is that whereas the film emulsion accumulates the effect of light, the human eye does not. Consequently the fluoroscopic screen must be brighter for viewing than is necessary for photography of the screen. In order to maintain the level of brightness necessary for viewing a higher voltage must be used than would be necessary for conventional radiography or fluorography and the screen contrast is therefore lower.

In Germany Dr. R. Berthold[7] photographed fluoroscopic images at a level of screen brightness appreciably lower than would be necessary for viewing and found that he obtained much higher penetrameter sensitivity than was possible fluoroscopically. He quotes a case where an aluminium casting was fluorographed at 65 kV while 100 kV was necessary for fluoroscopy. In this case the screen brightness was reduced to one-sixth of the level required for fluoroscopy, but, if necessary it can be reduced to as little as one-twentieth and even then only a 30-seconds exposure is needed. Such a technique could be used for fluorography of sections too thick or too dense for fluoroscopic viewing.

One of the major problems was to develop a screen with a much finer grain size than had hitherto been in use for fluoroscopy. As screen brightness could be reduced, thinner coatings of finer grain size could be used and a black backing was substituted for the normal white base. A fine-grain film, sensitive to the colour of the fluorescent light and fine-grain developer gave added sensitivity. Even so the conclusion which Dr. Berthold arrived at in 1943 was that fluorographic sensitivity was not comparable to that reached by conventional radiographic techniques except in cases where he photographed the enlarged image thrown on a fluorescent screen by an ultra-fine focus x-ray tube.

In the United States, Holland and the United Kingdom research is continuing and an improved screen has been developed. Progress has also been made in the design of cameras. Penetrameter sensitivity of 1 per cent. achieved under laboratory conditions has been reported.

Further research and development will be needed before the technique is sufficiently advanced to be accepted as an alternative to radiography.

12. Conclusions

Fluoroscopy in its present state is an inspection tool which can be used to ensure quality of goods for which radiography would be wasteful and where no permanent inspection record is needed. Especially in the light alloy foundry it can contribute to lower production costs by use as a preliminary to radiography.

Industrial fluorography has not yet emerged from the research laboratory but it seems likely that the technique will be applied on the shop floor within the foreseeable future. It must not be concluded, however, that fluorography will replace radiography in industry, any more than it has in medicine. It will have its own part to play in the non-destructive testing of materials.

13. References

1. D. T. O'Connor and D. Polansky, "Theoretical and practical sensitivity limitations in fluoroscopy," *Nondestructive Testing*, **10**, pp. 10-21, Fall 1951.
2. L. R. Curtis, "Fluoroscopic inspection of aluminium and magnesium castings," *Nondestructive Testing*, **8**, No. 3, pp. 25-30, Winter 1948.
3. B. Cassen and D. S. Clark, "Fluoroscopic examination of metallic objects," *Industrial Radiography and Nondestructive Testing*, **5**, No. 3, pp. 34-37, Winter 1946.
4. R. Berthold, "Grundlagen der Technischen Roentgendürchstrahlung," p. 69 (Verlag J. A. Barth, Leipzig, 1930).
5. R. Berthold and Otto Vaupel, "Zerstoerungfreie Werkstoffpruefung mit Roentgen- und Gamma-strahlen," p. 59 (Verlag Otto Sale, Frankfurt, 1944).
6. M. Iona, "New developments and applications of the fluoroscopic examination of metals and assemblies," *Industrial Radiography*, **2**, No. 4, pp. 20-28, Spring 1944.
7. R. W. Mayer, "Industrial fluoroscopy," *Industrial Radiography*, **2**, No. 4, pp. 37-39, Spring 1944.
8. R. Berthold, *Luftwissen*, **9**, No. 6, 1943.

DISCUSSION ON

"Industrial Applications of X-rays"*

Chairman: H. G. Foster, M.Sc. (*Member*)

Dr. A. Nemet (*Member*): Fluoroscopic examination in industry covers a rapidly expanding field and the most interesting, perhaps, of all is that of metals, particularly aluminium and steel. One of the snags with present methods is that the brightness of the screen is so very limited that the examination has to be carried out in total darkness. Also, in order to get the maximum brightness, the screen has to be a coarse-grained one, limiting the sensitivity of the method expressed in percentage of the total material thickness.

A fundamentally new method has been developed for intensifying the fluoroscopic x-ray image, and Fig. A shows the cross-section of the image intensifier tube. The total brightness intensification is of the order of 1000 : 1 and because of this large gain one can view the screen in dim daylight. The eye makes use of cone vision at the same time instead of the insensitive rod vision that is effective

*Discussion Meeting No. 7. Report of proceedings during the second part of Session 2 of the Industrial Electronics Convention on July 9th, 1954.
U.D.C. No. 621.386 : 67 2.

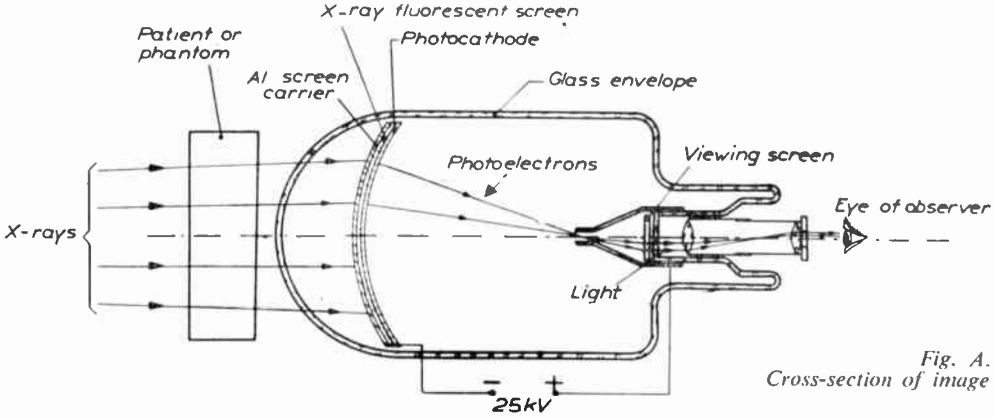


Fig. A.
Cross-section of image intensifier tube.

in normal fluoroscopy. Moreover, the screen has a fine grain and thus a resolution of 0·3 mm instead of the normally used fluoroscopic screens (0·8 mm). Fig. B shows a photograph of the tube.

Measurements show that the image intensifier affords very nearly radiographic sensitivities in fluoroscopy, especially when thick objects are viwed.

This tube opens up new possibilities for using x-ray fluoroscopy in production and development in cases where present-day fluoroscopy is not good enough and radiography is too expensive.

In conclusion, I would like to ask Mr. Miller what the electron energy of a linear accelerator is corresponding to minimum absorption in steel. We have seen a theoretical curve showing a minimum at about 9 MeV. Is this curve confirmed by experiments? I have an American publication with me referring to Betatron radiography, according to which the minimum is shown at 20 MeV. Could he explain the discrepancy?

C. W. Miller* (*in reply*): The theoretical curves of Fig. 2 are for monochromatic radiation and the energy scale shows, of course, the quantum energy of the radiation. It is shown in the paper (Section 2.2 and Fig. 3) that when x-rays are produced, the mean quantum energy of the radiation is appreciably below that of the energy of the electron stream producing the x-rays. The specification of an electron energy to give minimum absorption in practical radiography depends on several factors and, indeed, provided the electron energy is above 9 MeV, the radiation of minimum absorption will be selected when thick specimens are dealt with. Variation of electron energy then only varies the relative proportions of radiation of different energies and one might define an optimum from this point of view when electron energy is such that the peak of the radiation spectrum shown in Fig. 3 occurs at approximately 9 MeV.

Since the 20 MeV quoted for Betatron operation to give minimum absorption for steel will certainly refer to electron energy it can be seen from the general shape of Fig. 3 that it is not at variance with a monochromatic quantum energy of 9 MeV shown in Fig. 2.

Apart from differences in target construction leading to some difference in the shape of the x-ray spectrum, one would expect little difference between Betatron and linear accelerator as regards minimum absorption and an electron energy approaching 20 MeV might be expected to correspond to minimum absorption in both cases. Too much emphasis should not, however, be placed on the question of minimum absorption, partly because other factors have to be considered, as explained in the paper, and partly because the variation of absorption with energy is, in the relevant range, very slow. Under practical radiographic conditions the half-value layers for steel are for a 4-MeV and an 8-MeV accelerator 1 in and 1·14 in respectively, and increase of electron energy would only slightly improve these figures.

The curves of Fig. 2 can be checked by using monochromatic radiation from radioactive substances or using very thick specimens or filtration with x-ray generators, thus filtering out the softer

*J.Brit.I.R.E., 14, p. 361, August 1954.

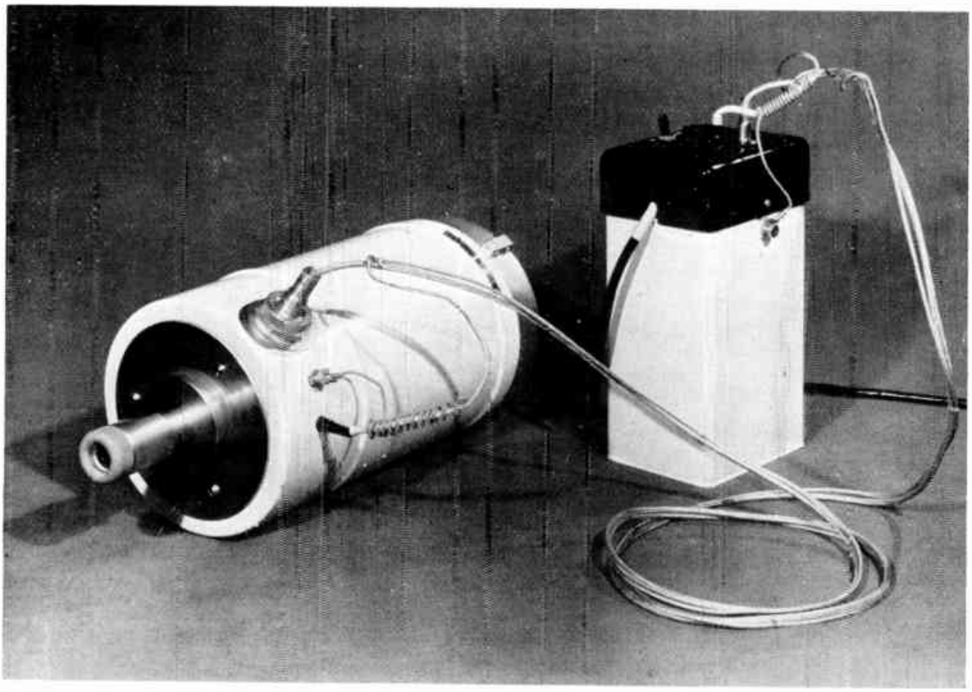


Fig. B.—Image intensifier tube and power supply.

radiations. In either case the effect of scattered radiation must be eliminated by good collimation.

G. Frater: Regarding fatigue of the operator, particularly in fluoroscopy, it should be possible to apply some form of television scanning system to identify flaws in material in the same way as it is possible to identify flaws in signatures of cheques. On scanning the screen a series of pulses would be obtained. One could thus detect some kinds of defect which occur by a less fatiguing method, by comparing the pulses with others produced by a standard specimen, through suitable coincidence and alarm circuits.

A. Hewett-Emmett* (*in reply*): There is, at the moment, work going on whereby a television type of camera, not sensitive to light but sensitive to x-rays, will be used. The camera transforms incident x-ray quanta into photons, and hence electrical signals. The scanning of an object that does not conform to standard, and its rejection is then fairly simple. With conventional x-ray intensities, low screen illumination, after-glow of the screen, and present handling speeds, it would be very difficult to mechanize the rejection of a product. Image intensification offers far more possibilities in automatic rejection.

*Mr. A. Hewett-Emmett presented the paper at the Convention on behalf of Mr. Kowol.

F. L. Steghart: I would like to ask Mr. Bliss to explain the reason for obtaining better pictures from soft radiation when dealing with material of small opacity.

J. J. Bliss (*in reply*): The fluctuating potential used for medical diagnosis produces a heterogeneous x-ray beam consisting of both hard and soft x-rays. The half-value layer of such a beam, that is the number of millimetres of some standard material such as copper required to reduce the intensity by half, is thus comparatively low and more detail appears on the film when radiographing the human body, in which opacity is low and differences in opacity only slight.

The industrial equipment described in the paper employs constant potential which produces an x-ray beam with a far higher proportion of hard x-rays and a correspondingly higher half-value layer. This provides better definition when dealing with comparatively opaque workpieces when contrast is between solid metal and gas- or slag-filled cavities. Any soft x-rays emitted would be valueless as they would not penetrate to the film.

A SIMPLE WAVEGUIDE DIRECTIONAL COUPLER*

by

P. Andrews, B.Sc. (*Graduate*)†

SUMMARY

This paper discusses an easily constructed waveguide-to-waveguide directional coupler. Constant coupling and directivity are obtained over a wide frequency band. The coupling factor and directivity are calculated for two typical cases for waveguide of external dimensions 3 in by 1½ in. In addition an expression is derived for the frequency sensitivity of the device, showing that the rate of change of coupling with frequency is almost zero in the middle of the useful range of the waveguide for H_{10} transmission.

1. Introduction

A directional coupler is a device which, when inserted in a transmission line along which there exist travelling waves, delivers to a pair of terminals located in an "Auxiliary Transmission Line," a voltage which is largely a function of the amplitude of the wave travelling in one preferred direction in the "Main Transmission Line," and relatively independent of the wave travelling in the opposite direction.

The "Coupling" of a directional coupler specifies that fraction of the power proceeding in the preferred direction in the main-line, which is delivered to the auxiliary line output terminals. This power is referred to as the "Coupled Power." The coupling is normally expressed as a power ratio (since this ratio is less than unity, when expressed in decibels it is a negative number. This negative sign is normally omitted for convenience), and is defined as the ratio (in decibels) of the power delivered to a matched detector at the auxiliary line output terminals to the power delivered to the input terminals of the main transmission line, provided that the main line output terminals are matched.

The auxiliary line terminals in the unpreferred direction should, under the above conditions, deliver no power. In practice, however, a small amount of "leakage power" will be delivered. The "directivity" of a directional coupler is defined as the ratio (in db) of coupled power to leakage power.

Waveguide-to-waveguide directional couplers are widely used in microwave systems. Much work has been done on the development of highly directive devices, which are usually not easy to manufacture since close tolerances are called for. For many applications high directivity is not necessary and for these applications an easily constructed directional coupler would suffice. The directional coupler which is described in this paper meets the following specification:—

(*a*) It is easy and cheap to manufacture.

(*b*) The coupling is substantially constant over the useful range of the waveguide.

(*c*) The coupler takes up only 3 in of waveguide when designed for WG10 (3 in × 1½ in ext.; 7·62 cm × 3·81 cm).

(*d*) The device will withstand high power without breakdown (more than two MW in WG10 even for a coupling aperture of 1 in diameter).

(*e*) The directivity is greater than 10 db.

The calculations and models which have been made are all based on waveguide WG10 (internal dimensions 2·84 in × 1·34 in). This work is, of course, equally applicable to other sizes of waveguide.

2. The Directional Coupler

Many types of directional coupler have been considered, but the majority fail to satisfy conditions (*a*), (*b*) and (*c*).

A directional coupler which was found to fit the specification is the orthogonally crossed waveguide type described below. This device has one circular coupling aperture placed on a diagonal of the area of intersection of the two waveguides, the aperture normally being tangential to two adjacent sides of the intersection

* Manuscript received June 30th, 1954, and in final form October 27th, 1954. (Paper No. 302.)

† Admiralty Signal and Radar Establishment, Cosham, Hants.

U.D.C. No. 621.372.8.

boundary (see Fig. 1). Greater directivity can be obtained if the aperture shape is made a slot or cruciform, and if two or more such apertures are used. Such modifications may in many cases be useful, since manufacture would not be greatly complicated thereby.

For the type of coupler considered here, if:—

Internal large dimension of waveguide $= a$,

Internal small dimension of waveguide $= b$,

Radius of aperture $= r$,

Distance of centre of aperture from narrow wall of waveguide $= x_0$,

Free space wavelength $= \lambda_0$,

Waveguide wavelength $= \lambda_g$,

A_1 = Voltage ratio coupling in forward direction between the waveguides, and

B_1 = Voltage ratio coupling in backward direction between the waveguides.

Then:[1]

$$A_1 = \frac{4\pi}{3} \cdot \frac{r^3}{ab}$$

$$\left[\frac{2}{a}\sin\frac{\pi x_0}{a}\cos\frac{\pi x_0}{a} - j\frac{\lambda_g}{\lambda_0^2}\sin^2\frac{\pi x_0}{a}\right] \quad ..(1)$$

and

$$B_1 = -j\frac{4\pi}{3}\,\frac{r^3}{ab}\left[\frac{\lambda_g}{\lambda_0^2}\sin^2\frac{\pi x_0}{a}\right]$$

$$\text{Therefore Directivity} = \left|\frac{A_1}{B_1}\right|^2$$

$$= 1 + 4\frac{\lambda_0^4}{a^2\lambda_g^2}\cot^2\frac{\pi x_0}{a} \quad ....(2)$$

$$\text{and Coupling} = |A_1|^2$$

The derivation of the expressions for A_1 and B_1 is adequately covered by Surdin[1] and Bethe.[2]

The equation (2) as given by Surdin[1] is obviously a misprint, since it is dimensionally incorrect. The formulae given above assume that the wall between the two waveguides is infinitely thin. In practice, the wall of one waveguide is cut away, the hole then being a waveguide-wall-thickness deep. It therefore acts as a short evanescent guide whose attenuation must be added to that obtained from equation (1).

The additional attenuation due to the finite thickness of the aperture can be calculated by taking into account the following expressions:—

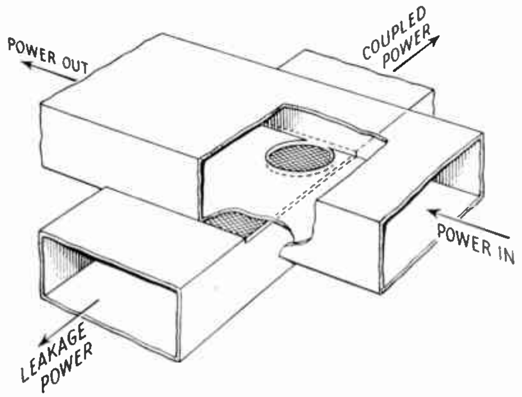


Fig. 1.—Cross-waveguide coupler.

Additional attenuation for magnetic coupling (db)

$$= C = 32\,\frac{t}{d}\sqrt{1-\left(\frac{1\cdot71\,d}{\lambda_0}\right)^2} \quad ....(3)$$

Additional attenuation for electric coupling (db)

$$= C_E = 41\cdot8\,\frac{t}{d}\sqrt{1-\left(\frac{1\cdot31\,d}{\lambda_0}\right)^2} \quad ......(4)$$

where t = total wall thickness at the aperture,
and d = diameter of the aperture.

Equation (3) affects the real part of (1), while equation (4) affects the imaginary part of (1) and equation (2). Due to the fact that wall thickness is normally kept to a minimum only equation (3) is of practical importance in our case.

Table 1 shows calculated values of coupling and directivity over a 600 Mc/s band when $a = 2{\cdot}84$ in, $b = 1{\cdot}34$ in, $r = 0{\cdot}50$ in, and $r = x_0$.

Table 1

Frequency Mc/s	Calculated Coupling in db	Calculated Directivity in db
2700	29·27	10·47
2800	29·36	10·56
2900	29.25	10·60
3000	29.23	10·60
3100	29·21	10·56
3200	29·18	10·51
3300	29·16	10·42

Table 2 shows calculated values of coupling and directivity over a 600 Mc/s band when $a = 2{\cdot}84$ in, $b = 1{\cdot}34$ in, $r = 0{\cdot}125$ in, and $r = x_0$.

Table 2

Frequency Mc/s	Calculated Coupling in db	Calculated Directivity in db
2700	75·73	23·02
2800	75·73	23·13
2900	75·72	23·17
3000	75·71	23·16
3100	75·71	23·13
3200	75·70	23·06
3300	75·69	22·98

There are 31 waveguide sizes in the United Kingdom, about a dozen of these being in common use. Design tables and curves are not, therefore, given in this paper since they are readily obtained from equations (1) and (2). Since the imaginary part of (1) is normally less than 10 per cent. of the real part, on rationalization the imaginary part contributes less than $\frac{1}{2}$ per cent. to the value of A_1. Thus, for any given value of x_0/a, A_1 is proportional to r^3/a^2b. With two exceptions all waveguides in common use have $a = 2b$ (for the two exceptions $a = 2{\cdot}12b$ and $a = 2{\cdot}25b$). Thus, in general, for any given value of x_0/a, A_1 is proportional to r^3/a^3. Thus, once data have been compiled for one size of waveguide, no further calculations need be made for new sizes unless extreme precision is required.

3. Frequency Sensitivity

The frequency sensitivity of the coupling can be calculated directly as follows:—

The frequency sensitivity is $\dfrac{\mathrm{d}|A_1|^2}{\mathrm{d}\lambda_0}$

A_1 is of the form $a - jf(\lambda_0)$,

Therefore

$$|A_1| = \sqrt{a^2 + f(\lambda_0)^2}$$

$$|A_1|^2 = a^2 + f(\lambda_0)^2,$$

and

$$\frac{\mathrm{d}|A_1|^2}{\mathrm{d}\lambda_0} = \frac{\mathrm{d}(a)^2}{\mathrm{d}\lambda_0} + \frac{\mathrm{d}[f(\lambda_0)^2]}{\mathrm{d}\lambda_0}$$

Therefore

$$\frac{\mathrm{d}|A_1|^2}{\mathrm{d}\lambda_0} = 2f(\lambda_0)\,.\,\frac{\mathrm{d}f(\lambda_0)}{\mathrm{d}\lambda_0} \quad \ldots\ldots\ldots\ldots\ldots (5)$$

Also $\lambda_g = \dfrac{\lambda_0\,\lambda_c}{\sqrt{\lambda_c^2 - \lambda_0^2}}$ where $\lambda_c = 14{\cdot}42$ cm for WG10.

Therefore, for WG10,

$$\lambda_g = \frac{14{\cdot}42\,\lambda_0}{\sqrt{208 - \lambda_0^2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (6)$$

From (1), (5) and (6),

$$\frac{\mathrm{d}|A_1|^2}{\mathrm{d}\lambda_0} = 2\left(\frac{4\pi}{3}\,.\,\frac{r^3}{ab}\,.\,14{\cdot}42\sin^2\frac{\pi x_0}{a}\right)^2.$$

$$\left(\frac{2\lambda_0^2 - 208}{\lambda_0^3(208 - \lambda_0^2)^2}\right) \ldots\ldots\ldots\ldots (7)$$

From (5) we can see that

$$\frac{\mathrm{d}|A_1|^2}{\mathrm{d}\lambda_0} = 0 \text{ when } (2\lambda_0^2 - 208) = 0$$

This is when $\lambda_0 = 10{\cdot}19$ cm. For $\lambda_0 < 10{\cdot}19$ cm $\dfrac{\mathrm{d}|A_1|^2}{\mathrm{d}\lambda_0}$ is negative, while for $\lambda_0 > 10{\cdot}19$ cm $\dfrac{\mathrm{d}|A_1|^2}{\mathrm{d}\lambda_0}$ is positive. Or more generally for any waveguide $\dfrac{\mathrm{d}|A_1|^2}{\mathrm{d}\lambda_0} = 0$ when $(2\lambda_0^2 - \lambda_c^2) = 0$.

This derivation once again assumes that the wall between the two waveguides is infinitely thin. From equation (3) it can be shown that the correction for the finite thickness of the aperture is also frequency sensitive.

$$C = \frac{32t}{d}\sqrt{1 - \frac{(1{\cdot}71d)^2}{\lambda_0}}$$

Then

$$\frac{\mathrm{d}C}{\mathrm{d}\lambda_0} = \frac{32t}{d}\,.\,(1{\cdot}71d)^2\,.\,\frac{1}{\lambda_0^2\sqrt{\lambda_0^2 - (1{\cdot}71\,d)^2}}$$

Thus, for small apertures

$$\frac{\mathrm{d}C}{\mathrm{d}\lambda_0} \propto \frac{1}{\lambda_0^3}$$

Since the correction (3), which has to be applied to (1) to allow for the wall thickness, is about 10 per cent. of the main coupling, the additional frequency sensitivity $\mathrm{d}C/\mathrm{d}\lambda_0$ is not serious.

4. Performance

Directional couplers have been made to this design and give satisfactory results. It has not been possible to check the calculations of Table 1 precisely, since the experimental apparatus available would not give the required degree of accuracy, but the practical models have agreed with theory within the experimental limits. In particular, two models were made in WG10 whose aperture diameters were specified as 1·000 in and 1·005 in. The difference in coupling was calculated as 0·25 db. A null method was used to measure the difference of coupling, which was 0·2 ± 0·1 db (see Appendix). The edges of the aperture are normally radiused, and a model so treated has been tested satisfactorily at 2 MW peak power. The input v.s.w.r. of these two models was better than 0·97 over the frequency range of Table 1. A photograph of one of the models is shown in Fig. 2.

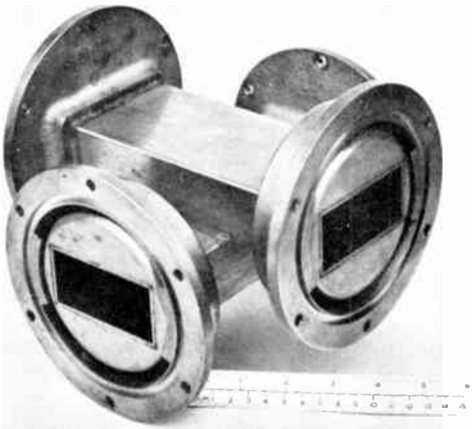


Fig. 2.—The directional coupler.

5. Acknowledgments

This paper is published by permission of the Admiralty.

6. References

1. M. Surdin. "Directive couplers in waveguides." *J.Instn Elect. Engrs*, **93**, Part IIIA, 1946, p. 725.
2. H. A. Bethe. "Theory of diffraction by small holes." *Physical Review*, **66**, 1944, p. 163.
3. C. G. Montgomery. "Technique of Microwave Measurements," M.I.T. Radiation Laboratory Series, Vol. 11. (McGraw-Hill, New York, 1947.)

7. Appendix: Measurement of Small Attenuation Differences

As described in this paper, two directional couplers were made with coupling holes having diameters of 1·000 in. and 1·005 in. The difference in coupling was calculated as being 0·25 db. The null method which was used to measure this difference in coupling is described below.

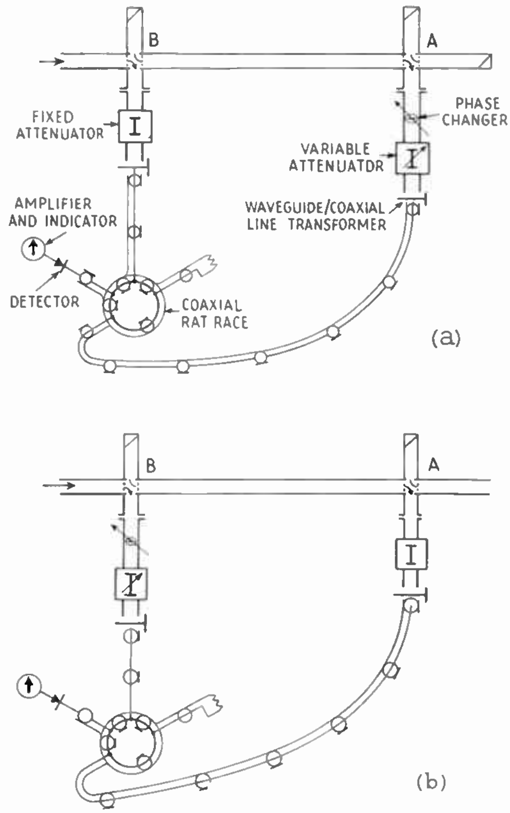


Fig. 3.—Apparatus for measurement of small attenuation differences.

The two directional couplers were set up as shown in Fig. 3a. Couplers A and B were connected in series in a terminated waveguide system into which was fed a modulated microwave signal. The coupled power of A was connected via a phase changer, a calibrated variable attenuator and a waveguide-to-coaxial line transformer into a coaxial rat-race. The coupled output arm of B was connected via a fixed attenuator and a waveguide-to-coaxial line transformer to the same coaxial rat-race. The output of the rat-race was detected by a crystal, amplified in a selective amplifier and displayed on a meter.

The phase changer and variable attenuator were successively adjusted until zero reading was obtained. The two inputs to the crystal must then be of equal amplitude and in anti-phase.

Let coupling coef. of B = B db.

Let coupling coef. of A = A db.

Let the fixed pad, the transformer, and cable attached to B have an attenuation of x db.

Let phase changer, transformer, and cable attached to A have an attenuation of y db.

Let the variable attenuator reading indicate a setting of p db.

Then:—

$$B + x = A + y + p \qquad \text{.............(1)}$$

All apparatus attached to the two coupler output arms was now interchanged (see Fig. 3b) and the phase changer and variable attenuator re-adjusted for zero reading.

Then if the new attenuator reading indicates a setting of q db

$$A + x = B + y + q \qquad \text{.............(2)}$$

Subtracting (2) from (1)

$$B - A = A - B + p - q$$

Therefore: $A - B = \frac{1}{2}(q - p)$

This method gave a difference of coupling coefficient between the two models of 0·2 db ± 0·1 db. The ±0·1 db limit in this result is the experimental resetting accuracy of the variable attenuator.

MANAGEMENT COURSES

The third of the extra-mural courses for managers in industry, arranged by the University of Cambridge, is to be held from June 27th to July 22nd this year at Madingley Hall.

The course will be concerned with problems common to management in all fields and special attention will be paid to changing conditions, whether initiated by management or by national or international circumstances. Attention will also be given to the importance of Trade Unions.

A condition of entrance to the course is that the candidates must have had some years' experience in industry and are under the age of 40 years.

Enrolment is taken from the whole field of industry (production, technical, sales, accounting, etc.), but engineering representatives comprised less than one-third of the attendance at the 1953 and 1954 courses. The report of the University states that "the line manager could with advantage have been more strongly represented." One of the problems of management is, of course, the difficulty encountered when making changes in working methods, and this indicates that the course also examines the problems sometimes resulting from the introduction of new techniques for improving and increasing production.

THE ANALYSIS OF BINARY GAS MIXTURES BY A SONIC METHOD*

by

E. W. Pulsford, B.Sc. (*Associate Member*)†

A paper presented during the Industrial Electronics Convention held in Oxford in July 1954

SUMMARY

The operation of the instrument described is based on the variation of the velocity of sound in a mixture of two gases, with the proportions of the mixture. The velocity of sound is determined at a fixed temperature by measuring the resonance frequency of a double Helmholtz resonator containing the mixture. The source of the sound is a diaphragm driven electromagnetically from a variable-frequency oscillator; a similar transducer followed by an amplifier and rectifier form the resonance detector.

The method of continuous recording of the mixture proportions, is also described. In this, the gas mixture flows through the resonator, and a simple servo-mechanism is used to maintain the condition of resonance and to operate the recorder.

1. Introduction

The instruments here described measure the proportions of a mixture of two known gases, making use of the necessary condition that the velocity of sound must be different in the two gases. In the mixture, the velocity of sound is intermediate between the velocities in the pure gases, its value for a particular mixture being related to the mixture proportions in a way to be shown later.

Instead of measuring this velocity directly, it is found more convenient to measure the resonant frequency of an acoustic cavity and to relate this to the mixture proportions. The resonant frequency at a given temperature is, for ideal gases, directly proportional to the velocity of sound in the gas, the constant of proportionality being a function of the linear dimensions only of the resonator. Once the instruments are calibrated for mixtures of the pair of gases, they become direct-reading devices: one is a manually controlled instrument, the other effects a continuous analysis and automatically records the results.

Briefly, the method involves the measurement of the resonant frequency of an acoustic cavity formed of two Helmholtz resonators coupled at the neck, and filled with the gas mixture under test, the whole being kept at a constant temperature. The resonator is driven by an electromagnetically operated diaphragm which, in turn, is driven by a sinusoidal waveform generated electronically. The state of resonance is detected by means of a similar diaphragm (in the other half of the resonator) acting as a microphone, the output of which is amplified, rectified and exhibited on a meter. The input frequency is adjusted for maximum response, and the result obtained from a knowledge of this resonant frequency. In the case of the automatic instrument, the tuning to resonance is done by a simple servo-mechanism, which also drives the pen of a chart recorder thereby keeping a continuous record.

These instruments are used in a process control plant where, because of the nature of the gases involved, it is not possible to extract samples for analysis away from the plant, and, therefore, a physical in situ method is essential. Further, the nature of the gases and the operational conditions affect the details of the design, particularly of the resonator. Firstly, the plant gas is, in this instance, at a pressure considerably lower than that of the atmosphere: secondly, the temperature is considerably above that of the normal ambient, and, thirdly, the plant gas is very corrosive, necessitating careful choice of the materials in which it is enclosed. The method of dealing with the first point appears later in consideration of the resonator design; the second point necessitates the use of a constant temperature enclosure for the resonator, and the third is a special matter not affecting the principle of the method.

* Manuscript received May 31st, 1954. (Paper No. 303.)

† Atomic Energy Research Establishment, Harwell.

U.D.C. No. 621.389:534.614:533.27.

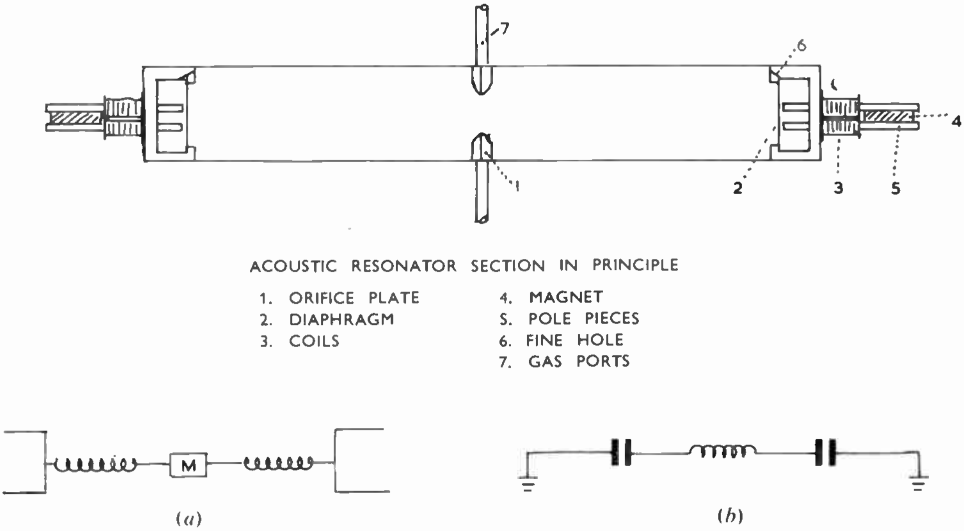


Fig. 1.—Sectional view of the resonator showing (a) mechanical analogue, and (b) electrical analogue.

On the process control plant, the resonator is situated in a by-pass line, and the gas mixture circulated by a small pressure difference. The electronic part of the instrument may be many yards away in a position convenient for the operating staff.

2. The Physical Principles Involved

In what follows it is assumed, for simplicity, that ideal gases are being considered, and the temperature is constant.

In an acoustic resonator of fixed linear dimensions it may be shown that the resonant frequency, f, of a particular mode of vibration (say the fundamental) is a linear function of the velocity of sound in the contained gas.

$$f = Bc$$

For the same resonator, successively filled with two gases in which the velocity of sound is c_1, c_2,

$$\frac{f_1}{f_2} = \frac{c_1}{c_2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

The classical expression for the velocity of sound in an ideal gas is

$$c = \sqrt{\frac{\gamma . p}{\rho}}$$

where γ is the ratio of the specific heats
p is the pressure.
ρ is the density.

Combining this expression with the gas equation $pv = RT$, and the density equation $\rho = M/V$ (M and V being the molecular weight and volume respectively) we obtain

$$c = \sqrt{\frac{\gamma RT}{M}}$$

Substituting similar expressions for c_1, c_2 in (1) we obtain

$$\frac{f_1}{f_2} = \sqrt{\frac{\gamma_1}{M_1} \cdot \frac{M_2}{\gamma_2}}$$

Suppose now that the resonator is filled with a mixture of the two gases. It would be expected, and is borne out by experiment, that the resonant frequency f_m would lie between f_1 and f_2, since M_m lies between M_1 and M_2 and γ_m lies between γ_1 and γ_2. The mean molecular weight of the mixture, M_m, is dependent entirely on the proportions of the two gases, and the same is true of the mean ratio of the specific heats, γ_m. It therefore follows that f_m is dependent only on these proportions, and a measurement of f_m enables us to determine them.

In an instrument embodying this principle there are two possible methods of calibration. Firstly, by introducing mixtures of known proportions and measurement of the resonant frequency, an empirical relationship between resonant frequency, and mixture proportions may be determined. This is probably the safer method, and is used with confidence by the plant workers. Secondly, by calculating M_m and γ_m a theoretical relationship may be obtained. The method of doing this is given in Appendix 1. This method depends for its accuracy on a number of factors: how closely the ideal gas laws are obeyed, and how accurately the physical constants M and γ are known. Such calculations were made for typical pairs of gases, and the results checked experimentally. The reasonably close agreement found is indicative of the reliability of the methods used in calculation, and of the relative insignificance of the corrections which ought to be applied to take account of the departure from ideality.[1]

3. The Resonator

The desirable characteristics of the resonator may be summarized as follows:—

(1) It should have a single resonant frequency. Other quasi-resonances it may have because of harmonic modes or diaphragm resonances must be relatively weak or far removed in frequency from the main resonance.

(2) The internal energy losses should be small, resulting in a sharply tuned resonance peak.

(3) The means by which the resonator is made to oscillate, and the resonance point detected, should both be of a simple nature.

(4) Provision should be made for working at pressures other than atmospheric, so that it may be included in closed gas-handling systems.

These characteristics are all possessed by a double form of the Helmholtz resonator; the basic features are illustrated in Fig. 1, and a photograph is given in Fig. 2. The resonator consists of two cylindrical cavities of approximately equal volumes connected by an orifice in the common separating wall. When resonance occurs, the gas in the region of the orifice oscillates under the influence of pressure changes of the gas in the remainder of the system. The whole may be considered analogous to a mass attached to two similar springs (Fig. 1a) or to an inductor connected to two capacitors (Fig. 1b) provided that the wavelength of sound is long compared to the linear dimensions of the resonator. When this condition is satisfied (as is easy to do in practice) the acoustic system, in common with its mechanical and electrical analogues has a single resonance frequency. There are other conceivable resonant modes, for example, when the half-wavelength of sound is equal to the distance between the orifice plate and the opposite wall of the cavity. Such a mode would have a frequency very much higher than the Helmholtz mode, and is easily placed outside the range of the oscillator. It should be said that this or any similar mode, is not observed in practice, and is, therefore, probably weak in comparison with the Helmholtz mode.

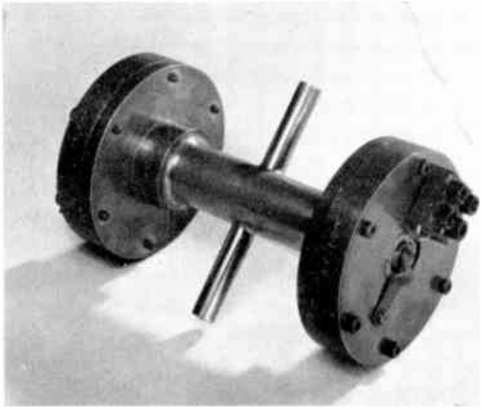


Fig. 2.—The resonator.

The method of calculating the resonance frequency of the acoustic system is given in Appendix 2.

In order to reduce the frictional and viscous losses in the system, and consequently to ensure a sharp resonance peak, the orifice between the two halves of the resonator is reduced to a simple, relatively large, circular hole, and the material of the orifice plate is polished. It is reasonable to suppose that, in general, the frictional losses are a reciprocal function of the diameter of the hole, while the oscillating mass is a cubic function. Therefore the ratio of the energy stored in the oscillating mass, to the energy lost per cycle of oscillation is increased rapidly by increasing the diameter of the hole. It must not be made too large or else the acoustic system may approach that of a double-stopped organ pipe, which has a harmonic series of resonances undesirable for our purpose.

Oscillations are set up in the gas in the resonator by means of a diaphragm forming one end of the tube. This diaphragm is driven electromagnetically, power being obtained from a variable-frequency oscillator. The driving coils are situated outside the body of the resonator to prevent any deterioration because of the attack of chemically active gases. Oscillations at resonance are detected by means of an exactly similar electromagnetic system at the other end, acting as a microphone, the output of which is amplified electronically, and used to indicate the resonance point, which is that of maximum response.

In order to work at pressures other than atmospheric, it is necessary to maintain the gas pressures equal on both sides of the diaphragm. This is done by having fine holes in the body of the resonator connecting the main cavity with a much smaller cavity behind the diaphragm. These fine holes are also relatively long, and so have a high acoustic impedance; when properly dimensioned, their presence does not affect the performance of the system, but if incorrectly dimensioned, there is a small chance that they may form, with the cavity behind the diaphragm, a Helmholtz resonator whose frequency lies within the instrument range. In such a case, as happened during development of the resonator, anomalous results are obtained. Small spurious resonances and double-humped main resonances have been observed, which have disappeared on altering the dimensions of these weep holes, and are, therefore, attributed to some such cause as that described.

The inlet and outlet ports for the gas are situated in the orifice plate, an antinodal point where there are no pressure changes. There are good reasons for this: if the ports are situated elsewhere, then their presence, together with that of the tubes conveying the gas, constitutes an additional indeterminate volume added to that of the resonator proper, and it is then found that the resonance frequency is somewhat dependent on the amplitude of oscillation; this is not the case when they are situated in the orifice plate. Also, by having them as shown, the initial response of the instrument to a change of proportions of the mixture is more rapid than it would be if they were elsewhere. This is because the elasticity of the gas in the body of the resonator is independent of its composition, while the resonant frequency depends on this, and also on the mass of a certain volume of gas in the region of the orifice, which mass is dependent on the mixture proportions. So, by changing the gas in the region of the orifice at the earliest opportunity, a rapid initial response to a change of mixture is obtained; the final equilibrium value is attained in a time depending on the diffusion of the gas until the contents have the same composition throughout and further diffusion does not result in alteration of the composition of the gas in the orifice. In practice this equilibrium is set up quite rapidly, and the indication of a change being imminent is virtually immediate.

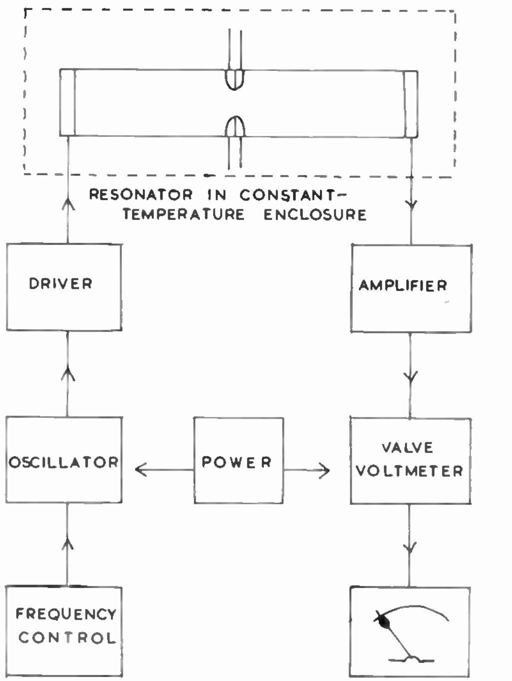


Fig. 3.—Block diagram of manually controlled system.

Two sizes of resonator have been tried. Their resonant frequencies when filled with dry air at 20°C. are of the order of 425 and 1,200 c/s. It is the smaller tube which figures in the illustration, and is in general use.

In the design of the smaller tube, use is made of the loaded diaphragms and electromagnetic system of existing deaf-aid earpieces, and these are found most satisfactory.

The connecting leads to the resonator are made of shielded twin cable, often many yards long so that the indicating unit may be conveniently situated for the operating staff.

4. Electronic Equipment

The block diagram (Fig. 3) shows that the resonator is driven from a calibrated variable-frequency oscillator via a driving stage, and the resonance point is detected by a microphone followed by an amplifier and peak-rectifier whose output is shown on a meter.

5. The Variable-Frequency Oscillator

The desirable characteristics of the variable-frequency oscillator may be summarized as follows :—

(i) It should produce a fairly pure sine wave, since the presence of harmonics might give rise to anomalous resonances.

(ii) Its frequency stability should be good especially if dial readings are to be used as indicators of mixture proportions.

(iii) Frequency control should be by variation of one, or, at the most, two circuit elements. The range of frequency control should be readily calculable to facilitate design.

(iv) It should work well in the range 100–1,500 c/s for which satisfactory resonators may be designed.

The Wien bridge oscillator meets all these requirements. The type used embodies thermistor control in the negative-feedback path, thus ensuring amplitude stability (required to a moderate extent only) and a sufficiently pure sine waveform.

The basic circuit is shown in Fig. 4. $R_1C_1R_2C_2$ is the frequency controlling network, and $Th.\ R_3$ is the negative-feedback control. It is well known that at a frequency given by

$$f = \frac{1}{2\pi\sqrt{R_1C_1R_2C_2}}$$

there is no phase change between the input to C_1 and the output from the junction of R_1 R_2, and also that the attenuation is moderate provided that there is not too great a difference in magnitude of R_1 and R_2, or C_1 and C_2. (If $R_1 = R_2$ and $C_1 = C_2$, the attenuation is $\frac{1}{3}$.) If the amplifier has zero phase change between input and output, and sufficient gain, sinusoidal oscillations are set up at this frequency. The output waveform is most free from harmonics when the gain of the amplifier is only just sufficient to maintain oscillation, and this condition is automatically provided by the thermistor in the negative-feedback path. The a.c. power fed to this part of the circuit causes a reduction of the resistance of the thermistor, resulting in an increase of feedback and reduction of amplifier gain until an equilibrium is set up in which the output amplitude is stabilized at that value which adjusts the resistance of the thermistor to the appropriate value to allow oscillation to be just maintained. This gives a good sine wave output, sufficiently free of harmonics for the purpose of exciting the resonator at one frequency setting only.

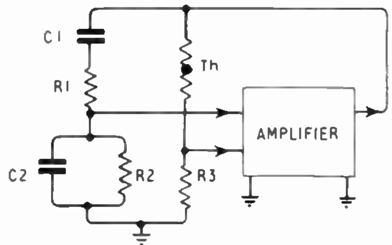


Fig. 4.—Wien bridge oscillator.

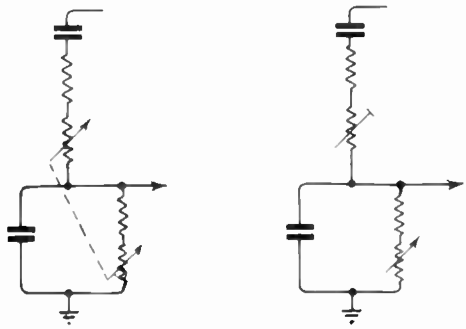


Fig. 5.—Oscillator networks.

The design of the frequency control network for any specific application, depends on the frequency and frequency range required. When the frequency range is large, variation of both resistors is used, as in Fig. 5a; when it is narrow, variation of one only is sufficient as in Fig. 5b. In the former case f is proportional to $\frac{1}{R}$, and in the latter, to $\frac{1}{\sqrt{R}}$. The frequency range depends on the ratio of the maximum to the minimum values of the variable arm(s) and is readily calculable, as also is the value of C for the frequency band required. The preset control in Fig. 5b is used to take up component tolerances in order to adjust the frequency band to the correct place; in the case of the network of Fig. 5a it is necessary to adjust the capacitors.

In general, the calibration of the frequency control dial in terms of mixture proportions is not linear, but it may be shown that for small proportions of a light gas in a heavy gas, the use of the network of Fig. 5b does result in a nearly linear calibration.

For the measurement of frequency, where the instrument is being calibrated or when high accuracy is required, it is convenient to use a standard scaler to count the number of oscillations in a known time. With this arrangement and accurate measurement of the timing interval by automatic methods applied to time pulses from a standard clock, it has been found that the long-term stability of oscillator is of the order of 1 in 10^4 to 1 in 10^5, and that the accuracy of repeated settings of a given resonance point is of the order of a few parts in 10^3. In general these figures imply an accuracy of the order of 1 per cent., or better, in measurement of mixture proportions, but the actual precision depends on the specific case.

6. The Driver Stage

The output of the oscillator, reduced in amplitude as required, is fed to the input of the driver stage which consists of a double triode connected as a type of push-pull power amplifier having the output transformer connected between the cathodes. The first triode acts as a phase splitter, whose anode waveform is fed to the grid of the second triode, resulting in push-pull action at the cathodes. The full circuit diagram, Fig. 6, shows this arrangement.

7. The Amplifier and Detector Stages

The desirable characteristics of the amplifier used in the detection of resonance may be summarized as follows:—

(i) Its gain over the frequency band in which resonances occur should be adequate and fairly constant.

(ii) Some control of gain should be provided.

(iii) The output should be as free as possible from any disturbances other than the amplified signal.

A simple two-valve amplifier employing pentode valves is found to answer these requirements: the circuit is given in the full circuit diagram, Fig. 6. Item (iii) is ensured by screening the input stage (leads and transformers) and by the use of a stabilized h.t. supply, otherwise some low-frequency fluctuations of the output meter may arise because of mains-voltage variation, thereby making it more difficult to tune for a resonance maximum.

The detector circuit consists of a d.c. restoring diode (which effectively doubles the available peak output) followed by an "infinite-impedance" detector with a backed-off meter to enable a sensitive indicator to be used although the total amplitude of the peak may be large. The low-pass filter in the meter circuit prevents signal frequency voltages and accidental surges from affecting the meter. The time-constant of this filter is not great enough to make the instrument so sluggish that tuning for resonance is difficult.

In practice, the preset gain control is set at about half-way, and the amplitude of the drive then adjusted so that a reasonably sized resonance peak is obtained: there is nothing critical about either of these mutually dependent adjustments.

8. Automatic Working

In the event of a continuous record being required of the composition of a mixture of two gases, it is convenient to arrange for automatic tuning of the oscillator to the resonance frequency of the acoustic tube, and to record this frequency on a suitable chart recorder which has previously been calibrated to interpret the resonance frequency as a mixture proportion.

The automatic instrument comprises much the same electronic circuits as those already described, with the addition of a recorder adapted to perform the tuning operation, and extra circuits to ensure that if the instrument "loses" the resonance peak, it automatically finds it again.

Because of the importance of the recorder mechanism in the automatic system, its mode of operation is first described before going into the details of the system as a whole.

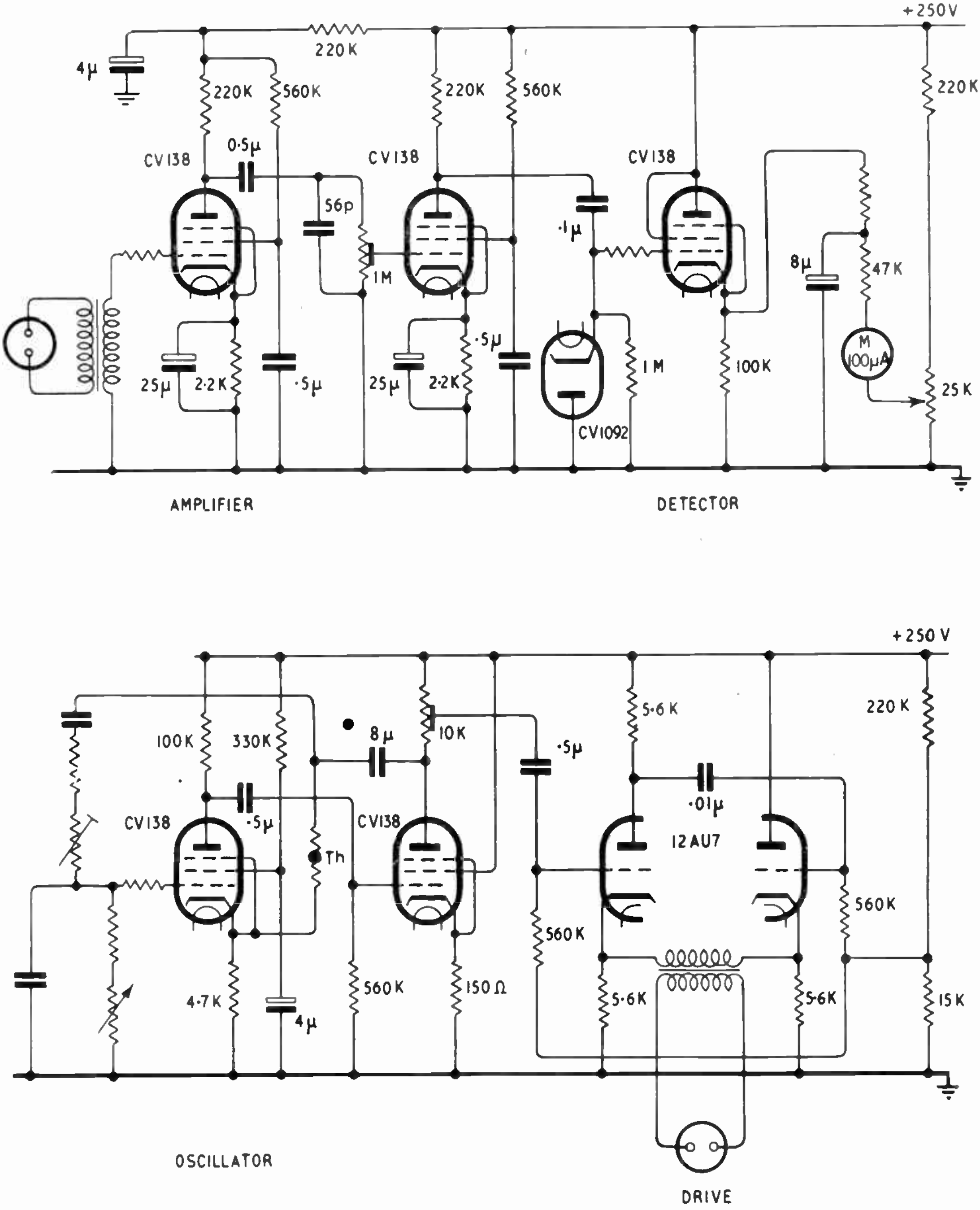


Fig. 6.—Circuit diagram of the manually controlled system.

It is found convenient to use the modified servo-mechanism of an industrial self-balancing potentiometric millivolt recorder. In this instrument the cycle of events described below takes place, the timing being done by cams on a mechanical shaft driven by a synchronous motor, which also supplies power for the adjustment of the variable resistor (used in the self-balancing function) and of the writing pen.

(i) Any difference between the input voltage and that existing on the potentiometer is detected by a suspended-coil pointer galvanometer, whose deflection is dependent on magnitude and sense of the difference.

(ii) When the damped galvanometer has come to rest the pointer is firmly clamped.

(iii) A pair of sensing fingers explores the position of the pointer. If it is not at zero, a rotation is imparted to a mechanism carrying a pair of surfaces on which the resetting wiper cams (also on the main timing shaft) can impinge.

(iv) The mechanism of (iii) is mechanically clutched to a shaft whose rotation results in alteration of the variable resistor and movement of the writing pen.

(v) The wiper cams reset the mechanism to a standard position, and in so doing adjust the variable resistance and writing pen. The circuit is thereby brought nearer to the balance point.

(vi) The galvanometer is unclamped and the mechanism declutched, whereupon the cycle repeats. The balance point is approached in a series of ever-decreasing steps.

This brief outline of the recorder operation supposes a voltage continuously applied to the galvanometer circuit. In the sonic analyser the galvanometer is used ballistically, a capacitor being discharged through the galvanometer at that point of the time cycle which allows the pointer to reach its extreme position at the instant of clamping. An adjustment of the variable resistor then takes place depending on the sense (and magnitude) of the ballistic throw.

When used for the purpose of the sonic analyser, the variable resistor in the recorder is made into the frequency control element of

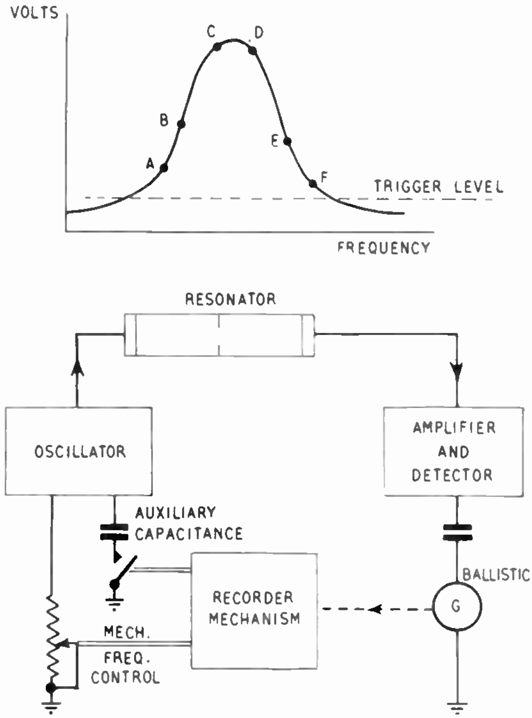


Fig. 7.—Block diagram of the automatic system.

the Wien bridge, and the ballistic mode of operation employed for adjusting it to resonance. An additional cam on the main timing shaft connects a small capacitor into the frequency control network at the appropriate instant. The insertion of this capacitor causes a small change of frequency, and hence, in general, a change in amplitude of the detector output. This amplitude change results in the passage of more or less charge through the ballistic galvanometer, with the result that the recorder mechanism alters the main frequency control in such a way that the frequency is brought nearer to resonance. The principle is illustrated in Fig. 7. Suppose the two frequencies are such as to produce outputs represented by points A and B (Fig. 7a) then, because connection of the auxiliary capacitor (point A) results in a change of output (point B), a charge represented in magnitude and direction by (B-A) is passed through the galvanometer, and matters are so arranged that the recorder movement increases the oscillator frequency. Similarly for points E and F, except that here the sense of the charge is reversed, and so the frequency is decreased. Equilibrium is set up when the two frequencies are symmetrically disposed about the peak, and the difference in outputs is zero. (C and D.)

In practice, because the response peak is fairly sharp, corresponding to a Q of 50 to 60, the amount of frequency change needed is quite small, being only about 1 or 2 per cent.

A block diagram illustrating the system is given in Fig. 7b.

In general, when the instrument is newly switched on, the frequency is not near the resonance point, and so automatic hunting to find the peak is provided. Should the detector output be less than a certain value, usually one-fifth to one-quarter of the peak value, a relay is operated via a trigger circuit. This causes a small direct current to pass through the galvanometer resulting in a unidirectional drift of oscillator frequency. As soon as the peak is found, the rising of the output voltage above the trigger level resets the trigger circuit, and thereafter the resonance point is found as already described. If the peak is not found before the recorder reaches the end of its scale, one of the two end-point limit switches comes into operation and reverses the current through the galvanometer, whereupon the instrument hunts back over the frequency range until the peak is found. "Losing" the peak for any length of time is thereby prevented.

A circuit diagram of the automatic instrument is given in Fig. 8.

9. Results

Because the details of the plant for which the sonic analyser was developed cannot be given at present, only some preliminary results, obtained in the development stages, are included. Graphs are given for mixtures of carbon dioxide and air, and for the vapour of the fluorcarbon C_7F_{14} and air. In Fig. 9 the full line gives the theoretical calibration obtained by the method of the appendix, and the marked points give the experimental results.

These examples are given to show the type of result which is expected and obtained.

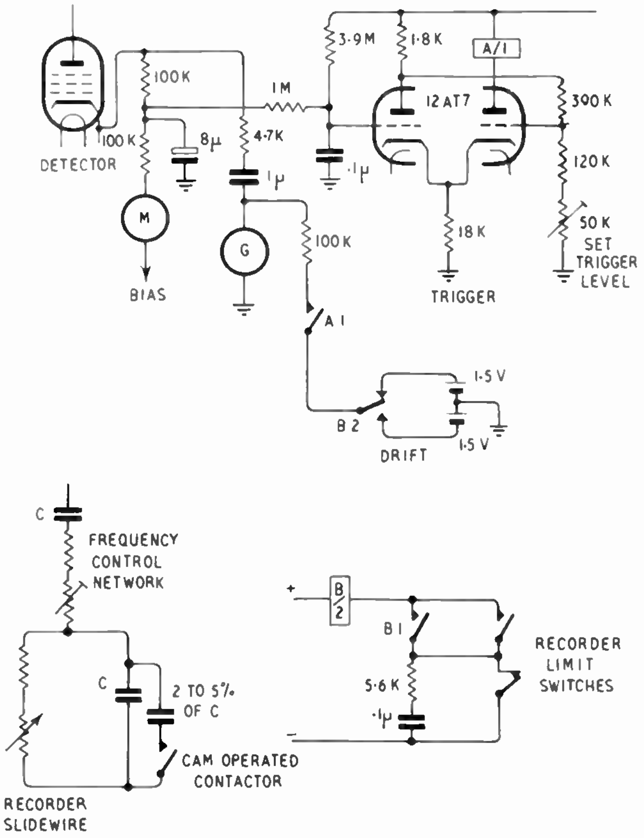
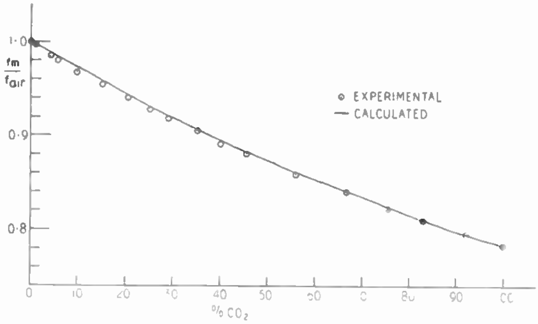


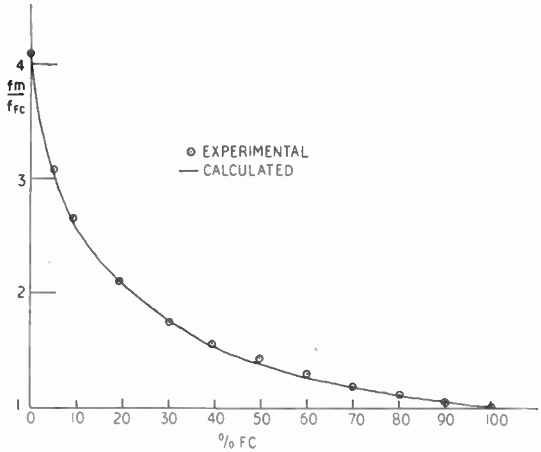
Fig. 8.—Circuit diagrams of the automatic system.

Clearly, for any particular pair of gases, it is necessary to know beforehand what results to expect, because the design of the instrument is influenced by them. The computation of the components of the frequency control network is a case in point.

It is interesting to note that, as well as being useful, though specialized, items of industrial plant equipment, these instruments also have their place in fundamental physical research, for it has been found possible to measure with the apparatus described, to a considerable degree of accuracy, the values of the constants in the equations of state which describe the behaviour of a number of gases and vapours.



(*a*)



(*b*)

Fig. 9.—(a) Results for air–carbon dioxide mixtures.
(b) Results for air–C_7F_{14} mixtures.

Figure 10 shows a test-rig for making and handling C_7 F_{14}—air mixtures. In this photograph, the larger of the two types of resonance tube is shown, together with the manometer for reading the partial pressures of the constituents of the mixture; the heating mat, thermostat head and stirring fan can also be seen.

10. Acknowledgments

The author's thanks are due to Mr. T. H. D. Attewell, who was closely associated with the development of the electronics and construction of the test-rig; to Messrs. J. P. Kerry and H. J. Edmunds, who built much of the electronic equipment; to Messrs. W. A. Kealy and H. J. Koller, for help in the production of experimental resonators, and to many others in the Atomic Energy project for useful advice, constructive criticism and help in testing the systems under plant conditions.

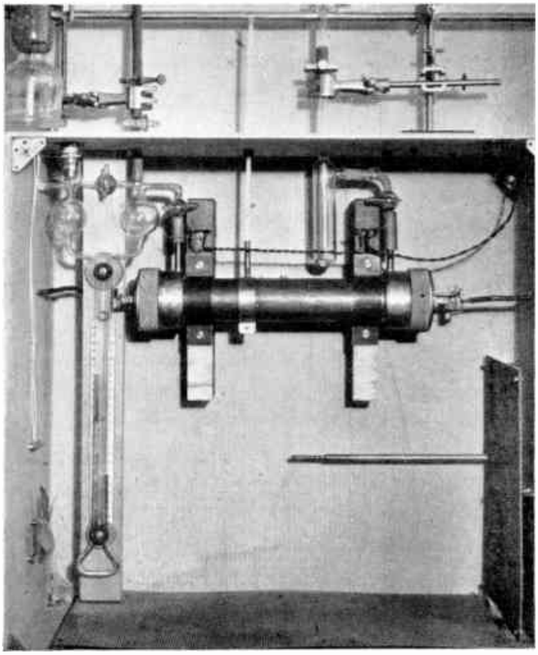


Fig. 10.—Inside view of the constant temperature enclosure of the test-rig

11. References

1. J. R. Partington and W. G. Shilling. "The Specific Heats of Gases." (Ernest Benn, London, 1924.)
   Chapters I and II of this book cover the velocity of sound in gases and mixtures, including all the fundamental work and the corrections for deviations from the ideal gas laws.
2. Alexander Wood. "Acoustics." Students' Physics, Vol. II. (Blackie, London, 1924.)
   Chapter IV deals with the Helmholtz resonator.
3. Lord Rayleigh. "Theory of Sound," Vol. II.
   Chapter 16 deals fully with the theory of resonances.
4. W. S. Tucker. "The determination of the velocity of sound by the employment of closed resonators and the hot-wire microphone." *Phil. Mag.*, **34**, 1943, p. 217.
   This paper deals with a method which is essentially similar to that used in the Sonic Analyser.

12. Appendix 1: The Velocity of Sound in a Mixture of Two Gases

The gases are assumed to obey the simple gas laws.

The velocity of sound (c_m) in the mixture may be written

$$c_m = \sqrt{\frac{\gamma_m RT}{M_m}} \quad \ldots\ldots(1)$$

Calculation of the Mean Molecular Weight, M_m.

For a mixture of two gases, whose molecular weights are M_1 M_2, in the proportions r_1 r_2 by volume, we have by Avogadro's Hypothesis that r_1 of the molecules are of the first gas and r_2 of the second.

The mean molecular weight is, therefore,

$$M_m = r_1 M_1 + r_2 M_2 \quad \ldots\ldots(2)$$

(Alternatively, if the partial pressures are p_1 p_2 then

$$M_m = \frac{M_1 p_1 + M_2 p_2}{p_1 + p_2} \quad \ldots\ldots(2a)$$

Calculation of Mean Ratio of Specific Heats, γ_m

It is shown in text books dealing with the theory of heat that, for an ideal gas*

$$Cp - Cv = R$$

where Cp and Cv are the molecular heats at constant pressure, and volume, respectively, and R is the gas constant. When Cp and Cv refer to the molecular weight in grams, R is shown to be 1·9875 calories/gm. mol.°C.

By combining this relationship with the expression defining γ, namely,

$$\gamma = \frac{Cp}{Cv}$$

We obtain

$$\gamma = \frac{Cv + R}{Cv}$$

$$= 1 + \frac{R}{Cv} \quad \ldots\ldots(3)$$

For a mixture of two gases, we may write

$$\gamma_m = 1 + \frac{R}{Cv_m} \quad \ldots\ldots(4)$$

* When the gases are not ideal, correction factors have to be applied as shown in Reference (1), p. 99 et seq.

Now Cv_m is found simply from Cv_1 and Cv_2; we have

$$Cv_m = r_1 Cv_1 + r_2 Cv_2 \quad \ldots\ldots(5)$$

From (3), we obtain

$$Cv = \frac{R}{\gamma - 1} \quad \ldots\ldots(6)$$

Combining (5) and (6)

$$Cv_m = R\left(\frac{r_1}{\gamma_1 - 1} + \frac{r_2}{\gamma_2 - 1}\right) \quad \ldots\ldots(7)$$

Combining (4) and (7)

$$\gamma_m = 1 + \frac{1}{\dfrac{r_1}{\gamma_1 - 1} + \dfrac{r_2}{\gamma_2 - 1}} \quad \ldots\ldots(8)$$

This may be written

$$\frac{1}{\gamma_m - 1} = \frac{r_1}{\gamma_1 - 1} + \frac{r_2}{\gamma_2 - 1} \quad \ldots\ldots(9)$$

Either of the forms (8) or (9) is suitable for computation.

Values of M_m (from (2) or (2a)) and of γ_m (from (8) or (9)) are used in (1) to calculate the velocity of sound in the mixture (c_m), or in the equation of paragraph (2) of the paper.

$$c_m = \sqrt{\left[1 + \frac{(\gamma_1 - 1)(\gamma_2 - 1)}{r_1(\gamma_2 - 1) + r_2(\gamma_1 - 1)}\right]\frac{RT}{r_1 M_1 + r_2 M_2}}$$

13. Appendix 2: Calculation of the Resonant Frequency of the Sonic Tube

Following the method used by Rayleigh,[3] we first calculate the resonant frequency of two cavities connected by a short neck. The result is then adjusted to the case when the length of the neck is small and the end corrections are found according to the method of Stephens and Bate.†

Assume two volumes V_1 V_2 connected by a tube of length l and of cross-sectional area S, filled with a gas of density ρ.

Suppose the gas in the neck to be moved a small distance dl towards V_1. The volume of gas V_1 is then compressed into a volume $V_1 - Sdl$. This compression is adiabatic, obeying the relation

$$pV_1^{\gamma} = \text{const.} = K \quad \text{(at a constant temperature)}$$

† R. W. B. Stephens and A. E. Bate, "Wave Motion and Sound," p. 163 et seq. (Arnold, London, 1950). This deals with the frequency of resonance of pipes having an orifice whose diameter is comparable with that of the pipe.

$$\text{or } p = KV_1^{-\gamma}$$

$$\text{and } dp = -\gamma p \cdot \frac{dv}{V_1}$$

$$\text{But because } dv = S dl,$$

$$dp = -\frac{\gamma p . S dl}{V_1}$$

The resulting force on the gas in the neck is given by $S dp$,

hence force due to compression of gas in V_1 is

$$-\frac{\gamma p S^2 dl}{V_1}$$

force due to expansion of gas in V_2 is

$$-\frac{\gamma p S^2 dl}{V_2}$$

total force due to both causes=

$$-\gamma p S^2 dl . \left(\frac{1}{V_1} + \frac{1}{V_2}\right)$$

The acceleration of the gas in the neck is, therefore, given by

$$\frac{d^2 l}{dt^2} = -\frac{\gamma p S^2 dl}{Sl\rho} . \left(\frac{1}{V_1} + \frac{1}{V_2}\right)$$

$$= -c^2 . \frac{S}{l}\left(\frac{1}{V_1} + \frac{1}{V_2}\right) dl, \text{ where } c = \sqrt{\frac{\gamma p}{\rho}}.$$

This is the differential equation of a simple harmonic motion of frequency

$$f = \frac{c}{2\pi}\sqrt{\frac{S}{l}\left(\frac{1}{V_1} + \frac{1}{V_2}\right)}$$

For a symmetrical resonator, $V_1 = V_2 = V$, and

$$f = \frac{c}{2\pi}\sqrt{\frac{2S}{lV}}.$$

S/l is termed the conductivity of the neck. The numerical value to be used in the computation of the resonant frequency of the type of resonator described is not simple to determine: we have to consider what end corrections to allow for the orifice, which is not much smaller in diameter than the body of the resonator with the result that the correction for an infinite flange is no longer accurate.

Writing the formula as

$$f = \frac{c}{2\pi}\sqrt{\frac{S}{l' + \beta} \cdot \frac{2}{V}}$$

where l' = length of the neck

β = end correction

and, following Stephens and Bate, writing

$$\beta = 2^{d/D} . d$$

where d and D are the diameters of the orifice and body respectively, we have

$$f = \frac{c}{2\pi}\sqrt{\frac{S}{l' + 2^{d/D} . d} \cdot \frac{2}{V}}$$

Although the value of l' is somewhat indeterminate because of the shape of the orifice plate (see Fig. 1), this formula gives a fairly close agreement with the measured frequency of the resonators when the length of the neck (l') is taken to be the thickness of the orifice plate at the place where the "rounding off" begins.

The above treatment is valid only if the wavelength of the resulting sound wave is long compared with the linear dimensions of the resonator. This condition is easy to realize in practice.

For the larger resonator, measurements are:—

V = 17·87 in.^{3}
d = 0·75 in.
D = 2·0 in.
l' = 0·25 in.
c = 1·343 × 10^4 in/sec. at 15°C.

Substituting these values in the formula gives

$$f = 429{\cdot}7 \text{ c/s.}$$

The measured value of f at 20°C. is 432 c/s which is equivalent to 429 c/s at 15°C.