RCA REVIEW

A Quarterly Journal of Radio Progress

Published in July, October, January and April of Each Year by
RCA INSTITUTES TECHNICAL PRESS
A Department of RCA Institutes, Inc.
75 Varick Street, New York, N. Y.

VOLUME VI January, 1942 NUMBER 3

CONTENTS

SUBSCRIPTION:

United States, Canada and Postal Union: One Year $2.00, Two Years $3.00, Three Years $4.00
Other Foreign Countries: One Year $2.35, Two Years $3.70, Three Years $5.05
Single Copies: 75¢ each

Entered as second-class matter July 17, 1936, at the Post Office at New York, New York, under the Act of March 3, 1879.

Printed in U.S.A.

NBC STUDIOS 6A AND 6B

By

GEORGE M. NIXON

National Broadcasting Company, New York

Summary—The NBC Studios 6A and 6B, located at Radio City, New York, were placed in operation early in November, 1941. The article describes the salient acoustical, electrical and architectural features of these studios. The photographs supplement the text in illustrating the more important details.

TWO new auditorium-type broadcast studios were placed in operation by the National Broadcasting Company at Radio City, New York, early in November 1941. These studios, designated as Studios 6A and 6B, incorporate the accumulated years of operating experience in the careful planning, constructing, and decorating of studios to coordinate properly all of the various operating requirements of network broadcasting.

The two studios are similar in size and modern architectural design, but distinctively different in decorative treatment to obtain a pleasant contrast in appearance. The overall length of each studio is approximately 100 feet; the width an average of 48 feet; and the ceiling height ranges from 13 to 19½ feet. The stage section is 37 feet deep and 45 feet wide; and its ceiling height is from 13 to 16 feet.

Figures 1 and 2 show the overall appearance of one of the studios—Studio 6A. Figure 1 as viewed from the rear of the elevated and stepped seating section shows the control booth, to the right of the stage, from which the occupants have an unobstructed view of the stage and of the major portion of the studio seating section. The opening for the clients' observation booth window may be seen in the upper right. From this booth the sponsors of the program may watch and listen to the progress of the program. Figure 2 shows the studio as viewed from the apron of the stage with the seats in the fore part of the studio removed. The opening of the spotlight booth is seen in the upper righthand corner of the picture.

The walls of Studio 6A and the stage ceiling are of a rich warm copper color. The chairs in this studio are of dark-green fabric upholstery, and the flooring has a rubber covering upon which the carpeting is laid. The walls of Studio 6B and the stage ceiling are of a bright cheerful silver color. This studio has a red rubber flooring, with red carpeting, and blue upholstered chairs.

Each studio accommodates an audience of about 450 in comfortably upholstered chairs which are arranged to provide an unobstructed view of the stage. One third of these chairs, 150, are removable and are located on the level floor area immediately in front of the stage. The remainder, 300, are fixed and arranged on a gradually stepped slope so that all seats are "good seats." The seats of the chairs are automatically retractable so that, when unoccupied, the seat slides downward and backward to provide a maximum of space between adjacent rows.

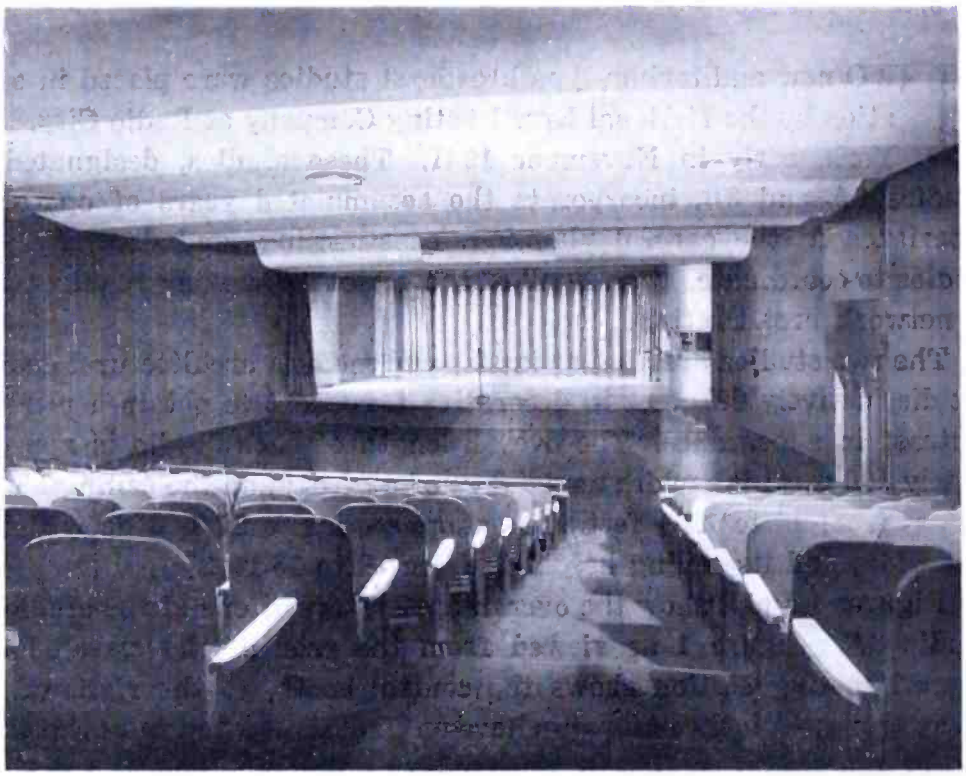


Fig. 1

The upholstered chairs, in addition to being comfortable, perform the important function of maintaining substantially the same acoustical conditions with and without an audience present in the studio. Program balance and microphone location may be determined and adjusted carefully during rehearsal, with the assuring knowledge that it will not be disturbed by the influx of the audience. The removable seats in the front of the stage permit use of this area by performing groups or, by the removal of two or three rows of seats and the use of platforms abutting the stage apron, the size of the stage may be augmented to accommodate very large performing groups.

The wall and ceiling surfaces of the stage are arranged to provide sound diffusion so that the reflected sounds are mixed properly and

the tonal quality of the performer or performing group is enhanced. The wall on the control booth side is plane, but the opposite wall is deeply "veed" in flat sections of wall, each 4 feet wide, to a depth of 2 feet, from the "crest" to the "trough" of the "vee" so that impinging sounds are reflected dispersely. The rear wall is constructed of a series of plaster semi-cylindrical convexly curved surfaces which provide additional diffuse reflection. In Studio 6A the rear wall is of vertical semi-cylindrical pillars and in Studio 6B these curved surfaces are in horizontal layers. A drapery is provided which may be drawn to cover

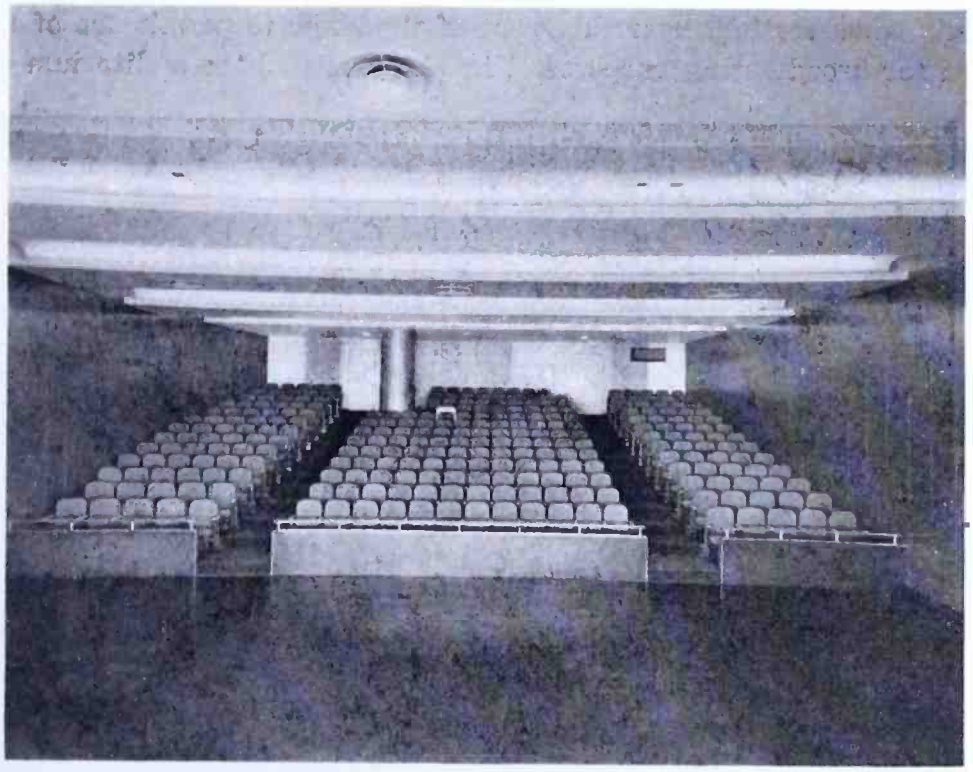


Fig. 2

these surfaces to alter acoustical conditions in the studio, as desired. Two other draperies are in each studio. One, for use as a "draw curtain" in the front of the stage or what would be the proscenium in a theater. The second is about midway along the side walls to reduce the apparent size of the stage when very small performing groups use the studio. The arrangement of the wall and ceiling surfaces may be seen in Figure 3.

The stage ceiling, shown in Figure 4, is in "broken sawtooth" fashion for two reasons, the first to provide a sound diffusing condition, and the second to conceal the border lights and "spots" from the eyes of the audience. This photograph, which is of Studio 6B, also shows the horizontal semi-cylindrical surfaces on the rear wall of the stage; a

closer view of the "veed" side wall and the method of concealing the draperies in "curtain pockets".

The acoustical treatment employed throughout the studio is rock wool 2 inches thick except for the rear portions of the stage ceiling where the rock-wool treatment is 4 inches thick. The acoustical treatment on the walls is covered with perforated asbestos board applied in 2 foot squares. The treatment on the stage ceiling is covered with perforated metal to provide greater absorption at the higher audio frequencies.

The ceiling of the auditorium section is untreated except in relatively small sections near the front of the stage to permit use of this area for broadcasting purposes. The side walls of the auditorium are

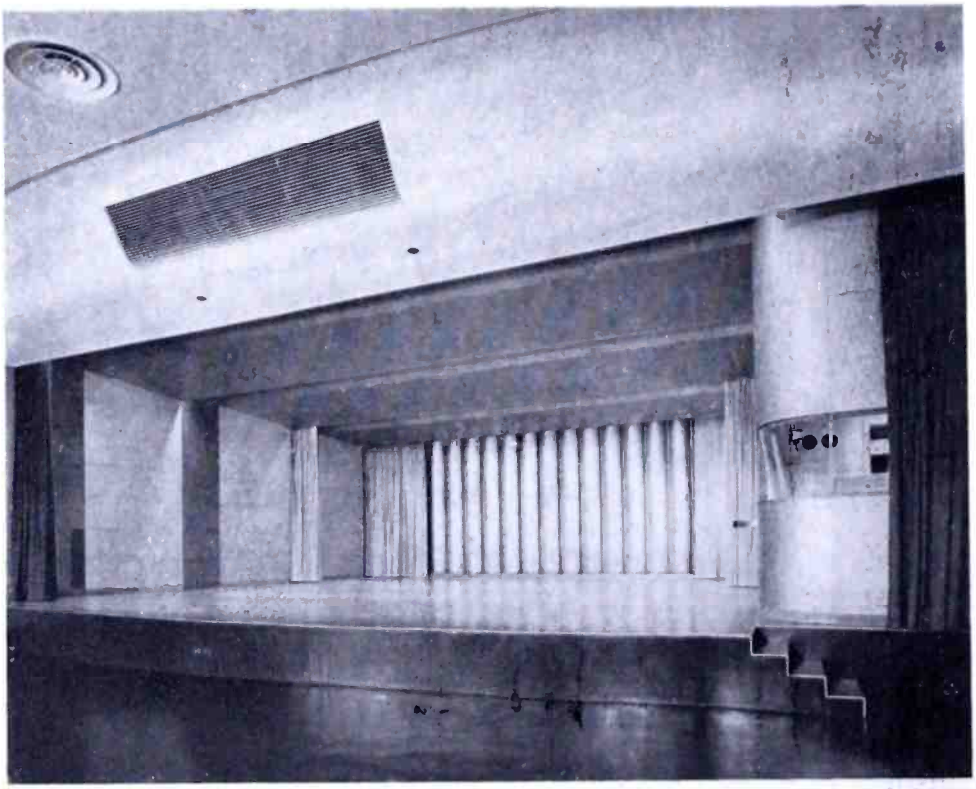


Fig. 3

likewise untreated except near the stage to permit microphone usage in this vicinity. The side walls are "splayed" horizontally in wide flat panels so that very few opposite parallel surfaces exist. The absence of acoustical treatment on the side walls and ceiling of the auditorium section provide beneficial reflections from these surfaces which aid in the transmission of sound from the stage to the listeners.

The rear wall of the auditorium is entirely treated with rock wool 2 inches thick covered with perforated asbestos board; in Studio 6A, the supporting surface is deeply serrated or "veed" and in Studio 6B, horizontally splayed. Impinging sounds are largely absorbed by the

rear wall of these studios; the small amount of energy reflected is diffused so that no annoying discrete, delayed reflections or "echoes" from this wall interfere with the enjoyment of listeners in the front seats.

The upholstered seats and broadloom carpeting lined with felt on the aisles also provide a highly absorbent area in this portion of the studio.

Sound isolation is provided by the use of double 6-inch solid-cinder block partitions to avoid sound transmission between the studio and adjacent spaces. The walls, floor, and ceiling of the stage section are supported on metallic springs damped by felt to reduce the transmission of vibration between the studio and the building structure.

Fig. 4

All necessary openings in the boundary surfaces have been given careful attention to insure maintenance of the sound isolation. The control booth and clients' booth windows are double glazed of 5/8-inch and 3/8-inch thick glass set in felt. The window frame between panes is lined with absorbent felt to decrease sound transmission and lessen any tendency of resonance due to the confined air volume. Entrance to the studios is effected through double doors separated by an acoustically treated vestibule. The doors are of solid wood 2 5/8 inches thick, provided with sponge-rubber gaskets on head and sides, and have

an automatic door closer of sponge rubber to seal the gap at the sill.

The air-conditioning ductwork is double wrapped with paper-covered rock-wool blanket at all points where studio partitions are pierced. The ducts are provided with cellular sound absorbers of Cabot's quilt (ell grass) for a distance of 16 feet. This material effectively attenuates sound that may travel through and along the ductwork.

The lighting system is not the least of the important features of these studios. The seating section is pleasantly illuminated by cove lighting in sweeping curves across the width of the studio. The light-

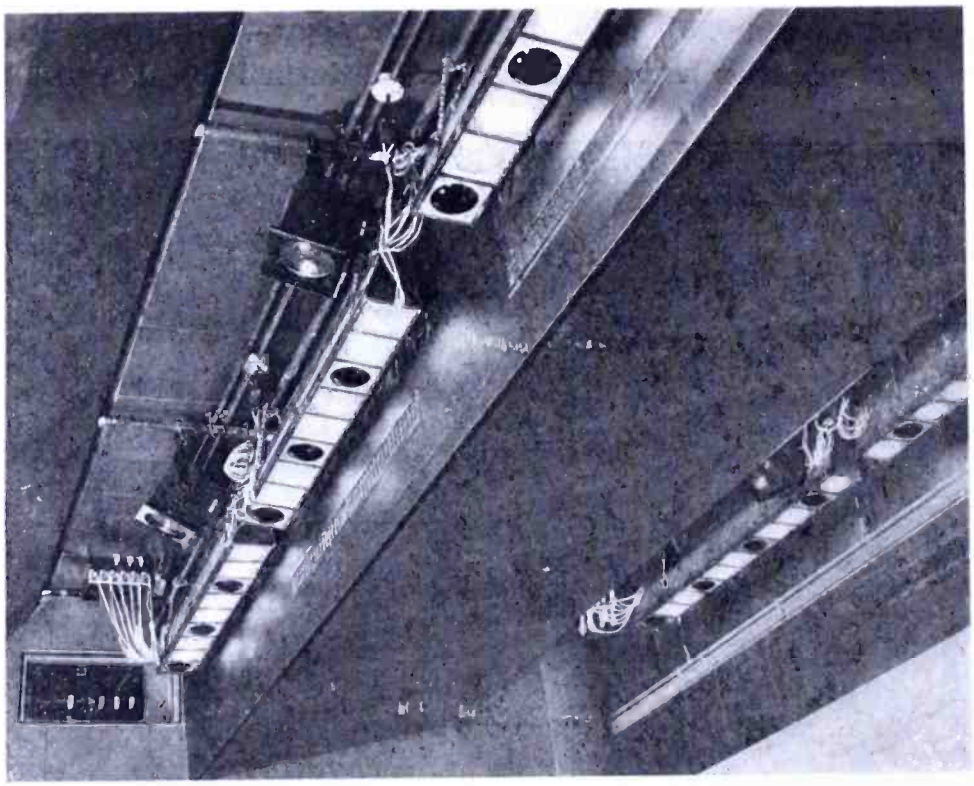


Fig. 5

ing forms an interesting pattern on the ceiling and provides indirect illumination in this portion of the studio. The stage lighting uses border lights with associated circuits arranged so that various color combinations may be obtained to provide an atmosphere in keeping with the program content. Spotlights are provided for individual artists or featured small groups on the program. Individual control of lighting sections together with master and sub-master controls permit flexible operation to obtain desired combinations. The auditorium lights are dimmed at the start of the program to heighten the enjoyment of the program by the visible audience. The stage lighting, border lights, and spots may be seen in Figure 5. The observation window of the electrician's booth is also seen in this photograph, and

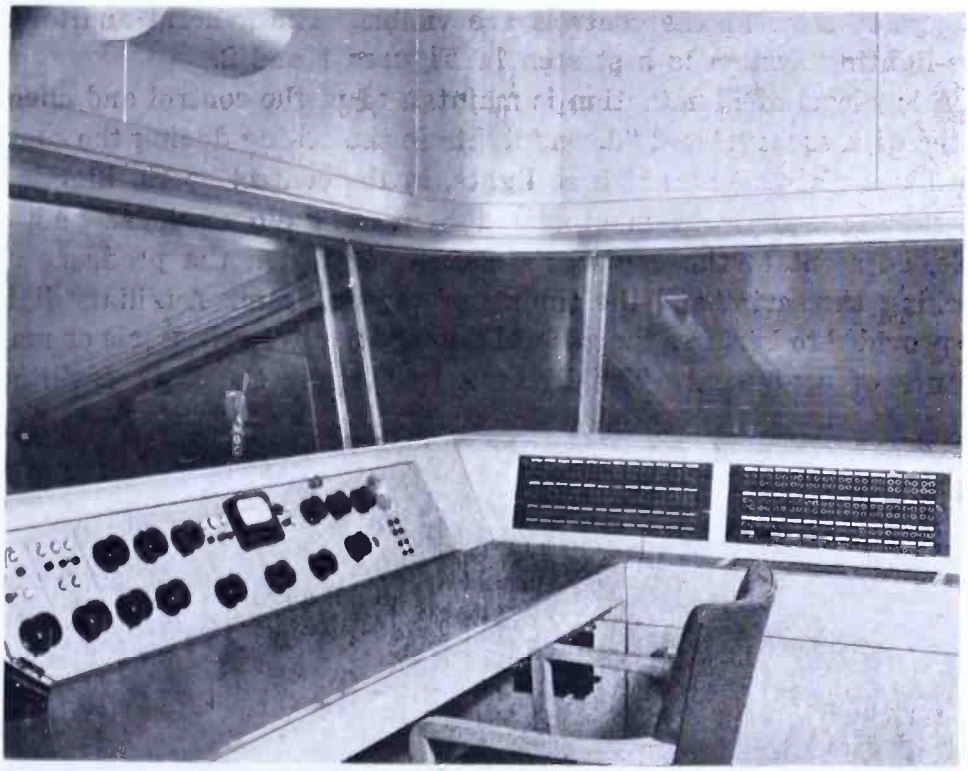


Fig. 6

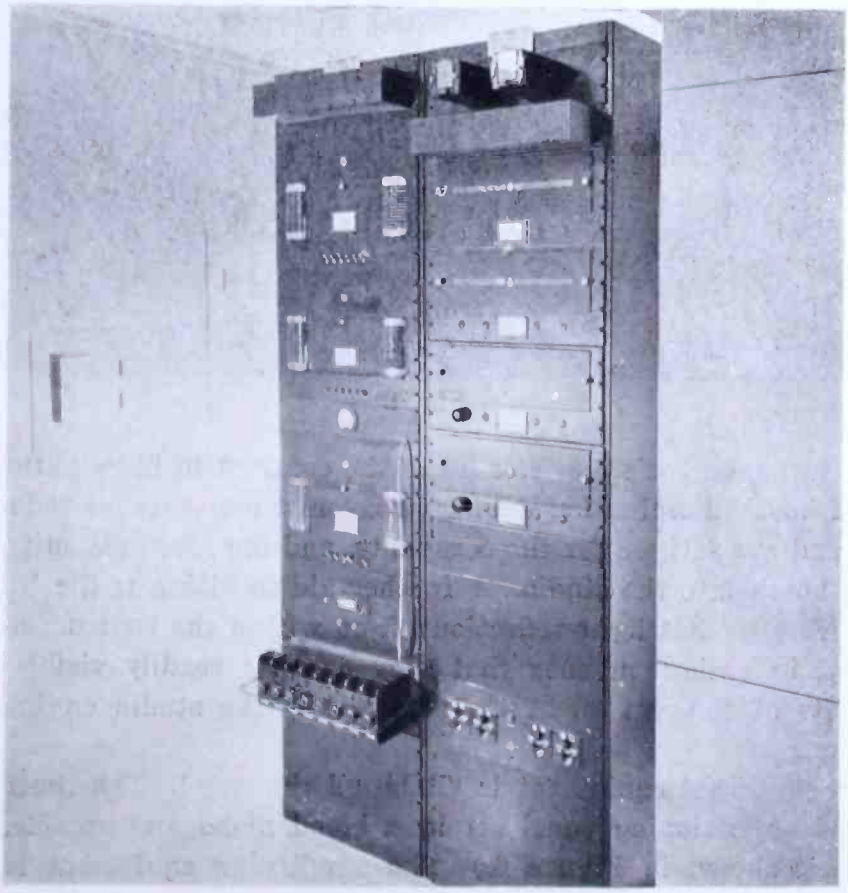


Fig. 7

through it some of the controls are visible. The general auditorium cove-lighting system is best seen in Figures 1 and 2.

A low level of illumination is maintained in the control and client's booths with concentrated "down" lights in the ceiling during the course of a broadcast program. These lights, in the control booth, illuminate the mixing equipment manipulated by the studio engineer and the table, adjacent to the engineer's console, at which the producer sits checking the script and the timing of the program. Auxiliary lights are provided to increase the general illumination for rehearsals or maintenance of equipment. The lighting in the client's booth facilitates

Fig. 8

concentration on the stage because of the contrast in illumination. The reduced booth illumination during program time tends to reduce eye-strain and eye fatigue for the occupants, and improves vision from the control booth into the studio. A further aid to vision is the tilting of the window so that light reflections from within the control booth are reflected in such a manner that they are not readily visible to the occupants of the control booth, particularly the studio engineer and producer.

The broadcast equipment is "RCA all the way". The controls are mounted on a sloping panel set in a blond mahogany console. (The console is shown in Figure 6.) The amplifying equipment is easily

accessible and located in a room just off the control booth, as shown in Figure 7.

The sound-reinforcing system for the studio is energized from the broadcast program circuit to permit the visitors to hear those artists who perform close to the microphone and who would be otherwise inaudible to a large portion of the audience. Another important reason for the use of a sound-reinforcing system is that the program is balanced for the microphone and all efforts are directed toward optimum broadcasting results. Consequently, in view of the arrangement of the orchestra, the choral group, sound effects, actors, etc., on the stage, the

Fig. 9

program as heard directly in the studio from the seating area would differ radically from that heard by the radio listener. The use of a sound-reinforcing system thereby provides the visible audience with the same program balance as heard by the radio listener. Volume control of the sound-reinforcing system is also provided at a small console located in the seating section and shown in Figure 8.

An interesting view of the studio taken during a program from mid-stage is shown in Figure 9.

The experience obtained from the operation of the studios should further the knowledge of studio design and provide a sound basis for future broadcast studio planning and construction.

The planning of these studios was under the direction of Mr. O. B. Hanson, Vice President and Chief Engineer of the National Broadcasting Company. The architectural treatment, decoration, and general construction details were the concern of W. A. Clarke, Manager of Technical Services. The design and supervision of the installation of the broadcast and electrical equipment were under the care of Mr. C. A. Rackey, Audio Facilities Engineer and his associates, J. Hastings and T. Nolen. The Operating and Architectural Departments of Rockefeller Center were responsible for the preparation of plans and general supervision of construction.

GENERATION AND DETECTION OF FREQUENCY-MODULATED WAVES

By

Stuart Wm. Seeley, Charles N. Kimball, Allen A. Barco

RCA License Laboratory, New York, N. Y.

Summary—This paper contains a description of a frequency detector which is inherently linear in operation and, consequently, adaptable for precision monitoring measurements. In conjunction with a mixer and heterodyne oscillator, the detector is capable of operation at signal frequencies encountered in the usual FM transmitter. The monitor detector is discussed in the first part of the article.

The availability of a detector capable of demodulating frequency-modulated waves with the introduction of less than 0.1 per cent distortion makes possible the investigation of other low-distortion frequency-modulated devices as, for example, a generator of phase or frequency-modulated waves. The second part of the paper contains a description of a method of producing low-distortion frequency-modulated signals. The system was developed, and its low-distortion capabilities realized, with the aid of the aforementioned linear detector circuit, in conjunction with an audio-frequency wave analyzer.

Part I—Linear Frequency-Modulated Monitor Detector

Principle of Operation

Introduction

Operation of the detector unit may best be described with the aid of the basic circuit diagram of Figure 1. e_{in} is, in practice, a frequency-modulated signal whose carrier frequency lies in the range between 100 and 300 kc. The amplitude of e_{in} is sufficient to overswing the limits of cut-off and zero bias of the 807 tube. During portions of the cycle when this tube is cut off, the plate potential rises to the $+B$ level. During positive grid portions of any cycle the plate current rises to a certain maximum value, beyond which it does not increase with further increase in positive grid swing. (See plate family curves of beam pentode tube characteristics.) Thus, the minimum value as well as the maximum value of the instantaneous plate potential is constant, and the output wave is squared off on top and bottom at definite fixed potentials, in spite of possible variations in amplitude of grid swing.

With a square wave of plate voltage, having a repetition rate equal to that of the input signal e_{in}, the action is as follows: when the plate potential of the 807 reaches its peak value (equal to E_B) the small condenser C (25 $\mu\mu f$ approx.) in Figure 1 has charged through diode d_1 to a potential equal to E_B, with the diode end of C negative. At that

part of the cycle when the plate potential of the 807 swings toward its lowest value, C discharges through diode d_2 in series with its load R_o; thus, one pulse of current flows through the load R_o of diode d_2 for each cycle of operation.

The total charge acquired by the condenser C once each cycle is CE_B (neglecting the contact potential of diode d_1). The portion of this total charge which passes through diode d_2 and R_o once each cycle is equal to that total charge (CE_B) minus the residual charge $Ce_{p\,min}$ (again, neglecting the contact potential of diode d_2). Then, as long as E_B and $e_{p\,min}$ remain constant and the time constants of the charge and discharge circuits are sufficiently small compared to the period of the input wave, the total quantity of electricity which flows through R_o in each cycle is constant. Thus, an increase in the repetition rate will increase the number of current pulses per unit of time and, therefore,

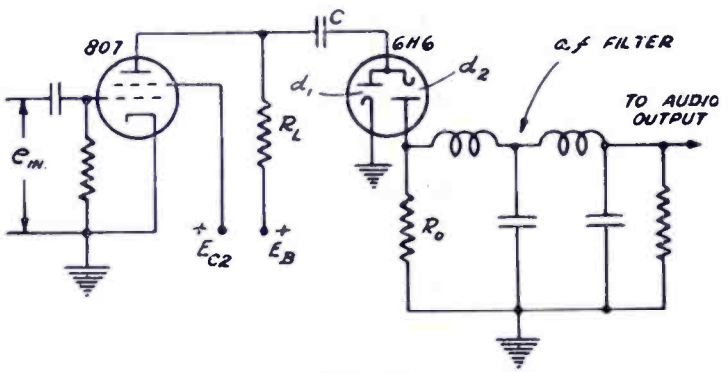


Fig. 1

increase the average value of current flowing through R_o. Conversely, a decrease in repetition rate decreases the average current through R_o, and the average potential across R_o is thus a perfectly linear function of the repetition rate or frequency of the input signal.

Obviously, then, dynamic operation of the detector will result in essentially distortionless detection over that part of the characteristic for which the previously discussed charge and discharge time-constant requirements hold.

Practical Detector Circuit

The circuit of the laboratory model of the monitor detector, shown in Figure 2, is essentially the same as the basic circuit of Figure 1, except that a 6AG7 amplifier stage has been added.

Certain pertinent factors relating to the development and use of the monitor detector are listed below.

(1) The sensitivity (audio volts developed per kc frequency deviation in the input wave) is approximately a linear function of the B voltage applied to the 807. Hence, variations in this B voltage will

introduce undesired components into the detector output. Therefore, the B supply must be well regulated.

(2) Approximately 20 peak-peak volts, at a center frequency of 150 kc, are required as input to the 6AG7, when the circuit of Figure 2 is used. Voltage at this frequency is obtained from that at any other signal frequency by means of the heterodyne oscillator and converter, shown in Figure 3. This particular circuit is designed for operation at signal frequencies in the 40-Mc range.

The a-c peak-to-peak grid voltage applied to the 807 must be at least 20 per cent greater than the value required to swing from zero bias to cut-off. The 3000-ohm series grid resistor in the 807 circuit aids in maintaining approximate equality of duration of the two halves of the plate-voltage wave, and also limits the peak grid current on positive grid swings.

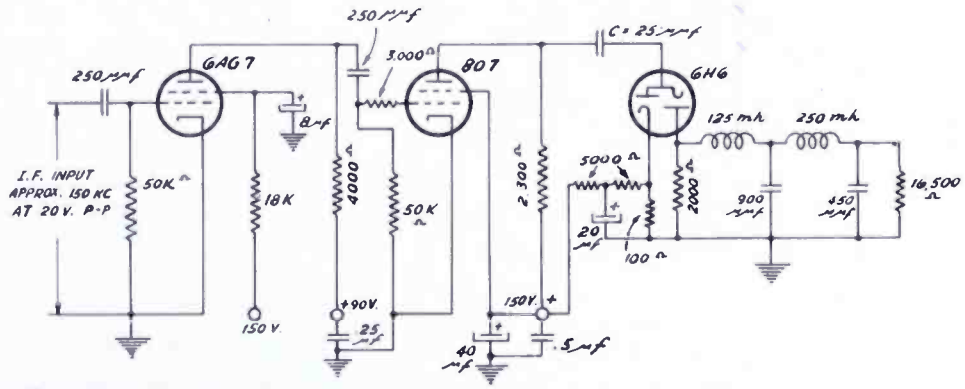


Fig. 2—90-volt and 150-volt supplies are regulated. One side of all heaters (supplied from 6.3-volt transformer) is grounded.

(3) The 807 plate load is adjusted to produce plate-current saturation at a peak grid signal just short of grid current.

(4) The value of capacitance C (in conjunction with other inherent circuit capacitances) is determined by the maximum permissible time constants in the charge and discharge circuits. These should both be sufficiently small compared to a half-period at the highest operating frequency to permit (a) e_p to rise to within 0.1 per cent of E_B in one-half cycle and (b) to permit e_p to fall to within 0.1 per cent of its normal minimum value, again within a half-cycle.

(5) The lower limit of operating frequency is determined by the cut-off of the low-pass filter in shunt to R_o (Figure 1). This filter is necessary both to remove undesired signal-frequency components from the output of the device and to allow the instantaneous potentials across R_o to rise and fall with each discharge-current pulse. If R_o were by-passed directly, to remove the signal-frequency components, the residual charge on the bypass condenser would act as a diode bias, and prevent complete discharge of condenser C. Then the discharge currents would

approach zero asymptotically as the frequency was increased, and the resultant curvature of the output voltage vs. input frequency would produce undesirable distortion.

The characteristic impedance of the filter is high compared to R_o. The filter is terminated at its output and where it may be connected to the grid of an audio amplifying tube.

(6) It has been found experimentally that a small bias voltage (approximately 1.5 volts) in series with the charge diode d_1 effectively overcomes the contact potential of both diodes and contributes somewhat to improved linearity of the device.

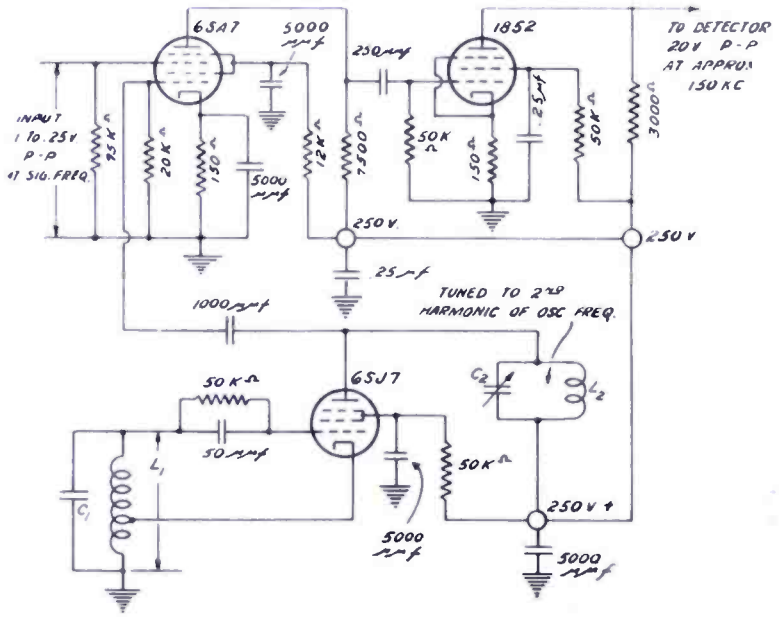


Fig. 3

Experimental Determination of Detector Linearity

One may expect representative linear performance from a monitor detector built according to the data presented. There may, however, be circumstances wherein an accurate knowledge of the capabilities of the device are required. For this reason a brief description of the calibration method employed is included here.

Figure 4 shows a circuit in which the monitor detector (without the heterodyne oscillator of Figure 3) is represented diagrammatically by a box with input and output terminals. A signal generator is employed (in conjunction with a calibrating crystal oscillator) to provide a point-to-point determination of d-c detector output vs. input frequency. An L. & N. Kohlrausch slide wire, fed from a 2-volt storage cell, is used in conjunction with a sensitive galvanometer in a potentiometer circuit which permits the d-c output of the detector to be

determined with a precision of about 0.02 per cent (approximate reading accuracy of the graduated scale of the slide wire).

The extreme sensitivity of the d-c measuring circuit requires the elimination of all sources of diode-current fluctuation. Hence, the *B* supplies feeding the 807 plate and 6AG7 screen are doubly regulated (for calibration purposes only), with the circuit shown in Figure 4, and the diode (d_1) bias obtained by bleeder current (as shown in Figure 2) is replaced by a 1.5-volt battery for determination of detector linearity.

The crystal oscillator of Figure 4 is used to check the calibration of the signal generator, by zero beat methods, with the detecting phones in the plate of the mixer tube. Calibration of the signal gen-

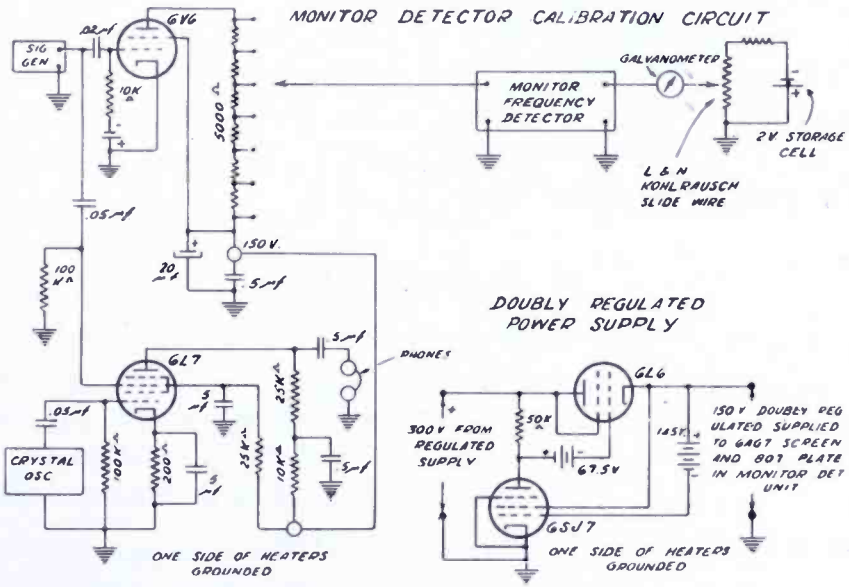


Fig. 4

erator at frequencies which correspond to fractional harmonics of the crystal is possible as a result of curvature in the mixer characteristics.

Measurements of d-c output voltage of the discharge diode in the detector circuit were made at regularly spaced input-frequency intervals. Any non-linearity in the plot of d-c volts vs. frequency was determined by plotting the readings, and comparing the resultant curve with a straight line having approximately the same mean slope. Then a curve of the difference in ordinates of the measured curve and the comparable straight line was again plotted against frequency.

By employing this method, it was found that the departure from linearity of the composite detector characteristic was less than 0.02 per cent over a frequency range of 50 to 250 kc. Thus, if a frequency-modulated wave of 100-kc deviation were applied to the detector at a carrier frequency of 150 kc, the detector would introduce only

approximately 0.02 per cent distortion in the audio output obtained from the detector.

Tests were also made to determine the susceptibility of the detector to amplitude modulation in the input signal. With 25 per cent amplitude modulation (and a mean value of input signal of 20 volts peak to peak) the output contains an undesired signal corresponding to less than 0.05 per cent of the audio voltage produced by 100-kc deviation.

The sensitivity of the detector is approximately one volt peak-to-peak (audio) for 100-kc deviation in the input signal (and 20 volts peak-to-peak applied signal). Therefore, if a single frequency signal of 100 kc is applied to the detector the d-c output should be approximately 0.5 volt. Knowledge of this sensitivity makes possible the adjustment of the heterodyne oscillator frequency (in Figure 3) to a value which causes the beat frequency to lie in the band for which the detector linearity is optimum. It is necessary merely to measure the d-c output of the detector as the heterodyne oscillator frequency is varied, and to adjust the latter to produce a d-c output voltage corresponding to 150-kc carrier input.

Applications

The monitor detector as thus constituted is an excellent instrument for studying the effects of the pass characteristics of a frequency-modulation receiver's r-f and i-f circuits on distortion, for checking the performance of frequency-modulation signal generators, for use as a monitor and peak deviation indicator at a frequency-modulation transmitter, and for many other applications.

Part II—Low Distortion Frequency-Modulation Generator

Introduction

Frequency-modulation signals can be derived from systems based on either frequency or phase modulation principles, provided that the audio potentials (or program material) be properly integrated before application to the phase modulator.

In a frequency-modulation system of the type employing a reactance tube modulator, the frequency deviation produced by a certain audio voltage is independent of the frequency of that audio voltage, and is determined only by the sensitivity of the reactance tube (kc deviation per audio volt), and by the subsequent frequency multiplication. Assuming that the circuits in the generator are sufficiently flat over the desired deviation band, the harmonic content of the output of a frequency-modulation generator of the aforementioned type is deter-

mined principally by curvature in the reactance tube characteristic, and is not a function of audio frequency.

Conditions obtaining in phase-modulation systems are quite different, since, for a given frequency deviation, the required peak-phase deviation varies inversely with audio frequency, i.e.

$$\phi \text{ in degrees at carrier freq.} = \frac{57.3 \times \text{freq. dev. } c/s \text{ at carrier freq.}}{\text{audio-freq. } c/s}$$

Thus, the audio voltage at the phase-modulator grid must vary inversely with audio frequency, to produce constant frequency deviation. If the distortion in the resultant frequency-modulated signal is caused by a non-linear relationship between the instantaneous phase of the generated signal and the instantaneous audio voltage, the distortion will generally vary in some inverse manner with audio frequency. Another source of distortion lies in the inability of the frequency multiplier circuits to pass all the required side components of the modulated signal. This deficiency will generally produce distortion increasing with audio frequency.

The distortion due to non-linearity in the phase characteristic can be reduced by employing frequency multiplication from a relatively low-frequency source of phase-modulated signals, to reduce the required phase deviation at the modulator. Side-band clipping distortion effects are suppressed by employing multiplier circuits having adequate band width.

Principle of Operation of Experimental Phase-Modulation Generator

In most generators employing phase modulation as a means for ultimately producing frequency-modulated signals, modulation by the audio components or program material is accomplished at a low value of signal frequency (generally about 100 kc), which is subsequently multiplied and heterodyned to produce the desired carrier frequency.

One system of phase modulation employed experimentally requires first the generation of amplitude-modulated signals at a low signal frequency, after which the carrier is shifted in phase with respect to the amplitude-modulated sidebands, to produce a phase-modulated signal. The inherently remanent amplitude modulation present in this developed signal is then removed by limiter action.

Another system of phase modulation is described in Patent No. 2050067, issued August 4, 1936 to Dr. Walter Van B. Roberts. This system makes use of an unique characteristic of a simple parallel tuned

circuit operated at $1/\sqrt{2}$ times its resonant frequency. An investigation of the operation of this circuit together with following frequency multipliers was made with the aid of the afore-described frequency-modulation monitor. The basic principle of operation of the circuit is as follows: A parallel resonant circuit, tuned to $\sqrt{2}$ times the frequency of an unmodulated low-frequency carrier (generally about 100 kc) is used as the plate load in a pentode whose grid is fed with voltage at the source frequency (100 kc). See Figure 5. The phase of the plate voltage of the pentode is then varied by changes in the series resistance of the resonant circuit. Figure 6 shows the manner of variation of ϕ with R; between the limits of $R = 0$ and $R = \infty$ the total phase change is 180°. One of the important features of the system is that the absolute value of the plate-load impedance is constant and

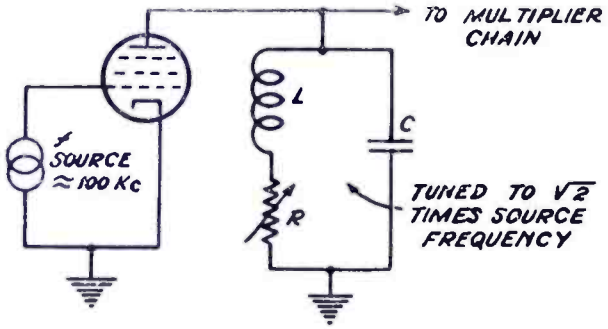


Fig. 5

equal in magnitude to the reactance of the circuit capacitance, for all values of R from zero to infinity. Hence, no amplitude modulation is present, regardless of the extent or amount of phase modulation. Appendix I contains a development of the equations of phase vs. R, and also a proof of the constancy of plate-load impedance as R is varied.

In practice, R in Figure 5 is the dynamic plate impedance of a tube, whose control grid is supplied with audio-frequency potentials. Since the variation of ϕ with R is non-linear, as shown by Figure 6, a linear relationship between the instantaneous phase of the pentode's plate voltage and the instantaneous value of modulator grid voltage is obtained only over an interval for which a curve of Eg vs. R_p (for the modulator tube) has the same shape and curvature of that of Figure 6. The problems associated with matching the curves of ϕ vs. R and R_p vs. Eg will be discussed in greater detail later.

The phase-modulated voltage (at a carrier frequency of 100 kc) is fed from the tuned circuit of Figure 5 through a chain of harmonic multipliers to develop a frequency-modulated or phase-modulated signal

at the desired carrier frequency. The phase deviation is, of course, multiplied by the same factor as is the center frequency.

The phase deviation required at the modulator is determined by the desired phase deviation and by the order of frequency multiplication. Thus, an expression for modulator phase deviation (at the 100 kc level) for any audio frequency is

$$\phi^\circ = \frac{57.3 \times \text{freq. deviation at ultimate carrier freq.}}{\text{audio frequency} \times \text{order of frequency multiplication}}$$

It is desirable to maintain the maximum required phase deviation in the modulator circuit at as low a value as possible, to minimize non-

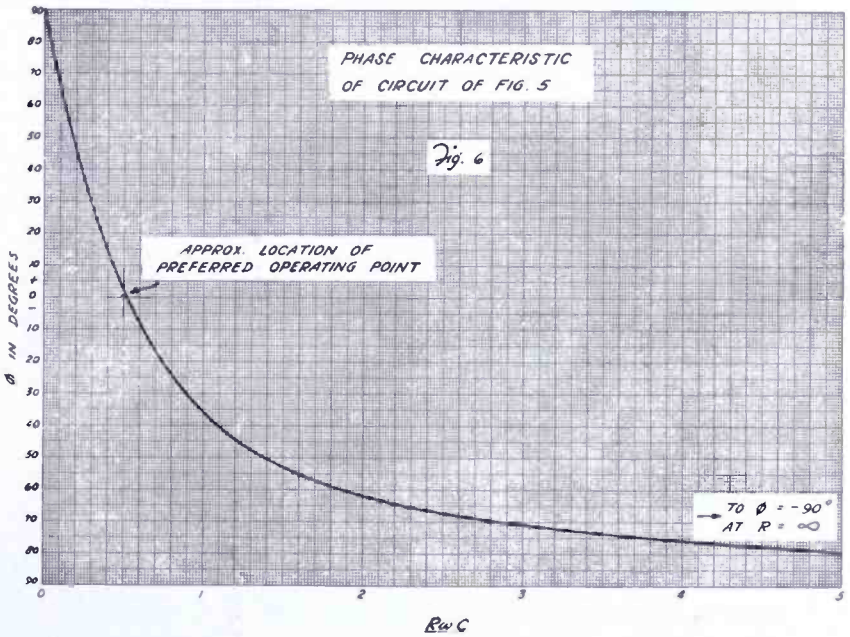


Fig. 6

linear effects at low audio frequencies (which require the greatest phase deviation for a given frequency deviation). Two methods of minimizing this low-frequency-modulator phase deviation are available. The first involves starting at a reduced value of signal frequency (lower than 100 kc) to obtain 2 or 3 times additional frequency multiplication, thereby reducing the required maximum modulator phase deviation by a factor of 2 or 3. The limiting low value of source frequency must be greater than the highest audio-modulating frequency.

A better way to reduce the low audio-frequency phase shift required in the modulator is to employ a heterodyne action at some point in the multiplier chain, by which the phase-modulated signal is heterodyned from say, 1500 kc, down to a lower frequency, after which additional multiplication is applied. This permits the realization of practically

unlimited multiplication of the phase deviation without resorting to very low source signal frequencies in the modulator circuit.

The heterodyne system was employed in the laboratory generator, as shown in the block diagram, of Figure 7. The source of signals is a crystal oscillator operating at 100 kc. This is fed through a filter to a 6SJ7 with the tuned circuit of Figure 5 as its plate load.

The grid of the 6L6 modulator tube (triode connected) is supplied with audio potentials, and its plate resistance constitutes the variable R of Figure 5. The resultant phase-modulated signal is multiplied successively by factors of 3 and 5 to produce a 1500-kc voltage, whose phase deviation is 15 times that effected by the modulator. In a

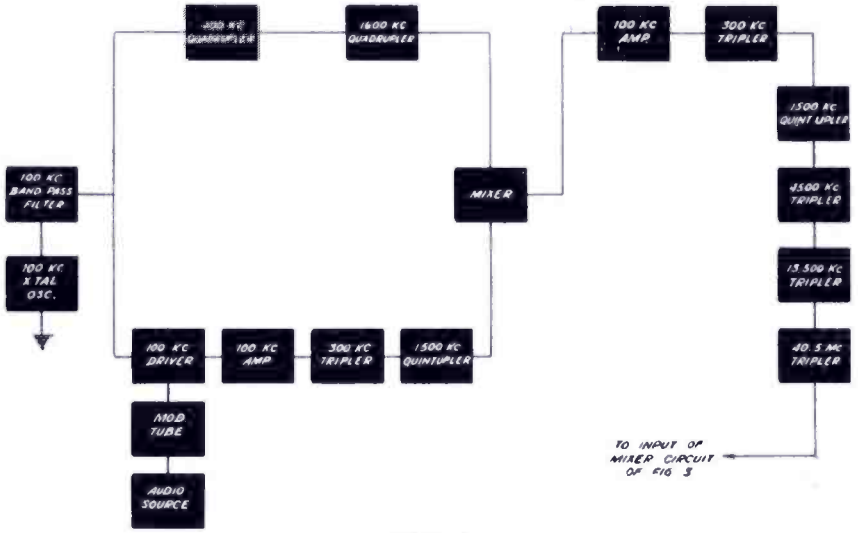


Fig. 7

parallel chain of multipliers, an unmodulated 100-kc signal (from the same crystal oscillator) is multiplied by factors of 4 and 4, to produce a 1600-kc voltage, with no phase deviation. The 1600-kc signal serves as a constant-frequency heterodyning source, converting the phase modulated 1500-kc signal to the difference frequency of 100 kc, which is taken from the mixer plate through a low-pass filter. The phase deviation in this 100-kc signal is increased over the modulator circuit deviation by a factor of 15. Hence, the apparent sensitivity of the modulator (degrees phase shift/audio voltage) is increased by 15 to 1, allowing a reduction of non-linear distortion products (at low values of audio frequency) by a factor greater than 15 to 1.

The 100-kc heterodyne signal is then fed through a subsequent chain of multipliers which raises the carrier frequency to 40.5 Mc, with a phase deviation of $405 \times 15 = 6075$ times that produced at the modulator. For distortion measurements this output signal is applied to the converter of Figure 3, which feeds into the linear monitor detector.

Figure 8 is a diagram of the phase-modulator circuit employed in the laboratory generator. The component circuits indicated by the blocks in Figure 7 are, in general, self-explanatory, but a few remarks determined from development experience are in order.

The coupling transformers for the 300 kc, 400 kc, 1500 kc, and 1600 kc stages are double-tuned transformers employing inductive coupling. The coupling and damping of each unit was adjusted to provide adequate band-pass characteristics. In this respect it must be noted that those stages following the heterodyne mixer require greater percentage band width than those preceding the mixer by a factor equal to the heterodyne ratio (15 in this case).

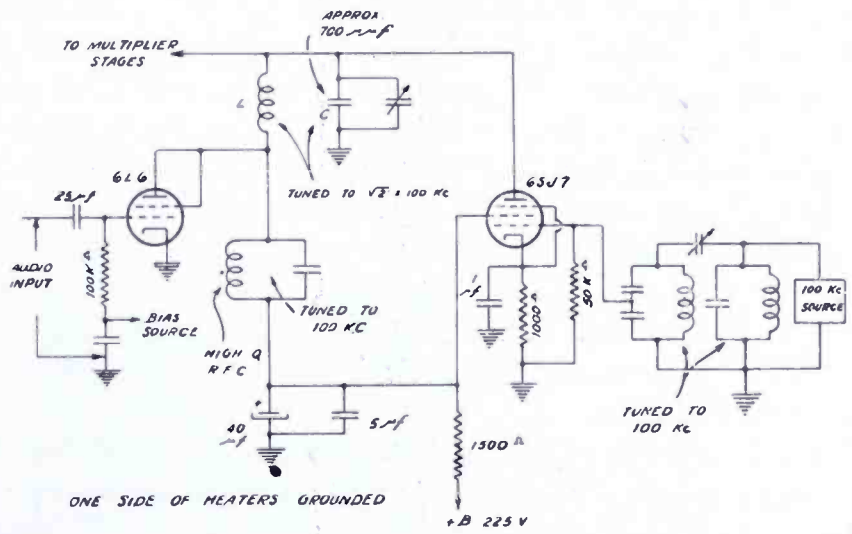


Fig. 8—Phase modulator circuit.

The 4.5 Mc, 13.5 Mc, and 40.5 Mc stages utilize single resonant circuits, each tuned with a 3-30 $\mu\mu f$ trimmer and all loaded or damped sufficiently to make the pass band essentially flat and thus to minimize the generation of amplitude modulation and phase distortion in these stages.

Jacks are provided in each stage to facilitate the measurement of d-c grid current, as a means for determining correct tuning conditions in preceding stages. The grid bias resistors are adjusted to produce maximum harmonic output in each stage. Thus, for example, in a tripler whose plate load is tuned to 300 kc, the tripling action is maximized by adjusting the tripler's grid leak to produce maximum grid current in the following tube. This bias adjustment is carried out in all stages. The grid-circuit time constants of all stages are made sufficiently low to permit the multiplier grids to follow any a-f variations in the amplitude of the signals, thereby producing limiting action.

Experimental Procedure

An essentially distortionless audio source is required for measurements of phase distortion in the frequency-modulation generator. This happens because the amplitude of the nth harmonic in the developed frequency deviation is proportional not only to the amplitude of the nth harmonic in the audio-source voltage, but also to the order of the harmonic, since the frequency deviation is proportional to the *rate* of phase deviation. Thus, 1 per cent 3rd harmonic in the audio input would produce (in an otherwise distortionless system) 3 per cent harmonic in the resultant frequency deviation, which would then appear as 3 per cent 3rd harmonic in the audio output of the monitor detector.

This problem may be circumvented in two ways, i.e., first integrate the audio voltage before application to the modulator, in which case the device produces frequency-modulated signals, requiring constant audio input for constant-frequency deviation as the audio frequency is varied. The source harmonics are then reduced in magnitude by an amount proportional to their order. The alternative method was used in the experimental work; an audio source of good waveform was used in conjunction with a series of band-pass filters, each capable of passing an octave of frequencies, starting at 30 cycles and ending at 8000 cycles. No integration was employed: hence, the generator produced phase-modulated signals.

A General Radio Type 636-A wave analyzer was used for all distortion measurements. The harmonic content of the audio source was found to be less than 0.05 per cent for all audio frequencies employed. The wave analyzer was then connected permanently across the monitor-detector output terminals. In parallel with this a 9-inch oscilloscope was connected to facilitate preliminary adjustments. In general the audio input to the modulator was adjusted to produce 100-kc deviation (for all audio frequencies) at the final carrier frequency of 40.5 Mc. The heterodyne oscillator in the detector unit was adjusted to produce 150 kc as the intermediate frequency for application to the monitor detector. This had previously been ascertained as lying in the region of optimum detector linearity.

As pointed out in a preceding section, there are two sources of distortion, in a generator of this type; the first occurs at low audio frequencies (less than 400 cycles) and is due to non-linear relationships in the modulator, which are accentuated by the large audio swings required.

This type of distortion is generally a function of the modulator bias, which controls the degree of matching between the curves of ϕ vs. R (Figure 6) and R_p vs. E_g of the modulator tube.

At higher audio frequencies (greater than 1000 cycles) the distortion is generally independent of modulator bias (since only a small audio swing is applied), but is determined, rather, by sideband clipping in the multiplier tuned circuits.

In adjusting the unit for minimum distortion products the audio input was set at 3200 cycles, and, with the monitor detector output indicating 100-kc deviation, and the wave analyzer reading 2nd harmonic distortion, adjustments were made by loading and tuning of the multiplier-load circuits until the distortion due to sideband clipping was minimized. This procedure was repeated at 7500 cycles. The overall distortion was reduced to less than 0.5 per cent rms, for 100-kc deviation, at these higher audio frequencies.

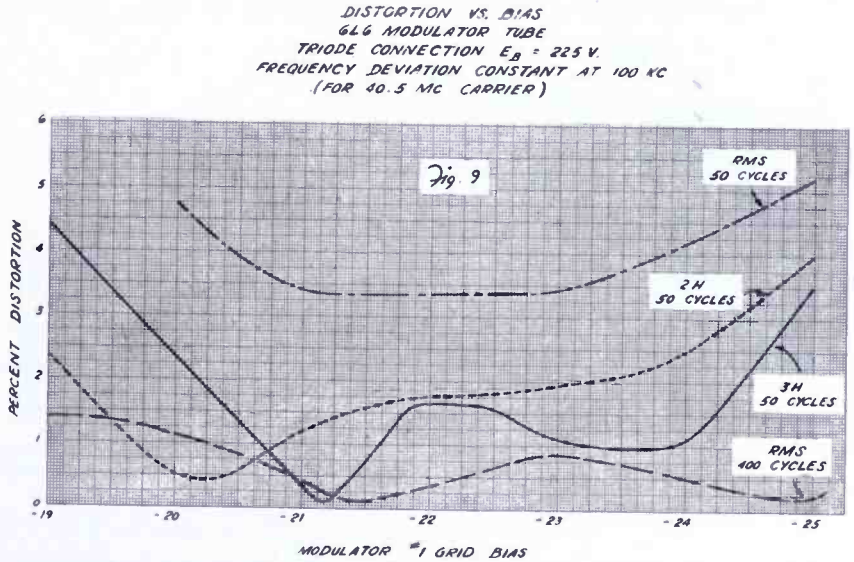


Fig. 9

The audio frequency was then reduced to 400 cycles, for distortion measurements at relatively larger modulator audio swings. With the frequency deviation adjusted for approximately 100 kc and the wave analyzer set for 2nd harmonic determinations, it was found that the amplitude of this harmonic was a function of modulator bias, and that a minimum value of less than 0.1 per cent could be attained with critical modulator bias adjustments. This sharply defined minimum is due to the dependence of the localized curvature of the R_p-E_g curve of the modulator upon its bias, and to the fact that, for a certain bias, correct matching between this curve and that of ϕ vs. R is obtained. Figure 9 shows curves of rms distortion vs. bias at 400 cycles modulating voltage.

At 50 cycles the applied audio voltage required for 100-kc deviation is 8 times that required at 400 cycles, and as a result, the degree of matching of the E_g vs. R_p and ϕ vs. R curve is less complete over an

audio cycle. However, the rms distortion at 50 cycles is relatively constant at 3.5 per cent over a 2-volt range in bias, for 100-kc deviation, as shown in Figure 9.

A large number of tube types (6J7, 6K7, 6V6, 6AC7/1852, 6AC7/1853, 6J5, 6L6) were used experimentally as modulators. The 6AC7/1852, for example, has higher sensitivity (kc deviation per audio volt), but the dependence of distortion on bias is more critical. The 6L6 was finally chosen as a result of its lower rms distortion at 50 cycles modulating frequency, and the relative constancy of distortion with bias.

Analysis of the modulator circuit shows that distortion produced at low audio frequencies, caused by mismatch of the R_p vs. E_g and ϕ vs. R curves, is due principally to insufficient curvature in the R_p vs.

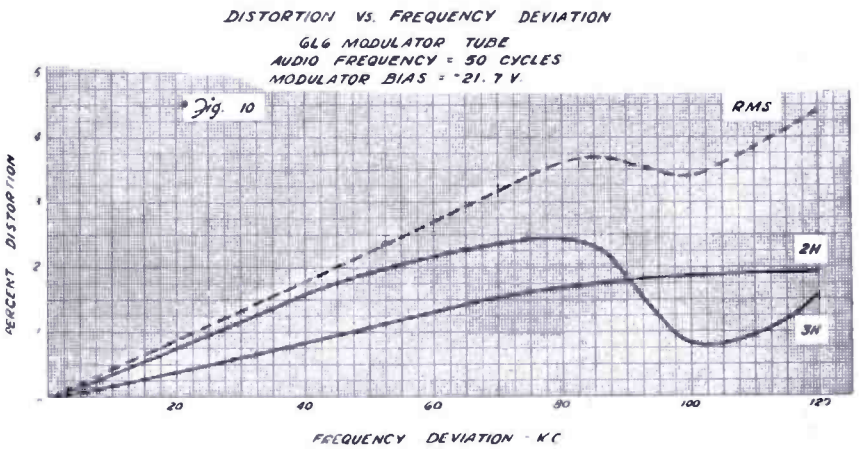


Fig. 10

E_g curve. This can be corrected, if desired, by some types of regenerative audio circuits, which tend to increase non-linearity in the modulator tube R_p curve, or by the use of an additional modulator to supply 2nd harmonic variations in R_p to produce additional curvature.

Experimental Results

Figure 10 shows the distortion incurred at 50 cycles modulating frequency vs. frequency deviation (in the 40.5 Mc carrier) taken at a value of bias for which the 2nd harmonic distortion is at a broad minimum.

Figure 11 is a plot of rms distortion vs. audio frequency for two values of frequency deviation. Note that the distortion is essentially independent of deviation except at frequencies below 200 cycles, for which the audio swings (approximately 20 peak-peak volts at 50 cycles) are large, causing considerable mismatch in the phase and R_p curves, which, at lower swings, are normally coincident.

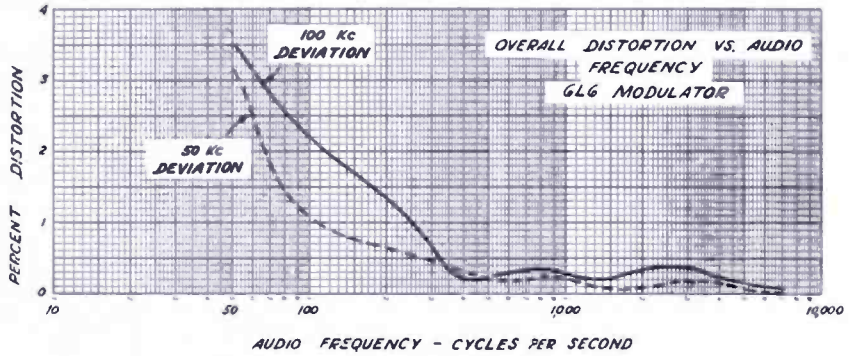


Fig. 11

The distortion at high audio frequencies is independent of modulator bias, but is very much a function of tuning, even with sufficient circuit damping. This is particularly the case in the 300-kc tripler following the mixer stage, which must accommodate a band of frequencies 15 times as great as that encountered in circuits directly following the modulator.

Attention must be directed toward minimizing amplitude modulation in the frequency-modulated wave, particularly at the lower signal frequencies, for which the phase deviation is small. As shown in Appendix II, n per cent amplitude modulation can produce n per cent 2nd harmonic distortion in the frequency deviation, due to sideband cancellations, if the phase deviation is sufficiently low to permit the representation of the wave by a carrier and a pair of first-order sidebands.

APPENDIX I

Analysis of Phase Modulator Load Circuit:

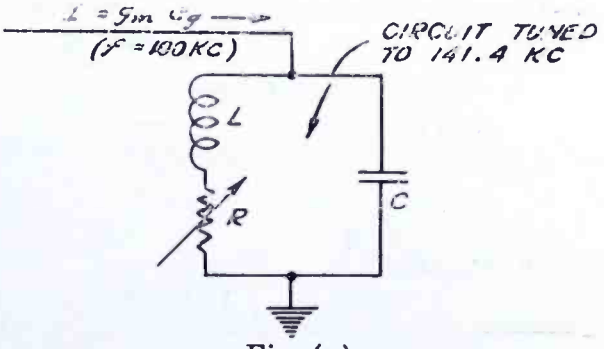


Fig. (a)

(1) To demonstrate the constancy of tuned impedance as R is varied. $\omega_s = 2\pi \times$ input frequency (100 kc)

$$Z = \frac{(R + jL\omega_s)\dfrac{1}{jC\omega_s}}{R + j\left(L\omega_s - \dfrac{1}{C\omega_s}\right)} = \frac{(RC\omega_s + jLC\omega_s^2)\dfrac{1}{jC\omega_s}}{RC\omega_s + j(LC\omega_s^2 - 1)}$$

The circuit is resonant at $\sqrt{2} \times$ the source frequency, i.e., C in the above figure is half the value required for circuit resonance at the source frequency. Therefore, $LC = \dfrac{1}{2\omega_s^2}$

and
$$\not{Z} = \frac{\left(RC\omega_s + j\dfrac{\omega_s^2}{2\omega_s^2}\right)\dfrac{1}{j\omega_s C}}{RC\omega_s + j\left(\dfrac{\omega_s^2}{2\omega_s^2} - 1\right)} = \left(\frac{2RC\omega_s + j1}{2RC\omega_s - j1}\right)\frac{1}{j\omega_s C}$$

The absolute magnitude of $\not{Z}$ is therefore $|\not{Z}| = \dfrac{1}{\omega_s C}$, which is independent of R.

(2) The phase vs. R equation of the circuit is readily determined from the above, viz:

$$\not{Z} = \left(\frac{2RC\omega_s + j1}{2RC\omega_s - j1}\right)\frac{1}{j\omega_s C} = \frac{1}{\omega_s C}\left(\frac{1 - j\,2RC\omega_s}{2RC\omega_s - j1}\right)$$

total phase shift ϕ = phase shift in numerator θ minus phase shift in denominator α

$$\phi = \theta - \alpha$$

$$\phi = \tan^{-1}(-2RC\omega_s) - \tan^{-1}\left(\frac{-1}{2RC\omega_s}\right)$$

$$\tan\phi = \frac{\tan\theta - \tan\alpha}{1 + \tan\theta\tan\alpha}$$

Hence,

$$\tan\phi = \frac{-2RC\omega_s + \dfrac{1}{2RC\omega_s}}{1 + \dfrac{2RC\omega_s}{2RC\omega_s}} = \frac{1}{2}\left(\frac{1}{2RC\omega_s} - 2RC\omega_s\right)$$

$$\phi = \tan^{-1}\left(\frac{1}{4RC\omega_s} - RC\omega_s\right)$$

This is plotted in Figure 6

APPENDIX II

Proof that amplitude modulation in a low-deviation phase-modulated wave can produce frequency-modulation distortion equal approximately to the per cent amplitude modulation, if this factor is small.

First assume that the phase deviation is sufficiently low to justify the use of only the first-order sidebands. The signal then has the form of Figure (b).

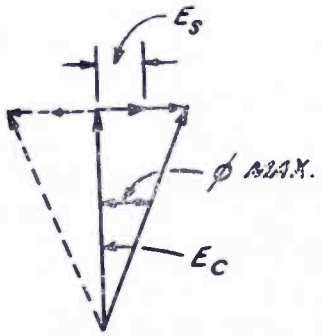


Fig. (b)

$$\phi = \phi_{max} \sin \omega_a t$$

With no added amplitude modulation

$$\phi = \tan^{-1} \frac{2E_s}{E_c} \sin \omega_a t$$

where E_s = phase-modulation sideband amplitude
E_c = carrier-voltage amplitude
$\omega_a = 2\pi \times$ modulating frequency

Now let the carrier be amplitude modulated at the same audio frequency, viz:

$$e_c = E_c(1 + m \sin \omega_a t)$$

The phase angle is given by

$$\phi = \tan^{-1} \frac{2E_s \sin \omega_a t}{E_c(1 + m \sin \omega_a t)}$$

Let $\frac{2E_s}{E_c}$ be a constant, determined by the amount of phase modulation.

Then $$\phi = \tan^{-1} K \frac{\sin \omega_a t}{1 + m \sin \omega_a t}$$

For small angles it is permissible to write

$$\phi = K \frac{\sin \omega_a t}{1 + m \sin \omega_a t}$$

also for m, (per cent amplitude modulation $\div$ 100) less than 0.10, it is sufficiently accurate (within 1 per cent) to write $\phi = K \sin \omega_a t \ (1 - m \sin \omega_a t) = K \sin \omega_a t - Km \sin^2 \omega_a t$. This can be expanded to give

$$\phi = K \left(\sin \omega_a t - \frac{m}{2} + \frac{m}{2} \cos 2\omega_a t\right)$$

The ratio of coefficients of the 2nd harmonic and fundamental audio terms is $m/2$, which implies the existence of $\frac{m}{2} \times 100$ per cent 2nd harmonic in the phase-modulated wave.

The frequency deviation is equal to $\frac{1}{2\pi} \frac{d\phi}{dt}$

Performing this operation we get

$$\frac{d\phi}{dt} = K \left(\omega_a \cos \omega_a t - 2\omega_a \frac{m}{2} \sin 2\omega_a t\right)$$

$$\Delta f = \frac{1}{2\pi} \frac{d\phi}{dt} = K f_a \left(\cos \omega_a t - m \sin 2\omega_a t\right)$$

where Δf is the frequency deviation (instantaneous) and f_a is the modulating frequency. It is seen that m per cent amplitude modulation will therefore produce m per cent 2nd harmonic distortion in the frequency deviation of the system.

A NEW CHEMICAL METHOD OF REDUCING THE REFLECTANCE OF GLASS

By

F. H. Nicoll

Research Laboratories, RCA Manufacturing Company, Inc., Camden, N. J.

Summary—A new chemical method of reducing the reflectance of glass is described. It is compared experimentally with previously known chemical methods of reducing the reflection, and is shown to be superior in many respects. The new method produces a tough hard film of very low reflecting power. The treatment involves exposure to hydrofluoric acid vapor and is applicable to large sheets of glass. The process requires neither vacuum nor expensive equipment and is suitable for many optical glasses. A number of possible uses of nonreflecting glass produced by this method are mentioned. Photographs are given of several examples of these applications, including cathode-ray-tube faces, ground-glass screens on cameras, and glass covers for photographs and pictures.

Introduction

THE use of an externally applied transparent film on glass, for the purpose of reducing the reflection of light from its surface, has been known for some time.[1,2,3] Considerable practical use has been made of evaporated films one-quarter of a wavelength thick.[4] These films have been used to increase the transmission of many optical devices employing lenses, prisms, etc., with very satisfactory results. Many obvious applications of such films have not, however, been realized, due either to lack of hardness of the film or to the high cost of vacuum evaporation.

In striving for improved contrast on television cathode-ray tubes, it has been observed that the contrast of the picture is always greater when the picture is viewed in the dark. This is partly due to the scattering of extraneous light from the fluorescent screen, but there is also the effect of specular reflections from the surface of the glass. Such reflections of illuminated surroundings noticeably reduce the contrast. In the case of monitors and oscilloscopes used on outdoor equipment, specular reflection of the observer himself may cause considerable annoyance. The elimination of such extraneous reflections is very desirable. With this and other applications in mind, an investigation has been carried out of chemical methods for reducing these surface reflections.

Experiments on the Old Chemical Methods of Treating Glass

Chemical methods of treating glass to reduce its reflection have been known for a very long time. Early workers such as Taylor,[5] Kollmorgan,[6] and Wright[7] have observed and investigated the effects of a number of chemical solutions and the films which they produced. Later Vasicek[8] measured the indices of refraction of such films and showed them to be quartz. Similarly, Blodgett[3] came to the conclusion that the film was quartz. She obtained reflections of 21 per cent and

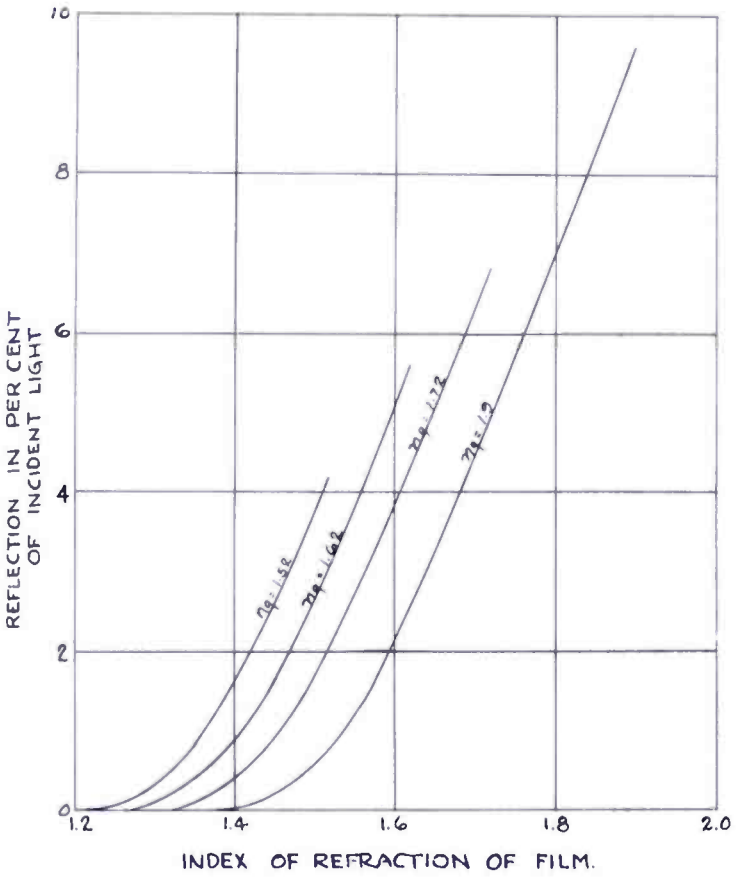


Fig. 1—Theoretical reflection from glass coated with a transparent film ¼ wavelength thick, in per cent of incident light.

30 per cent of the original reflection by chemically treating glasses having refractive indices of 1.72 and 1.61, respectively. Recently she has applied the technique to the making of gauges[9] for measuring the thickness of various monolayers. During the preparation of this paper, Jones and Homer[10] have published the results of some further experiments on the early chemical methods of etching lead glass.

The conditions for extinguishing reflection of monochromatic light have been given very fully by Blodgett[3] for the case of glass coated with a thin transparent film. If multiple reflections are neglected, and light of normal incidence only is considered, then for films of optical

thickness one-quarter of a wavelength viewed in air of refractive index 1, the resultant reflected intensity is given by

$$I_R = (B - C)^2$$

where $$B = (1 - n_f)(1 + n_f)$$

and $$C = (n_f - n_g)/(n_f + n_g)$$

The quantities n_f and n_g are the refractive indices of the film and glass respectively. The incident light intensity in this case is taken as unity. The resultant reflection I_R has been calculated for various

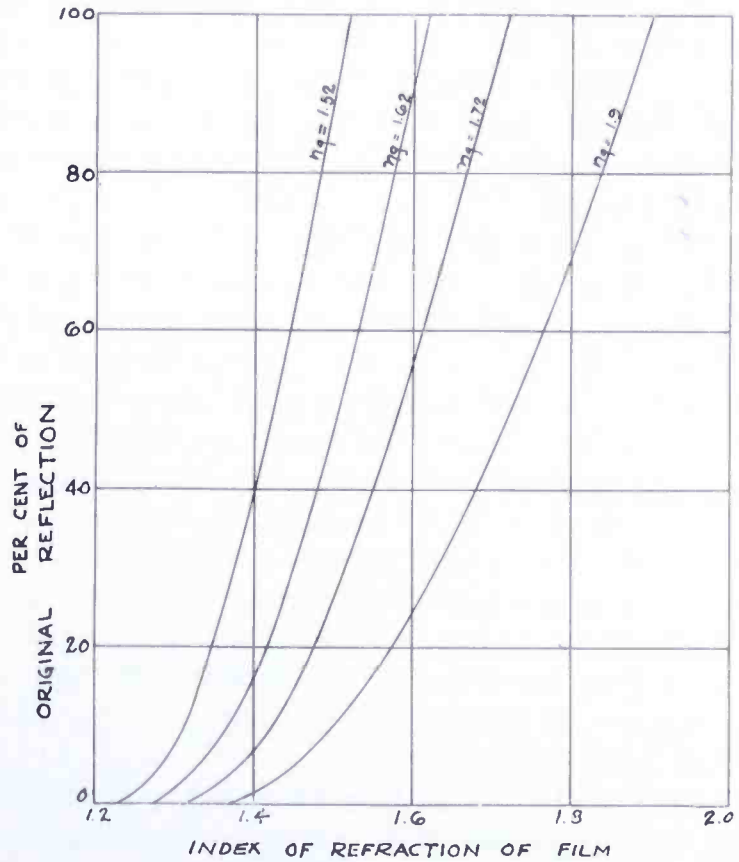


Fig. 2—Theoretical reflection from glass coated with a transparent film ¼ wavelength thick, in per cent of original reflection.

glasses and films of different indices of refraction. The curves are plotted in Figure 1 where the reflection is given in per cent of the incident light. They extend from values of $n_f = n_g$ down to the value of n_f corresponding to minimum reflection.

These curves are useful for comparing the results for different glasses. When one glass only is involved it is more convenient to have the reflected intensity in per cent of the reflection from the untreated glass. For this reason the curves have been replotted in Figure 2. With these curves it is possible to estimate the index of refraction of the film from the intensity of the reflected light. When the film thick-

ness has been adjusted to give minimum reflection on a piece of glass of known index of refraction, then the intensity of the reflected light gives a measure of the index of refraction of the film. This is most easily obtained from Figure 2 since the simple photocell method of measurement lends itself more readily to a comparison of the intensities of the reflected light from the treated and untreated glass, than to absolute measurements.

Preliminary experiments on the chemical method of producing transparent films were made using a number of inorganic solutions such as were employed by the early workers. Samples of various optical glasses were immersed in these liquids for varying lengths of time. The reflection was compared with that from the untreated glass using a photocell and a tungsten light source for the measurements. No filter was used, since experiments indicated that it was not necessary for an approximation to the sensitivity curve of the eye. It was found that with many solutions the reflection of flint glass of refractive index about 1.6 was reduced to about 25 or 30 per cent by prolonged treatment at 80 degrees Centigrade. Crown glass of index about 1.52 was more difficult to affect and the reflection was never reduced below about 70 per cent of its original value and in most cases was unaffected. Referring to Figure 2 we see that the results on both flint and crown glass correspond to a film having a refractive index between 1.4 and 1.5. This is in agreement with the assumption of other workers that the film was quartz. Because a greater percentage reduction was desired, it was clear that a heavier flint glass would have to be used. Since X-ray protection glass has an index of about 1.7 and has a very high lead content, it appeared to be very suitable. The high lead content is necessary since the portion of glass removed by the old chemical processes is mainly lead oxide.

Various inorganic solutions, such as were used by some of the early experimenters, were tried for their action on X-ray glass. It was immediately found that the majority of the solutions produced tough transparent films showing interference colors.

The amount of chemical action, and, hence, the thickness of the film, was found to be a function of time, temperature, and concentration. A limited range of these variables was investigated and no attempt was made to correlate them all in a systematic manner. The measured reflection for various film thicknesses was, however, obtained as a function of time. Figure 3 shows this reflection in per cent of the reflection from untreated glass, as a function of the time of immersion in a one per cent solution of sodium acid phosphate at 50 degrees Centigrade. The process caused the formation of a white film of soluble lead salts on the surface of the glass, in addition to the low reflection film of quartz.

The reflection measurements and estimate of color were made after this white deposit had been washed off. It will be noticed that the second-order minimum corresponding to a thickness of ¾ wavelength has been reached and passed. The colors of the film observed by reflected light are noted along the time scale. These indicate what colors are not being reflected. For instance, a purple film means red and blue are being reflected while green and yellow are not. A film of this color gives a minimum reflection for light most sensitive to the eye. The minimum of 14 per cent reflection fits in fairly well with the result expected from Figure 2 for a film of quartz. The film was very hard and could be rubbed off only with a rough lap. It could be completely removed by a quick dip in boiling NaOH. This process was, in fact, an excellent method of cleaning the samples before an experiment in order to get consistent results.

The time required to produce a purple film of minimum reflection is also a function of temperature. Figure 4 shows the approximate relationship for X-ray glass. It can be seen that the time is reduced by a factor of nearly ten as the temperature is increased from room temperature to 80 degrees Centigrade.

A number of additional inorganic solutions were also tried, all of which were effective to a greater or less degree. The table below gives some of the solutions used and the time necessary to produce a purple film.

TABLE I

Efficiencies of various solutions in producing a purple film of minimum reflection on X-ray protection glass.

Time of Exposure in Min.	*Solution*	*Concentration*	*Temperature °C*
7	$Fe_2(SO_4)_3$	1%	26°
2	$Fe_2(SO_4)_3$	1%	80°
50	NaH_2PO_4	1%	27°
15	NaH_2PO_4	1%	50°
9	NaH_2PO_4	1%	63°
5	NaH_2PO_4	1%	80°
45	H_3PO_4	1%	80°
No effect	H_3PO_4	Conc.	80°
100	$K_2Cr_2O_7$	2%	26°
2	$K_2Cr_2O_7$	2%	80°
5	$CuSO_4$	2%	80°
No effect in producing interference colors	NaOH	2%	27°

These chemical films, on being left exposed for a period of days, gradually increased in reflecting power, sometimes as much as two or three times. However, it was found that heating restored them to nearly their original values. This was thought to be due to the absorption and subsequent removal of water vapor. A similar effect has been noted in a recently published paper,[10] in which the authors mention that baking stabilizes the film.

There seems to be a very limited use for this type of chemical film where a reduction in reflecting power is desired. The fact that the films are very hard is an advantage but, on the other hand, the process is applicable only to the heavier flint glasses and the reduction in reflection is limited by the index of the film. This is a serious limita-

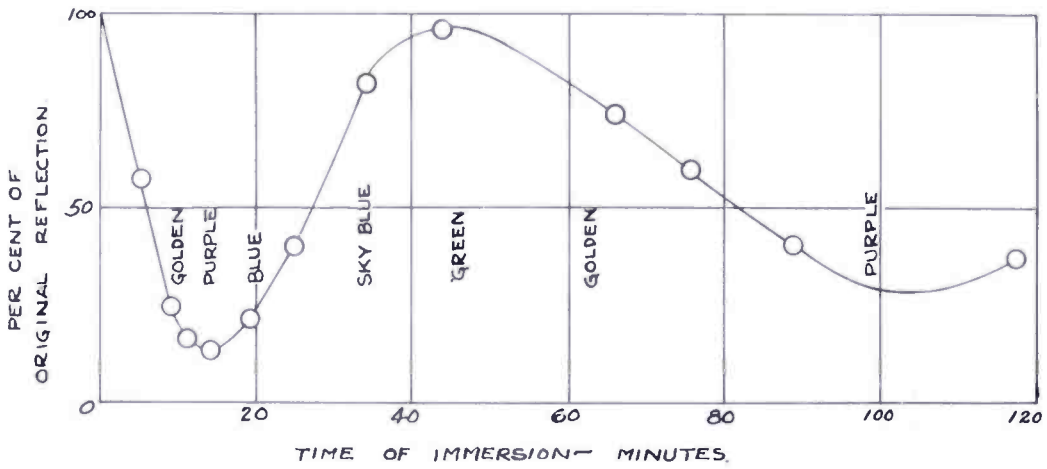


Fig. 3—Reflection as a function of time of immersion for X-ray protection glass in 1 per cent NaH_2PO_4 at 50° C.

tion for applications such as cover glasses for meters, viewing windows, etc. Because of the higher index of refraction of X-ray glass, the initial reflection is about 1.6 times that of window glass, as can be seen from Figure 1. It follows that even if the reflection of X-ray glass is reduced to about 15 per cent, this only corresponds to a reduction of window glass reflection to about 24 per cent. This in itself is not low enough for most purposes. In addition, the cost of X-ray glass is sufficiently high to prohibit its use for some applications.

For most applications requiring non-reflecting glass, a method of treating glasses of low refractive index is required. At the same time the method should be applicable to large pieces such as sheets of window glass. Further experiments were therefore undertaken in the hope of finding a chemical method which would satisfy these requirements.

A New Method of Producing Films of Low Reflecting Power on Glass

As long ago as 1900, Rayleigh[11] pointed out in one of his papers that hydrofluoric acid diluted one part in 200 of water removed a thickness of glass corresponding to about ¼ wavelength of light per hour. At the same time the glass, if agitated remained perfectly polished. The thickness corresponding to ¼ wavelength is of particular interest in the production of non-reflecting films. If some compounds of low refractive index were left on the glass while the remainder of the glass was dissolved away, then a dilute solution, such as Rayleigh used, could produce a film of low reflecting power.

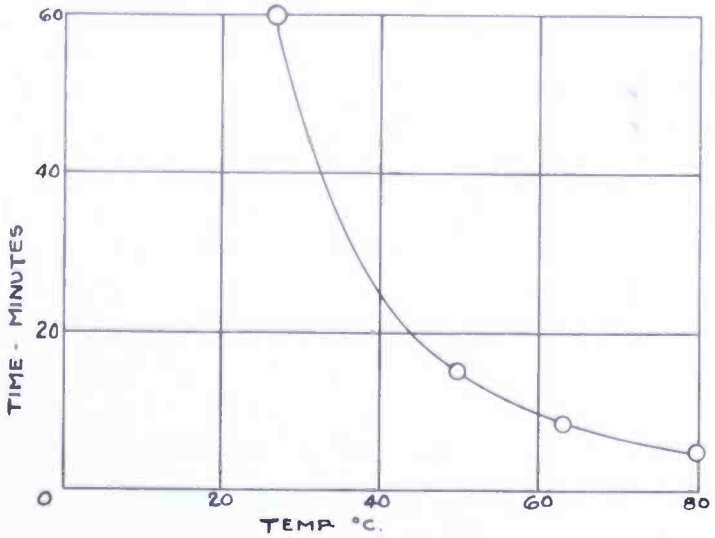


Fig. 4—Time of formation of purple films of minimum reflection as a function of the temperature of the solution, for X-ray protection glass in 1 per cent NaH_2PO_4.

With a view to investigating the effect of hydrofluoric acid liquid and vapor at the same time, some samples of window glass were placed half in the solution and half out. A solution of one part concentrated hydrofluoric acid (48 per cent) to 200 parts of water (¼ per cent HF solution) was used, the solution being contained in a 50 cc waxed beaker with a waxed glass lid. After 64 hours at room temperature, the glass in the liquid had been thinned appreciably, but was still polished in appearance. Immediately above the liquid, however, were interference colors ranging from blue near the liquid to straw color about 2 or 3 mm above the liquid. These colored films appeared to be perfectly transparent and the glass underneath did not appear to be etched or pitted by the acid vapors.

This result was sufficiently encouraging to justify further investigation. Experiments were therefore continued using 1, 2, and 4 per

cent HF solutions with the sample suspended horizontally over the liquid. The results were now more uniform and the films could be tested and examined more readily, since the area was much larger. It was observed that condensation of the water solution of HF on the film during its formation removed the film or made it very fragile. For this reason, it was found desirable to cool the bottom of the container to a temperature below that of the glass in order to prevent condensation. Water from the faucet at about 17 degrees Centigrade, about 10 degrees below room temperature, was found to be satisfactory.

The most satisfactory arrangement for producing uniform films seemed to consist of a shallow tray two or three inches high with the

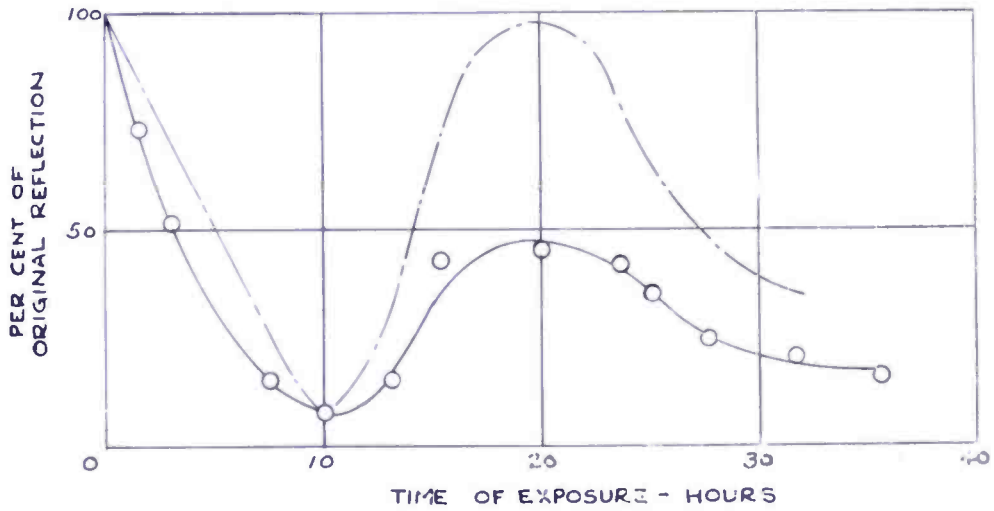


Fig. 5—Reflection as a function of time of exposure for window glass in the vapor of 1 per cent HF.

flat piece of glass to be treated placed over the top, its lower face exposed to the vapor of the acid. The tray itself was supported by blocks about ¼ inch high. Water to a depth of about ½ inch was circulated around and under the tray at a temperature of about 10 degrees below that of the room. The glass was usually cleaned with powdered chalk and water, although this was not always necessary. Under these conditions, with a 1 per cent solution, about 6 to 10 hours (depending on temperature) was required to produce a film of purple color on a piece of ordinary window glass. This film reduced visible reflections to a minimum, about 6 to 10 per cent of that from untreated glass.

Figure 5 shows a curve of reflection against time of exposure for one side of a piece of glass exposed on a tray 2 inches deep. The water bath was kept at about 15 degrees Centigrade and the average room temperature was 25 degrees Centigrade. The glass used was

⅛-inch Libbey Owens Ford "Double Strength" window glass, glazing quality. The acid solution was 1 per cent HF and about 50 cubic centimeters were used, covering the bottom of the tray to about 1/16 inch. Some of the products of the etching remain as a white film on the glass, but these are subsequently washed off with water. The reflection measurements for this curve were made with this white film remaining on the glass, but the approximate values of the reflection after the film was washed off are given by the broken line. The curve of Figure 5, like that of Figure 3, extends through the first-order minimum to the second-order minimum. The colors obtained corresponded to those given in Figure 2, but they were more vivid since the refractive index of the film was more nearly that required for extinction of the reflection at the minimum.

Some investigation was made of the effect of concentration and there was evidence that about 5 per cent HF was the greatest concentration that would produce a suitable hard film without heavy etching. The more concentrated the solution, the more rapid the production of the film. For instance, the vapor from the 48 per cent HF produced a purple film in about two minutes at room temperature, but the glass was so heavily etched that it became translucent.

These films produced by hydrofluoric acid vapor, were proved to be true interference films of high light transmission by comparing the transmission of glass with and without a film on it. The results were similar to those with evaporated transparent films. The transmission was considerably increased at the same time that the reflection was decreased. A pile of nine plates of window glass had a measured transmission of 51 per cent, which is quite close to the theoretical value. A similar pile of nine plates of treated window glass with the reflection at each face reduced to between 8 and 10 per cent gave a measured transmission of about 88 per cent, which is within a few per cent of the calculated value if the absorption in the glass is taken into account. In addition to the increase in transmission, an image viewed through the treated plates had much better contrast, and was practically free of multiple images. These results show that the transparency of the films produced is very high and that very little etching takes place.

The Behavior of Various Glasses Under Hydrofluoric Acid Treatment

A number of different types of glass were tested for their reaction to hydrofluoric acid vapor. In a number of cases non-reflecting films were produced and the results, where possible, were correlated with the composition of the glass. Libbey Owens Ford window glass produced very fine nonreflecting films. In particular, the "Single

Strength" and "Double Strength" types were satisfactory. Plate glass was usually satisfactory, but in some cases spurious marks were brought up, although no evidence of polishing marks was seen.

Several of the glasses forming good films by the above technique were known to contain more than 10 per cent calcium oxide. On the other hand, some of those not forming a film were known to contain less than 5 per cent calcium oxide. This led to the belief that the films were calcium fluoride formed by the action of the hydrofluoric acid vapor, the other products being removed as vapors or during washing in water. This view is borne out by reference to Figure 2. Glass with an index of about 1.52 has its reflection reduced to about 6 per cent at the minimum. This is seen to indicate a film of index slightly less than 1.3. Calcium fluoride is one of the few possible substances that could be formed from the glass and possess such a low index of refraction, unless a skeleton film is formed. This latter possibility seems to be doubtful since a drop of oil could be wiped off the film without affecting the reflection or color of the film.

A considerable number of small samples of various optical glasses were also tested. The compositions were not known, but the index of refraction n and the approximate values of the dispersion $(n_F - n_C)$ of the various samples are given below. n_F and n_C are respectively the indices of refraction for the F and C lines of the solar spectrum. The glasses are tabulated in three columns according to the ease with which the hydrofluoric acid vapor produced a nonreflecting film. Those listed as good formed hard films of low reflecting power with little or no etching. Those listed as fair tended to etch visibly and in some cases the films were only faintly colored, indicating a film of too high an index of refraction. The samples listed as poor usually produced heavy etching and no interference films, while in some cases only a faintly colored brownish film could be obtained.

TABLE II

Quality of Films Produced on Various Glasses

Good		*Fair*		*Poor*
n	$(n_F - n_C)$	n	$(n_F - n_C)$	n
1.5147		1.5232		1.5123
1.5158		1.619	.0169	1.525
1.5159	.0095	1.6214	.0172	1.7 (X-ray glass)
1.5166	.0088			1.47 (Pyrex)
1.5177				
1.5179				
1.5243	.0085			
1.6204				
1.6214	.0170			

The compositions of the various glasses in the above table were not known, but it is highly probable that a number of the optical glasses contained at least 10 per cent calcium oxide. On the other hand, X-ray glass which formed no interference film contained no calcium oxide. These results were sufficient to show that quite a wide range of optical glasses can be treated by the HF vapor process. In particular it seems to be applicable to those glasses which do not have their reflections reduced appreciably by the old chemical methods. The process, therefore, extends very greatly the usefulness of the chemical methods.

Properties of Films Produced by Hydrofluoric Acid Vapor

Several properties of this type of nonreflective film have been noted and in a number of cases they have been compared with those of an evaporated film of a satisfactory type. Some chemical solubilities have been tabulated below for convenience. Solubility and insolubility are indicated by the letters *S* and *I*.

Table III

Chemical Properties of HF-Produced Films and Evaporated Films

Solution	*HF-Produced Film*	*Evaporated Film*
Dilute HF	S	S
Boiling NaOH	S	S
$(NH_4)_2CO_3$ sol'n	I	S
Dil. H_2SO_4	I	S
Cold Dil. HNO_3	I	S
Cold NaOH	I	S
Conc. HCl	I	S
Cold H_2SO_4, Conc.	I	S
Boiling H_2SO_4, Conc.	I	S
Cold Chromic Acid	I	S
Boiling Chromic Acid	I	S

From this table it can be seen that the chemically produced nonreflecting film is very much more resistant to chemical action than the evaporated film. The solubility of the chemically produced film in dilute HF is quite a useful feature in some respects. If it is desired to remove the film from one portion of a glass sample, it is possible to do this and leave a sharp boundary by merely dipping the sample in 8 to 10 per cent HF for a few seconds. The film can also be removed in any desired pattern by swabbing with this solution.

In the case of physical properties the difference between these two types of film is also important. The chemically produced film was much harder and resisted rubbing better than the evaporated films. It could

Fig. 6—Photograph of framed picture with glass front, one-half of which is nonreflecting.

also be cleaned with water, alcohol, or some commercial window cleaning solutions. However, rubbing with powdered chalk, such as is present in some window cleaners, gradually removed the films. Both types of film withstood heating to red heat in air. The HF-produced

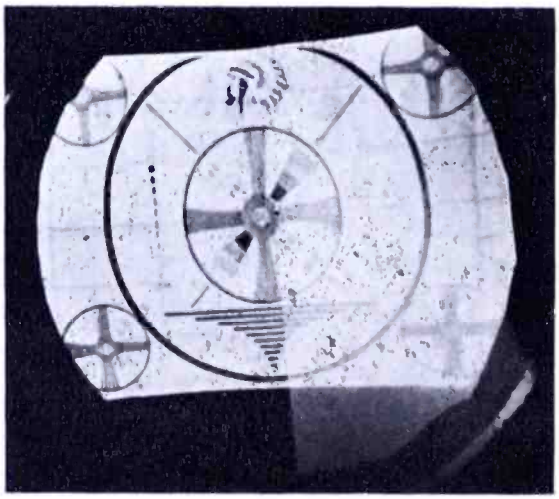


Fig. 7—Experimental cathode-ray tube face, one-half of which is nonreflecting both inside and out.

film on a piece of window glass was exposed to air in a furnace at 560 degrees Centigrade for about an hour, as a test of the ability of the film to withstand the temperature necessary to seal a glass window on the cone of a cathode-ray tube. This treatment altered the reflection of this particular sample by only 10 per cent.

Compared with the silica film produced by the old chemical method, the HF-produced film is somewhat less hard. However, the absolute reflection from one surface of a piece of window glass treated with HF may be as little as 0.25 per cent of the incident light, whereas the silica film produced on X-ray protection glass reflects about 0.94 per cent of the incident light. The HF-treated glass is therefore a marked improvement over the other type for purposes where a low reflecting power is desired. It should be pointed out here that these nonreflective films, like any other types, tend to increase their reflection as they accumulate dust. Occasional cleaning overcomes this difficulty. Similarly, finger prints are very noticeable since the presence of the oil increases the reflection considerably. It is best to wipe these off with alcohol or some similar solvent.

Uses of Non-Reflective Films Produced by Hydrofluoric Acid Vapor

The first use that suggests itself for these nonreflecting films is that of making lens surfaces nonreflecting, and of higher light transmission, as is now being done with evaporated films. Its use here, however, is limited by the fact that all optical glasses cannot be treated with hydrofluoric acid vapor to produce nonreflecting films. At the same time those which can be treated will take different times, depending on their composition. Nevertheless, its application to lenses should not be out of the question if suitable glasses are used.

This type of film can also be used to make meter fronts nonreflecting, as was demonstrated by Blodgett[3] with built-up films. These HF-produced nonreflecting films also gave a marked improvement in this case, particularly under glare conditions.

Store show windows may also be made nonreflecting by the HF vapor process since it is applicable to large sheets. However, no tests of its effectiveness in this application were made. On the other hand, photographs, etc., covered with nonreflecting glass had their appearance improved by the reduction of extraneous reflections. Figure 6 shows a framed picture, one-half of which has a nonreflective film on the glass front. In this case the reflection was about 10 per cent of that from the untreated glass.

Figure 7 shows the result of treating one-half of the face of an experimental cathode-ray tube. The face of the tube had the nonreflective film removed from one-half, but the other half was nonreflective on both inside and outside faces. The glass front on this tube was waxed on the cone of the cathode-ray tube. The photograph was taken with the tube face illuminated by diffuse light. On the one-half of the tube face, the nonreflective film has practically elimi-

nated the reflections, and the contrast of the scanned test pattern is much improved.

These fluoride films have also been tried on the ground-glass focusing screens used in cameras. As usually used, the unground side of the glass is exposed, and reflections from this surface interfere considerably in focusing a dim image on the screen. A nonreflecting film reduces this specular reflection so that it is practically invisible. This improves the contrast and makes it easier to focus accurately. A photograph of such a treated ground glass camera screen is shown in Figure 8. It was taken with the screen illuminated by diffuse light coming from the left. The lower half is seen to give a much clearer

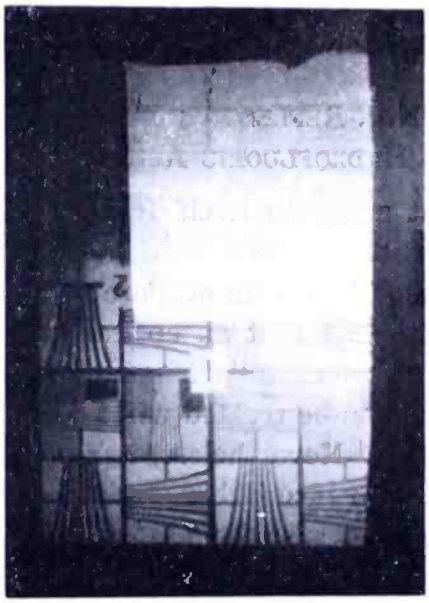


Fig. 8—Ground-glass focusing screen on a camera, one-half of which is nonreflecting.

image due to the presence of the nonreflecting film. At the left-hand edge of the upper half of the screen the image is also clear. This is because the light shield surrounding the screen prevented the diffuse light from reaching the ground glass. This is, therefore, the best image obtainable, since it corresponds to no stray illumination on the screen. It can be seen that the nonreflecting film produces practically the same result.

Similarly, a nonreflecting film has been found to improve the contrast on a ground-glass screen used for back-projected pictures such as those used for home movies. A sample screen was made about six inches square. An improvement was noted when an image was observed under ordinary room illumination with the unground surface treated, and a reduction in milkyness of the ground surface was noted when the ground side was treated.

BIBLIOGRAPHY

[1] J. Strong, "On a Method of Decreasing the Reflection From Non-Metallic Surfaces," *J.O.S.A.*, Vol. 26, p. 73, 1936.

[2] C. H. Cartwright and A. F. Turner, "Reducing the Reflection From Glass by Evaporated Films," *Phys. Rev.* 55, p. 595, 1939.

[3] K. B. Blodgett, "Use of Interference to Extinguish Reflection of Light from Glass," *Phys. Rev.* 55, p. 391, 1939.

[4] A. F. Turner, "Coating Lenses," *Phot. Technique*, p. 48, October 1940.

[5] H. D. Taylor, "The Adjustment and Testing of Telescope Objectives," *T. Cook*, York, England, 1896.

[6] F. Kollmorgen, "Light Transmission Through Telescopes," *Trans. Soc. Ill. Eng.* II, p. 220, 1916.

[7] F. E. Wright, "Characteristics of Optical Glass," *Ordnance Department Document*, No. 2037, p. 76.

[8] A. Vasicek, "New Method for Investigating the Refractive Index and Thickness of Interference Films on Glass," *Phys. Rev.* 58, p. 924, 1940.

[9] K. B. Blodgett, "Silica Gauge for Measuring Thickness by Means of Interference Colors," *Rev. Sci. Inst.* 12, p. 10, 1941.

[10] F. L. Jones and H. J. Homer, "Chemical Methods for Increasing the Transparency of Glass Surfaces," *J.O.S.A.*, 31, p. 34, 1941.

[11] Rayleigh, "On Polish," *Scientific Papers of Lord Rayleigh*, Vol. 4, p. 546.

AN ANALYSIS OF THE SIGNAL-TO-NOISE RATIO OF ULTRA-HIGH-FREQUENCY RECEIVERS

By

E. W. Herold

Research Laboratories, RCA Manufacturing Company, Inc., Harrison, N. J.

Summary—This paper presents an elementary analysis of the effect of the various sources of fluctuation noise on the signal-to-noise ratio of radio receivers. Because the noise induced in negative grids at high frequencies is included, the work is particularly applicable at ultra-high frequencies. It is found that the signal-to-noise ratio depends on the antenna noise; in addition, when bandwidth is not a consideration, it depends on the ratio of equivalent noise resistance to input resistance of the first tube, and, when bandwidth is a major consideration, on the product of input capacitance and equivalent noise resistance. The coupling from antenna to first tube is an important variable in receiver design and an optimum coupling is found which results in an improvement in signal-to-noise ratio. This optimum condition is often considerably different from the adjustment for maximum gain and, by its use, the noise induced in the grid becomes relatively unimportant. The noise from the second stage of the receiver is also evaluated. It is shown that the thermal noise from a wide-band interstage circuit may be made negligible by concentrating all the damping on the secondary side. Calculations of typical receiver arrangements using triode type 955 and pentode type 954 mixers are given for 300, 500 and 1,000 megacycles.

I—Introduction

THE useful reception of radio and other types of signals is limited, in the main, to those signals which exceed the unavoidable random fluctuation noise of the communicating system. Because there are other sources of noise in addition to normal random noise, such as static and man-made interferences, the signal must usually exceed the random noise by a considerable factor, under most conditions of reception. However, since random noise imposes a readily calculated absolute limit to sensitivity, it is usually used to designate receiver performance. This criterion is particularly appropriate when the noise which is inherent in the receiving antenna is included. At low frequencies, up to say 20 megacycles, the chief sources of noise in the usual receivers are due to atmospherics and allied causes, to thermal agitation in the antenna[1,2] and in the input circuits of the first tube,

[1] F. B. Llewellyn, "A Rapid Method of Estimating the Signal-to-Noise Ratio of a High Gain Receiver", *Proc. I.R.E.*, Vol. 19, pp. 416-420, March, 1931.

[2] R. E. Burgess, "Noise in Receiving Aerial Systems", *Proc. Phys. Soc.*, Vol. 53, Part 3, pp. 293-304, May 1, 1941.

and to shot-effect in the plate circuit of the first tube. At higher frequencies, however, the noise induced in the input grid by the passage of electrons[3,4,5] must also be considered as a contributing factor. When the gain of the first tube is low, it is also important to consider noise contributed by the coupling circuit to the following tube and also the noise contributed by the following tube itself.

In all cases, the most convenient evaluation of the total noise energy is made by referring all sources of noise to a given point in a receiver. For example, if the grid of the first tube is used as reference point, the noise voltage resulting from shot noise in the plate may be divided by the gain of the tube to give an entirely fictitious, but equivalent noise voltage at the grid reference point. This noise voltage may be added in the usual mean-squared manner to other independent noise voltages referred to this same point to give a total noise value. In estimating receiver performance, however, it is necessary to refer the received signal to this same point, if the signal-to-noise capability of the receiver is desired. In a companion paper,[6] D. O. North has shown that the absolute sensitivity of a receiving system depends upon a noise factor, N, of the receiver itself. The relations found in the present paper will show how N varies with tube and circuit constants and how receivers may be designed to minimize it.

Studies of signal-to-noise ratio in the input circuit of receivers have previously been made qualitatively by Llewellyn[1] and quantitatively by Williams[7] and by Fränz.[8] The present paper is an extension of the work of these men to include wide-band and ultra-high-frequency applications. In addition, a study of interstage coupling has been included for those cases when the gain of the first stage is low. This information is particularly necessary in ultra-high-frequency superheterodynes when a wide band is used with a low-gain converter stage.

II—Basic Relations

A tuned antenna of radiation resistance R_a will be assumed to be coupled to the receiver through a perfect transmission line of characteristic impedance $Z_o = R_a$. Coupling to the input of the first tube will

[1] loc. cit.

[3] Stuart Ballantine, "Schrot Effect in High-Frequency Circuits", *Journ. of Frank. Inst.*, Vol. 206, pp. 159-167, August, 1928.

[4] D. O. North and W. R. Ferris, "Fluctuations Induced in Vacuum-Tube Grids at High Frequencies", *Proc. I.R.E.*, Vol. 29, pp. 49-50, February, 1941.

[5] C. J. Bakker, "Fluctuations and Electron Inertia", *Physica*, Vol. 8, pp. 23-43, January, 1941.

[6] D. O. North, "The Absolute Sensitivity of Radio Receivers", RCA Review, January, 1942.

[7] F. C. Williams, "Thermal Fluctuations in Complex Networks", *J. I.E.E.*, Vol. 81, pp. 751-760, December, 1937.

[8] K. Fränz, "The Limiting Sensitivity in the Reception of Electric Waves and Its Attainability", *Elektrische Nachrichten-Technik*, Vol. 16, p. 92, 1939.

be assumed to be through a network which is the equivalent of a transformer whose leakage reactances are eliminated by tuning, and whose step-up, or effective turns ratio, is m*. An equivalent circuit then is as shown in Figure 1. The total input impedance of the tube with the antenna disconnected, as determined by circuit and lead losses as well as transit-time loading, is lumped as R_1 in the figure. The open-circuit antenna signal voltage is shown as e_a. Feedback effects which may change the signal-to-noise ratio in the tube will be neglected in this analysis. The noise sources which must be considered are:

1. Thermal and other noise in the antenna.
2. Thermal noise of circuit and leads.
3. Induced input electrode noise in first tube.
4. Plate noise of the first tube referred to the input circuit.
5. Noise of parts of receiver subsequent to the first tube, and also referred to the input circuit.

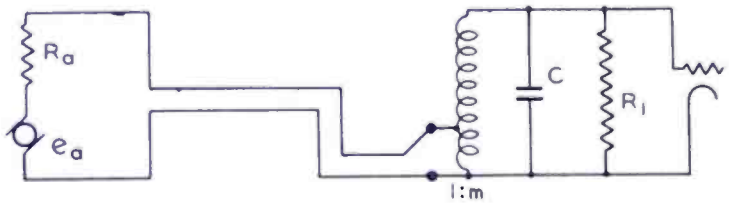


Fig. 1—Basic equivalent circuit of receiver input.

The noise of the antenna is conveniently introduced by considering this noise as if it were noise due to thermal agitation in the radiation resistance of the antenna and assigning to this resistance an effective temperature, *Ta*. Under ordinary circumstances, at frequencies below 20 megacycles, this effective temperature may be considerably higher than the ambient temperature.[9] At ultra-high frequencies, where directive antennas are employed and atmospherics are negligible, the value of *Ta* will more nearly approach the ambient temperature and may even go below it. When a conventional method of measuring receiver sensitivity is employed, a resistor at ambient temperature is substituted for the antenna resistance and the results obtained will correspond to *Ta* equal to the ambient temperature. These cases will be given special attention in this paper.

To distinguish between items 2 and 3 of the above list it will be convenient to divide R_1 into two parallel components, one having thermal noise and representing circuit and leads, and the other having induced grid noise and representing electronic loading. If the con-

* m is the ratio of output voltage to input voltage of the network.

[9] K. G. Jansky, "Minimum Noise Levels Obtained on Short-Wave Radio Receiving Systems", *Proc. I.R.E.*, Vol. 25, pp. 1517-1530, December, 1937.

ductances of these two components are g_Ω and g_e, respectively, then

$$g_\Omega + g_e = \frac{1}{R_1}$$

The noise induced in the input circuit by the passage of electrons has been found by North and Ferris[4] for tubes with an input control grid adjacent to the cathode. Their results for oxide-coated cathode tubes showed the noise to be equivalent to the thermal noise of a resistor of a value equal to the electronic loading and whose temperature is about 5 times room temperature. Furthermore, their work has shown that, theoretically, to a first approximation, induced grid noise may be added to the plate noise (referred to the grid) of the same tube as if they were independent sources of noise. Their results have made the application of this analysis to ultra-high frequencies possible.

With respect to induced grid noise in tubes whose control grid is not adjacent to the cathode, the problem is not yet completely solved, although a start has been made by Bakker.[5] The emphasis in the present paper will be on the more common tubes with oxide-coated cathodes and an adjacent control grid which is also the input grid. However, for the sake of completeness, the results which apply to the other forms of tube are given in an appendix.

One of the more common feedback effects in ordinary tubes at high frequencies is due to cathode lead inductance which introduces an input loading very similar in nature to the electronic loading. However, it has been shown[10] that this feedback leaves the signal-to-noise ratio unaffected, at least in the first approximation. In using the results of the present paper, therefore, input conductance due to cathode lead inductance should not be included in the evaluation of the electronic conductance, g_e. If the conductance due to the cathode lead is included in g_Ω as if it were an ohmic loss the error will be small in most practical cases.

The plate noise of the first tube will be referred to the grid in the usual manner by making use of an equivalent-noise-resistance concept.[11] In many practical cases, the gain of the first tube will be sufficiently high so that noise of succeeding stages will be negligible. However, to preserve complete generality, the other sources of noise will be assumed to be included by adding a second equivalent noise resistance which is

[4] loc. cit.

[5] loc. cit.

[10] M. J. O. Strutt and A. Van der Ziel, "Methods for the Compensation of the Effects of Shot Noise in Tubes and Associated Circuits", *Physica*, Vol. 8, pp. 1-22, January 1941.

[11] See Part V of B. J. Thompson, D. O. North and W. A. Harris, "Fluctuations in Space-Charge-Limited Currents at Moderately High Frequencies", RCA REVIEW, Vol. V, pp. 505-524, April 1941 and Vol. VI, pp. 115-124, July 1941.

calculated at the grid of the first tube by using the squares of the gains between this point and the actual sources of noise, in the usual manner. A later section will include the evaluation of this second equivalent-noise-resistance factor as applied to the second stage. The total equivalent noise resistance at the grid will be designated by R_{eq} and will be the sum of the noise resistance of the first tube and that calculated from all succeeding stages.

Using the above concepts, by assuming an overall receiver passband which is not wider than that of the circuits considered in the analysis, it now becomes possible to redraw the circuit to include all the noise sources. This is done in Figure 2 in which all noise sources are shown as constant-current generators except the equivalent-noise-resistance source, R_{eq} which is shown as a voltage. Throughout this analysis, Ta represents the effective antenna temperature, T_R represents ambient or room temperature in degrees Kelvin, k is Boltzmann's constant

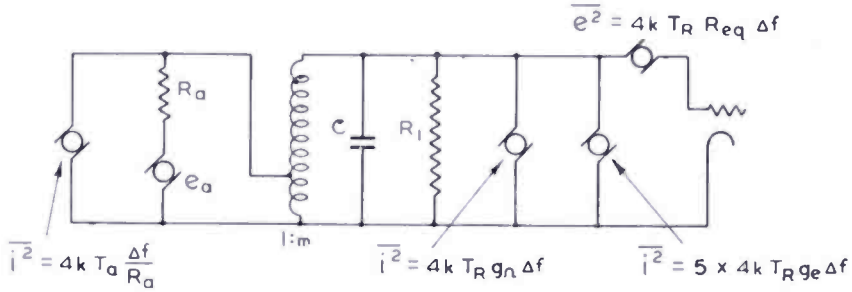


Fig. 2—Equivalent circuit of receiver input with noise sources included.

(equal to 1.37×10^{-23} joules per $^\circ K$) and $\triangle f$ is the overall effective passband of the receiver for noise purposes.[11] The induced noise in the input grid is shown in the form suitable for oxide-coated cathode tubes with control grid adjacent to the cathode.

Since the antenna resistance referred to the input circuit secondary is equal to m^2R_a (where m is the effective turns ratio, or step-up, of the input circuit), the antenna noise referred to the grid is just the thermal noise of a resistor m^2R_a at a temperature Ta in parallel with the resistance $R_1 = \dfrac{1}{g_\Omega + g_e}$ already there. Thus, another equivalent circuit can be shown in Figure 3a.

In order to simplify the analysis still further, let us replace the reflected antenna conductance, $1/m^2R_a$, by the symbol g_a. The two resistances of Figure 3a in parallel will then be equivalent to the single resistance of Figure 3b. Furthermore, by remembering that

[11] Thompson, North and Harris, loc. cit.

$$g_\Omega + g_e = \frac{1}{R_1}$$

we can rewrite the expression for the constant-current noise generator as shown in the latter figure. It will be well to keep in mind that the condition for maximum gain (impedances matched) corresponds to $g_a R_1 = 1$.

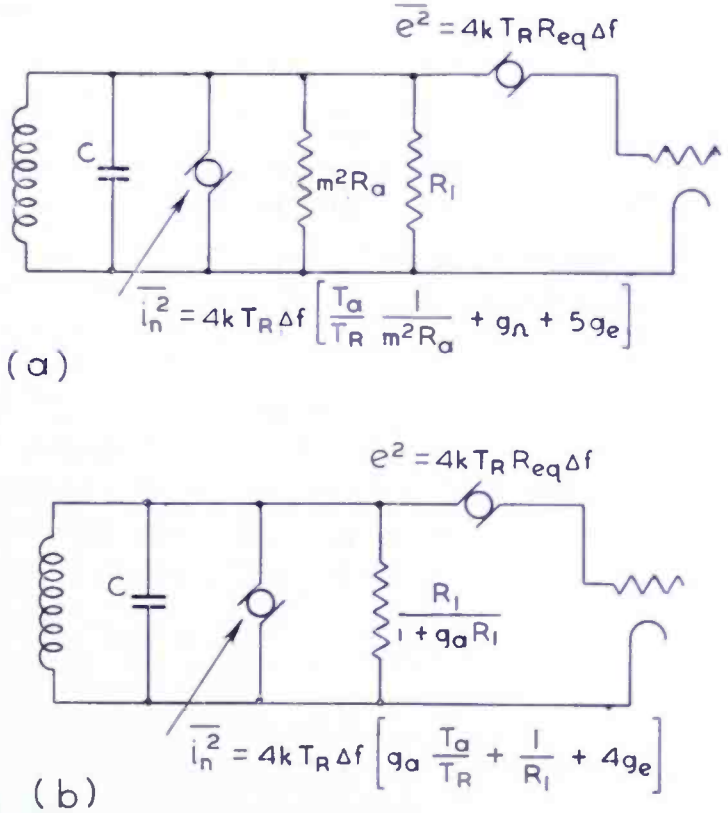


Fig. 3—(a) Equivalent circuit with antenna loading and noise shown as reflected values in secondary of antenna transformer; (b) Simplified equivalent obtained by introducing g_a for the reflected antenna conductance.

When the circuit is reduced to the simple one of 3b, the relations for the signal voltage and noise voltage are readily written down. As to the signal, in Figure 3b it could have been represented as a constant-current generator of value e_a/mR_a. In terms of the reflected antenna conductance, g_a, the signal voltage on the tube grid is then

$$e_{grid} = e_a \frac{\sqrt{g_a R_1}}{1 + g_a R_1} \sqrt{\frac{R_1}{R_a}} \qquad (1)$$

This expression reduces to the expected value for matched impedances if $g_a R_1$ is given the value unity.

The mean-squared noise voltage on the grid, from Figure 3b, is then

$$\overline{e_n^2} = 4k\,T_R\,\Delta f\left[R_{eq} + \left(g_a\frac{Ta}{T_R} + \frac{1}{R_1} + 4\,g_e\right)\frac{R_1^2}{(1+g_aR_1)^2}\right] \qquad (2)$$

The signal-to-noise ratio is given by dividing (1) by the square-root of (2) and is

$$\frac{e_{signal}}{\sqrt{\overline{e_n^2}}} = \frac{e_a}{\sqrt{4k\,T_R\,R_a\,\Delta f}} \times \left[\frac{1}{\frac{Ta}{T_R} + 2\frac{R_{eq}}{R_1} + (g_aR_1)\frac{R_{eq}}{R_1} + \frac{1}{(g_aR_1)}\left(1 + \frac{R_{eq}}{R_1} + 4g_eR_1\right)}\right]^{\frac{1}{2}} \qquad (3)$$

The denominator inside the radical of Equation (3) is the noise factor, N, whose importance in limiting the absolute sensitivity is discussed in a companion paper.[6]

For the sake of completeness, another relation, which is of some interest, is the bandwidth of the input circuit (which has been assumed equal to, or greater than, Δf). The total bandwidth, from points 3 *db* down on each side of the resonance curve, may be compared with that of a simple tuned circuit of capacitance C shunted by a resistance $\frac{R_1}{1+g_aR_1}$ and is

$$\text{Input band width} = \frac{1+g_aR_1}{2\pi\,C\,R_1}F \text{ (for 3 } db \text{ down)}$$

where F is equal to unity for an input circuit which is equivalent to a single-tuned circuit. For coupled circuits, or other band-pass arrangement, F may be somewhat greater.

Equations (1) to (4) are the fundamental relations which are applicable to the input circuit. Their interpretation will be made more clear by the subsequent discussion.

III—Conditions When Gain Is Maximum (Impedances Matched)

From Equation (1), maximizing with respect to g_aR_1, it is found that the maximum grid signal is given when $g_aR_1 = 1$ as, of course, is expected. This condition means simply that

$$m^2R_a = R_1$$

[6] D. O. North, loc. cit.

and the antenna impedance is exactly matched to the tube impedance by the input circuit. It is the usual condition of adjustment and warrants further consideration.

The signal-to-noise ratio (from Equation (3)), when $g_a R_1 = 1$, is

$$\left[\frac{e_{signal}}{\sqrt{\overline{e_n^2}}}\right]_{g_a R_1 = 1} = \frac{e_a}{\sqrt{\overline{e_t^2}}} \sqrt{\frac{1}{\left(\frac{Ta}{T_R} + 1\right) + 4\frac{R_{eq}}{R_1} + 4 g_e R_1}} \qquad (5)$$

where $\overline{e_t^2}$ is the open-circuit, mean-squared thermal noise voltage of the antenna at room temperature. This expression shows that the signal-to-noise ratio for a given signal, antenna noise and bandwidth depends only on the ratio of the tube equivalent noise resistance* to the total input resistance and on the fraction of this input resistance which is electronic in nature. The latter fraction, shown as $g_e R_1$ in the equation, must lie, of course, between 0 and 1. Thus, the signal-to-noise ratio must always lie between the two limits imposed by the above equation for $g_e = 0$ and for $g_e R_1 = 1$.

The minimum open-circuit antenna signal which will just equal the noise under the matched-impedance condition, is given by

$$\left[e_a\right]_{min} = \sqrt{\overline{e_t^2}} \sqrt{2 + 4\frac{R_{eq}}{R_1} + 4 g_e R_1} \qquad (6)$$

an expression which is obtained by setting $Ta = T_R$ as in receiver measurements with a dummy antenna. This minimum value may be used as a criterion for receiver performance.† Values of $\overline{e_t^2}$ computed for $Ta = 293°K$, $R_a = 75$ ohms, and various bandwidths are as follows:

Δf, kc	$\sqrt{\overline{e_t^2}}$, μV
10	0.11
200	0.49
2000	1.55
4000	2.2

* i.e., when noise of other tubes is small and R_{eq} is substantially due to first tube, only. Otherwise, R_{eq} includes noise of following stages also.

† As North has shown in Reference 6, $e_a/\sqrt{\overline{e_t^2}}$ cannot be changed by antenna design except by changing the directivity.

With the condition of matched impedances, which applies to Equation (6), the signal, when measured at the receiver input terminals, will be just half of the open-circuit signal, e_a.

A curve of the factor $(2 + 4\,R_{eq}/R_1 + 4\,g_eR_1)^{-1}$ plotted against R_1/R_{eq} is shown in Figure 4. The lower curve is for the case when $g_eR_1 = 0$ while the upper curve is for the case when $g_eR_1 = 1$. Results for all practical receivers, when adjusted for maximum gain, must lie between these two curves. Further discussion of these data will be given later in the paper.

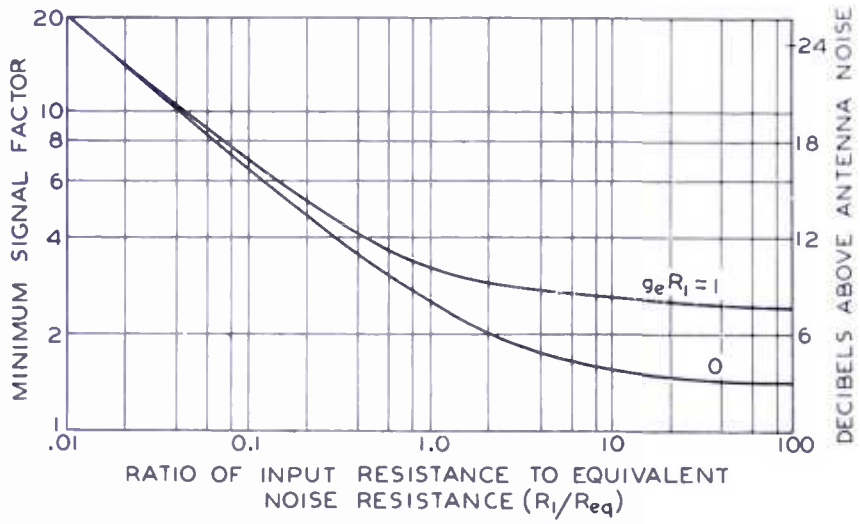


Fig. 4—The relative fluctuation noise of a receiver whose input circuit is adjusted for maximum gain.

IV—Conditions When Signal-to-Noise Ratio Is Maximum

Referring to Equation (3) of the basic relations, if this expression is maximized with respect to g_a, it will be found that best signal-to-noise ratio is given when

$$(g_aR_1)^2 = 1 + \frac{R_1}{R_{eq}}\,(1 + 4\,g_eR_1) \tag{7}$$

In terms of the step-up, m, of the input circuit the maximum signal-to-noise ratio is given when

$$m = \sqrt{\frac{R_1}{R_a}}\;\sqrt[4]{\frac{1}{1 + \frac{R_1}{R_{eq}}\,(1 + 4\,g_eR_1)}}$$

Substituting (7) in (3) gives the maximum signal-to-noise ratio as

$$\left[\frac{e_{signal}}{\sqrt{\overline{e_n^2}}}\right]_{max} = \frac{e_a}{\sqrt{\overline{e_t^2}}} \sqrt{\frac{1}{\frac{Ta}{T_R} + 2\frac{R_{eq}}{R_1} + 2\sqrt{\frac{R_{eq}}{R_1}\left(1 + \frac{R_{eq}}{R_1} + 4\,g_e R_1\right)}}} \tag{8}$$

It is again seen that the signal-to-noise ratio for a given signal, antenna noise, and bandwidth, depends only on the ratio R_{eq}/R_1 and on $g_e R_1$.

The minimum open-circuit antenna signal which will just equal the noise is then

$$\left[e_a\right]_{min} = \sqrt{\overline{e_t^2}} \sqrt{1 + 2\frac{R_{eq}}{R_1} + 2\sqrt{\frac{R_{eq}}{R_1}\left(1 + \frac{R_{eq}}{R_1} + 4\,g_e R_1\right)}} \tag{9}$$

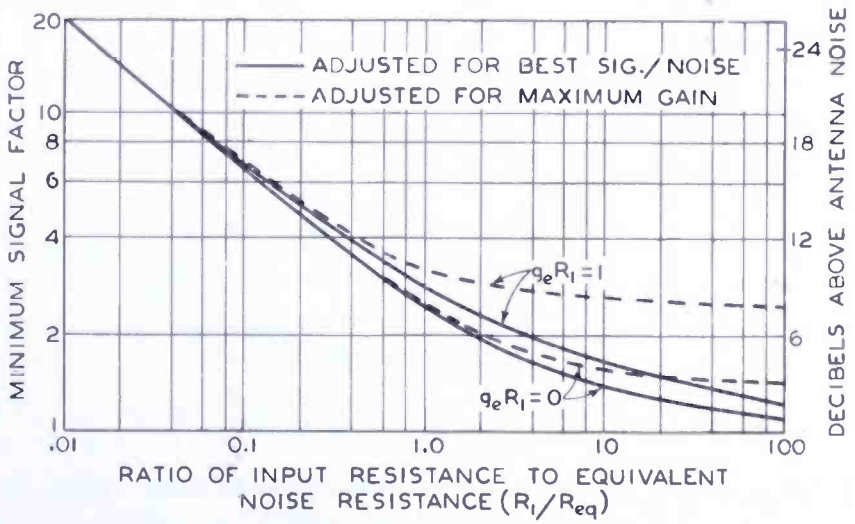


Fig. 5—The relative fluctuation noise of a receiver whose input circuit is adjusted for best signal-to-noise ratio as compared with one adjusted for maximum gain.

where Ta has been set equal to T_R for the lowest possible antenna noise (as with a dummy antenna). A curve of the factor included under the right hand radical plotted against R_1/R_{eq} is shown in Figure 5 for each of the two limiting cases $g_e R_1 = 0$ and $g_e R_1 = 1$. In order to facilitate comparison, the curves of Figure 4 are drawn with dotted lines.

It is, of course, necessary to sacrifice signal-voltage gain in order to achieve the improved signal-to-noise ratio. The ratio of the signal-voltage, antenna-to-grid gain (Equation (1)) with best signal-to-noise ratio to that with optimum-gain coupling is

$$\frac{\text{Actual Gain}}{\text{Max. Gain}} = \frac{2\sqrt{g_a R_1}}{1 + g_a R_1} \tag{10}$$

where $g_a R_1$ is the value given by Equation (7) for best signal-to-noise. Equation (10) represents, therefore, the gain-reduction factor.

Curves showing the reduction in gain as well as the improvement in signal-to-noise ratio* which are made possible with this coupling in comparison with optimum-gain coupling are shown in Figure 6. It is seen that over the range of values usually encountered ($R_1/R_{eq} < 10$), the signal-to-noise ratio improvement is appreciable though not startling even for $g_e R_1 = 1$. The reduction in gain is very nearly equal to the improvement in signal-to-noise ratio in this range. It should be noted that the bandwidth of the overall input circuit is *increased* by a change in coupling from that for best gain in the direction of that for best signal-to-noise ratio. The amount of the increase is readily found from Equation (4) using the value of $g_a R_1$ found from Equation (7).

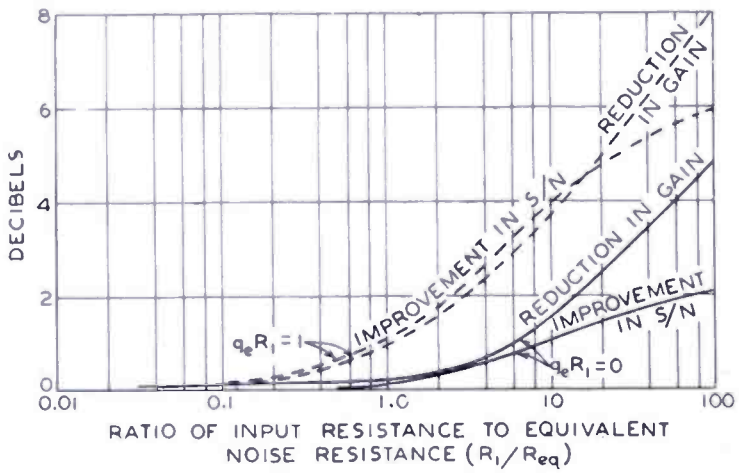


Fig. 6—Comparison of coupling for best signal-to-noise ratio, S/N, with coupling for best gain.

V—Discussion of Results of Sections III and IV

Perhaps the most interesting feature of the analysis has been that, no matter whether impedances are matched or whether best signal-to-noise ratio is desired, the signal-to-noise ratio depends mainly on the ratio of equivalent noise resistance, R_{eq}, to the total effective shunt resistance, R_1, of the input to the tube. When R_{eq} is chiefly due to the first tube (as it is when the gain of the first tube is reasonably high) and R_1 is also chiefly a result of the first tube (i.e., when circuit losses are negligible), a figure of merit for the first tube is R_1/R_{eq}. Unfortunately, this ratio varies with frequency. Its use, however, together with the curves of Figure 5, lead to a clear picture of receiver performance. It is seen, for example, that in an ultra-high-frequency

* i.e., the reduction in minimum signal which it is possible to receive.

amplifier or converter tube, an improvement in input resistance is just as valuable as a reduction in equivalent-noise resistance.

Referring to Figure 4, for the impedance-matched condition, the data show that little improvement is to be obtained in signal-to-noise by an increase in R_1/R_{eq} much greater than unity. Thus, if impedance matching is maintained, and a tube is found to have a minimum signal factor corresponding to $R_1/R_{eq} = 1$ it will not be possible to improve the performance much more, unless perhaps $g_e R_1$ can be reduced by a reduction in electronic loading so as to change from the upper curve towards the lower. However, if a mis-match is permitted, as shown in Figure 5, a definite improvement in signal-to-noise can be expected as R_1/R_{eq} is increased even up to 100.

At all times, the minimum thermal noise voltage of the antenna ($Ta = T_R$) sets a definite lower limit to receiver sensitivity. As an example, consider the minimum usable signal of a television receiver. A receiver with a 4-megacycle bandwidth designed for the reception of amplitude-modulated signals must have a signal-to-noise ratio of around 30 *db* for a satisfactory picture. Since the antenna thermal noise is 2.2 microvolts (for a 75-ohm antenna), a satisfactory picture requires a signal of 70 microvolts as a minimum and no amount of improvement in the receiver can possibly decrease this value.

One of the conclusions to be drawn from Figures 4 and 5 is that the presence of induced grid noise does not affect the signal-to-noise performance by a large amount. Under most practical high-frequency conditions, R_1/R_{eq} is well under 10 and if a tube and circuit whose major loading is electronic is compared with one whose loading is entirely resistive a maximum difference of only 1.4 db will be found when the coupling is adjusted for best signal-to-noise ratio.

The curves of Figure 6 indicate that adjustment for maximum gain (impedance-matching) is reasonable only as long as $R_1/R_{eq} < 1$. If the latter ratio can be improved markedly, by an improved circuit, tube, or both, so as to exceed unity by a significant amount it becomes profitable to mis-match somewhat to take advantage of the improved signal-to-noise ratio which is then attained.

There is still one other aspect of the signal-to-noise problem which is of importance in the normal broadcast band, for example. In a receiver design which is suitable for a wide variety of antennas, it is customary to couple the antenna very loosely to the input circuit. In this way variations of antenna impedance will have a minimum of effect on receiver line-up and performance. The condition of loose coupling corresponds to a very small reflected conductance, i.e., $g_a R_1 << 1$. This condition, when included in Equation (3), shows the signal-to-noise ratio to be

$$\frac{e_{signal}}{\sqrt{\overline{e_n^2}}} = \frac{e_a}{\sqrt{\overline{e_t^2}}} \sqrt{\frac{g_a R_1}{\left(1 + \frac{R_{eq}}{R_1} + 4\, g_e R_1\right)}} \qquad (11)$$

If $g_a R_1$ is held fixed, the ratio R_{eq}/R_1 is again seen to be of primary importance. Greatest improvement in signal-to-noise, however, will be accomplished by increasing the antenna coupling so as to increase g_a.

VI—Signal-to-Noise Ratio When Bandwidth Is of Paramount Importance

This section is concerned with those instances where the unavoidable tube or circuit capacitance is so great that the input circuit bandwidth is too narrow when the coupling is adjusted for best signal-to-noise (the bandwidth, in this case, will be even narrower for the impedance-matched condition). It will now be necessary either to adjust the reflected antenna conductance for the correct bandwidth, or to decrease the shunt resistance R_1 by adding a loading resistor. From Equation (4), it is seen that the relation between R_1, the bandwidth, the reflected conductance, and the capacitance is fixed as

$$g_a R_1 = \frac{2\pi \triangle f'\, C\, R_1}{F} - 1$$

$$= \frac{\triangle \omega C\, R_1}{F} - 1$$

where $\triangle\omega$ is introduced for $2\pi\triangle f'$. The bandwidth $\triangle f'$ is the *r-f circuit* bandwidth and is to be distinguished from $\triangle f$ which is the total effective bandwidth for noise purposes. When this expression for $g_a R_1$ is substituted in Equation (3) it is found that the *signal-to-noise* ratio is

$$\frac{e_{signal}}{\sqrt{\overline{e_n^2}}} = \frac{e_a}{\sqrt{\overline{e_t^2}}} \sqrt{\frac{1}{\frac{T_a}{T_R} + \frac{\triangle\omega C\, R_{eq}}{F} + \frac{1 + (\triangle\omega C\, R_{eq})/F + 4\, g_e R_1}{(\triangle\omega C\, R_1)/F - 1}}} \qquad (12)$$

It is now evident that, in this special case and once the r-f bandwidth is assigned, a low value of $R_{eq}\, C$ is the desideratum for best signal-to-noise. When R_{eq} and C are mainly contributed by the first tube, a possible figure of merit for the first tube is then $\frac{1}{2\pi\, C\, R_{eq}}$ which is expressible in cycles per second. It is also evident that, when R_1

approaches infinity, the best possible upper limit to the signal-to-noise ratio is obtained since the right hand term in the denominator then vanishes. The input circuit bandwidth is then obtained by the antenna loading only. Thus, the addition of a loading resistor to obtain bandwidth always results in a needless reduction of signal-to-noise ratio.

A point of some interest is that a pair of coupled circuits, with coupling adjusted for flat-top response and with R_1 approaching infinity gives a value of $F = \sqrt{2}$. With this coupled-circuit input arrangement, the minimum signal which is just equal to the noise is then

$$e_{min} = \sqrt{e_t^2}\sqrt{1 + 0.7(2\pi\Delta f')C R_{eq}}$$

where $\Delta f'$ is still defined by points 3 db down at each side of the resonance curve.

VII.—Effect of Output Circuit and Second Tube on Total Equivalent Noise Resistance, R_{eq}

A. Maximum gain: bandwidth not important—In some receivers, particularly those used at ultra-high frequencies with a low-gain amplifier or converter, it becomes necessary to consider the thermal noise and tube noise subsequent to the input stage. The calculation is most convenient if carried out in terms of an equivalent noise resistance due to the second circuit and tube, but referred to the grid of the *first* tube. The value so calculated must be added to the equivalent noise resistance of this first tube itself to obtain the total R_{eq} which has been assumed in the preceding analyses. The results are applicable whether the first tube be an amplifier or a converter. Induced grid noise will be neglected in this section.

In receiver arrangements where the bandwidth of the circuit coupling the first and second tubes is, a priori, wider than the overall bandwidth of the receiver, the coupling circuit is designed for maximum gain. With a single tuned circuit of resonant impedance R_2 and with a second tube of equivalent noise resistance R_{eq2} the mean-squared noise at the grid of the second tube is proportional to $(R_2 + R_{eq_2})$. This may be referred to the grid of the first tube by dividing by the square of the grid-to-grid gain. The total effective equivalent noise resistance at the grid of the first tube is then

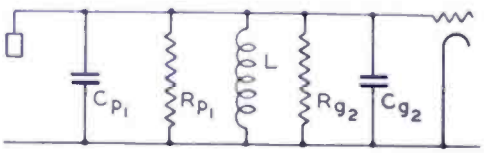
$$R_{eq} = R_{eq_1} + \frac{R_2 + R_{eq_2}}{(\text{GAIN})^2} \qquad (13)$$

where R_{eq_1} is the equivalent noise resistance of the first tube. When a pair of coupled circuits is used, with coupling adjusted for maximum gain, it is necessary to consider both primary and secondary im-

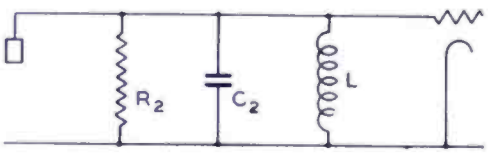
pedances. If we call the first of these R_{p_1} and the second R_{g_2}, the gain of the first tube will be proportional to $\frac{1}{2}\sqrt{R_{p_1} R_{g_2}}$. The mean-squared noise voltage at the grid of the second tube will be proportional to $R_{g_2}/2 + R_{eq_2}$ since the primary load with optimum coupling halves the value of R_{g_2} for noise-calculating purposes. Thus, the total effective equivalent noise resistance referred to the grid of the first tube is, in this case,

$$R_{eq} = R_{eq_1} + \frac{\frac{R_{g_2}}{2} + R_{eq_2}}{(\text{GAIN})^2} \tag{14}$$

In both Equations (13) and (14) the gain has been assumed to be the grid-to-grid gain. It will be found that for given first and second



(a) SINGLE-TUNED CIRCUIT



(b) EQUIVALENT CIRCUIT

$$C_2 = C_{p_1} + C_{g_2}$$

$$R_2 = \frac{R_{p_1} R_{g_2}}{R_{p_1} + R_{g_2}}$$

Fig. 7—Circuit and its equivalent for coupling first tube to second tube.

tubes, there is little to choose between the single-circuit and the coupled-circuit cases represented by Equations (13) and (14) as regards ultimate signal-to-noise ratio. This is in marked distinction to the wideband analysis to follow.

B. Wideband operation: bandwidth a major consideration—It now becomes necessary to consider the receiver in which the bandwidth of the coupling circuit between first and second tubes is an important, if not paramount, consideration. When the first tube is a converter or mixer, for instance, the bandwidth of the coupling circuit is almost always of primary concern since it is the first i-f circuit. When the coupling circuit is a single-tuned circuit, as in Figure 7, the effective bandwidth between points 1 db down on the curve is

$$\text{Bandwidth } \triangle f' = \frac{1}{4\pi C_2 R_2} \quad \text{(for 1 db down)} \tag{15}$$

When the value of R_2 from (15) is substituted in (13) and the bandwidth is $\triangle f'$, the total effective noise resistance becomes*

$$R_{eq} = R_{eq_1} + \frac{1 + 4\pi C_2 \triangle f' R_{eq_2}}{4\pi C_2 \triangle f' (\text{GAIN})^2} \tag{16a}$$

However, the gain is also expressible as $g_{m_1} R_2$ where g_{m_1} is the transconductance (or conversion transconductance) of the first tube. Thus (16a) can also be written (using again the value of R_2 from (15))

$$R_{eq} = R_{eq_1} + \frac{4\pi C_2 \triangle f'}{(g_{m_1})^2} (1 + 4\pi C_2 \triangle f' R_{eq_2}) \tag{16b}$$

It is seen clearly that, for the second tube as well as for the first, and when the bandwidth is a major consideration, the performance is

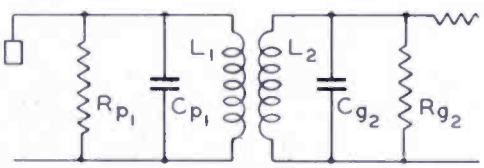


Fig. 8—Double-tuned circuit for coupling first tube to second tube.

specified by the product of a capacitance and the equivalent noise resistance, or $C_2 R_{eq_2}$.

For wideband applications it is highly advantageous to separate the two capacitances in an interstage coupler as is done by use of the pair of coupled circuits of Figure 8 or its equivalent T or π network. The fluctuation noise analysis in this instance is complicated for the first time by the possibility of a triangular instead of a rectangular noise spectrum. In a pair of coupled circuits (see Figure 8) adjusted for flat-top transfer impedance (i.e., second derivative of gain vs. frequency deviation set equal to zero), the impedance looking in at one pair of terminals may be decidedly double-humped if the circuits are of unequal Q. Thus the mean-squared thermal noise across such a pair of terminals must be obtained by integrating the resistive component of impedance over the pass band. Because the complete discussion is unimportant at this point, the derivation will be given in Appendix A and only the results discussed here. It is found that, by having a low-Q secondary circuit and a high-Q primary circuit, a

* It is assumed that the total thermal noise is $4kT R_2 \triangle f$. Actually if $\triangle f = \triangle f'$, due to dropping off of the tuned impedance near the edge of the pass band, this would be about 8 per cent too high, a negligible correction.

substantial reduction in thermal noise is made possible with no sacrifice in gain or circuit bandwidth.

The effective bandwidth of a pair of coupled circuits (Figure 8) taken between points 1 db down on each side and adjusted for flat response is

$$\text{Bandwidth } \triangle f' = \frac{1}{4\pi}\left(\frac{1}{C_{p_1}R_{p_1}} + \frac{1}{C_{g_2}R_{g_2}}\right)$$

$$= \frac{\omega_o}{4\pi}\left(\frac{1}{Q_1} + \frac{1}{Q_2}\right) \tag{17}$$

where ω_o is the angular frequency at the center of the band and Q_1 and Q_2 are primary and secondary Q's, respectively. As shown in the appendix, the equivalent noise resistance at the grid of the first tube is then

$$R_{eq} = R_{eq_1} + \frac{4\pi\triangle f' C_{p_1}}{g_{m_1}^2} \times$$

$$\left\{\frac{0.96\left(1+\frac{Q_1}{Q_2}\right)\left(1+0.158\frac{Q_1}{Q_2}\right) + \left(1+\frac{Q_1}{Q_2}\right)^2 2\pi\triangle f' C_{g_2}R_{eq_2}}{1+\left(\frac{Q_1}{Q_2}\right)^2}\right\} \tag{18}$$

where g_{m_1} is the transconductance or conversion transconductance of the first tube. It is easily shown that (18) is lowest when $Q_1 >> Q_2$ (see appendix) and this choice should always be made whenever possible. The part of the noise due to thermal agitation in the coupling network is reduced considerably by the choice. With this condition (i.e., $Q_1 >> Q_2$)

$$R_{eq} = R_{eq_1} + \frac{4\pi\triangle f' C_{p_1}}{g_{m_1}^2}(0.15 + 2\pi\triangle f' C_{g_2}R_{eq_2}) \tag{19}$$

Again, as in the previous cases, it is evident that the noise of the second tube is determined by the product $C_{g_2}R_{eq_2}$ which should be as small as possible. It is of considerable importance that the primary capacitance, C_{p_1}, (which is determined largely by the output capacitance of the first tube) should be as small as possible. If the condition $Q_1 >> Q_2$ is fulfilled, it will seldom be found that the thermal noise of the interstage circuit is appreciable compared with the noise produced by the

second tube. Thus, for most practical purposes, (19) may be replaced by

$$R_{eq} = R_{eq_1} + \frac{R_{eq_2}}{(\text{GAIN})^2} \tag{19a}$$

provided coupled circuits are used as indicated. Equation (19a) is obtained from (19) by neglecting the factor, 0.15, substituting $|Z_t|$ for its equivalent, and finally letting $g_{m_1} |Z_t| = \text{GAIN}$.

VIII—Practical Application to U-H-F Receivers

Examples of the application of the foregoing analysis are readily made to ultra-high-frequency receivers for operation above 300 megacycles. In the following applications, it must be emphasized that feedback effects have been neglected and that, to some extent, this neglect may lead to discrepancies between measured results and those calculated herein. Furthermore, since the purpose of this discussion is primarily illustrative, the tube data to be used will be only approximate. Exact tube data at frequencies above 300 megacycles or so are not available in any event.

The triode mixer, used in the converter stage of the receiver, will form an important part of the discussion and it may be well to offer a word of explanation. A triode in the converter stage, followed by a low-noise i-f system, seems to offer the greatest promise for a good overall signal-to-noise ratio with tubes now commercially available. In the ordinary triode mixer, signal-frequency feedback from the i-f circuit (which looks like a capacitance at the signal frequency) through the grid-plate capacitance may be very important in reducing the input resistance. However, in receivers designed to cover a very limited tuning range, a simple expedient may be used to greatly reduce feedback. If the i-f output circuit is designed so as to present a very low impedance to signal frequency, the signal-frequency voltage on the plate may be reduced so that feedback becomes negligible. As a simple example, consider the primary winding of the i-f transformer. At signal frequency, this winding is above its natural resonance, and behaves as a capacitance. If we add, in series with it, a small inductance (a loop of wire or a length of lead) and adjust the latter to series resonance with the effective distributed capacitance of the i-f coil, we obtain a series resonant output circuit at signal frequency without affecting, in any way, the i-f circuit. As a minor point in connection with the triode mixer, the input capacitance and output capacitance are each increased by the amount of the grid-plate capacitance.

The equivalent noise resistances which will be used are taken from formulas given in another paper.[12] The conversion transconductance is assumed to be adjusted to optimum value. The electronic input conductance figures are based on available data with fixed voltages applied and, in the case of the converter stage, are averaged over the oscillator cycle. The capacitance figures which will be used represent reasonable approximations.

The data will be presented in the form of a table. In the table, the electronic input conductance is expressed in micromhos per (megacycle)^2. Thus, to obtain the input conductance at any one frequency, the constant given must be multiplied by the square of the frequency in megacycles. In somewhat similar fashion, the non-electronic input loss is expressed in micromhos per megacycle. The figure given represents only a rough guess, but since this part of the loss is small in comparison with the electronic loss, its exact value is unimportant. Circuit losses, if any, will be assumed to be included in the non-electronic conductance figure. The Type 955 is assumed to be operated at a plate voltage of 180, the 954 at a screen voltage of 100. The column giving the peak oscillator voltage required is of importance in estimating the power required from the local oscillator. The grid bias of the mixer tube is assumed equal to the peak oscillator voltage in all cases.

TABLE I

First Tube and Method of Operation	Equiv. Noise Resist. R_{eq_1}, Ohms	Trans-conduct. g_m or g_c, μmhos	Elec-tronic g_e Per (Mc),^{2} μmhos	Non-Elec-tronic g_Ω Per Mc., μmhos	Approx. Peak Oscil-lator Volts
955 Mixer—At Oscillator Fundamental	5,200	700	0.0030	0.2	5
955 Mixer—At Oscillator 2nd Harmonic	12,000	350	0.0015	0.2	12
955 Mixer—At Oscillator 3rd Harmonic	15,000	230	0.0010	0.2	34
954 Mixer—At Oscillator Fundamental	30,000	700	0.0030	0.2	5
954 Amplifier	6,000	1,400	0.0050	0.2	—

Each example to be computed will be considered at three signal frequencies, 300, 500, and 1,000 megacycles. The intermediate fre-

[12] E. W. Herold, "The Operation of Frequency Converters and Mixers for Superheterodyne Reception", submitted for publication to *Proc. I.R.E.*

quency (i-f) is chosen as 30 megacycles and the overall bandwidth of the radio portion of the receiver is 7 megacycles, in order to minimize tuning difficulties and to permit wideband operation, as for television. However, when the relative sensitivities (signal equal to noise) are computed, they will be referred to an arbitrary bandwidth of 10 kilocycles. The actual sensitivity for other bandwidths following the second detector, is readily computed by multiplying by the square-root of the bandwidth ratio, in the usual way, with a negligible error.

A. The i-f tube and circuit—It is observed from Equation (19) that the best tube to select for the i-f amplifier is one with the lowest product of C_{g_2} and R_{eq_2}. It is easily determined that the type 6AC7/1852 is the best of the commercially available tubes in this respect, since it has a value for $R_{eq_2} = 780$ ohms and an effective C_{g_2} (including circuit and leads) which can be made as low as 20 micromicrofarads. This tube will, therefore, be chosen. In each example, the tube working into the i-f system will be an "Acorn" type and a total effective output capacitance of 5 micromicrofarads (C_{p_1}) will be assumed. This value is high enough to include circuit and leads. There is no significant difference between the 954 and 955 types in output capacitance because, with the latter, it is necessary to add the grid-to-plate capacitance as well. Because the i-f system is common to each example, we may compute its contribution to the noise immediately. The output plate impedance of a 955 mixer or converter is easily shown to be in excess of 20,000 ohms, while the 6AC7/1852 has an input impedance of around 5,000 ohms. Thus, if we use a pair of coupled circuits as the first i-f transformer, $Q_1 > 19$ and the secondary loaded by the 6AC7 only will have $Q_2' = 19$.† These Q's are so much higher than required to give our desired 7-megacycle circuit bandwidth that a damping resistance must be added.

Following the low-noise design exemplified by the discussion following Equation (18) we will add all the damping to the secondary. For simplicity, let us assume $Q_1 = 19$ in every case.‡ Then Equation (17) gives

$$\frac{1}{Q_1} + \frac{1}{Q_2} = 2\frac{\Delta\omega}{\omega} = 2\frac{7}{30} = 0.47$$

If $Q_1 = 19$ we see that Q_2 must be 2.4 and that we must add a shunt resistance of 730 ohms across the input to the 6AC7. Under these

† Obviously the i-f transformer losses can be neglected.

‡ This implies adding a 20,000-ohm shunt resistance across the primary of the i-f transformer when a pentode mixer is used.

circumstances, $Q_1/Q_2 = 7.9$ and the transfer impedance is 1,600 ohms, approximately. From Equation (18) we find the effective equivalent noise resistance at the grid of the converter stage is

$$R_{eq} = R_{eq_1} + \frac{523 \times 10^{-6}}{g_c^2}$$

Using the values from Table I for R_{eq_1} and g_c, we find R_{eq} for each of the converter stage possibilities.

B. Converter as first stage—It is first necessary to determine whether the input circuit bandwidth is sufficiently wide. The narrowest bandwidth of any example considered, will be obtained with the lowest frequency, 300 megacycles, and the highest value of input resistance, obtained with the 955 mixer operating at the third harmonic of the applied oscillator frequency. Using Equation (4) and letting $g_a R_1 = 1$ (the lowest value which should be considered in a low-noise design), the circuit bandwidth is at least

$$\Delta f' = \frac{1}{\pi C R_1} = \frac{1}{\pi \times 5 \times 10^{-12} \times \left(\frac{10^{+6}}{0.001 \times (300)^2 + 0.2 \times 300} \right)} = 9.6 \text{ megacycles}$$

Thus, it is found that the r-f bandwidth need not be increased in any of the examples and the receiver may be designed for best signal-to-noise ratio. Section IV of this paper is then applicable.

The calculations for each example are straightforward and only one will be carried out in detail. The others will be tabulated later in this section of the paper. We see from Equation (9) that we need only R_{eq}, R_1, and g_e. Let us choose for the detailed example the 500-megacycle receiver with a 955 mixer operating at oscillator fundamental. Then

$$R_{eq} = R_{eq_1} + \frac{523 \times 10^{-6}}{g_c^2} = 5200 + 1070 = 6270 \text{ ohms}$$

Also

$$R_1 = \frac{1}{g_e + g_\Omega} = \frac{10^6}{0.003 \times (500)^2 + 0.2 \times 500} = 1180 \text{ ohms}$$

Finally

$$g_e R_1 = 0.003 \times (500)^2 \times 10^{-6} \times 1180 = 0.88$$

Using Equation (9), we find $e_{a_{min}} = \sqrt{\overline{e_t^2}} \times 5.1$

For $R_a = 75$ and $\Delta f = 10$ kc, $\sqrt{\overline{e_t^2}} = 0.11$ microvolts so that the minimum open-circuit antenna signal to equal the noise is $e_{a_{min}} = 0.56$ microvolts. If a signal generator is used which has been calibrated

for the matched impedance condition, its calibration should indicate just half of the above value.

It is interesting to compare the sensitivity just computed with that which would obtain if the impedance-matched condition had been used. Using Equation (6), we find

$$\text{impedance-matched } e_{a_{min}} = 0.57 \text{ microvolts}$$

This represents a negligible reduction in signal-to-noise ratio. The result should have been expected from inspection of the curves of Figure 6 which are nearly coincident at the value of R_1/R_{eq} appropriate to this example.

C. Amplifier as first stage—Let us assume a 954 amplifier followed by a converter stage with the lowest noise equivalent (the 955 operated at oscillator fundamental). The calculations will be made neglecting the induced grid noise in the converter and assuming an idealized condition with no feedback in the r-f stage. In this example, the interstage circuit impedance will be given by half of the geometric mean of the 954 output impedance and the converter input impedance. Recent data on the 954 indicate an output conductance due to lead losses etc. roughly equal to

$$954 \text{ output } g = 0.0006 \times (\text{megacycle})^2 \text{ (in micromhos)}$$

For this relation, the 300-, 500-, and 1,000-megacycle interstage transfer impedances are then 3,700, 1,400, and 360 ohms, respectively. With $g_m = 1{,}400$ micromhos, the grid-to-grid gains are then 5.2, 2.0, and 0.5, respectively. The value of R_{eq_2} will be taken as 6,270 ohms (to include i-f noise) while R_{eq_1} (Table I) is 6,000 ohms. We find from Equation (14)

$$R_{eq} = R_{eq_1} + \frac{\frac{R_{g_2}}{2} + R_{eq_2}}{(\text{GAIN})^2} = \begin{cases} 6{,}300 \text{ ohms for 300 megacycles} \\ 6{,}700 \text{ ohms for 500 megacycles} \\ 32{,}000 \text{ ohms for 1,000 megacycles} \end{cases}$$

Using the conductance data from Table I, we find

$$R_1 = \begin{cases} 2{,}000 \text{ ohms for 300 megacycles} \\ 740 \text{ ohms for 500 megacycles} \\ 190 \text{ ohms for 1,000 megacycles} \end{cases}$$

The minimum open-circuit antenna signal is then readily computed. For Equation (9) (the condition for best signal-to-noise ratio), the relative open-circuit antenna signals ($R_a = 75$, $\Delta f = 10$ kilocycles) are

$$e_{a_{min}} = \begin{cases} 0.46 \text{ microvolts at 300 megacycles} \\ 0.70 \text{ microvolts at 500 megacycles} \\ 2.9 \text{ microvolts at 1,000 megacycles} \end{cases}$$

It is evident from these figures that a 954 r-f amplifier stage is inferior to the 955 converter stage, from a signal-to-noise point of view, even at 300 megacycles. This was already in evidence in Table I since both R_{eq_1} and g_e were lower for the 955 mixer than for the 954 amplifier.

TABLE II

First Tube	Minimum Open-Circuit Antenna Signals		
	300 Mc	500 Mc	1,000 Mc
955 Mixer—At oscillator fundamental	0.39μV	0.56μV	1.02μV
955 Mixer—At oscillator 2nd harmonic	0.45	0.66	1.20
955 Mixer—At oscillator 3rd harmonic	0.47	0.69	1.23
954 Mixer—At oscillator fundamental	0.75	1.15	2.22
954 Amplifier—Followed by 955	0.46	0.70	2.9

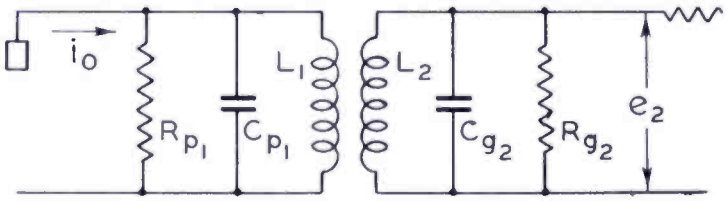


Fig. 9—Double-tuned circuit with constant-current drive.

D. Tabulated results for all examples—It is assumed that $R_a = 75$ ohms, $\triangle f = 10$ kilocycles and the design for best signal-to-noise is used (Equation (9)).

One of the more interesting aspects of the table is the fact that the 955 used as a mixer at 2nd or 3rd harmonic of the applied oscillator is less than 2 *db* poorer in signal-to-noise ratio than it is with the oscillator fundamental. It is, of course, necessary to apply a considerably larger oscillator voltage when harmonic conversion is desired.

Appendix A

Calculation of Noise in Wide Band Coupled Circuit Fed by Constant-Current Source

Because the behavior of coupled circuits is reasonably well known, no attempt to derive the basic relations for transfer impedance, bandwidth, etc. will be made here. The relations which will be used represent the usual approximations which hold exactly only for high Q's (i.e., bandwidth small compared to center frequency), but which are

qualitatively correct even for a low-Q circuit. Referring to the figure, if $Q_1 = \omega C_{p_1} R_{p_1}$, the primary circuit Q, and $Q_2 = \omega C_{g_2} R_{g_2}$, the secondary circuit Q, then the transfer impedance is

$$\left|\frac{e_2}{i_o}\right| = |Z_t| = \frac{\sqrt{2}}{\omega_o\sqrt{C_{p_1}C_{g_2}}} \frac{\sqrt{\frac{1}{Q_1^2}+\frac{1}{Q_2^2}}}{\left(\frac{1}{Q_1}+\frac{1}{Q_2}\right)^2} \sqrt{\frac{1}{1+\frac{2^6}{\left(\frac{1}{Q_1}+\frac{1}{Q_2}\right)^4}\left(\frac{\delta\omega}{\omega_o}\right)^4}} \tag{20}$$

where $\delta\omega$ is the deviation from the center angular frequency, ω_o. This equation assumes flat-top response, i.e., zero second derivative of $|Z_t|$ at center frequency. The total bandwidth, $\Delta f'$, for one *db* down on each side is then twice the deviation frequency for a one *db* decrease in $|Z_t|$ and is

$$2\pi\Delta f' = \Delta\omega = \frac{\omega_o}{2}\left(\frac{1}{Q_1}+\frac{1}{Q_2}\right) \tag{21}$$

Using (21) in (20) we get

$$|Z_t| = \frac{0.707}{\Delta\omega\sqrt{C_{p_1}C_{g_2}}} \frac{\sqrt{1+\left(\frac{Q_1}{Q_2}\right)^2}}{1+\frac{Q_1}{Q_2}} \sqrt{\frac{1}{1+4\left(\frac{\delta\omega}{\Delta\omega}\right)^4}} \tag{22}$$

The real part of the impedance looking into the C_{g_2}, R_{g_2} terminals is

$$Rl \text{ part of } Z_{\text{secondary}} = \frac{1}{\Delta\omega C_{g_2}\left(1+\frac{Q_1}{Q_2}\right)} \frac{1+2\frac{Q_1}{Q_2}\left(\frac{\delta\omega}{\Delta\omega}\right)^2}{1+4\left(\frac{\delta\omega}{\Delta\omega}\right)^4} \tag{23}$$

It is interesting to note that, if $Q_1 >> Q_2$, a curve of Equation (23)

is decidedly double-humped as shown qualitatively in Figure 10. Since the thermal noise is proportional to this variable, the noise spectrum is non-uniform and, in fact, the noise at center band is very small. This behavior results in a very low average thermal noise over the pass band when $\triangle f = \triangle f'$. It should be remembered that $\triangle f$, the effective pass band for noise purposes, is largely determined by the system following the final demodulation of an incoming signal whereas $\triangle f' = \frac{\triangle \omega}{2\pi}$ is the circuit bandwidth. Since the maximum noise will occur when $\triangle f = \triangle f'$, the calculations will be made on this basis. Since, as will be

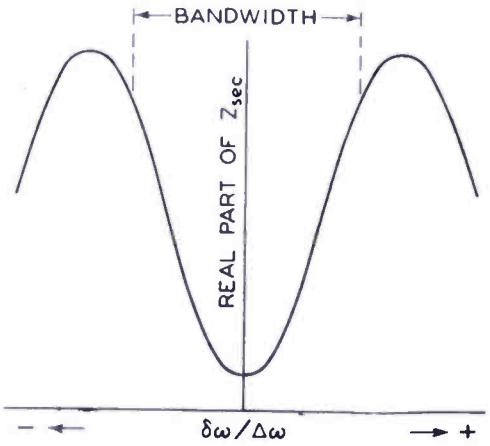


Fig. 10—The real part of the secondary impedance of a flat-top, double-tuned circuit for frequency deviations around mid-band and with primary Q much higher than secondary Q.

found, the thermal noise of the coupled circuits may be made very low, the correction which might have been applied when $\triangle f << \triangle f'$ would not have been of significance in any event.

We are now in a position to compute the thermal-noise voltage at the grid of the second tube, add to it the equivalent second-tube noise voltage, and, if we choose, refer the total back to the grid of the preceding tube by dividing by the square of the gain, $(g_{m_1} Z_t)^2$. The mean-squared noise of thermal agitation at the grid of the second tube is

$$\left(\overline{e_{t_2}^2}\right)\text{circuit} = 4kT_R \frac{1}{2\pi} \int_{\omega_0 - \frac{\triangle\omega}{2}}^{\omega_0 + \frac{\triangle\omega}{2}} Rl \text{ part of } Z_{sec}\, d\omega$$

where ω_0 is the mid-band angular frequency. This is equivalent to the thermal noise of a resistance at the second-tube grid of value

$$R_{t_2} = \frac{1}{\Delta\omega} \int_{\omega_o - \frac{\Delta\omega}{2}}^{\omega_o + \frac{\Delta\omega}{2}} Rl \text{ part of } Z_{\text{sec}}\, d\omega$$

$$= 2 \int_{o}^{\frac{1}{2}} Rl \text{ part of } Z_{\text{sec}}\, d\left(\frac{\delta\omega}{\Delta\omega}\right)$$

where $\delta\omega$ is the angular frequency deviation from the center-band point and $\Delta\omega$ is the angular bandwidth to points 1 *db* down. Using Equation (23), we obtain

$$R_{t_2} = \frac{2}{\Delta\omega C_{g_2}\left(1 + \frac{Q_1}{Q_2}\right)} \int_{o}^{\frac{1}{2}} \frac{1 + 2\frac{Q_1}{Q_2}x^2}{1 + 4x^4}\, dx \left(\text{where } x = \frac{\delta\omega}{\Delta\omega}\right)$$

$$= \frac{0.96}{\Delta\omega C_{g_2}} \frac{1 + 0.158\frac{Q_1}{Q_2}}{1 + \frac{Q_1}{Q_2}} \tag{24}$$

Equation (24) shows that R_{t_2} varies over a 6:1 range as Q_1/Q_2 is varied from one extreme ($Q_1 >> Q_2$) to the other ($Q_2 >> Q_1$). When $Q_1 >> Q_2$, the noise spectrum is non-uniform and, at mid-band frequency, the thermal noise is very much reduced over that at the edge of the band.* The transfer impedance (i.e., the gain) is, of course, constant over the band by the a priori relations.

The total effective noise resistance at the second-tube grid is ($R_{eq_2} + R_{t_2}$) where R_{eq_2} is the equivalent noise resistance of the second tube. The noise may be referred back to the first-tube grid by using (21), (22), and (24) to get

$$R_{eq} = R_{eq_1} + \frac{R_{eq_2} + R_{t_2}}{(g_m |Z_t|)^2}$$

* It should be emphasized that the thermal noise-reduction possibilities of coupled circuits cannot be utilized in the antenna-to-grid circuit in the same way.

$$= R_{eq_1} + \frac{2\Delta\omega C_{p_1}}{g_{m_1}^2} \times \left\{ \frac{0.96\left(1+\frac{Q_1}{Q_2}\right)\left(1+0.158\frac{Q_1}{Q_2}\right) + \left(1+\frac{Q_1}{Q_2}\right)^2 \Delta\omega C_{g_2} R_{eq_2}}{1+\left(\frac{Q_1}{Q_2}\right)^2} \right\} \quad (25)$$

The factor in the bracket of Equation (25) is determined by the choice of Q_1/Q_2 which is usually up to the designer. Curves of this factor are plotted in Figure 11. They show that Q_1/Q_2 should be high,

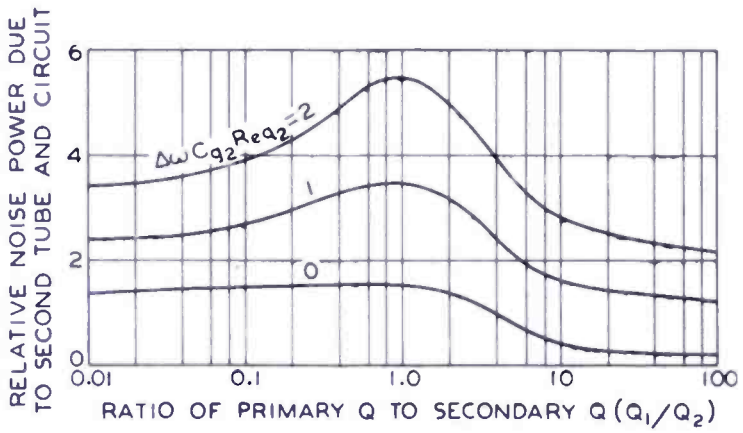


Fig. 11—The total fluctuation noise of a flat-top, double-tuned circuit of bandwidth $\Delta\omega$ whose secondary is connected to a tube of equivalent-noise-resistance, R_{eq}, and whose total secondary capacitance is C_{g2}.

i.e., the wide-band circuit damping should be concentrated in the secondary. The overall noise improvement to be obtained by such a choice is ordinarily considerably less than the theoretical maximum of 8 *db* because of the major role played by the first-tube equivalent noise resistance, R_{eq_1} not to mention the noise of the second tube as expressed in R_{eq_2}. However, since there appear to be no major disadvantages to the choice, it can be concluded that best design of the interstage coupling circuit will occur with the damping on the secondary only.

Appendix B

Signal-to-Noise Ratio for Mixer and Converter Tubes with Control Grid Not Adjacent to Cathode

In some mixer and converter tubes the input or signal grid is not adjacent to the cathode and the preceding formulas must be modified

before they are applicable. Furthermore, the electronic conductance of a tube of this kind is usually negative[12] so that some consideration must be given to this aspect. The analysis and measurements of induced grid noise in such tubes gives the noise in terms of an equivalent saturated diode current, I_d.[5] Thus the mean-squared grid noise current is

$$\overline{i_n^2} = 2\,e\,I_d\,\Delta f$$

and this noise current cannot be expressed in a simple manner in terms of g_e, the electronic input conductance, as when the control grid is adjacent to the cathode. However, at frequencies for which the transit angles are not too great, I_d increases with the square of the frequency just as does the magnitude of g_e.

The formulas developed in the body of the paper assumed the induced noise current to be $\overline{i_n^2} = (5 \times 4k\,T_R\,g_e\,\Delta f)$ so that, for the present case, we find that the quantity $4g_e\,R_1$ of the formulas becomes

$$4\,g_e R_1 = \left(\frac{i_n^2}{4k\,T_R\,\Delta f} - g_e\right) R_1 = (20\,I_d - g_e)\,R_1 \qquad (26)$$

For the tubes we are now considering, R_1 may be negative and it must be appreciated that stability requirements demand that the reflected antenna resistance be lower than R_1. An even more stringent requirement is that imposed by bandwidth. This requirement is given by Equation (4) and may be written

$$g_a \geqq \Delta\omega C - \frac{1}{R_1} \qquad (27)$$

where $\Delta\omega$ is $2\pi\Delta f'$ and a single-tuned circuit is assumed.

The value of g_a for optimum signal-to-noise ratio, irrespective of bandwidth, is given by Equation (7). This optimum coupling condition is rewritten (using Equation (26))

$$g_a^2 = \frac{1}{R_1^2} + \frac{1 - g_e R_1 + 20\,I_d\,R_1}{R_1\,R_{eq}}$$

$$= \frac{1}{R_1^2} + \frac{g_\Omega + 20\,I_d}{R_{eq}} \qquad (28)$$

If the value of g_a from Equation (28) is larger than that from (27) it is possible to obtain a signal-to-noise advantage by using (28). Otherwise, condition (27) must be used to assure proper bandwidth.

---

[5] C. J. Bakker, loc. cit.
[12] E. W. Herold, loc. cit.

Using (26) and the bandwidth condition (27), the minimum signal which is equal to the noise is found from Equation (3) to be

$$\left.e_a\right]_{\min} = \sqrt{4kT_R R_a \Delta f} \sqrt{\frac{T_a}{T_R} + \frac{g_\Omega + 20 I_d + (\Delta\omega C)^2 R_{eq}}{\Delta\omega C - \frac{1}{R_1}}} \quad (29)$$

If, on the other hand, the bandwidth is not a limiting consideration, we may use (26) and (28) in (3) to get the lowest possible noise and find

$$\left.e_a\right]_{\min} = \sqrt{4kT_R R_a \Delta f} \sqrt{\frac{T_a}{T_R} + 2\frac{R_{eq}}{R_1} + 2\sqrt{\frac{R_{eq}}{R_1}\left(\frac{R_{eq}}{R_1} + g_\Omega R_1 + 20 I_d R_1\right)}} \quad (30)$$

When R_1 is negative, as here contemplated, the sign of the inner square root must nevertheless be taken as positive if the result is to be valid.

List of Symbols

C = Total capacitance of input circuit to first tube.

C_{p_1} = Total capacitance of primary circuit of interstage transformer.

C_{g_2} = Total capacitance of secondary circuit of interstage transformer.

C_2 = Sum of C_{p_1} and C_{g_2}. Total shunt capacitance effective across a single-tuned circuit between first and second tubes.

e_a = Open-circuit signal voltage from antenna.

e_t^2 = Open-circuit mean-squared noise voltage from an antenna which can be considered at room temperature.

F = Factor by which double-tuned antenna transformer is better than single-tuned as regards bandwidth.

Δf = Overall receiver bandwidth for noise purposes (usually determined after final demodulation of signal).

$\Delta f'$ = Circuit bandwidth. This has a minimum value equal to Δf.

$\Delta\omega$ = Overall angular circuit bandwidth; equal to $2\pi\Delta f'$.

$\delta\omega$ = Angular-frequency deviation from center of a band-pass circuit.

ω_o = Center frequency of band-pass circuit.

g_a = Reflected antenna conductance in secondary of input circuit.

g_e = Electronic input conductance of first tube.

g_Ω = Circuit and "cold" tube loading conductance of input to first tube.

g_{m_1} = Transconductance or conversion transconductance of first tube.

k = Boltzmann's constant (1.37×10^{-23} joules per °K).

m = Effective turns ratio or step-up of antenna-to-grid circuit.

Q_1 = Ratio of shunt resistance to reactance of primary circuit of interstage transformer.

Q_2 = Ratio of shunt resistance to reactance of secondary circuit of interstage transformer.

R_a = Radiation resistance of antenna.

R_1 = Impedance of secondary side of antenna-to-grid circuit with tube connected but with antenna disconnected.

R_{p_1} = Shunt resistance across primary of interstage transformer.

R_{p_2} = Shunt resistance across secondary of interstage transformer.

R_2 = Shunt resistance across single-tuned circuit as interstage network.

R_{eq} = Total equivalent noise resistance at grid of first tube including noise due to first tube, interstage circuit, and second tube.

R_{eq_1} = Equivalent noise resistance of first tube at its grid.

R_{eq_2} = Equivalent noise resistance of second tube at its grid.

T_a = Temperature of a resistance equal to antenna radiation resistance and having same mean-squared noise as antenna.

T_R = Room, or ambient, temperature.

$|Z_t|$ = Magnitude of transfer impedance of interstage transformer.

THE ABSOLUTE SENSITIVITY OF RADIO RECEIVERS

By

D. O. North

Research Laboratories, RCA Manufacturing Company, Inc., Harrison, N. J.

Summary—The total random noise originating in a receiver has customarily been described in terms of the equivalent noise voltage at the receiver input terminals. A comparison of the signal-to-noise ratios of two receivers working out of identical antennas is thereby facilitated, but only so long as the coupling between antenna and receiver input is extremely loose.

This paper describes a method for rating and measuring the noise in complete receiving systems, antenna included. The proposed rating appears particularly applicable to ultra-high-frequency services and, more generally, to any service in which signal-to-noise ratio is made a prime consideration in receiver design and operation.

A portion of the study deals with the properties of receiving antennas, yielding as a by-product an alternative derivation of Nyquist's theorem concerning thermal fluctuations in passive networks.

A formula for absolute sensitivity is developed, which shows how the minimum usable signal-field strength is related to the operating wavelength, the antenna directivity, the local noise-field strength, the receiver bandwidth, and a number called the "noise factor", which is a basic measure of the internal noise sources of the receiver.

THE "Standards on Radio Receivers" adopted and published by the Institute of Radio Engineers in 1938, define the sensitivity of a radio receiver as "that characteristic which determines the minimum strength of signal input capable of causing a desired value of signal output".[1] So defined, the sensitivity is a measure of gain. The same report[2] takes cognizance of the fact that random fluctuations (noise) originating in the receiver set a limit to the useful sensitivity. Leaving a precise definition to those who write standards, we shall refer to this useful limit as the "absolute sensitivity", indicate in an elementary way its dependence upon certain properties of the receiver[3] and antenna, suggest methods for measuring it in the laboratory, and consider its modification by noise induced in the antenna.

[1] "Standards on Radio Receivers", 1938, definition 1R36.

[2] Page 42, section 13.

[3] A more thorough consideration of this aspect is the purpose of a companion paper by E. W. Herold, "An Analysis of the Signal-to-Noise Ratio of Ultra-High-Frequency Receivers", RCA Review, Jan. 1942.

The conventional method of describing receiver noise in terms of an equivalent-noise-side-band input, while possessing a limited utility, does not ever permit an immediate positive judgment of signal-to-noise ratio, nor even a comparison of two receivers, unless additional information is given concerning the antennas and the means for coupling. It will be seen, on the other hand, that a description of the absolute sensitivity of a receiver, antenna included, provides a direct basis for the judgment of a complete receiving system, and facilitates the inter-comparison of receiving equipments, no matter how diverse their characteristics.[4] A rating of this kind is particularly valuable in the ultra-high-frequency field, where antennas are designed as an integral part of the receiver, or, for that matter, in any service wherein maximum signal-to-noise ratio is an important factor in design and operation. It was during an attempt, two years ago, to find a sound basis for the comparison of television receivers, that the considerations reported below were crystallized.

Measurements in the Laboratory

Consider an ultra-high-frequency receiver constructed to work out of a specific antenna. This implies a specific radiation resistance, antenna reactance, and ohmic loss (such as might be present in a long down-lead). The important essentials of a laboratory simulation of the real antenna, therefore, appear to be

1—a signal generator providing a known radio-frequency voltage source, e, in series with

2—a resistor, R_a, whose resistance equals the radiation resistance of the prescribed real antenna;

3—tunable reactance to simulate the Q of the real antenna; perhaps

4—additional losses to simulate unavoidable losses in the real antenna, and possibly

5—a low-loss transformer to correct *minor* defects in the attempted similitude.

In many cases, the real antenna is prescribed to have a low Q, to permit reception over a wide range of frequencies. In this event, and provided the antenna Q is small in comparison with the Q of the receiver input circuit, item (3) is of minor importance. By the same token, emphasis upon "low-loss" in item (5) becomes less severe. Furthermore, item (5) will often be found incorporated in the receiver proper, in which case the responsibility for making it loss-free lies with the receiver designer.

[4] This same subject has recently received the capable attention of R. E. Burgess, "Noise in Receiving Aerial Systems", *Proc. Phys. Soc.*, Vol. 53, p. 293, May, (1941).

Consider next methods for measuring the noise output. Present I.R.E. standards prescribe measurement of noise after detection.[2] This method not only requires a known modulation for the signal generator, but also makes the noise measurement a function of the frequency response of the circuits which follow the detector. On the premise that post-detection filters are, or should be made sufficiently flat and broad to pass uniformly all frequencies offered by the pre-detection part of the receiver, it is here proposed to measure noise prior to detection. The natural point at which to insert a meter is just ahead of the detector. But occasionally latitude is permissible, and the demands made upon the meter and its position may thus be summarized.

1—It must be preceded by all the filters which materially determine the r-f and i-f selectivity, and must not, in itself, modify the selectivity.

2—It must be preceded by all of the significant sources of random fluctuations.

3—Its response to both noise and signal (preferably, to any admixture of these) must be known.

Item (1) is often too stringent, and may be relaxed in circumstances where it is known that the frequency spectrum of the noise power presented to the meter input terminals conforms essentially to the shape of the power-selectivity curve at that point. This point will be elaborated later.

A schematic of such a laboratory arrangement is shown in Figure 1, which, in addition displays *one* of the sources of noise, namely, thermal agitation in the resistor R_a which, at room temperature T_o, is simulating radiation resistance. The quantity k appearing in the formula,

$$\overline{e_i^2} = 4k\, T_o\, R_a\, \triangle f \tag{1}$$

is Boltzmann's constant. The quantity $\triangle f$ we will consider related to the overall selectivity curve in the usual way. That is, if f_o be the nominal operating frequency of the receiver, and if $G(f)$ be the response of a linear output meter to input signal $e(f)$, then

$$\triangle f = \int \frac{G^2(f)}{G^2(f_o)}\, df \tag{2}$$

By any of various methods which avoid overloading any portion of the system, it is possible to find a signal $e(f_o)$ which produces (as indicated by the meter) an output signal power equal in magnitude to the output noise power. Were the receiver ideal in the sense that it possessed no sources of noise save that originating in the dummy

[2] loc. cit.

antenna, it would be found that

$$\overline{e^2}(f_o) = 4k\,T_o\,R_a\,\Delta f \tag{3}$$

A fundamental description of the noise generated in the laboratory by a receiver equipped with a dummy antenna is, therefore, provided by the dimensionless measure

$$N = \frac{\overline{e^2}(f_o)}{4k\,T_o\,R_a\,\Delta f} \tag{4}$$

The number N we shall refer to as the "noise-factor". Expressed as a number, it states the ratio of actual noise power to that developed by an ideal receiver. Or, it may, of course, be given in decibels referred to the ideal receiver as a zero level. That is, the lower limit of N is unity, or zero decibels.

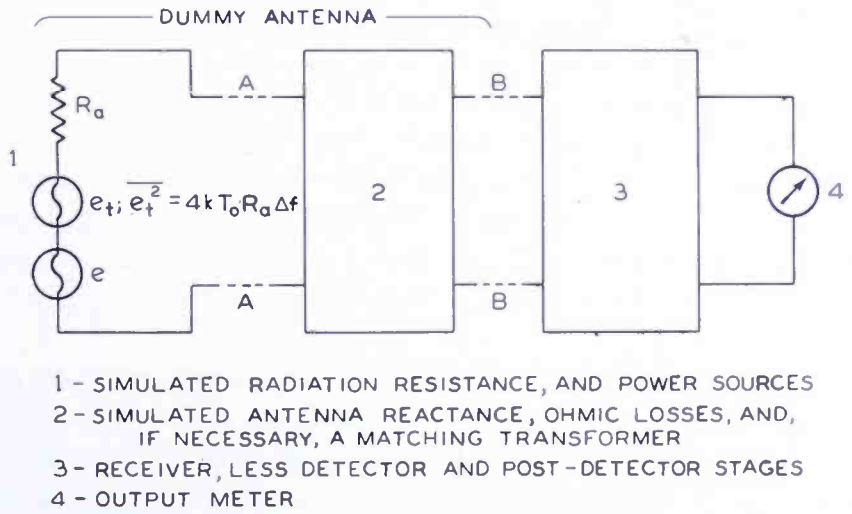


Fig. 1—Schematic of a receiving system with dummy antenna and output meter for laboratory measurements.

This proposed method for quoting laboratory measurement of receiver noise is particularly valuable for two reasons.

First, it is generally fairly insensitive to modifications in the selectivity, as was hinted above. For, unless the frequency spectrum of the noise power output differs materially from the power-selectivity curve, $\overline{e^2}(f_o)$ is proportional to Δf. The noise factor N thus provides a reasonably just basis for the comparison of equipments possessing selectivity curves widely different as to both shape and width.

Second, the quantity e^2/R_a is invariant in a transformation through a loss-free transformer. Consequently, whenever there are no prescribed ohmic losses in the antenna, and when the object of the measurement is not simply to measure N, but rather, to adjust the antenna coupling to an optimum which then produces the *lowest pos-*

sible N, one need not be limited to a value of R_a precisely equal to the radiation resistance of the prescribed antenna. One need only make sure that the prescribed antenna Q is provided, and even this requirement loses importance when the Q is too low to affect materially the overall selectivity. For all special, but highly significant measurements of this kind, it becomes unnecessary to know either the voltage calibration of the signal generator, or the precise value of R_a. Since only the quantity $\overline{e^2/R_a}$ appears in the expression for the noise factor, it is sufficient to know simply the power the combination will deliver to a matched load. In view of this property, it is seen that the minimum noise factor provides, also, a reasonably just basis for the comparison of equipments working out of antennas possessing widely different radiation resistances.

Some Properties of Receiving Antennas

Although we have obtained a method for determining and rating the noise of a receiver with dummy antenna, we are not yet in a position to describe its absolute sensitivity in operation. We must first learn two things: one, the connection between signal-field strength and voltage e produced thereby at the open terminals of an antenna of radiation resistance R_a; two, the rôle played by noise picked up by the antenna.

The rôle played by antenna reactance is altogether silent and will be ignored.

Consider two antennas as pictured in Figure 2. The reciprocity law (which is only a dignified way of saying that transfer impedances for passive networks are the same in both directions) assures us that if i is the current fed to one antenna, and e the resulting open-circuit voltage at the other, then

$$e = k\,i$$

and k is symmetrical in the two antennas. But if antenna (b) is transmitting, its power output is

$$P = i^2 R_b$$

And if antenna (a) is receiving some of that power, its open-circuit voltage will be proportional to the square-root of P. Therefore,

$$e \propto \sqrt{P} = \sqrt{R_b}\,i$$

It follows that k is proportional to $\sqrt{R_b}$, and, being symmetrical in the two antennas, must be proportional to $\sqrt{R_a}$ also. It is, in addition, proportional to a space-attenuation factor and to the product of the directivities of the two antennas.

We are thus enabled to write the relation between signal field strength and open-circuit antenna voltage. Let Ω represent the two angular coordinates which define a direction in space. Let ϕ represent the azimuth in a normal plane, to define the direction of polarization of a Poynting vector $S(\Omega,\phi)$ which describes the signal field. Let $D^2(\Omega,\phi)$ be the function which describes the power-directivity receiving pattern (identical with the radiation pattern in unencumbered

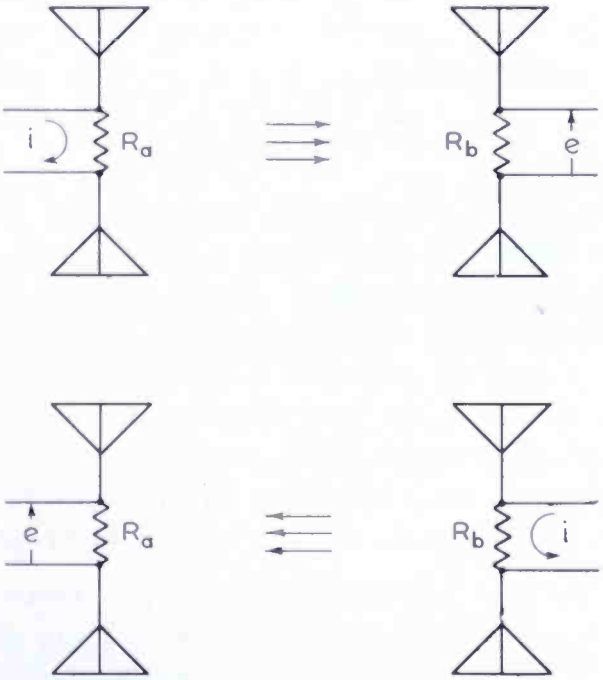


Fig. 2—The reciprocity law for two antennas with individual radiation resistances R_a and R_b.

space), and let it be normalized so that, averaged over all values of Ω and ϕ, $\overline{D^2(\Omega,\phi)} = 1$. It follows that

$$e^2 = A\, R_a\, D^2(\Omega,\phi)\, S(\Omega,\phi) \tag{5}$$

The coefficient A is evaluated in the Appendix by an appeal to Nyquist's theorem, for which an additional derivation is produced, proving (if, indeed, proof be needed) its valid application to radiation resistance as well as ohmic resistance. The result, Equation (15), is

$$A = \lambda^2/2\pi$$

where λ is the wavelength.

Hence,
$$e^2 = \frac{\lambda^2}{2\pi} R_a\, D^2(\Omega,\phi)\, S(\Omega,\phi) \tag{6}$$

When practical units are used, and λ is given in meters, the relation between S (watts per square meter) and its associated field strength E (volts per meter) is

$$S = E^2/120\pi \tag{7}$$

The open-circuit voltage produced at the antenna terminals by a signal field is then

$$e^2 = \frac{\lambda^2}{2\pi} \frac{R_a}{120\pi} D^2(\Omega,\phi) E^2 \tag{8}$$

The voltage produced by a noise field may be likewise expressed. It is certain that if the antenna were exposed solely to the isotropic equilibrium-radiation of an enclosure maintained at room temperature T_o, the noise voltage would be

$$\overline{e_i^2} = 4k\, T_o\, R_a\, \triangle f$$

It is equally certain that this is not the case. The strength of local noise fields has been, and will always continue to be, the subject of much serious speculation and measurement.[5] While it is known already that both man-made noise and atmospherics decrease rapidly with λ, the "resolving power" of receivers is not yet sufficient to determine what limit, if any, is approached. Furthermore, the development of new communications equipment for shorter wavelengths will always, of itself, implement the production of what must necessarily be regarded, from certain points of view, as new noise fields.

Local noise fields may be crudely divided into three groups:

1—Strictly random fluctuations similar to those generated within the receiver proper, and characterized by a proportionality between noise power and bandwidth.

2—Noise consisting of impulses, random enough, but occurring at a mean rate too low in comparison with the bandwidth to be included with (1), and, therefore, characterized by a lack of proportionality between noise power and bandwidth.

3—All other unwanted fields.

It is beyond our present purposes to discuss the relative amounts of service disturbance engendered by equal amounts of noise power in each of these categories. This vastly complex problem is receiving attention elsewhere.[6] Concerning ourselves with a measure of the

---

[5] For example, K. G. Jansky, "Minimum Noise Levels Obtained on Short-Wave Radio Receiving Systems", *Proc. I.R.E.*, Vol. 25, p. 1517, December, (1937).

R. K. Potter, "An Estimate of the Frequency Distribution of Atmospheric Noise", *Proc. I.R.E.*, Vol. 20, p. 1512, September, (1932).

G. Reber, "Cosmic Static", *Proc. I.R.E.*, Vol. 28, p. 68, February, (1940).

[6] C. M. Burrill, "Progress in the Development of Instruments for Measuring Radio Noise", *Proc. I.R.E.*, Vol. 29, p. 433, August, (1941).

power alone, we note that the difference between two determinations of the noise factor, conducted as described, one with a dummy antenna, the other with the real antenna in the absence of a signal field, provides just such a measure. We may go even further and, simply for convenience, express this measure in terms of an essentially fictitious temperature T_a of local space, writing

$$\overline{e^2} = 4k\, T_a\, R_a\, \Delta f$$

for the net open-circuit voltage produced by the local noise field at the antenna terminals, so that[7]

$$N\,(\text{operating}) = N + \frac{T_a}{T_o} - 1 \tag{9}$$

The total equivalent noise voltage at the antenna terminals of a system in operation may consequently be written

$$\overline{e^2} = 4k\, T_o\, R_a\, \Delta f \left[N + \frac{T_a}{T_o} - 1 \right] \tag{10}$$

Should it be ascertained that the local noise field consists entirely of the first kind, T_a will then be independent of bandwidth, but not necessarily independent of operating frequency; and if the noise field is not isotropic, T_a will naturally be a function of antenna orientation and directivity. To the extent that noise fields of the second and third kinds are present, T_a cannot even be presupposed independent of bandwidth.

Absolute Sensitivity

If, by the term "absolute sensitivity", we refer to the r-m-s signal field strength which produces at the open antenna terminals a mean-square voltage equal to the equivalent mean-square noise voltage, then, from (8) and (10), that field strength is

$$\overline{E^2} = \frac{240\pi^2}{\lambda^2} \cdot \frac{4k\, T_o\, \Delta f}{D^2(\Omega,\phi)} \left[N + \frac{T_a}{T_o} - 1 \right] \tag{11}$$

[7] E. W. Herold, Reference 3, exhibits the general functional relationship of N to the important sources of noise *within the receiver*, and to the circuit arrangements dictated by the service for which it is designed.

If E is conventionally expressed in *microvolts* per *meter*, λ in meters, $\triangle f$ in megacycles, and T_o set equal to 300 degrees Kelvin (room temperature),

$$\overline{E^2} = \frac{39\triangle f}{\lambda^2 D^2(\Omega,\phi)} \left[N + \frac{T_a}{300} - 1 \right] \qquad (12)$$

The expression shows clearly the distinct, prominent contributions of receiver proper, antenna directivity, and local noise fields.

The economy of efforts to improve the laboratory noise factor is seen to be bounded, in some cases sharply, by the existence of local noise. The use of tuned antennas for broadcast reception is, in many locations, a costly luxury. In some recent experiments, W. R. Ferris[8] and the author produced a noise factor $N = 3.2$ in an attempt to determine the lower limit for television service at 100 megacycles. Such a receiver would probably be unappreciated in metropolitan districts. On the other hand, at higher frequencies, T_a appears to decline, as stated before, while minimum N rises rapidly for a number of tube and circuit reasons.[3] There is surely much to be gained from improved N's at very short wavelengths.

The economy of antenna structures needs similar consideration. For half-wave dipoles in free space, D^2 is not a function of wavelength. For dipoles beamed by parabolas large in comparison with λ, D^2 (maximum) is proportional to A/λ^2, where A is the area of the parabola's aperture. Even were the noise factor to remain fixed as one moved towards shorter wavelengths, the absolute sensitivity would certainly suffer if one continued to use dipoles and would, indeed, still suffer, despite the switch to a parabola, unless the aperture of the parabola were given an area comparable to the square of the initial wavelength.

Conclusion

Because of the ease of interpretation, and the simplicity of measurement method, it is hoped that the material presented above may be considered as a basis for the adoption of a standard of "absolute sensitivity" for radio receivers.

Appendix

Thermodynamic reasoning alone is sufficient to justify the application of Nyquist's theorem[9] concerning thermal agitation to any

[3] loc. cit.

[8] Research Laboratories, RCA Manufacturing Company, Inc., Harrison, New Jersey.

[9] H. Nyquist, "Thermal Agitation of Electric Charge in Conductors", *Phys. Rev.*, Vol. 32, p. 110, July, (1928).

passive network. But it is, at least, interesting to develop the theorem in detail for a microscopic dipole exposed to equilibrium radiation.[10]

Using electrostatic units, we know that the radiation resistance of a dipole of length $l (l << \lambda)$ in vacuum is[11]

$$R_a = \frac{8}{3} \frac{(\pi f l)^2}{C^3} \tag{13}$$

where f is the operating frequency, C the velocity of light.

The energy density of equilibrium radiation at temperature T is[12]

$$\rho = \frac{8\pi\, hf^3\, df}{C^3} \left[\epsilon^{\frac{hf}{kT}} - 1 \right]^{-1}$$

wherein h is Planck's constant, k is Boltzmann's constant. But since $hf/kT \sim 10^{-4}$ for $f = 1{,}000$ megacycles and $T = 300$ degrees Kelvin, the Rayleigh-Jeans approximation is sufficiently accurate. Thus,

$$\rho = \frac{8\pi\, kTf^2}{C^3}\, df$$

But, being isotropic,

$$\rho = \frac{\overline{E^2}}{4\pi} = \frac{\overline{E_x^2} + \overline{E_y^2} + \overline{E_z^2}}{4\pi} = \frac{3\overline{E_z^2}}{4\pi}$$

Hence, the mean-square field strength at the antenna in its susceptible direction (z) is

$$\overline{E_z^2} = \frac{32(\pi f)^2}{3\, C^3}\, kT\, df$$

Therefore, the mean-square open-circuit voltage at the antenna terminals is, with the help of (13),

$$\overline{e^2} = \overline{E_z^2}\, l^2 = 4 \left[\frac{8(\pi f\, l)^2}{3\, C^3} \right] kT\, df = 4kT\, R_a\, df \tag{14}$$

* * * * *

To evaluate the coefficient A in (5), we submerge any antenna in equilibrium radiation and demand that the open-circuit voltage appear-

[10] A more general derivation was given by Burgess, Reference 4.

[11] For example, Hund, "High-Frequency Measurements", p. 399, (McGraw-Hill, 1933).

[12] For example, Ruark and Urey, "Atoms, Molecules, and Quanta", p. 58, Eq. 10, (McGraw-Hill, 1930).

ing at its terminals obey Nyquist's theorem. Since $D^2(\Omega,\phi)$ has been defined to have an average value of unity, and since (in *e.s.u.*)

$$\overline{S} = \frac{C\,\overline{E^2}}{4\pi}$$

it follows that

$$4kT\,R_a\,df = \overline{e^2} = A\,R_a\,\overline{S} = A\,R_a\,\frac{C\overline{E^2}}{4\pi} = \frac{2\pi A}{\lambda^2}\cdot 4kT\,R_a\,df$$

whence

$$A = \frac{\lambda^2}{2\pi} \tag{15}$$

This result enables one to state a perfectly general relationship between the "effective height" of an antenna, its radiation resistance, and its directivity. The effective height $h(\Omega,\phi)$ is defined

$$h = \frac{e}{E}$$

Hence, from (8), we have, in practical units,[13]

$$h^2(\Omega,\phi) = \frac{\lambda^2}{2\pi}\,\frac{R_a}{120\pi}\,D^2(\Omega,\phi) \tag{16}$$

We further find, in agreement with Burgess,[4] that the mean-square effective height averaged over all orientations with respect to a fixed signal vector is then

$$\overline{h^2} = \frac{\lambda^2}{2\pi}\,\frac{R_a}{120\pi} \tag{17}$$

---

[13] The quantity, 120π, carries the dimension, ohms. It is, in fact, the characteristic impedance of a strip x units broad of a transmission line consisting of two parallel infinite planes x units apart, supporting plane-wave transmission.

[4] loc. cit.

Finally, with reference to (6), it is to be noticed that the power an antenna can extract from a signal Poynting vector is proportional to

$$\frac{e^2}{R_a} = \frac{\lambda^2}{2\pi} D^2(\Omega,\phi)\ S(\Omega,\phi)$$

The radiation pattern of dipoles shorter than a half-wavelength is sensibly independent of the length; this leads to the curious observation that such dipoles extract sensibly equal amounts of power, no matter how short they are. In the language of physical optics this would be known as a diffraction phenomenon. In the language of atomic physics, we can only conclude that the capture-cross-section of short dipoles for passing Poynting vectors is of the order of λ^2.

AN OMNIDIRECTIONAL RADIO-RANGE SYSTEM

By

David G. C. Luck

Research Laboratories, RCA Manufacturing Company, Inc., Camden, N. J.

PART II — EXPERIMENTAL APPARATUS

Summary—*Experimental omnidirectional ranges have been developed and tested in flight at frequencies of 6425 kilocycles per second and 125 megacycles per second. In each case, a radiating system consisting of five vertical antennas and a metallic ground mat was used. Each transmitter was of a normal radiotelephone type, supplemented by a pair of balanced modulators, an impulse keyer, and a set of modulation controls. Full monitoring of the effect of all transmitter adjustments was provided. Essentially normal aircraft receivers and antennas were employed. Both cathode-ray azimuth indicators and pointer-type deviation from course indicators were provided.*

1—Introduction

THE previous section of this paper[1] compared the various methods available for radio guidance and showed the place in this art occupied by devices which indicate directly, continuously, and automatically the direction of mobile receiving craft from special fixed transmitters. We have called such devices "omnidirectional radio ranges." The previous part of the paper also gave a detailed explanation of the principles of operation of one form of omnidirectional range system and an analysis of the effects of various possible sources of instrumental error. In what follows, familiarity with this explanation is assumed in order to avoid repetition. The purpose of the present part of the paper is to describe apparatus developed in the course of experiments on this type of omnidirectional range. The results obtained in these experiments and some methods of using the instrument for aircraft guidance will be described in a third part, to appear later.

Experiments on an omnidirectional range system were begun at Camden early in 1936, a frequency of approximately 6½ megacycles per second being chosen for reasons of convenience. The first flight test, in November 1936, gave the expected automatic bearing indication

[1] "An Omnidirectional Radio-Range System. Part I—Principles of Operation" by David G. C. Luck, RCA Review, Vol. VI, No. 1, p. 55, July 1941. Hereafter referred to as I.

in all directions, but the accuracy attained was not of a practically useful order. Continuing work on details of apparatus and operation resulted, by the Spring of 1937, in usefully accurate and stable omnidirectional bearing indication. Field tests were continued through the summer of that year to complete the work.

In view of the strong general trend of aviation radio toward ultra-high frequencies and of the success of the early work, it was decided to undertake the development of an omnidirectional range to operate at 125 megacycles per second and to have instrumental errors not

Fig. 1—6425-kc omnidirectional range.

exceeding three degrees. Work was begun early in 1938 and equipment was built and installed by the latter part of that year. As there was at that time considerable interest in the project, an attempt was made to get the new system into operation early in 1939. As a result of the accuracy required and the problems of ultra-high-frequency technique encountered, however, quantitatively acceptable flight results were not attained until the late summer of that year. Flight testing continued until early in 1940 and the apparatus was kept in operative condition until dismantled upon the closing of Central Airport at Camden, New Jersey, at the end of July 1940. Testing being then practically completed, the system has not been set up again.

2—Description of Apparatus

A—Medium Frequency (6425 Kilocycles Per Second)

a—Antenna System

Figure 1 is an aerial photograph in which the radiating system used at the medium frequency is shown in relation to its surroundings. The two power lines in the background and the swamp in the foreground, in conjunction with the general slope of the ground toward the swamp, made the location far from ideal for accurate directional

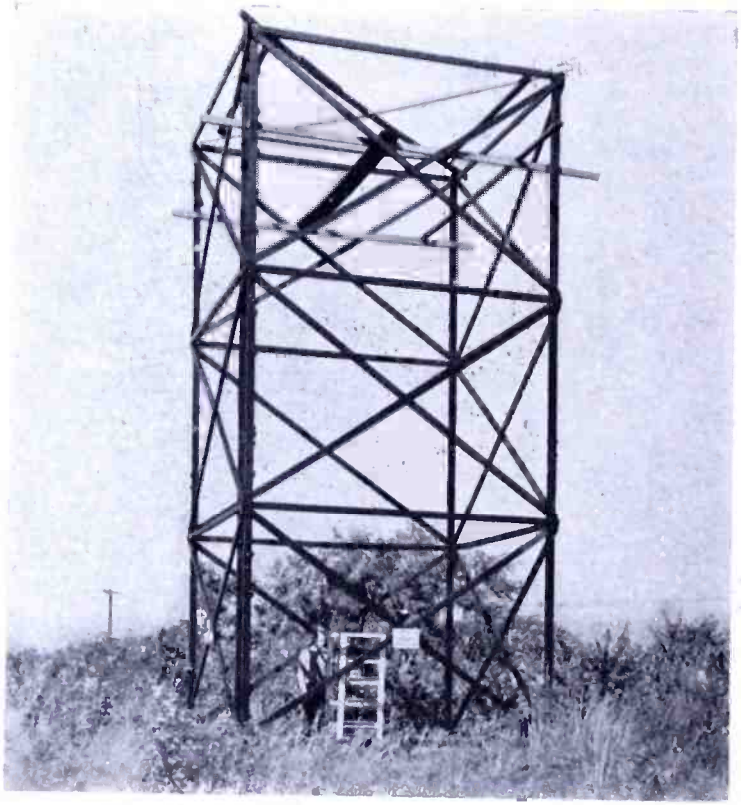


Fig. 2—6425-kc range antenna system.

radio work. The northeast corner of Central Airport, Camden was just off the lower left corner of the picture.

The wooden-lattice tower supported five 30-foot vertical wires accurately located at the corners and center of a 12-foot square, with diagonals accurately aligned in true north-south and east-west directions. A counterpoise, erected in a level plane tangent to the highest part of the sloping ground, comprised a 24-foot square of expanded copper sheet, fringed with 96 equally spaced radial wires extending 75 feet from the center of the antenna system. The circular area covered by the counterpoise shows clearly in the picture.

The transmitter was located directly at the center of the base of the antenna square, as Figure 2, taken before erection of the counterpoise, shows. The central, broadcast antenna terminated directly at

the top of the transmitter. Each corner antenna was connected to the center of the square, at ground level, by a short radial length of coaxial line with the outer conductor grounded immediately under the antenna, but otherwise insulated. Outer conductors were continuous along each full diagonal of the antenna square; the inner conductors were split at the center of the square and there fed by short balanced lines from the transmitter output terminals to ground level. These highly symmetrical directive antenna circuits were ungrounded, and they apparently maintained antenna balance to ground quite well. Antenna height

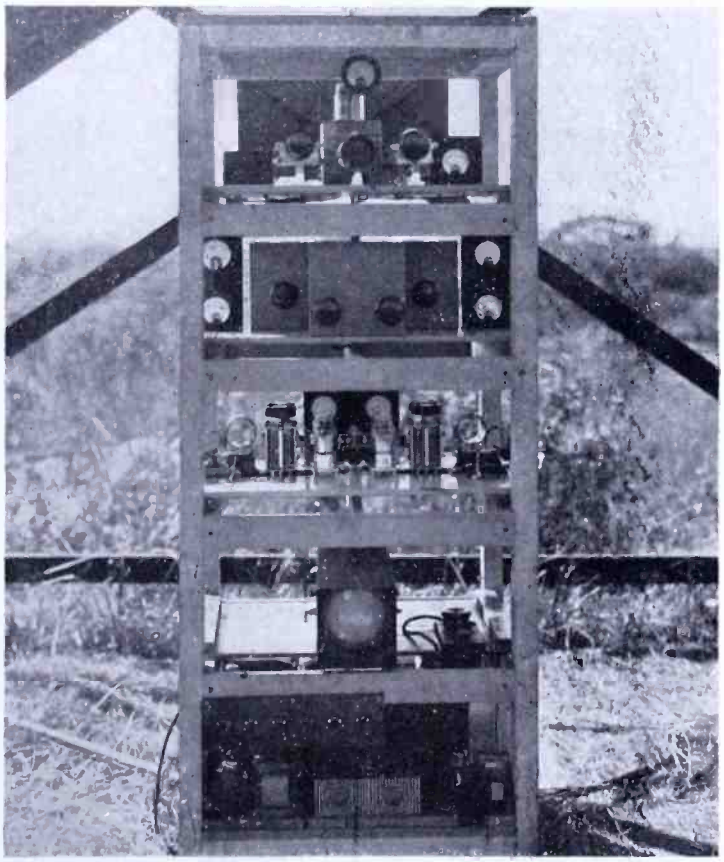


Fig. 3—6425-kc experimental range transmitter.

was about 1/5 wavelegth and spacing 1/10 wavelength at the operating frequency.

b—Transmitter

A very simple transmitter was used, shown by Figure 3. On the bottom shelf will be seen power-supply and impulse-keying equipment. On the shelf above, speech modulating equipment was installed later. The second shelf from the top carried a crystal oscillator and two driver amplifiers, feeding radio-frequency excitation upward to the 150-watt, broadcast-channel pentode amplifier on the top shelf and downward to the directive-channel balanced modulators on the middle shelf. Parallel grid excitation, push-pull plate modulation with only alternat-

ing plate voltage applied, and push-pull output were the operating conditions of the balanced modulators.

Two-phase, 60-cycle power with good waveform, for the balanced modulator plates, was derived from the single-phase supply line by shunting one plate-transformer primary with a capacitor and feeding the combination through a variable series inductor. This inductor was made from a small Variac by introducing an air gap into its core. The whole combination was fed from a standard Variac giving amplitude control. This scheme worked well, but the control of phase and amplitude by the two Variacs was not independent and readjustment was necessary whenever balanced-modulator tuning changes altered the load.

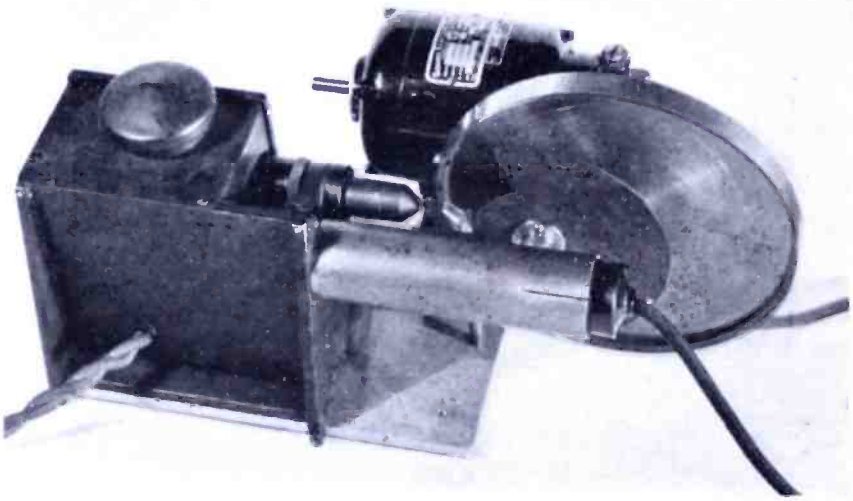


Fig. 4—Station-announcing phonograph.

A novel item in the speech equipment was the simple film phonograph shown in Figure 4, which repeated the 3-second announcement "Central Airport, Camden, New Jersey" at 15-second intervals. A normal 16-mm sound track projected beyond the edge of the large disc, around which it was wrapped. It ran, at normal linear speed for this type of film, past a fixed scanning line of light. Modulated light passing through the sound track was reflected, through a geared sector mask to cut out four of every five announcements, to a phototube in the cylindrical shield. A high-pass filter, with a 60-cycle rejection point, was interposed between the speech amplifier and speech modulator to minimize 60-cycle hum modulation of the broadcast carrier signal.[2]

c—Monitoring

Four carefully compared radio-frequency ammeters, installed at the bases of the four corner antennas, served both to show that antenna

[2] I—Sect. 3, C, p. 76.

symmetry was maintaining equal currents in the two antennas of each diagonal pair and to permit adjustment of north-south and east-west directive-pattern amplitudes to equality.[3] One of the two windings of a dynamometer meter was connected across each balanced-modulator plate-transformer primary, so that the desired condition of phase quadrature between the two modulating voltages[4] resulted in a null indication on this meter.

Phase of the reference keying was not controlled in the medium-frequency set-up, but was merely initially set to a convenient value, the particular value being of no importance in a single transmitter so long as it remains constant between receiving-indicator zero checks.[5] The desirability of being able to get such zero checks at any time was not at first realized, so no provision was made for non-directional transmission of calibrating modulation. Calibration was accomplished by turning off the balanced modulator delivering the smaller signal at the receiving point, thus causing the receiving indicator to read that cardinal bearing which lay nearest the true receiver bearing. Provision was made, however, for supplementing the keying impulses—normally occurring once per modulation cycle—with three others to give four equally-spaced impulses per cycle for use in making complete receiving indicator adjustments.

Toroidal pickup windings were placed around the bases of the antennas and connected directly to the plates of a cathode-ray oscillograph through short unloaded lines, for monitoring radio-frequency adjustments. By connecting diagonally opposite toroids together in parallel, their electrostatic coupling to the antennas was cancelled. Highly characteristic pattern shapes on the oscillograph indicated phase balance of the two tubes of each balanced modulator and phase quadrature between broadcast and directive antenna radio-frequency currents.[6] Radio-frequency adjustments were always made with the regular transmitter tuning controls.

d—Indicators

Regular communication antennas on several aircraft were used for reception, the receiver being an aircraft superheterodyne with crystal-controlled oscillator. This receiver was modified by being given a very flat automatic gain control, broadcast-type intermediate-frequency transformers to pass the rapid signal changes of the reference keying, and an audio-output transformer having a characteristic which was

[3] I—Sect. 3, C, p. 65.
[4] I—Sect. 3, C, p. 75.
[5] I—Sect. 3, C, p. 74.
[6] I—Sect. 3, D, p. 77.

flat down to 60 cycles. Otherwise it was an entirely normal aircraft receiver, which simply fed audio output from its headphone jack to the indicators.

In the cathode-ray indicator[7] two loosely-coupled circuits, tuned to 60 cycles per second, selected the pattern-rotation component from the receiver audio output. Output from this filter was applied directly to one set of deflecting plates of the electrostatically deflected cathode-ray tube, and through a 90-degree phase shifter of adjustable phase and amplitude characteristics to the other deflecting plates. Impulses were selected from the receiver output by a simple high-pass filter and applied to the grid of the cathode-ray tube, momentarily cutting off the electron beam and darkening its trace on the screen of the tube. Accelerating voltage was applied to the cathode-ray tube, from a vibrator unit with vacuum-tube high-voltage rectifier, through the normal brightness, focus, and pattern-centering controls.

Thus, the cathode-ray indicator contained only two tubes, the high-voltage rectifier and the cathode-ray tube itself. The latter was mounted in, and magnetically shielded by, a conical housing of thin Nicaloi. The indication given consisted of a circle of green fluorescent light, with a narrow dark gap moving around it to represent motion of the receiving craft around the transmitter. This indication was read against a circular engraved-celluloid scale pressed against the tube face. A green gelatine light filter was placed behind the scale, to improve contrast under adverse lighting conditions by suppressing much of the reflection from the white tube face without markedly dimming the green cathode-ray pattern. Setting of the scale zero was accomplished by turning the scale in front of the tube, or in the rough initial stage by turning the tube in its mounting. Accuracy of reading at points other than zero was assured by adjusting vertical and horizontal centering and deflection phase, with a special four-impulse-per-cycle calibrating transmission, until the four dark marks then present indicated correctly the four cardinal points; finally, relative amplitude of vertical and horizontal deflections was adjusted by eye to give the best-appearing circle.[8]

A deviation-from-course indicator was provided as a separate accessory unit.[9] Pattern-rotation sinusoid output, from the tuned filter of the cathode-ray indicator, was applied to a universal phase shifter and thence, in push-pull, to the suppressor grids of a pair of pentodes. Positive impulses were applied to the normally cut-off control grids of these tubes in parallel. A differential plate-current meter indicated deviation of the impulse phase from the instant of equal suppressor-

[7] I—Sect. 2, B, pp. 67-70.
[8] I—Sect. 3, F, p. 80.
[9] I—Sect. 2, C, pp. 70-72.

grid biases, which was in turn determined by the phase-shifter setting and by the bearing of the receiver from the transmitter.

An auxiliary arrangement for aural deviation indication was tried, but was not entirely satisfactory. This employed an unsymmetrical multivibrator to connect a pair of headphones alternately to each of the differential impulse tubes mentioned above. The connection was made for a relatively short time to one tube, and for a much longer time to the other. On course, a steady 60-cycle impulse tone was heard, while off course to one side telegraphic dots of this tone were heard predominantly, and off course to the other side dashes were heard. Avoidance of timbre changes in the on-course tone was difficult, and in any case the tone was not very markedly audible and was hard to distinguish from ignition interference.

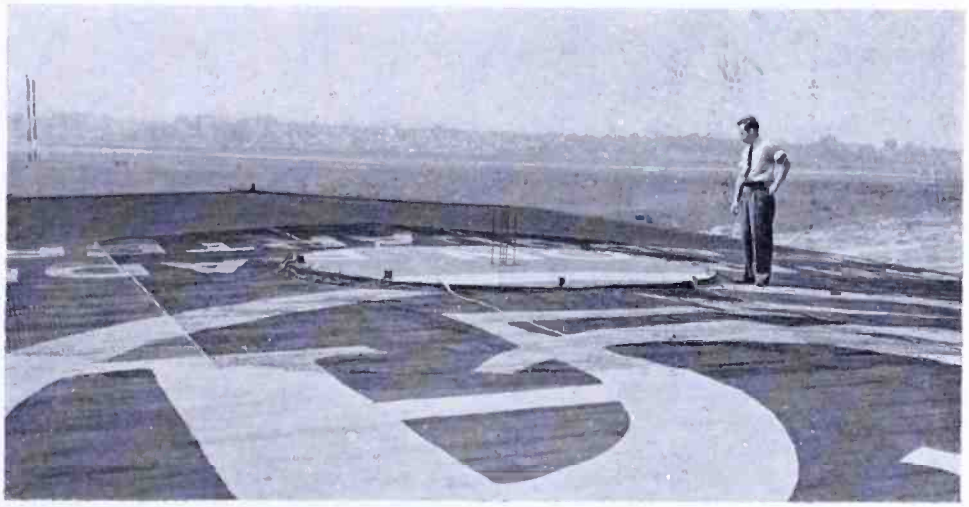


Fig. 5—125-Mc Adcock antenna system.

B—Ultra-High Frequency (125.0 Megacycles Per Second)

a—Antenna

At the ultra-high frequency chosen for our later tests, with a wavelength just under eight feet, the Adcock antenna system consisted of five two-foot rods, set at corners and center of a seven-inch square, and a fifteen-foot diameter copper ground disc, as shown in Figure 5. This system was set up in the middle of a hangar roof and, to avoid ill effects from imperfect bonding of the metal structure of the roof, was surrounded by a fringe of expanded copper sheet to give an overall ground-plane diameter of 45 feet. A vertically polarized antenna system was chosen because it was felt that a set of vertical rods could be more easily designed, built, adjusted, and fed than the corresponding set of horizontal loops which would have been necessary to provide a horizontally polarized signal.

Across the front of the hangar, about 60 feet from the antenna system, there was a brick parapet over ½ wavelength high and about

six wavelengths wide. To facilitate measurements on the roof, the energy approaching this barrier was deflected upward by a large inclined sheet of copper netting interposed between parapet and antennas. Metal uprights supporting obstruction lights were made demountable and taken down during operation to avoid re-radiation. A brick chimney, bearing about fifteen degrees from the antennas, also gave rise to standing waves on the roof, but no practical way to avoid this disturbance of the field pattern was found.

All antennas were insulated from ground at their bases and fed, through the ground plate, by the system of coaxial lines inside the

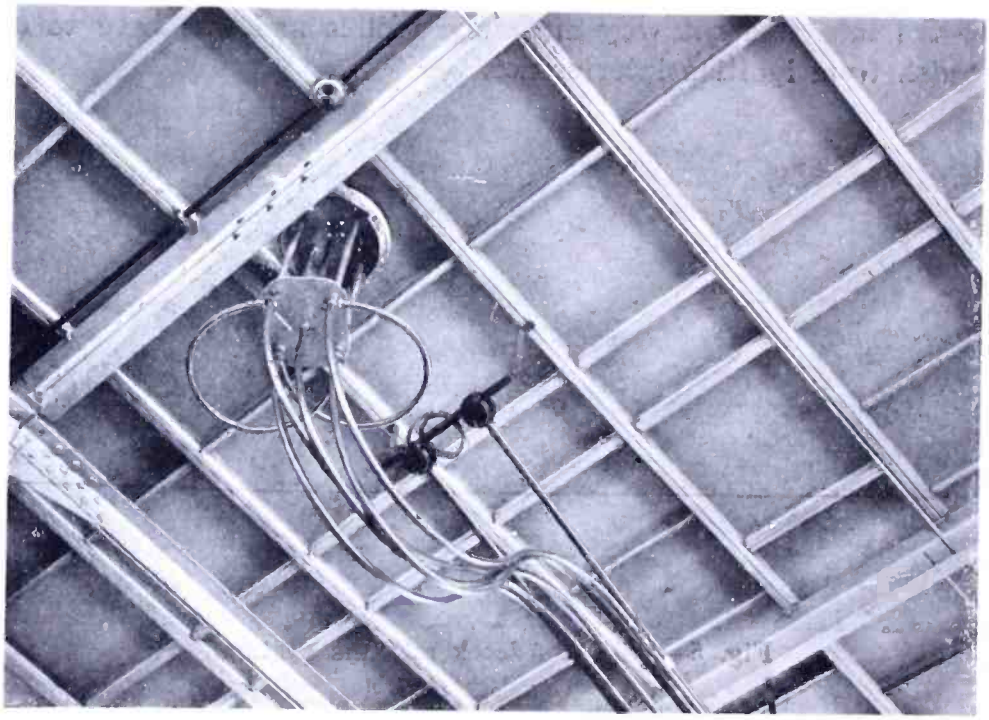


Fig. 6—125-Mc antenna feeding system.

hangar, as shown in Figure 6. The short straight line-segments shown immediately below the circular antenna base-plate were quarter-wave impedance transformers. With each antenna length adjusted to resonance, these transformers were designed to have the correct characteristic impedances to match the antenna base resistance (only three ohms for each directive antenna because of the proximity of the other antennas) to the approximately 60-ohm impedance actually characteristic of commercial ⅞-inch coaxial line. Curved line-sections seen joining the bases of diagonally opposite transformers were half-wave stabilizing and decoupling elements, operating as described below. The five lines running out of the picture were the matched feeders bringing power from the transmitter, which was about eight wavelengths distant.

Considering one directive antenna pair only, its feed system is shown in section by Figure 7(a). The antennas *AB* and *FG* were shorter than a quarter wavelength, by the small amount necessary to produce resonance or zero reactance at the base. The conditions to be fulfilled were that radio-frequency currents at the antenna bases *B* and *F* should always be equal in amplitude and opposite in phase.[10] Introduction of the quarter-wave line-sections *BC* and *EF* not only inverted the low antenna base resistance appearing at points *B* and *F* to a chosen higher value at the feeding points *C* and *E*, but also insured the fulfillment of the above antenna-current relations if equal and opposite radio-frequency voltages were maintained at these feed points.

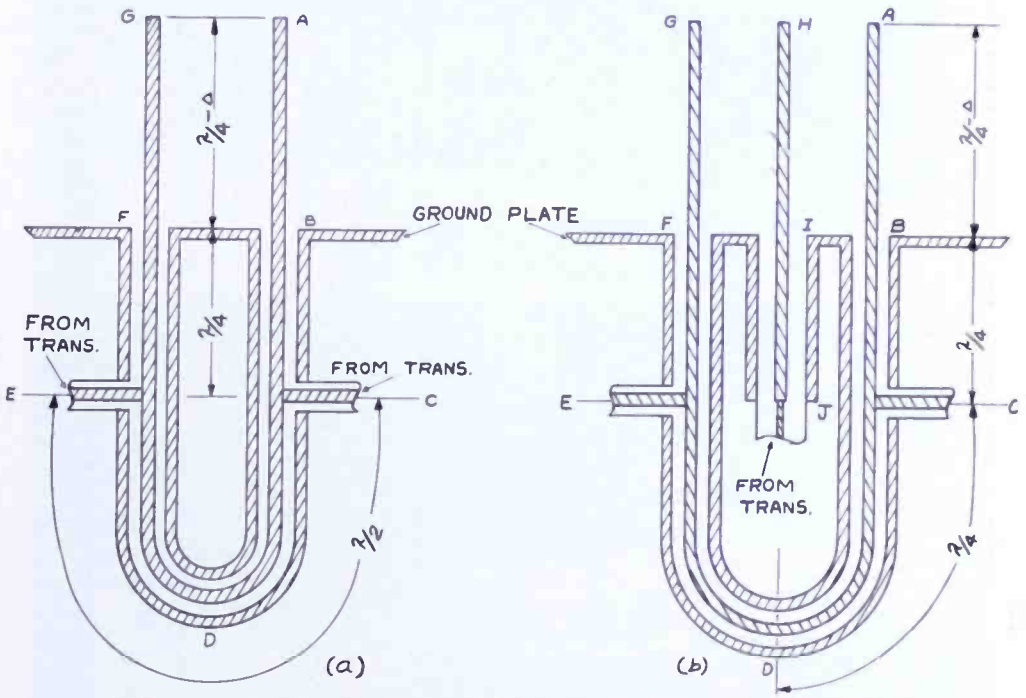


Fig. 7—Operation of ultra-high-frequency antenna feed.

By use of the well-known phase-inverting property of half-wave transmission lines, equal and opposite voltages were constrained to occur at points *C* and *E* by interconnecting them through the half-wave line section *CDE*. In the normal case, when the feed lines from the transmitter were well balanced and were supplying equal and opposite voltages at *C* and *E*, the line *CDE* simply bore a standing wave and transported no power. Should point *C*, for example, tend to be fed at a higher voltage than point *E*, the half-wave line would carry power from *C* to *E* to oppose this unbalance and maintain the required phase and amplitude relations.

Midway between the antennas of a directive pair there was also a non-directive antenna, as shown at *HI* of Figure 7(b). Current in this antenna tended to induce equal and co-phasal currents in the adjacent

[10] I—Sect. 3, D, p. 78.

antennas *AB* and *FG;* this result was not in itself particularly harmful, but was found quite annoying because it made the input impedance of the non-directive antenna depend on the circuits feeding the directive antennas. Consider for the moment that the half-wave stabilizing line *CDE* is cut at point *D* and the feed line to point *C* from the transmitter is disconnected. Then the quarter-wave line *CD*, open at *D*, appears as a short-circuit at *C;* thus, the impedance at *C* is zero inde-

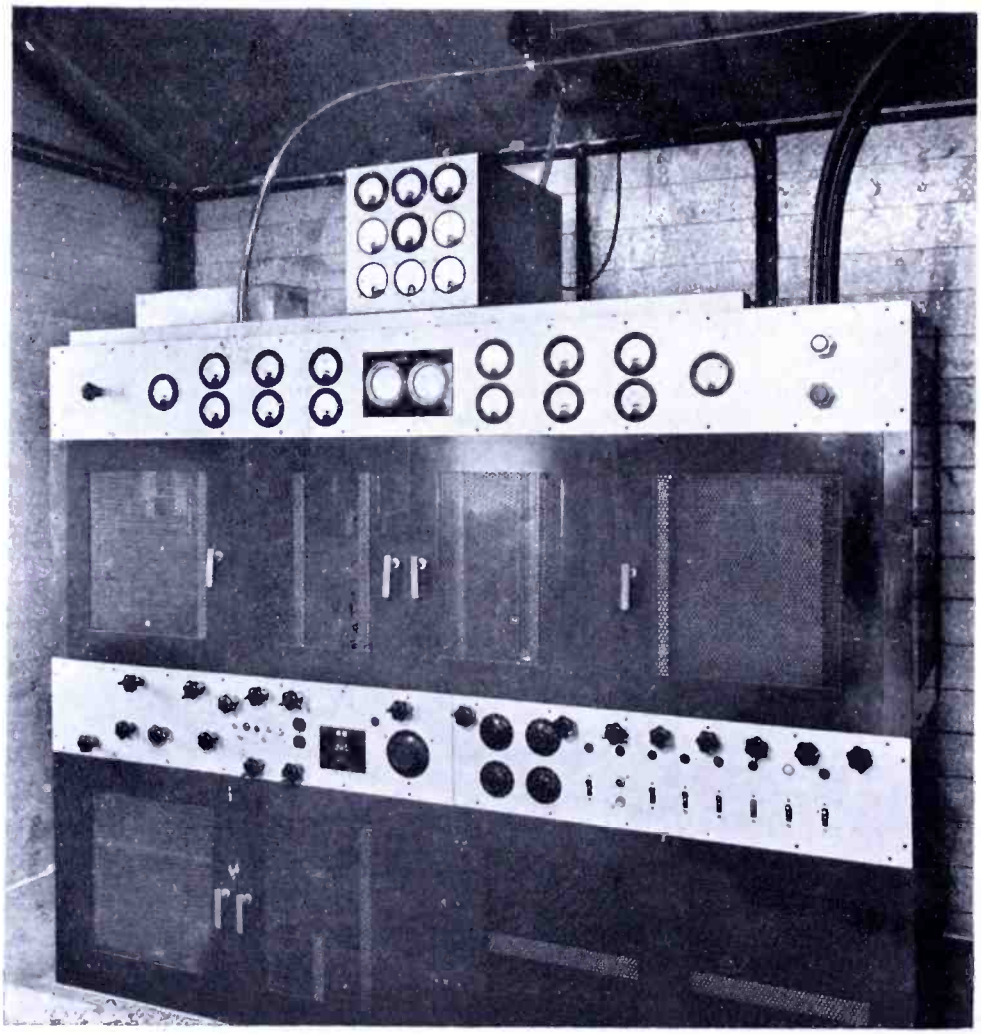


Fig. 8—U-H-F omnidirectional range transmitter.

pendent of any feed line which may be connected there. The quarter-wave line *BC*, effectively short-circuited at *C*, appears as an open circuit at the antenna base *B*, so that any voltage induced in the quarter-wave antenna *AB* can cause no antenna current to flow to ground. By symmetry, the voltages induced in antennas *AB* and *FG* by current in *HI* are equal and in phase, so that the induced voltages transferred to the hypothetical cut at point *D* by the lines *BCD* and *FED* are identical and this assumed cut in the line can be mended without altering anything. Thus, the half-wave line section acted to

short-circuit both point *C* and point *E* for any co-phasal voltages appearing there, at the same time as it equalized any contra-phasal voltages there applied.

b—Transmitter

Figure 8 shows the appearance of the 125-megacycle range transmitter. Its general arrangement is shown in the block diagram of Figure 9, where heavy lines represent radio-frequency paths, double lines modulation paths and light lines power supply. Above the dotted line appears a normal 200-watt, 125-megacycle radiotelephone transmitter. This employed a third-harmonic, low temperature-coefficient

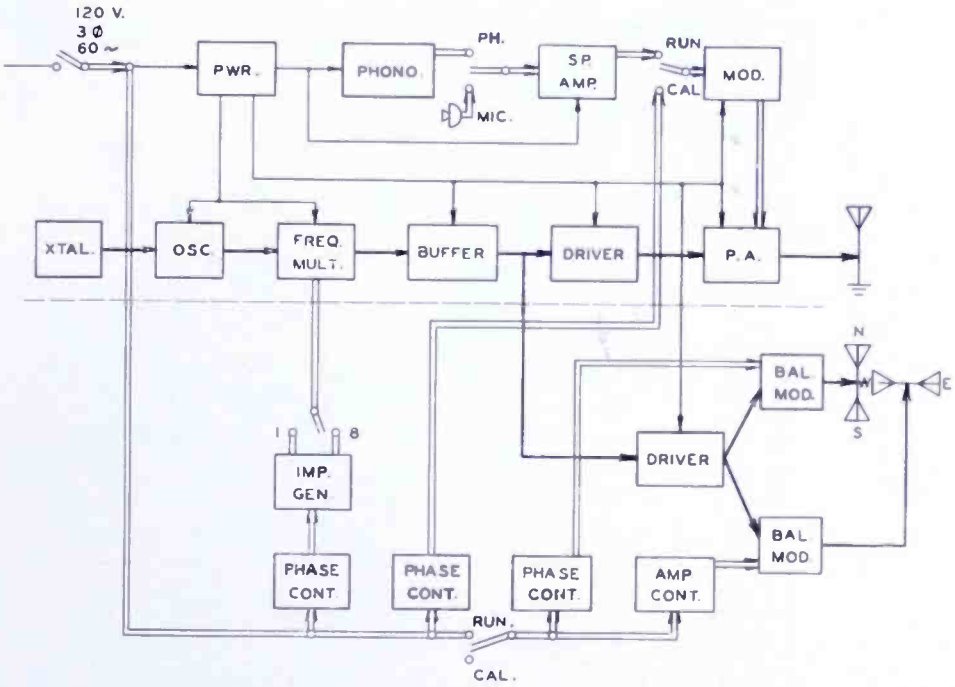


Fig. 9—Block diagram of omnidirectional range.

crystal controlling a pentode acting as oscillator and electron-coupled frequency doubler, a full-wave doubler stage, a push-pull tripler, and a neutralized push-pull buffer amplifier as its exciter system. This exciter energized neutralized driver stages operating at final frequency and a neutralized push-pull power-output stage was modulated up to 50 per cent by a class-A audio power amplifier. An RCA-804 was used as the oscillator, four 845's as the modulator, two 806's in the power-output stage, and pairs of RCA-834 tubes in all other stages. Grid-leak bias was used on all radio-frequency stages, with just enough cathode-resistor self-bias to prevent overloading in the absence of excitation. For tuning, all 125-megacycle circuits employed "trombone" variable inductances, resonated by tube-plate and neutralizing capacitances.

Below the dotted line of Figure 9 are indicated the elements distinguishing the omnidirectional range from an ordinary transmitter:

a pair of balanced modulators, an impulse keyer, and their associated power controls. Figure 10 illustrates the circuits of a neutralized amplifier and one of the pair of balanced modulators driven by it; the other modulator was identical with the one shown. The sliding contacts which permitted inductance variation may be seen to have been at positions of relatively low current in the standing waves on the electrically long inductance trombones. Only alternating voltage was applied to the balanced modulator plates. Small, differentially variable condensers were connected between modulator-tube plates and ground to permit balancing of the modulator output. These were adjusted by watching the pattern on a cathode-ray oscillograph tube with one pair

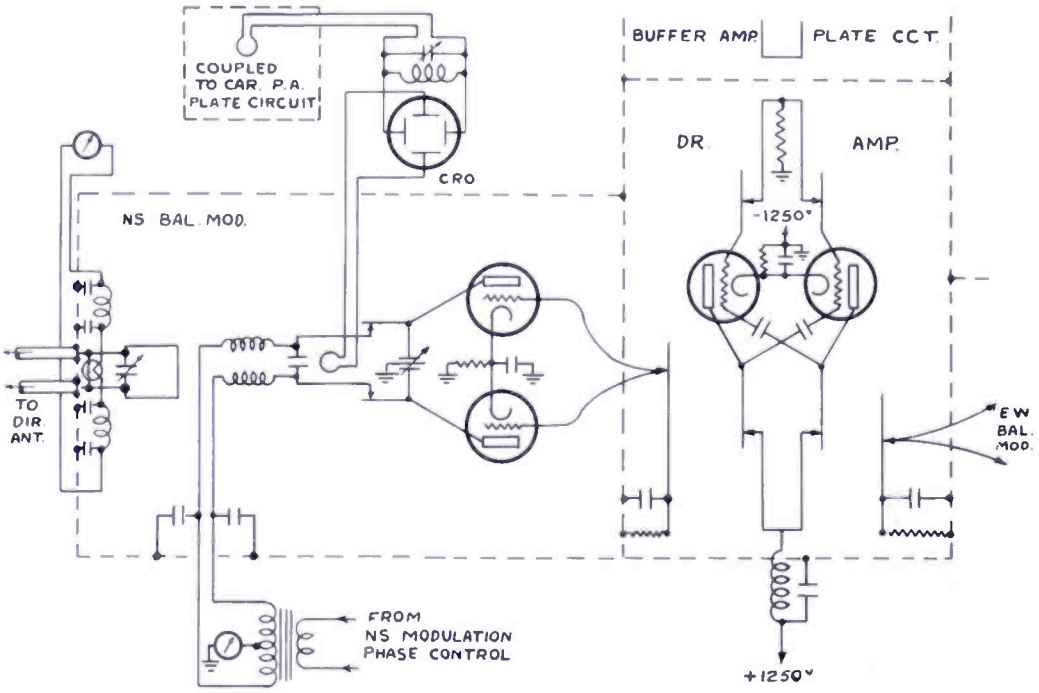


Fig. 10—Balanced modulator.

of plates coupled to the modulator output, as shown, and the other pair coupled to the carrier power-amplifier output. In general, the oscillograph pattern was a patch of dim light bounded by two crossed ellipses. Variation of the phase of the steady reference wave from the carrier amplifier, by adjustment of the tuned circuit shown in one of the deflecting circuits, ordinarily permitted either bounding ellipse (but not both together) to be collapsed to a straight line. Correct adjustment of the differential capacitor, giving equally phased outputs from both tubes of the modulator, was the only condition that permitted both ellipses to be closed at once.

Thermocouple radio-frequency voltmeters were connected across the feed lines to the directive antennas and served as antenna-circuit tuning indicators. Some care was required in installation to avoid spurious readings of these indicators caused by standing waves on the feeders

at harmonic frequencies. The output required from the balanced modulators was surprisingly low. Directivity, both vertical and horizontal, of spaced antenna pairs gave a geometrical power gain of 2½ with respect to the single carrier antenna; thus, an unmodulated directive signal of maximum strength equal to that of a 200-watt broadcast carrier would require only 80 watts into one pair of directive antennas. Since about 40 per cent modulation was found convenient, only 16 per cent of the latter power was required, or a peak output of 12.8 watts from each balanced modulator. Modulation being sinusoidal, average output was one-half the peak output, and the average power requirement was shared equally between two tubes in each balanced modulator. Thus, the average output power required per tube was only 3.2 watts,

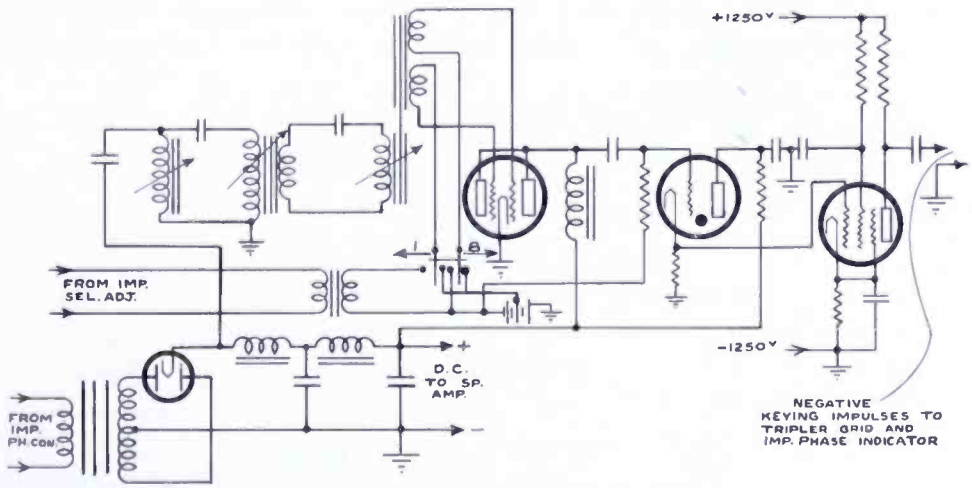


Fig. 11—Impulse keyer circuit.

and the peak per tube 12.8 watts, to provide 40 per cent modulation of the 200-watt broadcast carrier, successively in all directions. Actually, sufficient output to modulate a two-kilowatt carrier was easily obtained from the RCA-834 balanced modulators, without approaching limiting ratings or encountering serious instability. At the time this work was done, tube limitations at high frequencies prevented a two-kilowatt carrier output from being obtained economically at 125 megacycles per second and caused the adoption of the 200-watt figure. Directive modulation was set at the desired level, which was not in any way critical, by choice of balanced-modulator output coupling.

Figure 11 shows the circuit of the impulse-keying equipment. Keying phase was determined from the alternating component of the output of the full-wave rectifier supplying plate power to the speech amplifier. The rectifier output was connected to a series of circuits tuned to 240 cycles per second and terminating in a peaking transformer, of the usual type employing an easily-saturated lamination in one core leg, with push-pull output. Each side of the peaking-transformer output fed one grid of a twin triode, so that with both grids

biased just to cut-off in the absence of input, there appeared across the common plate-load inductance 480 equally-spaced voltage impulses of double polarity per second, or eight per modulation cycle. Accurate tuning of the 240-cycle circuits was essential to assure uniform spacing of these impulses. By biasing one triode completely off and adding a modulation-frequency component to the bias of the other triode, it was possible to select one of the eight impulses, which appeared alone 60 times per second, or once per modulation cycle, in the common triode output. These impulses gave sharp control of the discharge of a capacitor through a non-oscillating gas triode, which improved their shape without affecting their timing. Positive impulses from the gas-

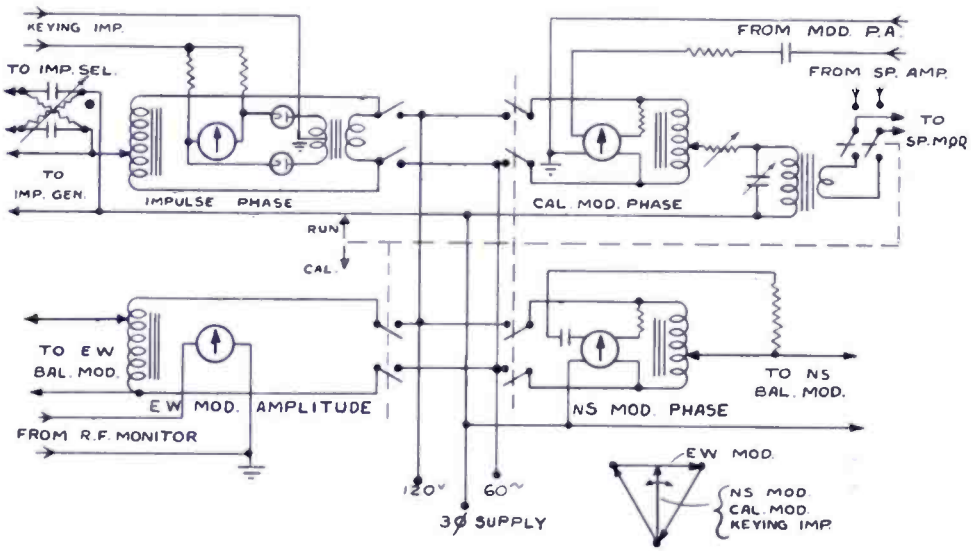


Fig. 12—Modulation control and monitoring circuits.

triode cathode circuit were amplified to give large negative impulses for keying the grids of the tripler stage. These rendered this stage inactive for periods of rather less than 100 microseconds, accurately recurring once or eight times per modulation cycle at will.

Two-phase supply for the balanced modulators was easily obtained in the case of the ultra-high-frequency transmitter, since this took its power from a three-phase line. As shown in Figure 12, the east-west balanced modulator was fed directly from one phase of the supply, through an amplitude adjusting Variac. The north-south modulator was connected between the variable tap of a Variac, across the east-west supply phase, and the third supply line; this Variac permitted phase adjustment over a ±30 degree range with the desired quadrature condition at midrange. The small vector diagram in the figure shows the phase and amplitude relations involved. Calibration and reference-keying phases were controlled in just the same way as the north-south

modulation phase.[11] It will be noted that the transformer feeding calibrating signal to the speech modulator was provided with a series resistor and shunt capacitor. These permitted initially bringing the proper Variac setting to midrange, despite phase shift in the speech-modulator circuits; they at the same time permitted setting the calibrating modulation at the same level as the normal directive modulation. In the circuit supplying the impulse-selecting bias voltage to the impulse generator, a resistance-capacitance network permitted initially selecting for normal range operation, from the eight impulses per modulation cycle provided by the peaking transformer, the one that allowed the impulse-phasing Variac to run nearest mid-range. This control system worked very smoothly in itself, but sudden line phase changes, probably due to changing load distribution on the supply system, were sometimes observed. All modulation switches were ganged, as indicated in Figure 12, to permit changing over from normal directive running to non-directive calibrating transmission with a single operation.

The simple film phonograph of Figure 4 was remodeled by removing the sector disc, adding to the phonograph disc a cam to key the Camden telegraphic identification signal KM, and adding a set of time switches. It then automatically transmitted a telegraphic identification signal on each minute, a telephonic identification on each intermediate half-minute, and a brief calibrating signal on each quarter-hour. The motor being synchronous, these signals were usable for checking clocks in aircraft. Telegraphic transmission, either by hand or by cam, was accomplished by energizing a relay to change the twin-triode bias, in the impulse generator of Figure 11, from the single impulse to the eight-impulse-per-cycle condition.

It will be seen from the above figures and description that an omnidirectional range transmitter is rather simpler than the pair of separate transmitters which make up a single present-day low-frequency "simultaneous" radio range. The omnidirectional range is, in effect, the voice channel alone of a simultaneous range, plus a pair of low-powered balance-modulator stages and an impulse keyer. The radio-frequency goniometer and course-bending artificial lines of the course-type range are not necessary with an omnidirectional range. It is, however, very convenient to be able to simulate their action, to a limited extent, to compensate for small differences between individual transmitters and installations and for slight variations which may occur from time to time in a given installation. This we did, without interrupting operation, by the simple set of modulation-frequency phase and amplitude controls described above.

[11] I—Sect. 3, C, pp. 74-76.

c—Monitoring

Instruments were provided to monitor separately the effect of each modulation-phase or amplitude control; these instruments as well as the control circuits are shown in Figure 12. For the north-south directive modulation and the calibrating modulation, simple center-zero dynamometer meters served to indicate phase quadrature with the east-west modulation, the desired operating condition, by the absence of deflection which is characteristic of such meters when their two windings are fed with currents 90 degrees apart in phase. Differently sensed deflections corresponded to phase differences more and less than 90 degrees. The meter circuit constants were chosen to make indications independent of small circuit variations and consequently highly reliable.

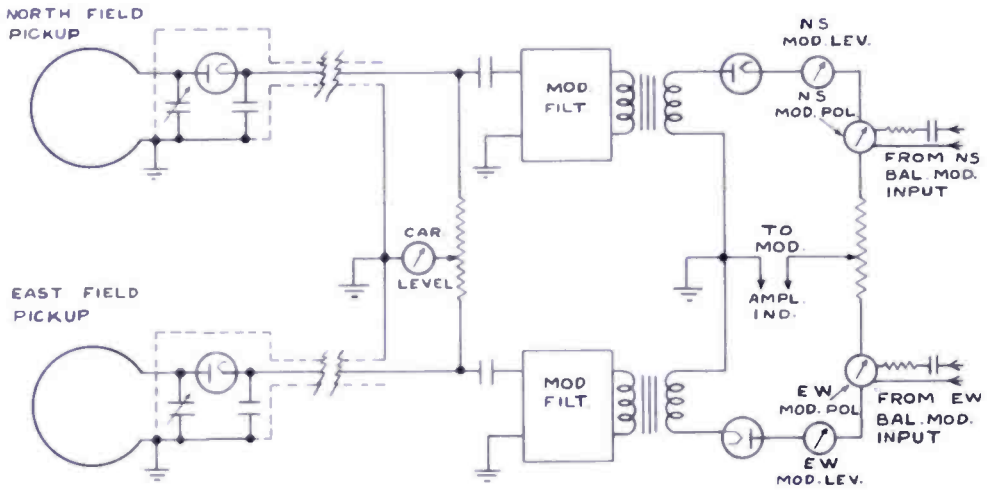


Fig. 13—Field monitoring circuits.

Because of the large ratio of root-mean-square to average value of the short, widely-spaced reference impulses, an attempt to indicate their phase by a dynamometer meter resulted only in overheating the meter without producing useful deflection sensitivity. However, a differential electronic circuit, of the type shown associated with the impulse control of Figure 12, gave good results. In proper operation, the reference-modulation impulses occurred just as the east-west modulating voltage passed through zero; that is, just when neither rectifier of the monitoring circuit passed any current from the sinusoidal supply. Thus, both rectifiers then passed the same impulse current, had the same average output and, therefore, produced no deflection of the differential meter. With improper impulse phasing, one rectifier output or the other predominated and the meter deflected in the corresponding sense.

The use of cathode-ray oscillographs to monitor radio-frequency phase equality of the outputs from the two tubes of each balanced modulator has been described above. It is also necessary to monitor the radio-frequency phase relations between directive and non-directive radiated fields.[12] Electronic wattmeter circuits located at and coupled to the antenna bases gave very sensitive and convenient indications—

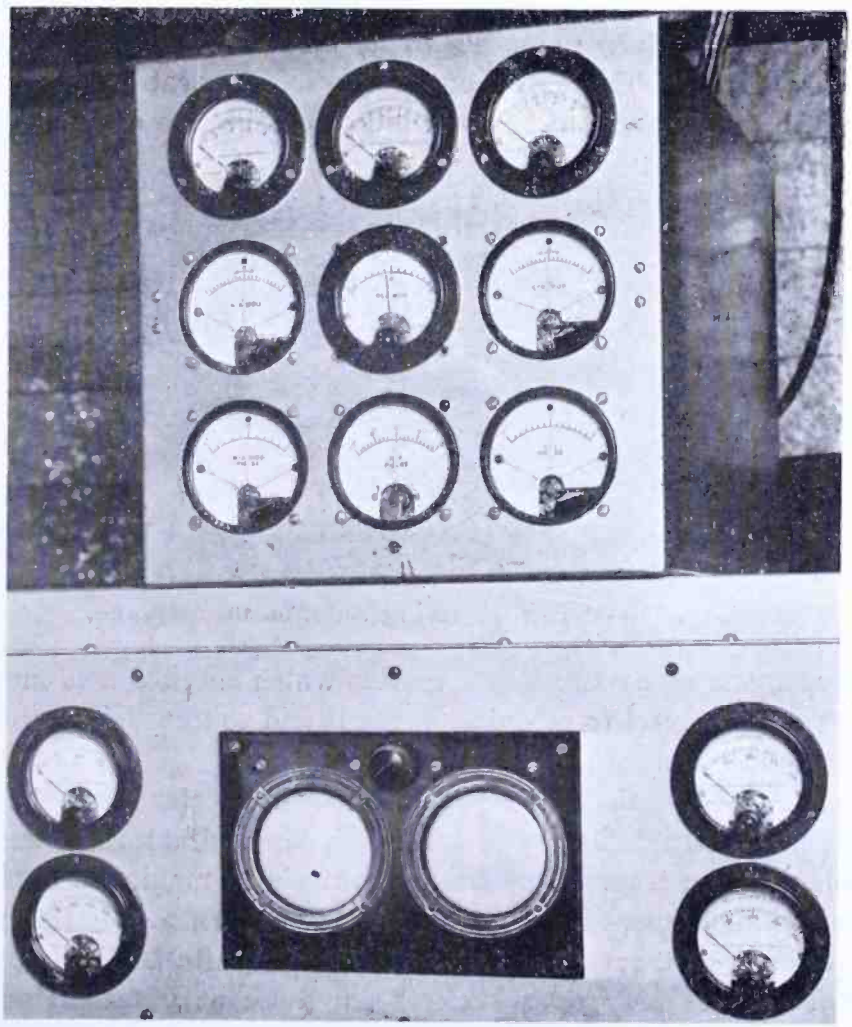


Fig. 14—Monitor unit.

of the zero-center sensed type—of departures from the desired condition. However, this system was found to be too dependent on indicator circuit conditions to have the degree of reliability necessary for satisfactory monitoring.

Antenna-current phase was finally monitored by the system shown in Figure 13. Small tuned-loop pickups with diode rectifiers were placed due north and east of the antenna system and distant by about

[12] I—Sect. 3, D, p. 77.

one wave-length, so that each was in the null of one directive signal and was subject only to the broadcast carrier and the other directive signal. Carefully grounded and shielded cables carried the low-frequency output from these units, visible at the edge of the ground disc in Figure 5, to the transmitter. There, the direct-current components from the two pickups were added, to give an indication of the non-directionally radiated carrier level, and the alternating components were fed (across differentially adjustable resistors to equalize the pickup unit sensitivities) to separate, modulation-frequency narrow-band-pass filters. The filter outputs were separately rectified and metered, by microammeters which indicated the two directive modula-

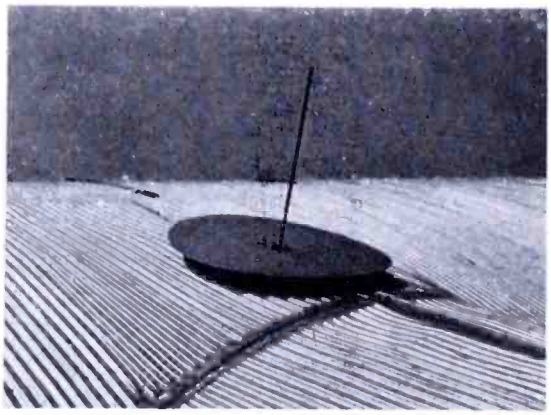


Fig. 15—U-H-F receiving antenna on airplane.

tion levels, and by dynamometer meters which checked field-modulation polarity with respect to modulating input and so guarded against gross radio-frequency phasing errors.

Modulation depth being greatest when carrier and suppressed-carrier modulated fields are in phase, phasing adjustments were made by means of the transmitter tuning controls to maximize the modulation-level indications while maintaining the correct modulation polarity. Close spacing of transmitting antennas and monitoring pickups caused the fields at the pickups to differ from those at great distances and consequently resulted in the radio-frequency phasing being incorrect by about 10 degrees. This effect, being the same for both directive fields, caused no direct bearing errors,[13] whereas greater pickup spacing resulted in excessive disturbance by nearby obstructions. Modulation-level meter sensitivities were equalized, during calibrating transmission when both pickups were known to receive the same modulation, by a differential adjustment of the meter-circuit resistances. These meters were also used in bringing calibrating and directive modulations to the same level. A differentially connected modulation-level meter, which

[13] I—Sect. 3, D, p. 77.

gave a central zero reading when the two systems were equalized in the calibrating operation, permitted bringing the two directive modulation depths accurately to equality by adjustment of the east-west amplitude-controlling Variac to give a zero reading during normal directive operation.

All monitoring instruments were grouped together, as shown in Figure 8 and more clearly in Figure 14. The cathode-ray oscillographs, the only monitoring instruments actually operating at radio frequency, were built into the transmitter proper. The four meters of Figure 12

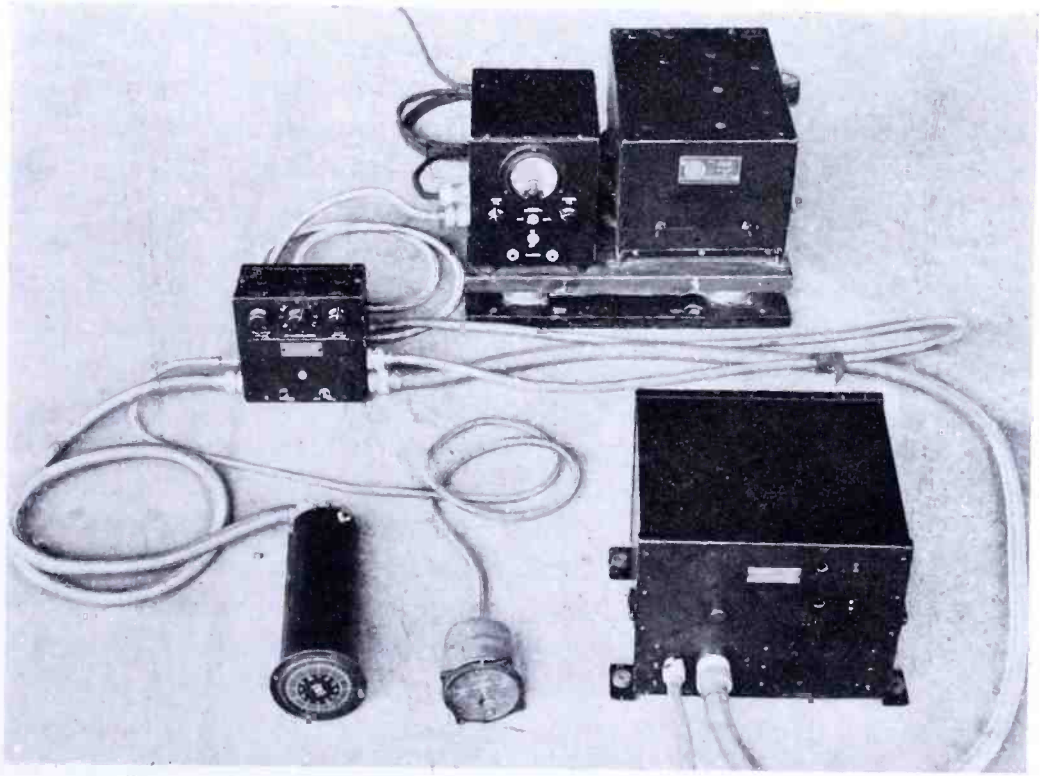


Fig. 16—Complete aircraft equipment (without antenna).

and the five of Figure 13, with their associated equipment, were mounted in the separate housing shown in Figure 14.

d—Indicators

For receiving, a vertical rod antenna, shown in Figure 15 mounted on the top of a transport airplane, was used. The circular ground plate at the antenna base was for the purpose of reducing interference from imperfect bonding of the airplane structure. A crystal-controlled superheterodyne, made for another purpose but having the flat automatic gain control and high audio fidelity which are characteristic of modern ultra-high-frequency aircraft receivers, was used. Audio output from a receiver phone-jack actuated the indicators, which also drew

tube plate power from the receiving dynamotor. Figure 16 shows the entire aircraft installation, less antenna. The large unit, which is the receiver, and the next smaller one, which contains the auxiliary apparatus for the indicators, were mounted out of the way. The three small units—a control box, a deviation-indicating meter, and a cathode-ray indicator in its cylindrical magnetic shield—constituted the cockpit installation. Both indicators were mounted in standard aircraft instrument holes.

The indicator system employed in the ultra-high-frequency work was rather more complex than the one previously used, in the ways enumerated below. A phase-shifting control was introduced into the audio line from receiver to indicators to avoid the need of mechanical rotation of the azimuth indicator scale for zero setting. To reduce the receiver power output required for indicator operation, an amplifier was incorporated into the filter for the sinusoidal signal component. This filtering amplifier was made of the selective-feedback type,[14] to avoid the weight of high-quality low-frequency reactors, and was equipped with an automatic gain control, to hold its output very constant and thereby relax the requirements on the receiver automatic gain control. A local impulse generator was incorporated into the impulse signal channel to improve the shape and equalize the size of the impulses applied to the indicator, as well as to compensate for the wide difference between the size of impulses received at sinusoidal modulation maximum (when receiver is north of transmitter) and at minimum (when south of transmitter).[15] This generator was a quiescent unsymmetrical multi-vibrator, operating through one cycle only when triggered by a received impulse. Finally, to improve the appearance of the cathode-ray indication by substituting a bright notch for the dark gap formerly used as the indicating mark on the bright circular cathode-ray trace, a simple radial-deflecting device for actuation by the impulses was added.[16]

As most of the circuits of the indicator, while not novel, are a little out of the run of ordinary radio circuits, they are shown in some detail in Figure 17, grouped according to function by the dashed lines. The filtering amplifier with its automatic gain control and the impulse shaper are strictly auxiliaries; they may be regarded simply as special, selective receiver-output stages. The vibrator-type high-voltage supply for the cathode-ray tube, having a vacuum-tube rectifier and a transformer winding to feed the heaters of the tubes with cathodes at high voltage, is not shown. Had the receiver not happened to have an over-

[14] "A New Type of Selective Circuit and Some Applications" by H. H. Scott, *Proc. I.R.E.*, Vol. 26, No. 2, p. 226, Feb. 1938.
[15] I—Figure 2.
[16] I—Sect. 2, B, p. 68.

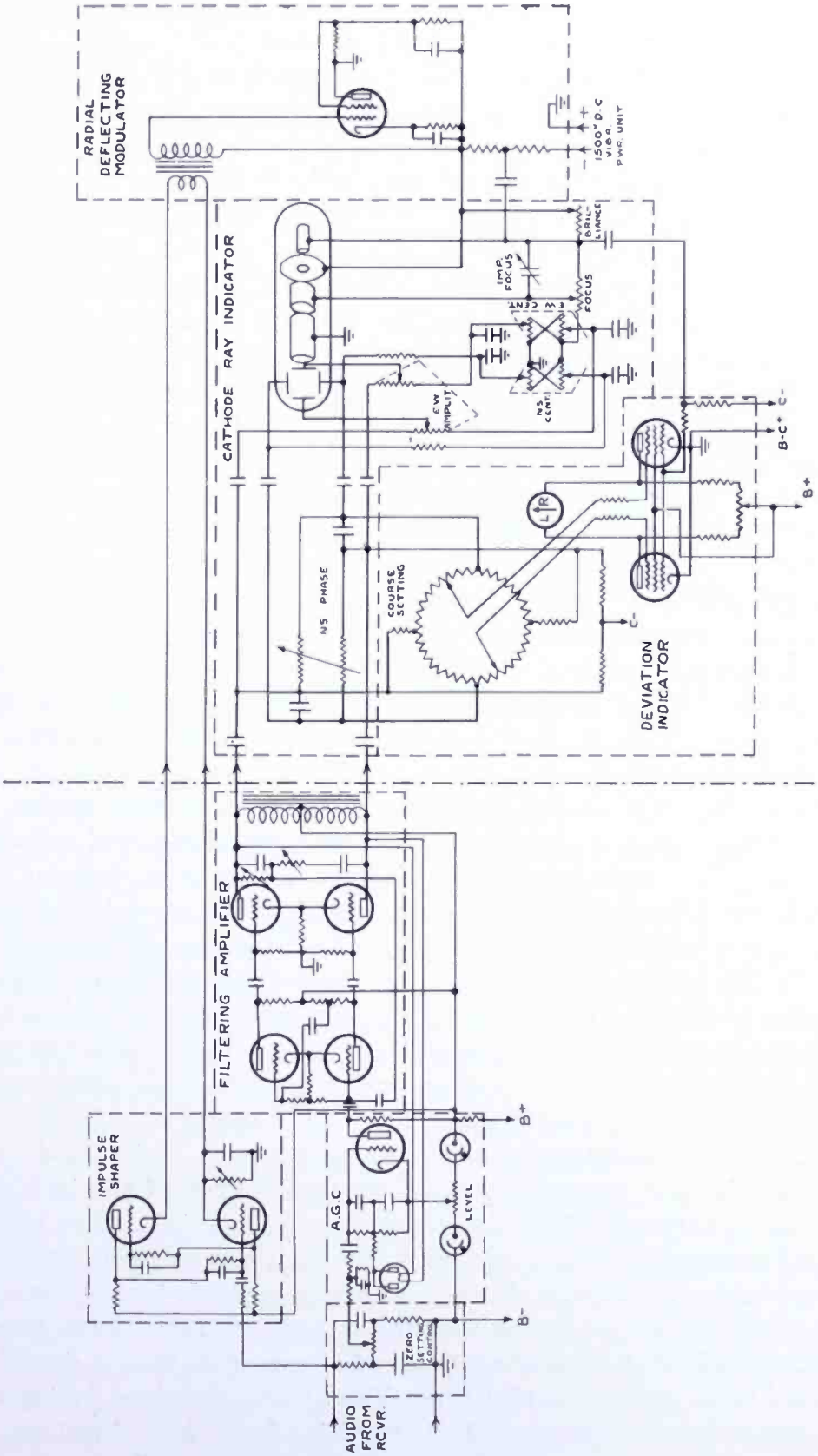


Fig. 17—Indicator circuit diagram.

size dynamotor, the vibrator supply would have been arranged to provide plate voltage for the above auxiliary tubes also.

To the right of the dot-dash line in Figure 17 appear the indicators proper. Placed ahead of the filter amplifier to avoid unproductive loading of the latter, the zero-setting phase control really belongs functionally in this part of the picture. The resistance-capacitance network between this control and the first amplifier grid served to keep impulse voltages out of the sinusoidal-current circuits and to compensate normal phase displacements in the receiver audio circuits and filter, so as to permit the control to work near the center of its ±15-degree range. Following the filter, a resistance-capacitance bridge was used to give the 90-degree phase shift necessary to produce a circular trace on the cathode-ray-tube screen; the four deflecting plates of the tube were connected to the four corners of this bridge network with a normal centering control for each pair of plates, and with sensitivity-equalizing potentiometers for one pair of plates. To the four corners of the bridge were also connected four equally spaced fixed taps on an ostensibly uniform, high-resistance carbon-composition ring. Two diametrically opposite, freely rotatable brushes bearing on this ring permitted picking off a voltage, balanced to ground and of reasonably constant amplitude, which could be continuously varied in phase throughout a 360-degree range. This control was used to set the deviation indicator to any desired course.

Operation of the radial deflecting system was directly similar to that of the familiar Heising constant-current modulator used in transmitters. The high-voltage supply fed the accelerating train of the cathode-ray tube and the plate-cathode circuit of a pentode, in parallel, through a common resistor. Normally cut off, the pentode had most of the time no effect on the operation of the cathode-ray tube. When a momentary positive voltage from the impulse shaper was applied to the pentode grid through a high-ratio isolating transformer, the pentode plate current in the common resistor dropped the accelerating voltage applied to the cathode-ray tube. This slowed the electrons in the cathode-ray beam and made them momentarily easier to deflect. The impulse thus appeared on the fluorescent screen of the indicator tube as an outward-pointing notch or jog on the bright circle of the normal cathode-ray trace. To compensate the dimming tendency of the extra radial spot velocity during the jog, a positive impulse voltage was applied to the control grid of the cathode-ray tube. This voltage was obtained by by-passing a portion of the common voltage-dropping resistor to the cathode of that tube. Focus was maintained during the jog by equalizing the time constants resulting from stray shunt capacitances, and by adjustment of a small trimming capacitor connected

between first and second anodes of the cathode-ray tube. Careful shielding was necessary to prevent the impulse voltages on the accelerating circuits of the cathode-ray tube from affecting the deflecting circuits.

Not the least of the features of the cathode-ray indicator system was the straightforwardness of the routine by which it was initially adjusted to give clear accurate indications at all parts of the azimuth scale. The indicator trace, before setting-up adjustment, was generally a defocused, decentered tilted ellipse. After rough preliminary adjustments to give proper pattern shape, size, focus, and brilliance were made, final adjustment for accurate indication was done as follows. With the range transmitting a calibrating signal with eight impulses per cycle, the zero-setting phase control was adjusted to give north and south indications in error by equal amounts and the east-west centering control was then set to give north and south readings of exactly zero and 180 degrees. Adjustment of the NS phase control to give equal errors at East and West was then followed by adjustment of the NS centering control to give east and west readings of exactly 90 and 270 degrees. Finally, the EW amplitude control was set to give the best average reading at the 45, 135, 225, and 315-degree points. Since there was very little interaction between these controls, adjustment with eight equally timed impulses per rotation cycle was easily made to bring the eight marks then visible to read exactly correct at the four cardinal points and correct on the average at the four primary intercardinal points of the scale. With commercial cathode-ray tubes the correctly adjusted pattern usually had a slightly elliptical shape, but because of the method of adjustment used, this did not seriously impair accuracy of indication at any part of the scale.

Zero-setting and deviation-indicator course-setting controls were the only ones, except for a power switch, brought out on the face of the control box (a third knob visible in Figure 16 actuated a manual volume control, later superseded by the filter-amplifier automatic gain control). The other controls affecting accuracy and appearance of indication were available under a dust cover on this unit. It will be seen from Figure 17 that all elements affecting readings were of a simple, stable sort and, except for the zero setter, were directly associated with the cathode-ray tube. The zero setter itself had to be used to compensate any phase-shift variations in the receiver and filter. While it was the custom to check and trim all adjustments frequently in the test work, it appeared that all the covered controls were very stable when not disturbed and that even the zero setting needed little attention during steady operation. The automatic calibrating transmissions at quarter-hour intervals gave a more than adequate zero check, and brief eight-mark periods forming part of these transmissions served to

show up and identify any inaccuracy which might develop in any of the other indicator adjustments.

To actuate the deviation indicator, sinusoidal voltage of freely variable phase, taken from the tapped-ring course-setting resistor described earlier, was applied in push-pull to the suppressor grids of two pentodes, as may also be seen from Figure 17. Impulsive voltage, taken from the accelerating train of the cathode-ray tube, was applied to the control grids of these tubes in parallel. The control grids were biased somewhat beyond cutoff and swung to zero grid; excess impulse voltage beyond that point was taken up by grid-current drop in a large series resistor. The suppressor grids were biased to the middle of their negative control range and were also fed through series resistors of high value. Their control of plate current was thus limited, on the one hand by cutoff and on the other by grid-current drop in the series resistors. This arrangement gave high phase sensitivity near the zero-point of the sinusoidal voltage, but avoided meter overloading for widely different phases.

A zero-center microammeter connected between the plates of the two pentodes, which had separate plate-load resistors, indicated the differential impulse current through these resistors, while a capacitor shunting the meter protected it from mechanical damage by the high peak currents. Because an impulse occurring at either of the two instants of zero sinusoidal voltage during each cycle would give equal plate currents and zero meter deflection, the sinusoid-phase control used for course setting was provided with a rough azimuth scale to avoid ambiguity. Because final course settings were made by bringing the meter to zero when the cathode-ray indicator showed the craft to be on the desired course, the accuracy of this scale was unimportant. Another result of this initial zero setting was that differing normal plate currents of the two pentodes were not directly harmful. However, to minimize the disturbing effect of noise impulses, it was found to be desirable to adjust the two load resistors to compensate for tube differences. This was easily done by using the eight-impulse-per-cycle transmission, the adjustment then being made for a zero meter deflection which was independent of sinusoid phase.

Weight is, of course, a major consideration in all aircraft equipment and the weight of our experimental apparatus was probably greater than it should be. An ultra-high-frequency receiver will be necessary equipment in any case; its weight need not be charged to the omnidirectional range. The weight of the deviation indicator with its cable was 1.8 pounds and is probably irreducible, as is the 3.3-pound weight of the cathode-ray indicator with its 9-wire cable from the control box. The control box itself weighed 3.8 pounds and, while some reduction

might be possible, it is probable that the total weight of the cockpit installation of control box, two indicators and their cables cannot be much less than 8 pounds. The auxiliary unit of the system, with its multiconductor cable to the control box, weighed 23.2 pounds. Here, design changes might make an appreciable reduction. In particular, if the auxiliary equipment were built integrally with a receiver, the added weight could be made much less than the above value. The degree of refinement and complexity justifiable in the indicator system can only be determined by experience in service; it is difficult to see how it can be much simpler than the early form described in Section 2*A*, *d*.

MEASUREMENT OF THE SLOPE AND DURATION OF TELEVISION SYNCHRONIZING IMPULSES*

By

R. A. MONFORT AND F. J. SOMERS

Engineering Department, National Broadcasting Company, New York

Summary—Satisfactory operation of television receivers in the field requires that the waveform of the transmitted synchronizing signals be held to narrow tolerances. It is therefore essential that suitable measuring equipment and techniques be available at the transmitter so that synchronizing waveshapes can be accurately and rapidly checked. This paper describes several measurement methods which have been found to be satisfactory under practical operating conditions.

1—GENERAL

THE operation of a television broadcast system on a regular public service schedule over a period of years has shown the need for various auxiliary measuring and testing equipment. Suitable equipment and methods for insuring that the synchronizing signals conform at all times to the prescribed standards are important factors in maintaining a high-quality service.

The tolerances prescribed for the steepness of wavefronts as well as the duration times of the various impulses that go to make up the composite synchronizing signal require measuring equipment capable of determining fractional microsecond time intervals. A further requirement, equally as important as high precision, is that the measuring equipment should be simple, and easy to use under actual operating conditions.

The purpose of this paper is to describe apparatus and methods developed by the National Broadcasting Company for synchronizing-signal measurements. These are the outgrowth of experience gained in our television broadcast operation and have been found to be a satisfactory solution to the measurement problem.

2—CHARACTER OF MEASUREMENTS

Figure 1, which is a copy of Drawing II of the Federal Communications Commission Standards†, will illustrate the character of the necessary measurements. These fall into two general classes: Slopes of leading or trailing edges, and the duration times of the various impulses. It will be noted that these measurements are to be made at the 10 per cent and 90 per cent amplitude points of the wave. This is to avoid any ambiguity which might arise due to the rounded corners on the waves resulting from finite bandwidth limitations.

* Presented at Pacific Coast I.R.E. Convention, Seattle, Aug. 20, 1941.

† Docket No. 5806, issued May 3, 1941. "Television Synchronizing Waveform for Amplitude Modulation."

An idea of the precision required may be gained from the fact that under the present standards of 525 lines, 60 fields, interlaced, the maximum time allowed for the transition from the 10 per cent to the 90 per cent amplitude point of the horizontal synchronizing wave is only 0.25 microsecond. The tolerance allowed for variations in the width or duration of horizontal synchronizing impulse is also a small value, being limited to 0.64 microsecond. Similar close tolerances have been specified for other parts of the wave.

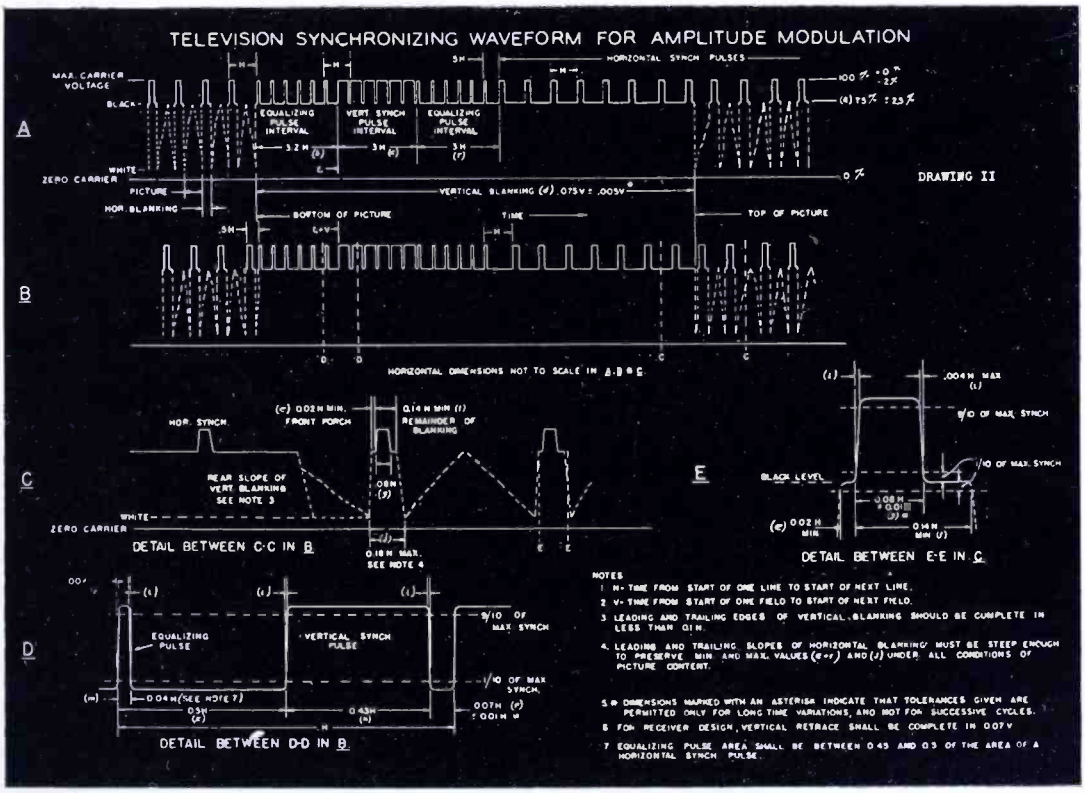


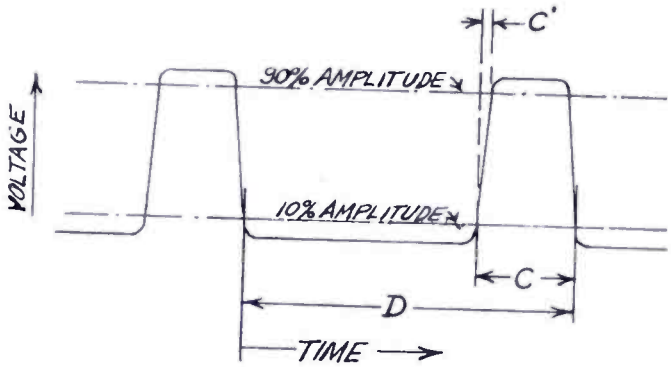
Fig. 1—Copy of standards.

Accurate measurements of this kind require the use of an indicating device having negligible inertia and a wide band frequency response. A cathode-ray oscilloscope or its equivalent is therefore generally employed. It can be shown (see Appendix I) that faithful reproduction of the wave of Figure 1E which has equal build-up and decay times of 0.5 per cent of a scanning cycle requires that the oscilloscope signal amplifier be responsive up to the 200th harmonic of the scanning-line frequency. Translated into terms of the present television standards, this means that the oscilloscope signal amplifier should have linear phase and flat amplitude response from below 60 cycles up to at least 3.15 megacycles. In addition, it is also required that the oscilloscope time-base amplifiers be free from compression or

non-linearity over the range of sweep amplitudes used in making measurements. Finally, in order to observe the finer details of the wave, a cathode-ray tube having a sharply focused spot and a large-diameter screen should be utilized. A tube having a screen diameter of at least five inches should be employed, but a nine-inch tube is preferable for accurate work.

3—Inadequacy of Sawtooth Sweep

As might be expected, ordinary methods of measurement fail when applied to the specialized waveforms of Figure 1. A common example of this is illustrated by Figure 2 which shows the measurement of



$$\text{Pulse Width} = W = \frac{C}{D} \times 100\%$$

$$\text{Slope of Leading Edge} = S = \frac{C'}{D} \times 100\%$$

Fig. 2—Sawtooth sweep measurement.

impulse widths and slopes using the usual oscilloscope sawtooth wave time base. The oscillogram shows one complete cycle and part of another. This is necessary since the finite retrace time of the sawtooth wave makes it impossible to see a single complete wave on the forward trace unless the sawtooth wave is a submultiple frequency of the wave being measured. In this procedure the dimensions C, C', and D would ordinarily be read on the oscilloscope screen with a transparent millimeter scale and the widths and slopes in percentage values would be

$$\text{Width} = W = \frac{C}{D} \times 100 \text{ per cent}$$

$$\text{Slope} = S = \frac{C'}{D} \times 100 \text{ per cent}$$

On a five-inch oscilloscope screen, the necessity of showing at least

two cycles means that D will be less than 100 mm. At the same time, the probable error in scale reading, taking account of the width of the oscilloscope trace may easily amount to ±1 mm. It can readily be seen that serious errors will occur if attempts are made to measure the slope of the leading edge of the horizontal synchronizing wave with this setup, since the value C', being less than 0.4 mm, cannot be accurately read with the scale. As a matter of fact, errors due to unsteadiness of the image and non-linearity of the sawtooth wave can easily add up to

$$\text{SIN } \frac{\theta}{2} = \frac{\frac{C}{2}}{\frac{D}{2}} = \frac{C}{D}$$

$$\% \text{ WIDTH} = \frac{\text{ARC}}{\text{CIRCUMFERENCE}} \times 100\% = \frac{\theta}{360^\circ} \times 100\% = \frac{\frac{\theta}{2}}{180^\circ} \times 100\%$$

$$\% \text{ WIDTH} = \frac{\text{SIN}^{-1} \frac{C}{D}}{1.8}$$

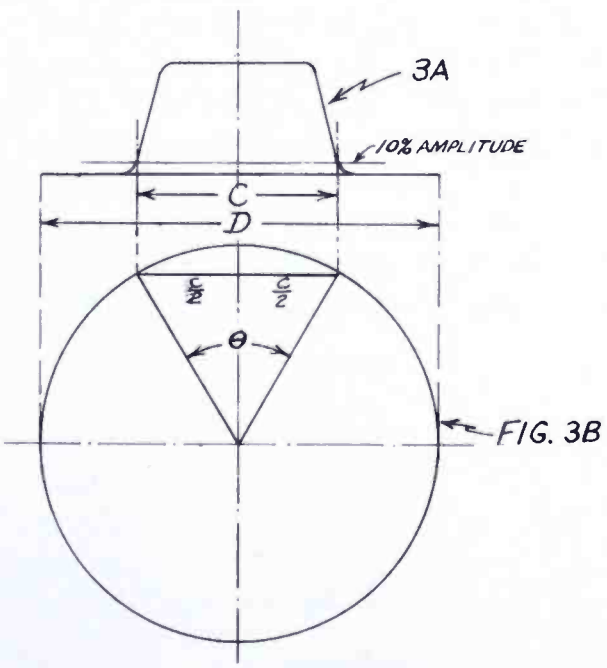


Fig. 3—Sine-wave measurement of widths.

more than 1 per cent of the scanning cycle, making the error greater than the value to be measured.

In view of these errors, it has been necessary to develop more accurate methods, keeping in mind the requirement that the equipment should be simple, and easy to use under practical operating conditions.

4—Sine-Wave Horizontal Sweep

A method of measurement which is at once simple and accurate within reasonable limits, uses an oscilloscope with sine-wave horizontal

deflection. The frequency of the sine wave used is either the same or a harmonic frequency of the impulse being measured. The pattern on the oscilloscope screen is stationary since the sine wave is derived either from the synchronizing generator or directly from the synchronizing impulse itself. A convenient phase-shifting device is provided so that the part of the wave of greatest interest can be moved to the center of the screen where it appears expanded horizontally. This feature of sinusoidal sweep tends to increase the accuracy obtained in measurement of impulse widths and slopes, since the dimensions to be measured are larger for a given width of horizontal time axis than if sawtooth waves are used. Also, since only one cycle appears on the screen, the effective horizontal time base can be wider than with sawtooth sweep, where, as already pointed out, more than one cycle must be included in the horizontal direction.

The procedure in making measurements with sine-wave horizontal deflection is illustrated by Figure 3A, which shows a train of horizontal driving impulses as they appear on the oscilloscope using 15,750-cycle sine-wave sweep. This type of impulse is utilized for energizing the horizontal sawtooth-wave generators of television cameras and video monitors in the television plant, and normally has a time duration of 6 per cent of a scanning cycle. The impulse is moved to the center of the screen by means of the phase-shifting device already mentioned and its width or duration can then be found by means of the following equations, derived with the aid of Figure 3B.

$$\text{Sin}\frac{\theta}{2} = \frac{\frac{C}{2}}{\frac{D}{2}} = \frac{C}{D}$$

$$\text{Per cent width} = \frac{\text{length of arc}}{\text{circumference}} \times 100 = \frac{\theta}{360^\circ} \times 100$$

$$\text{Per cent width} = \frac{\frac{\theta}{2}}{180^\circ} \times 100$$

$$\text{Per cent width} = \frac{\text{Sin}^{-1}\left(\frac{C}{D}\right)}{1.8}, \text{ where the angle is expressed in degrees.}$$

By measuring C and D in mm, with a transparent scale attached to the oscilloscope screen and applying the above equations, the duration

of the impulse can readily be found. Calculations may be eliminated by the use of the nomographic chart of Figure 4, which gives the duration of the impulse directly in per cent of a scanning cycle. Figure 5 is a photograph of the composite television signal as it appears on the oscilloscope for a 15,750-cycle sine-wave sweep. The lower-case letters refer to the impulses as shown in Figure 1. The width of any of the impulses indicated may be determined by shifting them to the center of the screen, measuring C and D, and referring to the chart. The width

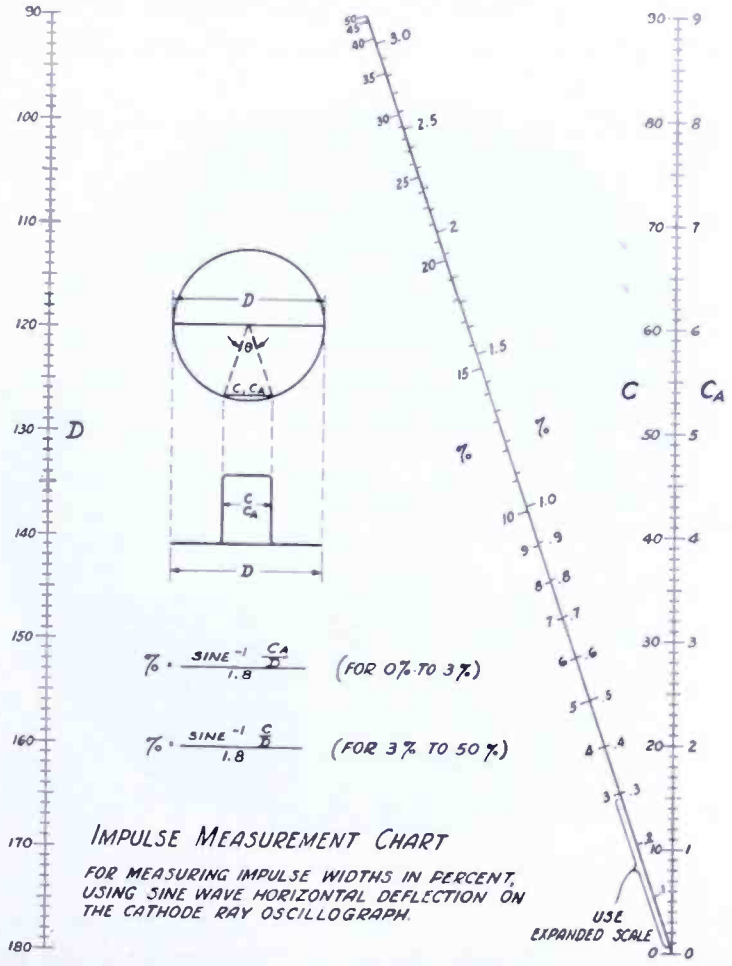


Fig. 4—Nomographic chart.

of vertical blanking may be found as shown in Figure 6, which is a photograph of the composite signal viewed on 60-cycle sine-wave sweep. The same chart may be used for this measurement, the result being expressed in per cent of one field-frequency scanning cycle.

One of the most serious sources of error in making measurements of this kind is in taking the millimeter scale readings. It can readily be shown, however, that errors in scale readings are minimized when sine-wave sweep is used, as compared to linear or sawtooth sweep. To give a concrete example, suppose we are measuring the width of the

horizontal synchronizing impulse which has a specified duration of 8 per cent ± 1 per cent of a scanning cycle and that the D dimension is 100 mm. For the sawtooth sweep, the value of C will be 8 mm. However, the error in reading the scale can easily be as much as ±1 mm. Consequently, reading D 1 mm too small and C 1 mm too large, gives an answer of 9.1 per cent which represents an error greater than the allowable tolerance for the width of the wave. Under the same conditions, using sine-wave sweep with D also 100 mm, reading D 1 mm too small and C 1 mm too large, gives an answer of 8.42 per cent which is within the prescribed tolerance. Another source of error is the presence of

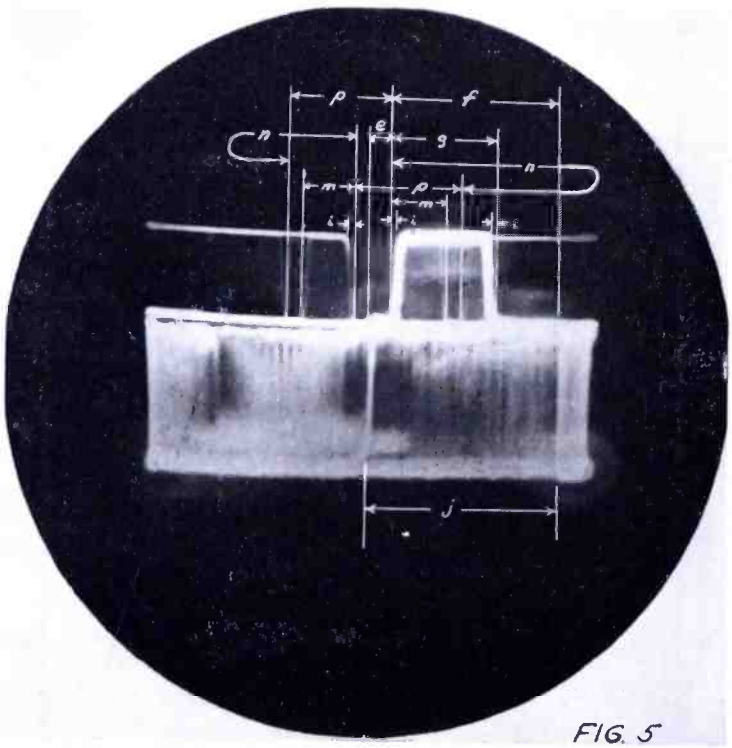


Fig. 5—Photo of composite signal on 15,750-cycle sweep.

harmonics in the sine wave. While a rigorous analysis has not been made for all possible harmonic combinations, it may be stated that a 1 per cent 2nd harmonic content, depending on its phase, can produce a maximum error of 0.22 per cent of a scanning cycle in the measurement of the width of the horizontal synchronizing signal. This amount of error will probably not be exceeded if the arithmetic sum of all harmonics present is held down to 1 per cent.

Another source of error is non-linearity or compression in the horizontal sweep amplifier of the oscilloscope. This should be investigated for the particular oscilloscope used, and measurements should be confined to the range of amplitudes within which the sweep is linear.

If reasonable care is taken to minimize the errors referred to, the sine-wave method gives impulse widths to an absolute accuracy of $\pm\frac{1}{2}$ per cent of a scanning cycle. While higher accuracy is desirable, the results are at least practical, since they are within the tolerance prescribed for these waves.

Increased accuracy in the measurement of the slopes of leading and trailing edges can be obtained by using a sine wave which is a harmonic of the scanning-line frequency. This stretches out the *C* dimension so that more accurate determinations can be made. The appearance of the synchronizing signal when viewed with 157.5-kc sine-wave deflec-

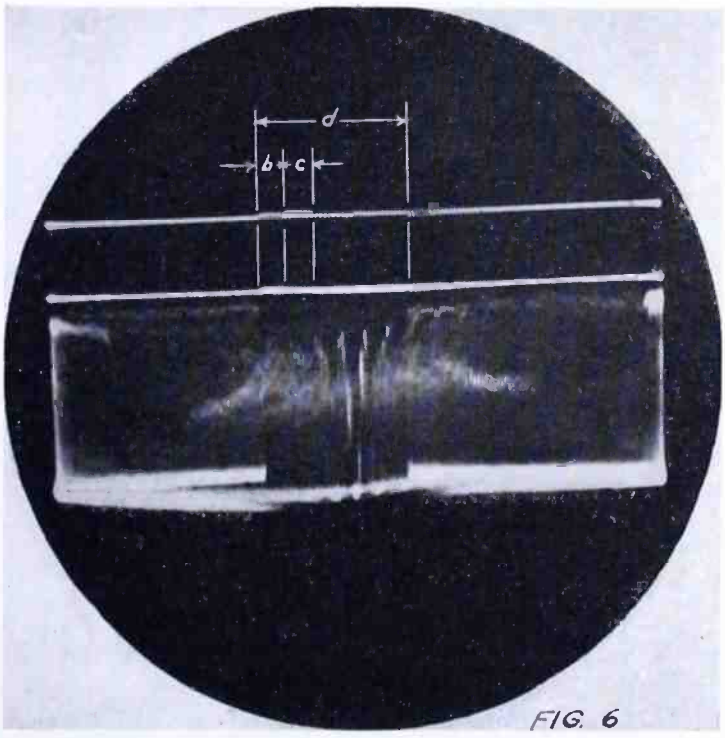


Fig. 6—Photo of composite signal on 60-cycle sweep.

tion is shown on Figure 7. Since this is the 10th harmonic, the *C* dimension is stretched out by a factor of ten. The same equations still hold and the same chart can be used, the answer obtained being divided by 10 to get the slope in per cent of a scanning cycle. Observing the same precautions as outlined for measurements with fundamental-frequency sine waves, impulse slopes may readily be measured by this method to an accuracy of ± 0.05 per cent of a scanning cycle, which is well within the prescribed tolerances.

5—Direct-Reading Sine-Wave Method

Increased accuracy can also be obtained with the modified measurement procedure shown in Figure 8, which eliminates the need for

charts or scales and gives impulse-width and slope values read directly from a calibrated dial. In this arrangement, the wave is brought into approximate position with respect to the cross hairs by means of the coarse phase adjustment. A fine phase adjustment provided with a dial calibrated in per cent of a scanning cycle is then used to shift the wave back and forth with respect to a fixed hair line in the center of the oscilloscope screen. The difference in dial readings for the 90 per cent amplitude point of the wave at position "A" and the 10 per cent amplitude point at position "B" gives the slope of the leading edge of the impulse directly in per cent of a scanning cycle. The slope of the

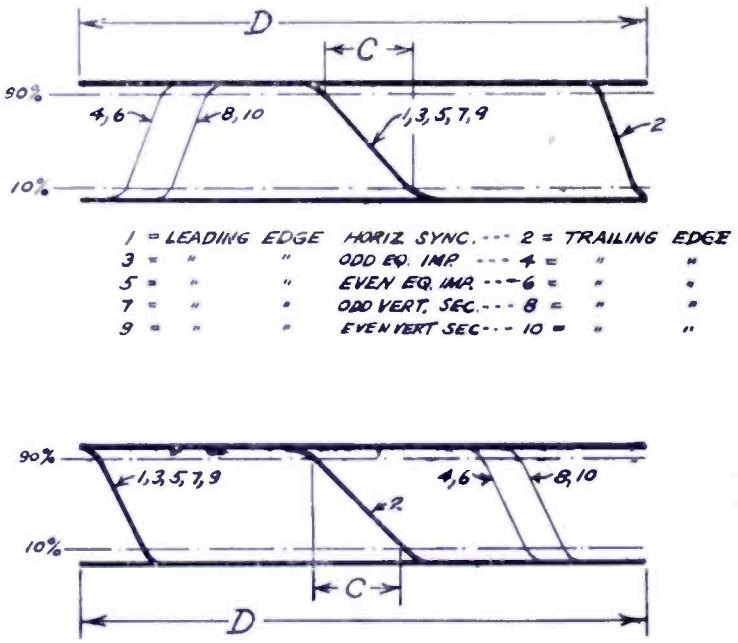


Fig. 7—Synchronizing signal viewed on 157.5-kc sweep.

trailing edge can be measured in the same way. Similarly, the duration of the impulse can be found by taking two readings at the intersection of the 10 per cent amplitude line and the vertical index mark. The phase-shifting arrangement shown is a modification of that described by Hartshorn[1] and gives a constant output voltage within the limits of voltage regulation of the amplifier feeding it. Over the small range required of the calibrated phase shifter, R and C can be chosen so that there is negligible change of amplitude with change of phase. Tests show that the results obtained with this arrangement are independent of moderate amounts of compression in the horizontal deflection amplifier of the oscilloscope, since even under these conditions the sweep will be linear in the central portion of the screen where observations are

made. This latter feature plus the elimination of scale-reading errors makes for increased accuracy.

6—Application to Time-Delay Measurements

An interesting application of the arrangement of Figure 8 is its use in determining the time delay of networks or lines carrying synchronizing impulses. The wave at the input to the network is lined up with the cross hairs on the oscilloscope, the setting of the fine phase

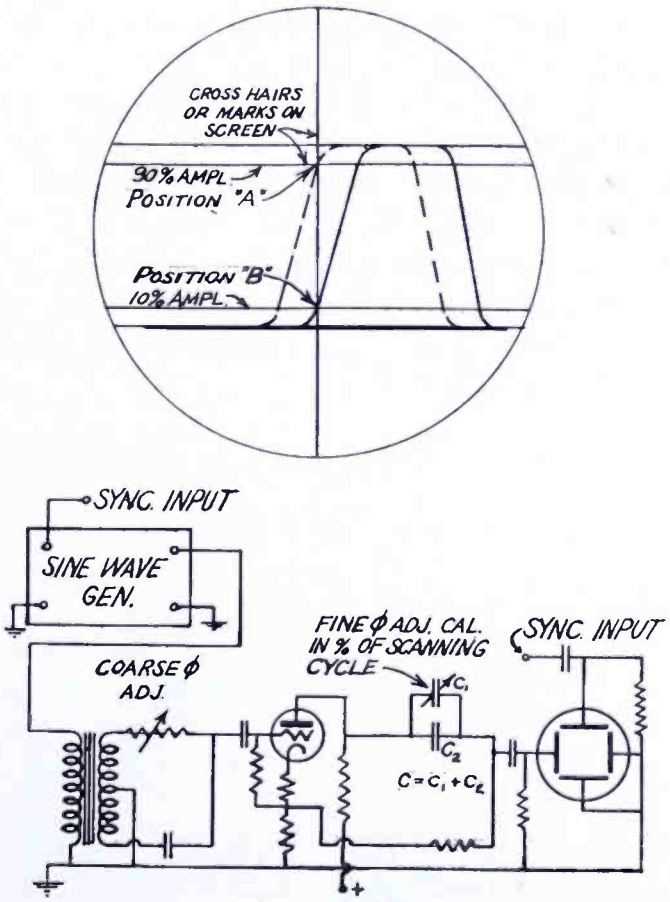


Fig. 8—Direct reading sine-wave method.

adjustment being noted. The oscilloscope vertical plates are then connected to the output of the network and the phase shifted to bring the wave back to its original position. Knowing the frequency of the sine wave and the phase shift in per cent of a cycle, the time delay in microseconds can easily be found. This same kind of delay measurement can, of course, also be carried out by the arrangement previously outlined for impulse width measurements using a millimeter scale and the chart of Figure 4.

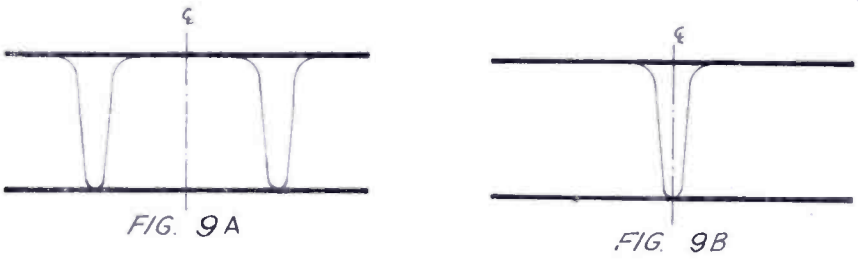


Fig. 9—Interlace check.

7—Use of Sine Wave in Checking Accuracy of Interlace

Line-frequency sine-wave sweep has also been found useful as a means of testing accuracy of interlace. For example, if 60-cycle driving impulses from the synchronizing generator are applied to the vertical plates of the oscilloscope and 15,750-cycle sine wave is applied to the horizontal plates, the pattern shown in Figure 9A will be obtained. It should be possible by shifting phase of the sine wave to cause these pairs of lines to cross each other in the center of the horizontal trace as shown in Figure 9B. If the lines cross anywhere else except the center, pairing of the scanning lines on the television picture (lack of interlace) will be indicated.

8—15,750-Cycle Sine-Wave Unit

The circuit diagram of the unit for generating the 15,750-cycle

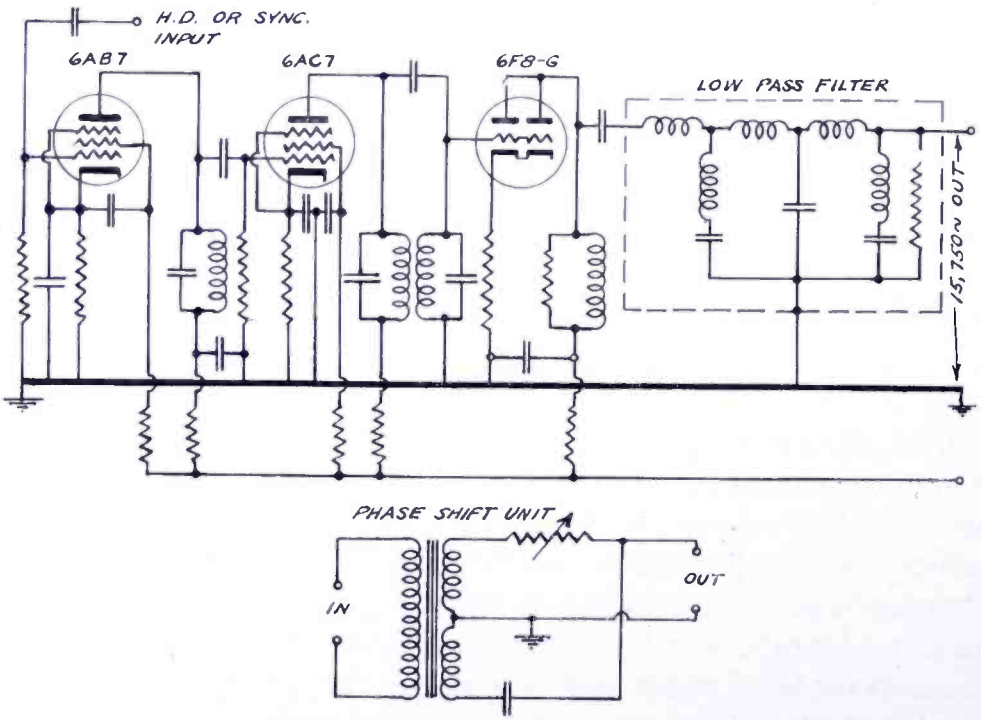


Fig. 10—Circuit of 15,750-cycle sine-wave unit.

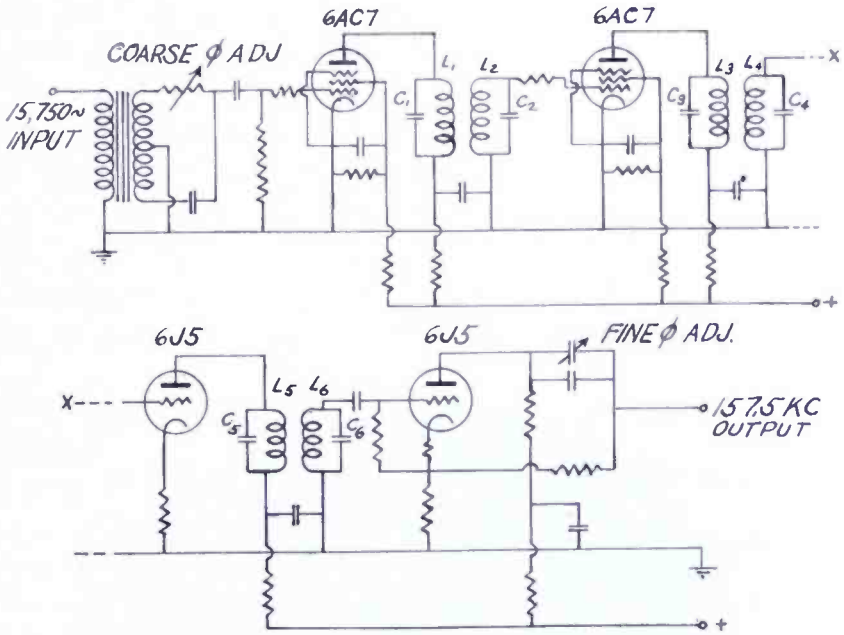


Fig. 11—Circuit of 157.5-kc sine-wave unit.

sine wave is shown in Figure 10. This consists essentially of an amplifier tuned to the scanning line frequency, with a low pass filter in its output. The filter is made up of one constant-K low pass section, an M-derived intermediate section designed for an infinite attenuation point at the second harmonic (31,500 cycles) and a terminal half-

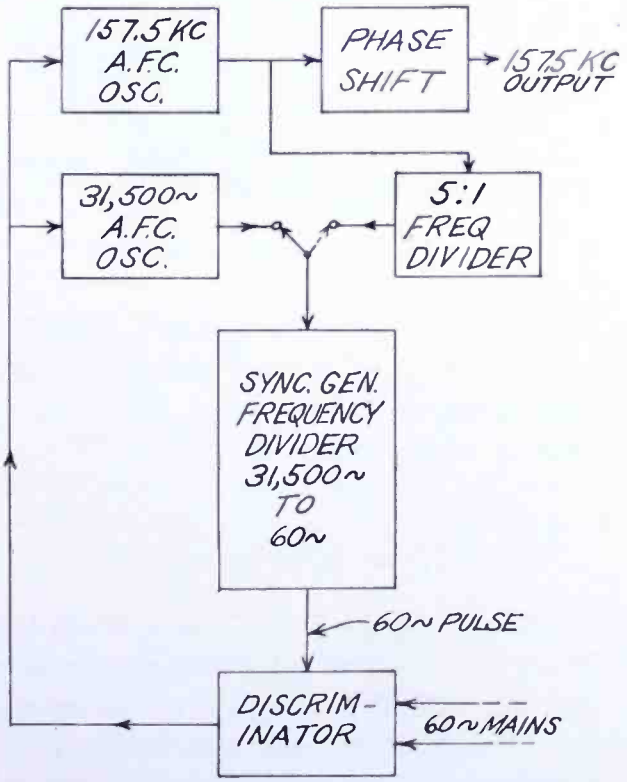


Fig. 12—Alternative method of obtaining 157.5-kc wave.

section designed for $M = 0.6$. The filter impedance is 1000 ohms and the cut-off frequency is 17,000 cycles. One type of variable phase-shifting network is also shown.

When measuring locally generated synchronizing pulses, this unit is actuated directly from the synchronizing signal generator. For signals originating at a remote point, as is the case with our mobile pick-up units, the sine-wave generator is energized from the synchronizing separator circuit of a local television receiver picking up the radiated signal. Measurement information can thus be made available

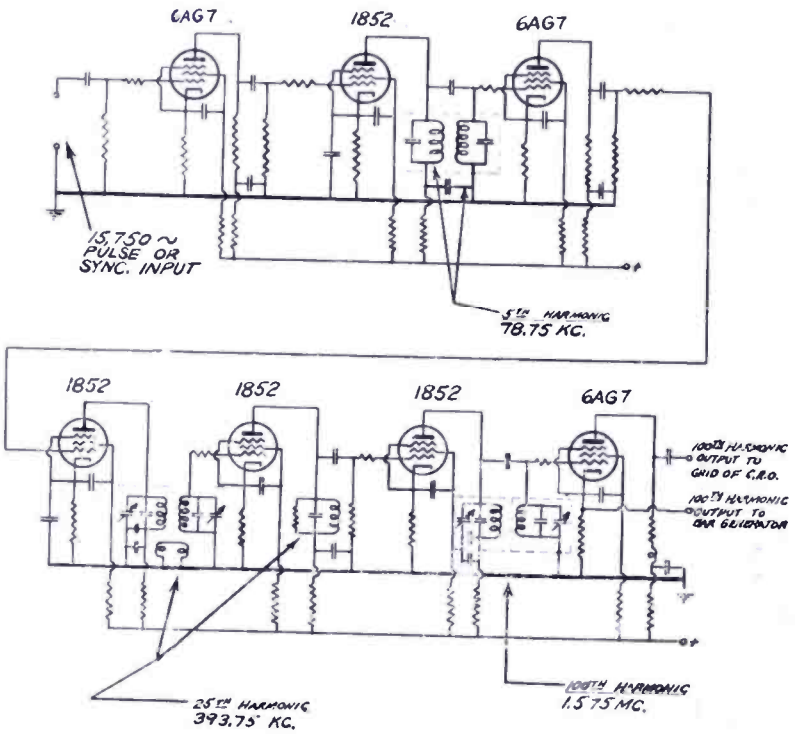


Fig. 13—Dot frequency generator.

to the mobile units via the order wire or radio link provided for communication purposes. This is a decided advantage, since space limitations make it impractical for the mobile units to carry a great deal of auxiliary measuring equipment.

9—157.5-KC SINE-WAVE UNIT

One method of generating the 157.5-kc sine wave is shown in Figure 11. This consists essentially of a harmonic generator and amplifier tuned to 157.5 kc and driven by the line-frequency sine-wave unit already described. To avoid serious amplitude modulation in the L_1C_1, L_2C_2 combination, coils having a Q in the neighborhood of 250

would be required. At this frequency such coils would be rather bulky. A practical answer was found in the use of small spool-wound coils, of 30 No. 36 litz, having a Q of 90. A signal of about 50 volts across L_2C_2 is fed to the grid of 6AC7/1852 which is operated with low screen voltage and zero bias to give limiter action. The plate of this tube feeding a second pair of tuned circuits delivers a signal having only a negligible amount of amplitude modulation. A slight frequency modulation results from the above process, but checks indicate it does not appreciably affect the accuracy of the measurements.

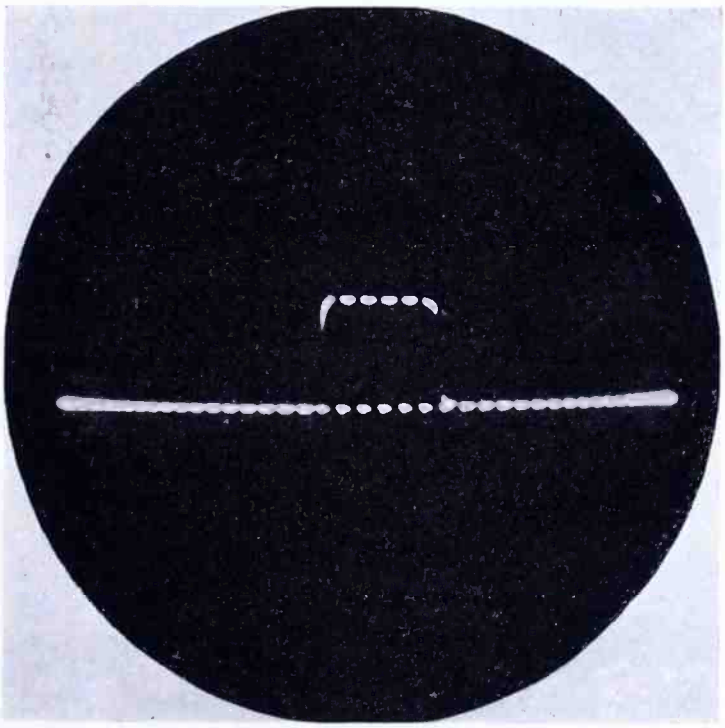


Fig. 14—H. D. modulated by dots.

An alternative method of obtaining the 157.5-kc sine wave is shown in block diagram form in Figure 12. This arrangement requires switching off the regular 31,500-cycle afc controlled oscillator of the synchronizing generator and locking in the 157.5-kc oscillator with the 60-cycle mains as shown. This method is suitable for test purposes only, as a switch of this kind could not be made during a program without causing a disturbance.

10—Dot Measurements

An alternative method of making the width and slope measurements already outlined is to modulate the grid of the oscilloscope with a high harmonic of the horizontal scanning frequency. Figure 13 shows the

circuit of a harmonic-generator unit for producing the 100th harmonic of 15,750 cycles. When this signal is applied to the oscilloscope grid, the wave is broken up into a series of dots as shown in Figure 14. In this case, one dot plus one space equals 1 per cent of a scanning cycle. Percentages can be read off directly, no calculations being necessary. An outstanding advantage of this system is that it gives results

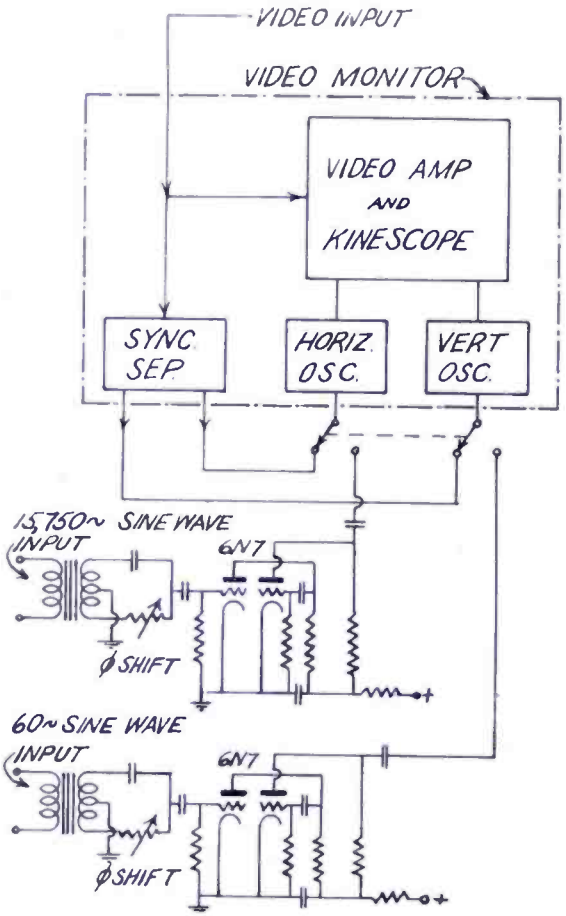


Fig. 15—Block diagram of pulse cross unit.

independent of sweep waveform or linearity. The wave shown in Figure 14 is a horizontal driving impulse having a ½ per cent slope on its leading and trailing edges. It is obvious that a larger number of dots, say 500 per scanning cycle, is needed for accurate measurement of steepness of wavefront of waves of this sort. At the present time, however, production of higher numbers of dots of sufficient stability requires removing the regular afc controlled 31,500-cycle oscillator from the synchronizing generator and switching in a stable fixed-frequency oscillator in place of it. If this is done, the 500th or higher harmonic of the line frequency will be sufficiently steady with respect to the

fundamental to allow very accurate width and slope determinations to be made.

11—Pulse-Cross Unit

While the equipment already described affords means of accurate measurement of various features of the synchronizing signal, it does not provide a check on such other items as the number of vertical

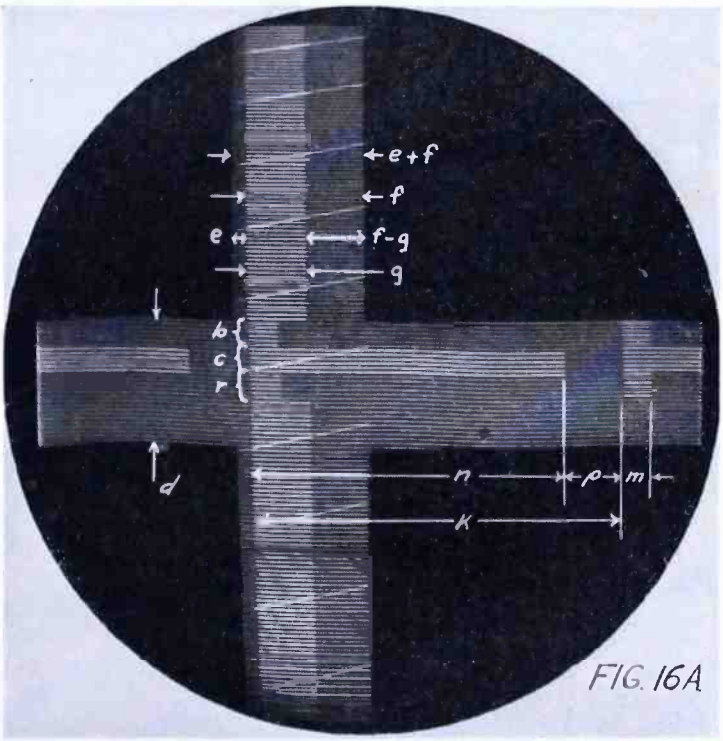


Fig. 16A—Pulse-cross correct.

synchronizing sections, number of equalizing pulses, etc. This information can readily be obtained by means of the arrangement shown in block diagram form in Figure 15. In this setup, which is similar to one described by Loughren and Bailey[2], sine waves of 60 cycles and 15,750 cycles are used to lock in the blocking oscillators provided for horizontal and vertical deflection of the video monitor. Phase shifters are arranged so that the horizontal and vertical blanking signals can be made to form a "pulse-cross" in the center of the kinescope screen. By applying the synchonizing signal to the kinescope grid in the proper polarity and increasing the vertical deflection, the pattern of Figure 16A is obtained. The values identified by the lower-case letters refer to the impulses shown on Figure 1. Several incorrect patterns resulting from faulty adjustment of the synchronizing generator are shown in Figures 16B, 16C, and 16D.

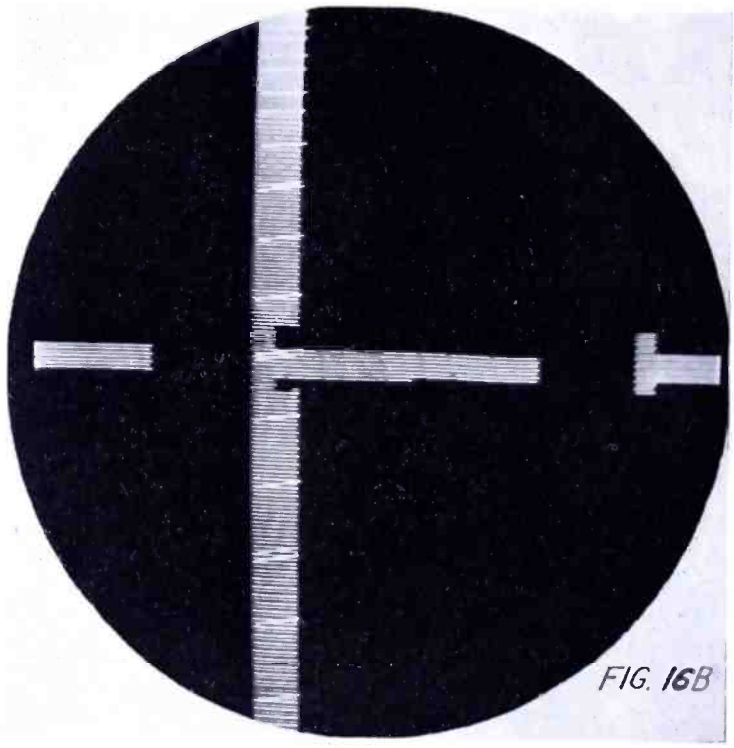


Fig. 16B—Pulse-cross incorrect.

While this pattern gives information as to the number and relative timing of the various impulses, it is not suitable for the accurate

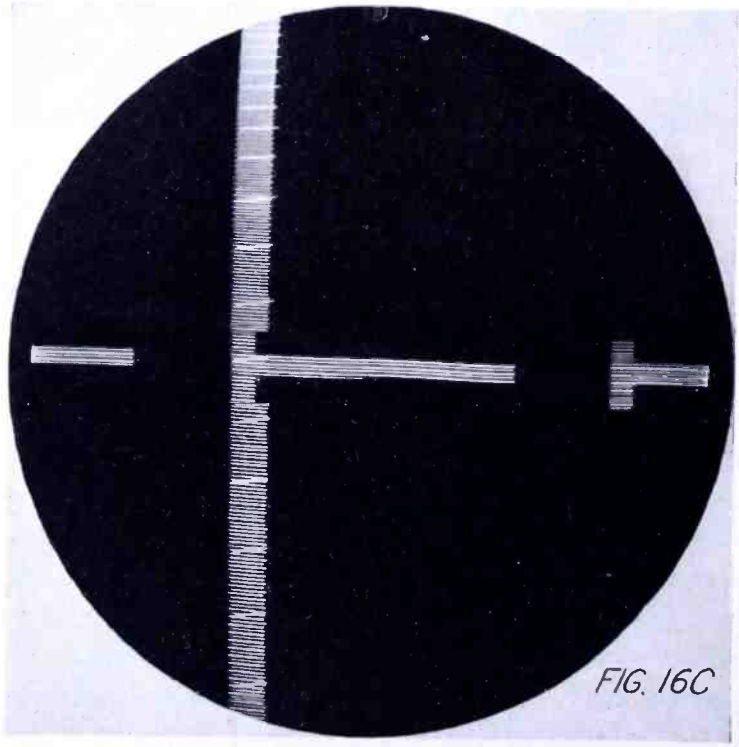


Fig. 16C—Pulse-cross incorrect.

determination of impulse widths and slopes since the 10 per cent and 90 per cent amplitude points specified in making these measurements cannot be found accurately on the pattern.

12—Conclusion

In conclusion it should be pointed out that, while the measurement methods presented here have been explained with respect to the specific synchronizing waveform of Figure 1, they are sufficiently generalized

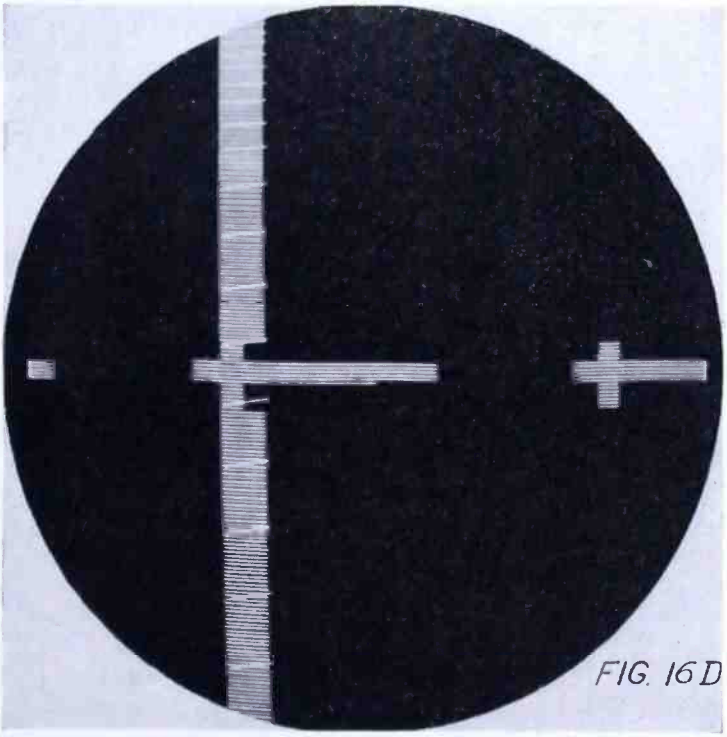


Fig. 16D—Pulse cross incorrect.

to be useful for other types of television synchronizing signals which have been proposed.

13—Acknowledgment

The authors wish to acknowledge the help of other members of the NBC Television Engineering Staff whose constructive suggestions and active cooperation contributed much to the success of this project.

Appendix I

It has been shown by Von Ardenne[3] that the general equation of a sawtooth wave having a finite retrace time T_α which is P per cent of the period T of the wave is

$$f(t) = \sum_{n=1}^{n=\infty} A_n \operatorname{Sin} n\omega t$$

$$\text{where } A_n = (-1)^{n-1}\left[\frac{Ak}{n^2\pi^2} \times \frac{100}{P} \times \text{Sin}\frac{(nP\pi)}{(100)}\right] \qquad (1)$$

and Ak is the amplitude of the sawtooth wave.

The amplitudes of the harmonics decrease more rapidly than $1/n$ and periodically become zero, the first zero position lying at the value of n which equals $100/P$. Inclusion of harmonics lying beyond the first zero contribute little to the shape of the sawtooth wave except to reduce the slight rounding at the peak of the wave. Therefore, the sawtooth wave

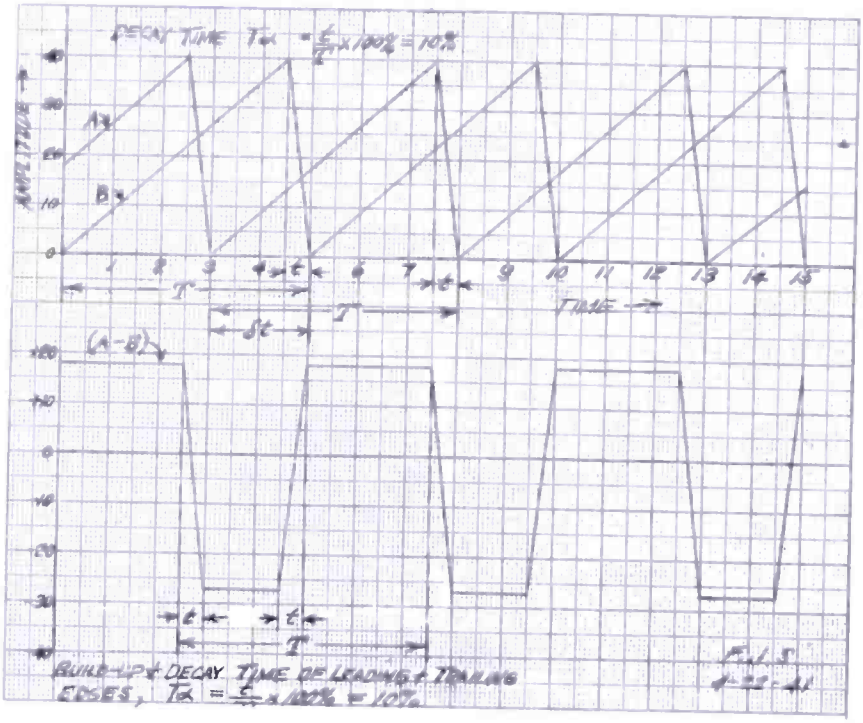


Fig. 17—Diagram for Appendix I.

is faithfully represented to a very close approximation by the summation of harmonics up to and including the first zero.

By a simple extension of this theory, it may be shown that the same conditions hold true for the number of harmonics necessary to represent a square-topped impulse type wave for which the build-up and decay time T_α are both equal to P per cent of the period T. In Figure 17 is represented the plot of two sawtooth waves of equal periods and times of decay, the wave B being delayed a time δ seconds behind the wave A. As shown in the diagram, subtraction of wave B

from wave A yields an impulse-type wave having a period T, the build-up and decay times, t, both being equal to P per cent of the period. The general equation of the resultant wave may be expressed as given in Equation (2), the value of A_n being identical with that above.

$$f(t) = \sum_{n=1}^{n=\infty} A_n[\text{Sin } n\omega t - \text{Sin } n\omega t(1+\delta)] \qquad (2)$$

The new wave has the same period as the sawtooth waves, its build-up and decay times T_α are both equal to the decay time of the original, and it also contains harmonics of identical frequency, and comparable amplitudes. It, therefore, follows that the impulse-type wave will also be represented to a close approximation by the inclusion of harmonic components up to and including the value of n equal to $100/P$.

REFERENCES

[1] L. Hartshorn, *Proceedings of the Physical Society* (London), Vol. 49, Part 2, March 1937.

[2] "Special Oscilloscope Tests for Television Waveforms," by A. V. Loughren and W. F. Bailey, presented at I.R.E. Rochester Fall Meeting, November 13, 1940.

[3] "Frequency Spectrum of Sawtooth Waves," M. Von Ardenne, *Television and Short Wave World*, January 1938.

OUR CONTRIBUTORS

ALLEN A. BARCO, a native of St. Louis, Missouri, received his degree of B.S. in Electrical Engineering from Washington University in 1937. Following his graduation he joined the engineering staff of the Radio Corporation of America License Laboratory, where he continues to be engaged. Mr. Barco is an associate member of Sigma Xi.

EDWARD W. HEROLD received a B.Sc. degree at the University of Virginia in 1930. He was employed in the research section of the Bell Telephone Laboratories from 1924 to 1927. During 1927 and the summers of 1928 and 1929, he worked in the engineering department of E. T. Cunningham, Inc. Since 1930, Mr. Herold has been in the Research and Engineering Department of the RCA Manufacturing Company at Harrison, N. J. He is a member of the Institute of Radio Engineers.

CHARLES N. KIMBALL received a B.E.E. degree from Northeastern University in 1931. From Harvard University he received his M.S. degree in 1932 and his Sc. D. degree in 1934. He spent two years with the National Union Radio Corporation and since 1936 has been connected with the RCA License Laboratory. Dr. Kimball is an associate member of the Institute of Radio Engineers.

DAVID G. C. LUCK was born July 1906 in Whittier, California. Bachelor of Science Mass. Inst. of Tech. 1927; Ph.D. 1932. Swope Fellow in Physics 1927-28, Malcolm Cotton Brown Fellow 1928-29. Assistant, Department of Physics, M.I.T., 1929-32. Research Division, RCA-Victor Co. 1932-35, Victor Division of RCA Manufacturing Co. 1935. Member, American Physical Society.

RAYMOND A. MONFORT attended the University of Kansas, and later, New York University. From broadcast station operator in Kansas City, he went to work for Western Electric in Kearny, New Jersey as Equipment Engineer. In 1932 he joined the National Broadcasting Company. Since 1935 he has been identified with NBC Television, and is now Television Maintenance Supervisor in the New York Studios. He is an associate member of the I.R.E.

FREDERICK H. NICOLL received a B.Sc. degree in Physics from Saskatchewan University, Canada, in 1929 and an M.Sc. in 1931. He held an 1851 Exhibition Scholarship to Cambridge University for three years research and received a Ph.D. from that university in 1934. He was a research physicist with Electric and Musical Industries, Ltd., in London from 1934 to 1939. Since that time he has been with RCA Victor Division of the RCA Manufacturing Company, Inc., as a research engineer. Dr. Nicoll is an Associate Member of the Institute of Radio Engineers.

GEORGE M. NIXON attended Pratt Institute and New York University. Since 1928 he has been engaged in general development work in the Development Group of the National Broadcasting Company. Mr. Nixon is a member of the Acoustical Society of America and the American Institute of Electrical Engineers.

DWIGHT O. NORTH received his B.S. degree from Wesleyan University in 1930 and his Ph.D. degree from the California Institute of Technology in 1933. Since 1934, Dr. North has been with the Research and Engineering Department of the RCA Manufacturing Company at Harrison, N. J., engaged principally in research studies of tube and circuit noise. He is a member of The Institute of Radio Engineers and a member of the American Physical Society.

STUART W. SEELEY received his B.Sc. Degree in Electrical Engineering from Michigan State College in 1925. He was an amateur experimenter and commercial radio operator from 1915 to 1924. Following this he joined the experimental research department of the General Electric Company, and a year later became Chief Radio Engineer for the Sparks Withington Company. Since 1935 he has been Section Engineer in the RCA License Laboratory.

FRANK J. SOMERS was born at San Jose, California, May 27, 1908. From 1922 to 1928 he was a licensed amateur radio operator. During the summer vacations of 1928 and 1929 he worked as a laboratory assistant on long wave receiver design for the Federal Telegraph Co. Receiving the degree of B.S. in E.E. from the University of Santa Clara in 1930, he joined the Bell Telephone Laboratories as a member of the technical staff in the Submarine Cable and Special Research Department. From October 1932 to June 1933 Mr. Somers attended the Graduate School of Electrical Engineering at Stanford University. From September 1933 to March 1935 he was employed as a research and development engineer by Television Laboratories Ltd. at San Francisco, Cal. From March 1935 until 1938 he was engaged in design and development work on U-H-F. Transmitters, television receivers and studio equipment for Farnsworth Television, Inc., at Philadelphia, Pa. From 1938 to date he has been a member of the Television Engineering Staff of the National Broadcasting Co. Mr. Somers is an Associate Member of the I.R.E., and is a licensed Professional Engineer in the State of New York.

TECHNICAL ARTICLES BY RCA ENGINEERS

Published Fourth Quarter, 1941

CARSON, B. R.—A Two-Side Non-Turnover Automatic Record Changer—*RCA Review*, October.

DEAL, HARMON B.—Receiver Control by Transmitted Signal—"Alert Receiver"—*RCA Review*, October.

DUKE, VERNON J.—A Method and Equipment for Checking Television Scanning Linearity—*RCA Review*, October.

FERRIS, W. R.—see WAGNER and FERRIS.

FLETCHER, L. E. and C. L. KENNEDY—A Modern Control Room for a Commercial Radio Transmitter Central—*RCA Review*, October.

FOLKERTS, H. F. and P. A. RICHARDS—Photography of Cathode-Ray-Tube Traces—*RCA Review*, October.

FOSTER, D. E. and J. A. RANKIN—Intermediate-Frequency Values for Frequency-Modulated-Wave Receivers—*Proceedings of the I. R. E.*, October.

HILLIER, J.—A Discussion of the Fundamental Limit of Performance of an Electron Microscope—*Physical Review*, October.

——— see ZWORYKIN, HILLIER, and VANCE.

KENNEDY, C. L.—see FLETCHER and KENNEDY.

KORMAN, N. I.—Coupled Resonant Circuits for Transmitters—*Proceedings of the I. R. E.*, October.

LLEWELLYN, F. B.—Book Review of "Electron Inertia Effects"—*Proceedings of the I. R. E.*, October.

MALTER, L.—The Behavior of Electrostatic Electron Multipliers as a Function of Frequency—*Proceedings of the I. R. E.*, November.

MORTON, G. A.—A Survey of Research Accomplishments With the RCA Electron Microscope—*RCA Review*, October.

RANKIN, J. A.—see FOSTER and RANKIN.

RICHARDS, P. A.—see FOLKERTS and RICHARDS.

——— see WALLER and RICHARDS.

THOMAS, H. E.—The Development of a Frequency-Modulated Police Receiver for Ultra-High-Frequency Use—*RCA Review*, October.

THOMPSON, B. J.—Voltage Controlled Electron Multipliers—*Proceedings of the I. R. E.*, November.

VANCE, A. W.—see ZWORYKIN, HILLIER, and VANCE.

WAGNER, H. M. and W. R. FERRIS—The Orbital-Beam Secondary-Electron Multiplier for Ultra-High-Frequency Amplification—*Proceedings of the I. R. E.*, November.

WALLER, L. C. and P. A. RICHARDS—A Simplified Television System for the Radio Amateur and Experimenter—*RCA Review*, October.

ZWORYKIN, V. K., J. HILLIER, and A. W. VANCE—A Preliminary Report on the Development of a 300-Kilovolt Magnetic Electron Microscope—*Journal of Applied Physics*, October.