RCA REVIEW

a technical journal

RADIO AND ELECTRONICS
RESEARCH • ENGINEERING

VOLUME IX JUNE 1948 NO. 2

RCA REVIEW

GEORGE M. K. BAKER
Manager

CHAS. C. FOSTER, JR.
Business Manager

SUBSCRIPTIONS:

ıted States, Canada and Postal Union: One Year $2.00, Two Years $3.50, Three Years $4.50
Single Copies: 75¢ each

Other Countries: One Year at $2.40, Two Years $4.30, Three Years $5.70
Single Copies: 85¢ each

Published quarterly in March, June, September, and December by Radio Corporation of America, RCA Laboratories Division, 30 Rockefeller Plaza, New York 20, N. Y.

Editorial and General Offices: RCA REVIEW, Radio Corporation of America, RCA Laboratories Division, Princeton, New Jersey

Entered as second class matter April 3, 1946, at the Post Office at New York, New York, under the act of March 3, 1879

RADIO CORPORATION OF AMERICA
DAVID SARNOFF, *President*
LEWIS MACCONNACH, *Secretary* ARTHUR B. TUTTLE, *Treasurer*

PRINTED IN U.S.A.

RCA REVIEW

a technical journal

RADIO AND ELECTRONICS
RESEARCH · ENGINEERING

Published quarterly by
RADIO CORPORATION OF AMERICA
RCA LABORATORIES DIVISION
in cooperation with

RCA VICTOR DIVISION
RADIOMARINE CORPORATION OF AMERICA
RCA INTERNATIONAL DIVISION
RCA COMMUNICATIONS, INC.
NATIONAL BROADCASTING COMPANY, INC.
RCA INSTITUTES, INC.

VOLUME IX JUNE 1948 NUMBER 2

CONTENTS

RCA REVIEW

BOARD OF EDITORS

COMPARATIVE PROPAGATION MEASUREMENTS; TELEVISION TRANSMITTERS AT 67.25, 288, 510 AND 910 MEGACYCLES*

By

GEORGE H. BROWN, JESS EPSTEIN AND DONALD W. PETERSON

Research Department, RCA Laboratories Division,
Princeton, N. J.

Summary—*In order to study propagation and multipath effects over a wide range of frequencies under typical broadcast conditions, comparative propagation measurements have been made using television transmitters at 67.25, 288, 510 and 910 megacycles. These measurements were taken along two radials from New York City. One radial extended slightly north of west over extremely hilly country with a number of suburban towns, large homes, many trees and elevations ranging from sea level to 1200 feet. The second line ran southwest over fairly level terrain with very few hills, the highest of which was 230 feet. Regular television broadcasts were used on 67.25 megacycles. A special laboratory transmitter and Turnstile antenna were used for measurements on 288 megacycles. Experimental low-power laboratory transmitters were used with directional antenna arrays for making the measurements on 510 and 910 megacycles.*

The influence of hilly terrain on propagation is clearly illustrated by comparison of the data along the two radials. The best agreement with theoretical values at all frequencies was obtained along the comparatively smooth southwest line. There was closer agreement with the theoretical curves at 67.25 megacycles than at 288 megacycles, while the measured values at 510 and 910 megacycles were usually far below the theoretical. The data has been analyzed to assist in forming an overall picture of the situation.

Shadowing from hills and other obstructions increases steadily as the frequency increases, thus requiring higher power at the higher frequencies. A basis for estimating power requirements as a function of frequency is offered.

Multipath effects are present at both 67.25 megacycles and 288 megacycles, but are usually too slight to be serious. In obstructed or hilly areas, multipath at 510 and 910 megacycles is severe. However, in most places, a clean picture can be obtained by orienting the receiving antenna. It was generally possible to find several responses which gave a good picture. This was true when the receiving antenna was a large array having a narrow beam and a large front-to-back ratio or a single dipole and reflector with a low front-to-back ratio and a broad pattern.

INTRODUCTION

IN THE PAST, surveys have been made of the coverage of individual transmitters operating in the high frequency region. With the advent of black-and-white television on frequencies lying between

* Decimal Classification: R583.16.

50 and 216 megacycles and with growing interest in the possibilities of color television in the region between 500 and 900 megacycles, the need for comparative measurements over a wide range of frequencies became apparent. Accordingly, plans were made early in 1946 to carry out a series of experiments using the Empire State Building in New York City as the location for the transmitters. Field measurements and observations were planned to show attenuation as a function of distance over rough terrain and comparatively smooth terrain and to show the

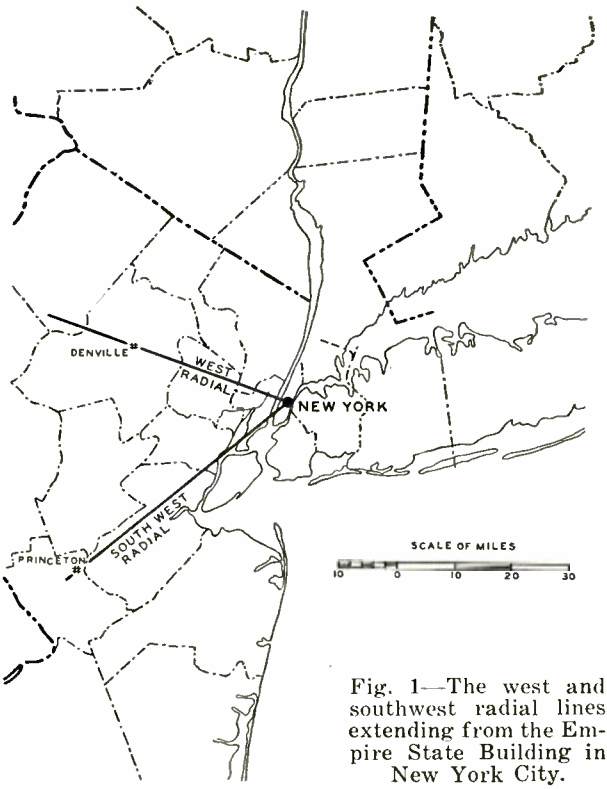


Fig. 1—The west and southwest radial lines extending from the Empire State Building in New York City.

magnitude and differences of multipath effects at the various frequencies.

Since it was evident that it would not be possible to make extensive measurements over the entire service area, two radial lines were selected which could be conveniently reached by the field truck and which presented a variety of terrain. One radial extended slightly north of west from the Empire State Building and will be referred to as the "west radial." This line extended over extremely hilly country, with many suburban towns generously supplied with large homes and multi-

tudes of trees. Elevations ranged from sea level to 1200 feet. The second line was southwest with very few hills, the highest of which was 230 feet.

The locations of these radial lines are shown in Figure 1, while Figure 2 gives the profiles along the radials.

Description of the Transmitting Facilities and the Field Measuring Equipment

Early in 1946, the WNBT transmitter began operating to furnish television broadcasting service with a picture carrier frequency of

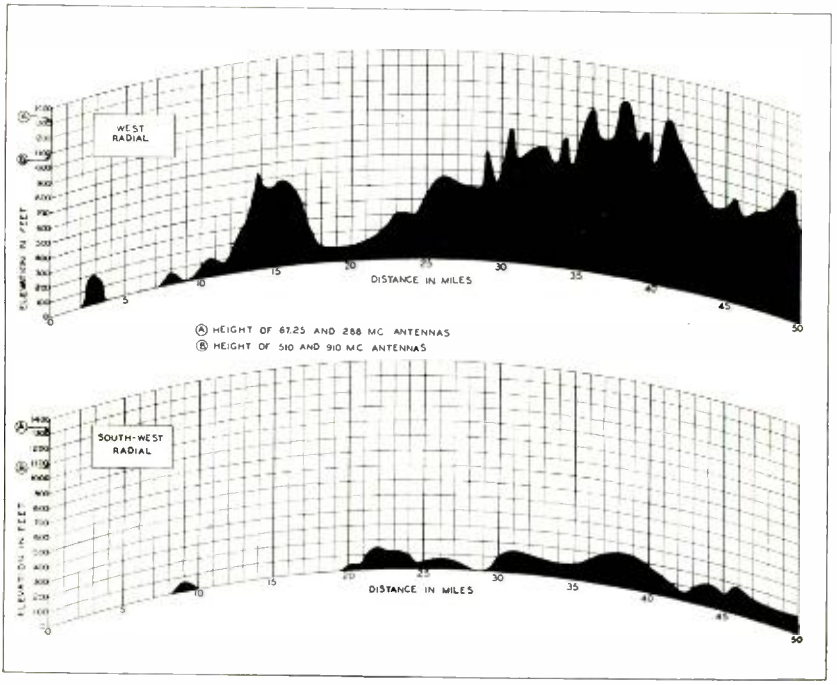


Fig. 2—The profiles along the west and southwest radials.

67.25 megacycles. The transmitter was available with test pattern during most of the daylight hours. This transmitter was located on the eighty-fifth floor of the Empire State Building. Another transmitter, especially constructed at the Laboratories for these propagation tests and operating at a frequency of 288 megacycles, became available and was moved to the eighty-fifth floor of the Empire State Building.

Separate Turnstile antennas for each transmitter were mounted on a single pole on the top of the building. Here, for the first time, an

excellent opportunity was afforded for making comparative tests of propagation at widely-separated frequencies with comparable radiated power and with the transmitting antennas at the same location and height.

The entire antenna system was erected at the Laboratories for purposes of measurement and test before erection on the Empire State Building. Figure 3 shows this antenna system. Closest to the roof may be seen the four-layer Turnstile antenna used to transmit the picture signal with a carrier frequency of 67.25 megacycles and the sound signal at 71.75 megacycles. Directly above this Turnstile antenna is a two-layer loop antenna system for transmission of a

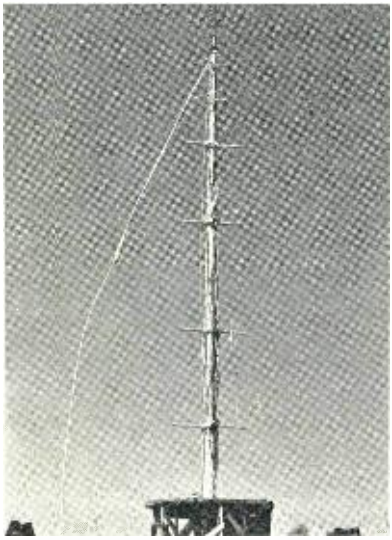


Fig. 3—The multiple antenna system erected at the Laboratories.

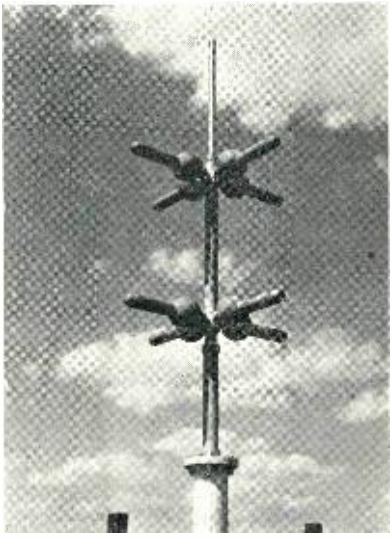


Fig. 4—The 288-megacycle Turnstile antenna.

frequency-modulated signal at 97.3 megacycles. This transmission constituted a broadcasting service which was not utilized in this survey. The 288-megacycle Turnstile antenna is located at the top, above the FM antenna, and may be seen in greater detail in Figure 4.

Power gain measurements were made for both Turnstile antennas when located as shown in Figure 3. In addition, the horizontal radiation patterns of both antennas were obtained by rotating the pole about its vertical axis while reading a field intensity meter located a few thousand feet distant. The vertical pattern of the 288-megacycle Turnstile was obtained by mounting this antenna in a horizontal position on a turntable. The 67.25-megacycle antenna was much too

large to handle in this fashion. However, the vertical pattern was obtained during the development stage by using the ground-plane setup shown in Figure 5. Here the vertical pattern of the final antenna was obtained by measuring the horizontal pattern of the arrangement shown.

During the course of the field measurements at the two frequencies mentioned, the effective radiated power of the 67.25-megacycle transmission was maintained at 3.63 kilowatts, while the effective radiated power at 288 megacycles was 2.6 kilowatts.

The field measuring equipment used a modified TRK-9 receiver. A single stage preamplifier preceded the receiver for 67.25-megacycle

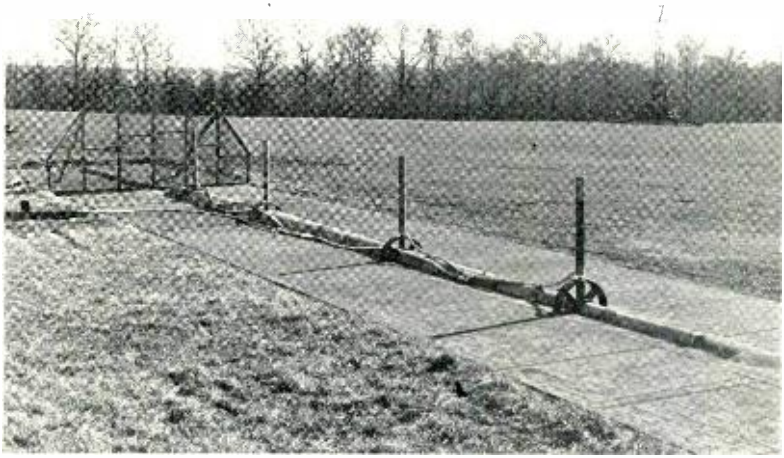


Fig. 5—Ground-plane for the 67.25-megacycle Turnstile.

reception. A 288-to-67.25 megacycle converter was used ahead of the preamplifier for 288-megacycle reception. The TRK-9 automatic-gain-control voltage was measured for field intensity information. All measurements were made with the same test pattern impressed on the transmitters at all times so that a known relation existed between the observed voltage and the value of peak signals. Voltage calibration of the receiver was obtained from a Model 80 Signal Generator.* The apparatus was installed in a panel truck with a portable power supply. A mast mounted on the truck roof was raised to place the two receiving dipole antennas 30 feet above the earth. Both receiving antennas were rotatable.

Measurements were made at approximately two-mile intervals along

* Manufactured by Measurements, Inc.

the two radials shown in Figure 1. Receiving sites were chosen in a variety of surroundings: in open fields, in wooded areas, along highways, atop hills, and in valleys. An effort was made to stay away from electric lines and large buildings. Whenever possible, the antenna was moved a wavelength or more toward the Empire State Building while observing test pattern and gain-control voltage. All data were taken between 10 a.m. and 4 p.m. in clear weather so that even at distances of forty miles or more the measurements may be expected to yield normal values.

The observations at 67.25 and 288 megacycles were carried out between late July and the end of October, 1946. At this time, an

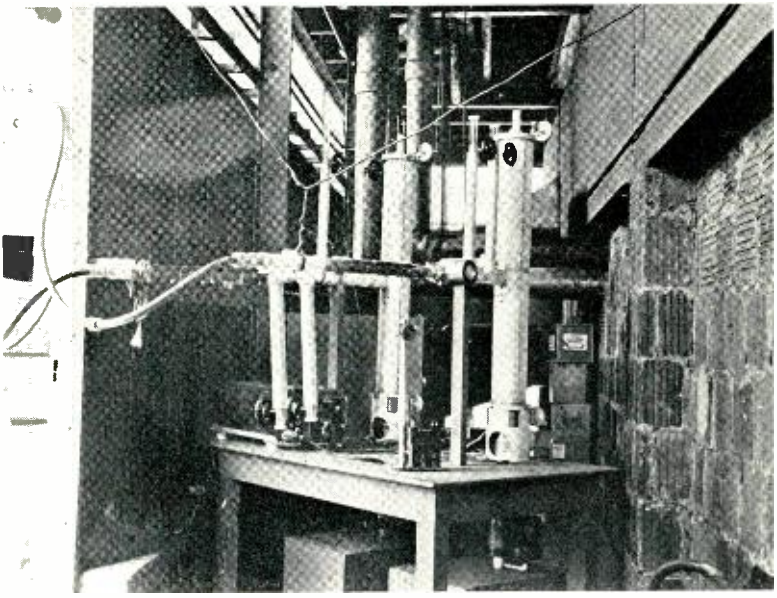


Fig. 6—The self-excited oscillators used in the 510- and 910-megacycle tests.

experimental tube capable of delivering a few hundred watts at frequencies as high as 900 megacycles became available. The measurements just completed indicated that a few hundred watts radiated in a broadcast fashion at 500 and 900 megacycles would yield field intensities too weak to measure easily and certainly too weak to observe multipath effects. Hence directional arrays were constructed and so mounted that the beams could be pointed along either radial. A beam antenna and a self-excited oscillator were provided for operation on a frequency of 510 megacycles, and a similar set of equipment was constructed for operation on 910 megacycles. There was not room to mount the equipment on the roof of the Empire State Building so the

oscillators were placed in a room on the eighty-seventh floor, with the antennas on a balcony at the same level. The antenna elevation was then 1061 feet above ground and 1109 feet above sea level.

The two oscillators are shown in Figure 6. The two directional antenna systems may be seen in Figure 7. Each array consisted of two cophased frames stacked one above the other. The radiation pattern of each system was then 35 degrees wide at the one-half voltage points in the horizontal plane, while the corresponding width in the

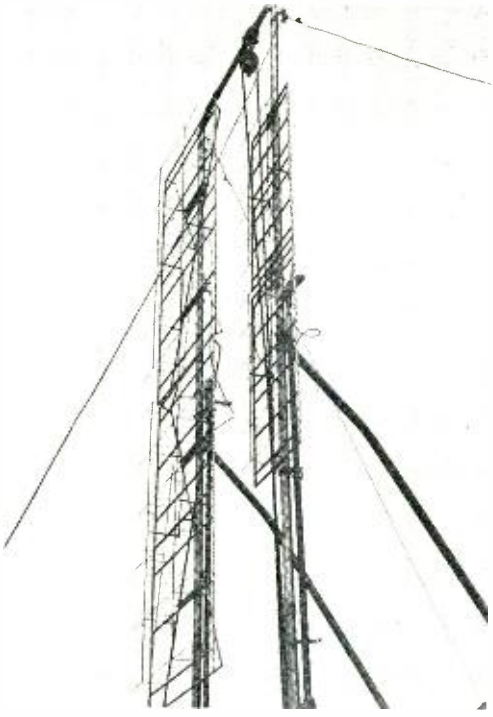


Fig. 7—The directional arrays used in the 510- and 910-megacycle tests.

vertical plane was 8.5 degrees. Another view of the antennas, looking down from above, is given in Figure 8 (page 184).

A modified AN/APR-4 receiver was used for the field intensity measurements. Two specially-constructed signal generators were used for calibration purposes. The receiving antenna for 510 megacycles was a single dipole in front of a flat metal reflector. For 910 megacycles, a similar antenna was provided. In addition a directional antenna was used which consisted of a single frame identical in construction to the two frames used for transmission. The three antennas mounted on the field truck are shown in Figure 9 (page 184).

Fig. 8—The 510- and 910-megacycle antennas on the eighty-seventh floor balcony.

Fig. 9—The field truck with the 510- and 910-megacycle receiving antennas in place.

The two transmitters were operated simultaneously so that a measurement could be made at each frequency by a single visit to a measuring point. Measurements were made on each radial at exactly the same locations used in the previous survey at 67.25 and 288 megacycles. Measurements were also made at many additional locations.

These field intensity measurements were made during the summer of 1947. When the measurements were completed, the oscillators were modulated with a test pattern. A television receiver, Model 630TS, with an appropriate converter, was then used in the field truck to observe the effects of multipath propagation and the measuring points were again visited. For these latter observations, a frame similar to the directional array shown in Figure 9 was also used at 510 megacycles.

Measured Field Intensity Data

The principal factors involved in establishing the field intensity at high frequencies are the effective radiated power, the distance between the transmitting and receiving antennas, the heights of the transmitting and receiving antennas, the earth's curvature, the conductivity and dielectric constant of the earth, refraction resulting from changing dielectric constant of the air with height, and irregularities of the terrain. The last-named factor is not accounted for in theoretical treatments. If the earth is assumed to be a smooth sphere having uniform electrical properties and if the dielectric constant of the air is assumed to be a linear function of height, the field intensity calculation may be carried out by the methods described by K. A. Norton[1]. In order to review the effects of frequency where the ideal or theoretical conditions exist, Figure 10 was prepared. In these calculations, the following factors were used:

Height of transmitting antenna	= 1250 feet
Height of receiving antenna	= 30 feet
Conductivity of the earth	$= 50 \times 10^{-15}$ electromagnetic units
Dielectric constant of the earth	= 15
Effective radiated power	= 100 kilowatts.

An inspection of Figure 10 reveals the increase in field intensity that accompanies an increase in frequency when the distance is less than

[1] A Report Prepared by K. A. Norton on the Calculation of Ground Wave Field Intensity over a Finitely Conducting Spherical Earth, and presented at the Hearing before the FCC in the Matter of Aural Broadcasting on Frequencies above 25,000 Kilocycles, March 18, 1940.

60 miles. At distances greater than 70 miles, the fields for the high frequencies fall far below those for the low frequencies. A line showing twice the free space field is also included on this figure. This line shows the limiting value above which the field intensity never rises, even when the direct wave and the ground reflected wave are in phase.

Figure 11 shows the measured field intensity of the 67.25-megacycle signal along the west and southwest radials as a function of distance. The calculated field intensity is shown in each case by the solid curve. The measurements along the southwest radial show close correlation

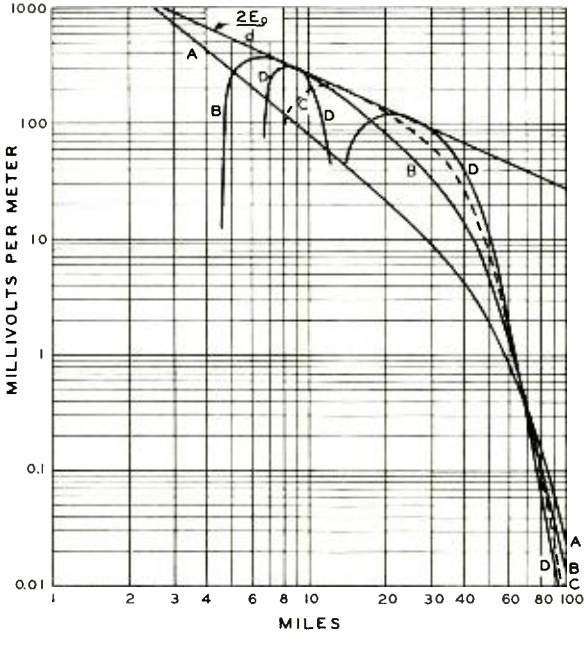


Fig. 10—Theoretical field intensities over a smooth spherical earth.
A. 67.25 megacycles. C. 510 megacycles.
B. 288 megacycles. D. 910 megacycles.

with theory. The comparatively smooth terrain is responsible. Data taken along the west radial shows much more scattering. An idea of the difference in terrain may be obtained by inspecting Figure 2.

The measured data taken at 288 megacycles along the same radials is displayed in Figure 12. The theoretical curve is again shown here. The scattering is pronounced. Several points on the charts are connected by vertical lines. This is to indicate that the two values thus

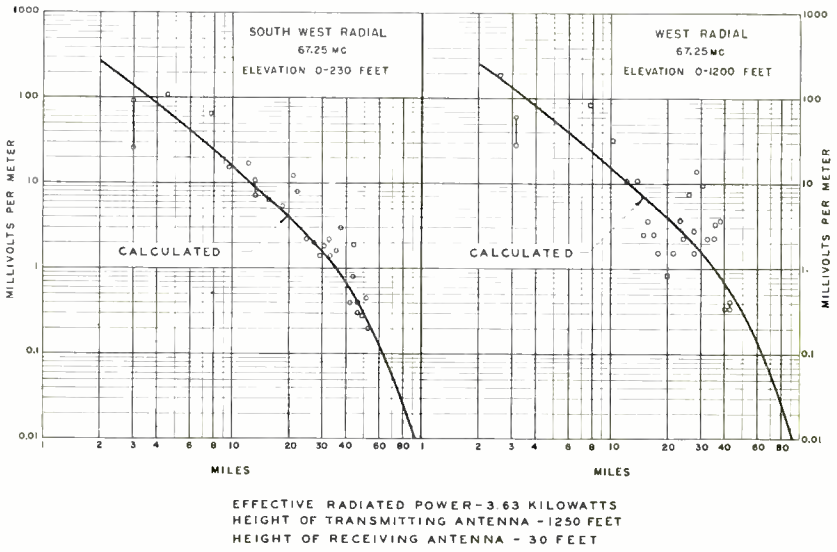


Fig. 11—Measured field intensity at 67.25 megacycles.

linked represent the maximum and minimum signals encountered in the vicinity of the measuring point. The asterisks on the figure indicate the location of points where the field intensity was below 0.1 millivolt per meter and was not measurable because of the noise level of the receiver.

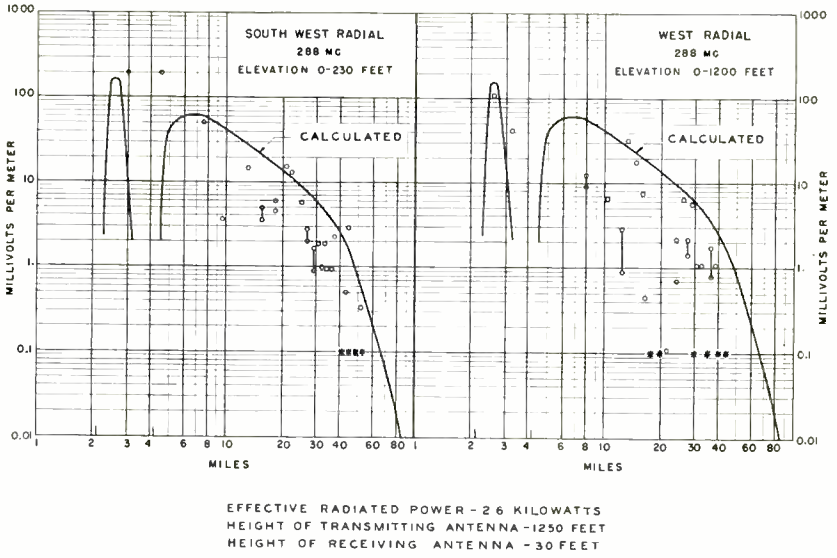


Fig. 12—Measured field intensity at 288 megacycles.

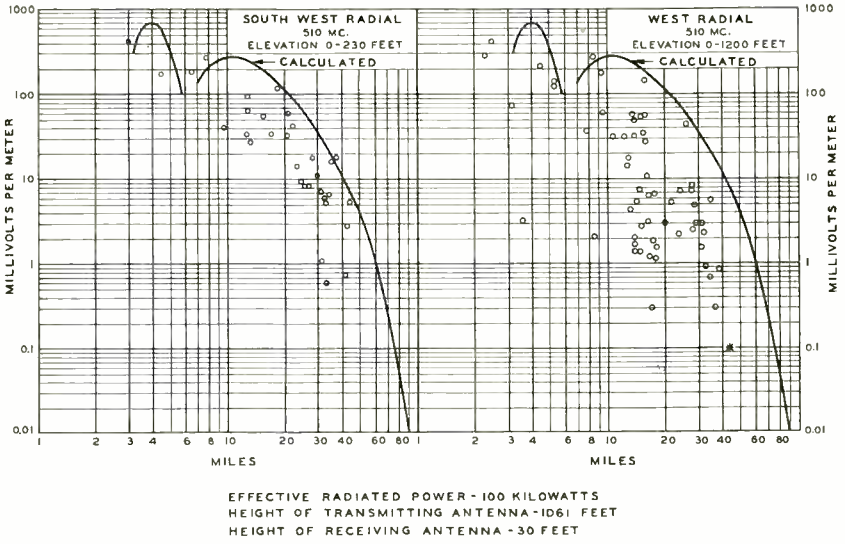


Fig. 13—Measured field intensity at 510 megacycles.

The measured field intensity along the southwest and the west radial at a frequency of 510 megacycles is shown in Figure 13. Before plotting in this figure, the measured values were corrected to an effective radiated power of 100 kilowatts. It should be noted that in this figure, the theoretical graph is based on a transmitting antenna height of

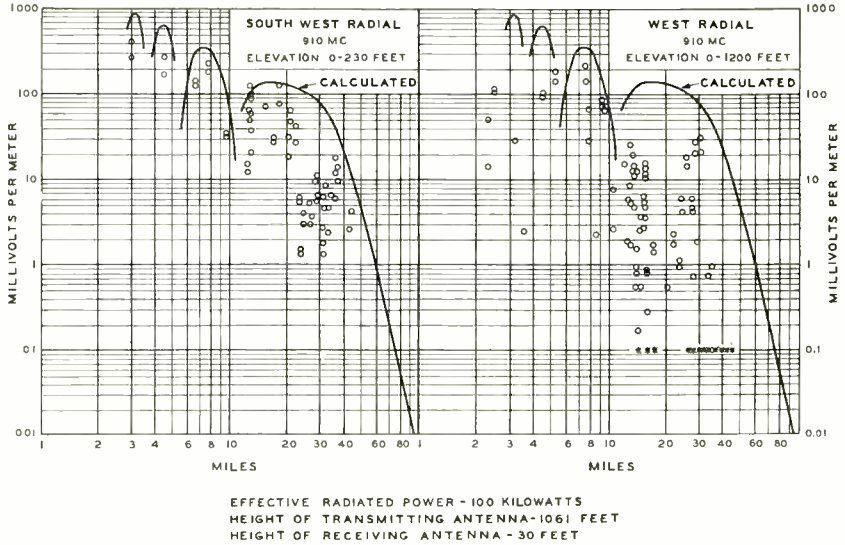


Fig. 14—Measured field intensity at 910 megacycles.

1061 feet, since for this transmission the antenna was on the eighty-seventh floor balcony. The effect of the roughness of the terrain on the west radial is becoming apparent at this frequency.

Figure 14 shows the measured values obtained at a frequency of 910 megacycles, again with the measured values corrected to an effective radiated power of 100 kilowatts. The calculated curve is based on a transmitting antenna height of 1061 feet. On the west radial, many locations are indicated where the field intensity was too low to be measured.

When measuring at 510 megacycles, the receiving antenna was a single dipole placed in front of a flat metal sheet. At many measuring points, when the receiving antenna was rotated the receiver voltage varied in such a way that it was evident that multipath signals were present. The same effect was noted to be more pronounced at 910 megacycles. At this higher frequency, measurements were made with a dipole in front of a plane reflector and with an array. The receiving array was 26 inches wide and 56 inches high. This array may be seen

Table I—Percentage Distribution of the Ratio V_A/V_D in Terms of Locations (Frequency = 910 Megacycles)

The ratio V_A/V_D lies between:	Percentage of locations where the ratio lies between the indicated limits.		
	Southwest radial	West radial	Both radials
0-1	0	18.	10.1
1-1.5	0	0	0
1.5-2	0	0	0
2-2.5	0	7.7	4.4
2.5-3	3.33	15.5	10.1
3-3.5	20.	12.9	15.9
3.5-4	10.	20.5	15.9
4-4.5	23.3	12.9	17.4
4.5-5	16.6	2.5	8.7
5-5.5	16.6	5.	10.1
5.5-6	3.33	0	1.5
6-6.5	0	0	0
6.5-7	6.66	0	2.9
7-7.5	0	2.5	1.5
7.5-8	0	0	0
8-8.5	0	2.5	1.5

Average V_A/V_D

On west radial — 3.3
On southwest radial— 4.38
On both radials — 3.75

in Figure 9. In many shadowed locations, it was possible to rotate the array and observe strong signals from many directions. This was accepted as evidence of strong multipath signals.

An additional disturbing effect was found at 910 megacycles. A coaxial switch was used to switch the receiver rapidly from the dipole to the array. In open locations, particularly on the southwest radial, the voltage on the receiver terminals was substantially greater when the array was used. However, in shadowed regions on the west radial, it was often found that the receiver voltage was lower with the array than the dipole. This was probably due to the fact that the field was badly

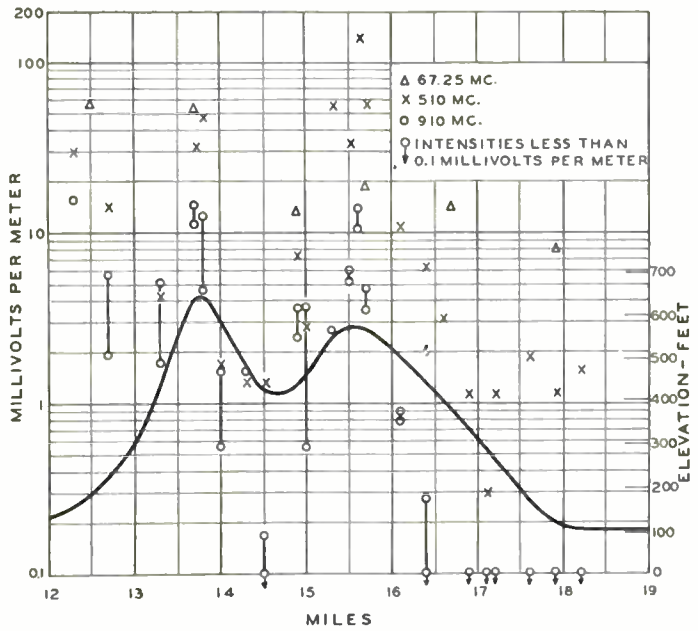


Fig. 15—The effect of a ridge of hills demonstrated at three widely-separated frequencies. (The measured field intensities have all been corrected to an effective radiated power of 100 kilowatts. The connected circles for the 910-megacycle signal indicate the two values of field intensity obtained at each location, one with the dipole antenna, the other with the array.)

distorted with large phase changes across the aperture of the array. Unfortunately, this effect was found only in the shadowed regions where the signal was weak and a boost in signal such as one might expect with an array was badly needed. Table I shows the number of places, expressed percentagewise, where the ratio V_A/V_D lies between the indicated limits. (V_A is the voltage on the receiver terminals when the array is used, while V_D is the corresponding voltage when the dipole is connected.) It is interesting to note that on the southwest

radial, no values less than 2.5 were obtained for V_A/V_D, while on the west radial 18 per cent of the observations show the ratio to be less than unity.

The manner in which the shadows due to large hills change with frequency is shown in Figure 15. Here the measured field intensities for 67.25, 510 and 910 megacycles have all been corrected to an effective radiated power of 100 kilowatts. The 67.25-megacycle signal is lowered somewhat by the ridge of hills, while the 510-megacycle signal fluctuates at a great rate. The shadows cast by the two peaks are quite apparent. However, even this large fluctuation does not compare with the effect at 910 megacycles. Behind the second peak, a region over three miles in extent was found where it was not possible to measure the weak signal of the 910-megacycle transmitter.

Analysis of the Field Intensity Data

While an inspection of Figures 11 to 14 inclusive shows the general trend of attenuation as the frequency is increased, it seems appropriate to attempt a more quantitative analysis of the data. With this in mind, Figures 16 and 17 were prepared. Figure 16 was constructed directly from the measured values and the theoretical curves shown in Figures 11 to 14. Thus, in Figure 16, a curve for a given frequency relates the measured field intensities at that frequency to the theoretical curve for the same frequency. However, a glance at Figure 10 reveals that the theoretical values for widely-divergent frequencies are not related by a simple law. To compare the results measured at four frequencies, it is necessary to establish a base. This has been done in preparing Figure 17. First, the curve for 67.25 megacycles on Figure 16 was transferred directly to Figure 17. Then on a copy of Figure 12 which shows the measured field intensities for 288 megacycles, a theoretical curve was plotted for the field intensity at 67.25 megacycles, using transmitting and receiving antenna heights of 1250 feet and 30 feet, respectively. The effective radiated power was assumed to be 2.6 kilowatts, the same value used in the 288-megacycle measurements. Then the measured 288-megacycle field intensities were compared to the theoretical curve for 67.25 megacycles. The result of this operation is shown by the 288-megacycle curve of Figure 17. Next, on Figure 13 was plotted a theoretical field intensity curve for a frequency of 67.25 megacycles, a receiving antenna height of 30 feet, and an effective radiated power of 100 kilowatts. The transmitting antenna height was taken as 1061 feet, to correspond to the height of the antenna used for 510 megacycle transmission. The measured field intensities at 510 megacycles were then compared to this new 67.25 megacycle theoretical

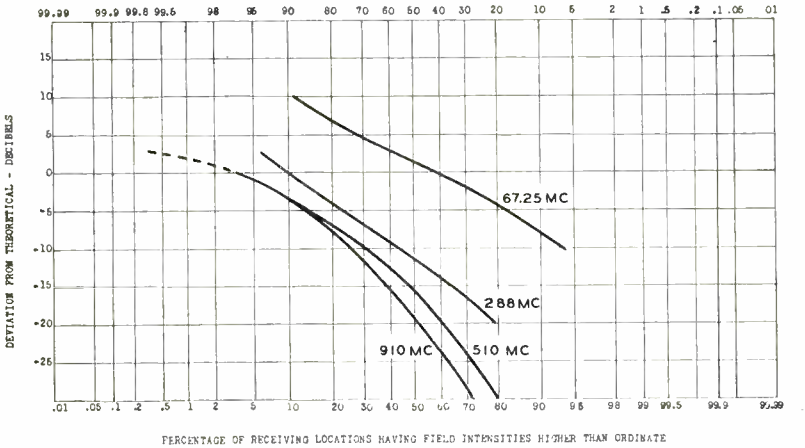


Fig. 16—An analysis of the data shown in Figures 11 to 14. (A curve for a given frequency relates the measured field intensities at that frequency to the theoretical curve for the same frequency.)

curve to form the 510-megacycle curve on Figure 17. The 910-megacycle curve on Figure 17 was constructed in the same manner.

The following interesting tabulation may immediately be made from Figure 17:

Frequency (megacycles)	Percentage of receiving locations having intensities greater than the 67.25-megacycle theoretical values	Percentage of receiving locations having intensities less than the 67.25-megacycle theoretical values
67.25	58.	42.
288.	52.	48.
510.	35.	65.
910.	22.	78.

Figures 16 and 17 have been replotted in a more demonstrative manner in Figures 18 and 19, respectively. A glance at Figure 19 shows that the whole group of distribution curves is reasonably flat from 67.25 megacycles to 288 megacycles. Thus, it may be concluded that for practical reception conditions such as were encountered in this survey, essentially the same field intensity in millivolts per meter may be expected to result for any frequency between 67.25 megacycles and 288 megacycles, for a given transmitting antenna height and transmitter power. As the frequency is increased still further, expectations of a given field intensity are notably decreased.

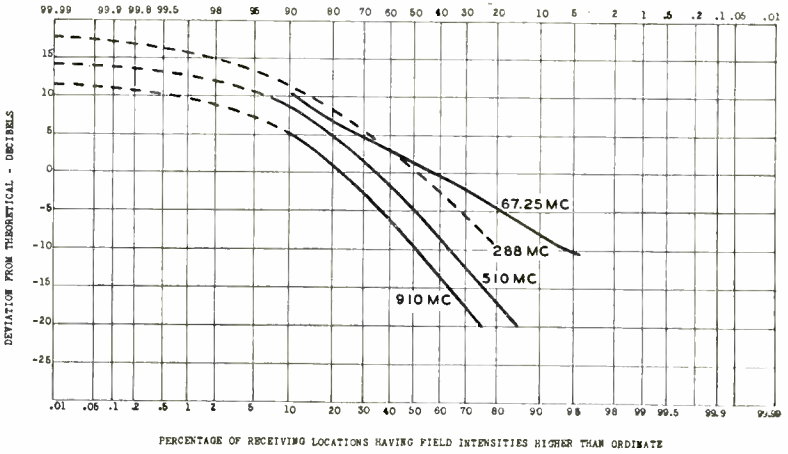


Fig. 17—A second analysis of the data shown in Figures 11 to 14. (The theoretical curve referred to on the ordinate scale is the theoretical curve for 67.25 megacycles. A curve for a given frequency relates the measured field intensities at that frequency to the theoretical curve for 67.25 megacycles at a corresponding antenna height.)

Figure 19 now affords a means of drawing some conclusions concerning the power requirements at the frequencies suggested for color television, namely the region between 500 and 900 megacycles. Assume that a signal is being received on a half-wave dipole and this dipole

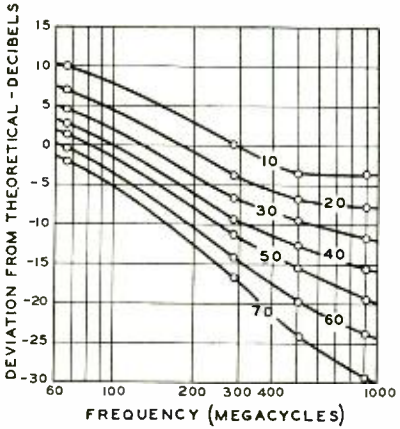
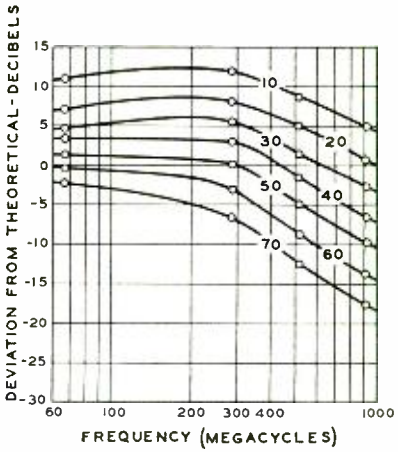


Fig. 18—Figure 16 replotted to demonstrate the phenomena more clearly. (The numbers on the individual curves indicate the percentage of the receiving locations having field intensities higher than the ordinate values shown on the left-hand scale.)

Fig. 19—A replot of Figure 17. (The measured field intensities are compared to the theoretical values at 67.25 megacycles.)



in turn feeds the signal to a transmission line through a suitable matching section. Then if it is also assumed that the receiver input impedance is the same as the characteristic impedance of the transmission line, the voltage on the transmission line (essentially the same as the voltage on the receiver terminals for a low-loss line) is

$$V_{TL} = \frac{\lambda E}{2\pi} \cdot \sqrt{\frac{Z_c}{R_r}}$$

where Z_c = the characteristic impedance of the transmission line,

R_r = the radiation resistance of the half-wave dipole,

λ = the wavelength, in meters,

E = the field intensity in millivolts per meter,

V_{TL} = voltage on the transmission line, in millivolts.

Since it is assumed that the antenna is cut to be one-half wave in length at the frequency in question, the voltage on the transmission line will decrease accordingly as the wavelength is decreased. If the frequency is multiplied by a factor n, the field intensity E must be also multiplied by the same factor to keep the voltage on the transmission line constant. That is, if the frequency is n times 67.25 megacycles, the power used at 67.25 megacycles must be multiplied by a factor n^2 to offset the effect of the shorter antenna at the new and higher frequency. A second multiplying factor is then necessary to account for the difference in propagation characteristics at the two frequencies in question. Suppose that at 510 megacycles, one wishes to insure that in at least 50 per cent of the receiver locations the signal on the transmission line will be at least as great as the signal enjoyed by 50 per cent of the receiver locations when the frequency is 67.25 megacycles. On Figure 19, the curve indicated for 50 per cent is examined and a difference in signal level of 6.5 decibels is found for the two frequencies. The power must be multiplied by a factor of 4.46 to take care of the difference in propagation. For this example, n^2 is 57.3, so the total factor on the power is 255. In other words, to insure that the receiver input voltage at 50 per cent of the receiving locations using a frequency of 510 megacycles will exceed the value exceeded by 50 per cent of the receiver locations using 67.25 megacycles, the power used at 510 megacycles must be 255 times the power used at 67.25 megacycles.

The two upper curves of Figure 20, constructed in the manner just outlined, apply to 70 and 50 per cent of the locations when it is assumed that half-wave dipoles are used.

If instead of half-wave dipoles as receiving antennas directional arrays are used with power gains of n^2, the power multiplying factor consists only of the factor obtained from Figure 19. The two lower curves on Figure 20 result.

It seems appropriate at this point to discuss the matter of power gain as applied to the receiving antenna. In the case of a directional or beam antenna in which the aperture or mouth opening has a given area, one may quickly estimate the maximum possible power gain by

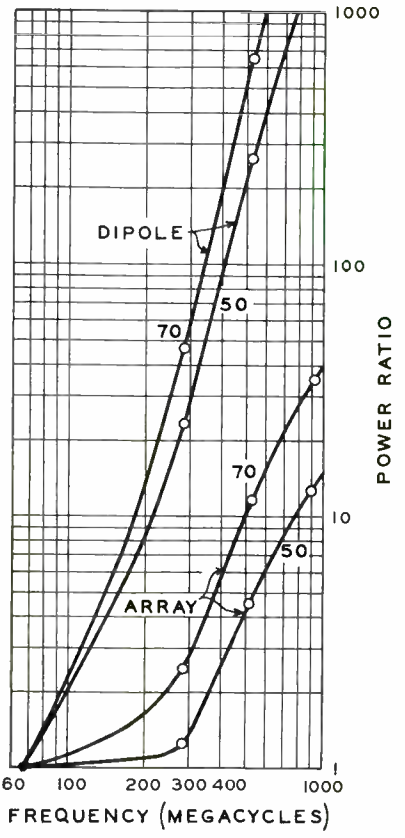


Fig. 20—An estimate of the power requirements for television broadcasting versus frequency. (The ordinate scale is the factor by which the power used at 67.25 megacycles must be multiplied, when moving to a new frequency, in order to obtain the same voltage on the receiver transmission line. The upper two curves assume the use of a half-wave dipole at each frequency, while the lower two curves postulate the use of an array with a power gain sufficient to overcome the loss due to the shortening of a half wave dipole as the frequency is increased. The numbers on each curve refer to the percentage of receiving locations considered.)

expressing the area in terms of the wavelength. Then the power gain is 7.65 times the area measured in square wavelengths. As the frequency increases and the wavelength decreases, the power gain increases proportionally as the square of the frequency, since for a fixed area the number of square wavelengths included in the area increases in the same manner. However, this power gain is expressed in terms of the power received by a half-wave dipole at the particular frequency under consideration. But as the frequency increases the reference half-wave

dipole becomes shorter. The practical result is that to obtain a given voltage on the terminals of the transmission line, where the field intensity remains constant in millivolts per meter as the frequency increases, use must be made of an antenna of constant physical aperture. The effective aperture or capture area of a half-wave dipole at 67.25 megacycles is approximately 28 square feet. Hence to obtain the same voltage on the transmission line at 510 megacycles, one must use the same area for the directive antenna, and at 910 megacycles one must again use an antenna with an aperture which is over 5 feet on a side. This hypothetical antenna with a mouth opening of 28 square feet will have a power gain of 57.5 at 510 megacycles and a power gain of about 180 at 910 megacycles.

The fact demonstrated in Table I, namely that on the extremely hilly radial the directional receiving array failed to improve the signal in 18 per cent of the locations has already been indicated. These were at locations where the signal was weak and the gain was badly needed. Hence the true picture of the needed power ratio versus frequency lies somewhere between the two groups of curves shown in Figure 20.

Multipath Observations

Multipath observations on the 67.25- and 288-megacycle channels were carried out during the summer of 1946 at the same time that the field intensity measurements were made. The observations were all made outside of Manhattan, so that extreme conditions of multipath effects were not observed. Multipath effects were usually different at the two frequencies. Where usable field intensities were obtained on both frequencies, it was generally possible to orient the receiving antennas and receive good pictures on either frequency, with multipath effects entirely absent or of such low magnitude as to be quite unimportant. At two sites approximately 2.5 miles from the transmitter on the west radial, the picture received on 67.25 megacycles with a receiving antenna height of 30 feet showed both positive and negative ghosts. When the receiving antenna was lowered to 6 feet, the picture was materially improved. Only two sites, at about 30 miles on the west radial, were found where the 67.25-megacycle picture obtained with the dipole antenna was not usable due to multipath. At both sites, no 288-megacycle signal could be found.

On the comparatively smooth southwest radial, very little evidence of multipath was noted on either frequency except in a few locations in the vicinity of large fields of oil storage tanks.

In the early fall of 1947, the 510 and 910 megacycle oscillators were modulated with the test pattern. On the southwest radial, the pictures

were clean and free from multipath. When the directional arrays were used, a strong clean picture was obtained when the receiving antenna was pointed in the direction of the Empire State Building. When the antenna was rotated to point in other directions, weak reflected signals were sometimes noted but these reflected signals were usually too weak to use. In the vicinity of the oil storage tanks mentioned above, the reflected signals compared in intensity to the main signal but there was no real difficulty in obtaining a clean picture on either frequency, using either an array or a dipole in front of a screen.

On the west radial, particularly in the shadowed areas but also in many open spaces flanked by hills or buildings, signals arriving by many paths were numerous. As the directional arrays were rotated, it was generally possible to find several positions where strong signals were received. Surprisingly, many of these reflected signals were as strong as the main signal and produced as clean a picture. The difference between multipath effects at 510 megacycles and 910 megacycles is simply that the multipath signals are more profuse at the higher frequency, and, of course, in shadowed areas it was harder to receive a signal strong enough to give a satisfactory picture at the higher frequency. At many points where an acceptable picture was obtained at 510 megacycles, the signal was too weak to give a good picture at 910 megacycles.

While multipath propagation is very much in evidence at both frequencies, it appears to be a simple matter in most instances to orient the receiving antenna to obtain a clean picture free of multiple images. This conclusion should be qualified because of two factors which may play an important part in successful television broadcasting at these higher frequencies. First, while it seems generally possible to find a receiving antenna position which will produce a clean picture, it is quite possible that this will not be the best position for another transmission at some other frequency in the band between 500 and 900 megacycles. Hence the receiving antenna installation problem might be complex for a television service established in this band of frequencies. Secondly, in the tests at 510 megacycles and 910 megacycles, directional transmitting antennas were used in order to obtain adequate signals. This directivity at the transmitting antenna prevents multipath signals which arise from reflections from buildings on Manhattan Island. For example, if a signal were truly broadcast from the Empire State Building, a receiver on the west radial might be subject to multipath signals originating by reflections from buildings lying to the north or south or even east of the Empire State Building, while these signals could not be present when directive transmitting antennas are used.

Hence the conclusions drawn from the tests concerning multipath effects in the 500- to 900-megacycle band may be somewhat optimistic when applied to a broadcast service.

Conclusion

The influence of hilly terrain on propagation is clearly illustrated by comparison of the data along the west and southwest radials. The best agreement with theoretical values at all frequencies was obtained along the comparatively smooth southwest line. There was closer agreement with the theoretical curves at 67.25 megacycles than at 288 megacycles, while the measured values at 510 and 910 megacycles were usually far below the theoretical. The data has been analyzed and presented in Figures 16 to 19 to assist in forming an overall picture of the situation.

Shadowing from hills and other obstructions increases steadily as the frequency increases, thus requiring higher power at the higher frequencies. Figure 20 offers a basis for estimating power requirements as a function of frequency. In general, the service area depicted by a series of contours of constant field intensities will be about the same size for the high frequencies as for the lower frequencies, but the service area for the high frequencies will be spotted with local areas where the signal is low or nonexistent. To insure signals of usable strength in some of these local shadow areas, at frequencies between 500 and 900 megacycles, it would be necessary to increase the radiated power by a fantastic ratio. On the other hand if the entire service area of a 900-megacycle transmitter consisted of terrain similar to that encountered on the southwest radial, one might expect to provide a very substantial service with a radiated power of the order of 100 kilowatts.

The effect displayed in Table I, namely the failure of high-gain directive receiving antennas to function properly in shadowed areas where the field is badly distorted, cannot be over-emphasized since this makes it impossible to employ a simple means of making use of weak signals.

Multipath effects are present at both 67.25 megacycles and 288 megacycles, but are usually too slight to be serious. In obstructed or hilly areas, multipath at 510 and 910 megacycles is severe. However, in most places, a clean picture could be obtained by orienting the receiving antenna. It was generally possible to find several responses which gave a good picture. This was true when the receiving antenna was a large array having a narrow beam and a large front-to-back ratio or a single dipole and reflector with low front-to-back ratio and

a broad pattern. It is quite likely that the best position and orientation of the receiving antenna for one station operating in the frequency band between 500 and 900 megacycles will not prove to be best or even suitable for one or more other stations operating on other channels in the band. Indeed, receiving antennas may be required which are rotatable and even this added luxury may prove to be insufficient. At least, it seems evident that the receiving antenna problem will be of primary importance in establishing a successful television broadcasting service at these higher frequencies.

ACKNOWLEDGEMENT

The assistance rendered by others in bringing this extended investigation to a successful conclusion is gratefully acknowledged. In particular, mention should be made of the operating staff of Television Station WNBT who operated the test transmitters and kept detailed records of operation while carrying out their normal duties, as well as the several groups of engineers at RCA Laboratories who constructed the 288-megacycle transmitter, developed the tubes and designed the oscillators for the 510 and 910 megacycle transmissions, and produced the converters to use with the 630TS receivers.

APPENDIX

System Calibration for Determining Field Intensities

In the process of measuring the field intensities at 510 and 910 megacycles, it was necessary to determine a number of physical quantities at the transmitting end. The effective radiated power was found by estimating the power gain of the directive array, reducing this figure slightly to allow for power loss in the transmission line, and finally multiplying these factors by the actual power output of the oscillator. The power output was determined by replacing the antenna by a water-cooled load resistor arranged so that water flow and temperature rise of the cooling water through the resistor could be measured. A thermal type milliammeter with a square-law scale was loosely coupled to the transmission line and its deflection was proportional to power; the proportionality constant was determined by the test with the water-cooled resistor. The power-indicating meters may be seen on the two transmission lines shown in Figure 8.

On the receiving end, the effective length and gain of the receiving antenna was measured. These figures, together with transmission line loss and signal generator calibration, related the signal generator reading to the field intensity.

When the field measurements were completed, the entire set of transmitting equipment used at 510 and 910 megacycles was moved to the Laboratories in Princeton, where the directive transmitting antennas were erected on the roof of a one-story building. The field intensity measuring equipment in the field truck with the receiving antenna 30 feet above ground was used to measure the signal in the beam of the transmitting antenna at a number of points a few hundred feet from the transmitting antenna. The vertical beam width of the transmitting antenna was narrow enough so that for a distance of over 400 feet the receiving antenna was subjected to only the direct wave, with no ground-reflected signal. Only two quantities were necessary in this determination. The first was the deflection of the power-indicating meter on the transmission line feeding to the transmitting antenna. This meter reading may be referred to as D. Then in the field truck, the receiver was tuned to the operating frequency and the deflection of the output meter noted. The antenna was then replaced by the signal generator and the signal generator output adjusted until the output meter on the receiver showed the same deflection as before. The signal generator output voltage was then recorded. This latter quantity is designated as V_{sg}.

The free-space field for P kilowatts of radiated power is

$$E \text{ (millivolts per meter)} = \frac{137\sqrt{P}}{d_{\text{miles}}} = \frac{724{,}000\sqrt{P}}{d_{ft}} \tag{1}$$

But

$$\mathrm{E} = V_{sg}K_1 = \frac{724{,}000\sqrt{K_2 D}}{d_{ft}} \tag{2}$$

where K_1 = the product of receiving-antenna gain, effective length, and the loss factor for the transmission line; and

K_2 = factor on the power-indicating meter to convert to kilowatts, lumped with the power gain of the transmitting antenna, including transmission line loss.

It should be remarked that it is not even necessary that the signal generator be calibrated in absolute value, but only necessary that the linearity be established and an unknown coefficient of linearity be assumed. This latter coefficient may then be lumped in K_1.

The solution of Equation (2), gives

$$\frac{K_1}{\sqrt{K_2}} = \frac{724{,}000\sqrt{D}}{V_{sg} \cdot d_{ft}} \tag{3}$$

Following this procedure, the ratio $K_1/\sqrt{K_2}$ was determined at both 510 and 910 megacycles, using a dipole in front of a flat reflector and a directional array for receiving at each frequency. Thus there were obtained four values of this ratio for reference use in determining actual values from the field data obtained along the west and southwest radials.

The field intensity at some point along a radial is proportional to the square root of the effective radiated power. This field, called E′, is

$$E' \text{ (millivolts per meter)} = F \cdot \sqrt{K_2 D'} \qquad (4)$$

where D' is the reading of the power meter at the transmitter, shown by the log book to be the value existing at the approximate time of measurement of field intensity. K_2 is the same as the coefficient used in Equation (2) because the same transmitting equipment was used in both cases. The quantity, F, is a function of the transmitting and receiving antenna heights, the frequency, and a number of other variables, including the effect of hills and trees on the propagation. F could be determined analytically only by possessing unlimited mathematical ability. In any event, it is seen that F is field intensity in millivolts per meter for one kilowatt of radiated power, the quantity that was the object of the survey.

With the receiving equipment, the signal generator reading is determined that corresponds in strength to the signal found on the receiver terminals. This reading is V'_{sg}. Then the field intensity is

$$E' = V'_{sg} K_1 \qquad (5)$$

Equating (4) and (5), the quantity F is found to be given by the relation

$$\mathrm{F} = \frac{V'_{sg}}{\sqrt{D'}} \times \frac{K_1}{\sqrt{K_2}} \text{ (millivolts per meter for one kilowatt)} \qquad (6)$$

Thus to obtain the quantity F, one needs simply the measured signal generator output, V'_{sg}, the power output meter deflection D', and the appropriate ratio determined in Equation (3). The experimental values shown in Figures 13 and 14 were derived from the field observations in this manner.

It is interesting to see that the method outlined in this Appendix yields the desired field intensity data, namely millivolts per meter for one kilowatt of effective radiated power, without it becoming necessary actually to measure the power or to calibrate the receiving equipment or signal generator.

MOTION PICTURE PHOTOGRAPHY OF TELEVISION IMAGES*

By

Robert M. Fraser

Television Development, National Broadcasting Company, Inc., New York, N. Y.

Summary—*The permanent recording of television programs, for documentary, historical, legal, or critical purposes and as an aid to networking may be accomplished by motion picture photography of the television image, making use of the practical arts developed by motion picture engineers for recording the sight and sound of a television broadcast. This paper describes the apparatus and methods developed for the photographing of the television cathode-ray image. Development is traced from the first attempts in 1938 through the experimental cameras to the commercial camera system now in use. Subsequent sections deal in some detail with 16- and 35-millimeter equipment considerations, kinescope phosphors and film spectral characteristics, resolving power of films, exposure of film, processing and printing of kinescope film, the photographic monitor, and sound recording.*

Introduction

A METHOD of permanently recording the otherwise transient video signal is desirable in the advancement of television art. Motion picture photography of the television image is a method of doing this. Use is made of the practical arts developed by the motion picture engineers in recording the sight and sound of a television broadcast. The reasons for recording television programs are manifold and are similar to the reasons for recording the sound of a standard broadcast program. Television recordings may be made for documentary purposes, to preserve an historical event, for legal purposes, and for critical purposes. These recordings may be used for a delayed or a repeat broadcast or for syndication to other television stations unable to obtain network programs because of the lack of coaxial cable or microwave relay connections with network sources. Television recordings of auditions are useful in the marketing of programs or talent.

It is the purpose of this paper to describe the apparatus and methods developed for the motion picture photography of the television cathode-ray image.

In 1938, the first attempts were made to photograph the television image on motion picture film. Kinescopes or cathode-ray tubes at that time used low efficiency phosphors and operated on relatively low

* Decimal Classification: R583 × R582.

second-anode voltages compared with present day practices. The amount of light obtainable from these cathode-ray tubes was not enough to produce a full exposure on the fastest films then obtainable with an exposure of one thirtieth of a second in a sixteen frame-per-second camera. By photographing the cathode-ray tube at eight frames per second with an exposure time of one fifteenth of a second, recognizable images were obtained on the motion picture film. Of course, these films when projected on a twenty-four frame projector show an unnatural rapidity of motion and are considered to have nothing more than historical interest.

The cameras used in these early experiments were spring-motor driven and the shutter rate was therefore nonsynchronous with the frame rate of the television system. This gave rise to phenomena termed "shutter bar", or banding — a black or a white bar which in a nonsynchronous system moves across the film image when projected, at a rate dependent on the difference in frequency between that of the television system and the frame rate of the motion picture camera. The width of the bar depends on the shutter angle of the camera. If the shutter angle and the frame rate of the camera combine to give an exposure time of less than one thirtieth of a second, less than a full television frame is photographed and a black or under-exposed section of the image results. If the exposure time is greater than one thirtieth of a second, there is an overlapping of the television image frame, i.e., a full frame plus part of the succeeding frame are photographed. This results in a white or over-exposed section on the film frame. It is apparent from this that the shutter speed of any camera used to photograph the television image should be precisely one thirtieth of a second or a multiple of one thirtieth such as one fifteenth or one tenth of a second if "shutter bar" is to be avoided. This rule applies to still cameras as well as to motion picture cameras. The degree of contrast between the under-exposed or over-exposed portion or banded section of the image and the correctly exposed portion decreases with the multiple increase in exposure in units of one thirtieth of a second. If the exposure is less than one thirtieth of a second, the error in exposure will be fifty per cent; for an exposure just under a fifteenth of a second, the error will be twenty-five per cent; and for an approximate tenth of a second exposure, the error will have decreased to twelve and one-half per cent. Since it is impossible to find still cameras with a shutter accuracy of the degree necessary to photograph the television image at a thirtieth of a second, better results can be obtained at the slower shutter speeds in respect to uniformity of the photograph. The problem of photographing rapid motion on the television screen with

a still camera is serious because of the distortion of motion, but no more so than in the case of direct photography.

Motion picture photography of television images was undertaken during the war to record television transmissions from cameras installed in aircraft and in guided missiles. Due to the conditions under which the television images generated by the Block[1] and Ring[2] equipment used in these tests were recorded, nonsynchronous cameras operated by batteries or spring motors at approximately sixteen or eight frames per second were used. "Shutter banding" was noticeable in these films which did not destroy their value in the studies then underway.

Some further work was done with an Eastman Cine-Special driven with a synchronous motor at fifteen frames per second. The shutter on this camera is open for 170 degrees which results in an exposure just under one thirtieth of a second. By phasing the motor drive so that the shutter opened and closed during the vertical blanking period of the television image, acceptable results without banding were obtained. When these fifteen frame-per-second recordings were projected through a standard silent projector at sixteen frames per second, no undesirable results due to change of speed were noticeable.

For a complete recording of a television program it is necessary to record the sound portion. Present day motion picture practice is to record sound at a twenty-four frame per second rate. It is desirable that recordings of television programs be capable of being played back on standard motion picture sound projectors. This necessitates the adoption of the standard twenty-four frame rate to record television programs.

Experimental Camera

A method has been devised of recording twenty-four frames of the standard thirty frame television signal.[3] The equipment now in use uses this method. A shutter driven by a sixty-cycle synchronous motor at twenty-four cycles per second is utilized. This shutter has a closing angle of 72 degrees and an opening of 288 degrees. At the twenty-four cycle per second rate these angles represent a closing time of 1/120 of a second and an opening time of 4/120 or 1/30 of a second, which is the time for one full television frame. Figure 1 shows the time sequence of such a shutter in relation to the television scanning cycle. The camera is driven by a synchronous motor from the same source of 60

[1] M. A. Trainer and W. J. Poch, "Television Equipment for Aircraft", *RCA REVIEW*, Vol. VI, No. 4, pp. 469-502, December, 1946.

[2] R. E. Shelby, F. J. Somers and L. R. Moffett, "Naval Airborne Television Reconnaissance System", *RCA REVIEW*, Vol. VI, No. 3, pp. 303-337, September, 1946.

[3] D. W. Epstein—U. S. Patent No. 2,251,786.

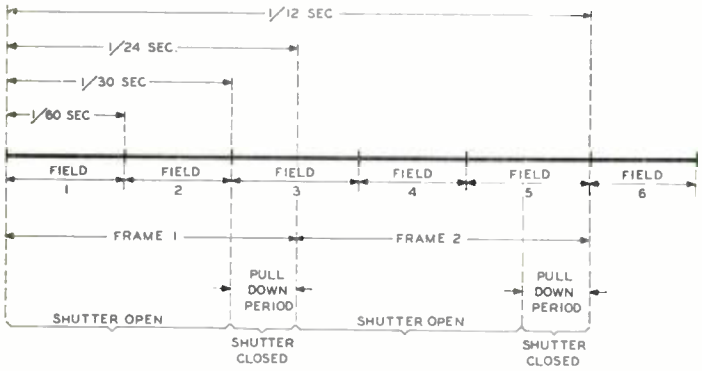


Fig. 1—Time sequence of exposure and pulldown timing of the camera in relation to the field rate of the television image.

cycle current as is used for the television synchronizing generator. If the phasing of the motor is such that the camera shutter opens on the beginning of one field, it remains open for that and the succeeding field, then closes. The shutter remains closed for 1/120 of a second or half the third field, while the film is advanced, then opens. It remains open for this last half of the third field, for the full succeeding fourth field, and for the top or first half of the fifth field. The shutter then closes at the same point that it opened in relation to the scan of

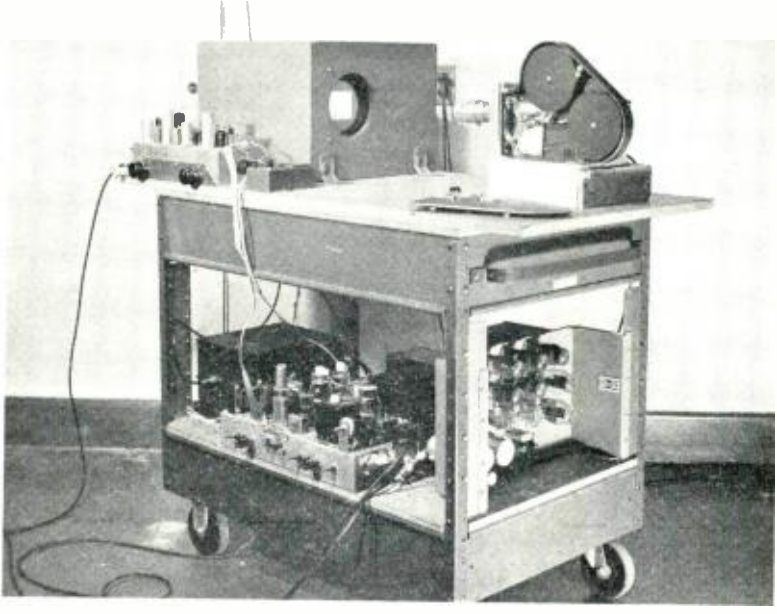


Fig. 2—"Breadboard" camera and photographic monitor.

the image. A half field later the shutter opens completing the cycle. Pull-down of the film occurs during this time that the shutter is closed.

The one hundred twentieth of a second time allowed for pulldown is considerably less than is found in standard motion picture cameras. The normal pulldown time of a standard camera is between a fortieth and a fiftieth of a second. To achieve this high rate of pulldown in the television recording camera without undue strain on the film and to maintain registration of the film image is quite an undertaking. That it has been solved satisfactorily is to the credit of the motion picture engineers.

The Eastman Kodak Company constructed a "Breadboard" camera using the foregoing principles. This camera is shown in Figure 2. The camera was capable of a 200-foot load of 16-millimeter film, allowing the recording of 5½ minutes of program time. The satisfactory operation of this camera proved the 288-degree shutter to be practical in the photography of the thirty-frame television image at a twenty-four frame rate.

Commercial Camera

The Eastman Kodak Company in cooperation with the National Broadcasting Company, encouraged by the successful operation of the breadboard camera, began the design of a commercial recording camera capable of recording a half hour of program with a 1200-foot load of 16-millimeter film. The design of this camera was quite complicated by a number of factors. It had been determined through tests with the breadboard camera that the shutter has to rotate with a low flutter rate. A slight change in angular speed of the shutter results in banding of the film image. In severe cases this banding alternates from black to white on alternate film frames. It is therefore necessary to design the shutter drive to have the utmost constancy of angular speed. This is accomplished by using an 1800 revolutions-per-minute synchronous motor to drive the shutter at the necessary 1440 revolutions-per-minute rate through a set of precision gears. Another synchronous motor of larger capacity is employed to drive the film transport mechanism and the Geneva intermittent. The two motors are kept in step during the starting and stopping periods by a phase coupling device which allows the stronger of the two motors to assist the weaker until they both reach synchronous speed. The coupling then floats so that there is no physical connection between the motors.

The shutter motor then drives the shutter independently of any varying or intermittent change of load in the camera. The armature of this motor acts as a flywheel to damp out any tendency to flutter in the dynamically balanced shutter blade.

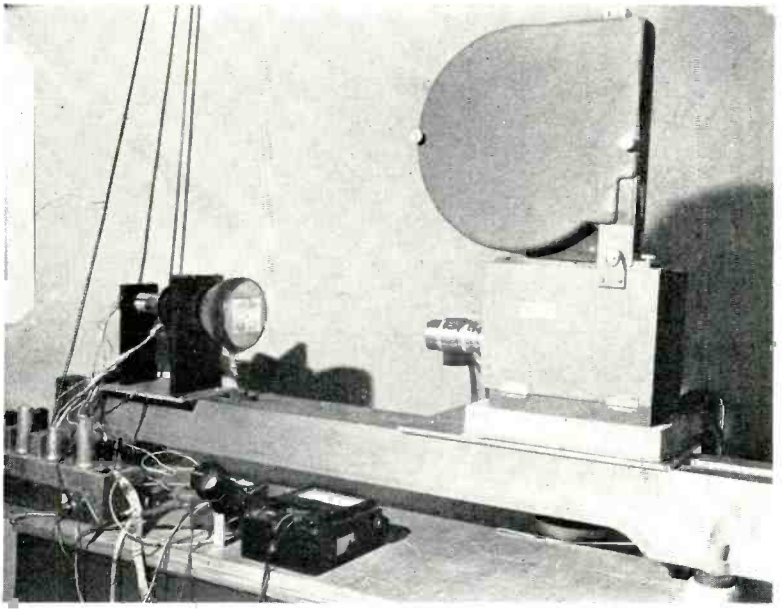


Fig. 3—The Eastman Kodak television recording camera.

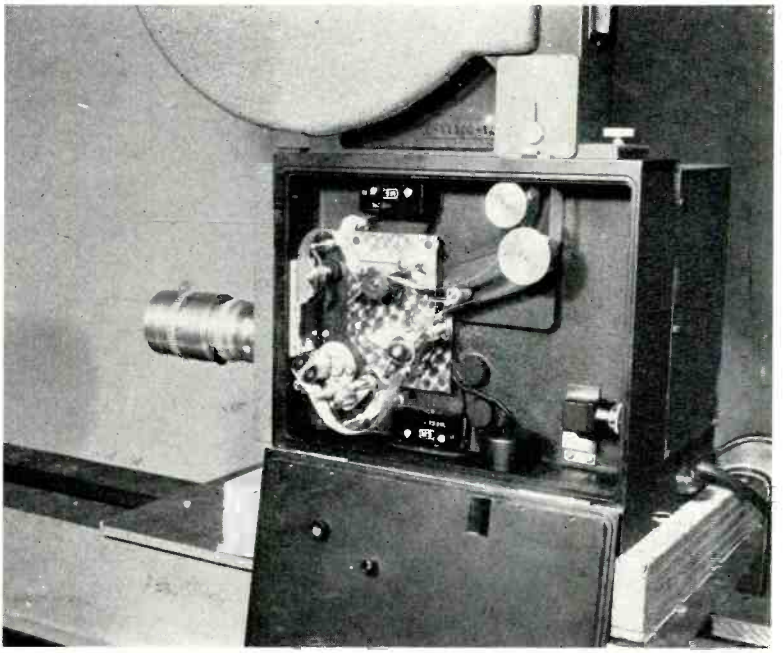


Fig. 4—Eastman Kodak camera interior showing film threading.

An eight tooth sprocket pulldown actuated by an accelerated Geneva star is employed for film pulldown. The pulldown angle of approximately 57 degrees is obtained by means of a spline-and-slot type of accelerator interposed between the constant speed shaft and the geneva driver.

The Eastman Kodak television recording camera is shown in Figures 3 and 4. Nylon is used in the film gate and pressure pad to minimize emulsion pile up, a potent source of trouble in motion picture cameras.

All friction points in the takeup side of the 1200-foot magazine are equipped with ball bearings so that take-up of film progresses smoothly from the two-inch core diameter to the full ten-inch diameter of the full 1200-foot roll. Loop loss indicators are provided and actuate microswitches in the event of loop loss, lighting a warning lamp mounted on the base of the camera.

Focussing and framing of the picture frame is done by means of a right-angle view finder equipped with a magnifying lens. This unit is snapped under the pad spring in place of the pressure pad. Visual focusing is done by means of this finder and checked by exposing film at several different settings about this visual optimum. The processed film is examined under a microscope to determine actual best focus.

The lens equipment of the camera is a 2-inch Eastman Anastigmat f1.6. Apertures of f2.0 to f2.8 are normally used to avoid sharpening the shadow of the shutter in the film plane during the opening and closing time, thereby reducing the possibility that a slight timing error may result in banding. It has been determined that at apertures of f5.6 and above, the cutoff of the shutter in the image field becomes abrupt and causes banding on the order of two or three television lines.

16- and 35-Millimeter Equipment Considerations

There are several reasons for the choice of 16-millimeter film for Kinescope recording rather than 35-millimeter. The main reason is that the cost of 35-millimeter is somewhat more than three times the cost of 16-millimeter for the same period of recording. The current quality of television images, which will undoubtedly undergo gradual refinement, is considered to be roughly equivalent to 16-millimeter home movies, although actually somewhat better with reference to contrast and detail. No marked improvement, however, is to be had by recording on 35-millimeter rather than 16-millimeter at the present time. With the use of fine grain high resolution 16-millimeter film emulsions, no loss of resolution in recording the television image is noticeable.

Fire regulations covering the use of 35-millimeter film, which apply regardless of whether the 35-millimeter film is acetate safety base or

the combustible nitrate base, are rigorous. The cost of providing space that meets these regulations for the use of 35-millimeter film is extremely high, and the changes needed in existing space are difficult to accomplish. 16-millimeter films are available only in acetate safety base which is classified by the Underwriter Laboratories as having a safety factor slightly higher than that of newsprint. The use of 16-millimeter films, therefore, are not restricted by fire regulations. It should be noted that in New York City these restrictions apply to space in which equipment capable of operating with 35-millimeter film is installed, so in order to forestall trouble, all equipment should be single purpose 16-millimeter equipment rather than dual purpose 35- or 16-millimeter equipment.

Another factor in the choice of 16-millimeter film is the high cost of 35-millimeter projection equipment. Most television stations are providing projection facilities for 16-millimeter film only for this reason. In order to service these stations with syndicated programs photographed from the kinescope, 16-millimeter prints will be needed.

Kinescope Phosphors and Film Spectral Characteristics

The spectral sensitivity of the film emulsion can be matched to the phosphor spectral characteristic for the greatest actinic efficiency. There are three general classifications of film emulsions in terms of their spectral characteristics:

1. *Panchromatic,* sensitive from the ultraviolet (4000 Angstrom Units (A°)) through the red (7000 A°). The spectral response of these emulsions correspond approximately to that of the eye and so are generally used for direct photography;

2. *Orthochromatic,* sensitive from the ultraviolet through green, (5700 A°) is used in direct photography where it is desirable to reduce the red sensitivity;

3. *"Ordinary"* or non-color sensitive emulsions, nonsensitized, responding to the ultraviolet and blue portions of the light spectrum. This type of emulsion is used in coating films and papers generally employed in making positive prints from negatives. In 16-millimeter form it is economical in comparison to the panchromatic and orthochromatic types. Another advantage is the ease of handling as relatively bright safelights may be used.

To match these film characteristics, kinescope phosphors are available with light output ranging from the ultraviolet through the entire visible spectrum. Three types of phosphors in common use in television techniques are as follows:

1. *P1, green fluorescence,* commonly used in oscillographic work.

It is the most efficient visually, but has poor actinic efficiency.

2. *P4, white fluorescence*, used for black and white reproduction of television images in most home receivers. This phosphor has a high output in the blue and green portions of the spectrum, but is down in the red. It has the advantage in kinescope photography that picture quality is most readily judged visually.

3. *P5 and P11*, these two phosphors are blue with high ultraviolet output. Photographically, they are very efficient. There is the difficulty in using a blue phosphor in judging the quality of image visually, due to the fact that the human eye has a low response in the blue region and cannot evaluate the quality of the ultraviolet component of the image light output at all.

Tests have been made on the P11, zinc sulphide, phosphor as to the relative actinic efficiency to the panchromatic, orthochromatic, and "ordinary" non-color sensitive emulsions. A Weston exposure meter was used to determine the light output of the aluminized P11 phosphor kinescope. A series of exposures were made of the image on the tube and the correct exposure as judged by visual inspection of the negative was chosen. The Weston meter was then set with this exposure data to find the Weston rating of the type of film for the P11 light output. This rating was then compared with the Weston rating for daylight as given by the manufacture of the film used. It was found that the exposure required for the P11 phosphor image for panchromatic film was one sixth that required for white light of equal intensity, for orthochromatic it was one twelfth, and for the ordinary or non-color sensitive stock, the exposure ranged from 1/16 to 1/32 of the exposure needed for white light.

With these facts it is apparent that for recording of television images a zinc sulphide, blue-fluorescing screen is desirable since it makes possible the use of high resolution, low cost, positive type of film stocks.

Resolving Power Of Films

Present day television systems operating on a 30-frame, 525-line standard, are capable of resolving 483 lines in the vertical direction, and inside the studio plant, before being transmitted on the 4.5 megacycle channel of the radio frequency transmitter, of resolving over 600 television lines horizontally.

Manufacturers of photographic film rate the resolution of their products in lines resolved per millimeter. By dividing the 7.2-millimeter height of the 16-millimeter frame, into 483 lines and dividing the result by two to convert from television lines to photographic lines per millimeter, it is found that in order to resolve the television scan-

ning lines, a resolving power of better than 33 lines per millimeter at contrast ranges below 1 to 10 is required. To resolve 600 television lines on 16-millimeter film, the emulsion must have a resolving power of 42 lines per millimeter.

The subject contrast of the test charts used to determine film resolution is of the order of 1 to several hundred times. Super X, a panchromatic emulsion, is rated by the manufacturer at 55 lines per millimeter. In television terms there would be resolved 792 television lines in the 16-millimeter frame. It might be thought that such a film would be suitable for photography of the television image. This is not so, for this film, used in photography of the television image, does not fully resolve the scanning lines.

The reason for this discrepancy lies in the different method used by the manufacturer to rate the film. As pointed out above, the resolution is determined by photographing a chart that has a contrast ratio between the black lines and the white spaces of several hundred times. Resolution is then determined by the point at which these lines are barely resolved on the film, or at a point where the contrast ratio is slightly greater than unity.

The resolving power of film, of television pickup tubes, and of image reproducing tubes, falls off with decrease in the subject contrast of the test target. A film rated at 55 lines per millimeter at a subject contrast of several hundred times may have only a resolving power of twenty lines per millimeter when the contrast is in the order of 1 to 10.

It has been pointed out in the literature[4,5] that the television image may have contrast ratios of fifty times in large areas, falling off to contrast ranges less than 1 to 10 in fine detail. A film emulsion rated at 90 lines per millimeter under normal test conditions has the necessary resolving power at the lower contrast ranges to resolve the required 42 lines per millimeter.

Suitable emulsions with resolving powers in excess of 90 lines per millimeter are to be found in the fine grain sound recording and print stocks. Both low and high gamma emulsions are available. For recording a positive image on the kinescope, the low gamma variable density type of emulsion is used. When recording from a reversed negative image on the cathode-ray tube, a high-gamma variable area or print type of emulsion is used.

Some improvement in the quality of resolution of the photographic

[4] A. Rose, "A Unified Approach to the Performance of Photographic Film, Television Pickup Tubes, and the Human Eye", *Jour. Soc. Mot. Pic. Eng.*, Vol. 48, No. 10, October, 1946.

[5] O. H. Schade, "Electro-Optical Characteristics of Television Systems: Introduction; Part I—Characteristics of Vision and Visual Systems", *RCA REVIEW*, Vol. IX, No. 1, pp. 5-37, March, 1948.

image is made by the use in the video feed of a phase and amplitude equalizer.* The increase in amplitude or contrast of fine detail that can be obtained by boosting the high frequencies in the kinescope image compensates somewhat for the normal falling off of fine detail contrast in the film image. Phase correction is used to reduce the transient white that follows black. The amount of high frequency peaking that can be used is limited by the noise component of the video signal.

Exposure Of Film

Present day aluminized kinescopes operating at high second-anode voltages in the order of 27 kilovolts are capable of brightness ratios of several hundred times in large areas. Most films used in normal photography can handle a range of this order. However, to make full use of such a latitude requires a very accurate exposure. Generally the object brightness range under controlled lighting arrangements never exceeds a ratio of 1 to 30 in direct photography. It is, therefore, necessary that the contrast of the photographic kinescope be maintained within the limits set by the latitude of the particular film used and that the brightness range be set to duplicate the 1 to 30 ratio used in direct photography so as to duplicate the printing contrast of a normal negative.

This is most conveniently achieved by the following method. A plain raster is used on the kinescope such as would be obtained by the use of the blanking signal or pedestal without picture modulation. The brightness of this raster is varied by means of the video gain control or kinescope grid bias control. The beam current of the kinescope is measured by means of a microammeter. Since the light output of the tube is dependent on the watts input to the screen, the measure of beam current affords a measure of the brightness of the tube. Film is exposed to this raster at beam currents varied by steps. The density of the film processed as a normal negative is measured and plotted against the logarithm of the beam current. A normal negative developed to a gamma of 0.65 which has been exposed to an object with a brightness range of 1 to 30 or in logarithmic units, a range of 1.5, should have a density range from 0.25 in the shadows to approximately 1.4 in the highlights. The change in beam current necessary to produce such a range on the kinescope can be read from the plot of the log beam current and film density. The average brightness of the kinescope with picture then would be set by using a beam current that produces a density in the middle of the above range. The video signal is adjusted

* Designed by E. D. Goodale, Television Development, NBC. It is planned that a paper on the equalizer will be published in the September 1948 issue of *RCA REVIEW*.

to a level that will put the blanking level of the composite signal just at visual cutoff of the kinescope. A picture signal judged to have an a-c axis of 50 per cent should be used for this adjustment. This method is largely empirical, but, with experience on the part of the operator, can be made to give consistent results.

Processing and Printing of Kinescope Film

A number of tests have been made in cooperation with the film manufacturers on the processing and printing of films exposed to the kinescope. Both reversal processing and negative processing of the original film were tried. Results show that standard processing methods result in optimum picture quality. Negatives exposed to television images originating in iconoscope cameras are developed to a gamma of 0.7 as determined by a standard IIB sensitrometric test. Film of orthicon pickups gives best results when processed to approximately 0.6 to 0.65. These are interim values as tests on the processing of films have not been completed.

Printing is done according to standard motion picture laboratory practice. Step printing in which the print stock and negative are exposed to the printing light a frame at a time is preferred over continuous printing, where the negative and print stock run past an illuminated slit at a continuous speed. There is a sufficient amount of slippage between the negative and print stock in the continuous printing process to degrade the resolution of the television image. Contrary to the opinion held by many workers, the fact that the film image of a television image is poorer in resolution than in the case of direct photography does not mean that less care in the handling of the film in printing and in projection can be used. The fact is that the utmost care must be taken to maintain the original quality inherent in the film negative throughout the printing process and in the projection of the resulting print.

Films of iconoscope programs can be printed at one printer light setting: i.e., the densities and contrast range of the film resulting from the recording of the outputs of a number of iconoscope cameras does not change sufficiently to warrant changes in the intensity of the printing light.

In film recordings of programs picked up by orthicon cameras the picture negative must be timed for printing. There is considerable difference in the contrast range between different orthicon cameras. Light changes in the order of 100 per cent are sometimes required when a switch between cameras occurs.

This is an undesirable condition as the timing of negatives is an

expensive, time consuming, procedure. The cure for this situation is in better control of the output levels of the various orthicon cameras in the studio. Much of this change can be charged to the fact that the spectral characteristics of the orthicon vary from tube to tube. An orhicon with excessive infrared response has a different tonal graduation as compared to an orthicon with no or little response in this region. If the spectral characteristics of the orthicon can be standardized within closer limits than is now done, much of this timing difficulty in the film recordings may disappear.

In kinescope recordings meant for retransmission through the television system a print gamma of 2.2 and a maximum density of 2.4 are recommended. Further tests may show the desirability of changing these recommendations, but to date the best results in the televising of release prints have been obtained under such conditions.

Emulsion position in the final print is of great importance in television because films may be spliced with other films for special purposes. The use of a non-standard emulsion position requires a change of focus in the film projector when interspliced with films using standard emulsion position. This would require the constant attention of the projectionist to maintain optimum focus throughout the spliced film, therefore it is advantageous to insist upon a standard emulsion position for all film to be used in television. The Society of Motion Picture Engineers' standard for 16-millimeter film is emulsion "toward the screen."

In the recording of television images there are several methods of obtaining the final print:

A. the use of reversible film stock in photographing a positive cathode-ray-tube image (A dupe negative is made of this material and prints are made from this negative. The final prints then have standard emulsion position.);

B. the photography of the cathode-ray-tube image using high contrast positive stock and a negative kinescope image resulting in a positive print from which dupe negatives are made for prints having standard position; and

C. the use of a positive image, photographing with a negative type of film from which final prints are made, resulting in a non-standard emulsion position (However, by reversing the direction of horizontal scanning the original negative may be made to have the same emulsion position as that of a dupe negative. Prints made from this negative, then have a standard emulsion position.)

Other factors must be considered in determining the method of recording. Where it is expected that a great number of prints will be

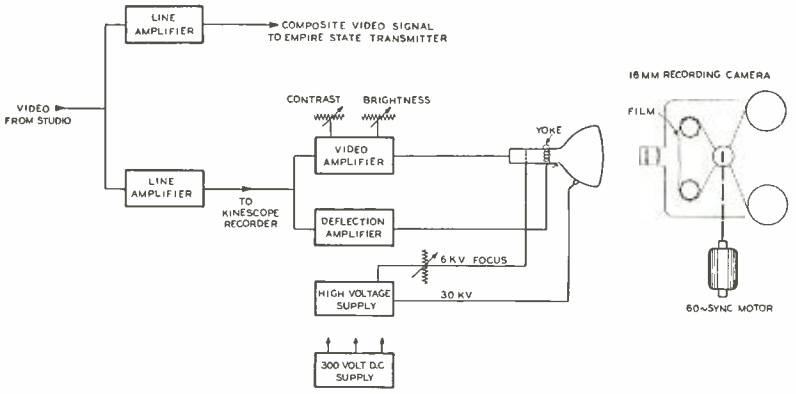


Fig. 5—Kinescope photographic monitor block diagram.

required, methods A or B would be desirable because of the protection of the original material.

Photographic Monitor

The various sizes of cathode-ray tubes may be used for television recordings. In order to insure adequate exposure on the fine grain postive type emulsions, the voltages used with these tubes should be on the order of 20,000 to 30,000 volts. All other factors being equal, such as relative spot size, uniformity of focus over the picture area, brightness and contrast, the smaller tubes offer advantages in the size of the photographic setup. A five-inch kinescope with a flat screen, aluminized P11 phosphor, and the same general type of electron gun

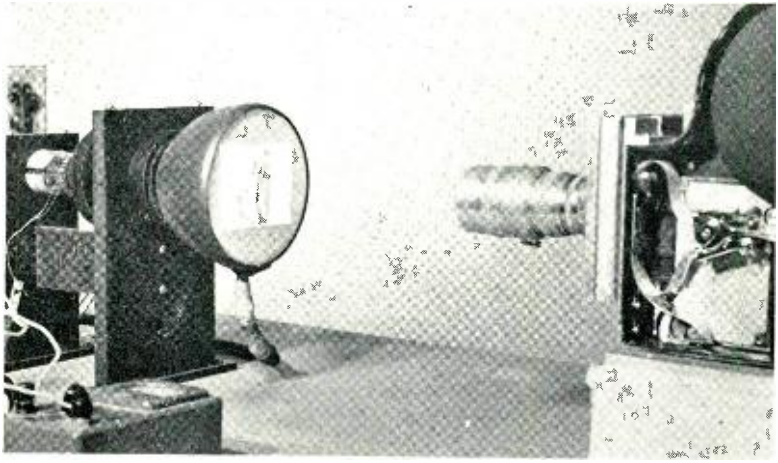


Fig. 6—Close-up of the kinescope with cover removed. (On the tube face is an actual television image of a baseball pickup.)

construction as is used in the 5TP4, gives adequate resolution, contrast and brightness.

A block diagram of the experimental photographic monitor setup is shown in Figure 5 and a photograph of the unit in Figure 6. The video amplifier is flat to eight megacycles, and down ten per cent at ten megacycles. A radio-frequency high voltage power supply delivers 29 kilovolts to the second anode of the kinescope. It should be mentioned that these tubes with the P11 phosphor have a high X-ray output at these voltages as compared with the 5TP4 at the same voltages, and that it is necessary to use more shielding than is evident in the photograph to safeguard personnel.

The deflection circuits follow conventional design. Provision is made to switch the direction of horizontal sweep in order to obtain a "dupe" negative emulsion position.

The camera is mounted on one end of a five-foot lathe bed isolated from the table with shock absorbers to absorb camera vibrations. The five-inch kinescope is mounted at the other end of bed. Isolation of the tube from vibrations transmitted through the lathe bed is accomplished by means of rubber shims around the deflection yoke and around the neck of the tube. The deflection chassis, the radio-frequency power supply, and the 285-volt regulated power supply are mounted on the lower shelf.

Sound Recording

Recording of the sound portion of a television program is done with standard 16-millimeter sound-on-film recording equipment operating with a synchronous drive at 24 frames per second. A switch operates both the motion picture cameras and the sound recorders simultaneously. Synchronizing marks are momentarily injected into both the sound and video channels by means of a remote switch. The marks are produced by a 420-cycle tone generated by an oscillator. The tone in the video channel produces bars in the kinescope image and an easily identified modulation of the sound track. The picture bars and the track tone are lined up for synchronization of the picture negative with the sound track negative.

Conclusion

A practical method of recording television programs, both sight and sound, has been developed. The recordings can be retransmitted through the television system with acceptable results. As the quality of the television image improves, the quality of recording in kinescope pho-

tography will be improved to a degree where the average viewer will be unable to tell if the program he is seeing is "live" or "canned." Certainly, these recordings made from wide-band channels available within the studio plant, unlimited by the restricted band width of the radio-frequency transmitter, will compare favorably with live programs as reproduced on the home television screen.

General References

G. L. Beers, E. W. Engstrom and I. G. Maloff, "Some Television Problems from the Motion Picture Standpoint", *Jour. Soc. Mot. Pic. Eng.*, Vol. 32, No. 2, February, 1939.

C. F. White and M. R. Boyer, "A New Film for Photographing the Television Monitor Tube", *Jour. Soc. Mot. Pic. Eng.*, Vol. 47, No. 2, August, 1946.

BARIUM TITANATE AND BARIUM STRONTIUM TITANATE RESONATORS*

By

H. L. Donley

Research Department, RCA Laboratories Division,
Princeton, N. J.

Summary—*Values of the equivalent circuit elements are given for polarized polycrystalline barium titanate and barium strontium titanate resonators vibrating normal to the direction of the polarizing and radio frequency fields. An increase in the overall or effective piezoelectric constant with increasing polarizing field is observed for these resonators. At 25 volts per mil an average piezoelectric constant of about 200 × 10⁻⁸ e.s.u. (electrostatic), a Q of about 80 and an electromechanical coupling factor of 0.2 is found for the barium titanate resonators. Lower activity but higher Q results when the strontium content is increased in the barium strontium titanate composition type resonators.*

The change in resonant frequency of these longitudinal mode titanate resonators follows the change in length in a linear fashion. A frequency constant of 211 kilocycle-centimeters is found for the barium titanate resonators which constant increases to 275 kilocycle-centimeters as the strontium content increases to 30 per cent in the barium strontium titanate composition.

The piezoelectric activity of these polarized polycrystalline titanate resonators essentially ceases as the Curie temperature of the particular composition is approached. A large temperature coefficient of frequency of roughly 1 part in 400 per degree Centigrade at room temperature is observed for the barium titanate resonators.

Introduction

IT IS WELL known that piezoelectric elements are obtained by cutting these elements from a single crystal of the piezoelectric substance in definite directions and orientations depending upon the characteristics desired. Barium titanate and barium strontium titanate, although in polycrystalline form, also display an average or induced piezoelectric activity when a direct current polarizing field is applied, apparently a consequence of the piezoelectric properties of barium titanate.[1] After an initial drop in activity upon removal of the polarizing field the particular titanate sample remains apparently permanently piezoelectric as indicated by observations made over a period of a year. Furthermore, upon removal of the polarizing field and initial charge the

* Decimal Classification: R214.2.

[1] S. Roberts, "Dielectric and Piezoelectric Properties of Barium Titanate", *Phys. Rev.*, Vol. 71, pp. 890-895, June, 1947.

titanate sample for a given applied thrust shows a reversal in sign of the charge developed as the sample is reversed analogous to an x-cut quartz crystal. Also, in analogy to magnetic phenomena, the piezoelectric activity of these titanates follows somewhat a hysteresis loop as the magnitude and direction of the polarizing field is altered. The results obtained depend upon the maximum polarizing field employed as well as the time of exposure to this field. Resonance vibrations resulting from this apparent piezoelectric behavior can be excited preferably by superimposing a small alternating voltage on a larger direct current bias or applying the exciting voltage to a previously polarized sample. These resonance phenomena tend to disappear at the Curie or critical temperature at which the dielectric constant is maximum for the particular barium strontium titanate composition.

The rather sharp resonances observed for these polycrystalline materials immediately suggest their possible practical application as resonators. The present experimental study was undertaken to determine the resonant frequencies and the equivalent electrical constants of these titanate resonators as well as the dependence of these constants upon bias voltage and temperature.

With the assumption that polarized polycrystalline specimens of $BaTiO_3$ are piezoelectric in origin then an overall or average piezoelectric constant for the material can be determined from the measured equivalent electrical constants of resonators of this material.

Form and Mounting of Samples

The samples tested were in the form of long, narrow, thin plates to assure the wide separation of the frequencies corresponding to the various possible modes of motion as well as minimizing the coupling between these modes. The plates, all 20 mils thick, included fired-on silver electrodes[2] which covered in general the major surfaces of the plate. The plates were cut by means of a bonded carborundum wheel to the desired length and width from a larger plate which had been fired to the proper thickness.

The plate was mounted at its center in a simple knife bar or moderate tension spring-type holder which provided both a means of connection to the fired-on silver electrodes and a means for exciting the plate into lengthwise vibration.

The direct current bias was applied parallel to the thickness dimension of the resonator and superimposed on the radio-frequency voltage.

[2] The firing of the ceramic bodies and their processing is adequately discussed elsewhere. See, for example, A. von Hippel, R. G. Breckenridge, F. G. Chesley and L. Tisza, "High Dielectric Constant Ceramics", *Ind. Eng. Chem.*, Vol. 38, pp. 1097-1109, November, 1946.

Experimental Method

If one assigns to these polycrystalline titanate resonators an effective or overall piezoelectric constant, d, associated with the particular mode of motion studied then this constant may be determined from the values of the equivalent circuit elements of the vibrating plate. In the neighborhood of the fundamental resonance of the plate, with electrodes adherent to the surface of the plate, an equivalent circuit for the vibrating plate consisting of a series R, L, C in shunt with a capacity, C_o, may be assumed.[3] In e.s.u. the elements of this equivalent circuit have the following values for lengthwise vibrations.

$$C_o = e\,bl/4\pi\,t \qquad (1)$$

$$L = \frac{s}{8Y^2 d^2}\,\frac{l\,t}{b} \qquad (2)$$

$$C = \frac{8Yd^2}{\pi^2}\,\frac{l\,b}{t} \qquad (3)$$

$$R = \frac{\pi^2 s\,r}{8Y^2 d^2}\,\frac{t}{b\,l} \qquad (4)$$

where l, b, t are respectively the length, breadth, and thickness of the resonator in centimeters.

e = dielectric constant, s = density,

Y = Young's modulus for the material,

d = the effective or average piezoelectric constant for the material,

r = a "viscosity" or mechanical loss factor for the material.[4]

From the above relations the resonant frequency for the plate in lengthwise vibration is given by,

$$f_o = \frac{1}{2\pi\sqrt{LC}} = \frac{1}{2l}\sqrt{\frac{Y}{s}} \qquad (5)$$

The element values for these titanate resonators were found by a study of the crevasse cut in the response curve of a parallel resonant circuit, the well known inverted resonance curve method as used by Watanabe[5] and others.

A battery operated signal generator provided a good constant fre-

[3] W. G. Cady, PIEZOELECTRICITY, McGraw-Hill Book Co., New York, N. Y., 1946.

[4] This follows Cady's assumption that the mechanical loss is only because of the viscosity of the material. See W. G. Cady, "The Piezoelectric Resonator", *Proc. I.R.E.*, Vol. 10, pp. 83-114, April, 1922.

[5] Y. Watanabe, "The Piezoelectric Resonator in High Frequency Oscillation Circuits", *Proc. I.R.E.*, Vol. 18, pp. 695-717, April, 1930.

quency source. The Q measurements were made in the neighborhood of 250 kilocycles where frequency differences were accurately measured by means of a mixer amplifier whose output, consisting of the beat frequency between a 250-kilocycle quartz crystal oscillator and the signal generator frequency, fed an audio frequency meter.

To avoid the pronounced non-linear effects possible with these materials[6], the radio-frequency voltage applied to the plates was kept at 1 volt for the measurements.

Experimental Results

Similar to quartz resonators, a resonator of polarized barium titanate displays a whole series of resonant frequencies. For example, a particular sample of $BaTiO_3$ with a polarizing field of 13 volts per mil with $l = 0.384$ inch, $b = 0.173$ inch, $t = 0.02$ inch showed approximate low frequency resonances at 223, 428, 440, 490, 539, 614, 710, 788, 918 kilocycles with the lowest resonance being the most prominent. When the width (b) dimension of this same resonator was decreased to 0.104 inch the fundamental frequency become 222.2 kilocycles with additional responses at 707, 830, 832, 867, 977 and 1168 kilocycles showing that the width dimension was instrumental in shifting the secondary frequency spectrum, but had little if any influence upon the fundamental frequency. No attempt has been made to identify the frequencies of the various possible modes of motion. The lowest observed resonant frequency was taken as the fundamental frequency for the longitudinal mode and the measurements made were based upon this prominent resonant frequency.

Typical results of the effect of an increase and decrease in direct current bias on the various equivalent circuit element values for a particular fully plated barium titanate resonator is shown in Figure 1. The remanent piezoelectric activity is brought out in Figure 1 by the fact that the maximum piezoelectric constant, d, corresponding to maximum polarizing field remains even after the polarizing field is reduced to zero. Both the equivalent series resistance, R, and the Q of the resonator decrease with increasing bias although R decreased some 65 times as compared to about 2 to 1 drop in Q in going from 3 to 25 volts per mil for the results shown in Figure 1. The observed decrease in Q with increasing bias means, of course, an increase of mechanical losses at higher polarizing fields and the more rapid drop in R means an increase in piezoelectric activity with increasing polarizing field

[6] H. L. Donley, "Effect of Field Strength on Dielectric Properties of Barium Strontium Titanate", *RCA REVIEW*, Vol. VIII, No. 3, pp. 539-553, September, 1947.

which outweighs the increase in mechanical losses. Presumably the reason for the low Q is the inherent hysteresis effect in this material as evident in Figure 1.

Within about ± 1 per cent, the frequency constant ($f_o l$ in Equation (5)) for the fully plated barium titanate resonator measured was found to be 207 kilocycle-centimeters which gave a value of Young's modulus of 9.5×10^{11} dynes per square centimeter using the published

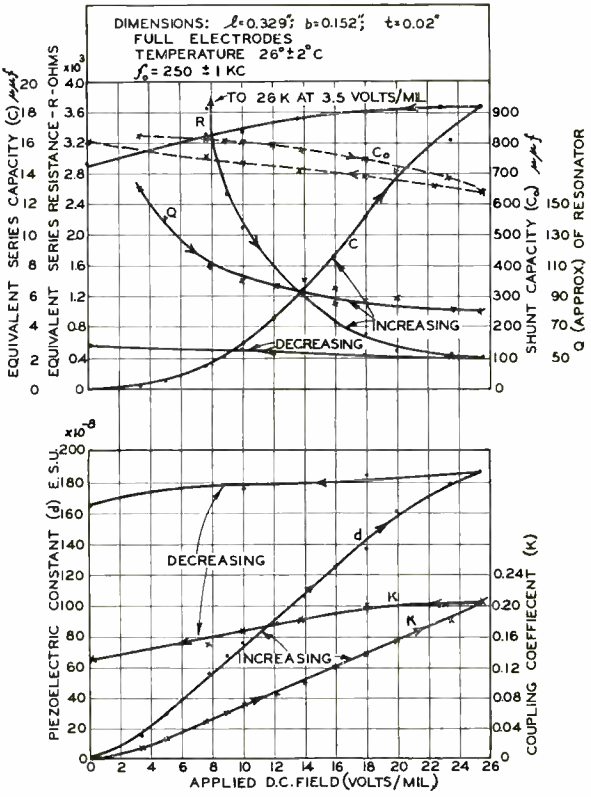


Fig. 1—Variation of constants of barium titanate resonator with polarizing field.

density value[2] of 5.52 for $BaTiO_3$. Various resonators with electrodes covering only the central portions (about $\frac{1}{8} \times \frac{1}{8}$ inch) showed about a 3 to 4 per cent increase in frequency constant with f_o approximately at 250 kilocycles.

The values of the apparent piezoelectric constant, d, determined from the measured resonator constants is also shown in Figure 1 for increasing and decreasing polarizing field. The increase of piezoelectric activity for $BaTiO_3$ with an increase in polarizing field as

shown in Figure 1 has also been observed by static tests, where for a given pressure the direct current voltage developed increased with increasing polarizing field.

As shown by Mason[7] the figure of merit of a resonator is determined by the capacitance ratio $(C_0/C = h)$, and h is related to the electromechanical coupling coefficient k by the following formula,

$$\frac{C_0}{C} = \frac{\pi\, e}{32 Y d^2} = h = \frac{\pi^2}{8}\left(\frac{1 - k^2}{k^2}\right) \tag{6}$$

In Figure 1 are also plotted values of k calculated from Equation (6) using the measured values of C_0 and C also shown in Figure 1. The

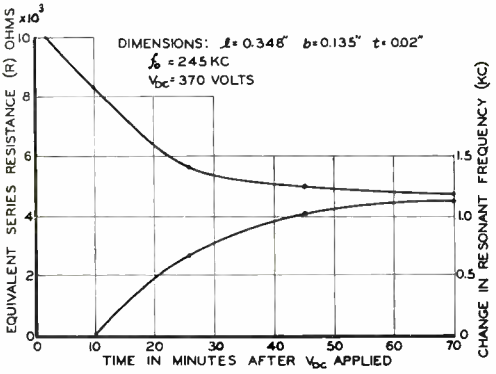


Fig. 2—Effect of length of polarizing time on barium titanate resonator.

value of k for $BaTiO_3$ at 25 volts per mil is about twice that for the longitudinal mode in quartz.

In addition to the increase of piezoelectric activity with increasing bias this activity is also dependent to a lesser extent upon the length of time the bias is maintained as shown, for example, in Figure 2 where the equivalent series resistance, R, being proportional to $1/d^2$ finally reached a constant value after approximately an hour.* A slight increase in resonant frequency, less than 0.5 per cent is also to be noted. Here again, static tests have also shown an increase in piezoelectric activity with duration of polarizing field.

[7] W. P. Mason, "Properties of Monoclinic Crystals", *Phys. Rev.*, Vol. 70, pp. 705-728, November 1, 1946.

* Judging the piezoelectric activity from the behavior of R neglects the small change in Y and the increase in losses with time (see Equation 4). Because of changing R the accurate determination of Q was difficult but a slight decrease in Q was noted in this case.

In Figure 3 is presented evidence to show that the piezoelectric activity of $BaTiO_3$ essentially ceases when the Curie temperature is reached. This is shown in Figure 3 by the very rapid increase in the equivalent series resistance, R, of the resonator as the Curie temperature (120-125 degrees Centigrade for $BaTiO_3$) is approached. The position of the Curie point on the temperature scale is shown approximately by the peak in shunt capacity, C_o, in Figure 3. However, the resonant effect does not cease abruptly as the Curie temperature is passed, as is shown by the plot of f_o vs. temperature, although a very rapid increase in resonance frequency occurs at the Curie temperature.

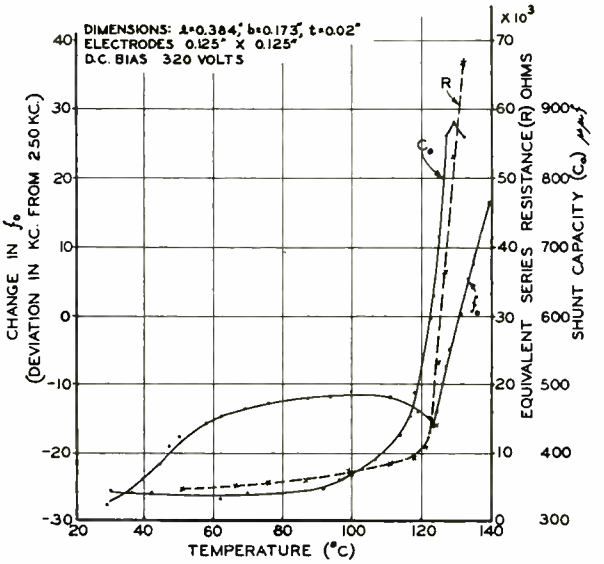


Fig. 3—Effect of temperature on barium titanate resonator.

Resonators of various compositions of barium strontium titanate were also studied. Since strontium titanate is non-ferroelectric[2] and presumably non-piezoelectric, an increase in the strontium content of the barium strontium titanate composition yielded resonators of lower piezoelectric activity than those made from 100 per cent $BaTiO_3$. Static tests also confirmed this decrease in activity as the strontium percentage in the titanate composition was increased. In fact, piezoelectric activity, as revealed by both static and dynamic measurements, ceased entirely at room temperature for the 50/50 (Ba/Sr)TiO_3 composition. Although the piezoelectric constant progressively decreased for an in-

crease in strontium content of the barium strontium titanate composition the Q's increased accordingly. Table I summarizes the observations made on different plates of the four compositions studied where only the central portions of the resonator were equipped with fired-on silver electrodes about $\frac{1}{8} \times \frac{1}{8}$ inch. With the limited electrode area lower d values and higher Q values were obtained than would have been the case for full sized electrodes. In fact, a comparison of the values given in Table I for 100 per cent $BaTiO_3$ with the Q and d values in Figure 1 show that the full electrode resonator gave a four fold increase in d at 20 volts per mil and one half the Q of the resonators with the limited electrode area.

When the length of these long, narrow, thin titanate resonators was gradually shortened, it was found that the resonant frequency corresponding to a given length accurately (within ½ per cent) followed an inverse linear relationship with the resonator length. Such a plot showing that the resonance frequency is inversely proportional to the length is given in Figure 4 for measurements taken on three different plates of the composition 80/20 $(Ba/Sr)TiO_3$ over the frequency range 36 to 317 kilocycles. The electrodes used for these measurements covered only the central portion of the plate.

Similar to the 100 per cent $BaTiO_3$ resonators, a rapid increase in R and f_o took place as the Curie temperatures of the 70/30 and 80/20 $(Ba/Sr)TiO_3$ composition resonators was passed. Here again, the resonance property persists for a considerable temperature range above the Curie temperature (65 degrees) as revealed in Figure 5 for the 80/20 $(Ba/Sr)TiO_3$ composition.

Table I—Some Approximate Average Values Observed for Barium Strontium Titanate Resonators at 20 Volts per Mil D-C Field with Electrode Area About ⅓ Total Area of Plate

Composition		Frequency Constant		Density[2]	Y	d	Q
% $BaTiO_3$	% $SrTiO_3$	kc. in.	kc. cm.	S	dynes/cm.^{2} $\times 10^{11}$	e.s.u. $\times 10^{-8}$	
100	0	83	211	5.52	9.7	40	150
80	20	95	241.3	5	11.7	30	350
70	30	108	274.5	5	15	15	600
50	50	No Activity					

Discussion of Results

The variation of piezoelectric constant with polarizing field (25 volts per mil maximum) as shown in Figure 1 for the barium titanate longitudinal mode resonator agrees with the recent results of Mason[8] for the electrostrictive constant of the radial mode of barium titanate disks as a function of polarizing field. His results show a 40 per cent increase in electrostrictive constant in going to 30 kilovolts per centimeter from 10 kilovolts per centimeter, the highest field used here to get the values given in Figure 1. Also, a 2 to 1 increase in constant is observed by Mason[8] when the thickness mode is used.

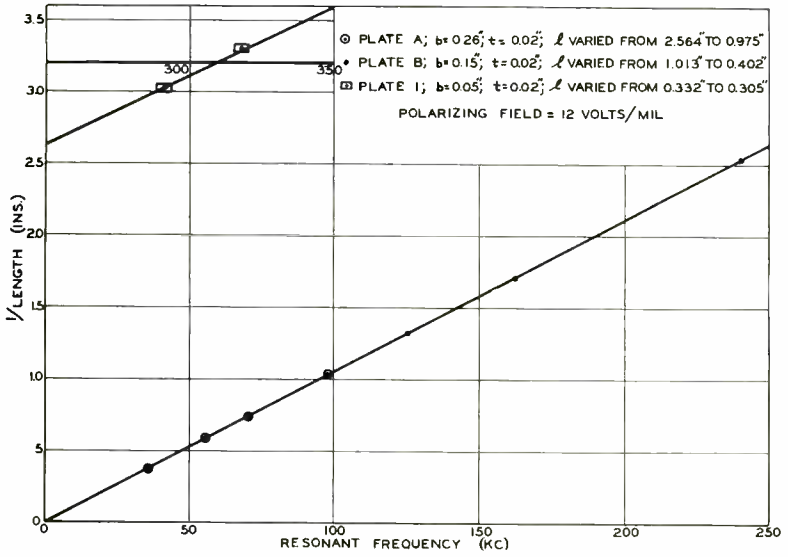


Fig. 4—Relation between resonant frequency and length for longitudinal mode 80/20 (Ba/Sr) TiO_3 Resonators.

The effect of increase in piezoelectric activity with length of time of application of the polarizing field as shown qualitatively in Figure 2 and magnitude of field is in general agreement with the observations of Cherry and Adler.[9]

The large temperature coefficient of frequency of roughly 1 part in 400 per degree Centigrade at room temperature as well as the low Q's for the polarized barium titanate resonators as compared to quartz

[8] W. P. Mason, "Electrostrictive Effect in Barium Titanate", *Phys. Rev.*, Vol. 72, pp. 869-870, November 1, 1947.

[9] W. L. Cherry, Jr. and Robert Adler, "Piezoelectric Effect in Polycrystalline Barium Titanate", *Phys. Rev.*, Vol. 72, pp. 981-982, November 15, 1947.

resonators rather limits their use as resonator elements in most practical applications. Although the Q's increase for resonators of the barium strontium titanate compositions the poor thermal characteristics still remain.

On the other hand, on the basis of data given here and from unpublished results of this laboratory, barium titanate has interesting possibilities as a pick-up device. In this essentially high impedance application the lower reactance offered by the barium titanate ceramic

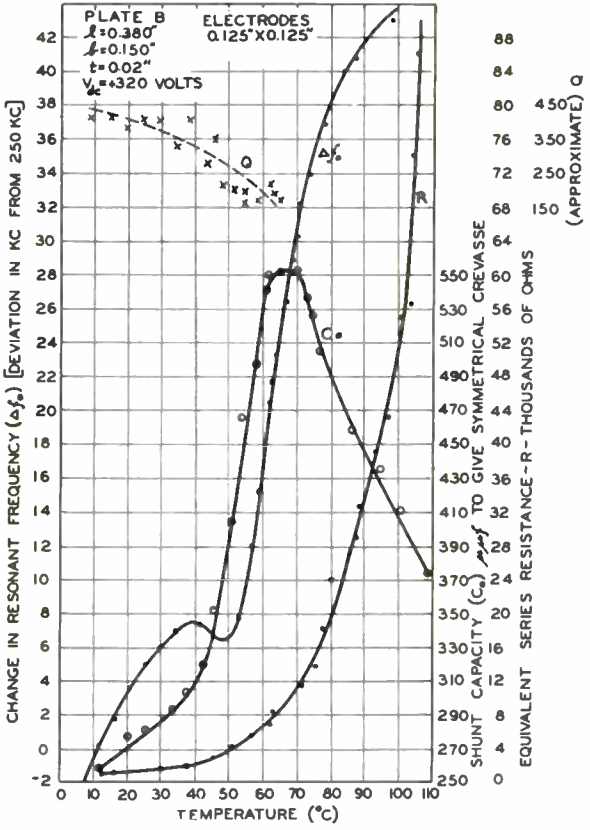


Fig. 5—Effect of temperature on characteristics of 80/20 (Ba/Sr) TiO_3 Resonator.

over the universally used Rochelle-salt crystal pickup is a distinct advantage. In particular, Rochelle-salt has a dielectric constant of about 100 and that of barium titanate is about 1200 at room temperature which gives barium titanate a tenth the reactance of Rochelle-salt at a given frequency and for the same size units. However, because the piezoelectric activity of Rochelle-salt is large we assume a perfect

transfer of mechanical to electrical energy or electromechanical coupling factor of unity as compared to at least 0.2 for polarized barium titanate, a 5 to 1 ratio in favor of Rochelle-salt. The 10 to 1 reactance ratio favoring barium titanate offsets this latter advantage so that from a most conservative estimate the barium titanate has a figure of merit roughly twice Rochelle-salt for this particular application. Even on the basis of equal performance, the fact that this low-cost, easily manufactured material maintains its apparent piezoelectric activity over a wide temperature range extending to roughly 100 degrees Centigrade as compared to the upper useful temperature limit of 45 degrees Centigrade for Rochelle-salt decidedly favors the use of the barium titanate in this particular application.

ACKNOWLEDGMENT

Appreciation is extended to C. Wentworth who supplied the fired samples tested.

SUNSPOTS AND RADIO WEATHER*†

By

AUDREY ARZINGER,‡ H. E. HALLBORG‡ AND J. H. NELSON#

Summary—A new criterion for the positioning and correlation of sunspots is shown with respect to their effects upon worldwide high-frequency circuits. The apparent influence of the reported polarity and intensity of the spots is noted. Photographs of typical disturbance-producing spots are illustrated. The solar east-west movements are related to vertical incidence F-layer recordings at Washington, D. C., and transatlantic signal recordings at Riverhead, L. I., N. Y. The applications of these data for short-period radio transmission condition forecasts are described.

INTRODUCTION

THE paper describes means which have been employed to improve the accuracy of short-period radio weather forecasting. A 6-inch refracting telescope mounted on the roof of the RCA Communications, Inc., building at 66 Broad Street, New York City, is used for daily solar observations. Observed sunspots are projected upon a map of the sun, properly oriented to account for relative inclinations of the axes of the sun and the earth. The sunspots are registered on the map in linear co-ordinates about the sun's axis of rotation. Meridians on the map are located at 13.3-degree intervals, with respect to the east limb of the sun. This procedure accounts for one day's travel of a spot, when the rotational period is 27 days.

The map, in addition, carries a critical zone, which is herein defined and localized. The relationship of spots with respect to the critical zone, and their effects upon radio circuits is studied, both by traffic reports, and signal recordings. The effects of the polarity and intensity of spots is indicated. Spots are functions of the solar mechanism, which produces that complex ionospheric phenomenon, which is called "Radio Weather" in this paper.

RADIO WEATHER DEFINED

The term "Radio Weather", as used herein, relates to the propagation variables as observed on a world-wide network of circuits operated

---

* Decimal Classification: R113.4 × R800(523.74).

† Presented at the joint meeting of the Institute of Radio Engineers and the International Scientific Radio Union in Washington, D. C. on May 3, 1948.

‡ Research Department, RCA Laboratories Division, New York, N. Y.

Engineering Department, RCA Communications, Inc., New York, N.Y.

by RCA Communications, Inc. The ionosphere is the normal highway for this world-wide, high-frequency network. It responds to daily changes in the solar surface, and particularly to the sunspots which cross that surface. The effects on the earth are measurable by ionosphere stations in many parts of the world. The effects are also measurable in terms of signal variability on globe encircling circuits. This variability is radio weather. It may be normal, subnormal, or supernormal in its effects upon circuit performance.

Since the ionosphere is the communications highway, a convenient medium-latitude ionosphere measuring station, such as Washington, D. C., may be used as a radio weather pilot station. This has been accomplished by plotting hourly values of F-layer critical frequencies against the sun's 27-day rotational period. A running average of the 27-day rotational period, or one solar circumference, provides a smoothed picture of an average day, when processed hour by hour. Against this smoothed average are plotted the observed day-by-day, hourly measurements. Hours below the average are indicated by solid fills, above the average by a solid line. The result is a radio weather map for successive rotations of the sun, as shown in Figure 1.

Pertinent Characteristics of Sunspots

Sunspots have been observed for more than 150 years, and volumes have been written about them. It is not the purpose of this paper to invade the field of the astronomer, but rather to cooperate with him in the application of solar knowledge. There will be described here only such observed characteristics of sunspots as concern radio circuit performance, and short-period—2 days at most—radio weather forecasting.

(a) *Effect of size, shape and locations of spots.*

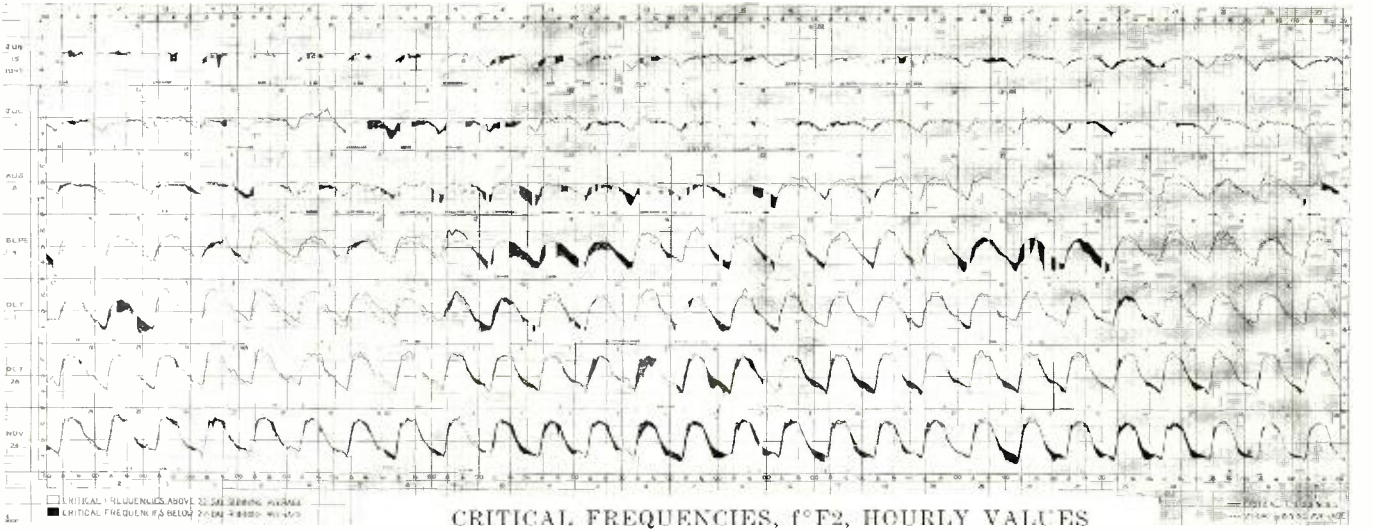
Mere size of a spot is a meaningless criterion. A huge spot, consisting mostly of penumbra, may be a complete dud in its effect upon radio circuits. A small spot, mostly umbra, can be quite deadly. The larger the umbra the more potent its effects upon radio circuits. Rapidly changing spots, indicative of activity, are the most lethal in their effects upon radio circuits.

Their relative positions are important. They affect circuits to a major extent only when about 2-days travel east of the central meridian, in what will presently be described as "the critical zone".

(b) *Sunspot polarity effects.*

Mount Wilson regularly reports the polarity of umbras as red

A—Characteristic Not Measurable Because of Blanketing by Sporatic or Abnormal E.
B—Characteristic Not Measurable Because of Loss of Trace Due to Absorption.
C—Characteristic Not Measurable Because of Loss of Trace Due to Failure of Equipment.
D—Critical Frequency Higher Than Upper Frequency-Limit of Recorder.
E—Critical Frequency Less Than Lower Frequency-Limit of Recorder.
F—Spread-Echoes Present.
G—f°F2 Equal to or Less than f°F1.
H—Stratification Observed Within the Region.
J—Ordinary-Wave Critical Frequency Deducted from Measured Extraordinary-Wave Critical Frequency
K—Ionosphere Storm in Progress.
L—Critical Frequency, muf, or *muf* Factor for F1 Layer Omitted.
N—Unable Make Logical Interpretation.
R—F2 Incoherent Near Critical Frequency.
S—Interference Obliterated Screen—No Record.



CRITICAL FREQUENCIES, f°F2, HOURLY VALUES
COMPARED TO 27-DAY RUNNING AVERAGE
AND THE SOLAR CYCLE
From CRPL data, Washington, D. C.
75°W MEAN TIME

Fig. 1—Washington, D.C. as a "Radio Weather" pilot station — chart for seven solar rotations. (June 15 to December 20, 1947)

(positive) and violet (negative), as well as the relative intensities of red and violet. Polarities have been found to have a significant effect upon circuits in the northern and southern hemispheres. No conclusive figures can be presented at the present time. It may be stated, however, that reds have a preponderant effect in the northern hemisphere, violet in the southern hemisphere. Reds depress frequencies in the northern hemisphere, raise them in the southern, and vice versa for violet. This phenomenon may be due to the earth's polar attractions for solar radiations of opposite polarity, causing recombination and lowering of ionization near the poles for one polarity, and intensification of ionization for the other.

The study is complicated by the fact that many world-wide circuits have reflection points in both hemispheres. The effect is proportional to the reported intensities of red or violet, or the preponderance thereof. Here is a fertile field for the solar and ionosphere physicist.

Solar Critical Zone

Intensive observations of the locus of circuit disturbance producing spots, has narrowed to a semi-circle of about 26 degrees in radius, centered at the sun's optical center, on the eastern hemisphere, and terminating at the axis of rotation. The spots have maximum effect when they cross either the circle, or the central meridian. While within the circle, they cause erratic conditions. On North Atlantic circuits reds tend to lower frequencies. However, since both red and violet accompany most spot groups, there is a mixed effect when both polarities are within the circle.

The critical zone determination is an outgrowth of observations* started on a home-made 6-inch reflecting telescope. Since that time, a standard telescope installation has been used on a recognized procedural basis. This work includes the preparation of daily maps, and the issuance of daily radio weather forecasts.

Further consideration and data with regard to the critical zone will be found in the section—Delineation of the Critical Zone.

Typical Sunspot Photographs

Through the courtesy of the U. S. Naval Observatory, several photographs are presented of the sun taken at, or near, the times of severe disturbances on the North Atlantic circuits of RCA Communications, Inc. Linear co-ordinates about the sun's axis of rotation have been drawn on the photographs, and the critical zone is indicated. The

* Conducted by J. H. Nelson.

relative locations of sunspots, with respect to the critical zone, and to the axis of rotation, may, therefore, be observed. Relative intensities and spot polarities, are also indicated.

The date, and the hour, at which the photograph was taken is indicated for each case. Radio circuit conditions, prevailing at, or near, the time of the solar pictures, as reported by the Central Radio Office of RCA Communications, Inc. at Broad Street, New York City, or by the Riverhead, L. I., Receiving Center, are correlated with each picture. An attempt is also made to show the effect on the circuits of sunspot polarity.

A solar photograph taken at 1:14 p.m. on September 25, 1947, is shown in Figure 2 (page 235). At this particular time the northern trailing spot (red) was about 5 hours east of the central meridian. A southern spot, also red, was about 16 hours east of the central meridian.

The Riverhead log for September 25th reads:

"***Midnight — 8 a.m.***

Conditions from Europe poor to unreadable until 5 a.m. when 13 and 14 megacycles started to come up. Europeans generally unheard until 12:30 a.m., and then night frequencies that did come up, good only for aural reception at slow speeds. By 6 a.m. day frequencies were about normal. West Coast also very bad, while South America poor most of the night.

8 a.m. — 4 p.m.

Conditions from Europe poor all watch, very low intensities and moderate noise levels. The high frequencies poor and not usable. Early transition to 10 megacycles from Europe at 1:05 p.m. South America fair most of watch except for weak period from 9:00 to 9:40 a.m.; then fair. West Coast and Tangier fair to good.

4 p.m. — Midnight

Europeans poor until 7 p.m., when they improved to fair. Others good."

This report indicates an improvement in circuit conditions at 7 p.m., which is about the time that the northern (red) spot passed the central meridian. The storm ended at 5 a.m. on September 26th, which corresponds closely to the time when both spots are west of the meridian. (See Figure 10 (page 240) for recorded disturbed day conditions on September 24, 1947.)

Figure 3 (page 235) is a photograph showing spot locations at 12:17 p.m. on October 2, 1947. At the time the photograph was taken a trailer spot (violet) of one group was just crossing the central merid-

ian, and the leader (red) of the next easterly group was about to cross the circle. This is a spot combination that normally causes generally erratic conditions.

The Riverhead log for October 2, 1947, reads as follows:

"***Midnight — 8 a.m.***

Europeans fair until 4 a.m., then weak on all frequencies until 5 a.m., fairly good conditions balance of watch. Others good.

8 a.m. — 4 p.m.

Conditions fair to good from 8 a.m. to 1 p.m., when fair with early transition to night frequencies. Other directions good.

4 p.m. — Midnight

Conditions fair from east until around 7:30 p.m. when mostly poor. Flutter fading with conditions erratic throughout the watch. Transitions were early, and lower frequencies never did recover normal strength. Tangier poor in last hour from 12 to 7 megacycles. South good. West Coast fair."

This is an example of generally erratic conditions while a red is in the critical zone. This is a continuing storm, which did not end until the leader (red) of the following group crossed the meridian about midnight of October 3rd.

Figure 4 (page 235) is a photograph of sunspot locations taken at 10:28 a.m., February 3, 1948. At the time the leader (red) was within the circle, and the follower (violet) about 10 hours east of the circle.

Extracts from the Broad Street Traffic Chief's log are quoted as follows:

"***Midnight — 8 a.m.***

Poor and subnormal until daylight, all Europe and South America.

8 a.m. — 4 p.m.

North Europe subnormal in morning, improved in afternoon. Berne multiplex poor 1 p.m.-2 p.m. due to weak signals. Stockholm unreadable until 2:30 p.m.

4 p.m. — Midnight

Generally normal except London multiplex poor 7 p.m. to 10 p.m. Signals weakened to South America 9:30 p.m."

This is another example of erratic conditions on Northern Circuits with red in the critical zone. South America went weak at about the time the follower (violet) crossed the circle.

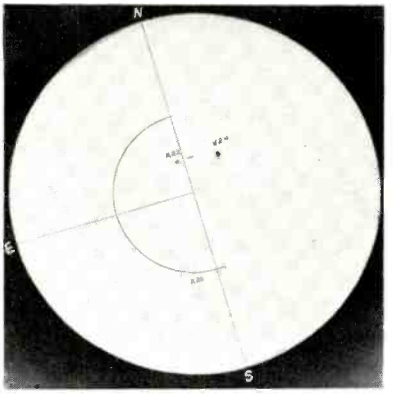


Fig. 2—Sept. 25, 1947; 1:14 p.m.

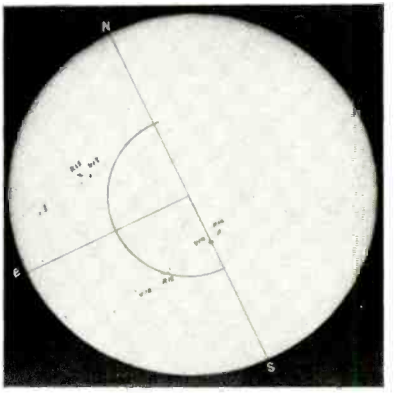


Fig. 3—Oct. 2, 1947; 12:17 p.m.

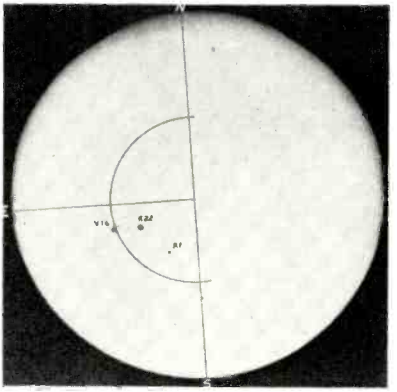


Fig. 4—Feb. 3, 1948; 10:28 a.m.

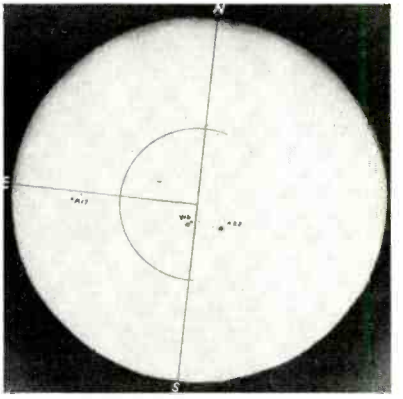


Fig. 5—Feb. 5, 1948; 11:31 a.m.

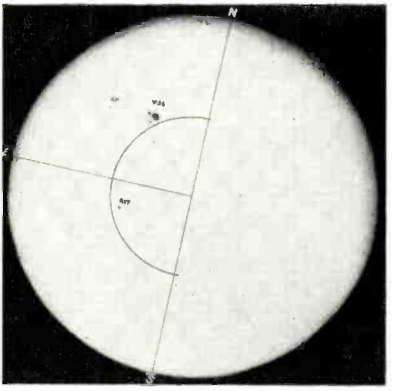


Fig. 6—Mar. 1, 1948; 11:07 a.m.

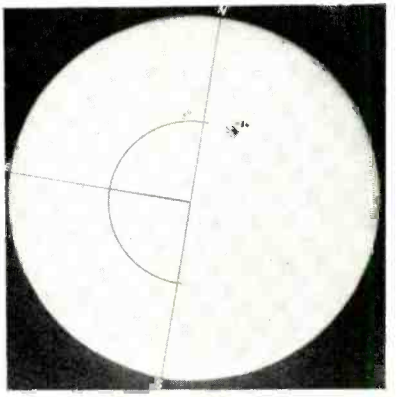


Fig. 7—Mar. 3, 1948; 10:53 a.m.

U. S. Naval Observatory photographs of the sun.

Figure 5 is a photograph of sunspot locations at 11:31 a.m. on February 5, 1948. At this time the trailer spot (violet) is about 3 hours east of the central meridian.

Quotations from the Broad Street Traffic Chief's log for this date read as follows:

"***Midnight — 8 a.m.***

Generally normal except day frequencies made late contacts.

8 a.m. — 4 p.m.

North Europe normal in morning, erratic in afternoon. London multiplex poor 11 a.m. to 2 p.m., and normal thereafter. Erratic to Caracas and Bogotá, unable read his normal high frequencies due high noise level. Rome multiplex good today except 1:30 p.m. to 3:30 p.m.

4 p.m. — Midnight

Conditions very good, or excellent to all points."

The violet spot may have caused the increased absorption between 11 a.m. and 3 p.m. Conditions became normal at about the transit time of the central meridian by the violet spot.

Figure 6 is a photograph of spot locations at 11:07 a.m. on March 1, 1948. At this time a heavy violet umbra is about 9 hours east of the critical zone circle, while a red umbra had crossed about 11 a.m.

Extracts from the Broad Street Traffic Chief's log read as follows:

"***8 a.m. — 4 p.m.***

Europeans and Central American circuits erratic beginning about 11 a.m. Bad dip on Europeans at 3 p.m. South America not affected.

4 p.m. — Midnight

North Europe poor. Central Europe generally poor. South Europe and South America erratic."

The large violet spot came in after sundown. Its ionizing effects were not felt in full until the following day, when South American circuits were reported poor.

This spot group is of particular interest since there had been forebodings of serious communication disturbances in the press.

Figure 7 is a photograph of the sun taken at 10:53 a.m. on March 3, 1948. Photographic conditions were poor; but the picture is of interest since the exposure was made near the time at which the March 1st storm had subsided.

The ending of the storm may be described in the words of the Traffic Chief's log at Broad Street, for March 3, 1948, as follows:

"***Midnight — 8 a.m.***

Poor to Oslo (North Europe) until 4:30 a.m. At 4:30 a.m. Oslo started short waves. Tangier (South Europe) unreadable on short waves until 4 a.m., and unable read shorts from us until 5:30 a.m. London improved to fair about 12:30 a.m.

8 a.m. — 4 p.m.

North Europe fair in general, with Oslo reading short waves throughout.

4 p.m. — Midnight

All areas normal."

The records indicate that Oslo came up at 4:30 a.m., and Tangier at 5:30 a.m. The red spot crossed the central meridian between 4 and 5 a.m.

Probability of a Solar Ionosphere

The existence of a critical zone on the sun poses an interesting speculation with regard to the existence of a solar ionosphere, or the equivalent thereof. This might be the corona or the prominences overlying sunspots.

If the 26-degree semi-circle delineates a region from which radiations affecting the earth's ionosphere are confined, these radiations penetrate the solar ionosphere only when within about 7 minutes of arc of the line connecting sun and earth. Radiations of lower angle do not reach the earth, or are returned to the sun. The criticalness of the solar radiation angle is indicated by the fact that, considered from the east limb of the sun, this angle represents only 16 minutes of arc.

If subsequent observations confirm that spots west of the axis of rotation do not affect the earth's ionosphere, such a result may be due to the centrifugal movement of the gases forming the prominence, or the solar corona. A displacement of only about 7 minutes of arc would render such radiations inoperative.

North Atlantic Signal Recordings

Recordings of London and Tangier signals simultaneously at Riverhead, N. Y., provide further means of correlating solar, ionosphere and signal conditions.

The recording dates selected represent normal, above normal, and subnormal radio weather conditions, referred to the radio weather map

of Figure 1. A normal day, on this map, is one in which daily measured critical frequencies are close to the 27-day moving average. Supernormal days will then be above the moving average, and subnormal days (shaded) will register below moving average.

In addition to the radio weather condition on the days selected, there is shown the distribution of light and darkness over the circuit paths, and the positions of sunspots within the critical zone, with their polarities and relative intensities (from Mount Wilson reports). The

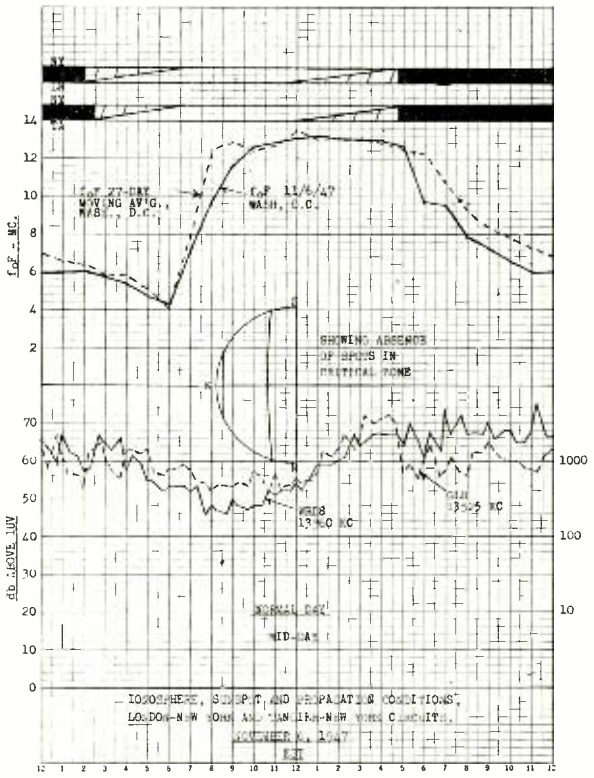


Fig. 8—North Atlantic signal recordings — normal day conditions.

time is Eastern Standard, and the mid-point of light distribution over the circuits will indicate noon at mid-circuit.

Normal day conditions in November are shown in Figure 8. The date of recording is November 6, 1947. The circuits compared are GLH, London, 13,525 kilocycles, and WRD8, Tangier, 13,360 kilocycles. There were no spots in the sun's critical zone on November 6, 1947, and the f_0F recordings at Washington, D. C., compared with the 27-day running average, as shown.

Lowest signal intensities occurred on this normal day at about mid-day at the mid-point of the path, and highest at midnight, with Tangier signals of the order of 2,500 microvolts, and London, 1,000 microvolts. Both signals were usable for 24-hour operation, a normal condition for these frequencies.

Supernormal conditions are illustrated by recordings made on October 6, 1947, plotted in Figure 9. It will be noted that on this date there is a preponderance of violet spots in the critical zone, at

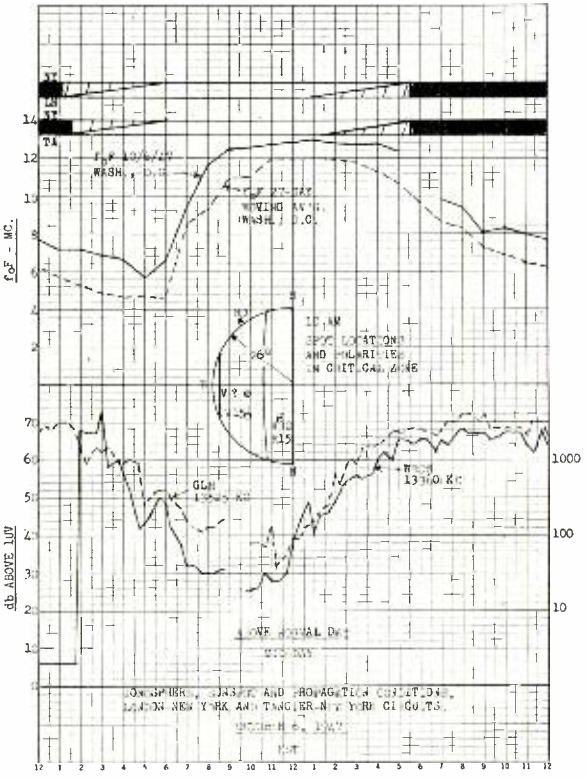


Fig. 9—North Atlantic signal recordings — supernormal conditions.

positions and relative intensities as indicated in the center diagram. The f_0F recordings for the day at Washington, D. C., are seen to be above normal for the entire 24 hours.

The absorption dip again occurs at noon at the mid-point of the circuits, but is deeper for the low-latitude circuit. The high-latitude circuit shows higher night intensities. This is a commonly observed characterstic for supernormal days on North Atlantic circuits.

Subnormal, disturbed day, conditions are illustrated by recordings made on September 24, 1947, Figure 10. On this day f_oF critical frequencies at Washington, D. C., were far below average throughout the 24 hours. As shown in the critical zone diagram of Figure 10, there were two predominant red spots in the zone. These normally cause a decrease in ionization, erratic conditions, depressed frequencies, and early transitions on European circuits. Violet, on the other hand, tends to increase ionization.

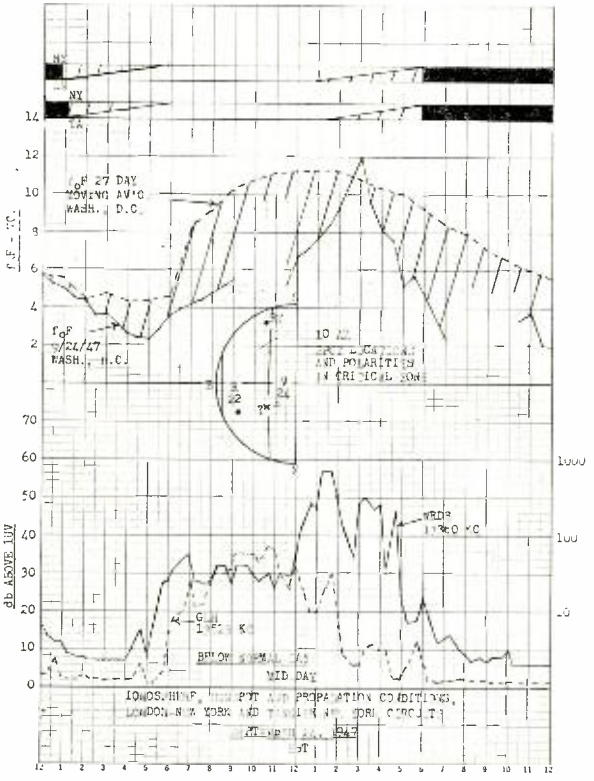


Fig. 10—North Atlantic signal recordings — subnormal, disturbed day, conditions.

A reversal of the normal day trend is indicated with highest signal levels during daylight hours. A momentary rise in critical frequencies and signal levels occurs about 3 p.m., possibly due to the violet. The low-latitude circuit, Tangier, is consistently better than the London circuit during this severely disturbed day. This recording confirms the conclusion that red spots decrease ionization, whereas violet spots tend to increase it, on North Atlantic circuits.

Delineation of the Critical Zone

A series of spot position checks, at disturbed hours, over a considerable period of time, gave strong evidence of falling upon a circle, about 26 degrees in radius from the optical center of the sun, on its eastern hemisphere. Linear co-ordinates about the sun's axis of rotation are the basis of reference. These co-ordinates are shown on the solar photographs.

A more complete delineation of the critical zone was obtained by noting simultaneously the positions of individual spots, at times when the London circuit rating, during the period July 17 to December 15, 1947, dropped to 5, or below. The London multiplex is erratic at a rating of 5. When multiple spots appeared near the zone, those of maximum intensity and size were given first consideration. If one of these, of maximum intensity, was on the circle, it only was considered. If spots of nearly equal intensities, and questionable identity were in the zone, they were interconnected by lines. Positive spots (red) are indicated by a solid circle, negative spots (violet) by a plain circle, and unknown polarities by crosses. The results of spots so identified are shown in the map, Figure 11 (page 242).

The polarity relationship of spots, causing disturbances of rating 5 or below on the London-New York circuit, for the 6-months period, July 17 to December 15, 1947, were as follows:

Red (Positive)	42	(Per cent of Total = 48)
Violet (Negative)	31	(Per cent of Total = 35)
Unknown polarity	15	(Per cent of Total = 17)
Total in Zone	88	

The preponderance of red in this tabulation is not conclusive due to the 17 per cent of spots of unknown polarities. The relative intensity of radiation at the respective polarities is also a factor. Data available have been incomplete. It is hoped that more conclusive evidence will be obtained in subsequent polarity data from Mount Wilson, and by continued observation.

A summary of lost traffic hours and spot positions on the London-New York circuit, during the same 6-months period provides more conclusive evidence as to localization of the critical zone, as follows:

Total lost hours	405	(Per cent of Total = 100)
Total lost hours with spots in critical zone	359	(Per cent of Total = 89)
Total lost hours with no spots in critical zone	46	(Per cent of Total = 11)

This is preponderant evidence of localization of solar disturbance to the boundaries herein described.

Microwave Solar Noise and the Critical Zone

An article[1] by A. E. Covington of the National Research Council, Ottawa, Canada, provides an additional critical zone correlation, based upon microwave solar noise measurements, during the eclipse of November 23, 1946.

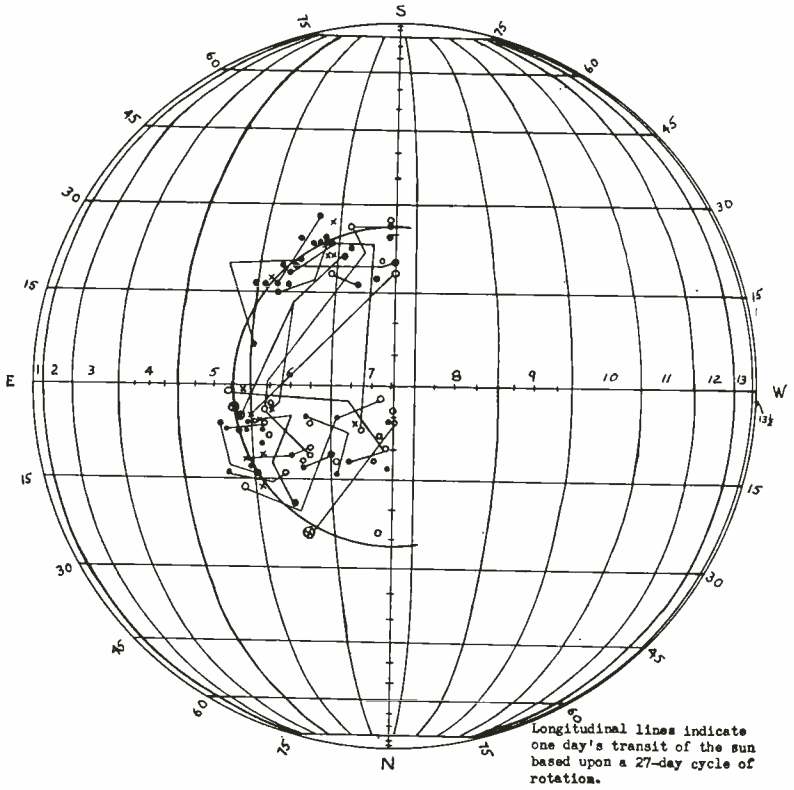


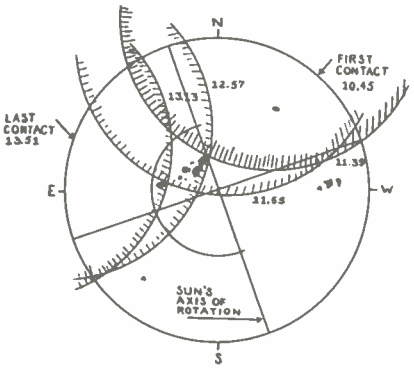
Fig. 11—Solar critical zone (locus of disturbing spots — July 17 to December 15, 1947.)

Figures 1 and 2 of this article are reproduced as Figure 12, the only modification being that the critical zone of the sun, as herein described, has been superimposed. It will be observed that there are spot groups in the north-west, south-west and south-east quadrants of the sun, as well as the prominent group within the critical zone.

[1] A. E. Covington, "Microwave Solar Noise Observations During the Partial Eclipse of November 23, 1946," *Nature*, Vol. 159, No. 4038, pp. 405-406, March 22, 1947.

The north-west spot was eclipsed at approximately 11:15 a.m., but there is no indication of a change in noise level due to this occlusion. On the other hand, spot groups in the south-west and south-east zones were not eclipsed. Their contributions to the noise level are doubtful.

The group of spots in the critical zone was obscured at 11:39 a.m. and uncovered at 12:57 a.m., which correlates quite accurately with the recorded bay in the noise curve.

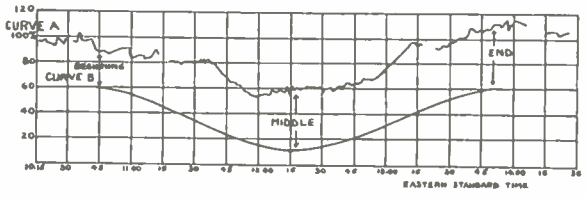


"Fig. 1. LUNAR SHADOW POSITIONS OUTLINING ACTIVE REGION. DATA FROM DOMINION OBSERVATORY PHOTOGRAPHS"

Fig. 12—Figures 1 and 2 of Reference 1 with solar critical zone superimposed.

"Fig. 2. CURVE A, VARIATION OF SOLAR NOISE DURING ECLIPSE, NOVEMBER 23, 1946.

CURVE B, ECLIPSED AREA OF SUN."



Propagation Significance of the Critical Zone

It has been generally observed that very low correlation factors, or the entire absence thereof, have been obtained by comparisons of daily sunspot numbers, and daily circuit performance.

The correlations have always improved when smoothed averages, over long periods have been used. Such a result would logically follow, since the majority of all spots, at certain times, pass through the critical zone.

The further localization, and confirmation of the critical zone, and

the confinement of sunspot counts to inhabitants of this zone, will not only increase the day-to-day correlations of ionosphere and circuit performance, but also provide a more realistic basis for frequency predictions over the sunspot cycle.

Practical application of the critical zone, as here indicated, has produced a dependable forecasting service, with many hits registered within 15 minutes. Its use is now an accepted routine service, and an unquestionable aid to traffic handling.

ACKNOWLEDGMENT

In the course of this work, the following organizations lent active cooperation: the McMath-Hulbert Observatory at Pontiac, Michigan; the Harvard Astronomical Observatory at Climax, Colorado; the Mount Wilson Observatory at Pasadena, California; the U. S. Naval Observatory at Washington, D. C.; and the Central Radio Propagation Laboratory of the National Bureau of Standards.

The cooperation of Dr. U. S. Lyons and Mrs. L. T. Day of the U. S. Naval Observatory in furnishing the photographs accompanying this paper, and for daily reports of sunspot areas and positions is deeply appreciated.

The work is a further step in co-ordinating the efforts of the astronomer, the ionosphere physicist, and the radio communications engineer, to the end that all peoples may have better world-wide communications.

ELECTRO-OPTICAL CHARACTERISTICS OF TELEVISION SYSTEMS*

By

Otto H. Schade

Tube Department, RCA Victor Division,
Harrison, N. J.

NOTE: This paper consists of an Introduction and four parts: Part I — Characteristics of Vision and Visual Systems; Part II — Electro-Optical Specifications for Television Systems; Part III — Electro-Optical Characteristics of Camera Systems; Part IV — Correlation and Evaluation of Electro-Optical Characteristics of Imaging Systems. The Introduction and Part I appeared in the March 1948 issue of *RCA REVIEW;* a summary is reprinted herewith for reference purposes. (A limited number of copies of the March 1948 issue is still available for those who desire a complete file on this paper.) Part II is included in this issue. The remaining parts are scheduled for publication in the September and December issues of Volume IX of *RCA REVIEW* during 1948.

INTRODUCTION; PART I — CHARACTERISTICS OF VISION AND VISUAL SYSTEMS

(Reprinted from RCA REVIEW, March 1948)

Summary—*The optical and electro-optical conversion processes in television systems are examined as intermediate stages of a multi-stage process by which optical information at the real object is "transduced" into sensory "response" at the brain. The characteristics of the human eye and vision in the final stage of the process determine the requirements and standards for preceding stages. When expressed on a unified basis by "transfer" and "aperture response" characteristics, the properties of the process of vision can be correlated with those of external imaging and transducing processes. It is shown that image definition, or the corresponding information from optical or electrical image-transducing stages, can be specified by the characteristics of an equivalent "resolving aperture." These characteristics may be computed and measured for all components of the system.*

Quantitative data from measurements permit definite quality ratings of optical and electrical components with respect to theoretical values. A subjective rating of the resolution in an imaging process external to the eye such as a television system is derived by establishing a characteristic curve for the relative "sharpness" of vision as affected by the "aperture response" of the external imaging process.

A general review of the material and the broad methods of analysis employed are given in the Introduction. Following this, Part I treats characteristics of vision and visual systems. In this part, viewing angle, sensation characteristics, color response, persistence of vision, flicker, resolving power, response characteristics, and steady and fluctuating bright-

* Decimal Classification: R138.3 × R583.11.

ness distortions are discussed and related to the characteristics of external imaging systems and the television process.

* * * * * * * * * * * * *

There follows the second paper in this series: Part II—Electro-Optical Specifications for Television Systems.

The Manager, RCA REVIEW

PART II — ELECTRO-OPTICAL SPECIFICATIONS FOR TELEVISION SYSTEMS

Summary—*The ability of an image-forming device to reproduce fine detail can be specified basically by the size and flux distribution of the small light spot formed as the image of a point source of light. It is shown that the defining ability of practical image transducers is specified more accurately by response characteristics obtained by scanning a test object with the elemental point image which represents the "resolving aperture" of the imaging device. Methods of computing and measuring the "aperture flux response" of practical image transducers are developed for correlation of optical and electrical system components.*

The television raster is treated as a sampling process and its effect on the system resolution is evaluated as an aperture process. Brightness, repetition rate, and flicker in television images are treated in relation to the screen materials and performance of practical kinescopes.

A. The Analysis of Image Definition and Structure and Their Specification by the Flux Response Characteristics of Equivalent Resolving Apertures

The ability of an imaging device or process to form a clear image of fine detail is specified basically by the size of the small light spot formed by the device when imaging a point source of light. The elemental point image (known as a "figure of confusion", a "diffraction disc", a "picture element", or, less specifically, as an "elemental area") depends in size on certain characteristics of the particular transducing process. Optical images contain a very large number of elemental areas which can be formed simultaneously or in rapid sequence by one or many beams of light or electrons. The element controlling the cross section of these beams is usually a physical aperture. The specific action of a transducer in defining or changing the size of an elemental area (even without a physical aperture) has, therefore, been termed an "*aperture process*". The action of defining the elemental area in any transducing process can be assigned, in a transferring sense, to an equivalent resolving aperture. This equivalent "aperture" is equal in size, shape, and flux distribution (or transmission) to the elemental area itself when a "point" signal is transduced.

The effect of the "aperture" on the definition of images can be evaluated by the process of scanning in the form of a "frequency response" characteristic or "aperture response" characteristic because the scanning motion identifies spatial variations in image flux as different frequencies according to the detail size in the image. It will be confirmed by optical experiments that the "aperture response" of simultaneous figures of confusion in optical images is identical with the aperture response obtained by scanning the image with one aperture equal to one figure of confusion.

It is quite adequate to specify the definition in an image by the size and flux distribution of its resolving aperture rather than by the aperture effect. The accuracy in measuring small "spot" sizes and their flux distribution is limited, however, in practical cases by fluctuation phenomena (noise) which usually do not permit intensity measurements to be made with a precision better than one to two per cent. It will be shown in the section on complex apertures that errors of this order can cause a 50 per cent change in the response factor of practical apertures.

Aperture response characteristics computed by scanning "square-wave" line patterns are, therefore, especially suitable for rating the performance of practical imaging devices because they permit a direct comparison of theoretical and measured data. The combined aperture response of several aperture processes in cascade can be established from individual aperture characteristics by repeating the scanning process.

1. Aperture Response Characteristics

The aperture effect on definition is determined by the simple geometric process of moving the aperture over a "flux pattern" and plotting the integrated flux passing through the aperture as a function of displacement. A unit-function flux change which is optically a sharp boundary between two brightness levels B_1 and B_2, a single rectangular pulse or a series of pulses forming a square flux wave are useful test patterns.

To pass over a unit-function flux change, the aperture must progress one aperture diameter, δ. (Figure 14). The form of the "transition curve" depends, obviously, on the shape and flux distribution within the aperture. A square aperture with uniform density gives a straight-line transition; a round aperture with uniform or cosine-square density has an S-shaped transition. The transition curve has symmetric halves and is given for several apertures in Table I of the Appendix.

The most useful pattern (for obtaining a standard detail test signal) is, for many reasons, the square-wave flux pattern, which, optically, is the standard line or bar test pattern. Groups of lines or a tapered

wedge are used to vary line width or line number with respect to the aperture diameter δ. (See Introduction, and Part I).

For computation of aperture response curves, the aperture may be considered as a three-dimensional body such as a cube, cylinder, cone, etc. The base of this body represents the aperture size and shape; the height represents the flux density; and a vertical central cross section represents the density distribution. Finding a point on the response curve or flux wave generated by passing the aperture over a flux boundary or a number of boundaries is, hence, the problem of determining the total volume within the vertical sections of the aperture body located between boundaries containing signal flux. The aperture flux

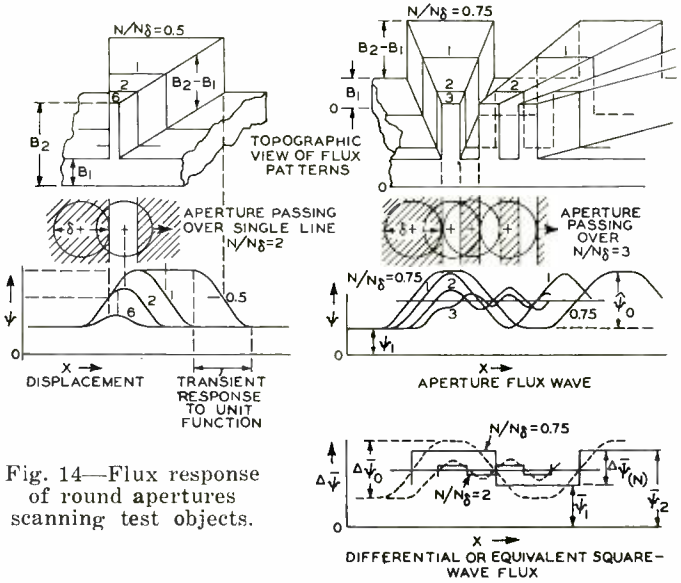


Fig. 14—Flux response of round apertures scanning test objects.

waves are the plot of the active flux volume as a function of aperture displacement. The flux wave generated by passage of the aperture over square-wave flux patterns is characterized by its amplitude and its wave form. (See Figure 14).

The amplitude response factor $r\Delta\hat{\psi}$* is defined as the ratio of the flux amplitude $\hat{\psi}_N$ at a line number N to the amplitude $\hat{\psi}_o$ at $N = 0$.

$$r\Delta\hat{\psi} = \hat{\psi}_N/\hat{\psi}_o \tag{23}$$

The response factor $r\Delta\hat{\psi}$ as defined by Equation (23) is single valued

* It should be pointed out that the response factors derived in this paper differ by definition from those obtained by a two dimensional Fourier analysis; see References 1, 2 and 3 of Part I.

and independent of the contrast of the test object. It is unity in the range from $N = 0$ to the line number $N\delta$ at which the aperture diameter (δ) is equal to one line width. (The line number N in this paper is expressed in television lines; both dark and light lines are counted).

The waveform of the flux wave changes in this range from a square at $N = 0$ to a trapesoid for $N > 0$, the top of the trapezoid shrinking to a point at $N = N\delta$. At $N = N\delta$, the wave becomes a triangle (square aperture) or a sine wave (round aperture). Beyond $N\delta$ it decreases in amplitude towards zero which occurs at a particular value $N = N_c$. (The wave changes again towards a square at $N = N_c$ for the square aperture).

The computed response characteristics, Figure 15, show that sharply

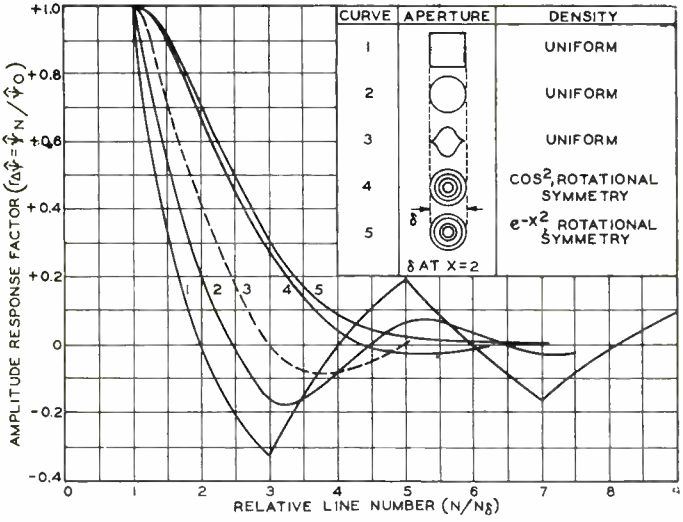


Fig. 15—Amplitude response of various apertures to square-wave flux pattern.

defined apertures have multiple zero points. These curves are analogous to the transient response of electrical networks with sharp cutoff. The response beyond the first zero is a spurious signal having one line in the pattern missing after each zero. The negative sign refers to the polarity of the flux with respect to its a-c axis and indicates a sudden phase reversal; black lines continue as white lines and vice versa. The "overshoot" is a maximum for a square aperture with uniform flux density and zero for a round aperture with exponential flux distribution.

The Differential Flux Response Factor (Equivalent Square Wave)

The detail contrast in the optical signal from the aperture is a maximum when $r\Delta\hat{\psi} = 1$ and when the flux wave is rectangular. The

average flux change obtained by intergration of half waves specifies the degree of equivalent contrast. The ratio of $\Delta\bar{\psi}_N$ at the line number N to the flux change $\Delta\bar{\psi}_o$ of the perfect square wave at $N = 0$ is the differential flux response factor

$$r\Delta\bar{\psi} = \Delta\bar{\psi}_N/\Delta\bar{\psi}_o \tag{24}$$

This response factor is also single valued and independent of the test object contrast. The detail contrast itself is easily derived from $r\Delta\bar{\psi}$ but depends on absolute values (See below). The differential flux response is by definition the *equivalent square-wave amplitude response*. It is unity only at $N = 0$, decreasing linearly past the 70-per cent value

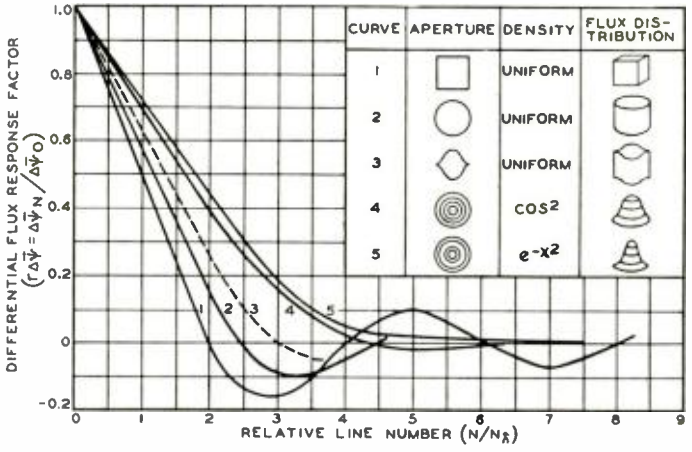


Fig. 16—Equivalent square-wave response of various apertures to square-wave flux pattern.

at $N \simeq N\delta$* and reaches its first zero at $N = N_c$. Round apertures with a cosine-square or exponential flux distribution with rotational symmetry (lenses and electron beams) generate flux waves substantially sinusoidal in shape in the range from $N\delta$ *to* N_c permitting the useful approximation:

$$\text{for } N > N\delta; \qquad r\Delta\bar{\psi} \simeq 2r\Delta\hat{\psi}/\pi$$

The flux response factors $r\Delta\bar{\psi}$ computed for various aperture types are shown in Figure 16. (See Table II in Appendix).

In certain cases, the aperture characteristic may be described by the *energy response factor* which is the response to absolute signal

* The square aperture is an exception.

values. For optical objects the energy response is expressed by the contrast ratio

$$C = \bar{\psi}\ (\text{white})/\bar{\psi}\ (\text{black}) = \bar{B}_2/\bar{B}_1 = (\bar{\psi}_1 + \Delta\bar{\psi})/\bar{\psi}_1 \qquad (25)$$

or the "per cent contrast" $\%C = 100\ (1 - 1/C)$ (26)

The energy response factor is expressed by

$$rC\delta = \%C_N/\%C_0. \qquad (27)$$

A test object with substantially 100 per cent contrast furnishes the value:

$$rC\delta = \%C_N = [\,(1 - (1 - r\Delta\bar{\psi}))/(1 + r\Delta\bar{\psi})\,]\,100. \qquad (28)$$

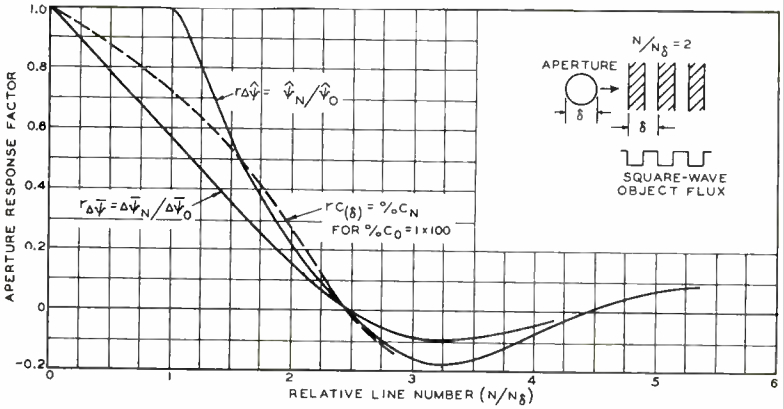


Fig. 17—Flux response of round aperture with uniform flux density or transmission.

The three response curves for a round aperture are shown in Figure 17.

2. Complex Apertures

Practical apertures have, in many cases, a flux distribution following a complex law. The complex aperture is considered as a superposition of a number of simple coaxial apertures and the complex aperture response is obtained as the arithmetic sum of the response curves of the component apertures each having a certain flux (ψ_n), density distribution, and diameter δ_n. In round apertures, the flux ψ_n in component apertures of equal shape is proportional to the product of their density (height) and base area (δ_n^2). Figure 18 illustrates the graphic construction of the response characteristic for a round aperture with cosine-square flux cross section as the sum of 5 coaxial round apertures

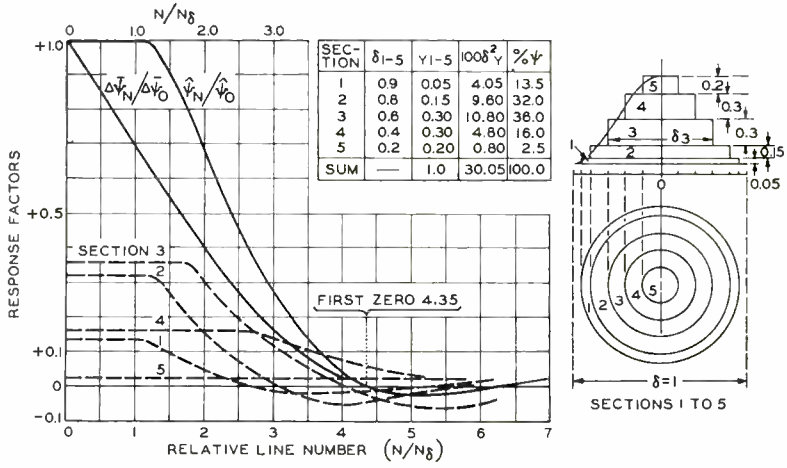


Fig. 18—Synthesis of flux response of a round aperture with cosine-squared flux density or transmission.

with different diameters and uniform flux density. It should be noted that the response curves of the larger sections are continued to the value N_c of the smallest section because their "spurious" response values beyond the first zero must be included when adding up individual response curves.

The aperture response of a fine light spot surrounded by a relatively low-intensity diffusion disc is illustrated by the constructions in Figure 19. The insert drawing shows the cross-sectional flux distribution resulting from two coaxial cosine-square apertures with the diameter ratio $\delta_2/\delta_1 = 5$ and the flux amplitude ratio $\hat{\psi}_2/\hat{\psi}_1 = 0.05$. The corresponding aperture response $r\Delta\bar{\psi}$ is that of Curve 1. Because of the large diameter (δ_2) the flux ψ_2 in the "diffusion" disc is 57 per cent of the total flux and causes a pronounced "kink" in the response curve, which

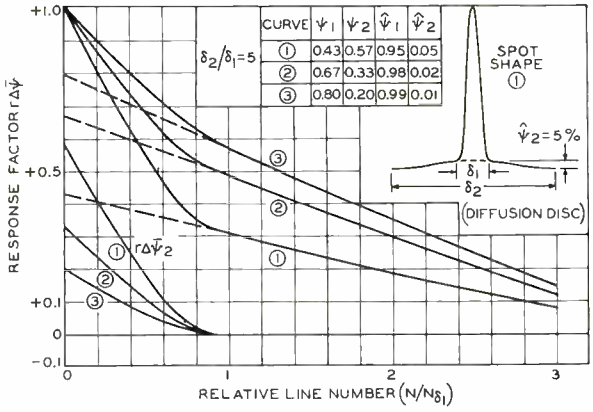


Fig. 19—Effect of aberrations (diffusion disc) on the aperture response.

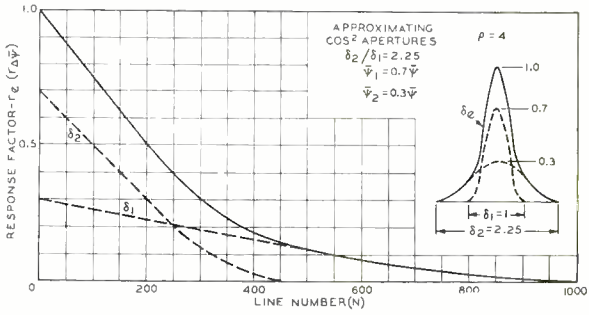


Fig. 20—Formalized aperture response characteristic of the eye for $\rho = 4$.

is still noticeable for a 1-per cent diffusion intensity (Curve 3). The similarity of the eye and kinescope response characteristics as shown in Figures 20 and 13*. These and other measured characteristics may thus be analyzed for aberrations by the reverse graphic procedure of subtracting the normal extrapolated fine-aperture response curve from the total curve. Table III in the Appendix is useful for curve extrapolation.

3. *Simultaneous Imaging by Multiple Apertures*

The waveforms, zero points, and spurious components computed for various aperture types are not a pecularity of the scanning process but occur in simultaneous imaging with multiple apertures or figures of confusion as well. The effects are, therefore, accurately duplicated by out-of-focus imaging with lenses. Figure 21 shows that a point of light focused on a screen at a distance d is produced by a light beam which has a cross section shaped by the lens aperture. When the screen is moved out of focus (d_1 or d_2), the light spot increases and assumes the cross section and flux distribution of the lens aperture. The spot image is the "figure of confusion". The upper right-hand corner of

* Part I—page 35.

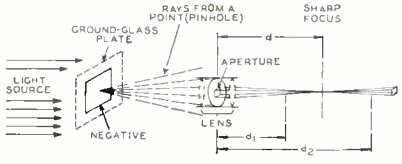


Fig. 21—Aperture effect on shape of point-imaging beam in out-of-focus imaging.

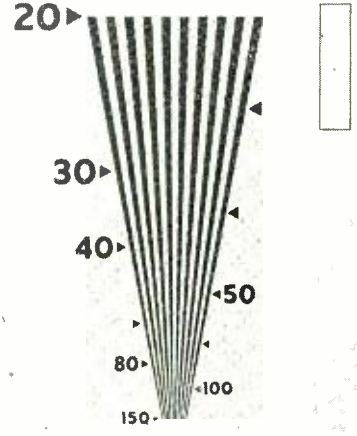


Fig. 22—Wedge test pattern and pinhole apertures in sharp focus.

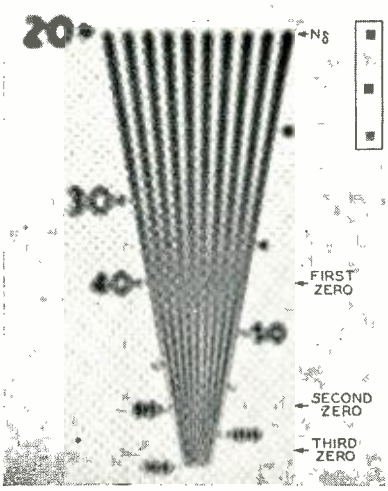


Fig. 23—Test pattern image formed by square aperture (defocused lens) as shown by pinhole images.

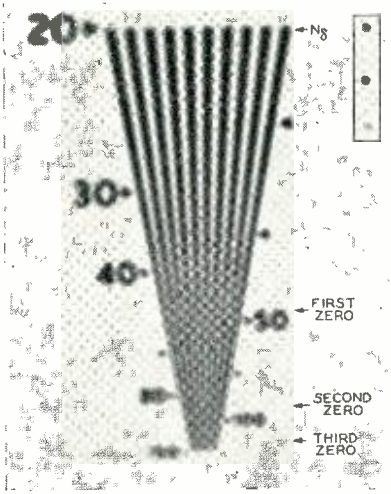


Fig. 24—Test pattern image formed by round aperture (defocused lens) as shown by pinhole images.

Figures 22 to 25 show the respective size and shape of the generating aperture or figure of confusion produced by three pinholes in the negative. In the wedge patterns of these photographs, the multiple zeros and spurious signals are plainly visible. Note the sharp crests (triangular wave) in Figure 23 at $N\delta$, the square shape near the zeros, and the phase shift and loss of one wedge line at every zero, all of which

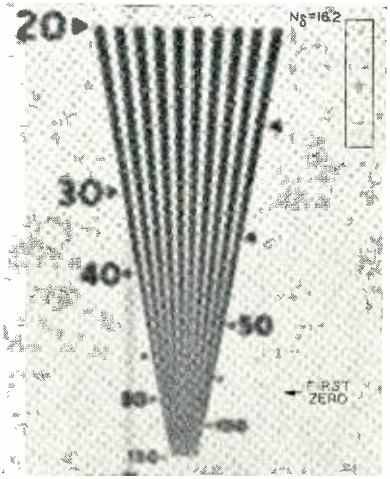


Fig. 25(a) — Test pattern image formed by round aperture with cosine-squared flux (defocused lens) as shown by pinhole images.

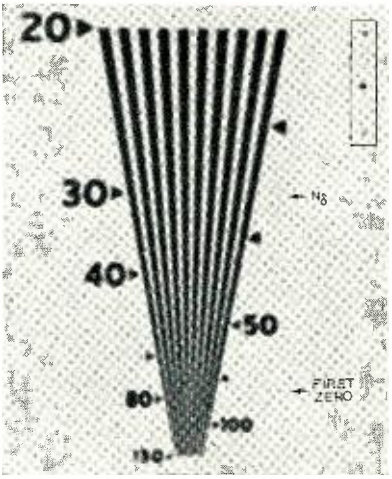


Fig. 25(b) — Test pattern image formed by round aperture with uniform flux computed to give a cutoff equal to that of Figure 25(a).

are in perfect agreement with computed values. The cosine-squared aperture with rotational symmetry (Figure 25(a)) was synthesized by multiple exposures with the following lens apertures.

$1/\delta \propto$ f:	5	6	6.8	8	9.6	12	14.4	19.2	32
exposure t (seconds)	0.5	1.	1.5	1.5	1.5	1.	1.	1.	1.

A 10-second exposure was made first at f/8.6 with a single round aperture ($N\delta = 30$) of equivalent flux, computed to have substantially the same cutoff point as the cosine-squared aperture with $N\delta = 16.2$ (Figure 25(b)). It served to set the out-of-focus lens adjustment which remained the same for the cosine-squared exposures. This cosine-squared distribution has rotational symmetry and applies to electrical

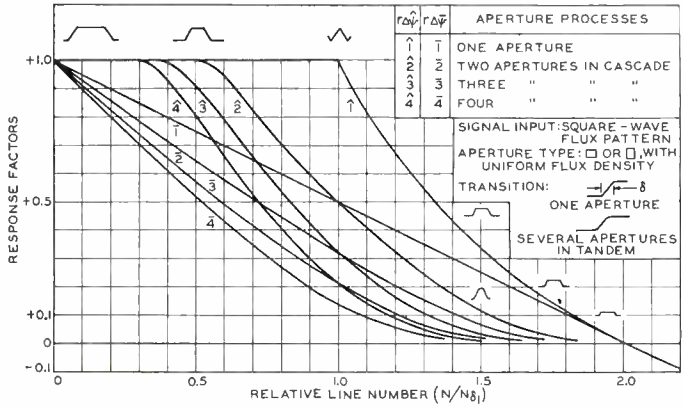


Fig. 26—Flux response of identical square apertures in cascade.

beams and lenses. It should not be confused with the shaped aperture shown in Figures 15 and 16 by Curve 3.

4. Aperture Processes in Cascade

The degradation of detail signals in multi-stage transducing processes is an effect caused by a progressive integration of the signal flux within aperture areas modified by their transmission or density characteristics. To obtain the over-all response characteristic of aperture processes in cascade, it is necessary to repeat the process of scanning with apertures. While the test pattern signal in the second operation is modified in waveform and amplitude by the first process, the signal to the third scanning operation is modified by the second process, and so forth. When starting with a square-wave flux pattern, each succeeding process rounds off the flux waveform and "stretches" the transition

curve produced by a unit function boundary in the original pattern. After several repetitions of the process, the response characteristics of all aperture types resemble each other closely as shown by the curve families Figure 26 and Figure 27 computed for two dissimilar aperture types (See Table IV in Appendix). The nature of the process and the substantially uniform contraction of the response characteristic occurring in each additional process, suggest a trial of the rule of squared sums for an approximation of progressive response characteristics. Cawein* observed that the scale contraction δ_p of (sine wave) response characteristics for a number of equal aperture processes in cascade follows this rule fairly well as expressed by

$$\delta_p{}^2 = \delta_1{}^2 + \delta_2{}^2 + - - - \delta_n{}^2. \tag{29}$$

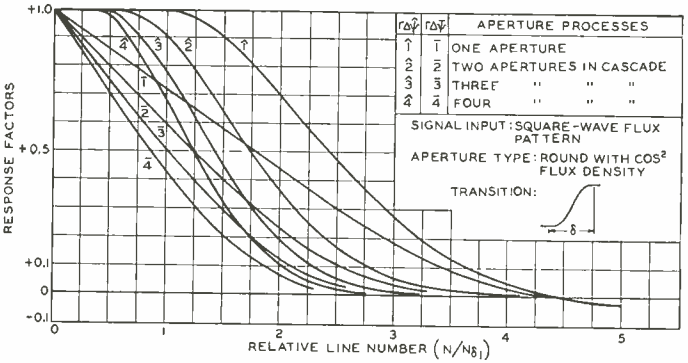


Fig. 27—Flux response of identical round apertures (cosine-square flux) in cascade.

In this expression δ_p is the "width index" or the equivalent aperture diameter of the over-all process and δ_1 to δ_n are the width indices or aperture diameters of the component aperture processes. The equivalent aperture of two identical processes in cascade, for example, is $\sqrt{2}$ times larger in diameter than the aperture of each single process. Examination of the true characteristics Figure 26 and Figure 27, discloses that this general rule does not apply accurately to strongly curved sections of the response characteristics $r\Delta\bar{\psi}$ or $r\Delta\hat{\psi}$, but holds well in the substantially linear part of the differential flux response curves $r\Delta\bar{\psi}$ even for widely different aperture types. The over-all aperture response of a number of aperture processes in cascade is obtained, therefore, with good accuracy from measured or computed characteristics $r\Delta\bar{\psi} = f(N)$ of individual processes by observing the following rules. An aperture response value $r\Delta\bar{\psi}$ is selected which

* Part I—Reference 3.

occurs in the substantially linear part of any one individual response characteristics. As characteristics of practical aperture processes have often a pronounced "knee" (See Figure 19) a point-by-point evaluation or an approximation of the characteristic by an equivalent linear characteristic with lower cutoff may be required if the "knee" occurs within the range of observation. Cutoff values (N_c) should never be used. The selected response $r\triangle\overline{\psi}$ (usually greater than 0.3) furnishes particular resolution line numbers N_1, N_2, $---N_n$ for the individual stages. The line number N_p at which the same response is obtained from the cascaded process is computed by the rule of squared sums:

$$N_p = 1/\sqrt{1/N_1^2 + 1/N_2^2 + --- 1/N_n^2}. \tag{30}$$

5. The Aperture Response in Sampling Processes

a. The Sampling Principle

Practical imaging processes analyze and reproduce the continuous light-flux distribution in optical images by a sampling method. The spatial distribution of the light-flux samples may be random, as in photographic film (grain), or it may be arranged to follow a regular pattern termed a "raster" as in printing and television processes.

In principle, samples of the image flux are taken by the "analyzing" aperture of a photosensitive transducer at image points specified by the raster. The integrated aperture flux is transduced into signal "amplitudes" which are transduced again into flux samples by the "synthesizing" aperture of the image-reproducing device. Proper spatial location of the samples in the re-created image is obtained by matching or "synchronizing" the rasters in the two processes. The amplitude of the reproduced flux samples is controlled by the signal amplitudes, but the area occupied by the sample is determined by the size of the synthesizing aperture. (In some cases, it is also a function of signal amplitude.)

The accuracy of the process in sampling the flux amplitude at any raster point is expressed by the aperture response of the analyzing aperture, but information on the flux between raster points is inaccurate or missing. To restore the continuity of signals between sampling points in the recreated image and to avoid interference by the structure of the raster, the missing information is approximated by increasing the dimensions of the aperture "depositing" the flux samples at the raster points (or lines) in proportion to the raster pitch distance $\triangle x$. With respect to image definition, the lack of information due to sampling as well as the increase of aperture dimension in the reproducing aperture cause loss of detail and can be expressed by respective aperture response factors. (See below)

To be useful, all imaging processes are followed by the process of vision. The aperture process of the eye can therefore perform the required widening operation and restore continuity of the image field when the raster pitch distance ($\triangle x$) is decreased below certain values. Small apertures with high response factors ($r\triangle\bar{\psi}$) can then be used to advantage in both transducers. The flux response of the three basic apertures of the sampling process in television systems will be treated in a subsequent section.

b. Line Number and Threshold Visibility of Raster Fields

The "vertical" aperture displacement in the television line scanning process occurs in small discrete steps $\triangle V$ which are inversely proportional to the "active" scanning line number N_V in the image. The stepwise displacement of the aperture path causes an image structure termed a "line raster".

Visibility of the raster in the image field depends on the aperture size of the image-forming transducer as well as on the aperture sizes of all succeeding processes through which the (kinescope) image is observed.

Continuity of the image field termed a "flat" field* or the illusion of seeing a flat field is produced by spaced "samples" when the sampling point spacing coincides with a zero response value N_c of the transducing or observing aperture. The equivalent television line number of a raster is $N_R = 2n_x$, where $n_x = 1/\triangle x$ is the number of raster points or lines in the vertical picture dimension. A minimum flux variation occurs for the condition

$$N_R = N_c = 2n_x. \tag{31}$$

Complete null points or flat fields without residual flux "ripples" are obtained only with certain apertures*. The out-of-focus projections of a fine line structure in Figure 28 demonstrate flat-field conditions (center and right side) for three apertures with uniform density. Inaccuracies in aperture shape caused residual ripples and slight shifts of the theoretical null points of the square and diamond apertures. Apertures with sharp cutoff have several critical flat field conditions, while apertures with gradual cutoff, such as the eye, maintain a flat field for $2n_x \geqq N_c$. The line width or aperture diameter (δ) of the kinescope can therefore be given any value when the raster line number $N_V = n_x$ satisfies Equation (31) where N_c is the limiting resolution of the eye. For a viewing ratio $\rho = 4$, the eye response as given in Figure 20 (See also Figure 8‡) is limited to $N_c = 1000$ lines. *A televi-*

* Part I—Reference 1.

‡ Part I—page 26.

sion line raster field with an active line number $N_v = 500$ lines will therefore appear continuous at $\rho = 4$.

It is apparent that the number n_x in a raster should be increased (by $\sqrt{2}$ for a 45-degree angle) to maintain continuity or eliminate visible serrations at the edges of slanting contours*. A step structure will therefore be visible at certain image flux distributions when the raster is dimensioned according to Equation (31). (See next Section.)

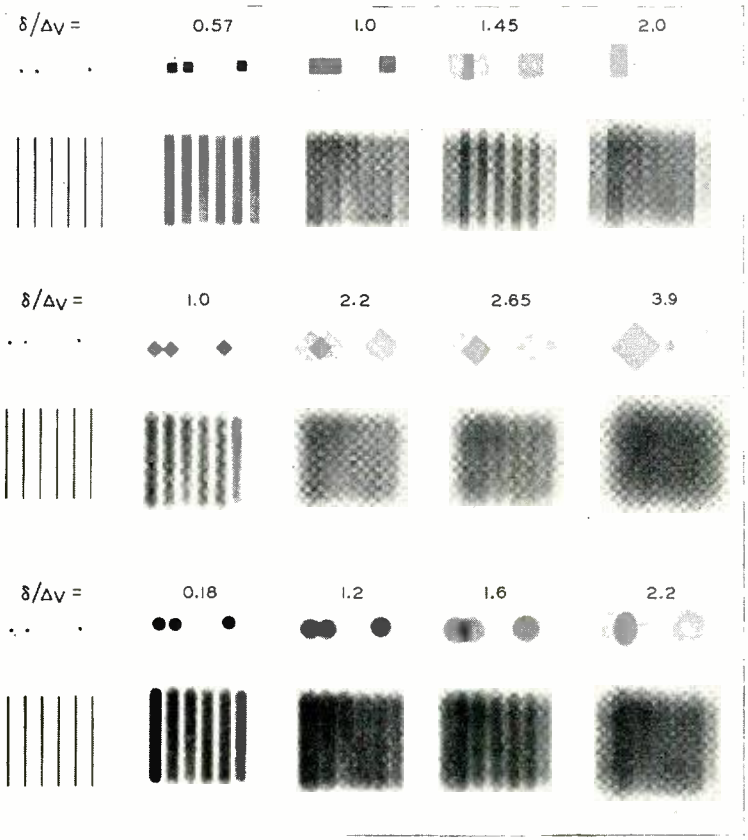


Fig. 28—Widening of raster samples by various aperture types and sizes.

c. The Sampling Process and Aperture Response in Television Systems

The signal amplitude from a photoelectric transducer with constant aperture size is proportional at any instant to the integrated light flux within the aperture area. When this "analyzing" aperture (A_1) is displaced in finite incremental steps $\triangle x$ or $\triangle V$, the generated signals

* Part I—Reference 1.

are exact "samples" of the aperture flux wave obtained by continuous displacement of the same aperture.

The "modulation envelope" of the sampling pulse wave is therefore specified by the normal aperture response factors $r\triangle\hat{\psi}$ and $r\triangle\bar{\psi}$. It may appear that the vertical aperture response of television scanning apertures is a special case because a continuous aperture displacement

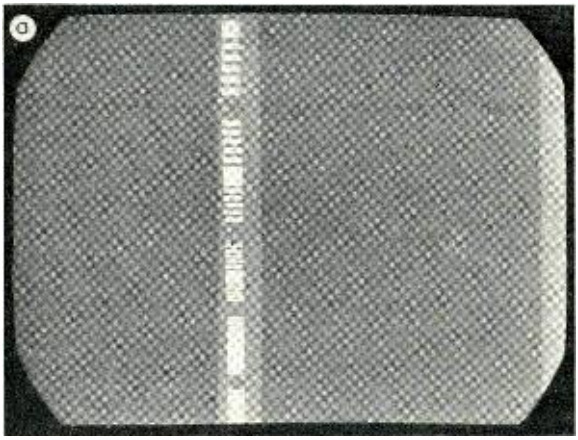
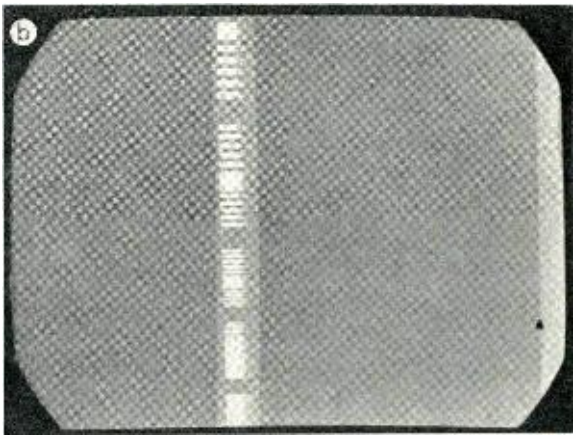


Fig. 29 — Kinescope images of "vertical" aperture-response samples.



occurs simultaneously in the horizontal direction. A line element $\triangle H$ of a cosine-square aperture, for instance, has a rectangular base and the shape given in Table I of the Appendix. Calculation of the (vertical) aperture response of this line element however discloses no change from the horizontal flux response of the aperture.

Normal aperture response characteristics can, therefore, be measured by incremental (vertical) scanning of square-wave test patterns. Their equivalence with characteristics obtained by continuous scanning

is demonstrated as follows. The beam of a 5-inch projection kinescope with short decay phosphor is used as a scanning aperture for generating signals in a light-spot scanning system. The aperture signals of the light spot are transduced by a 931-A multiplier phototube into electrical waveforms which are observed on an oscilloscope with cross section selector. The test object is a transparency of parallel-line groups which can be inserted vertically or horizontally into the light-spot scanner. The image of this object can be seen on a monitor kine-

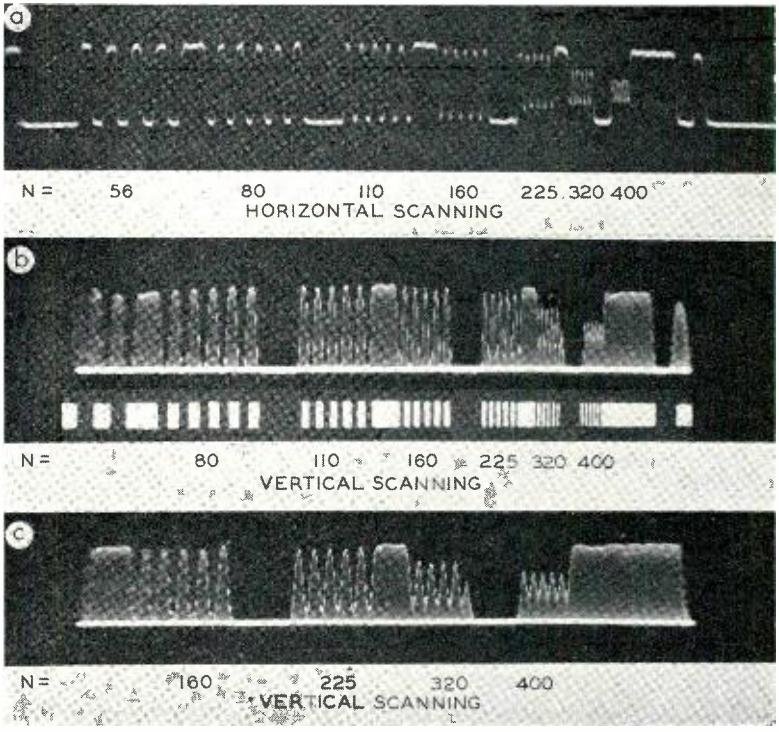


Fig. 30—Oscillograms of aperture response of a kinescope light spot ($N\delta \simeq 160$).

scope and appears for vertical scanning as shown in Figure 29. Trace (a) in Figure 30 is an oscillogram showing the light-spot aperture response as the modulation of the continuous (horizontal) line trace; while the traces (b) and (c) show the response of the same aperture in the vertical direction as the envelope of the scanning pulse wave generated by the step displacement $\triangle V$ of the scanning aperture. The ratio N_V/N of the active scanning line number to the test-object line number is the number of samples per test pattern line. In trace (c) Figure 30, the ratio N_V/N is considerably greater than unity for all

line groups, while in trace (b) it has been reduced to $N_V/N \simeq 1.25$ for the finest line group.

For better interpretation of waveforms, the number of samples can be increased by compressing the scanning raster in the light-spot scanner. This compression does not affect the aperture response or calibration which is set by the optical magnification of the lens system.

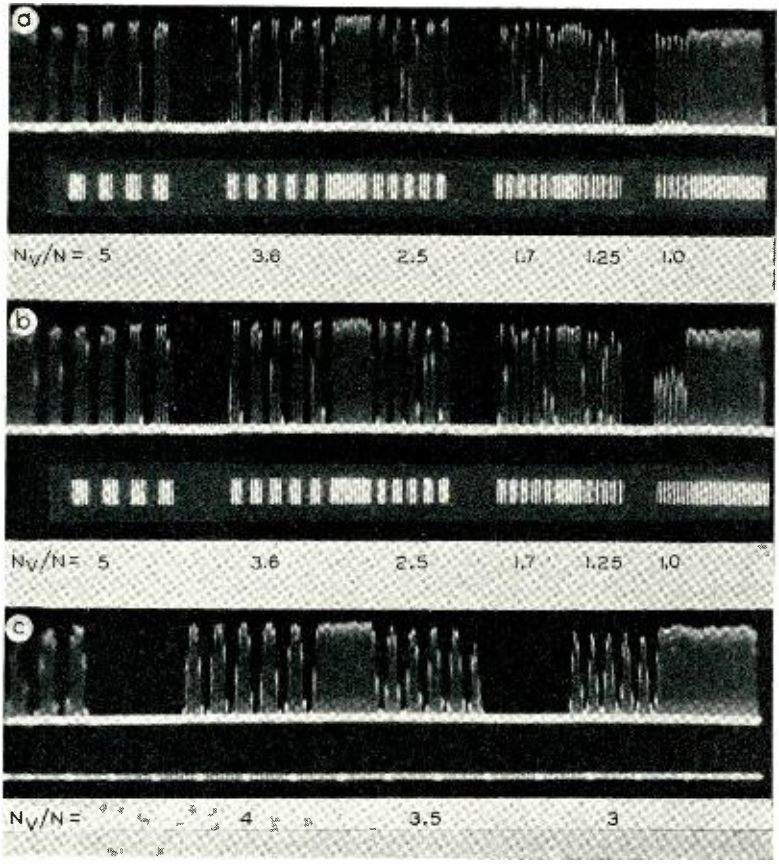


Fig. 31—Oscillograms of aperture response of a kinescope light spot ($N\delta \simeq N_v$).

The increased aperture response of a fine scanning aperture is shown in Figure 31. The traces (a) and (b) and the corresponding kinescope images also illustrate the effect of phase differences at $N = N_V$. Trace (b) shows the critical condition of sampling the null points of a square wave with a small aperture. Both figures show spurious signal components which are beat frequencies between N_V and N. Their pattern is most evident in the oscillograms because the pulse top is

recorded strongly by the oscilloscope, the fainter amplitude lines being the horizontal pulse sides. The amplitude lines which are invisible on the oscilloscope with sharp horizontal definition (Figure 29(a)) are made visible by decreasing the horizontal frequency response by a shunt capacitance (Figure 29(b)). The scanning tests verify that the normal aperture response of the scanning aperture can be determined accurately by the vertical sampling process.

The cascaded aperture effect of the principal apertures in a sampling system in which the eye serves as a "pulse lengthener" is illus-

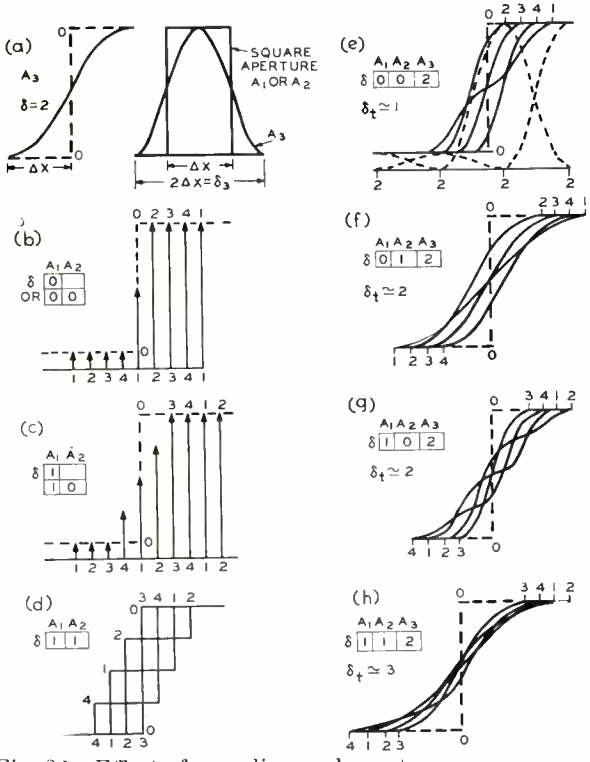


Fig. 32—Effect of sampling and aperture response on a unit function transition.

trated by a series of graphs in Figure 32. Aperture numbers and sizes are listed in the left corner of each graph. The symbols A_1 and A_2 indicate the apertures of the photoelectric (pickup) and electro-optical (kinescope) transducers. A_3 is an observing aperture simulating the eye. The aperture diameter δ of A_1 or A_2 is listed as zero for an aperture much smaller than the raster pitch and as unity when a square aperture $\delta = \Delta x$ is used. The aperture simulating the eye is shown as a line element with cosine-square cross section (See Figure 32(a)) to

produce a perfect flat field for samples of constant amplitude. Its diameter, therefore, is $\delta_3 = 2(\Delta x)$. The test object is a unit function flux change marked as $0 - - - 0$.[19]

The effects of changing the aperture size δ_1 from near zero to unity are shown in Figure 32(b) and (c) by the four phasing conditions (1 to 4). The small aperture produces only one intermediate signal amplitude (1) in the four positions shown. The amplitude is nearly always correct, but the phase may vary within one line pitch (Δx). This position error is the "raster effect" which may rate on the average an aperture value $\delta_r = 0.5$ (See later). The square aperture $\delta_1 = 1$ is correct in amplitude only once (3) and in error most of the time, the average transition length (envelope of amplitude points) rating one line pitch ($\Delta x = \delta_1 = 1$); it also has the same phase error ($\delta_r = 0.5$). Figure 32(d) is the effect of two square apertures $\delta_1 = 1$ and $\delta_2 = 1$ in cascade. It is correct in amplitude only once (3); it widens the transition to $\Delta x = 1$ on the average; and it has the phase error. The case of two small apertures is the same as (b).

More significant and readily interpreted differences are seen in the overall processes $A_1 + A_2 + A_3$ shown by the transitions (e) to (h) for which the respective transition length δ_t is indicated. The broken lines in (e) show the four sample points (2) from a small δ_2 widened to the aperture size δ_3. Their sum produces the curve 2. It is a curious fact that the transition δ_t is only one half as wide as that of A_3 alone ($\delta_3 = 2$); (the effect can be observed by experiments with real apertures and with different objects) but a position error within $\Delta x = 1$ is still in evidence. When one larger aperture is used, the transition is lengthened to $\delta_t \simeq 2$. Condition (g) may be rated as a little shorter transition, but it is more distorted than that shown by (f), where the sequence in sizes is reversed. Two larger apertures (Figure 32(h)) cause a further lengthening to $\delta_t = 3$. The phase- or position-error is progressively reduced. The reduction is a normal effect expected from the aperture addition in quadrature (Equation (29)). Computation of δ_t by Equation (29) furnishes the value

$$\delta_t = \sqrt{\delta_1^2 + \delta_2^2 + \delta_3^2 + \delta_r^2} = \sqrt{1 + 1 + 2^2 + 0.5^2} = 2.5$$

which represents the effective transition length quite well if the fact that the cutoff points are not accurately following Equation (29) in normal cascaded aperture processes (Figure 26) is considered. Similar

[19] The vertical brightness transition for a specified kinescope aperture and raster has been similarly evaluated: see R. D. Kell, A. V. Bedford and G. L. Fredendall, "A Determination of Optimum Number of Lines in a Television System", *RCA REVIEW*, Vol. V, No. 1, pp. 8-30, July, 1940.

answers have been obtained with other test objects (sloping transitions, irregular waves, and square waves). It may be concluded that, *on the average, the sampling process does not change the normal aperture response.*

d. The Aperture and Response Characteristics of the Raster

The phase or position error caused by the sampling process is better known as a "spurious" signal component or a "beat pattern". It is especially noticed in images having parallel slanting lines, circular line groups, wedges, etc. The position error disappears with larger apertures or line numbers when the aperture process preceding and including the first scanning aperture has a cutoff $N_c < n_x$ (See reduced phase error in central portion of Figure 32(h)). The effect of the phase error on sharpness may be rated as an aperture factor. In the directions of the raster pitch the phase uncertainty is, on the average, $\delta_r = 0.5\Delta x$. As the raster point- or line-distance Δx changes with angle, a 45-degree angle may be considered as an average value. The average phase error in the image is, therefore, 0.5 $\sqrt{2}$ times the raster pitch (Δx). If the eye rates a small position error as an equivalent decrease in sharpness, the phase error may be expressed by an equivalent "raster aperture". *The equivalent symmetric "raster aperture" of the sampling process is, hence, a square aperture with a tentative size*

$$\delta_r \text{ square} \simeq \Delta x/\sqrt{2}. \tag{32a}$$

The response factor $r\Delta\bar{\psi}$ of this aperture is 50 per cent at a line number N equal to $\sqrt{2}$ times the raster line number (n_x) and the aperture response curve (See Figure 16) is a straight line from $r\Delta\bar{\psi} = 1$ at $N = 0$ to

$$r\Delta\bar{\psi} = 0.65 \text{ at } N = n_x = N_V \tag{32b}$$

and has a limiting resolution $N_c = 2\sqrt{2}nx$. (32c)

The effect of the raster structure varies with picture content. It can be evaluated by subjective tests with normal image subjects on television systems without electrical channel limitation and known apertures. Such measurements have been made by Baldwin[5] with a television system employing two square apertures $\delta_1 = \delta_2 = \Delta x$. Equations (29) and (32a) furnish the value $\delta_p = \sqrt{1 + 1 + 1/2}\,\Delta x = 1.58$ for this process. Baldwin found by optical comparison that the equiva-

[5] See Part I.

lent optical square figure of confusion of his television images was $\sqrt{2.67}$ or $\delta_p = 1.63\ \triangle x$. His raster contained only 240 lines and in view of the difficulties of maintaining a flat field without visible structure with large square apertures, the agreement with the computed value can be considered as excellent.

The sampling process itself (with infinitesimal apertures) is a process with constant amplitude response from $N = 0$ to its nominal *sharp cutoff* at $N_{co} = n_x$. These characteristics are significant and outstanding, and cannot be duplicated by normal aperture processes. They are approached, however, quite closely by phase-corrected electrical filters with sharp cutoff. This and other similarities permit the deduction that an electrical channel can be replaced by an equivalent two-dimensional aperture.* The response characteristics of the electrical channel and its equivalent aperture are discussed in a subsequent section.

6. Random Fluctuations and the Aperture Effect of Grain Structures

Random fluctuations (noise) of the image signal have, in principle, no effect on the aperture response of an imaging system. For a given finite time of observation, a high fluctuation level will obscure detail signals to a greater extent than large area signals, but it does not change the *ratio* of detail to large area signals which is specified by an aperture response factor. A high fluctuation level may limit a response factor measurement by uncertainty of observation to relatively large values (for example to $r\Delta\overline{\psi} > 0.5$) while a low "noise" level permits accurate measurement of the cutoff resolution $r\Delta\overline{\psi} = 0$.

The elemental structure found in all signals (unit charge, electron, diffraction disc, grain), however, causes an aperture effect which can be observed with sufficient "magnification". The number of elemental samples or signal "units" occurring per unit area and time of the pass band is a measure of the fluctuations in the signal, i.e., it determines the signal-to-noise ratio. In the case of electrical signals the aperture effect of the elemental unit is so small that it is negligible within the limits of optical pass bands. Fluctuations, therefore, may be observed but there is no aperture effect due to electron size. When the elemental unit is of a size which can be resolved by the optical pass band (grain or particle sizes of photographic emulsions, pickup tube mosaics, kinescope screen materials) its aperture effect is noticeable and affects the aperture response of the imaging process. The existence of an aperture effect does not exclude the possibility of a low fluctuation level as it does not necessarily specify the number of grains per unit area or

* This treatment is permissible when the asymmetry of the aperture of the over-all process remains within certain limits; see Reference 5, Part I.

time, which determines the signal-to-noise ratio. The aperture response of a number of grain structures has been measured with a television microphotometer and will be discussed further in Part III of this paper.

B. General Specifications for Television Systems

1. *Line Raster and Aperture Size*

The analysis of the sampling process has shown that it is unnecessary to limit the diameter of the scanning apertures in television systems to one line width ($\triangle V$) when the eye can perform the required widening operation to obtain a "flat" field. For a viewing ratio $\rho = 4$, Equation (31) states that an active line number $N_V \geqq 500$ is satisfac-

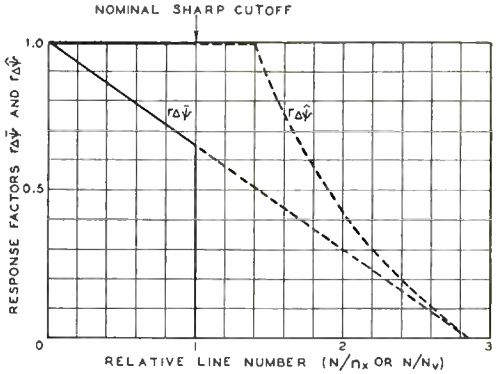


Fig. 33—Flux response characteristics of the equivalent two-dimensional square aperture of a raster process.

tory for use with small-diameter scanning apertures. The equivalent aperture of the raster itself has been evaluated in terms of a *two-dimensional square aperture* $\delta_r = 0.7/N_V$ (Equation (32)) with the aperture response shown in Figure 33. The increase in sharpness obtained with small apertures as compared to apertures covering a flat field in both transducers is illustrated by Figures 34 and 35 which should be viewed at a distance which gives threshold visibility of the line structure in Figure 34.

2. *The Equivalent Optical Aperture of the Frequency Channel and Theoretical Limits of Television Systems*

The definition of television systems is limited in the horizontal dimension by an electrical frequency channel ($\triangle f$). Electrical filters can be designed to have a considerably sharper cutoff than normal optical apertures. The square-wave response factors of the theoretical band pass filter with sharp cutoff can be derived from the Fourier components of a square wave, which contains only odd harmonics ($n = 1,3,5---$) decreasing in amplitude according to order

Fig. 34—400-line television picture with small scanning apertures in camera and kinescope.

Fig. 35—400-line television picture with scanning apertures increased to give a "flat" field in camera and kinescope.

$$Y_n = 4/\pi n. \tag{33}$$

A finite frequency channel limits the frequency components in the square wave. A square line repetition occurring at one-fifth cutoff frequency can therefore generate a signal containing only the first, third and fifth harmonic. For repetition frequencies greater than one-third cutoff frequency, the channel will pass only the fundamental frequency.

The theoretical wave shapes for uniform phase delay up to cutoff frequency are shown in Figure 36(b). The square wave from a finite

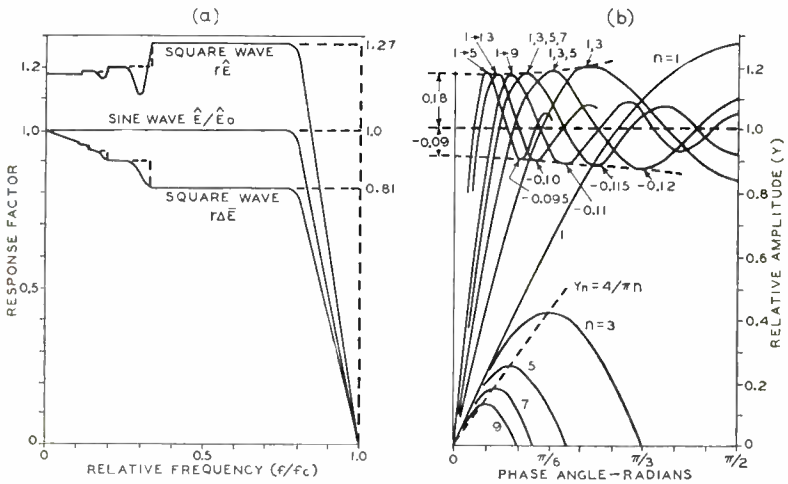


Fig. 36—Square-wave response characteristics of limited frequency channels.

frequency channel exhibits, therefore, "underswing" and "overshoot" preceding and following the rise or drop in current because of missing high-frequency components. These peaks change or disappear when the channel exhibits non-uniform phase delay or a gradual cutoff.

The amplitude response $r\hat{E}$ of the theoretical sharp-cutting channel is, therefore, greater than unity as shown by the dashed curve in Figure 36(a), exhibiting discontinuities where f_c/f is an odd number. The effect of a sloping cutoff (solid curve in Figure 36(a)) having no phase error and beginning at $f/f_c < 1$ results in dips in the amplitude response at the harmonic points. Because of phase shift errors, these dips usually do not occur with practical cutoff filters.

The equivalent square-wave response $r\Delta\bar{E}$ obtained by averaging the response per half cycle, remains at 81 per cent after an initial drop

in the first third of the frequency channel until the amplitude response begins to drop. The similarities of the electrical response characteristics and the vertical sampling process are apparent and it may be expected that the equivalent symmetric aperture of the electrical channel is in the same order as that of the sampling process.

The overshoot and following transient oscillations impose a restriction on the minimum horizontal line number N_H because the third and following odd number peaks cause multiple contours in the image and should be near threshold visibility at $\rho = 4$. Lower line numbers require reduction of the equivalent square-wave brightness step, $\triangle B$, between the second and third transient peaks by a more gradual cutoff in the electrical channel. Threshold visibility of a peak-to-peak brightness fluctuation, $\triangle B$, is obtained (Equation (11)*) when $\triangle B/B = 0.018/r_e$.

The value $\triangle B$ is proportional to the equivalent square-wave value $\triangle \hat{E}_{2,3} = (2/\pi)\, \hat{E}_{2,3}$ of the second and third transient peaks, which is 56 per cent[20] of the first and second peak-to-peak value $\hat{E}_{1,2}$ as expressed by

$$\triangle B \propto 1.12\, \hat{E}_{1,2}/\pi$$

A brightness value $B = \hat{B}/2 = E/2$ may be considered as an average value for unit function steps of various amplitudes and levels, when E is the signal voltage corresponding to a unit function from zero to $\hat{B}$. The threshold visibility requirement is hence expressed by

$$\triangle B/B \propto 2.24\, \hat{E}_{1,2}/\pi E = 0.018/r_e$$

$$\hat{E}_{1,2}/E \simeq 0.025/r_e. \qquad (34)$$

Equation (34) states the requirement for a television system with unity aperture response. When the response is limited by finite aperture sizes in the camera chain and kinescope, the frequency response of the channel can be corrected to compensate the decrease in horizontal signal components due to the aperture effect of the transducers. The correction requires a rising frequency response of the video amplifier (high-peaking circuits) and appropriate phase correction and is termed "aperture correction". The over-all response in the horizontal direction is then again determined by the requirements for unity aperture response (Equation (34)).

The degree of correction obtainable in practical systems depends on the signal-to-noise ratio. A high aperture response is of course still desirable as this aperture correction is not effective for vertical signal

* Part I—page 29.

[20] Jahnke, Emde, TABLES OF FUNCTIONS (Si (X) Function). Dover Publications, New York, N. Y., 1943.

components. (The effects of aperture asymmetry on the over-all response will be discussed in Part IV.)

The present standard television channel $\triangle f = 4.25$ megacycles accommodates the line numbers $N_V = 490$ and $N_H = 340$. The response factor of the eye is $r_e = 0.225$ (Figure 20) at 340 lines and Equation (34) furnishes the maximum transient ripple percentage $\hat{E}_{1,2}/E = 11.0$ per cent.

The relation between sine-wave amplitude and transient response of electrical filters has been treated in detail by various authors[19,21,22]. Typical response curves and corresponding ripple percentages ($\hat{E}_{1,2}/E$) are shown in Figure 37. Curve 4 indicates the approximate channel response for a (standard) television channel with $N_H = 340$ lines giving a threshold transient ripple of 11.0 per cent.

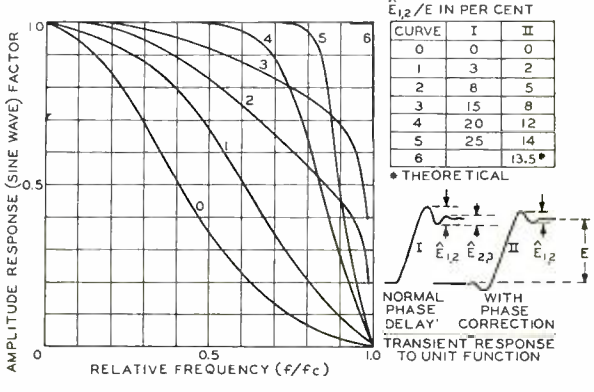


Fig. 37—Amplitude response and transient response of electrical channels.

An equivalent aperture size of the frequency channel is indicated by the length and shape of the transition curve obtained for a unit function signal. The S-shaped transition curve indicates a cosine-square aperture type (Table II, Appendix) with a diameter $\delta_H = 1/N_H$ for the square-cut frequency response curve. The aperture diameter increases with decreasing frequency response of the channel and has been plotted in Figure 38 as a function of transient ripple voltage (data from Kell et al. and Kallmann et al.[19,22]). The aperture diameter is indicated by the line number $N\delta_H$ at which the flux response factor $r\triangle\overline{\psi}$ is 71 per cent. *The equivalent two-dimensional aperture* δ_f *of the frequency channel* is specified by the horizontal component of a 45-degree

[21] A. V. Bedford and G. L. Fredendall, "Analysis, Synthesis, and Evaluation of the Transient Response of Television Apparatus", *Proc. I.R.E.*, Vol. 30, October, 1942.

[22] Kallmann, Spencer & Singer, "Transient Response", *Proc. I.R.E.*, Vol. 33, March, 1945.

line*

$$\delta_f = \delta_H/\sqrt{2} \tag{35a}$$

with the equivalent line number

$$N\delta_f = N\delta_H\sqrt{2} \tag{35b}$$

The equivalent optical aperture of a (standard) television channel with $N_H = 340$ lines is, hence, a cosine-square round aperture with $r\Delta\bar{\psi} = 0.71$ at $N\delta_f = 0.75 \times 340\ \sqrt{2} = 360$ lines. 50-per cent response (See Figure 16) occurs at $N_{(0.5)} = 630$ lines.

The equivalent optical aperture of a television system with ideal components is the *cascaded* value of the raster aperture δ_r and the channel aperture δ_f. For the line number $N_r = 490$ and $N_H = 340$ of the standard 525-line system, the line number at $r\Delta\bar{\psi} = 0.5$ computed with Equation (32) and Figure 33 is $N_r = 490\ \sqrt{2} = 690$ for the raster

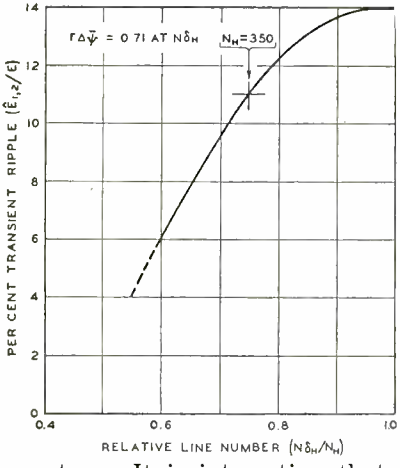


Fig. 38—Aperture size ($1/N\delta_H$) of of frequency channels as a function of transient ripple.

aperture. It is interesting that it is nearly equal to the channel response which was found to be 630 lines for this response factor.

The equivalent optical aperture of the standard 525-line television system and frequency channel ($\Delta f = 4.25$ megacycles) has, therefore, a response $r\Delta\bar{\psi} = 0.5$ at a line number $N = 465$ lines as a theoretical limit. (Computed with Equation (30).)

3. Brightness, Repetition Rate, and Flicker in Television Images

a. Brightness and Flicker of the Image Field

Adequate discrimination of detail is known to require a high-light image brightness equal to or greater than 15 foot-lamberts, because the average brightness $\overline{B}$ should be greater than 2 foot-lamberts for operation with a normal transfer factor g_e of 12 on the eye transfer characteristic (Figure 8‡).

* This choice is indicated by the addition of apertures in quadrature.
‡ Part I—page 16.

The illusion of seeing the cyclic light-spot synthesis of television image fields without brightness fluctuations depends on the persistence of vision and requires a certain minimum field repetition rate or "critical flicker frequency" (C.F.F.).

The C.F.F. for intermittent light flashes having a fixed exponential decay times (t) can be obtained by replotting Figure 6‡, as shown by Figure 39. The three frequency scales on the ordinate of Figure 6

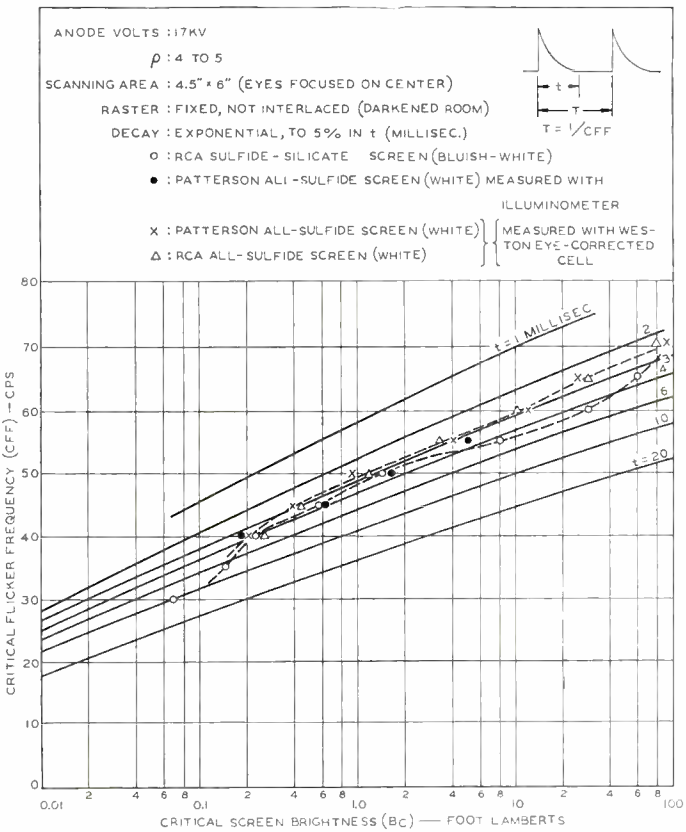


Fig. 39—Critical screen brightness of kinescope fields (white light) as a function of field frequency.

indicate the effect of a change in viewing ratio (ρ). A low ratio causes exposure of the rod-populated peripheral region of the retina which is more sensitive to flicker. These curves may be expected to indicate, in general, the slope of kinescope flicker curves, but it can hardly be expected that the impulse light wave produced by a single high-speed scanning spot is directly equivalent to a simultaneous impulse excita-

‡ Part I—page 22.

tion of the entire screen area. The peak brightness of the kinescope light spot is extremely high and saturation effects are likely to occur at high average values of brightness, $\overline{B}$, in the eye as well as in measuring devices. Furthermore, the decay of sulfides and mixtures of sulfides with silicates, deviates from an exponential law, and is also a function of current density and excitation time. *The time constants and C.F.F. of practical phosphors for the generation of white or colored light must, therefore, be determined by observation under actual operating conditions.*

The C.F.F. for unmodulated white light from various kinescopes (uniformly bright screen) has been measured and is indicated in Figure 39. This figure shows that all-sulfide screens permit a critical flicker brightness (B_c) of 10 to 13 foot-lamberts at 60 cycles, while sulfide-silicate mixtures permit a B_c of approximately 30 foot-lamberts.# These values increase by a factor of two for large viewing ratios (See ordinate scales of Figure 6). All measurements were made with a viewing ratio of 4 to 5 with eyes fixed at the center of the scanned area (4.5×6 inches) of actual television rasters. Care was taken to eliminate all spurious electrical fluctuations such as 60-cycle ripple, line fluctuations of the raster, and spurious interline flicker. The scanning raster was generated without interlacing but with interlocked frequencies in order to obtain a stationary line raster of 400 to 500 lines. Variation of spot size from 0.010 to 0.1 inch showed little effect on the C.F.F. in 500-line rasters. The screen brightness or rather "luminosity" B_c was measured directly with a Weston eye-corrected cell which follows closely the luminosity curve of the eye (Figure 5‡). These readings are somewhat lower than illuminometer readings on white light.* Sufficient time was allowed for adaptation of the eye and adjustment of pupil diameter. Repetition of measurement was advisable for consistency of observation, especially at higher field-repetition rates.

The critical flicker brightness of kinescope images is a function of the light distribution in the image. Modulation of the kinescope light-spot intensity breaks up the flickering area into smaller areas, for which viewing ratio (ρ) and critical brightness have higher values. The critical peak brightness $\hat{B}_c$ of the television image may, therefore, be considerably higher, especially when the high-light areas are small.

The indicated decay time is only an *equivalent* value and does not indicate that the light from the material has actually decreased to 5 per cent after t.

‡ Part I—page 21.

* The discrepancy is possibly an integration error of the cell due to its saturation characteristic. For the same reason, illuminometer readings made with a brightness reducing filter may give a higher value than the actual luminosity to the eye.

It has been observed on normal television pictures that the average brightness level of the image can be raised to a value equaling, roughly, that of the critical flicker brightness B_c for unmodulated light without obtaining objectionable flicker of the image. The average screen brightness $\bar{B} = B_c = 10$ to 30 ft-L for present phosphors and for a field frequency of 60 cycles per second can therefore, be considered as a satisfactory value which permits a high-light brightness ($\hat{B}$) in the order of 50 to 150 foot-lamberts in normal television images.

b. Color Flicker

It has been checked by observation on scanned rasters that B_c and the C.F.F. remain substantially constant for a fixed light-color and repetition period independent of the time of scansion (T_v) of the raster, as long as the decay time (t) is less than the repetition period (T). (See sketch in Figure 39.)

Image fields scanned within the time T_v equal to or smaller than T (say $T_v \leq 1/60$ second) but repeating after the same time T (say $T = 1/60$ second), appear to have substantially the same values of B_c and C.F.F.‡ In a sequential color transmission system using three primary colors, one bright field is scanned in $T_v = T/3$. The critical color field brightness of the primary colors (red, blue, green) generated with composite sulfide screen materials has been checked on television rasters by viewing and measuring the luminosity of the kinescope light (white) through Wratten color filters.** (Figure 40). It should be emphasized that the luminosity differences in Figure 40 apply to the particular screen material and may be caused by differences in decay time of the components. Relative energy and luminosity (standard eye response) of the components for white light from fluorescent "daylite" lamps and kinescopes are given in Table II in round figures as obtained with a monochrometer and an eye-corrected cell, respectively.

The color filters reduce the energy from the kinescope to 12 per cent, i.e., their effective equalized filter factor is 8.3 (approximately). The energy passed by any one single color (say green) is thus 4 per cent of the kinescope output because transmission occurs only every third field in one color cycle.

The critical brightness for unmodulated light in a sequential color

‡ Observations were made with $T = 1/60$ second, $T_v = 1/60$ second and 1/180 second for white and green light by four observers. One required a 25 per cent reduction of B_c at $T_v = 1/180$ second, but all reported no change in the white to green ratio, which was found to be unity (See W/G in Figure 40).

** Measurements of B_c were made with the Weston eye-corrected cell because color-comparison readings by illuminometer vary widely with different observers.

Table II—Relative Energy and Luminosity of Color Primaries for Two Light Sources.

	Type of light	Transmission through Wratten Filters			
Wratten Filter		#47(B)	#58(G)	#25(R)	#26(R)
Relative Energy	Fluorescent "daylite"	11%	11%	22%	20%
Relative Luminosity	Fluorescent "daylite"	3.8	26	14.8	13

		Approximate values for obtaining white light from kinescope through rotating filters.		
Wratten Filter		#47(B)	#58(G)	#26(R)
Relative Energy	Kinescope light	12%	12%	12%
Relative Luminosity	Kinescope light	3.5	28	7

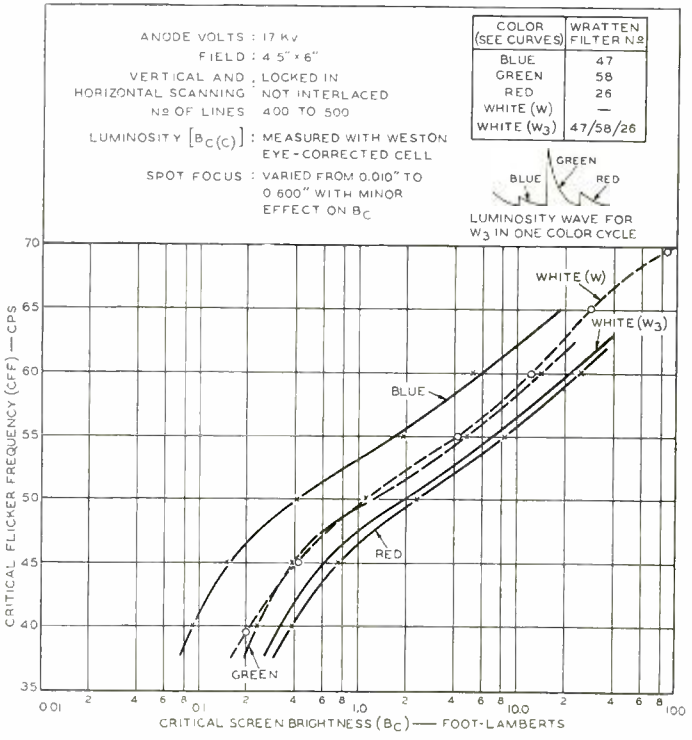


Fig. 40—Critical screen brightness of kinescope fields (colored and white light) as a function of field frequency.

transmission with a C.F.F. of 40 color cycles is restricted for a viewing ratio of 4 to the values $B_{c(G)}$ of 0.2 foot-lamberts and a synthetic white $B_{c(W3)}$ of 0.32 foot-lamberts which must be considered inadequate for easy vision. Normal values, $B_{c(G)} \geqq 6$ foot-lamberts and $B_{c(W3)} \geqq 9.6$ foot-lamberts, require a C.F.F. $\geqq 56$ color cycles (168 fields per second).

c. Interline Flicker

Interline flicker may occur due to sequential displacement of part rasters by one or several lines. In a raster interlaced 2 to 1, black-and-white lines interchange position in alternate fields at a frequency equal to one half the field repetition frequency. The critical brightness $B_{c(N)}$ per half raster is one half of the average screen brightness $B_{c(frame)}$ for unmodulated white light and is obtained by dividing the screen brightness (B_c) from Figure 39 or 40 by the resolution factors of eye and kinescope. The kinescope response factor is to be multiplied by $\sqrt{2}$ as it is effective only in one dimension.* The critical frame brightness is then given by

$$B_{c(frame)} = 2B_{c(N)}/\sqrt{2}\, r_e r_k \tag{36}$$

For a black-and-white transmission with $1/T_f = 30$ cycles per second, $N = 500$ per half raster and Figure 39 and Equation 36 furnish $B_c(500) = 2 \times 0.1/0.04 = 5$ foot-lamberts for the kinescope resolution shown in Figure 13‡. This value increases with increases of viewing ratio when the kinescope light is modulated.

d. Detail Flicker

A detail flicker effect caused by the interlacing process occurs upon reproduction of high-contrast detail having a vertical dimension in the order of one scanning line or less. A narrow dark line in a white background, for example, located substantially parallel to a scanning line will be reproduced only in one interline field and, therefore, the white lines adjacent to the dark line appear with a flicker frequency equal to the frame frequency because the interline remains dark. A relatively short line is reproduced with full visibility, as $N = 75$ for 1/100 of the horizontal line length. Near the short line, therefore, the critical brightness is reduced to less than 1 foot-lambert without benefit of filter factors ($r_e = 1$ and $r_k = 1$).

For an interlace ratio of 2:1, the flicker will disappear when the line object is scanned in both interlaced fields: i.e., when the vertical resolution is decreased to one-half by defocussing the camera lens or the scanning beam of the pickup tube. It is not eliminated but actually

* The kinescope is here treated as an object.

‡ Part I—page 35.

increased in intensity by defocussing the kinescope scanning beam as the flickering detail area is enlarged. This type of flicker is often confused with inaccurate or unstable interlacing because it causes the effect of a vibrating contour or line structure. (observable in the horizontal line wedge of a test pattern.)

Elimination of the effect without loss of resolution requires reduction of B, a kinescope phosphor with longer persistence, or an increase in frame frequency.

4. Kinescope Characteristics

a. Kinescope Transfer Characteristics

The transfer characteristic ($B = f(E)$) of normal kinescopes follows approximately a 3rd-power law in its central section. Current distribution (masking) and/or saturation effects of screen materials may cause a decrease at higher brightness values as exemplified by Figure 41. In the low brightness range, the characteristic is normally exponential unless modified by "variable-μ" effects of the control element.

It is desirable to standardize a "normal" characteristic for the kinescope because its inverse determines the signal compression required at the transmitter for maintaining a specified over-all transfer characteristic for the system.

The brightness range in directly viewed or projected television images is greatly reduced by illumination of the screen by ambient front or rear light. The amount of such contrast losses depends on the reflection factor of the screen.[23,24] The ranges indicated in Figure 41 have been measured in a dark room on a kinescope with settled screen and a test pattern image containing roughly 60 per cent black areas. The larger range 70:1 is the contrast measured between the white background and a point ¼ inch inside a black rectangle (1 × 2 inches) in an 8 × 10 inch image. The smaller range 50:1 was obtained close to the edge of the rectangle.

This relatively good contrast range decreased to 13:1 when the image contained only 10 per cent black in a white background because of increased rear illumination by light reflected from the bulb walls.

Recently developed kinescopes with aluminum backing[25] eliminate the rear illumination completely and in a dark room maintain a constant

[23] R. R. Law, "Contrast in Kinescopes", *Proc. I.R.E.*, Vol. 27, pp. 511-524, August, 1939.

[24] V. K. Zworykin and G. A. Morton, TELEVISION, John Wiley and Sons, New York, N. Y., 1940.

[25] D. W. Epstein and L. Pensak, "Improved Cathode-Ray Tubes with Metal-Backed Luminescent Screens", *RCA REVIEW*, Vol. VII, No. 1, pp. 5-10, March, 1946.

large-area contrast range of over 100 to 1 which is independent of the ratio $\hat{B}/\overline{B}$. For normal peak-to-average ratios $\hat{B}/\overline{B} = 4$ to 6, the ambient rear illumination is quite small in correctly designed kinescopes and aluminum-backed screens show, therefore, little, if any increase in general contrast for a given screen material.

Aluminizing of the screen rear surface has no effect on the contrast reduction caused by external "front" illumination of the viewing screen. The contrast loss due to front illumination (E_o) is controlled by the reflection factor (K) of the screen material and the optical transmission (τ) of materials (glass, filters) interposed between the screen and the source of ambient light. The effect of ambient room illumination (E_o)

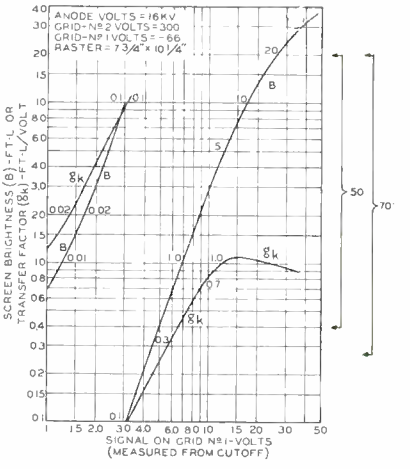


Fig. 41—Kinescope transfer characteristics.

on the contrast C obtained for $E_o = 0$ is given by

$$C' \simeq (\hat{B} + K\tau E_o)/(B_{\min} + K\tau E_o + 0.05E_o) \qquad (37)$$

with $\hat{B}/\overline{B}$ min = Contrast in dark room with $E_o = 0$

$\hat{B}$ = Peak brightness K = Reflection factor of screen material

τ = Transmission factor of filters placed over screen (neutral or color)

$0.05\,E_o$ = Reflection from filter or glass surfaces.

The factor K varies from $K = 0.6$ to 0.8 for present kinescope screens. A neutral filter or coating on the tube surface attenuates the ambient light by its transmission factor τ, but it also attenuates the kinescope light. Table III shows the effect of room illumination ($E_o = 1$ and 3 foot-candles) for a tube with a contrast C_o of 100 to 1, K of 0.7, and for various values of τ. The spectral qualities of kinescope

Table III—Effect of Ambient Room Illumination (E_a) and Filter Transmission (τ) on Brightness ($\hat{B}'$) and Image Contrast (C').

	$E_a = 1$		$E_a = 3$		$E_a = 1$		$E_a = 3$		$E_a = 1$		$E_a = 3$		$E_a = 1$		$E_a = 3$	
$\hat{B}$	C'	$\hat{B}'$	C'	$\hat{B}'$	C'	$\hat{B}'$	C'	$\hat{B}'$	C'	$\hat{B}'$	C'	$\hat{B}'$	C'	$\hat{B}'$	C'	$\hat{B}'$
20	22	21	10	22	34	10			47	5			60	2.5		
40	35	41	17	42	50	20	26	20	64	10	38	10	75	5		
80	52	81	28	82	67	40	40	40	78	20	54	20	86	10	67	10
120											64	30	90	15	75	15
	$\tau = 1$				$\tau = 0.5$				$\tau = 0.25$				$\tau = 0.125$			

and room light are assumed to be identical. The kinescope brightness is then reduced by the same factor τ to $\hat{B}' = \tau\hat{B}$.

The underlined values show the gain in contrast range (C') for a normal transmitted peak brightness ($\hat{B}'$) of 20 foot-lamberts when the (filter) transmission (τ) is decreased at the expense of kinescope power.

It is seen, however, that the image contrast, when $\hat{B}'$ is 20 foot-lamberts and τ is 0.25, differs little from that when $\hat{B}'$ is 40 foot-lamberts and τ is 0.5. The same kinescope power is required for both values ($\hat{B}$ = 80 foot-lamberts), but the brightness is twice as great for $\tau = 0.5$ and the image contrast, therefore, will appear just as good or even better to the eye because of the higher transfer factor of the eye characteristic (Figure 3(a)*).

The last column in which τ is equal to 0.125, is representative of the conditions for sequential color transmission and indicates that the kinescope brightness $\hat{B}$ must be in the order of 100 foot-lamberts for a normal image brightness.

b. The Aperture Flux Response of Kinescopes

Detail contrast and aperture-response characteristic of the kinescope are a function of the electron-beam diameter, electron- and light-scattering effects in the screen material, and reflections on glass surfaces (halation). The aperture response factor r_k ($r\Delta\bar{\psi}$) can be measured by the threshold visibility method (See Part I for discussion of response characteristic of eye) by fading-in a test pattern or sine-wave signal. The eye is used as an indicator to detect a fixed value of thresh-

* Part I—page 16.

old contrast. The screen is, therefore, viewed at a close range and the viewing angle, including the signal lines or bars, is maintained substantially constant. A magnifying glass is used for higher line numbers to insure a fixed threshold-contrast value for the eye. The response characteristic of a kinescope measured by this method is shown in Figure 13.*

The curve shape can be analyzed as the sum of two cosine-square aperture characteristics (See Figure 19) from which the corresponding beam-current distribution, line cross section, and transition curves are easily computed. (See Section on Complex Apertures).

A second and more reliable method to measure kinescope aperture characteristics is by signal generation in a light spot scanning system. The general principle has been outlined in the section on aperture processes. Because normal kinescope screen materials have a relatively long decay, measurements are made at low frequencies by observing the vertical aperture response. (See Figures 29 to 31). The signal amplifier must be corrected for effects of the phosphor decay which is compensated to allow decay of the light output within one line length. Trailing should, in general, not exceed ½ scanning line length and the scanned bar length (slit width) should be constant for all values N. A vertical slit mask is placed over the kinescope or the test pattern bars for observation of their cross section (See Figure 29). The slit has a constant width of 1/10 to 1/30 line length. The optical magnification ratio between kinescope "spot" diameter and test bar width determines the line number calibration. If, for example, the slide has a bar pattern containing 10 lines per millimeter and the kinescope screen has a vertical size $V = 180$ millimeters, an aperture response test at $N = 500$ requires the optical reduction $\frac{1}{M} = \frac{180}{500} \times 10 = 3.6$ between kinescope and test pattern. With M properly adjusted, the actual scanning line number N_V and vertical deflection on the kinescope under test are relatively unimportant and the raster can be compressed for obtaining a sufficiently continuous "modulation envelope" on an oscillograph operating with a "vertical" deflection speed. The oscillograph beam is keyed on by a short "horizontal" pulse (double-frequency synchronizing pulse) occurring in the center of scanning lines to eliminate all oscilloscope traces except for a vertical strip through the center of the scanned area. (The combination of vertical slit mask and oscilloscope pulse is a vertical line selector). The signal response from kinescopes used for image reproduction should be confined to visible light. A light filter (Wratten #2A) eliminating ultraviolet

* Part I—page 35.

radiation may be required because the ultraviolet tends to give increased resolution from certain phosphors (not the case for standard sulfide screens).

Several high-voltage kinescope characteristics measured by this method are shown in Figure 42. Curve 1 shows the beneficial effect of high cathode "loading" (high E_{C_2} voltage) which causes a considerable increase of $r\Delta\bar{\psi}$ in this tube. This experimental tube operated at 17 kilovolts represents a close approach to the ideal kinescope for a 500-line picture although an increase in response can still be obtained if aberrations can be eliminated. Curves 2a, b, c, illustrate that the aperture size may increase as a function of beam current, a well-known fact which may cause "blooming" (overlapping of traces) at high intensities.

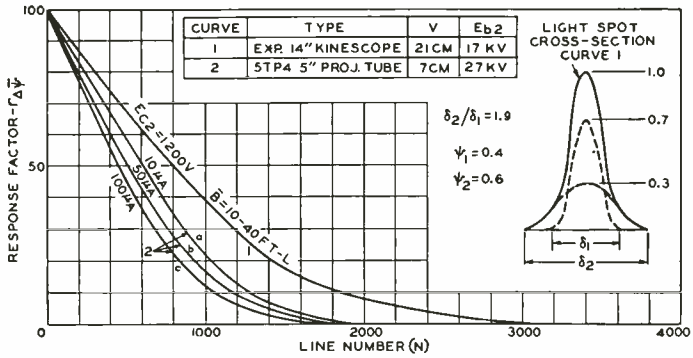


Fig. 42—Aperture flux response of two kinescopes.

5. Threshold Signal-to-Noise Ratios Required for High Quality

Threshold signal-to-noise ratios (R) at the control grid of the kinescope or at the transmitter (for a constant transfer factor to the kinescope grid) can be specified on the basis of a standard kinescope transfer characteristic. It is obvious from Equation (22)‡ that R maximum increases with kinescope brightness ($B = f(E)$) and also with kinescope resolution* because the noise filter factor m (See Part I) changes from m_{ek} to m_e for an ideally sharp kinescope which may be approached by pre-emphasis of high frequencies in the video channel. (In the horizontal direction).

Signal-to-noise ratios $R_{\max}$ for threshold visibility as a function of average field brightness level B_0 (foot-lamberts) have been computed

‡ Part I—page 37.

* This relationship applies also to increases of motion picture brightness and resolution.

Table IV—Threshold Signal-to-Noise Ratio $\hat{R}_{max}$ at the Kinescope Grid (Fig. 41) for a Channel Resolution ($\bar{N}_{co}$) of 400 Lines, a Peak Brightness ($\hat{B}$) of 32 Foot-Lamberts, and Constant Noise

Average brightness level B foot-lamberts	g_k	m_e 0.22	m_{ek} 0.1	m_e 0.45	m_{ek} 0.3	computed $m_{ek} = 0.41$	observed with m_{ek} threshold	good
0.2*	0.15	60	27	121	81	111	100	67
3.	0.85	62	28	126	84	115	115	67
5.	1.0	44	20	90	60	82	80	57
10.	1.1	24	11	50	33	45	63	40
		peaked channel		flat channel			flat channel	
		$\rho = 4$				$\rho = 2$		

* The factor *25* in Equation (22) (Part I—page 37) is reduced at $B_o = 0.2$ by the reduced transfer factor (g_r) to $25 \times 1.8/3.2 = 19.5$ (See Figure 3a) (Part I—page 16).

from Equation (22*) on the basis of the kinescope transfer characteristics given in Figure 41 for a peak brightness ($\hat{B}$) of 32 foot-lamberts ($\hat{E} = 40v$) and a 400-line channel (See Figure 11‡) and are given in Table IV. The values for the eye and an ideal kinescope response are listed in the columns m_e. The columns m_{ek} are computed for the eye and kinescope response given in Figure 13.**

Signal-to-noise ratios observed for $\rho = 2$ are in excellent agreement with computed values. If 5 foot-lamperts is selected as a representative average value for B_o (for $\hat{B} = 32\ Ft\text{-}L$) it may be stated that good quality at $\rho = 4$ in a 400-line channel *with constant noise level* requires the following values:

Signal-to-noise ratio in:	$\hat{R}_{max}$	$\|R\|_{max}$
flat channels	50-90	300-500
peaked channels	17-40	100-250

These values are to be measured at a point in the system where the video signal is related to the kinescope grid voltage by a constant factor.

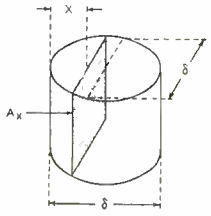
Signal-to-noise ratios at the television camera differ from the above

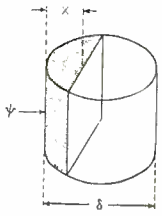
* Part I—page 37. ‡ Part I—pages 32-33. ** Part I—page 35.

values when the gradation scale of the system is changed in the electrical channel. The camera signal-to-noise ratio will be discussed in Part III of this paper. It is also pointed out that a kinescope screen material of much longer persistence or a storage-type reproducer will reduce the above values by increasing the number of simultaneously visible fluctuations or "flux samples". A reduction in the threshold signal-to-noise ratio would be indicated by a characteristic causing no flicker at frame repetition rate (30 cycles per second).

APPENDIX

Table I—Area and Volume of Aperture Sections





X/δ	*Area* Ax/A_{max} Cos²	Uniform	Transition curve (one half) Cos²	$\int_{x=o}^{x} \psi dx$ Uniform
0.	0.	0.	0.	0.
0.05	0.011	0.438	0.00054	0.019
0.1	0.0575	0.6	0.0043	0.0525
0.15	0.147	0.714	0.015	0.097
0.2	0.277	0.8	0.038	0.142
0.25	0.433	0.866	0.0775	0.195
0.3	0.60	0.915	0.132	0.252
0.35	0.756	0.954	0.205	0.312
0.4	0.885	0.98	0.294	0.375
0.45	0.97	0.995	0.393	0.405
0.5	1.0	1.0	0.50	0.50

(See page 285 for Table II)

Table III—Useful Aperture Flux Response Factors

Aperture	Density	$N_r/N\delta$	N_r/N for a response $r\Delta\overline{\psi} = 0.3$	0.5	0.7
Square	Uniform	2.	1.43	2.	3.34
Round	Uniform	2.47	1.48	2.08	3.5
Round	Cos²	4.35	1.74	2.46	4.15

Table II—Aperture Response Factors (Figures 15 and 16)

Aperture type	Square		Round			Round			Round		
Flux	Uniform		Uniform			$\cos^2$			e^{-x^2}		
$N/N\delta$	$r\triangle\hat{\psi}$	$r\triangle\bar{\psi}$	$N/N\delta$	$r\triangle\hat{\psi}$	$r\triangle\bar{\psi}$	$N/N\delta$	$r\triangle\hat{\psi}$	$r\triangle\bar{\psi}$	$N/N\delta$	$r\triangle\hat{\psi}$	$r\triangle\bar{\psi}$
0	1.	1.0	0	1.	1.0	0	1.	1.0	0	1.	1.0
0.5	1.	0.75	0.5	1.	0.785	0.5	1.	0.855	0.5	1.	0.86
0.75	1.	0.625	0.75	1.	0.68	0.75	1.	0.78	0.75	1.	0.78
1.0	1.	0.50	1.0	1.	0.57	1.	1.	0.715	1.	1.	0.72
1.25	0.6	0.375	1.25	0.80	0.47	1.25	0.98	0.65	1.25	0.98	0.655
1.5	0.34	0.25	1.5	0.56	0.36	1.5	0.91	0.58	1.5	0.92	0.59
1.75	0.14	0.125	1.75	0.37	0.255	1.75	0.81	0.51	2.	0.72	0.46
2.	0.	0.	2.	0.215	0.155	2.	0.70	0.44	2.5	0.52	0.33
2.5	—0.2	—0.135	2.25	0.09	0.065	2.5	0.48	0.31	3.	0.31	0.20
3.	—0.33*	—0.165	2.47	0.	0.	3.	0.28	0.18	3.5	0.18	0.115
3.5	—0.14	—0.105	2.75	—0.095	—0.006	3.5	0.13	0.09	4.	0.09	0.057
4.	0.	0.	3.	—0.16	—0.09	4.	0.04	0.025	4.5	0.05	0.032
4.5	+0.11		3.25	—0.18	—0.10	4.35	0.	0.	5.	0.03	0.019
5.	+0.2*		3.5	—0.16	—0.09	5.	—0.28		6.	0.01	0.0064
5.5	+0.09		4.	—0.08	—0.05	5.5	—0.25				
6.	0		4.47	0.	0.	6.3	0.				
			5.	+0.07							
			5.5	+0.07							
* Sharp maxima			6.5	0.							

Table IV—Aperture Response Factors of Identical Apertures in Cascade (1 to 4)

a) Square apertures with uniform density (Fig. 26)

$r\Delta\bar{\psi}$	$N/N\delta_1$ for Aperture Number Indicated 1	2	3	4	$r\Delta\hat{\psi}$	$N/N\delta_1$ for Aperture Number Indicated 1	2	3	4
1.0	1.	1.	1.	1.	1.0	1.	0.5	0.36	0.3
0.8	0.4	0.28	0.23	0.2	0.9	1.05	0.64	0.51	0.42
0.7	0.6	0.42	0.35	0.3	0.8	1.11	0.72	0.59	0.5
0.6	0.8	0.56	0.47	0.41	0.7	1.18	0.80	0.67	0.575
0.5	1.0	0.71	0.59	0.515	0.6	1.26	0.90	0.74	0.65
0.4	1.2	0.87	0.72	0.63	0.5	1.34	1.0	0.83	0.72
0.3	1.4	1.03	0.86	0.76	0.4	1.43	1.1	0.92	0.80
0.2	1.6	1.21	1.02	0.89	0.3	1.54	1.22	1.025	0.90
0.1	1.8	1.42	1.22	1.08	0.2	1.67	1.36	1.15	1.02
0.05	1.9	1.56	1.36	1.23	0.1	1.82	1.53	1.32	1.18
					0.05	1.91	1.66	1.44	1.30

b) Round apertures with $\cos^2$ density (Fig. 27)

$r\Delta\bar{\psi}$	$N/N\delta_1$ for Aperture Number Indicated 1	2	3	4	$r\Delta\hat{\psi}$	$N/N\delta_1$ for Aperture Number Indicated 1	2	3	4
1.	1	1	1	1	1.0	1	0.6	0.45	0.35
0.8	0.70	0.50	0.41	0.35	0.9	1.52	1.0	0.77	0.66
0.7	1.07	0.75	0.62	0.52	0.8	1.77	1.25	0.97	0.83
0.6	1.42	1.0	0.80	0.70	0.7	2.0	1.42	1.15	0.96
0.5	1.78	1.25	1.03	0.85	0.6	2.2	1.6	1.3	1.12
0.4	2.14	2.52	1.25	1.06	0.5	2.45	1.76	1.45	1.25
0.3	2.52	1.8	1.5	1.27	0.4	2.67	1.94	1.6	1.4
0.2	2.82	2.12	1.76	1.53	0.3	2.93	2.15	1.75	1.55
0.1	3.42	2.60	2.15	1.88	0.2	3.25	2.4	2.0	1.75
0.05	3.8	2.9	2.4	2.1	0.1	3.7	2.77	2.3	2.02
					0.05	3.96	3.1	2.6	2.25

THEORETICAL ANALYSIS OF VARIOUS SYSTEMS OF MULTIPLEX TRANSMISSION*†

By

Vernon D. Landon

Research Department, RCA Laboratories Division,
Princeton, N. J.

Summary

Specification of Systems:

A systematic method of classifying and specifying systems of multiplex transmission has been devised. A frequency division system is one using a separate sub-carrier for each channel with spaced sub-carrier frequencies. Sub-carriers may be amplitude modulated (AM), frequency modulated (FM), or single side band modulated (SS). A group of sub-carriers may be used as a modulating signal to amplitude modulate or phase modulate a higher frequency carrier. Thus FM-AM means a group of frequency modulated sub-carriers are used to amplitude modulate a carrier.

A time division system is one using sub-carriers each of which consists of a series of d-c pulses. Each pulse sub-carrier has the same frequency but the pulses are spaced in carrier angle so that the pulses are interleaved but do not overlap. Thus the summation of all the pulses is itself a series of d-c pulses. Each pulse sub-carrier may be amplitude modulated (PAM), phase modulated (PPM), width modulated (PWM) or numbers modulated (PNM). The sub-carriers may be used to amplitude modulate or frequency modulate a carrier. Thus PPM-AM means that a group of phase modulated pulse sub-carriers (with time division) are used to amplitude modulate a carrier.

The above systems may be called double modulation systems. A group of double modulation waves with spaced carrier frequencies (now called intermediate sub-carriers) may be used to modulate a still higher frequency carrier. This is called a triple modulation system. As an example SS-PM-FM means that a group of SS-PM waves with spaced carrier frequencies are used to frequency modulate a carrier. The symbol XX is used to mean that the type of modulation is unspecified.

Wide Band Gain of the Various Systems:

The wide band gain R_o is defined as the ratio of signal noise ratios (on random noise) comparing a single channel of a given multiplex system to a one channel AM system. A formula for the factor is found for each system in terms of the width of the audio band of each channel f_m, the

---

* Decimal Classification: R460 × R148.

† This paper, summarizing work done under contract with the U. S. Army Signal Corps, is being published in two parts in consecutive issues of *RCA REVIEW*, as follows:

Vol. IX, No. 2, June 1948—Complete Summary: Introduction (including Outline, Specification of Systems, Definitions and List of Symbols); Signal Noise Analysis.

Vol. IX, No. 3, September 1948—Minimum Signal Requirements; Bandwidth Economy; Impulse Noise, Cross Modulation and Interference Analyses; Propagation Considerations; Experimental Results.

number of channels n, and the radio-frequency band width B. The formulas indicate that the best systems (in the order of merit) are:

PAM-FM (Slow), SS-PM, PAM(±)-FM, PWM-FM, and PPM-AM. PAM(±) indicates bidirectional (positive and negative) pulses. This arrangement is used to avoid displacing the carrier with dissymmetrical frequency modulation. PAM-FM (Slow) indicates that the time division period is below the audible frequency range rather than above it. Systems PNM-XX and PAM-PCM-XX are not shown in the table as the signal noise ratio for these systems is infinite if the signal amplitude is above threshold. These systems are not necessarily the best, however, because of threshold effects discussed below. (PCM stands for pulse code modulation.)

The Threshold of the Various Systems:

The threshold of any system is the received average power required to realize the wide band gain. Formulas are given for the threshold levels of the various systems. To the extent that they have a high crest factor, the pulse systems have lower threshold levels than the other systems.

Minimum Received Signal Power Required for a Specified Signal Noise Ratio:

For most of the better systems the signal noise ratio increases with bandwidth as long as the signal strength is above threshold. However, the threshold level also increases with bandwidth. As the bandwidth is varied, minimum input power is required at the point where the specified signal noise ratio is obtained at threshold. This critical bandwidth has been determined for setups having 4, 12, 48, and 72 channels, with the required signal noise ratio having the values 46, 56, 66 and 66 decibels respectively. If this theoretically best bandwidth is over 10 megacycles, the figure 10 megacycles was used. The required signal power for the specified signal noise ratio at that bandwidth was determined and the results tabulated.

The results indicate that the eight best systems are PAM-FM(slow), PNM-FM, PAM-PCM-FM, PNM-AM, PAM-PCM-AM, SS-PM, PAM(±)-FM, PWM-FM, and PPM-AM in the order of merit for large values of n. For small values of n, PPM-AM is superior since it can utilize a wider band before running into threshold (because of its high crest factor). However, if the band-width of the PPM-AM system should be limited to the value used for the other listed systems, then the other systems would be better regardless of the number of channels.

Signal Noise Ratio on Impulse Noise:

On impulse noise of less than signal amplitude the wideband gain is the same as for random noise. On strong impulse noise it is assumed that a limiter is used on XX-AM systems and on the standard single channel AM system. Under these conditions the wideband gain on strong impulse noise is roughly the same as for random noise for systems XX-XX and PAM(±)-FM. For PXX-XX (not including PAM(±)-FM) however, the signal noise ratio on strong impulse noise is unity or worse. Also loss of synchronism may be experienced.

Interchannel Cross Modulation:

The actual signal to cross modulation ratio obtained in each system is more a function of the skill and technics used in making a setup than a function of the type of system.

In frequency division systems, cross modulation is caused by non-linearity

in amplifiers and by envelope non-linearity in modulators and de-modulators. In time division systems cross modulation is caused by a slight overlapping of adjacent pulses because of insufficient bandwidth (or by over modulation in PWM or PPM systems). A discussion of the relative feasibility of the various systems in this respect must depend on an engineering and experimental study. It may be found that for long relay chains, a triple modulation system will be required to avoid cumulative cross modulation.

Interchannel Cross Talk:

Cross talk, other than that caused by cross modulation, should never be a problem in time division systems if proper shielding is used. In frequency division systems such cross talk is governed by sub-carrier spacing and the design of the frequency selective filters. The design of such filters is well known.

Mutual Interference Between Multiplex Systems Operating on the Same Assigned Frequency:

In general the ratios of signal to interference ratios for the various systems correspond roughly to the ratios of signal to noise ratios on random noise. An exception is the fact that AM, XX-AM and PXX-AM systems are more susceptible to the serious fixed-pitch beat note interference. Also, threshold conditions are more apt to be exceeded if the interfering wave is pulsed and less likely to be exceeded if the desired signal is pulsed, especially if the duty factor is small.

Mutual Interference Between Multiplex Systems Operating on Adjacent Frequency Assignments:

The XX-AM systems produce little sideband splash in the adjacent channel. The PXX-FM systems (not including PAM(±)-FM) are probably the worst in this respect. The remainder of the systems are roughly equal in potential performance, but may vary greatly in practice according to the skill and technics used in making the setups.

Not considering sideband splash but simply the ability to produce interference because of the imperfect selectivity of the receiver on the adjacent multiplex channel, the order of merit of the interfering systems is about the same as above.

The susceptibility to sideband splash places the systems in the same order of merit as signal to random noise ratio considerations, except for cases where the interference may exceed the threshold level. In these cases, the PXX-AM systems have the advantage insofar as they have a high crest factor.

In susceptibility to adjacent channel interference due to imperfect selectivity (neglecting sideband splash) the XX-PM systems have a definite advantage.

Propagation, Susceptibility to Selective Fading:

On the basis of a small number of tests at 4000 megacycles, it appears that path length differences are much less than one meter. If this conclusion is correct, the problem of selective fading disappears almost completely with respect to atmospheric discontinuities.

If high points in mountainous country are chosen for terminal positions

difficulties may be had with ground reflections having larger path length differences.

INTRODUCTION

Outline

Introduction

Specification of Systems
Definitions and List of Symbols
Basic Types of Modulating Signal
Gain Due to Pre-Emphasis

Signal Noise Ratios

Frequency Division Systems
Time Division Systems
Triple Modulation Systems

Minimum Signal Requirements

Thresholds of the Various Systems
Required Received Signal Powers

Bandwidth Economy

Impulse Noise Analysis

Nature of Impulse Noise
Weak Impulse Noise
Strong Impulse Noise
Wave Form of Impulse Noise

Cross Modulation Analysis

Frequency Division Systems
Time Division Systems
Interchannel Crosstalk

Interference Analysis

Multiplex Systems on Same Frequency
Multiplex Systems on Adjacent Frequencies

Propagation Study

Experimental Results

Specification of Systems:

The possible permutations of frequency, amplitude, and pulse modulation with and without sub-carriers and with frequency division and time division leads to an unbelievably large number of possible methods of multiplexing. It seems almost impossible to cover all possible methods because there is no upper limit to the degree of complexity which may be involved. However, an attempt has been made to cover all of the simpler systems and all of the practical systems which have been suggested. Probably any system of greater complexity would not show improved signal noise ratio, but there is no definite assurance on this point since it is very difficult to assign limits to the results of future human ingenuity. An example may be found in PNM (pulse numbers modulation) which at first thought is not a particularly desirable method of modulation. However, it has been discovered that an almost perfect signal noise ratio may be obtained with this system by employing gating with the gate narrower than the signal pulse. This is explained in detail later.

In Table I, the various systems are classified in a logical manner. There are three practical methods of multiplexing through frequency division and four methods using time division. Thus there are seven basic types of wide band signals which may be called the modulating signal.

In addition to the seven basic types of modulation mentioned in the table, there are various hybrid types of modulation in which different channels are modulated in different ways on the same carrier (or pulse) thereby obtaining band width economy and (possibly) signal noise ratio improvement at the expense of a critical adjustment to obtain cross talk elimination. The various hybrids are as follows:

In frequency division a given sub-carrier may be amplitude modulated with one channel and frequency modulated with another, thus halving the number of sub-carriers required. Single side band may not be so combined.

In time division a given set of pulses may be modulated three ways simultaneously, PAM, PPM, and PWM, or any two of the three methods may be used together. Numbers modulation does not go well with the other three, although it might be possible to make it work with PAM by using for PAM a detector with a very fast response and a slow recovery to fill in the omission spaces. The same detector is required for PAM when PWM is present. In general, hybrids are not desirable because of cross talk difficulties.

The modulating signal may be applied to the carrier or inter-

mediate sub-carrier as amplitude modulation (AM) or frequency modulation (FM). The intermediate sub-carriers may be omitted, or may be applied as AM, or as FM, or as AM pulses or as pulses frequency modulating the carrier. As shown by the table this leads to 70 possible systems, not including hybrids. Many are quite impractical and are presented only for completeness. In each system involving FM

Table I—Multiplex Systems Classification

		Multiplex sig. applied as	No intermediate sub-carrier	Intermediate SC applied as AM	Intermediate SC applied as PM	Intermediate SC applied as AM pulses	Intermediate SC applied as FM pulses
	Column numbers →		1	2	3	4	5
Frequency Division	S.C. for each channel using AM	AM	AM-AM	AM-AM-AM	AM-AM-PM	AM-PAM-AM	AM-PAM-FM
		PM	AM-PM	AM-PM-AM	AM-PM-PM	AM-PPM-AM	AM-PPM-FM
	S.C. for each channel using FM	AM	FM-AM	FM-AM-AM	FM-AM-PM	FM-PAM-AM	FM-PAM-FM
		PM	FM-PM	FM-PM-AM	FM-PM-PM	FM-PPM-AM	FM-PPM-FM
	S.C. for each channel SSB and carrier supr.	AM	SS-AM	SS-AM-AM	SS-AM-PM	SS-PAM-AM	SS-PAM(±)FM
		PM	SS-PM	SS-PM-AM	SS-PM-PM	SS-PPM-AM	SS-PPM-FM
Time Division	Pulse amplitude modulation	AM	PAM-AM	PAM-AM-AM	PAM-AM-PM	PAM-PAM-AM	PAM-PAM-FM
		FM	PAM(±)FM	PAM-FM-AM	PAM-FM-PM	PAM-PPM-AM	PAM-PPM-FM
	Pulse phase modulation	AM	PPM-AM	PPM-AM-AM	PPM-AM-PM	PPM-PAM-AM	PPM-PAM-FM
		FM	PPM-FM	PPM-FM-AM	PPM-FM-PM	PPM-PPM-AM	PPM-PPM-FM
	Pulse numbers modulation	AM	PNM-AM	PNM-AM-AM	PNM-AM-PM	PNM-PAM-AM	PNM-PAM-FM
		FM	PNM-FM	PNM-FM-AM	PNM-FM-PM	PNM-PPM-AM	PNM-PPM-FM
	Pulse length modulation	AM	PWM-AM	PWM-AM-AM	PWM-AM-PM	PWM-PAM-AM	PWM-PAM-FM
		FM	PWM-FM	PWM-FM-AM	PWM-FM-PM	PWM-PPM-AM	PWM-PPM-FM

1st group of letters indicates type of multiplexing.

2nd group indicates how multiplex signal is applied to carrier or sub-carrier.

3rd group indicates how sub-carrier is applied to carrier.

AM = amplitude modulation

PM = phase modulation

FM = frequency modulation

PAM = pulses amplitude modulated

PPM = pulses phase modulated

SS = single sideband carrier suppressed

PNM = pulse numbers modulation

PWM = pulse length modulation

S.C. = sub-carrier

or PM, the appropriate alternative is chosen rather than putting in both types.

It is believed desirable to include the following definitions to avoid possible ambiguities.

Definitions:

1. **Wave.** A wave is a disturbance varying with time or space or both.

2. **Sine Wave.** A sine wave is a wave whose amplitude varies as the sine of a linear function of time, or space, or both.
   NOTE: It is often convenient to represent a sinusoidal wave by a rotating vector.

3. **Angle or phase of a sine wave.** The Angle or phase of a sine wave is the measure of the progression of the wave in time or space from a chosen instant or position.
   NOTE: In a sine wave represented by a rotating vector, the angle is the angle through which the vector has progressed.

4. **Carrier.** A carrier is a wave suitable for modulation by a modulating wave.
   NOTE: Examples of carriers are a sinusoidal wave and a recurring series of pulses.

5. **Pulse.** A pulse is a single disturbance characterized by the sudden rise and decay of a normally constant d-c potential.

6. **Unidirectional series of pulses.** A unidirectional series of pulses consists of a series of pulses in which every pulse rise is in the positive direction or every pulse rise is in the negative direction with respect to the potential between pulses.

7. **Bidirectional series of pulses.** A bidirectional series of pulses consists of a series of pulses in which some of the pulses rise in the positive direction and the remainder of the pulses rise in the negative direction with respect to the potential between pulses.

8. **Subcarrier.** A subcarrier is a carrier which is modulated and applied as modulation on a second carrier.

9. **Intermediate sub-carrier.** An intermediate sub-carrier is a carrier which is modulated by one or more sub-carriers and which is used as a modulating wave for another carrier.

10. **Carrier frequency.** The carrier frequency is the frequency of the carrier as defined in (4).

11. **Frequency of a pulse carrier.** The frequency of a pulse carrier is to be identified with the pulse repetition rate (PRF).
    NOTE: A carrier consisting of a series of pulses is not suitable for

space radio transmission. Hence pulse carriers are usually treated as sub-carriers and used to modulate higher frequency sinusoidal carriers. However pulse carriers may be transmitted over wire lines.

12. **Carrier angle.** The carrier angle is the angle of the carrier wave.
NOTE: The angle of a pulse carrier is to be identified with the angle of the fundamental a-c component of the pulses.

13. **Modulation.** (a) Modulation is the process by which some characteristic of a carrier is varied with time in accordance with a modulating wave.
(b) Modulation is the variation of some characteristic of a carrier with time.

14. **Modulating wave.** The modulating wave is the variation with time impressed upon some characteristic of the carrier.

15. **Modulated wave.** A modulated wave is a wave some characteristic of which varies with time in accordance with the instantaneous value of a modulating wave.

16. **Sideband.** (a) A sideband is the group of components of a modulated wave that lie on one side of the carrier frequency.
(b) A sideband is one of the components of a modulated wave at other than carrier frequency.

17. **Amplitude modulation (AM).** Amplitude modulation or AM is modulation in which the amplitude of a wave is the characteristic subject to variation.

18. **Angle modulation.** Angle modulation is modulation in which the angle relative to the carrier angle of a wave is the characteristic subject to variation.

19. **Phase modulation (PM).** Phase modulation or PM is angle modulation in which the angle of a wave is caused to depart from the carrier angle by an amount proportional to the instantaneous amplitude of the modulating wave.

20. **Instantaneous frequency.** Instantaneous frequency is the time rate of change of the angle.
NOTE: If the angle is measured in radians, the frequency in cycles is the time rate of change of the angle divided by 2π.

21. **Frequency modulation (FM).** Frequency modulation or FM is angle modulation in which the instantaneous frequency of a wave is caused to depart from the carrier frequency by an amount proportional to the instantaneous amplitude of the modulating wave.
NOTE: Phase and frequency modulation are particular forms of angle modulation.

NOTE: When the modulating wave is a single frequency sinusoid, it is impossible to distinguish between PM and FM by an examination of the modulated wave alone. In practice because of pre-emphasis the modulation is often a combination of PM and FM, but is called frequency modulation.

22. **Single sideband modulation (SS).** Single sideband modulation or SS is modulation resulting from the partial or complete elimination of the carrier and all components of one sideband from an amplitude modulated wave.

23. **Pulse amplitude modulation (PAM).** Pulse amplitude modulation or PAM is modulation in which a modulating wave is used to amplitude modulate a carrier wave consisting of a series of pulses.

24. **Pulse time modulation (PTM).** Pulse time modulation or PTM is modulation in which the modulating wave is used to modulate the time of occurrence of pulses in a carrier wave consisting of a series of d-c pulses.

25. **Pulse phase modulation (PPM).** Pulse phase modulation or PPM is modulation in which a modulating wave is used to phase modulate a carrier wave consisting of a series of d-c pulses.

26. **Pulse frequency modulation (PFM).** Pulse frequency modulation or PFM is modulation in which the modulating wave is used to frequency modulate a carrier wave consisting of a series of d-c pulses.

NOTE: Pulse phase modulation and pulse frequency modulation are particular forms of pulse time modulation.

27. **Pulse numbers modulation (PNM).** Pulse numbers modulation or PNM is modulation in which the pulse density per unit time of a pulse carrier is varied in accordance with a modulating wave by making systematic omissions without changing the phase or amplitude of the transmitted pulses.

NOTE: An example of PNM would consist of the omission of every other pulse, which would correspond to zero modulation. Reinserting some or all pulses corresponds to positive modulation. Omitting more than every other one corresponds to negative modulation.

28. **Pulse width modulation (PWM).** Pulse width modulation or PWM is modulation in which a modulating wave is used to modulate the time duration of the individual pulses of a pulse carrier. (Sometimes called pulse length modulation or pulse duration modulation.)

29. **Pulse time duration.** The pulse time duration is the area under the pulse energy wave form divided by the pulse energy amplitude.

NOTE: This usually approximately equals the time interval during which the pulse exceeds 50 per cent of peak energy amplitude.

30. **Duty factor or the duty cycle.** The duty factor or the duty cycle of a pulse carrier is the product of the pulse time duration and the pulse repetition frequency.

31. **Crest factor of a pulse carrier.** The crest factor of a pulse carrier is the ratio of the peak pulse amplitude to the root-mean-square amplitude.

32. **Limiter.** A limiter is a transducer, the output of which is approximately proportional to the instantaneous input amplitude up to a certain level but which has an almost constant output for input amplitudes exceeding this value.

    NOTE: A limiter may be used to remove amplitude modulation and transmit angle modulation.

    NOTE: In the I.R.E. Standards on Electrocoustics (1938) a transducer is defined as follows:

    > 1A45 *Transducer.* A transducer is a device by means of which energy can flow from one or more transmission systems to one or more transmission systems.

33. **Clipper.** A clipper is a transducer, the output of which is zero or a fixed value for instantaneous input amplitudes up to a certain value, but which has an output that is a function of input for amplitudes exceeding the critical value.

34. **Clipper limiter.** A clipper limiter is a transducer the output of which is a function of the instantaneous input amplitude for a range of values lying between two predetermined limits but is approximately constant at one level, for input values below the range and approximately constant at another level, for input values above the range.

    NOTE: This is sometimes called an amplitude gate.

35. **Transfer factor.** The transfer factor of a transducer is the ratio of the differential output current or voltage to the differential input current or voltage.

36. **Time gate.** A time gate is a transducer whose transfer factor as a function of time has the wave form of a series of pulses taking the value zero between pulses.

37. **Synchronous gate.** A synchronous gate is a gate whose pulses of transfer factor are synchronized with incoming signal pulses.

38. **Gating.** Gating is the process of selecting a portion of a wave.

39. **Compound modulation.** Compound modulation is modulation in which one or more signals are used to modulate their respective sub-carriers and these sub-carriers are used to modulate the carrier, or one of a second group of sub-carriers, etc.

40. **Double modulation.** Double modulation is compound modulation with only one group of sub-carriers.

41. **Triple modulation.** Triple modulation is compound modulation in which a number of double modulation waves are used as a modulating wave to modulate a carrier wave.

42. **Multiplex modulation.** Multiplex modulation is modulation in which more than one signal is transmitted using a common carrier.

43. **Frequency division multiplex.** Frequency division multiplex is the process or device in which each signal modulates a separate sub-carrier and the sub-carriers are spaced in frequency in a manner to avoid overlapping of the sub-carrier sidebands.

44. **Time division multiplex.** Time division multiplex is the process or device in which each signal modulates a separate pulse sub-carrier, the pulses sub-carriers being spaced in carrier angle so that the summation of all the sub-carriers is a series of pulses.

45. **Frequency division multiplex modulation.** Frequency division multiplex modulation is the result of modulation of a carrier by the output of a frequency division multiplex device.

NOTE: A double modulation system with frequency division may be designated by the symbol XX-XX where it is desired that the types of modulation should not be specified. To refer to specific systems, the specific symbols should be used. For example SS-PM means that one or more signals are used to produce single sideband modulation of their respective sub-carriers which are spaced in frequency. The single sideband waves are used to phase modulate a carrier.

46. **Time division multiplex modulation.** Time division multiplex modulation is the result of modulation of a carrier by the output of a time division multiplex device.

NOTE: A double modulation system with time division may be designated by the symbol PXX-XX where it is desired that the types of modulation should not be specified. Again, to refer to specific systems, the specific symbols should be used. For example, PPM-AM means a system in which one or more signals are used to phase modulate their respective pulse sub-carriers (with time division) and the pulses are used to amplitude modulate a carrier.

Similarly, triple modulation systems may be designated by the symbols XX-XX-XX or XX-PXX-XX, etc. An example is SS-PM-PM which designates a system in which a number of SS-PM waves are used to phase modulate a carrier.

To be even more specific, a numerical subscript may be used to indicate the number of sub-carriers. Thus FM_m-AM_n-PM means a system in which a number m of signals are used to frequency modulate their respective sub-carriers which are spaced in frequency and which are used to amplitude modulate a higher frequency sub-carrier. A group of n similarly modulated higher frequency sub-carriers are used to phase modulate a carrier. The total number of channels is then mn.

47. **Hybrid multiplex modulation.** Hybrid multiplex modulation is modulation in which the type of modulation used is different for different signal channels.

NOTE: Usually a given sub-carrier is modulated two different ways by different signals.

48. **Modulator.** A modulator is a device to effect the process of modulation.

49. **Frequency-shift modulation.** Frequency-shift modulation is a form of frequency modulation in which the modulating wave abruptly shifts the carrier frequency between predetermined limits.

50. **Amplitude modulator.** An amplitude modulator is a device to effect the process of amplitude modulation.

51. **Phase modulator.** A phase modulator is a device to effect the process of phase modulation.

52. **Frequency modulator.** A frequency modulator is a device to effect the process of frequency modulation.

53. **Frequency deviation.** Frequency deviation (in frequency modulation) is the difference between the instantaneous frequency of the modulated wave and the carrier frequency.

54. **Frequency swing.** Frequency swing (in frequency modulation) is the difference between the maximum and minimum values of the instantaneous frequency.

55. **Phase deviation.** Phase deviation is the difference between the instantaneous angle of the modulated wave and the angle of the carrier.

56. **Deviation ratio.** Deviation ratio (in a frequency-modulation system) is the ratio of the maximum frequency deviation to the maximum modulating frequency of the system.

57. **Modulation index.** The modulation index (for a sinusoidal modulating wave) is the ratio of the maximum frequency deviation to the frequency of the modulating wave.

58. **Detection.** Detection is the process by which a wave corresponding to the modulating wave is obtained in response to a modulated wave.

59. **Frequency-modulation detector.** A frequency-modulation detector is a detector whose voltage or current output is a function of the frequency deviation of a modulated wave.

60. **Balanced frequency-modulation detector.** A balanced frequency modulation detector is a frequency modulation detector so arranged that the response to amplitude modulation is a differential output which is zero in the absence of frequency modulation.

61. **Discriminator.** A discriminator is a device in which amplitude variations are derived in response to frequency variations.

62. **Slope filter.** A slope filter is a filter having a rising or falling response over a given frequency range.

    NOTE: A slope filter may be used as a frequency discriminator.

63. **Phase-modulation detector.** A phase-modulation detector is a detector whose output is a function of the phase deviation of a modulated wave.

64. **Differentiating network.** A differentiating network is a network whose output is the time derivative of its input wave form.

    NOTE: Such a network preceding a frequency modulator makes the combination a phase modulator; or, following a phase detector makes the combination a frequency detector. Ideally, its ratio of output amplitude to input amplitude is proportional to frequency and it has a phase lead of 90 degrees.

65. **Integrating network.** An integrating network is a network whose output waveform is the time integral of its input waveform.

    NOTE: Such a network preceding a phase modulator makes the combination a frequency modulator; or, following a frequency detector, makes the combination a phase detector. Ideally, its ratio of output amplitude to input amplitude is inversely proportional to frequency, and it has a phase lag of 90 degrees.

66. **Pre-emphasis network.** A pre-emphasis network is a network inserted in a system in order to emphasize one range of frequencies with respect to another.

    NOTE: The usual purpose is to emphasize components corresponding to the higher frequencies in the signal input.

67. **De-emphasis network.** A de-emphasis network is a network inserted in a system in order to restore the pre-emphasized frequency spectrum to its original form.

68. **Noise.** Noise is an unwanted disturbance whose energy is distributed among a large number of frequency components within the useful frequency band.

69. **Impulse noise.** Impulse noise is noise characterized by transient disturbances separated in time by quiescent intervals.

70. **Random (or fluctuation) noise.** Random (or fluctuation) noise is noise characterized by a large number of overlapping transient disturbances occurring at random.

71. **Signal-to-noise ratio.** The signal-to-noise ratio is the ratio of the amplitude of the signal to that of the noise.

NOTE: This ratio is a function of the bandwidth of the transmission system.

NOTE: This ratio is usually in terms of peak values in the case of impulse noise and in the terms of root-mean-square values in the case of random noise.

72. **Signal-plus-noise-to-noise ratio.** The signal-plus-noise-to-noise ratio is the ratio of the amplitude of the signal plus the noise to that of the noise.

73. **Carrier-to-noise ratio.** The carrier-to-noise ratio is the ratio of the amplitude of the carrier to that of the noise after selection and before any non-linear process such as amplitude limiting and detection.

74. **Improvement threshold.** The improvement threshold (in a frequency-modulated receiver) is the condition of unity for the ratio of peak carrier voltage to peak noise voltage after selection and before any non-linear process such as amplitude limiting and detection.

75. **Intelligence bandwidth.** The intelligence bandwidth of a multiplex system is the sum of the audio (or video) frequency bandwidths of the individual channels.

76. **Wide-band ratio.** The wide-band ratio of a multiplex system is the ratio of the frequency bandwidth utilized to the intelligence bandwidth.

77. **Wide-band improvement.** The wide-band improvement of a multiplex system is the ratio of signal noise ratios, comparing a single channel of the system under discussion to a single channel AM system having the same average radiated power.

List of Symbols:

Symbols for complete transmission systems:

AM—Audio from a single source is used to amplitude modulate a carrier wave.

FM—Audio from a single source is used to frequency modulate a carrier wave.

SS—Same as AM, except that the carrier and one sideband are suppressed.

XX—Any of above simple types of transmission.

AM-AM—A number of complete AM waves with spaced carrier frequencies are used to amplitude modulate a carrier wave.

XX-XX—A number of complete XX waves are used to XX a carrier wave.

PAM—Audio from a single source is used to amplitude modulate a carrier wave consisting of a series of d-c pulses.

PAM-AM—A number of PAM signals (with time division) are used to amplitude modulate a c-w (continuous wave) carrier.

Further possible extensions:

AM-PAM-FM means that a number of AM waves with spaced carrier frequencies are used to amplitude modulate a carrier wave consisting of d-c pulses. A number of these systems with time division are used to frequency modulate a carrier wave. See Table I for further examples of systems.

Other symbols:

K_1 = a constant associated with the first modulation.

$= \sqrt{3}$ for FM; = unity for AM, SS or PM.

M_1 = modulation index of the signal on each channel.

$= \text{deviation ratio } \frac{f_{d1}}{f_m}$ for FM; = unity for AM and SS.

K_2 = a constant associated with the second modulation.

$= \sqrt{2}$ 2 for AM-PM, FM-AM, and FM-PM; unity for SS-AM and SS-PM.

M_2 = the modulation index of each sub-carrier or individual signal applied to the main carrier; $= f_{dm}/f_{scm}$ (for AM-PM or FM-PM).

f_{dm} = the frequency deviation of the main carrier caused by the unmodulated m^{th} sub-carrier.

f_{scm} = the frequency of the m^{th} sub-carrier.

d_m = the peak deviation caused by the modulated m^{th} sub-carrier. $= 2f_{dm}$.

$d_1 = 2f_{d1}$; $= d_m$ for the first channel.

m = the identification number of a given channel.

n = the total number of channels.

n_o = any given value of n.

f_m = the highest modulation frequency for each channel.
= 3500 cycles for all numerical work.

f_c = the space between sub-carriers for SS-AM; = ½ space between sub-carriers for AM-PM.
= 5000 cycles for all numerical work.

B = the effective noise bandwidth of the receiver through the radio- and intermediate-frequency circuits.

d_s = the peak deviation obtained by adding $d_1 + d_2 + - - - + d_m - - - + d_n$ (for linear summation).

d_s(rms) = the root-mean-square deviation of the main carrier for all channels.

d_s(peak) = $4d_s$(rms).

f_{d1} = the peak frequency deviation of a given sub-carrier caused by the modulating signal in FM-AM or FM-PM systems.

$y = f_m/f_c = 0.7$ in numerical work.

M_2(rms) = the root-mean-square deviation of the main carrier caused by one channel, divided by the mean signal frequency of that channel (in the SS-PM system).

M_2(peak) = M_2.

f_d(rms) = the root-mean-square deviation of the main carrier in the SS-PM system.

f_d(peak) = the peak deviation of the main carrier; = $4f_d$(rms).

D = the duty factor of one channel of a pulse system.

nD = the duty factor of the n channel pulse system.

r_1 = the minimum value of $1/nD$ if cross modulation is to be avoided.
= 2 (assumed value for PAM(±)-FM system).

ϕ = the angular deviation or modulation index due to the modulation of the pulse frequency by the signal in the PPM-AM system.

f_p = the pulse frequency for a single channel.
= $3f_m$ (for PAM, PPM, PWM).

a = time duration of a given pulse.
= $1/B$ in cases using the minimum pulse length for a given bandwidth.

S = the number of harmonics of f_p passed; = $B/2f_p$.

e_n = the instantaneous value of a noise component.

$\triangle$ = the infinitesimal amplitude of a single noise component.

j = any integer from 1 to S.

$w_o = 2\pi$ times the carrier frequency.

$w_p = 2\pi f_p$.

$w_a = 2\pi$ times the audio frequency difference between the frequency of the noise component and the frequency of the j^{th} harmonic sideband of f_p.

θ = a random phase angle.

e_s = the instantaneous voltage of a PPM-AM system.

E = the amplitude of a sine wave having the same average power as e_s.

E_s = the envelope of e_s.

E_s^1 = the slope of E_s.

p = the number of amplitude gradations used in a PAM-PCM-AM system.

PCM = pulse code modulation.

FCM = frequency code modulation.

q = the number of pulse positions in the code for a certain amplitude in PAM-PCM-AM system.

K_3 = the ratio by which the average pulse time duration is larger than the minimum pulse time duration (in PWM systems) ; $= aB$.

K_4 = the ratio of the pulse rate to the highest transmitted signal frequency for pulses assigned to one channel.

APM = a hybrid of AM and PM on the same sub-carriers.

$PAPWM$ = a hybrid of PAM, PPM and PWM on the same pulses.

$PPWM$ = a hybrid of PPM and PWM on the same pulses.

$r = B/nf_m$.

G_a = gain due to pre-emphasis in an AM system; = 2.8 decibels or a voltage ratio of 1.38.

G_f = gain due to pre-emphasis in an FM system; = 5.4 decibels or a voltage ratio of 1.86.

P_r = the received power required to produce a specified signal noise ratio.

P_t = the received signal power at threshold for any system.

$= KTBF_nC_f$.

P_{10} = the value of P_r for 10-megacycle bandwidth.

P_B = the value of P_r for the bandwidth B.

K = Boltzmann's constant; $= 1.37 \times 10^{-23}$ joules per degree Kelvin.

T = the absolute temperature; = 293 degrees Centigrade approximately.

n_1 = the number of sub-carriers used to modulate an intermediate sub-carrier.

n_2 = the number of intermediate sub-carriers used to modulate a carrier.

F_n = the noise factor of the receiver; = 12 decibels or 16 to 1 power ratio in numerical work.

C_f = the ratio of signal power to noise power at threshold for a particular system.

R_r = the specified minimum signal-noise ratio.

R_t = the signal-noise ratio at threshold; $= R_o R_a$ for greatest range.

$R_o = \dfrac{S/N\ (XX\text{-}XX)}{S/N\ (AM)}$ = The ratio of signal-to-noise voltage ratios, comparing one channel of the XX-XX system to a single channel AM system.

R_a = $S/N\ (AM)$ at the signal strength required for threshold in the $(XX\text{-}XX)$ system; $= \left(\dfrac{B}{2f_m}\right)^{1/2} \left(\dfrac{C_f\ (XX\text{-}XX)}{C_f\ (AM)}\right)^{1/2}$

C = peak carrier in intermediate-frequency circuit.

e = peak noise voltage in a one cycle band width in intermediate-frequency circuit.

B_o = minimum value permissible for B.

F_v = equivalent noise frequency band of circuits for pulse signals (video frequency band).

F_{vo} = minimum value of F_v.

t_g = gate duration time.

e_v = peak voltage in a one cycle band in the circuit following the f-m discriminator.

b = any integer.

$k = F_v / n f_p$.

f_x = the difference in frequency of a noise component and the carrier.

e_{an} = the peak amplitude of a noise component in the audio frequency band.

E_n^1 = the peak voltage at a channel output due to the noise in the frequency band O to F_t ahead of the gate.

E_{noise} = the peak noise voltage in the audio frequency band f_m at a channel output due to the gating of a triangular noise spectrum which extends from O to F_p cycles per second.

B_t = the value of B for which $P_r = P_t$.

e = the instantaneous voltage of the response of an amplifier to unit impulse.

A = the amplification of the amplifier at resonance.

$\alpha = R/2L$ for each tuned circuit of the amplifier.

m_1 = the number of tuned circuits of an amplifier.

R_I = the ratio of the desired to the undesired carrier amplitude on a root-mean-square basis.

K_z = the ratio of crest factors for the interfering signal and for random noise.

Detailed Discussion of the Seven Basic Types of Modulating Signal:

Using frequency division it is essential that a sub-carrier frequency be assigned to each channel. For amplitude modulation these sub-carriers must be spaced in frequency a little more than twice the highest audio frequency utilized in each channel. For frequency modulation the sub-carriers should be spaced $B/4n$ cycles apart starting $B/4n$ from zero frequency, where B is the available bandwidth and n is the desired number of channels. For single sideband modulation the sub-carriers should be spaced apart only slightly more than the highest audio frequency utilized in each channel.

Using time division the pulses are divided into groups, each group having a number of consecutive pulses equal to n, the number of channels. The first pulse in each group belongs to the first channel, the second pulse in each group, to the second channel and so on for n channels. For PAM the pulses of channel 1 are amplitude modulated with the signal of channel 1, etc. For PPM the pulses are phase modulated with the signal. For pulse width modulation PWM, the width of the pulses is made a linear function of the modulating voltage. For pulse numbers modulation PNM, the fundamental pulse rate and phase are held constant but the pulse density per unit time is varied by making systematic omissions. Omitting every other pulse corresponds to zero modulating voltage. No omissions corresponds to maximum positive

voltage and no pulses corresponds to maximum negative voltage. Thus omitting every fourth pulse corresponds to 50 per cent positive modulation.

For PAM the pulse frequency for each channel must be greater than twice the highest audio frequency. Usually 2½ to 1 or 3 to 1 is used. The pulses must be sufficiently narrow so that one pulse is substantially zero before the next has built up to half amplitude (to prevent cross-talk) in all time division systems.

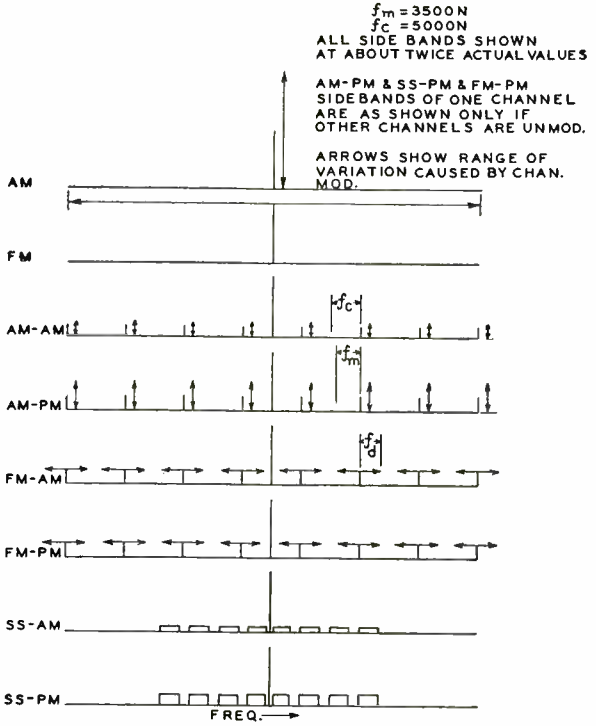


Fig. 1—Sideband spectra of frequency division systems.

For PPM the pulse frequency for each channel must be enough greater than the highest audio frequency so that none of the sidebands produced by the modulation extend into the audio frequency spectrum with appreciable amplitude.

For pulse width modulation the pulse rate requirement is the same as for PAM.

For pulse numbers modulation the pulse rate must be larger than for any of the other systems in order to reduce distortion. It is esti-

mated that the required pulse rate is about 20 times the highest audio frequency.

The sideband spectrum of several systems is shown in Figures 1 and 2. Figure 2 also shows the modulating wave form for PAM, PPM and PAM(±). PAM refers to uni-directional pulses and is most appropriate for PAM-AM. PAM(±) refers to bi-directional pulses and is most appropriate for PAM(±)-FM, to avoid displacing the carrier during modulation. The sideband spectrum of PWM and PNM are too involved to include in the figures.

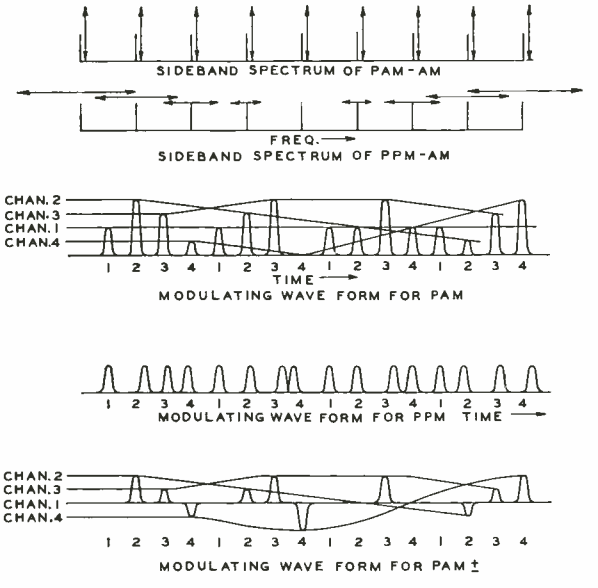


Fig. 2—Sideband spectra and wave forms.

Gain Due to Pre-emphasis:

A certain amount of gain in signal noise ratio for voice may be obtained for every system by the use of pre-emphasis and de-emphasis networks. Single sideband systems are assumed to have the same figure as AM systems. The gain is also a function of the audio bandwidth, being greater, the greater the width of the band. The gain is obtained only on the first application of the signal to a carrier or sub-carrier. Pre-emphasis and the approximate value of the gain is discussed later.

To avoid carrying along a constant factor in all systems, the gain due to pre-emphasis is omitted from all the following formulas.

Signal Noise Ratios for Various Systems Compared to AM

Frequency Division Systems

For any frequency division system (*XX-XX*) the formula may be written,

$$R_o = \frac{\text{S/N (XX-XX)}}{\text{S/N (AM)}} = K_1 M_1 M_2 K_2 \quad \ldots\ldots\ldots\ldots \quad (1)$$

where K_1 is a constant associated with the first modulation. K_1 is always unity except where the first modulation is FM, when $K_1 = \sqrt{3}$.

M_1 is the modulation index of the signal on each channel. (Deviation ratio if FM) It is unity for AM and SS, but is f_{d1}/f_m for FM.

M_2 is the modulation index of the sub-carrier or individual signal applied to the main carrier. It will be determined in each instance.

K_2 is a constant associated with the second modulation. It is required in order to take care of the following: The amplitude of a sub-carrier sideband is $E \cdot M_2/2$. The signal-noise ratio would be the same as that of a single sideband, except for the fact that the signal from the two sub-carrier sidebands adds linearly while the noise adds up as the square root of the sum of the squares. This improves the signal ratio by $\sqrt{2}$. Thus for (AM-AM), $K_2 = \sqrt{2}/2$ as it is for (AM-PM), (FM-AM), and (FM-PM). For SS-AM and SS-PM $K_2 = 1$.

AM-AM

For this system,

$$K_1 = 1; \quad M_1 = 1; \quad M_2 = 1/2n; \quad K_2 = \sqrt{2}/2$$

where n = the number of channels.

The value of M_2 cannot be $1/n$ because there must be room for each sub-carrier to be 100 per cent modulated. Then,

$$\frac{\text{S/N (AM-AM)}}{\text{S/N (AM)}} = \frac{\sqrt{2}}{4n} \text{ (for linear summation)} \quad (2)$$

(The 2 in $1/2n$ allows for the fact thata each sub-carrier may go to double amplitude when modulated. In $K_2 = \sqrt{2}/2$, the denominator corresponds to the 2 in the expression for the sideband amplitude $E_b = EM_2/2$. The $\sqrt{2}$ is the gain in signal noise ratio when the two sidebands are added.)

If it is assumed that the modulations add directly then $M_2 = 1/2n$ as above. This value should be used if n is small.

However, if n is large, the complex wave resulting from all the

sub-carriers may be compared to a noise wave in which the crest factor[1] is known to be about 4 and independent of bandwidth. The crest factor is defined as the ratio of the amplitude of the highest peaks to the r-m-s (root-mean-square) value. (Higher peaks than four to one above r-m-s do occur but they are so infrequent as to be negligible.) The r-m-s value of a sine wave is $\sqrt{2}/2$ times the peak amplitude. Hence the highest peaks of the complex wave from all the sub-carriers is the mean or unmodulated amplitude of one of them times $\sqrt{2}/2 \times \sqrt{n} \times 4$. Then $M_2 = \dfrac{\sqrt{2}}{4n^{\frac{1}{2}}}$ for r-m-s summation. Using the new value of M_2 in Equation (1) gives,

$$\frac{\text{S/N (AM-AM)}}{\text{S/N (AM)}} = \frac{\sqrt{2}}{2} \times \frac{\sqrt{2}}{4\sqrt{n}} = \frac{1}{4n^{\frac{1}{2}}} \text{ for r-m-s summation} \quad (3)$$

(Note, the modulation of the sub-carriers is assumed to have no effect on the second calculation.)

AM-PM

The formula is the same as for AM-AM except that the modulation index f_{dm}/f_{scm} is substituted for M_2[2].

$$\frac{\text{S/N (AM-PM)}}{\text{S/N (AM)}} = \frac{\sqrt{2}}{2} \frac{f_{dm}}{f_{scm}} \quad (4)$$

where f_{dm} is the mean (no signal) deviation of the main carrier produced by the m^{th} sub-carrier and f_{scm} is the frequency of the sub-carrier.

It can be seen that here, the signal noise ratio is a function of the available bandwidth. However, the fraction of the available bandwidth, that should be assigned to each channel is not a constant if each channel is to have the same signal noise ratio. Instead, the modulation index f_{dm}/f_{scm} should be kept constant for each channel so that the frequency deviation available for a given channel, is proportional to the sub-carrier frequency of that channel. If it is assumed that the peak

[1] For a discussion of crest factor in noise, see V. D. Landon, "The Distribution of Amplitude with Time in Fluctuation Noise", *Proc. I.R.E.*, Vol. 29, pp. 50-55, Feb. 1941. Also discussion on above paper by K. A. Norton and V. D. Landon, *Proc. I.R.E.*, Vol. 30, No. 9, pp. 425-429, Sept. 1942. Also S. O. Rice, "Mathematical Analysis of Random Noise", *Bell Sys. Tech. Jour.*, Vol. 23, pp. 282-332, 1944, and Vol. 24, pp. 46-156, 1945.

[2] For a detailed discussion of noise in FM systems see Murray G. Crosby, "Frequency Modulation Noise Characteristics", *Proc. I.R.E.*, Vol. 5, No. 4, April 1937.

frequency deviations of the various channels add directly and their sum should equal $B/2$ where B is the total available bandwidth, then the frequency deviation allotted for each channel can be solved for in terms of B and n the total number of channels.

Let the peak deviation of the first channel be d_1. Let f_c be ½ the effective channel width of any channel, that is f_m plus a small guard band. Then the sub-carrier frequencies are:

$$f_c, 3f_c, 5f_c \text{ - - - - } (2m-1)f_c \text{ - - - - } (2n-1)f_c$$

Since the modulation index should be the same for each channel the frequency deviation should be proportional to the sub-carrier frequency so that for the m^{th} channel $d_m = (2m-1)\, d_1$. Then the total deviation added linearly is,

$$d_s = d_1(1+3+5 \text{ - - - } (2n-1)) = d_1(2n-1+1)\frac{n}{2} = d_1 n^2 = B/2$$

$$d_1 = B/2n^2.$$

However, the mean or no signal deviation of the m^{th} sub-carrier is one-half the peak deviation and is equal to:

$$f_{dm} = \frac{1}{2} d_m = \frac{B(2m-1)}{4n^2}.$$

The modulation index is,

$$M_2 = \frac{f_{dm}}{f_{scm}} = \frac{B(2m-1)}{4n^2} \div [f_c\,(2m-1)\,] = \frac{B}{4n^2 f_c}. \qquad (5)$$

Then $$\frac{\text{S/N (AM-PM)}}{\text{S/N (AM)}} = \frac{\sqrt{2}}{8}\frac{B}{n^2 f_c} \quad \text{(linear summation)} \qquad (6)$$

On the other hand for large values of n it may be assumed that the frequency deviations add as the square root of the sum of the squares. For telephone work, the r-m-s value of an amplitude modulated carrier is changed only a negligible amount by the modulation. The r-m-s value of the deviation of the m^{th} channel is then $\frac{\sqrt{2}}{2}\frac{d_m}{2} = \frac{\sqrt{2}\, d_m}{4}$.

The r-m-s summation of the deviation for all channels is,

$$d_s(\text{r-m-s}) = \frac{\sqrt{2}}{4}\sqrt{d_1^2 + d_2^2 + d_3^2 \text{ - - - } d_n^2}$$

$$= \frac{\sqrt{2}}{4} d_1 \sqrt{1^2 + 3^2 + 5^2 \text{ - - - } (2n-1)^2}$$

$$= \frac{\sqrt{2}}{4} d_1 \sqrt{\frac{4n^3 - n}{3}}$$

(This equivalence is exact. The proof consists in, first showing that it is true for one value by a numerical substitution, and second by showing that if it is true for $n = n_m$, it is true for $n = n_o + 1$.)

$$= \frac{1}{2} \sqrt{\frac{2}{3}} d_1 n^{3/2}$$

(Dropping the n as negligible.) (7)

If n is large the crest factor of the deviation will be four, just as for noise.

$$d_s(\text{peak}) = 4d_s(\text{r-m-s}) = 2\sqrt{\frac{2}{3}} d_1 n^{3/2} = \frac{B}{2}$$

$$d_1 = \frac{1}{4}\sqrt{\frac{3}{2}}\frac{B}{n^{3/2}}$$

$$f_{dm} = \frac{d_m}{2} = \frac{1}{8}\sqrt{\frac{3}{2}}\frac{B\,(2m-1)}{n^{3/2}} \tag{8}$$

$$\frac{\text{S/N (AM-PM)}}{\text{S/N (AM)}} = \frac{\sqrt{2}}{2}\frac{f_{dm}}{f_{sum}} = \frac{\sqrt{2}}{2}\frac{1}{8}\sqrt{\frac{3}{2}}\frac{B}{n^{3/2}}\frac{(2m-1)}{f_c(2m-1)}$$

$$= \frac{\sqrt{3}}{16}\frac{B}{n^{3/2} f_c} \text{ (for r-m-s summation).} \tag{9}$$

FM-AM

For this case the sub-carriers should have the following values when unmodulated:

$$\frac{B}{4n}, \frac{3B}{4n}, \frac{5B}{4n} \cdots \frac{B}{4n}(2m-1) \cdots \frac{B}{4n}(2n-1)$$

The signal noise ratio on each channel, compared to AM is,

$$\frac{\text{S/N (FM-AM)}}{\text{S/N (AM)}} = K_1 M_1 K_2 M_2 = \sqrt{3}\frac{f_{d1}}{f_m} \times \frac{\sqrt{2}}{2} M_2 \tag{10}$$

The value of f_{d1} is $B/4n$, as any larger deviation would cause the

channels to overlap. Due to the fact that all sub-carriers have equal amplitude, it is not considered necessary to provide a guard band. Each channel selectivity curve should be down to about 20 per cent at channel band limits, using a limiter to flatten the response.

Side bands of appreciable amplitude will be produced by each channel in the frequency bands allocated to the adjacent channels but the interference produced will not be appreciable because of the wide-band gain of each frequency modulated channel, if M_1 is large.

The factor M_2 corresponds to the percentage by which each sub-carrier is permitted to amplitude modulate the main carrier. Assuming linear addition of these percentages, then,

$M_2 = 1/n$ (not $1/2n$ because the sub-carriers are not amplitude modulated). Then

$$\frac{\text{S/N (FM-AM)}}{\text{S/N (AM)}} = \sqrt{3}\,\frac{B}{4n f_m}\,\frac{\sqrt{2}}{2n} = \frac{\sqrt{6}}{8}\,\frac{B}{n^2 f_m} \text{ (linear summation)} \quad (11)$$

For larger values of n it may be assumed that the percentages of modulation add up in r-m-s fashion. In this case $M_2 = \sqrt{2}/4\sqrt{n}$. This may be derived as follows:

M_2 = the peak value of the modulation index due to a modulating sine wave sub-carrier.

$\frac{M_2}{\sqrt{2}}$ = the r-m-s value of the modulation index.

$\frac{n^{\frac{1}{2}} M_2}{\sqrt{2}}$ = the r-m-s value of the complex modulation index.

$\frac{4n^{\frac{1}{2}} M_2}{\sqrt{2}}$ = the peak modulation index of the complex modulating wave providing n is large enough so that the complex wave has the same crest factor as noise.

But $\frac{4n^{\frac{1}{2}} M_2}{\sqrt{2}}$ = unity or 100 per cent modulation; then, $M_2 = \frac{\sqrt{2}}{4\,n^{\frac{1}{2}}}$

Then $$\frac{\text{S/N (FM-AM)}}{\text{S/N (AM)}} = \frac{\sqrt{3}}{4n}\,\frac{B}{f_m\,4\sqrt{n}}$$

$$= \frac{\sqrt{3}}{16}\,\frac{B}{f_m\,n^{3/2}} \text{ (r-m-s summation)} \quad (12)$$

FM-PM

$$\frac{\text{S/N (FM-PM)}}{\text{S/N (AM)}} = K_1 M_1 \times K_2 M_2 = \sqrt{3}\, M_1 \times \frac{1}{\sqrt{2}} \frac{f_{dm}}{f_{scm}} \qquad (13)$$

Let M_1 take any value between unity and $\dfrac{B}{4n f_m}$. Then the m^{th} sub-carrier frequency should be, $f_{scm} = M_1 f_m (2m-1)$.

For linear summation the peak frequency deviation of the m^{th} sub-carrier is,

$d_m = \dfrac{B}{2n^2} (2m-1)$ just as for AM-PM.

However, for FM-PM $f_{dm} = d_m$ (not $d_m/2$ as in AM-PM where allowance had to be made for amplitude modulation)

$$\frac{f_{dm}}{f_{scm}} = \frac{\dfrac{B}{2n^2}(2m-1)}{M_1 f_m (2m-1)} = \frac{B}{2n^2 M_1 f_m} \qquad (14)$$

$$\frac{\text{S/N (FM-PM)}}{\text{S/N (AM)}} = \frac{\sqrt{3}}{\sqrt{2}} \text{M}_1 \frac{B}{2n^2 M_1 f_m} =$$

$$\sqrt{\frac{3}{2}} \frac{B}{2n^2 f_m} \text{ (for linear summation)} \qquad (15)$$

For large values of n it may be assumed that the frequency deviations caused by the sub-carriers add as the square root of the sum of the squares. If d_m is the peak deviation of the m^{th} channel, the r-m-s deviation is $d_m/\sqrt{2}$. The r-m-s summation of the deviations for all the channels is,

$$d_s(\text{r-m-s}) = \frac{1}{\sqrt{2}} \sqrt{d_1^2 + d_2^2 - - - d_n^2} = \frac{d_1}{\sqrt{2}} \sqrt{1^2 + 3^2 + 5^2 - - - (2n-1)^2}$$

$$= \frac{d_1}{\sqrt{2}} \sqrt{\frac{4n^3 - n}{3}} = \sqrt{\frac{2}{3}}\, d_1\, n^{3/2} \text{ (approximately)}$$

If n is large the crest factor of the deviation will be four just as for noise.

$$d_s(\text{peak}) = 4d_s(\text{r-m-s}) = 4\sqrt{\frac{2}{3}}\, d_1 n^{3/2} = \frac{B}{2}$$

$$d_1 = \sqrt{\frac{3}{2}}\,\frac{B}{8n^{3/2}}$$

$$f_{dm} = d_m = \sqrt{\frac{3}{2}}\,\frac{B}{8n^{3/2}}\,(2m-1)$$

$$\frac{f_{dm}}{f_{scm}} = \frac{\sqrt{\frac{3}{2}}\,\frac{B}{8n^{3/2}}\,(2m-1)}{M_1 f_m\,(2m-1)} = \sqrt{\frac{3}{2}}\,\frac{B}{8\,M_1\,n^{3/2} f_m}$$

$$\frac{\text{S/N (FM-PM)}}{\text{S/N (AM)}} = \frac{\sqrt{3}}{\sqrt{2}} M_1 \sqrt{\frac{3}{2}}\,\frac{B}{8\,M_1\,n^{3/2} f_m} = \frac{3}{16}\,\frac{B}{n^{3/2} f_m} \qquad (16)$$

SS-AM

The signal noise ratio for SS-AM is better than for AM-AM by the following factor. The factor is made up of several contributing components. Twice the amplitude may be used on one side band when the other is omitted, but this has no net effect. Omitting the carrier (the sub-carrier in SS-AM) allows twice as much modulation to be used. Receiving from one half the bandwidth reduces the noise by $\sqrt{2}$. The net gain in S/N $= 2\sqrt{2} = 4/\sqrt{2}$.

Then $$\frac{\text{S/N (SS-AM)}}{\text{S/N (AM)}} = \frac{4}{\sqrt{2}} \times \frac{\sqrt{2}}{4n} = 1/n \text{ (for linear summation)} \qquad (17)$$

This is fairly obvious directly, inasmuch as (SS-AM) for the lowest possible channel becomes simple AM. For the other channels it is simple AM with sidebands displaced in frequency.
For large values of n:

$$\frac{\text{S/N (SS-AM)}}{\text{S/N (AM)}} = \frac{1}{\sqrt{n}} \text{ (for r-m-s summation)} \qquad (18)$$

No crest factor need be considered if the crest factor of each signal separately is about the same as that of the complex resultant of all the signals, namely, 4 to 1. For two-way telephone service the effective crest factor of each channel is greater than four so that the S/N ratio

is better than indicated by Equation 18 by a factor discussed under Minimum Signal Requirements.

SS-PM

In this system, each channel is pure PM. The signal noise ratio for each channel compared to the signal to noise ratio for a single AM channel is the modulation index of that channel. This is independent of the overall bandwidth of the receiver and whether or not this band is utilized with further modulation by the other channels.

To secure the maximum possible modulation index for each channel without overshooting the given bandwidth, the channel sub-carriers should be spaced as close together as practical from a filtering standpoint.

Let f_m = highest audio frequency in each channel

$$f_c = \frac{f_m}{y} = \text{sub-carrier spacing}$$

$$y = \frac{f_m}{f_c} = \text{from .5 to .8 (= .7 in later numerical work)}$$

Then the sub-carrier frequencies are as follows:

$$0, f_c, 2f_c, 3f_c, - - - (n-1)f_c$$

If y is nearly unity then after suppressing the sub-carrier and one sideband the signal energy may be imagined to be concentrated at the centers of the spaces between sub-carriers, i.e. at

$$f_c/2, 3f_c/2, 5f_c/2, 7f_c/2 - - - (2n-1)f_c/2$$

The above assumption is quite accurate for the higher channels, but less so for the lower ones. However, the net effect is that a slightly too great allowance is made for the modulation of the first few channels which reduces the signal noise ratio for all channels by a very small amount. In general no correction is necessary.

If the modulation index is M_2 for each channel when applied to the main carrier, then the frequency deviation for the channels is:

$$M_2f_c/2, 3M_2f_c/2, 5M_2f_c/2 - - - M_2(2n-1)f_c/2$$

The linear summation of the frequency deviation is:

$$\frac{M_2f_c}{2}[1+3+5 - - - + (2n-1)] = \frac{M_2f_c}{2}n^2$$

If this is set equal to one half the radio-frequency band width B, then:

$$B = M_2f_c n^2$$

$$M_2 = \frac{B}{(f_c\, n^2)} = \frac{\text{S/N (SS-PM)}}{\text{S/N (AM)}} \text{ (linear summation)} \qquad (19)$$

However, if the number of channels n is large the frequency deviations should be added up r-m-s. For simple r-m-s summation assuming all channels active, we have:

$$f_d(\text{r-m-s}) = \frac{M_2(\text{r-m-s})\, f_c}{2} \sqrt{1^2 + 3^2 + 5^2 - - - + (2n-1)^2}$$

$$= \frac{M_2(\text{r-m-s})\, f_c}{2} \sqrt{\frac{4}{3} n^3 - \frac{n}{3}}$$

(This equivalence is exact, but the proof is involved and need not be repeated here.)

If n is large the term $n/3$ is negligible compared to $4n^3/3$.

So $$f_d(\text{r-m-s}) = \frac{M_2(\text{r-m-s})\, f_c}{\sqrt{3}}\, n^{3/2}$$

$$M_2(\text{r-m-s}) = \frac{\sqrt{3}\, f_d(\text{r-m-s})}{f_c\, n^{3/2}}$$

(where M_2(r-m-s) is the r-m-s deviation of one channel divided by its mean signal frequency)

If the crest factor of the signal in each channel is the same as that of the complex summation resultant (which should be 4 just as for noise) then the same equation applies for peak values.

$$\frac{\text{S/N (SS-PM)}}{\text{S/N (AM)}} = M_2(\text{peak}) = \sqrt{3} \frac{f_d(\text{peak})}{f_c\, n^{3/2}}$$

$$= \frac{\sqrt{3}\ B}{2 f_c\, n^{3/2}} \text{ (r-m-s summation)} \qquad (20)$$

Summary of Derivation for Frequency Division Systems

The derivation of the formulas for the frequency division systems is summarized in Table II (Page 317). The value of

$$\frac{\text{S/N (XX-XX)}}{\text{S/N (AM)}} = K_1 \times M_1 \times M_2 \times K_2$$

K_1 is the coefficient connected with the first modulation. It is $\sqrt{3}$ if this modulation is FM. It is unity otherwise.

	K_1	M_1	M_2	K_2	$\frac{\text{S/N (XX-XX)}}{\text{S/N (AM)}} = $	M_2	$\frac{\text{S/N (XX-XX)}}{\text{S/N (AM)}}$
	coefficient for triangular noise spectrum		modulation index for second modulation	factor required because the two sub-carrier sidebands add signals directly but add noise r-m-s	Formula for S/N (XX-XX) / S/N (AM)	modulation index for second modulation (substitute for value used for linear summation)	Formula for S/N (XX-XX) / S/N (AM)
	For Linear Summation					For r-m-s Summation	
AM-AM	1	1	$\frac{1}{2n}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{4n}$	$\frac{\sqrt{2}}{4n^{\frac{1}{2}}}$	$\frac{1}{4n^{\frac{1}{2}}}$
AM-PM	1	1	$\frac{B}{4n^2 f_c}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{8}\frac{B}{n^2 f_c}$	$\frac{1}{8}\sqrt{\frac{\sqrt{3}}{2}\frac{B}{n^{3/2} f_c}}$	$\frac{\sqrt{3}}{16}\frac{B}{n^{3/2} f_c}$
FM-AM	$\sqrt{3}$	$\frac{f_d}{f_m} = \frac{B}{4n f_m}$	$\frac{1}{n}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{6}}{8}\frac{B}{n^2 f_m}$	$\frac{\sqrt{3}}{4n^{\frac{1}{2}}}$	$\frac{\sqrt{3}}{16}\frac{B}{n^{3/2} f_m}$
FM-PM	$\sqrt{3}$	$1 < \frac{f_d}{f_m} < \frac{B}{4n f_m}$	$\frac{B}{2n^2 f_m M_1}$	$\frac{\sqrt{2}}{2}$	$\sqrt{\frac{3}{2}\frac{B}{2n^2 f_m}}$	$\frac{1}{8}\sqrt{\frac{\sqrt{3}}{2}\frac{B}{M_1 n^{3/2} f_c}}$	$\frac{3}{16}\frac{B}{n^{3/2} f_m}$
SS-AM	1	1	$\frac{1}{n}$	1	$\frac{1}{n}$	$\frac{1}{\sqrt{n}}$	$\frac{1}{\sqrt{n}}$
SS-PM	1	1	$\frac{B}{n^2 f_c}$	1	$\frac{B}{n^2 f_c}$	$\frac{\sqrt{3}}{2}\frac{B}{n^{3/2} f_c}$	$\frac{\sqrt{3}}{2}\frac{B}{n^{3/2} f_c}$

Table II—Frequency Division Systems

M_1 is the modulation index or deviation ratio for each channel separately. It is unity for AM or SS, but is $\frac{f_d}{f_m} = \frac{B}{4n f_m}$ for FM, except for FM-PM, where it may take any value from 1 to $\frac{B}{4n f_m}$ without changing S/N since M_2 is made inversely proportional to M_1.

M_2 is the modulation index for the second modulation. For linear summation, it is $1/n$ for FM-AM and SS-AM, but $1/2n$ for AM-AM because room must be left for 100 per cent modulation in the first modulation. For AM-PM it is $\frac{B}{4n^2 f_c}$. For FM-PM it is $\frac{B}{2n^2 M_1 f_m}$. For SS-PM it is increased another 2 to 1 because the frequency of the center of each band in SS is one half the corresponding sub-carrier frequency in the AM-PM system, so that a given frequency deviation means twice the modulation index.

K_2 is a factor required because the signal from the two sub-carrier sidebands adds directly while the noise adds as the square root of the sum of the squares. It is $\sqrt{2}/2$ for all frequency division systems except (SS-XX) where it is unity.

For r-m-s summation M_2 is $1/\sqrt{n}$ for (SS-AM), if it is assumed that the crest factor of each signal is 4 just as in a complex wave of random phases. For AM-AM, M_2 is $\frac{\sqrt{2}}{4n^{\frac{1}{2}}}$ ($\sqrt{2}$ is the crest factor of the sub-carrier sine wave. The modulation is assumed to have a negligible effect on the r-m-s value. The crest factor of the complex wave is 4). M_2 for FM-AM is the same.

The value of M_2 for AM-PM is $\frac{1}{8}\sqrt{\frac{3}{2}}\frac{B}{n^{3/2} f_c}$ as proved in the text. For FM-PM it is $\frac{1}{8}\sqrt{\frac{3}{2}}\frac{B}{M_1 n^{3/2} f_m}$. For SS-PM the value is larger than for AM-PM by $\frac{4}{\sqrt{2}} \times 2$. The $\frac{4}{\sqrt{2}}$ is the ratio of crest factors and the 2 is the ratio of sub-carrier frequencies for the two cases.

Time Division Systems

PAM-AM

In this and the other time division systems, the fundamental pulse rate is nK_4f_m where n is the number of channels.

K_4 is the ratio of the pulse rate to the highest transmitted signal frequency for pulses assigned to one channel. Theoretically in a PAM-AM system K_4 may have a value as low as 2. However, if it does it becomes necessary to have an infinitely sharp filter to avoid the distortion components which come in at frequencies just above the frequency of f_m. To allow for filtering difficulties it is more usual to let $K_4 = 2.5$ or 3. In this report K_4 will be given the value 3 for PAM, PPM, and PWM systems.

f_m is the highest audio signal frequency in each channel.

The pulses are separated into groups of n consecutive pulses per group (or $n+1$ if a synchronizing pulse is used). The first pulse in each group is assigned to the first channel, the second to the second channel, etc.

In PAM each pulse is amplitude modulated with the appropriate signal.

The signal noise ratio tends to be poorer than for straight AM because of the extra noise picked up by the wide radio-frequency band necessary to transmit the short pulses. On the other hand, the signal noise ratio tends to be better because the instantaneous value of the signal during the signal peaks is higher than for a continuous-wave (C.W.) system by a factor equal to $1/\sqrt{D}$ where D is the duty factor, or the ratio of the duration of a pulse to the time between pulses. These two effects exactly cancel each other. However, to transmit n times as many pulses requires n times as much power. Or with a given amount of power, n times as many pulses means $1/\sqrt{n}$ times the amplitude.

Then
$$\frac{\text{S/N (PAM-AM)}}{\text{S/N (AM)}} = \frac{1}{\sqrt{n}}. \qquad (21)$$

Note that here there is no question of "type of summation" or division of the modulation. The S/N ratio is the same as for SS-AM assuming root-mean-square summation.

*PAM±FM**

A PAM ± FM system is one which multiplexes several channels on

* O. E. Dow furnished most of the theory in this section.

to a single radio-frequency carrier on a time division basis. The amplitude modulated pulses frequency modulate the radio-frequency carrier. The channel pulses are modulated plus and minus about a zero mean value.

The receiving multiplex unit consists essentially of n gates, each of which registers with a particular channel pulse. Associated with each gate is a low pass filter and an audio amplifier. The gate will be considered as an amplifier whose gain is increased from zero to unity and back to zero at a rate equal to the pulse repetition rate of each channel. The gate is considered to be "open" when the gain of the amplifier is unity. The duration of each "open" time will normally be equal to or less than $1/nf_p$. The time required for the gate to open or close will be considered zero.

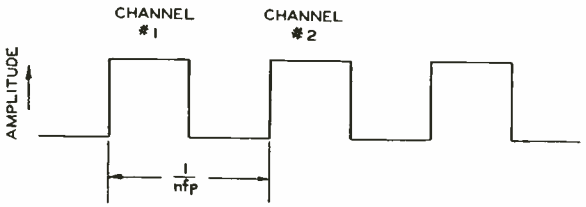
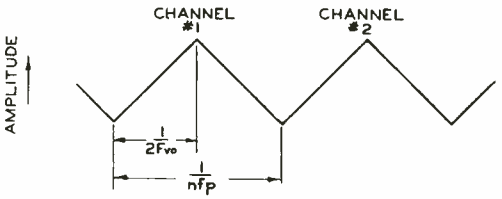


Fig. 3—Normal multiplex transmitter output.

Fig. 4 — Multiplex signal through minimum frequency band.



The initial shape of the channel pulses generated by the transmitting multiplex unit is shown in Figure 3. Only fifty per cent of the total available time is utilized, however, if the pulses are passed through a filter with a build-up time of $1/2nf_p$ the pulses will be distorted as shown in Figure 4. (The actual shape of the pulses will be considerably different, but the simpler triangular pulse illustrates the point.) The duration of the pulse above 50 per cent amplitude is $1/B_0$ where B_0 is the intermediate-frequency bandwidth corresponding to the modulation band F_{vo}. For a double sideband system $2F_{vo} = B_0$. For the triangular pulse the build-up time will be equal to $1/2F_{vo}$. The minimum modulation band it is permissible to transmit the signal pulse through will be set as $F_{vo} = nf_p$. It is realized that this minimum is not fundamental, but until more experience has been obtained with crosstalk balancing circuits this value will be used. If the duration of the gate is set at $1/nf_p$ the signal received will not vary as the pulse shape is

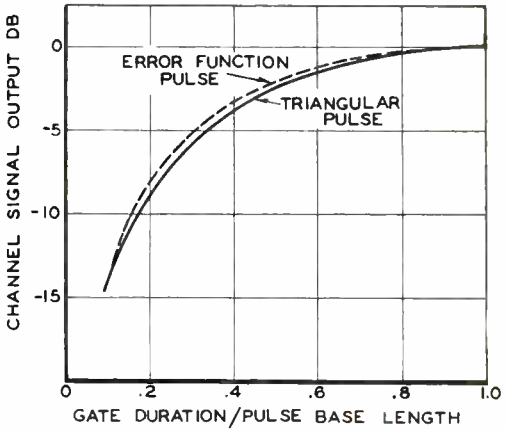


Fig. 5 — Variation of signal with gate width in a PAM system.

changed from that of Figure 3 to that of Figure 4 by a successive reduction of the transmission frequency band. Here again the use of such a wide gate will depend upon the skill and ingenuity exercised in reducing crosstalk between channels. More will be said later about the present status of crosstalk balancing technique.

The signal in any particular channel is the average of the voltage for that channel pulse. In Figure 3, if the peak amplitude corresponding to 100 per cent modulation is one, the peak signal in each channel output is $1/2n$. In Figure 4, if the gate is $1/nf_p$ seconds wide, the signal in each channel is still $1/2n$. The variation of the signal, as the gate open time is reduced, is as shown in Figure 5. The error function pulse approximates in shape the pulses generally obtained. It will be observed from Figure 5 that the signal output variation with change in width of gate is approximately the same for the two pulse shapes. Also, if the base of the pulse is equal to the channel interval time, approximately one decibel of signal is lost in operating the gate at $\frac{2}{3}$ the total channel interval. The crosstalk will be reduced to a much greater degree by operating with the narrower gate.

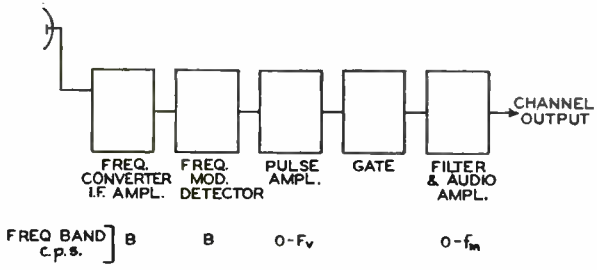


Fig. 6—Receiving circuit frequency band widths.

The noise which appears at the output of any one channel will now be determined. The intermediate-frequency band will be independent of the pulse frequency band, F_p, but equal to or greater than twice F_c (Figure 6). The signal must at all times exceed the threshold value. The value of the peak noise in the channel output is derived as follows:

Channel Output Noise in a Time Division Multiplex System

The peak amplitude of the carrier in the intermediate-frequency amplifier is taken as C and the peak amplitude of the noise in a one cycle per second band at the same point in the intermediate-frequency amplifier is e. Since C will be large compared to e the phase modulation of C by e will be e/C, and the corresponding frequency modulation is $(e/C)f_x$, where f_x is the difference in frequency between the noise component and the carrier. If the useful width of the intermediate-frequency amplifier is B cycles per second the peak frequency deviation corresponding to 100 per cent modulation is $B/2$. The fractional part of peak modulation for the one cycle band of noise is $(e/C)f_x(2/B)$ (peak). The resulting incremental voltage following the frequency modulation discriminator is:

$$e_v = \sqrt{2}\,\frac{e}{C}\,\frac{2f_x}{B}\ \text{(peak)}.$$

The factor $\sqrt{2}$ is required because e_v is considered to be the noise in a one cycle bandwidth at a frequency interval f_x on each side of the carrier. The discriminator output is taken to be unity at peak deviation. This assumption does not affect the signal-to-noise ratio.

To determine the effect of the gate on this noise voltage it will be considered as the real part of $E(t)$. $E(t) = e_v\epsilon^{j2\pi f_x t}$

The gate will be represented as the function $F(t)$

$$F(t) = f_p t_g \left[1 + \sum_{b=1}^{\infty} \frac{\sin(b\pi t_g f_p)}{\pi b t_g f_p} \left(\epsilon^{+j2\pi b f_p t} + \epsilon^{-j2\pi b f_p t} \right) \right]$$

where t_g = gate duration time, b = any integer. The noise voltage at the gate output due to the voltage e_v at its input will be:

$$F(t)E(t) = e_v f_p t_g \left[\epsilon^{j2\pi f_x t} + \sum_{b=1}^{\infty} \frac{\sin(\pi b t_g f_p)}{\pi b t_g f_p} \left(\epsilon^{j2\pi(bf_p + f_x)t} + \epsilon^{-2\pi(bf_p - f_x)t} \right) \right]$$

The components which fall within the audio pass band 0 to f_m are:

$$e_r f_p t_g \left[\epsilon^{j2\pi f_x t} + \sum_{b=1}^{\infty} \frac{\sin(\pi b t_g f_p)}{\pi b t_g f_p} \cdot \epsilon^{-j2\pi(bf_p - f_x)t} \right]$$

$$\text{where } -f_m < (bf_p - f_x) < +f_m$$

If $f_p = 2f_m$, a component of noise voltage will fall within the audio pass band for all values of f_x, and the peak amplitude of these components may be found from the following equation,

$$e_{an} = e_r f_p t_g \frac{\sin \pi b t_g f_p}{\pi b t_g f_p} \qquad \begin{matrix} b = 0 \text{ to} \\ b = \infty \end{matrix} \qquad \text{(peak)}$$

by letting b vary continuously (instead of by unity steps from 0 to F_v and letting $bf_p = f_x$. Making this substitution,

$$e_{an} = e_r f_p \frac{\sin \pi f_x t_g}{\pi f_x} = \sqrt{2}\, \frac{e}{C} \frac{2}{B} \frac{f_p}{\pi} \sin \pi f_x t_g$$

The peak voltage E'_n at the channel output due to the noise in the frequency band $0 - E_r$ ahead of the gate is found by summing up the noise component e_{an} as follows:

$$E'_n = \sqrt{\int_0^{F_r} e_{an}^2 df_x}$$

$$(E'_n)^2 = \frac{8}{\pi^2} \left(\frac{f_p}{B}\right)^2 \left(\frac{e}{C}\right)^2 \int_0^{F_v} \sin^2(\pi t_g f_x)\, df_x$$

$$= \frac{8}{\pi^2} \left(\frac{f_p}{B}\right)^2 \left(\frac{e}{C}\right)^2 \frac{F_r}{2} \left(1 - \frac{\sin 2\pi t_g F_r}{2\pi t_g F_r}\right)$$

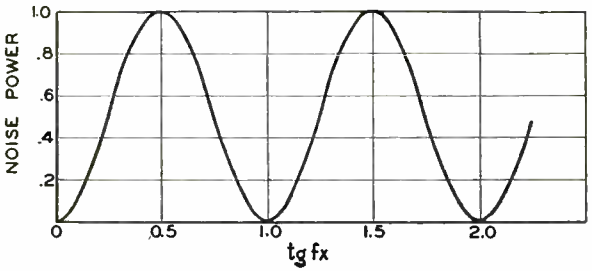
$$E'_n = \frac{2}{\pi} \frac{f_p}{B} \sqrt{F_v \left(1 - \frac{\sin 2\pi t_g F_r}{2\pi t_g F_r}\right)}^{1/2} \frac{e}{C} \qquad \text{(peak)}$$

In practice it will be necessary to have f_p greater than $2f_m$ hence some of the components of noise will fall outside the audio pass band The noise components represented by the unshaded portion below the bottom curve of Figure 7 will not fall in the audio pass band f_m. The proportion of the noise power represented by $(E'_n)^2$ which will fall

within the audio pass band is $2f_m/f_p$. Hence E_n must be reduced by the factor $\sqrt{2f_m/f_p}$

$$E_{noise} = E_n \sqrt{\frac{2f_m}{f_p}}$$

$$= \frac{\sqrt{8}}{\pi} \frac{\sqrt{f_m f_p}}{B} \sqrt{F_v} \left(1 - \frac{\sin 2\pi t_g F_v}{2\pi t_g F_v}\right)^{1/2} \frac{e}{C} \text{ (peak)} \quad \text{(a)}$$



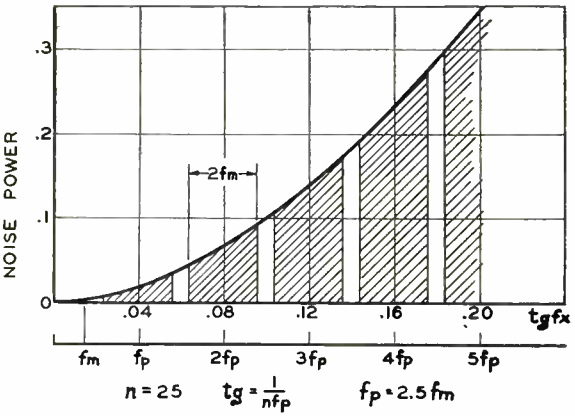


Fig. 7 — Noise power vs. video frequency f_v in PAM ± FM system.

E_{noise} is the peak noise voltage in an audio frequency band f_m at a channel output due to the gating of a triangular noise spectrum which extends from 0 to F_v cycles per second. The gate repetition rate is f_p pulses per second and the gate duration is t_g seconds. The intermediate-frequency amplifier bandwidth ahead of the frequency modulation detector is B cycles per second. The peak carrier amplitude in the intermediate-frequency amplifier is C. The peak noise amplitude at the same point in the intermediate-frequency amplifier is $\sqrt{B}\, e$. The gain of the audio amplifier is not considered.

In Equation (*a*) the value of the term in the parenthesis is represented in the curve of Figure 8 and indicates the variation of noise with the duration of the gate if F_v is maintained constant.

The value of F_v will depend upon the number of channels, n, and the pulse repetition rate, f_p, of each channel. Thus, $F_v = knf_p$, where k is a multiplier determined by crosstalk considerations. The length of each gating interval t_g is $1/nf_p$. Upon substitution of these values for F_v and t_g Equation (a) becomes:

$$E_{noise} = \frac{\sqrt{8}}{\pi} \quad \frac{\sqrt{n f_m}}{B} \quad f_p \left[k \left(1 - \frac{\sin 2\pi k}{2\pi k} \right) \right]^{1/2} \quad \frac{e}{C} \text{ (peak)} \qquad (b)$$

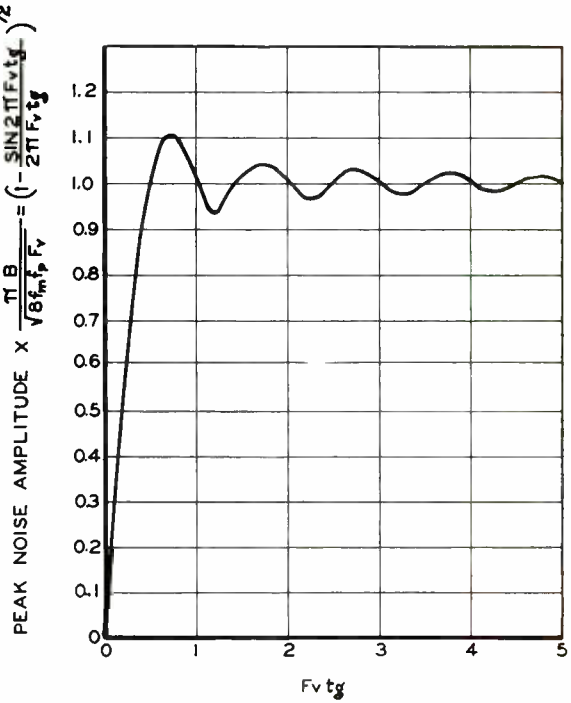


Fig. 8—Channel output noise vs. gate duration.

The terms outside the square brackets are fixed by the system specifications. This is only true of B if the peak frequency deviation is greater than F_v. A plot of the term within the square bracket is shown on Figure 9. The smaller the value of k the lower the noise level at the channel output terminals. The first zero slope point above zero occurs at $k = 1$ $(t_g = 1/nf_p)$. This value of k will also produce a value of crosstalk which may readily be balanced out. If k is set equal to one, Equation (b) becomes:

$$E_{noise} = \frac{\sqrt{8}}{\pi} \quad \frac{\sqrt{n f_m}}{B} \quad f_p \quad \frac{e}{C} \quad \text{(peak)} \qquad (c)$$

Under the conditions represented in Figure 4 the peak signal (in the output of a low pass filter with cut off at f_m) is $1/2n$, hence the signal-to-noise ratio is:

$$S/N\ (PAM \pm FM) = \frac{\pi}{\sqrt{32}\, n^{3/2}} \frac{B}{f_m^{1/2} f_p} \frac{C}{e}\ (peak) \qquad (d)$$

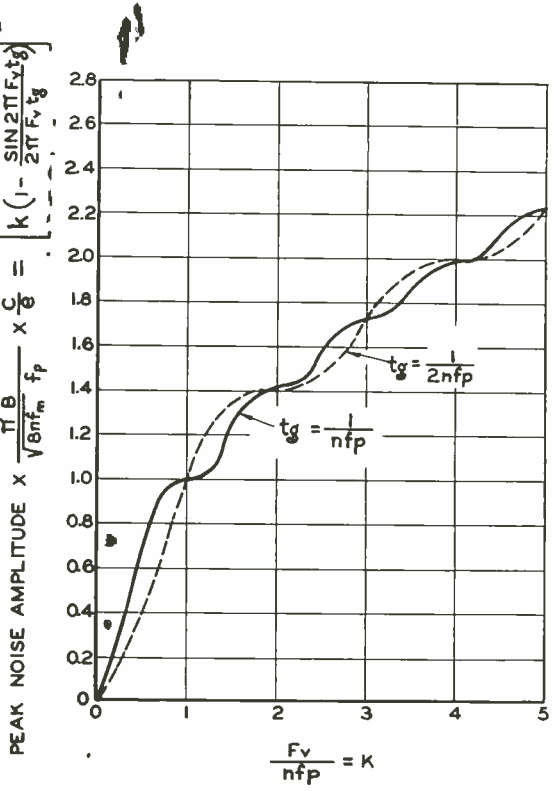


Fig. 9—Channel output noise vs. video frequency band F_v.

The signal-to-noise ratio of a single channel amplitude modulation system is:

$$S/N\ (AM) = \frac{1}{\sqrt{2f_m}} \frac{C}{e}\ (peak) \qquad (e)$$

Dividing Equation (d) by Equation (e) gives the improvement factor,

$$R_0\ (PAM \pm FM) = 12^{-1}\pi B n^{-3/2} f_m^{-1} = 0.262\, n^{-1/2} r \qquad (f)$$

The ratio,

$$\frac{R_0\ (SS - PM)}{R_0\ (PAM \pm FM)} = \frac{0.866\, y}{0.262} = 3.3y = 2 \quad \begin{matrix}(\text{assuming} \\ y = 0.6)\end{matrix}$$

The relations derived from Equation (*c*) to Equation (*f*) inclusive have been for the condition that $F_r = nf_p$ or $k = 1$. It has been demonstrated in the laboratory that F_r may be reduced to $nf_p/2$ and still obtain a crosstalk ratio of 57 decibels between adjacent channels if the delay line method of crosstalk balancing is used. The signal noise ratio for the condition $F_r = 0.5\, nf_p$ will be derived. The noise is given in (Equation (*b*). The buildup time of the signal pulse will be $1/nf_p$, and the average of the pulse will be $1/n$, since the gate must be restricted to a duration of $1/nf_p$, and the average of the pulse will be $1/n$. Since the gate must be restricted to a duration of $1/nf_p$, or 1/2 the base of the pulse the average within the gate period will be 2½ decibels less than $1/n$. (See Figure 5.) Thus the peak signal will be $1/1.33n$. The signal-to-noise ratio will be,

$$\text{S/N (PAM} \pm \text{FM)} = \frac{.75\pi}{\sqrt{8}\,\sqrt{f_m\, k\, n^3 f_p}} \frac{B}{e} \frac{C}{\text{(peak)}} \qquad \text{(g)}$$

Dividing Equation (g) by Equation (e) gives the improvement factor:

$$R_o \text{ (PAM} \pm \text{FM)} = \pi\sqrt{2}\ \ 8^{-1} B\, f_m{}^{-1}\, n^{-3/2}$$
$$= 0.555\, B\, f_m{}^{-1}\, n^{-3/2} \qquad = 0.555\, n^{-\frac{1}{2}} r$$

(It is assumed that $F_r = 1/2\, nf_p$ and $f_p = 3\, f_m$, and $K = 1/2$)

PPM-AM

PPM is chosen rather than PFM because the phase deviation is limited to about $2\pi/3n$ radians for multiplex operation even with low-frequency modulation. (See page 334.)

To obtain the signal noise ratio for this system compared to AM, the actual signal noise ratio is calculated for each system at a signal level corresponding to a peak carrier noise ratio in the radio-frequency circuit of 2. The same bandwidth of radio-frequency circuits are assumed for both systems in the derivation.

To solve for the peak time displacement of the pulses caused by the noise at threshold consider Figures 10 and 11. Figure 10 shows how the use of a clipper and a limiter prevents noise between pulses and noise in the middle of pulses from affecting the output even if a gate is not used.

It also shows the origin of the factor 2 given in the previous paragraph since noise between pulses must fall just short of reaching half the signal amplitude to avoid piercing the clipper level, while noise occurring during the pulse must fall just short of half the signal ampli-

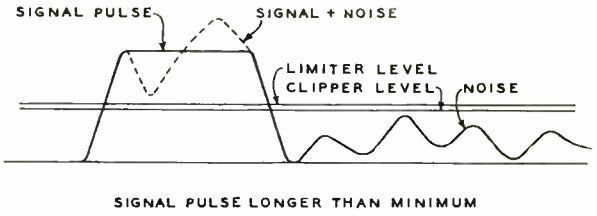


Fig. 10—Signal pulse longer than minimum.

tude to avoid piercing the limiter level. The figure is drawn for a signal pulse which is longer than the minimum value which the system would pass. The pulse has a flat top in the absence of noise.

In Figure 11 minimum-length pulses are represented. Being of minimum length these pulses are roughly triangular in shape. The dotted lines are used to represent the wave of signal plus noise.

On the left the signal and noise are assumed to have the same radio-frequency phase. On the right the signal and noise are assumed to be 180 degrees out of phase. It should be noted that the minimum duration of a noise pulsation is the same as that of the signal. If a noise pulse has its peak near the beginning or the end of a signal pulse, it has the effect of shifting the timing of the start or finish of the current pulse in the clipper-limiter output circuit. A noise pulse of a given amplitude has the greatest effect on the timing when the noise peak occurs at the instant the wave of signal plus noise pierces the clipper level. Note that a noise pulse of near threshold amplitude, may affect the timing of both the start and the finish of the signal pulse. For weaker noise, and for noise which is in phase with the signal the shift of the start and of the finish are independent.

The line AC represents the start of a pulse. If a noise peak (whose radio-frequency phase is the same as that of the signal pulse) occurs at point A with amplitude AF it can change the start point of the current in the clipper output circuit from P to F. Now the minimum width of a pulse passed by an amplifier of frequency bandwidth B is $a = 1/B$, where a is measured at 50 per cent of peak amplitude. Also

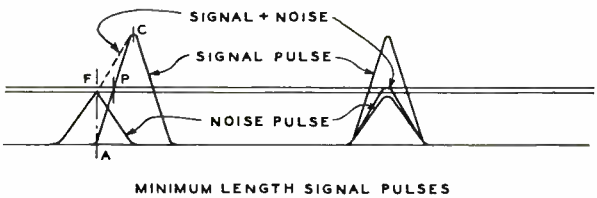


Fig. 11—Minimum length signal pulses.

the time of rise[3] is approximately a. Then $PF = a/2 = 1/2B$. If the radio frequency phase of the noise is reversed the start of the current is shifted in time an equal amount in the opposite direction. Both the beginning and end of the pulse are subject to this shift and since they may both be shifted in the same direction it has been assumed in some reports that the peak time deviation at threshold is $1/2B$.

Actually there is a small error involved in this assumption. The peak deviation occurs on both sides of a given pulse simultaneously so seldom that it may be neglected. The noise at the beginning of a pulse is almost independent of that at the finish of the pulse. The noise at the beginning deviates the center of gravity of the pulse one half as much as it deviates the start, or $1/4B$ at threshold. While the peak value for noise on the finish of the pulse is the same, the two values add as the square root of the sum of the squares. Thus, the peak time deviation at threshold is $1/2\sqrt{2}\,B$.

Most of the systems in common use utilize only the leading edge of the pulse to determine the effective pulse phase. This practice involves a sacrifice in signal noise ratio of $\sqrt{2}$. Unfortunately it is difficult to devise a system in which the effective pulse phase is determined by the center of gravity of the pulse. It can be done in effect by building two complete demodulation systems, one working on the leading edge of the pulse and one on the trailing edge of the pulse, and adding the signal outputs.

For the AM case, assume a C.W. AM signal of such an average power that the carrier amplitude is twice the peak amplitude of the noise. Since the signal carrier is by far the largest single component in the signal and noise wave, the components in the output of the signal detector are largely the result of the other radio-frequency components heterodyning the carrier. Beats between other components are negligible. Then the noise components have a frequency spectrum through the audio and supersonic bands up to the frequency $B/2$. The higher audio and supersonic components may be filtered out leaving the amplitude of the remaining noise the same as though the radio-frequency band width had been reduced to the value $2f_m$. Then the signal-to-noise ratio for AM on a peak basis at the chosen signal level is:

$$\text{S/N (AM)} = 2\left(\frac{B}{2f_m}\right)^{1/2} \tag{24}$$

For the same average power in a PPM-AM system the peak signal noise ratio in the receiver output may be written,

[3] E. A. Giullemin, COMMUNICATION NETWORKS, Vol. II (p. 487, Fig. 207), John Wiley and Sons, New York, N. Y., 1935.

$$\text{S/N (PPM-AM)} = \phi \times \frac{2B}{\sqrt{2}\pi f_p} \left(\frac{1}{naf_p}\right)^{\frac{1}{2}} \left(\frac{f_p}{f_m}\right)^{\frac{1}{2}} \tag{25}$$

In this equation the factor ϕ represents the angular deviation or modulation index due to the modulation of the pulse frequency by the signal.

The factor $\sqrt{2}\,\pi f_p/2B$ represents the peak angular deviation which would result from the noise if the noise peak amplitude were just one half the pulse amplitude. (The time deviation of the pulse from its normal position would be $a/2\sqrt{2}$ which is equal to $1/(2\sqrt{2}B)$. The angular deviation is then $\sqrt{2}\,\pi f_p/2B$).

Actually since the average radiated power is being kept constant the pulse amplitude is greater than twice the noise amplitude by a factor equal to one over the square root of the duty factor (the crest factor of the pulsed signal). The duty factor is,

$$a \Big/ \frac{1}{nf_p} = anf_p.$$

The crest factor is then $(1/anf_p)^{\frac{1}{2}}$.

The noise in each channel of the receiver output has a uniform frequency spectrum up to the frequency f_p. The noise is reduced by a law pass filter by the factor $(f_p/f_m)^{\frac{1}{2}}$.

At first sight it might seem that the noise spectrum would extend up to the frequency $B/2$ so that the factor $(f_p/f_m)^{\frac{1}{2}}$ should be changed to $(B/2f_m)^{\frac{1}{2}}$, just as in the PAM-AM case. That this is not true may be seen from the following argument. Consider the pulses of a single channel in the output of a receiver after the first demodulation and commutation. Each pulse is deviated from its normal position by random noise, each by an amount independent of that of its neighbor. If the chain of pulses is put through a low pass filter and the cut-off frequency gradually varied from $B/2$ to lower values, the pulses will decrease in amplitude and increase in time duration, but each will retain the same noise phase deviation until the pulses begin to overlap. This occurs when the cut-off frequency approaches f_p in value. Thus, $(f_p/f_m)^{\frac{1}{2}}$ is the correct factor.

The ratio of signal-to-noise ratios is then,

$$\frac{\text{S/N (PPM-AM)}}{\text{S/N (AM)}} = \frac{\phi \quad B}{\pi f_p n^{\frac{1}{2}}} \tag{26}$$

(In this equation it is assumed that the minimum pulse length is used so that $1/a = B$ and $(B/a)^{\frac{1}{2}} = B$).

PPM-AM From the Sideband Standpoint

It is of some interest to derive a similar formula by analyzing the pulses into their sideband components and calculating the contribution of the noise components near each sideband. To simplify the work, the assumption will be made that each pulse has the familiar shape $\sin x/x$ corresponding to each sideband component having equal amplitude. Let the number of harmonics of f_p passed be

$$S = \frac{B}{2 f_p}. \tag{27}$$

Then the radio-frequency components consist of the carrier with S sidebands on each side, or a total of $2S + 1$ components each of equal amplitude. If E is the amplitude of a sine wave of power equivalent to that of n channels, then the amplitude of each component for a single channel is $\frac{E}{n^{\frac{1}{2}}(2S+1)^{\frac{1}{2}}}$.

The instantaneous voltage for the pulses of one channel may be represented by,

$$e_s = \frac{E}{n^{\frac{1}{2}}(2S+1)^{\frac{1}{2}}} \cos w_o t\,(1 + 2\cos w_p t + 2\cos 2\,w_p + - - - + 2\cos S w_p t). \tag{28}$$

The voltage of a single radio-frequency noise component within an audio frequency of the j^{th} harmonic sideband may be represented by,

$$e_n = \triangle \sin\,(w_o t + j w_p t + w_a t + \theta)$$

where $\triangle$ = the infinitesimal amplitude of a single noise component.

j = any integer from 1 to S $\quad$ θ = a random phase angle.

w_o = 2π times the carrier frequency $\quad$ $w_p = 2\pi f_p$

w_a = 2π times the audio frequency difference between the frequency of the noise component and the frequency of the j^{th} harmonic sideband.

The value of e_n may be written

$$e_n = \triangle[\sin w_o t \cos\,(j w_p t + w_a t + \theta) + \cos w_o t \sin\,(j w_p t + w_a t + \theta)].$$

When e_n is added to the signal the first term of the bracket has a negligible effect on the shape and position of the radio-frequency envelope affecting the radio-frequency phase, but the second term shifts the envelope to one side. The instantaneous value of the envelope is,

$$E_s = \frac{E}{n^{\frac{1}{2}}(2S+1)^{\frac{1}{2}}}(1 + 2\cos w_p t + 2\cos 2w_p t + 2\cos S\, w_p t) + \Delta \sin (j\, w_p t + w_a t + \theta). \qquad (29)$$

The slope of the envelope is,

$$E_s^1 = \frac{E\, w_p}{n^{\frac{1}{2}}(2S+1)^{\frac{1}{2}}} 2\,(\sin w_p t + 2\sin 2\, w_p t - - - + S \sin S\, w_p t) + (w_p + \frac{w_a}{j})\, \Delta j \cos (j\, w_p t + w_a t + \theta). \qquad (30)$$

The peak of the envelope occurs at the instant $E_s^1 = 0$. Maximum time displacement due to the noise component occurs on any pulse where $w_a t + \theta = 0$. For such a pulse at the peak,

$$\frac{E \quad w_p}{n^{\frac{1}{2}}(2S+1)^{\frac{1}{2}}} 2\,(\sin w_p t + 2\sin 2\, w_p t - - - + S \sin S\, w_p t) = -(w_p + \frac{w_a}{j})\, \Delta j \cos (j w_p t). \qquad (31)$$

Where $w_p t$ is very small, $\sin j w_p t = j w_p t$ and $\cos j w_p t = 1$. Then,

$$\Delta j = \frac{2E\, w_p t_1}{n^{\frac{1}{2}}\,(2S+1)^{\frac{1}{2}}} (1 + 2^2 + 3^2 - - - + S^2) \frac{w_p}{w_p + \frac{w_a}{j}}$$

$$= \frac{2E\, w_p t_1}{n^{\frac{1}{2}}(2S+1)^{\frac{1}{2}}} \left(\frac{S^3}{3} + \frac{S^2}{2} + \frac{S}{6}\right) \frac{w_p}{w_p + \frac{w_a}{j}}. \qquad (32)$$

(This equivalence is exact. The complete verification consists in proving it for one value by a numerical substitution, and then proving that if it is true for $S = S_o$ it is true for $S = S_o + 1$.)

The phase modulation index produced by the noise component is $w_p t_1$, where t_1 is the time when the overall envelope slope goes to zero. The maximum amplitude modulation index for a similar noise component near the carrier in a simple AM case is $\triangle$, E. The ratio of the

$$\text{two indices is } \frac{w_p t_1 E}{\triangle} = j \frac{n^{\frac{1}{2}}(2S+1)^{\frac{1}{2}}}{2\left(\frac{S^3}{3}+\frac{S^2}{2}+\frac{S}{6}\right)} \frac{w_p + w_a/j}{w_p} \qquad (33)$$

In the AM case it is only necessary to consider noise components near the carrier but in the pulse case it is necessary to consider the noise components near each side band, then,

$$\frac{\text{Noise mod. index } (PPM\text{-}AM)}{\text{Noise mod. index } (AM)} = \frac{n^{\frac{1}{2}}(2S+1)^{\frac{1}{2}}}{2\left(\frac{S^3}{3}+\frac{S^2}{2}+\frac{S}{6}\right)} \left(2\left[1+2^2+3^2---+S^2\right]\right)^{\frac{1}{2}}$$

$$= \frac{n^{\frac{1}{2}}(2S+1)^{\frac{1}{2}}}{2^{\frac{1}{2}}} \frac{\left(\frac{S^3}{3}+\frac{S^2}{2}+\frac{S}{6}\right)^{\frac{1}{2}}}{\left(\frac{S^3}{3}+\frac{S^2}{2}+\frac{S}{6}\right)} = \left[\frac{n(2S+1)}{2\left(\frac{S^3}{3}+\frac{S^2}{2}+\frac{S}{6}\right)}\right]^{\frac{1}{2}}$$

$$= \left(\frac{3n}{S^2}\right)^{\frac{1}{2}} \text{(dropping minor terms)} \qquad (34)$$

(Note: In the summation the factor $\frac{w_p + \frac{w_a}{j}}{w_p}$ may be neglected since w_a is positive for some terms and negative for others.)

$$\text{Then, } \frac{\text{S/N (PPM-AM)}}{\text{S/N (AM)}} = \frac{S}{3^{\frac{1}{2}} n^{\frac{1}{2}}} \phi = \frac{1}{2\sqrt{3}} \frac{B}{n^{\frac{1}{2}} f_p} \phi \qquad (35)$$

The pulse slope method gave $\frac{1}{\pi} \frac{B}{n^{\frac{1}{2}} f_p} \phi$

The coefficient for the side band method is $1/2\sqrt{3} = 1/3.464 = 0.289$. The coefficient for the pulse slope method is $1/\pi = .318$. This is a fair check considering the wide difference of method.

It is now desirable to clear the expression of ϕ and f_p, substituting functions of f_m, n and B.

The value of f_p should be minimized. It seems reasonable to assign $f_p = 3f_m$, because this is the smallest value that can be assigned and still keep the lower second harmonic sideband out of the audio band. The time between adjacent pulses for different channels is $1/nf_p$. The time deviation of a given pulse by the signal is about $\frac{1}{3}$ of this as a maximum or $1/(3n\,f_p)$. The exact value of this numerical coefficient is dependent on the duty factor and on the amount of cross modulation that can be tolerated.

The angular deviation is then $\phi = 2\pi f_p \left(\dfrac{1}{3n\,f_p}\right) = \dfrac{2\pi}{3n}$ (36)

(The maximum value of ϕ giving negligible second harmonic sidebands is about 0.5 radians. If $\phi = 0.5 = 2\pi/3n$, then $n = 4\pi/3$. Thus $\phi = 2\pi/3n$ should be used only if n is 4 or larger.)

Substituting $\phi = 2\pi/3n$ and $f_p = 3f_m$ we have, $\dfrac{\text{S/N (PPM-AM)}}{\text{S/N (AM)}}$

$$= \frac{1}{2\sqrt{3}} \frac{B}{3f_m} \frac{2\pi}{3n} \frac{1}{n^{\frac{1}{2}}} = \frac{\pi}{9\sqrt{3}} \frac{B}{f_m n^{3/2}} \quad \text{(sideband method)} \quad (37)$$

$$\text{or } \frac{1}{\pi} \frac{B}{3f_m} \frac{2\pi}{3n} \frac{1}{n^{\frac{1}{2}}} = \frac{2}{9} \frac{B}{f_m n^{3/2}} \quad \text{(pulse slope method)} \quad (38)$$

In later tabulations the pulse slope method solution will be used.

PPM-FM

The derivation of the improvement factor for this system is simplified by making use of the factor derived for PPM-AM.

If it is assumed that the same ratio peak carrier to peak noise exists in the PPM-AM and PPM-FM cases, then in the video circuits, the peak signal to peak noise ratio will be the same except for the factor $\sqrt{3}$ which accounts for the triangular noise spectrum in the *FM* case.

(The deviation ratio is unity because the video circuits are designed to pass all frequencies up to $B/2$). However for equal radiated power the peak carrier voltage in the PPM-AM case should exceed that in the PPM-FM case by the square root of the duty factor, of the n channel system, $\sqrt{nD} = \sqrt{\frac{3n f_m}{B}}$

Therefore

$$\frac{\text{S/N (PPM-FM)}}{\text{S/N (AM)}} = \frac{2}{9} \frac{B}{f_m n^{3/2}} \sqrt{3} \sqrt{\frac{3n f_m}{B}} = \frac{2 B^{\frac{1}{2}}}{3 f_m^{\frac{1}{2}} n} \qquad (39)$$

PNM-AM

No signal noise formula can be assigned to this system because when the system is properly used there is theoretically no way in which incoming noise can affect the output circuits at all. It is necessary that the signal exceed the threshold just as in other systems.

In the typical system, the transmission consists of pulses of constant amplitude, frequency and phase except for systematic omissions. Omitting every other one corresponds to zero modulation. Re-inserting one or more pulses corresponds to positive modulation. Thus the pulse density per unit time is proportional to the modulating voltage. It is similar to PFM-AM in this respect, but differs in that the pulse phase is constant.

At the receiver the modulating wave is recovered by amplifying, detecting, limiting, and clipping and then passing the resulting pulses through a low pass filter. To obtain the perfect noise performance, however, a more complicated receiver must be used. In addition to the signal channel mentioned above, the received pulses are used to synchronize a local oscillator at the fixed pulse frequency. Coupling should be weak enough that noise does not modulate this oscillator. The oscillator output is limited, differentiated, and clipped and limited so that very narrow square pulses result. These are added to the limited signal pulses and the resultant is limited and clipped. If the locally generated pulses occur in the center of the flat top of the limited signal pulse, there is no way for noise to affect the output.

The bandwidth requirement is somewhat greater than for PAM-AM, because the number of pulses must be greater than f_m by a larger factor to avoid distortion. The actual ratio required will have to be discovered by experiment, but it seems probable that it should be about 20. The figure used in calculations is 21 so as to be the same for PAM-PCM-AM considered below. Since the figure is 3 for PAM-AM the bandwidth

required is 7 times as large for PNM-AM, the actual bandwidth being $21 \times n \times f_m \times 3$. (Assuming $B/3$ is the highest frequency pulses that may be transmitted and considered to be independent).

Note that this system has exceptionally low cross modulation for the same reason that it has low noise.

PNM-FM

This system has a perfect signal noise ratio for the same reason as PNM-AM. The threshold is poorer however, since it is a C.W. system, and yet it requires at least as great a bandwidth.

PAM-PCM-AM

A system which does not appear on Table I and which might have important applications, might be designated, PAM-PCM-AM, where the C stand for code. The system starts off with a sampling of the signal wave at a rate of about $3f_m$ just as in PAM. The resulting pulses are rated by amplitude into "p" gradations. This is spoken of as quantizing. A code is assigned to each of the "p" gradations of amplitude. The code consists of a group of pulses. There is a time position for each of "q" pulses, each of which may be present or omitted. The number of possible combinations of "q" pulses omitting from o to q is 2^q. Thus $p = 2^q$ is the number of pulse amplitude gradations that may be sent over the system. For example if q is 7, it is possible to distinguish between 128 different amplitudes. The actual pulse rate required for this type of transmission is then $7 \times 3nf_m = 21nf_m$. This excells straight PNM-AM which would only be able to transmit 21 gradations of amplitude (for an audio frequency equal to f_m) with the same number of pulses. For larger values of q the advantage of PAM-PCM-AM is even greater.

Like PNM-AM the system is essentially noise free above threshold. The threshold is determined by the total number of pulses that must be transmitted just as in PNM-AM. For $q = 7$, $p = 128$ the threshold is the same as for PNM-AM with $f_p/f_m = 21$.

PAM-FCM-AM

The last mentioned system PAM-PCM-AM brings to mind, by analogy, another system not shown in Table 1, which might be called PAM-FCM-AM, the FCM standing for frequency code modulation. It is different from PAM-PCM-AM, only because the possible "q" pulses are transmitted simultaneously (by q separate channels with frequency division using AM-AM) instead of being transmitted consecutively. The separate channels only have to be $9nf_m$ wide to transmit pulses at a maximum rate of $3f_m$. Thus the overall bandwidth is $2\ q\ 9nf_m$ for a

system with n intelligence channels, or 9 q nf_m if the sub-carriers are put on the main carrier single sideband. This is exactly the same bandwidth as that required for PAM-PCM-AM for the same number of channels.

PWM-AM

For this system the average pulse time duration is larger than the minimum possible time duration by a factor K_3. The minimum time duration is $1/B$. The mean time duration a is then equal to K_3/B. At threshold, noise at the beginning of the pulse can increase or decrease the time duration by the interval $1/2B$. The noise at the end of the pulse can do the same. Since the two effects are not correlated, the noise adds up rms. Then the noise can change the pulse time duration by the interval $1/\sqrt{2}\,B$. Modulation can change it by the interval K_3/B. Thus the signal noise ratio at threshold is $\sqrt{2}\,K_3$. It should be noted however that the value of the average signal power at threshold is a function of K_3. The signal noise ratio at a power level corresponding to the carrier equal to twice noise peaks in the AM case is as follows:

$$\text{S/N (PWM-AM)} = \frac{\sqrt{2}\,K_3}{D^{\frac{1}{2}}}\left[\frac{f_p}{f_m}\right]^{\frac{1}{2}} = \frac{\sqrt{2}\,K_3}{(anf_p)^{\frac{1}{2}}}\left[\frac{f_p}{f_m}\right]^{\frac{1}{2}}$$

$$= \sqrt{\frac{2\,K_3\,B}{n\,f_m}} \quad \left(\text{substituting } a = \frac{K_3}{B}\right) \qquad (40)$$

$$\frac{\text{S/N (PWM-AM)}}{\text{S/N (AM)}} = \sqrt{\frac{2K_3B}{n}\,\frac{1}{f_m}\,\frac{1}{2}\left(\frac{2\,f_m}{B}\right)}^{\frac{1}{2}}$$

$$= \sqrt{\frac{2\,K_3}{n}\,\frac{\sqrt{2}}{2}} = \left(\frac{aB}{n}\right)^{\frac{1}{2}} \qquad (41)$$

The maximum value of K_3 (or aB) is one that results in a duty factor of 0.5 in the absence of modulation. The duty factor is,

$$D = a\,nf_p = \frac{K_3}{B}\,nf_p = \frac{1}{2}\ ; \qquad K_{3(\max.)} = \frac{B}{2n\,f_p}, \text{ then}$$

$$\frac{\text{S/N (PWM-AM)}}{\text{S/N (AM)}} = \sqrt{\frac{K_3}{n}} = \sqrt{\frac{B}{2n\,3f_m\,n}} \quad (\text{assuming } f_p = 3f_m)$$

$$= \frac{1}{\sqrt{6}\, n} \left(\frac{B}{f_m} \right)^{1/2} \qquad = \frac{1}{\sqrt{6}\, n^{1}} \left(\frac{B}{n f_m} \right)^{1/2} \qquad (42)$$

The above was written with the assumption that the pulse width could be varied from zero to $1/nf_p$.

Actually the modulation can never change the pulse duration by quite as much as K_3/B. A more accurate figure would be $(K_3 - 1)/B$. If K_3 is large the difference is quite small. If K_3 is not far from unity the error is quite large. In general the value of K_3 will be kept large by using a wide bandwidth as this will result in improved signal noise ratio.

If n is large the signal noise ratio may be improved by the factor $\sqrt{2}$ by varying the pulse width by changing only one side. Keeping one side of the pulse at a fixed position the noise on that side can be eliminated by gating at the receiver. This reduces the noise by the factor $\sqrt{2}$. The signal is unaffected because the width may be varied by the same total amount. A difficulty is encountered if an attempt is made to utilize this single-side-variable system with n less than 4. The peak phase deviation of the variable side is π/n. The phase modulation of the fundamental pulse frequency is one half as much or $\pi/2n$. To avoid the production of second harmonic side bands the peak phase displacement of the fundamental pulse frequency should be kept below 0.5 radians. Then $\pi/2n = .5$, $n = \pi$.

This is the minimum value of n for which the pulse width may be varied over 100 per cent of the available time interval without danger of beat interference produced by second harmonic side bands.

This limitation is removed if both sides of the pulse are varied in position because in this case frequency modulation of the fundamental pulse frequency is not involved.

PWM-FM

This system is one in which the PWM modulating signal of the above system is used to frequency modulate a carrier. In the absence of channel modulation, the resulting wave is a carrier, frequency modulated with a symmetrical square wave of frequency nf_p or $3\, nf_m$. The commutation between channels takes place in the center of the bottom (or top) of the square wave.

If the ratio of peak signal to peak noise in the radio-frequency circuits is unity then when one of the highest noise peaks has a radio-frequency phase angle, 90 degrees from the carrier phase angle, the

phase of the carrier will be modulated by roughly one radian (just as in PPM-FM, page 334). This phase deviation occurs in a time $1/B$, which is the time of rise of the shortest pulse the radio-frequency circuits will pass. This corresponds to a frequency shift of the angle (one radian) divided by the time ($1/B$) divided by 2π or $B/2\pi$. The frequency shift of the signal is B, so the ratio of amplitudes of the detected signal pulses to the noise pulse is 2π.

It should be noted that although peak phase displacement due to the noise pulse occurs in a time $1/B$ from the start of the pulse, peak frequency deviation occurs at half pulse amplitude at time $1/2B$. The curve of frequency deviation is the derivative of the envelope of the radio-frequency noise pulse.

The time of rise of the signal pulse is about the same in the FM case as in the AM case.[1] For the signal noise ratio under discussion however, the noise pulse is $1/2\pi$ times the signal pulse amplitude (instead of 1/2), so the shift is $1/\pi$ times as great as in the PWM-AM case at threshold.

$$\text{However,} \quad \frac{\text{threshold (PWM-AM)}}{\text{threshold (PWM-FM)}} = 4Dn \text{ (power ratio)}$$

where $nD = 1/2 =$ the duty factor.

$$\text{Then} \quad \frac{\text{S/N (PWM-FM)}}{\text{S/N (AM)}} = \frac{\text{S/N (PWM-AM)}}{\text{S/N (AM)}} \times (4nD)^{\frac{1}{2}}\pi$$

$$= \frac{1}{\sqrt{6}\,n}\left(\frac{B}{f_m}\right)^{\frac{1}{2}}\sqrt{2}\,\pi \qquad = \frac{\pi}{\sqrt{3}\,n}\left(\frac{B}{f_m}\right)^{\frac{1}{2}} \tag{43}$$

PAM-FM (Slow)

A scheme has been suggested which does not appear in Table I. This scheme is a time division system in which the sample blocks of signal are below the lowest frequency in the audio spectrum instead of being above the highest audio frequency. The essential idea is to break up the signal of each channel into blocks of a suitable length corresponding to a sub-audible block repetition rate. (Each block will contain many audio cycles.) The blocks are transmitted n times as fast as the signal naturally occurs, thus multiplying frequency components by n. The blocks of the various channels are placed end to

[1] H. Salinger, "Transients in Frequency Modulation", *Proc. I.R.E.*, Vol. 30, No. 8, p. 378, August, 1942.

end in a similar manner to ordinary PAM. Methods of obtaining this type of transmission and the problem of reassembling the channels at the receiver will be discussed briefly. This seems desirable because this system turns out to have the best performance of all, in spite of being little known.

The fundamental idea is to record the signals of n channels simultaneously for a short period of time, and then transmit the signals in sequence at n times normal speed. In the meantime the signals are recorded for a second adjacent short period (no gap). When the n first short periods are completed the second short periods are sent in sequence. Since the speed of transmission is n times normal, the process is continuous and no portion of any signal is omitted. At the receiver the signal for a given channel is selected by a synchronous commutator. For two way telephone service, however, the "short periods" must be quite short (of the order of 0.1 second) and at the receiver a method of slowing down the signal to the normal rate must be found. The various blocks of signal must be fitted together without any appreciable discontinuity.

Using Storage Tubes

A method which has been suggested would make use of special storage tubes at both transmitter and receiver. In the storage tube a target sheet of dielectric material would be scanned by $n + 1$ electron beams. For the transmitter tube one of the beams would scan the target exactly as for a television picture (though not so fast). This beam would take off the signal stored by the other n beams. These n beams would each scan a single line of the target. The intensity of each beam would be modulated by the signal from one of the channels. A pattern of charges corresponding to the signals would thus be placed on the target. The n single-channel put-on beams would move at $1/n^{th}$ the speed of the take-off beam.

At the receiver a similar storage tube would be used except that the fast beam would put the signals on the target, the slow beams taking them off.

Using a Wire Recorder

The system could also be operated with wire recorders replacing the storage tubes. The wire recorder could be arranged as follows: A continuous loop of wire would be arranged over a large wheel and several idler pulleys so as to form a figure 8. For the recorder at the transmitter the n put-on recording heads would be mounted on the large

wheel and spaced at equal intervals around the circumference. The single take-off head would be mounted at some point on the wire loop external to the large wheel. The wire would move continuously at such a rate that a given point would move once around the large wheel in one scanning cycle. The wheel would rotate in the same direction as the wire but at a slightly slower speed so that the length of wire passing a given recording head in one revolution would be $1/n$ times the circumference. When a given recording head completes one revolution it automatically finds itself at a new section of wire one circumference removed from the previous point. It utilizes $1/n^{th}$ of this length for recording the signal of this channel. Thus the wire entering the take-off head has the signals of the n channels recorded in sequence. One scanning cycle is completed in a length of wire equal to the circumference of the wheel. The wire passes the fixed take off head n times as fast as it passes the recording heads. Hence the transmission will be continuous, the n channels being transmitted in sequence each being transmitted at n times the normal rate. The modulation frequencies will all have n times their normal values.

At the receiver a similar tape recorder would be used except that the functions of the put-on and take-off heads would be interchanged.

To avoid a discontinuity between blocks of the reconstructed signal the blocks should overlap up to the point of 50 per cent response. This could be obtained by having the wire entering the multi-head wheel, nearly parallel to the wire leaving it. This makes it necessary to use one or more idler pulleys.

The recorded signals should be wiped off after passing the take off head in the transmitting unit, and immediately preceding the put-on head in the receiving unit. The length of wire in the circuit external to the wheel should be kept to a minimum to avoid time delay.

Obviously the best way to transmit such a signal for best signal noise ratio is as straight FM. The gain over AM is,

$$\frac{\text{S/N (PAM-FM) Slow}}{\text{S/N (AM)}} = \sqrt{3}\,\frac{B}{2nf_m}. \tag{44}$$

Various Hybrids

APM-AM

A hybrid may be defined as a system in which different channels use different systems of modulation. The system in which every other channel amplitude modulates its individual carrier and the remainder

of the channels utilize the same sub-carriers with phase modulation may be called APM-AM. The signal noise ratio is the same as for AM-AM except that $n/2$ is used for n, thus,

$$\frac{\text{S/N (APM-AM)}}{\text{S/N (AM)}} = \frac{\sqrt{2}}{2n} \text{ for linear summation.} \tag{45}$$

$$\frac{\text{S/N (APM-AM)}}{\text{S/N (AM)}} \frac{1}{4(n/2)^{\frac{1}{2}}} \text{ for root-mean-square summation.} \tag{46}$$

This represents a net gain of 2 to 1 (or $\sqrt{2}$ for root-mean-square summation), but the system is not very practical because of the critical balance required to eliminate cross talk.

Single sideband does not seem to hybridize with the other types of frequency division systems.

Time Division Hybrids

In time division multiplexing a system is possible in which the first channel is PAM, the second PPM and the third PWM all on the same set of pulses. By time division other channels are placed on other pulses in a similar manner. This might be called PAPWM-AM. The poorer channels will have the same signal noise ratio as a PAM-AM system except that $n/3$ may be used for n, thus,

$$\frac{\text{S/N (PAPWM-AM)}}{\text{S/N (AM)}} = \sqrt{\frac{3}{n}} \frac{\sqrt{2}}{4}. \tag{47}$$

Before the above can be relied on, a check is necessary to see if the signal noise ratio of any one of the three types of channels is harmed by the presence of the other types of modulation.

For the PAM channels the value of K $(= aB)$ must be larger than the ideal minimum value because of the use of PWM on the same pulses. Fortunately the signal noise ratio for PAM-AM is independent of the value of K. However, the degree of PPM must be reduced to allow room for PWM and the duration of the pulse must be less than $1/2nf_r$ to allow room for PPM. However, a 2 to 1 sacrifice for both PWM and PPM would leave both better than PAM which is the limiting factor. Because of PAM the system is not competitive. If either PWM or PPM is omitted the system remains not competitive because of PAM.

PPWM-AM

If the PAM is omitted the system might be called PPWM-AM. Since approximately a two to one sacrifice is necessary in the degree of modulation of both types of modulation to allow room for the other, the hybrid has no advantage. In any of the possible hybrids it is very difficult to eliminate cross modulation.

Triple Modulation Systems

AM-AM-AM

Where the carriers of several AM-AM systems with spaced carrier frequencies, are considered to be sub-carriers and used to modulate the main carrier, the resulting system is called AM-AM-AM. Going from AM (with 1 channel) to AM-AM with n channels impairs the signal noise ratio by the factor,

$$\frac{\text{S/N (AM-AM)}}{\text{S/N (AM)}} = \frac{\sqrt{2}}{4n} \text{ for linear summation} \tag{2}$$

$$\text{or } 1/4n^{\frac{1}{2}} \text{ for root-mean-square summation.} \tag{3}$$

In an exactly similar fashion going from AM-AM with n_1 channels to AM-AM-AM with $n_1 \times n_2$ channels impairs the signal noise ratio by the factor,

$$\frac{\text{S/N (AM-AM-AM)}}{\text{S/N (AM-AM)}} = \frac{\sqrt{2}}{4n_2} \text{ for linear summation.} \tag{48}$$

$$\text{So } \frac{\text{S/N (AM-AM-AM)}}{\text{S/N (AM)}} \frac{\sqrt{2}}{4n_2} \times \frac{\sqrt{2}}{4n_1} = \frac{1}{8(n_1 n_2)}$$

$$= 1/8n \text{ for linear summation.} \tag{49}$$

The net loss is the factor $2\sqrt{2}$ regardless of the ratio of n_1 to n_2. Similarly for large values of n,

$$\frac{\text{S/N (AM-AM-AM)}}{\text{S/N (AM)}} = \frac{1}{4n_1^{\frac{1}{2}}} \frac{1}{4n_2^{\frac{1}{2}}} = \frac{1}{16n^{\frac{1}{2}}} \text{ for}$$

$$\text{root-mean-square summation.} \tag{50}$$

The net loss is 4 to 1 regardless of the ratio of n_1 to n_2. Of course the method only applies when n_1 and n_2 are both large.

XX-XX-AM

Adding one or more sub-carriers in the above manner to any cw system has a similar effect reducing the signal noise ratio by the factor $2\sqrt{2}$ for linear summation. It should be noted, however, that the bandwidth is slightly more than twice as great.

XX-XX-PM

Adding one or more sub-carriers as PM on the main carrier the signal noise ratio is multiplied by a factor $\frac{1}{\sqrt{2}}\frac{f_d}{f_{sc}}$, where f_{sc} is the frequency of the sub-carrier and f_d is the mean peak frequency deviation produced by the sub-carrier. Note, however, that a wider band is required than in the XX-XX modulating system.

SS-PM-PM (with a single intermediate sub-carrier)

According to the above the signal noise ratio compared to AM is,

$$\frac{\text{S/N (SS-PM-PM)}}{\text{S/N (AM)}} = \frac{B_{sc}}{f_c n^2} \times \frac{1}{\sqrt{2}} \frac{f_d}{f_{sc}} \text{ (for linear summation)}.$$

Where B_{sc} refers to the bandwidth required for the single intermediate subcarrier and its complex modulation, and is not the final radio-frequency bandwidth B of earlier formulas. The overall bandwidth B is of course much wider. To get the formula in terms of the new bandwidth consider the following:

At first sight it would appear that B_{sc} could be made equal to $2f_{sc}$ and $f_d = B/2$. If this could be done the net loss caused by using the intermediate sub-carrier would be the factor $\sqrt{2}$. Actually B_{sc} cannot be as large as $2\,f_{sc}$. The instantaneous frequency must be kept several times larger than the highest modulating. The exact limiting ratio is not known but 4 seems about right. Also f_d cannot be as large as B/2. In a practical case,

$nf_m =$ 150 kilocycles $\quad f_{sc} =$ 1000 kilocycles $\quad f_d =$ 1400 kilocycles

f_{sc} (min. mod. value) $=$ 600 kilocycles

$B_{sc} =$ 800 kilocycles $\quad B =$ 4000 kilocycles

This leads to the value,

$$\frac{\text{S/N (SS-PM-PM)}}{\text{S/N (AM)}} = \frac{B}{f_c n^2} \frac{1}{\sqrt{2}} \frac{800}{1000} \frac{1400}{4000} = \frac{B}{f_c n^2} \times \frac{1}{5}. \quad (51)$$

Where the B now refers to the total radio-frequency bandwidth

$$\frac{\text{S/N (SS-PM-PM)}}{\text{S/N (AM)}} = \frac{\sqrt{3}}{10} \frac{B}{f_c n^{3/2}} \text{ for root-mean-square summation.} \quad (52)$$

Thus for either the root-mean-square summation or linear summation equations, the use of the intermediate sub-carrier involves a sacrifice in signal noise ratio of something like 5 to 1 over what could be obtained if the same pass band were fully utilized with an SS-PM system. This figure would vary somewhat with changes in n, f_m and B.

SS-PM-FM

If the final modulation of the above system is applied as FM instead of PM it becomes SS-PM-FM which is the system used in one type of microwave radio relay equipment.[5] An accurate analysis of this system is more complicated than for SS-PM-PM because of a lack of symmetry in the side bands. However, the final result is only slightly different from that given above. An advantage is that FM is much easier to produce and to demodulate than PM. The engineers also feel from practical experience that there is value in the fact that this system concentrates the sideband energy nearer the main carrier. The higher frequencies of the final modulating voltage corresponding to complex modulation sidebands produce a lower modulation index with FM than with PM (assuming the same modulation index for the sub-carrier frequency). Many of these higher sidebands may be cut off by the receiver selectivity without doing any harm since their mates on the other side of the sub-carrier produce a much higher modulation index. Engineers familiar with this equipment report that this factor makes SS-PM-FM better than SS-PM-FM by about the ratio 10 to 7. For the purpose of this paper the same formula as SS-PM-PM will be used.

XX-PAM-AM

From the data on PAM-AM, it can be seen that taking a given modulating signal and applying it to a PAM-AM system, the signal noise ratio is the same as though an AM system were used. If the modulating system is a group of AM channels the resulting system is AM-PAM-AM the signal noise ratio is the same as for AM-AM. If the modulating signal is a group of XX channels the resulting system is XX-PAM-AM and has the same signal noise ratio as XX-AM (XX denotes any type of modulation desired).

[5] G. G. Gerlach, "A Microwave Relay Communication System", *RCA REVIEW*, Vol. VII, No. 4, pp. 576-600, December, 1946.

This is somewhat misleading because the bandwidth required is much greater. The pulse frequency must be at least 2.5 times the highest modulating frequency and the bandwidth required to pass the pulses is about 3 times the pulse frequency. Thus the bandwidth is increased by a factor of 7.5 to 1. If reduced to the original bandwidth by modifying the modulating signal, a signal noise ratio sacrifice will be entailed but the value will differ for the different systems.

SS-PPM-AM

This system is one in which a group of n_1 channels of SS is applied as modulation to each of the n_2 channels of a PPM-AM system. Then the ratio

$$\frac{\text{S/N (SS-PPM-AM)}}{\text{S/N (SS-AM)}} \text{ corresponds to } \frac{\text{S/N (PPM-AM)}}{\text{S/N (AM)}} \text{ with } n = n_2 \text{ and } f_m = nf_c.$$

$$\text{Then, } \frac{\text{S/N (SS-PPM-AM)}}{\text{S/N (AM)}} = \frac{\text{S/N (SS-PPM-AM)}}{\text{S/N (SS-AM)}} \times \frac{\text{S/N (SS-AM)}}{\text{S/N (AM)}}$$

$$= \frac{2}{9} \frac{B}{n_2^{3/2}(n_1 f_c)} \times \frac{1}{n_1^{\frac{1}{2}}} = \frac{2}{9} \frac{B}{(n_1 n_2)^{3/2} f_c}. \tag{53}$$

$$\text{Thus, } \frac{\text{S/N (SS-PPM-AM)}}{\text{S/N (PPM-AM)}} = y. \tag{54}$$

Similarly, for any system having $n^{-\frac{1}{2}}$ in the "alternate form" of the equation for R_n (page 350), SS may be added out front (in the system designation) with no change in the formula except for the additional factor y. An exception is (FM-XX). This also works if the FM is changed to PM so as to equalize the S/N in all channels.

SS-PAM±FM

For this system a number n_1 of SS systems are used to modulate each channel of an n_2 channel PAM ± FM system giving a system with $n = n_1 n_2$ channels. Then

$$\frac{\text{S/N (SS-PAM±FM)}}{\text{S/N (AM)}} = \frac{\text{S/N (SS-PAM±FM)}}{\text{S/N (SS-AM)}} \times \frac{\text{S/N (SS-AM)}}{\text{S/N (AM)}}$$

$$= \frac{\pi}{12} \times \frac{B}{n_2^{3/2}\,(n_1 f_c)} \times \frac{1}{n_1^{\frac{1}{2}}} = \frac{\pi}{12} \frac{B}{n_2^{3/2} n_1^{3/2} f_c} = \frac{\pi}{12} \frac{B}{n^{3/2} n_1 f_c} \qquad (55)$$

Thus $$\frac{\text{S/N (SS-PAM} \pm \text{FM)}}{\text{S/N (PAM} \pm \text{FM)}} = y$$

FM-PPM-AM

Let a number n_1, of FM systems with spaced sub-carriers (as in FM-AM) be used to modulate each of the n_2 channels of a PPM-AM system. Then,

$$\frac{\text{S/N (FM-PPM-AM)}}{\text{S/N (AM)}} = \frac{\text{S/N (FM-PPM-AM)}}{\text{S/N (FM-AM)}} \times \frac{\text{S/N (FM-AM)}}{\text{S/N (AM)}}$$

$$= \frac{2}{9} \frac{B}{n_2^{3/2} B_0} \times \frac{\sqrt{3}}{16} \frac{B_0}{n_1^{3/2} f_m}$$

where B_0 is the bandwidth of each FM-AM system

$$\frac{\text{S/N (FM-PPM-AM)}}{\text{S/N (AM)}} = \frac{2}{9} \frac{B}{n_2^{3/2}} \frac{\sqrt{3}}{16 n_1^{3/2} f_m} = \frac{\sqrt{3}}{72} \frac{B}{n^{3/2} f_m} \qquad (56)$$

The net result of placing the FM out front (in the system designation) is that signal noise ratio is poorer by the factor $\sqrt{3}/16$. The above is also true for any other system having the same algebraic factors in the formula. (Except (FM-XX) which must be changed to (PM-XX) to keep the signal noise ratio the same in all channels.)

XX-XX-PM

Calculations similar to the above show that adding PM on the end (in the system designation) of a double modulation system in general reduces the signal noise ratio by a factor of the same order of magnitude as the above. AM-AM is of course an exception as it can be improved by making it AM-AM-PM. This is not as good as AM-PM however. In general no triple modulation system ending in PM is as good as a similar double modulation system. SS-PM-PM (with n_2 large enough to use root-mean-square summation) is nearly as good as PM-PM, but not as good as SS-PM.

Thus, in Table I, no system in columns 2 or 3 is as good as a similar system in column 1.

Columns 4 and 5 (of Table I)

Of the systems in columns 4 and 5 only SS-PPM-AM, SS-PAM±FM, and SS-PPM-FM have merit each being nearly equal to the corresponding system omitting the SS.

AM-PXX-XX

None of these four systems is as good as the corresponding PXX-XX system

FM-PXX-XX

The corresponding PXX-XX system is better in each case except for FM-PAM-AM which is not as good as FM-AM.

SS-PXX-XX

This system is covered on page 346.

PAM-PXX-XX

Each of these four systems has a poorer signal noise ratio than the corresponding PXX-XX system.

PPM-PXX-XX

These four systems are completely impractical because a continuous variation of the phase of the PPM pulses cannot be represented by the small available number of shorter PXX pulses per pulse. The phase would only be variable in steps as each shorter pulse came into or out of operation.

PWM-PXX-XX

These four systems are also completely inoperative. The width modulation cannot be continuously variable when used to modulate other shorter pulses with a limited number of pulses per pulse.

PNM-PXX-XX

These four systems are operative but each has a higher threshold than the corresponding PNM-XX system because of the wider band required. Above threshold, signal noise ratio $= \propto$ just as in PNM-XX.

Summary of Triple Modulation Systems

All triple modulation systems are inferior to a related double modulation system with the exception of SS-XX-XX which is nearly equivalent to XX-XX in many cases.

Gain in Signal Noise Ratio by the Use of Pre-emphasis

The signal noise ratio of all systems can be improved by the use of pre-emphasis in the modulator circuit of the transmitter with a corresponding de-emphasis circuit in the receiver. The gain in an AM system is called G_a and in an FM system G_f. A PM system has the same factor as an AM system. Thus the gain is G_f for FM, FM-AM, FM-PM, PAM(±)FM(slow). The gain is G_a for AM, AM-AM, AM-PM, SS-AM, SS-PM, PAM-AM, PPM-AM, PWM-AM, PWM-FM. In no case can the gain of pre-emphasis be obtained twice. In applying the several sub-carriers (of the frequency division systems) as modulation of a main carrier, the modulation must be carried out so that the same signal noise ratio is obtained on each channel. For XX-PM systems this might be looked at as XX-FM with pre-emphasis but regardless of this, the matter is already taken care of in the formulas.

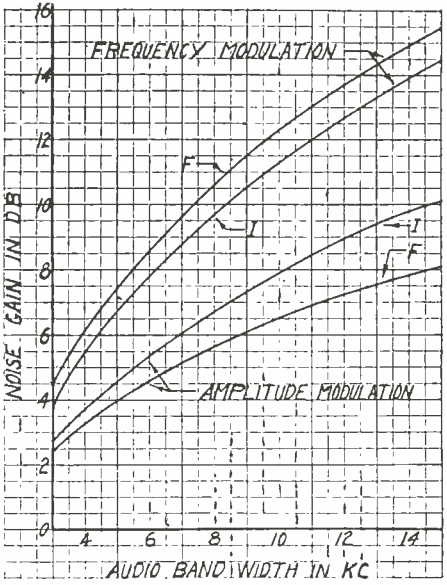


Fig. 12 — Calculated noise gains effected by the de-emphasis noise.

The gain due to pre-emphasis is plotted as a function of audio bandwidth for both AM and FM in Figure 12.[6]

It can be seen from the figure that for $f_m = 3500$ cycles/second,

$G_a = 2.8$ decibels or a voltage ratio of 1.38

$G_f = 5.4$ decibels or a voltage ratio of 1.86

Summary of Signal Noise Ratios on Random Noise:

The signal noise ratio formulas are summarized in Table 3. All the frequency division formulas are presented in two forms, for linear summation and for root-mean-square summation. Assuming a large number of channels is desired, the root-mean-square summation should be used. For all time division systems the formula is the same whether n is large or small.

[6] M. G. Crosby, "The Service Range of Frequency Modulation", *RCA REVIEW*, Vol. IV, No. 3, p. 362, January, 1940 (Figure 6).

Table III—Formulas for System Signal Noise Ratios
(Compared to Single Channel AM as a Standard)

Formulas for R_o

System	Linear summation	Root-mean-square summation	Alternate form for root-mean-square summation	Additional Factor Obtainable from Pre-emphasis if $f_m = 3500$
AM	1			
FM	$2^{-1}\sqrt{3}\ f_m^{-1} B$			1.86
AM-AM	$\sqrt{2}\ 4^{-1} n^{-1}$	$4^{-1} n^{-\frac{1}{2}}$	$0.25\ n^{-\frac{1}{2}}$	1.38
AM-PM	$\sqrt{2}\ 8^{-1} Bn^{-2} f_c^{-1}$	$3^{\frac{1}{2}}\ 16^{-1} Bn^{-3/2} f_c^{-1}$	$0.108\ n^{-\frac{1}{2}} ry$	1.38
FM-AM	$\sqrt{6}\ 8^{-1} Bn^{-2} f_m^{-1}$	$3^{\frac{1}{2}}\ 16^{-1} Bn^{-3/2} f_m^{-1}$	$0.108\ n^{-\frac{1}{2}} r$	1.86
FM-PM	$3^{\frac{1}{2}}\ 2^{3/2} Bn^{-2} f_m^{-1}$	$3 \times 16^{-1} Bn^{-3/2} f_m^{-1}$	$0.188\ n^{-\frac{1}{2}} r$	1.86
SS-AM	n^{-1}	$n^{-\frac{1}{2}}$	$n^{-\frac{1}{2}}$	1.38
SS-PM	$Bn^{-2} f_c^{-1}$	$3^{\frac{1}{2}}\ 2^{-1} Bn^{-3/2} f_c^{-1}$	$0.866\ n^{-\frac{1}{2}} ry$	1.38
SS-PM-FM	$5^{-1} Bn^{-2} f_c^{-1}$	$5^{-1}\ 3^{\frac{1}{2}}\ 2^{-1} Bn^{-3/2} f_c^{-1}$	$0.173\ n^{-\frac{1}{2}} ry$	1.38
PAM-AM		$n^{-\frac{1}{2}}$	$n^{-\frac{1}{2}}$	1.38
PAM(±)-FM		$\pi \cdot 12^{-1} Bn^{-3/2} f_m^{-1}$	$0.262\ n^{-\frac{1}{2}} r$	1.86
PPM-AM		$2 \times 9^{-1} Bn^{-3/2} f_m^{-1}$	$0.22\ n^{-\frac{1}{2}} r$	1.38
PPM-FM		$2 \times 3^{-1} B^{\frac{1}{2}} f_m^{-\frac{1}{2}} n^{-1}$	$0.66\ n^{-\frac{1}{2}} r^{\frac{1}{2}}$	1.38
PNM-AM		∞		
PNM-FM		∞		
PAM-PCM-AM		∞		
PAM-FCM-AM		∞		
PWM-AM		$6^{-\frac{1}{2}} B^{\frac{1}{2}} n^{-1} f_m^{-\frac{1}{2}}$	$0.41\ n^{-\frac{1}{2}} r^{\frac{1}{2}}$	1.38
PWM-FM		$\pi\ 3^{-\frac{1}{2}} B^{\frac{1}{2}} n^{-1} f_m^{-\frac{1}{2}}$	$1.8\ n^{-\frac{1}{2}} r^{\frac{1}{2}}$	1.38
PAM-FM(Slow)		$3^{\frac{1}{2}}\ 2^{-1} Bn^{-1} f_m^{-1}$	$0.866\ r$	1.86

Note: $r = Bn^{-1} f_m^{-1}$, $y = f_m f_c^{-1}$

Even in the simplified form shown in the column on the right it is not immediately apparent which system is best because the values of n, r and y are not specified. However, n and r are always unity or larger. The value of y is usually only slightly below unity. The values of n and r may range from unity to fairly large numbers. Keeping this in mind and examining the formulas for the frequency division systems, it can be seen that SS-PM, FM-PM, FM-AM and AM-PM all have the

same formula $n^{-\frac{1}{2}}r$ except for a numerical coefficient and y. (Note that $y = f_m/f_r$ and $r = B\,n^{-1}\,f_m{}^{-1}$) The above sequence of systems is arranged in the order of merit.

Among the time division systems the systems PAM±FM and PPM-AM again have the formula $n^{-\frac{1}{2}}\,r$ except for the numerical coefficient. These two systems are very nearly equal.

Each of these systems is about two to one below SS-PM in signal noise ratio. This ratio is inverted if triple modulation is used in the frequency division case (SS-PM-FM).

The systems PWM-FM, PPM-FM and PWM-AM all have the formula $n^{-\frac{1}{2}}\,r^{\frac{1}{2}}$ except for the numerical coefficient. They are arranged above in the order of merit. The change in algebraic formula indicates less advantage from extra bandwidth. The formulas indicate a high position for these three systems for low values of r. Thus PWM-FM is better than PPM-AM if r is less than 67, but it must be kept in mind that PWM-FM is inoperative if r is less than 24.

The system PAM-FM(slow) is indicated as being best of all, being particularly meritorious for large values of n, since n is missing from the formula (except insofar as it is inherent in r).

The systems AM-AM, SS-AM and PAM-AM are the poorest in the list because of the absence of r from the formula.

The systems PNM-XX and PAM-PCM-AM are theoretically completely insensitive to random noise when the received signal is above threshold. Unfortunately the characteristics of the systems are such that a new type of noise is generated in the system itself independent of thermal noise. The magnitude of this noise is outside the scope of this paper.

Because of threshold effects and because the values of r and n may vary over wide limits in different specifications. Table 3 is not completely satisfactory as a means of comparing the systems for noise performance. More detailed comparisons are developed in subsequent sections.

(ED. NOTE: It is planned to publish the remainder of this paper in the September, 1948 issue of *RCA REVIEW*.)

FREQUENCY-MODULATED RADAR TECHNIQUES*†

By

Irving Wolff and D. G. C. Luck

Research Department, RCA Laboratories Division,
Princeton, N. J.

NOTE: The previous paper in this series, entitled PRINCIPLES OF FREQUENCY MODULATED RADAR, was published in *RCA REVIEW* in March, 1948. The summary of this first paper is reprinted herewith for reference purposes. A limited number of copies of the March 1948 issue is still available for those who desire a complete file on the FM radar series. It is planned that the third paper, entitled SOME APPLICATIONS OF FREQUENCY-MODULATED RADAR, will be published in the September 1948 issue of *RCA REVIEW*.

PRINCIPLES OF FREQUENCY-MODULATED RADAR

(Reprinted From RCA REVIEW, March, 1948)

Summary—*The principles of operation of FM radar systems are developed for the determination of range and relative speed of reflecting objects. Quantitative expressions are derived relating the radar output frequency to the range and speed being measured and the transmitted-signal characteristics. It is shown that range and speed may be independently determined. In single-target systems, measurement of output frequency by means of cycle counters has been frequently used. This leads to a type of error which may appear in laboratory calibration of the equipment but which is usually averaged out in field operation. Search operation against more than one target requires either a multiplicity of selective range gates or scanning in range with a single gate. The former system leads to rather complicated apparatus and the latter to unduly slow operation. For the same average transmitted power, pulse and FM radar are capable of similar maximum ranges; however, the simple FM system which scans in range with a single gate uses time inefficiently and either requires longer than the pulse system to obtain data or operates at reduced maximum range. In general, operating with the same average power both FM and pulse systems have their maximum ranges determined by the time allowed to obtain the complete range data and their range resolution by the usable radio-frequency band. FM radar avoids need of high peak power by using a narrow noise band, and facilitates use of a wide radio-frequency band for high resolution, but requires great care in protecting reception against its own transmission. Techniques for automatic control on the basis of range and speed, described in later papers, are very simple, but equipment yet devised for search is either complex or slow in action.*

* * *

There follows the second paper in this series—FREQUENCY-MODULATED RADAR TECHNIQUES.

The Manager, RCA REVIEW

---

* Decimal classification: R537 × R148.2

† This paper covers work initiated in 1938 by RCA Laboratories Division and carried on since 1941 with the support of the U. S. Navy, under Contracts NOs-87822, NXsa-25337, and NXsa-35042.

Summary—The FM radar development program has given rise to a number of interesting techniques. Two superheterodyne systems have been devised which remove the transmitted modulation from the received intermediate-frequency signal. An electromechanical device using a vibrating capacitor and an electronic device using a beam of electrons have proved useful for frequency modulating radar oscillators. Several adaptations of the well-known cycle-rate counter for utilization of beat frequencies as data have by their simplicity made FM radar very useful, where automatic indication of or control by range and speed of a single target is needed.

I. Introduction

RADAR measurement of the range and speed of a reflecting target can be accomplished by using frequency-modulated radio signals. The principles of FM radar and some of their consequences have been discussed in a previous paper.[1] Practical application of these principles has required the development of some novel techniques and has resulted in a number of complete radar systems for special uses. Brief descriptions of these techniques are given here, while the complete systems will be taken up in a later paper.

Techniques for the production and use of radio-frequency signals at very-high and ultra-high frequencies were everywhere developing rapidly during the work on FM radar. Developments along such lines made in the course of this work paralleled those elsewhere, so need not be considered here. Normal triode oscillators are used as transmitters in FM radar equipment operating at 440 and 1500 megacycles and magnetrons are used at 4000 megacycles.

Normal antennas, ranging from a simple streamlined dipole for 440-megacycle altimeters and the dragless "slot antenna"[2] developed to replace it, to dipole-excited paraboloidal reflectors for 4000 megacycles, are used on aircraft in locations dictated by the specific operating conditions to be met. The necessity for protecting a continuously sensitive receiver from the direct output of a transmitter operating continuously on the same channel gives to the antenna problem for FM radar its only novel aspect. The solution generally has been the use of separate antennas for transmission and reception, placing them so as to minimize coupling or *feed through* between them. The necessity of avoiding loose contacts in portions of the aircraft exposed to both antennas intensifies the problem of reducing microphonics that is encountered with all sensitive airborne equipment. Crossed-polarization

---

[1] I. Wolff and D. G. C. Luck, "Principles of FM Radar," *RCA REVIEW*, Vol. IX, No. 1, pp. 50-75, March, 1948.

[2] N. E. Lindenblad, "Slot Antennas," *Proc. I.R.E.*, Vol. 35, No. 12, pp. 1472-1479, December, 1947.

duplexing tests at 4000 megacycles have shown promise of permitting single-antenna operation.

Receiving equipment, ranging from a simple balanced mixer for use at very-high frequencies to two specialized superheterodyne systems, has shown more novelty. Novel techniques and devices for modulating transmitter frequency have also resulted from the FM radar work, as have special methods of utilizing the information (Section II(C) of Reference 1) contained in the beat note developed by mixing transmitted and received signals. These developments are described in the following sections.

II. Receivers

Early FM altimeters,[3,4] as well as the FuG-101 altimeter used by Germany during the war, made use of a simple detector to develop a

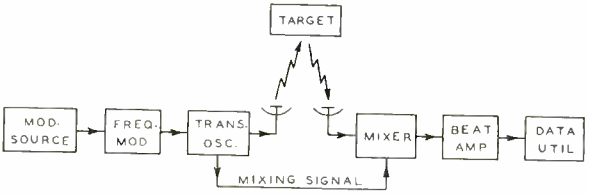


Fig. 1—Basic FM radar system.

beat-note output, by combination of the target-reflected received signal with the direct feed-through signal reaching the receiving antenna more or less fortuitously from the transmitting antenna. Throughout the work considered here, however, it has been found preferable to minimize fortuitous feed through and to rely instead on an intentional mixing-signal coupling path provided within the equipment. Figure 1 is a block diagram of the basic FM radar system, as used in some complete equipments. Use of a doubly balanced mixer or detector is quite valuable in reducing the disturbing effect of modulated interfering signals, as well as that of unintentional amplitude modulation which is always present in the transmitter.

Freedom from disturbance by amplitude modulation can also be had, at the expense of simplicity but without the criticalness of the doubly balanced mixer, by using a superheterodyne receiver. The wide swings of transmitted frequency necessary to provide high resolution in range (Section IV(B) of Reference 1) do not in themselves give useful

[3] S. Matsuo, "A Direct-Reading Radio-Wave Reflection-Type Absolute Altimeter for Aeronautics," *Proc. I.R.E.*, Vol. 26, No. 7, pp. 848-858, July, 1938.

[4] L. Espenschied and R. C. Newhouse, "A Terrain Clearance Indicator," *Bell. Sys. Tech. Jour.*, Vol. 18, No. 1, pp. 222-234, January, 1939.

information; they should therefore be eliminated from the intermediate-frequency portion of the superheterodyne to improve its design. In other words, the local oscillator of the receiver should track the modulation of the transmitter frequency, a feature which will also prevent continuous reception of fixed-frequency interfering signals. Two methods of accomplishing this have been developed to the point of satisfactory operation, giving rise to receiving systems that may be called respectively a side-band superheterodyne and a signal-following superheterodyne. Both require that the transmitted signal shall be directly available at the receiver, separately from the signal to be received.

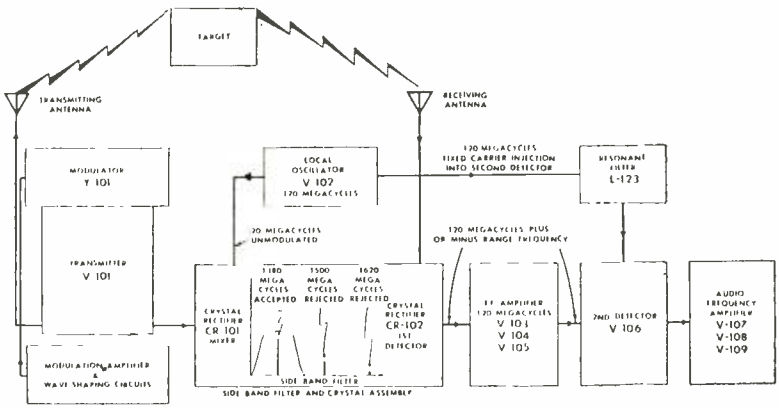


Fig. 2—Side-band superheterodyne system.

Figure 2 is a block diagram illustrating the operation of a side-band superheterodyne. The "local oscillator" signal in the usual superheterodyne sense is in this case a side band developed by modulating a portion of the transmitter-output signal, the original transmitter signal and the unwanted side band being removed by a band-pass filter. The actual local oscillator, which modulates a signal bled from the transmitter, operates at the same fixed intermediate frequency as does the intermediate-frequency amplifier. The useful intermediate-frequency signal finally developed is displaced in frequency from the local oscillator by the same range-beat and speed-beat frequencies that characterize the displacement of the received radar signal from the transmitted signal. By mixing the oscillator signal with the intermediate-frequency signal in a final detector, the desired range and speed-beat frequencies are recovered for utilization as data.*

* A side-band superheterodyne has also been used in continuous-wave radar equipment developed elsewhere.

Figure 3 is a block diagram illustrating the operation of a signal-following superheterodyne. Here, there is an actual local oscillator in the usual sense, and its frequency is automatically controlled so as to maintain the beat between it and the transmitter always approximately centered on the fixed frequency characteristic of a discriminator in an auxiliary or control intermediate-frequency channel. The beat between local oscillator and received signal, appearing in the main or signal intermediate-frequency channel, departs slightly from the nominal center of the intermediate-frequency band both because of the desired

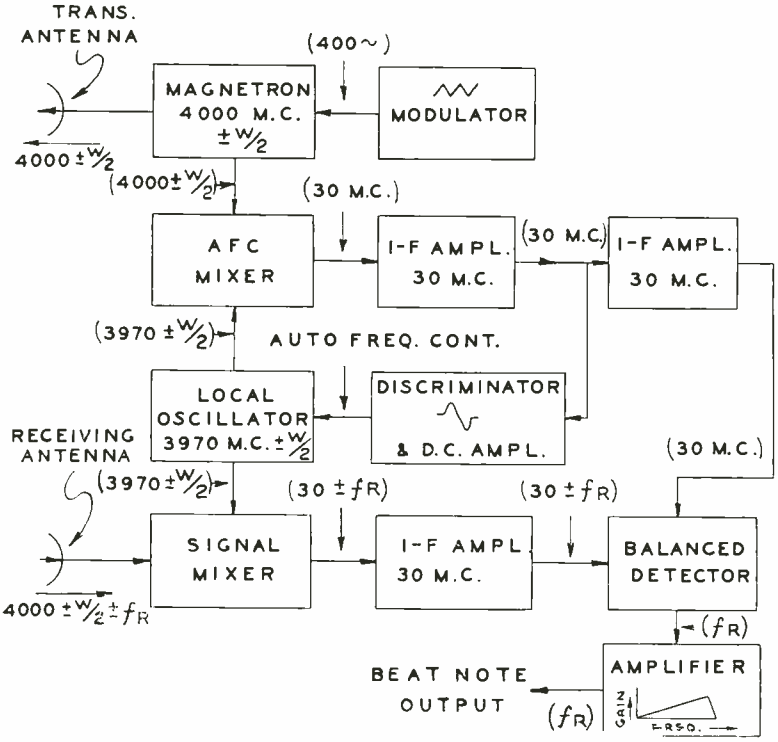


Fig. 3—Signal-following superheterodyne system.

radar-beat frequencies and because of imperfect frequency control of the local oscillator. By mixing the control-channel and signal-channel outputs in a final detector, the true radar beat frequencies are recovered for utilization, independently of the degree of exactness with which the local-oscillator frequency follows the modulation of the transmitter.

III. Modulation

Frequency modulation of signals for radar use has always been

accomplished in the most direct way, by varying a reactance in the circuit of an oscillator; the variable-reactance element may be either electromechanical or purely electronic. Rotary variable capacitors have been used successfully.[4,5] These require careful design and precise construction; they must run at an accurately controlled speed in order to maintain the desired rate of change of transmitted frequency, and thereby the range-sensitivity calibration of the radar (Section III(A) of Reference 1). The rotary-capacitor modulator is rather inflexible when the range sensitivity of the radar must be varied freely during operation.

The modulator used in the systems later described is a capacitor with one electrode fixed and the other vibrated by a dynamic loud-

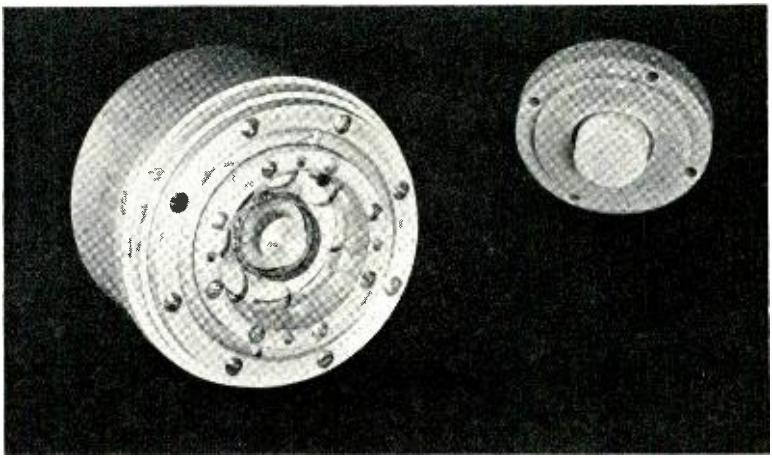


Fig. 4—Vibrating-capacitor frequency modulator.

speaker mechanism. Such a modulator is seen partially disassembled in Figure 4. The fixed, insulated electrode of the right-hand piece engages the voice-coil sleeve in the main structure at the left, to form a cylindrical variable capacitor when fully assembled. In order to produce a triangular or symmetrical-sawtooth variation with time of capacitance and thereby (for small swings) of oscillator frequency, the driving force on the moving electrode and therefore the voice-coil current must vary as indicated in Figure 5. The triangular component at (a) is required to overcome suspension stiffness, the square-wave component at (b) (negligible in practice) to overcome viscous or frictional resistance, and the impulsive component at (c) to overcome

[5] P. Giroud and L. Couillard, "Sondeur Radioélectrique pour la Mésure des Hauteurs des Aéronefs au-dessus du Sol", *Ann. Rad. Elect.*, Vol. 2, No. 8, pp. 150-172, April, 1947.

inertia and reverse the motion of the coil and electrode suddenly at the ends of its travel. The over-all current wave required therefore has the form of Figure 5(d).

Modulator drive of the required form is easily obtained from a square-wave signal source, by integrating the square wave electrically to develop a triangular component, differentiating to develop a pulse component, and mixing these components in suitable proportions. This method of obtaining the modulator-driving signal produces a triangular-wave output with amplitude reciprocally related to frequency. The product of modulation frequency and width of transmitter-frequency band swept, and therefore the radar range sensitivity (Section

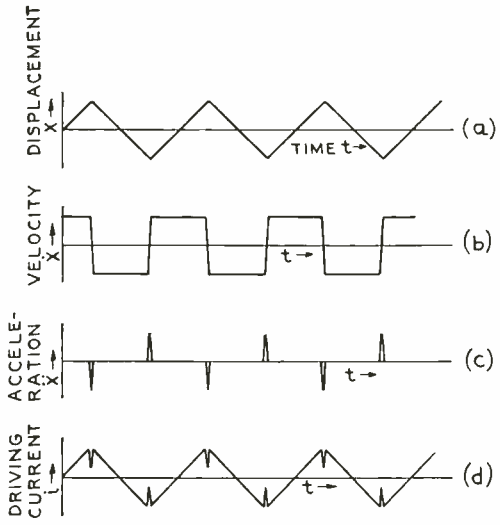


Fig. 5—Driving force on vibrating capacitor.

III(A) of Reference 1), is thus automatically compensated for variations in the square-wave frequency. Radar range sensitivity is easily controlled by adjustment of square-wave amplitude.

Magnetrons have been frequency modulated by varying their operating anode current, but it is difficult to obtain in this way a sufficiently linear frequency sweep and sufficiently small amplitude modulation. It is possible to modulate somewhat the frequency of a magnetron oscillator by varying either the tuning or the losses of one or more external resonant circuits coupled to it, provided that this coupling either is very tight or is separated from the normal load coupling; practical frequency modulators can be made in this way. Magnetron oscillators have also been frequency modulated quite successfully by auxiliary beams of electrons introduced directly into their anode cavities. Unlike mechanical methods of varying circuit reactance to

modulate oscillator frequency, this electronic method permits extremely rapid modulation.[6,7,8]

IV. Utilization of Radar Data

In the FM radar systems for use against single targets that have been developed to a serviceable stage, the beat-note frequency data from the radar proper must be so converted as to produce a dial indication or control action. Such conversion might be accomplished by means of frequency-selective circuits, but another method seemed initially more promising and the urgency of war-time development did not permit much exploration of alternatives.

The device actually applied in several forms is the cycle-rate counter, now widely used in commercial audio-frequency meters. In this device a fixed capacitor is repeatedly charged and discharged, receiving an equal charge on each cycle of the frequency to be determined, and

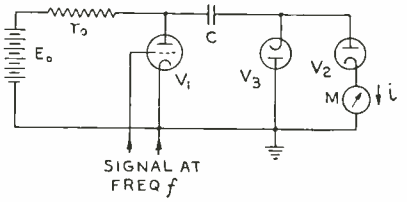


Fig. 6—Basic cycle-rate counter.

the average charging current is measured. Figure 6 is the basic circuit of such a cycle-rate counter, with charge and discharge of capacitor C controlled by an electronic switch or *limiter* V_1 (actually a pentode, though a triode is shown for simplicity), and with charge and discharge current paths segregated by diodes V_2 and V_3. The current through meter M is simply fCE_0. With limiter tubes of convenient size, the current developed by this counter is inconveniently small. Replacement of the meter by a load resistor to permit the use of an amplifier leads, however, to a non-linear characteristic. Two special modifications to avoid these difficulties have proved valuable additions to FM radar technique.

A non-linear relation of output voltage to input frequency occurs, if the meter M of Figure 6 is replaced by a load resistor paralleled with

---

[6] L. P. Smith and C. Shulman, "Frequency Modulation and Control by Electron Beams," *Proc. I.R.E.*, Vol. 35, No. 7, pp. 644-657, July, 1947.

[7] G. R. Kilgore, C. I. Shulman and J. Kurshan, "A Frequency-Modulated Magnetron for Super-High Frequencies," *Proc. I.R.E.*, Vol. 35, No. 7, pp. 657-664, July, 1947.

[8] J. S. Donal, Jr., R. R. Bush and H. R. Hegbar, "A 1-Kilowatt Frequency-Modulated Magnetron for 900 Megacycles," *Proc. I.R.E.*, Vol. 35, No. 7, pp. 664-669, July, 1947.

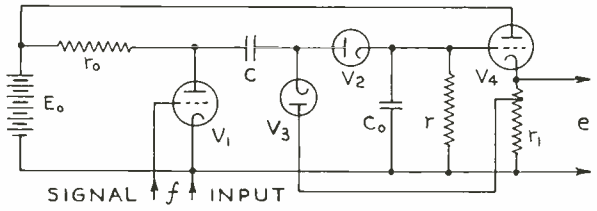


Fig. 7—Linear-output counter.

a smoothing capacitor, because the output voltage developed across this load effectively subtracts from the voltage swing of the limiter plate and so reduces the charge transferred by C during each cycle. This may be overcome by insuring that the cathode voltage of diode V_2 does not change in relation to the anode voltage of diode V_3 during the counter operation. Figure 7 shows how a cathode follower may be used to meet this condition and so provide an output voltage e varying substantially linearly with input frequency f. It is necessary to return V_3 to a point tapped down on the cathode resistor r_1, as shown, to prevent steady current flow in the diodes as a result of their own internal voltages (electron emission velocities, contact potentials, etc.).

Another solution is given by the null or slide-back counter of Figure 8, in which a balancing voltage from potentiometer r_2 is applied to the counter load r to reduce the counter-output voltage at point A to zero. The voltage-division ratio x of the potentiometer for balance varies linearly with frequency. This arrangement has the useful property that its null indication is independent of supply voltage E_o. A servo mechanism may easily be arranged to maintain the null condition by automatic adjustment of r_2, thus producing as final output a shaft rotation proportional to input frequency. Counter diodes may be connected to provide at will either a positive output voltage, as in Figure 7, or a negative output, as in Figure 8. When fed the beat-note output of an FM radar, these counters provide directly useful data output proportional to target range.

In the case of a radar using symmetrical-sawtooth frequency modulation, the frequency of the beat between transmitted signal and signal

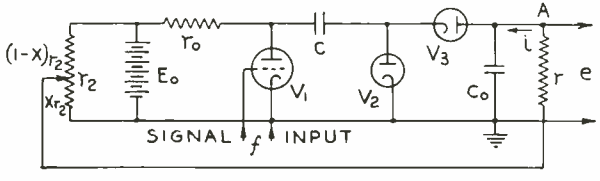


Fig. 8—Null-type counter.

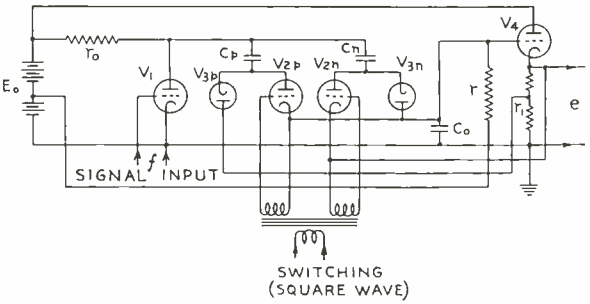


Fig. 9—Switched counters.

received after reflection from a moving target will alternate between two fixed values, $f_R + f_S$ on the downward modulation sweep and $f_R - f_S$ on the upward sweep (Section II(C) of Reference 1). Target range is represented by f_R and target speed by f_S (Section II(A) of Reference 1). To make full use of this information, two counters, one with positive and one with negative output, are required; these are switched into operation alternately, in synchronism with transmitter modulation. Switching is accomplished, as indicated in Figure 9, by adding a control grid to one diode of each counter, and using it to disable that counter by preventing diode action when the counter is not needed.

Response to the range frequency f_R, present equally during both up and down sweeps, just cancels out over the complete modulation cycle if the two counter capacitors C_p and C_n are just equal. This leaves a net switched-counter output proportional on the average only to the speed frequency f_S, which reverses its effect between upsweep and downsweep, and such an arrangement may be called a speed counter. Instantaneous voltage out of the speed counter will vary during the radar modulation cycle as shown in Figure 10, each cycle of the beat frequencies delivering its increment of charge to the storage capacitor C_o. If C_p and C_n are made unequal, the effect of the range-beat fre-

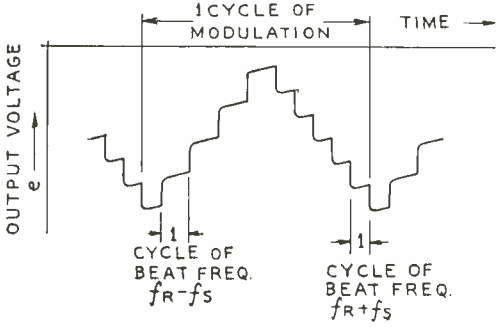


Fig. 10—Output voltage of switched counters.

quency will not cancel over the modulation cycle and the average counter output will vary in proportion both to target range and target speed.

The *range sensitivity* h_R of an unsymmetrical switched counter, with positive counter acting during modulation upsweep, is $\frac{1}{2} r(C_p - C_n)E_o$ volts per cycle per second of range-beat frequency; its *speed sensitivity* h_S is $\frac{1}{2} r(C_p + C_n)E_o$ volts per cycle of speed-beat frequency. With a radar range sensitivity of k_R cycles per second per unit range and radar speed sensitivity of k_S cycles per unit speed (as given by Equations (10) and (11) of Reference 1) and with counter-load resistor r returned to a bias voltage e_o as indicated in Figure 9, the total counter-output voltage e is then in time average

$$e = k_R h_R R - k_S h_S S + e_o, \tag{1}$$

where R is range or distance between and S is relative speed of approach of radar and target. This is the typical FM radar data output as noted in Reference 1; it is in a form suited for immediate utilization. If operating conditions are so chosen that speed frequency f_S exceeds range frequency f_R, the situation is reversed from the above description. The simple unswitched counter then responds to target speed and the symmetrical switched counter to target range; the unsymmetrical switched counter again provides combined range and speed data.

Data from the counters may be used in various ways. A meter connected as in Figure 6, or in series with r_1 of Figure 7, will give directly an indication of range (or of speed if f_S exceeds f_R). If the potentiometer r_2 of the null counter of Figure 8 is preset by hand, the voltage at point A can, by means of relays, actuate warning signals to indicate whether range exceeds, equals, or is less than the preset value. Alternatively, the relays can control motion of the craft bearing the radar so as to maintain automatically the preset range. If the null counter is kept balanced by a servo, the servo shaft provides suitable input data for many kinds of computers. In the case of the switched counter, passage of the combined counter-output voltage of Equation (1) through a preset value can automatically release a missile at the proper instant to insure a hit on an isolated target, or the total output can be indicated or used in computation or control as with the simpler counters.

RCA TECHNICAL PAPERS†

First Quarter, 1948

Any requests for copies of papers listed herein should be addressed to the publication to which credited.

"Antenna Design for Low-Angle FM Propagation", O. O. Fiet, *Tele-Tech* (February) 1948

"Antenna Facts", R. L. Rod, *CQ Magazine* (January) 1948

"Barrier Grid Storage Tube and Its Operation", A. S. Jensen, J. P. Smith, M. H. Mesner, and L. E. Flory, *RCA REVIEW* (March) 1948

"Broadcasting Studio Pickup Technique", H. M. Gurin, *Audio Engineering* (February) 1948

"Circuit Design Precautions to Prevent Internal Arcs from Damaging Kinescopes", *RCA Application Note AN-128*, RCA Tube Department, Harrison, N. J. (January 15) 1948

"A Developmental Pulse Triode for 200 Kw. Output at 600 Mc.", L. S. Nergaard, D. G. Burnside, and R. P. Stone, *Proc. I.R.E.* (March) 1948

"Electro-Optical Characteristics of Television Systems: Introduction; Part I — Characteristics of Vision and Visual Systems", O. H. Schade, *RCA REVIEW* (March) 1948

"Electronic Timers Employing Thyratrons 2D21 or 2050", *RCA Application Note AN-131*, RCA Tube Department, Harrison, N. J. (March 1) 1948

"Excitation and Emission Phenomena in Phosphors", H. W. Leverenz, Section of PREPARATION AND CHARACTERISTICS OF SOLID LUMINESCENT MATERIALS, John Wiley & Sons, Inc., New York, N. Y. 1948

"Free Space Microwave Propagation", A. L. Hammerschmidt, *RCA REVIEW* (March) 1948

FREQUENCY MODULATION, Volume I, *RCA REVIEW*, RCA Laboratories Division, Princeton, N. J. (January) 1948

"Frequency Stabilization with Microwave Spectral Lines", W. D. Hershberger and L. E. Norton, *RCA REVIEW* (March) 1948

"High-Speed Servicing", A. Liebscher, *Radio Service Dealer* (February) 1948
RCA Radio Service News (March-April) 1948

"How to Convert a Standard Television Receiver into a Direct-Viewing Video Extension Monitor", E. K. Price, *Broadcast News* (March) 1948

† Report all corrections or additions to RCA REVIEW, Radio Corporation of America, RCA Laboratories Division, Princeton, N. J.

"Liquid Dielectrics for Variable Condensers", S. Wald, *Radio News* (April) 1948

"A New Variable-Area Recorder Optical System", J. L. Pettus and L. T. Sachtleben, *Jour. Soc. Mot. Pic. Eng.* (January) .. 1948

"On the Investigation of Specimen Contamination in the Electron Microscope", J. Hillier, *Jour. Appl. Phys.* (March) 1948

"Optical Problems in Theatre Television", I. G. Maloff, *RCA Service Company News* (February) 1948

"Planning the Transmitter Building", J. Vassos and S. W. Pike, *Broadcast News* (March) 1948

"Principles of Frequency-Modulated Radar", I. Wolff and D. G. C. Luck, *RCA REVIEW* (March) 1948

"Progress in Television", V. K. Zworykin, *Radio Craft* (January) 1948

"Pulse-Operated High-Voltage Power Supply for Television Receivers", *RCA Application Note AN-130*, RCA Tube Department, Harrison, N. J. (February 16) 1948

"Quality Control of Miniature Tubes", W. L. Van Keuren, *RCA Rad. Serv. News* (March-April) 1948

"Quick Changing of 8D21 Tube", E. H. Potter, *Broadcast News* (March) 1948

"Radio-Frequency Performance of Some Receiving Tubes in Television Circuits", R. M. Cohen, *RCA REVIEW* (March) 1948

"RCA 'Special Red' Tubes for Industrial Applications", *RCA Application Note AN-129*, RCA Tube Department, Harrison, N. J. (February 2) 1948

RCA TECHNICAL PAPERS — INDEX, Volume II(b) (1947), *RCA REVIEW*, RCA Laboratories Division, Princeton, N. J. (March) 1948

"A Remote Control for Radio Tuning", S. Wald, *Radio News* (March) 1948

"Review and Preview of Television", David Sarnoff, *Inter. Photo.* (March) 1948

"The Sensitivity Performance of the Human Eye on an Absolute Scale", A. Rose, *Jour. Opt. Soc. Amer.* (February) 1948

"Servicing FM Receivers; Part IV — Special Circuits", M. Kaufman, *Radio Maintenance* (January) 1948

"Simultaneous Field Strength Recordings on 47.1, 106.5, and 700 Megacycles", W. L. Carlson, *RCA REVIEW* (March) 1948

"Some Observations on Light Absorption of Powdered Luminescent Materials", F. H. Nicoll, Section of PREPARATION AND CHARACTERISTICS OF SOLID LUMINESCENT MATERIALS, John Wiley & Sons, Inc., New York, N. Y. 1948

"Sound Reinforcement in the Hollywood Bowl", M. Rettinger and S. M. Stevens, *Audio Engineering* (February) 1948

"Sound Service is not a Luxury", E. Stanko, *Theatre Catalog 1948* (January) 1948

"Stereoscopic Viewing of Cathode-Ray Tube Presentations", H. Iams, R. L. Burtner, and C. H. Chandler, *RCA REVIEW* (March) 1948

"Tape Relay System for Radiotelegraph Operation", S. Sparks and R. G. Kreer, *Teleg. and Teleph. Age* (March and April) 1948
Reprinted from *RCA REVIEW* (September) 1947

"Teleran", W. W. Watts, *Signals* (January-February) 1948

"Television DC Component", K. R. Wendt, *RCA REVIEW* (March) 1948

"Television Today", F. W. Wentker, *Transmitter* (Spring) 1948

"Theater Television — a General Analysis", A. N. Goldsmith, *Jour. Soc. Mot. Pic. Eng.* (February) 1948
Inter. Project. (May) 1948

"Theatre Television Prospects", A. N. Goldsmith, *Inter. Project.* (January) 1948

"Triplexed Antenna System of WNBW/WRC-FM", L. J. Wolfe, *Broadcast News* (March) 1948

"TT-5A Television Transmitter", C. D. Kentner, *Broadcast News* (March) 1948

"A Tube Complement for AC/DC AM/FM Receivers", *RCA Application Note AN-127*, RCA Tube Department, Harrison, N. J. (January 2) 1948

"WNBW and WRC-FM, NBC, Washington", R. F. Guy, *Broadcast News* (March) 1948

"Working in Research", a pamphlet, RCA Laboratories Division, Princeton, N. J. (February) 1948

"Zinc: Cadmium-Sulfide: Selenide Phosphors—Part I: Synthesis and Constitution", H. W. Leverenz, E. J. Wood, S. Lasof, and R. E. Shrader, Section of PREPARATION AND CHARACTERISTICS OF SOLID LUMINESCENT MATERIALS, John Wiley & Sons, Inc., New York, N. Y. 1948

"Zinc: Cadmium-Sulfide: Selenide Phosphors—Part II: Photoluminescence and Cathodoluminescence as a Function of Temperature; Crystal Structure; Diffuse Reflectivity; and Photolysis", S. Lasof, R. E. Shrader, and H. W. Leverenz, Section of PREPARATION AND CHARACTERISTICS OF SOLID LUMINESCENT MATERIALS, John Wiley & Sons, Inc., New York, N. Y. 1948

"Zinc: Cadmium-Sulfide: Selenide Phosphors—Part III: Spectral Emission Characteristics and Relative Intensities", R. E. Shrader, S. Lasof, and H. W. Leverenz, Section of PREPARATION AND CHARACTERISTICS OF SOLID LUMINESCENT MATERIALS, John Wiley & Sons, Inc., New York, N. Y. 1948

NOTE.—Omissions or errors in these listings will be corrected in the yearly index.

Corrections:

The following corrections refer to the paper entitled "Small-Signal Analysis of Traveling-Wave Tube", published in the December, 1947 issue, Vol. VIII, No. 4, pp. 585-611:

p. 597, Equation (22)—the symbol i before h_m should be deleted.

p. 599, Equation (28) should read—

$$V_0 = \frac{1}{2}\frac{m}{e}v_p^2 = \frac{1}{2\dfrac{e}{m}}\frac{c^2}{\left(1+\dfrac{x^2\nu^2}{\mu^2}\right)}$$

p. 602, Equation (28a) and page 610—the quantity $\left(\dfrac{\lambda}{2\pi b}\right)$ should be inverted.

p. 605, Equation (40) should read—

$$(-1)^m p_m(ix) = \frac{\Gamma_n}{x^2}\frac{\pi b^1}{\lambda}\left\{ \text{etc.}\right.$$

p. 606, Equation (45) should read—

$$(F)_{m=n} = 1 + \frac{\alpha^2 ec}{2kt}\frac{(\pi^7 I_0)^{1/3}}{(e/m)^{2/3}}\frac{b}{\lambda\mu}Q_0\left(ix, \frac{a}{b}\right) \quad ; \quad \left(\frac{a}{b} \leqq 1\right)$$

p. 607, Equation (46) should read—

$$(F)_{m=0} = 1 + \frac{\alpha^2 ecb}{4kt\lambda}\left[\frac{\pi^4 I_n}{(4e/m)^2}\right]^{1/3}\frac{e^{-(2/3)\mu}}{\mu^{7/3}}J_n^{2/3}\left(i\mu, \frac{a}{b}\right)\left(4+\frac{1}{\mu}\right)$$

p. 609, Equation (49) should read—

$$F = 1 + \frac{\alpha^2 ec}{2kt}\frac{(\pi^7 I_0)^{1/3}}{(e/m)^{2/3}}\frac{b}{\lambda\mu}Q_0\left(ix, \frac{a}{b} \geqq 1\right)$$

p. 609, Equation (51) should read—

$$(F)_{m=0} = 1 + \frac{\alpha^2 ecb}{4kt\lambda}\left[\frac{\pi^4 I_n}{(4e/m)^2}\right]^{1/3}\left(\frac{b}{a}\right)^{1/3}\frac{e^{-(2/3)\mu\frac{a}{b}}}{\mu^{7/3}}\left(4+\frac{1}{\mu}\right)$$

AUTHORS

AUDREY M. ARZINGER was graduated from the Senior High School of Rutherford, New Jersey, in 1938. Since 1942, she has been associated with the Radio Systems Research Laboratory of RCA Laboratories Division in New York, N. Y., as statistician on terrestrial magnetism, solar and key-circuit performance investigations.

GEORGE H. BROWN received the B.S. degree at the University of Wisconsin in 1930; the degree of M.S. in 1931; the Ph.D. degree in 1933; and his professional degree of E.E. in 1942. From 1930 until 1933 he was a Research Fellow in the electrical engineering department at the University of Wisconsin, and from 1933 to 1942 he was in the research division of the RCA Manufacturing Company at Camden, N. J. Since 1942, he has been at RCA Laboratories Division, Princeton, N. J. Dr. Brown is a Member of Sigma Xi, the American Institute of Electrical Engineers, New York Academy of Sciences, and a Fellow of the Institute of Radio Engineers, and the American Institute of Electrical Engineers.

HUGH L. DONLEY received the B.S. degree from Hobart College in 1930, the M.Sc. and the Ph.D. degrees from Brown University in 1932 and 1935 respectively. From 1930 until 1933 he was assistant in Physics at Brown University, and from 1933 to 1934 he was a Fellow in Physics at Brown University. From 1935 to 1942 he was in the research division of RCA Manufacturing Company at Camden, N. J. Since 1942 he has been with RCA Laboratories Division, Princeton, N. J. Dr. Donley is a member of Phi Beta Kappa, Sigma Xi, and a Senior Member of the Institute of Radio Engineers.

JESS EPSTEIN received the E.E. degree and the M.S. degree in Physics from the University of Cincinnati in 1932 and 1934. From 1934 to 1935, he was an instructor in Physics at the Cincinnati College of Pharmacy. In 1935, he joined the research division of RCA Manufacturing Company at Camden, N. J., and was transferred to RCA Laboratories Division at Princeton, N. J., in 1942, where he is now engaged in work on antennas, transmission lines and propagation. Mr. Epstein is a Member of Sigma Xi, and a Senior Member of the Institute of Radio Engineers.

ROBERT M. FRASER graduated from the University of Minnesota with a B.A. in physics and mathematics in 1935. He joined the National Broadcasting Company in 1937 as a staff photographer and was transferred to Television Operations as a maintenance engineer in 1939. In 1942, he was assigned to the Development Group, working on airborne television applications, and since 1945 has been engaged in the development of television recording. Mr. Fraser is an Associate Member of the Society of Motion Picture Engineers and a Member of the Institute of Radio Engineers.

HENRY E. HALLBORG graduated from Brown University in 1907 with a degree of B.Sc. in electrical engineering. After graduation, he joined the General Electric Company, first as a student engineer and then in the Testing Department. Following this, he was associated with the following organizations: 1909-1912, National Electric Signaling Company; 1912-1915, Marconi Wireless Telegraph Company of America; 1915-1923, U.S. Navy Department as an expert on radio aids; and 1923-1925, consulting engineer, C. Brandes Company. In 1925, he joined the Radio Corporation of America, and until 1932 worked on short-wave transmitter development. Since 1932, he has been performing research on terrestrial magnetism and ionosphere and space circuit analysis, first with RCA Communications, Inc., and since 1942, with the Radio Systems Research Laboratory of RCA Laboratories Division, New York, N. Y. Mr. Hallborg is Junior Past-President and Director of the Brown Engineering Association and a Fellow of the Institute of Radio Engineers.

VERNON D. LANDON attended Detroit Junior College. From 1922 to 1929, he was in charge of the Radio-Frequency Laboratory of the Westinghouse Electric and Manufacturing Company. In 1930 he was Assistant Chief Engineer of the Radio-Frequency Laboratories and in 1931 he became Assistant Chief Engineer of the Grigsby Grunow Company. In 1932, he joined RCA Manufacturing Company in Camden, N. J. Since 1942, he has been with RCA Laboratories Division at Princeton, N. J. engaged in research work on radio circuits. Mr. Landon is a Member of Sigma Xi, and a Senior Member of the Institute of Radio Engineers.

DAVID G. C. LUCK received his B.S. degree from Massachusetts Institute of Technology in 1927 and his Ph.D. in 1932. In 1927-28 he was a Swope Fellow in Physics, and a Malcolm Cotton Brown Fellow in 1928-29. He was an Assistant in the Department of Physics at Massachusetts Institute of Technology in 1929-32. He joined the Research Division of RCA Victor Company in 1932 and was with the Victor Division of RCA Manufacturing Company from 1935 to 1941, working on pulse communication, direction finding, and radio guidance of aircraft. He was transferred to RCA Laboratories Division upon its formation and has remained with the Princeton laboratories of the Division since that time. Dr. Luck is a Member of the American Physical Society, Sigma Xi, and the Institute of Radio Engineers.

JOHN H. NELSON received his schooling at West Newbury High in Massachusetts, and the Massachusetts University of Extension in Boston. Since 1923, he has been with RCA Communications, Inc., as Operator, Traffic Chief, and Supervisor at the Central Radio Office. In 1946, he was promoted to Assistant to the Controller of Traffic Facilities and since that time has specialized in short-wave radio propagation and ionospheric disturbance research.

DONALD W. PETERSON received the B.S. degree in E.E. from the University of Wisconsin in 1936. In that year, he joined the Service Department of RCA Manufacturing Company, Camden, N. J., shifting to the research division of that company in 1939. In 1942 he was transferred to RCA Laboratories Division in Princeton, N. J., where he is currently engaged in work on antennas, transmission lines and propagation. Mr. Peterson is a Member of Sigma Xi and the Institute of Radio Engineers.

OTTO H. SCHADE graduated from the Reform-Real-Gymnasium, Halle, Germany, in 1922. From 1922 to 1924 he was with the Telephonfabrik A. G. vorm. J. Berliner, Berlin and Dusseldorf; from 1924 to 1925, in charge of the laboratory in the radio manufacturing company "Ratag" in Berlin; and from 1926 to 1931, in the engineering department of the Atwater Kent Manufacturing Company. Since 1931 he has been with the Tube Department, RCA Victor Division at Harrison, N. J. He received a Modern Pioneer Award from the Radio Manufacturers Association in 1940. Mr. Schade is a Senior Member of the Institute of Radio Engineers.

IRVING WOLFF received the B.S. degree in physics from Dartmouth College in 1916 and the Ph.D. degree in physics from Cornell University in 1923. He was an instructor in physics at Iowa State College during 1919 and at Cornell University from 1920 to 1923, where he was a Hechscher Research Fellow in 1924. He joined the Technical and Test Department of the Radio Corporation of America in 1924. From 1930 to 1941, in the Research Division of the RCA Manufacturing Co., he worked on problems in microwaves, radar, and aviation. Since 1941 he has been connected with the RCA Laboratories Division, Princeton, N. J., where he is now Director of the Radio Tube Research Laboratory. Dr. Wolff is a Fellow of the Acoustical Society of America and the Institute of Radio Engineers, and a Member of the Physical Society, American Association for the Advancement of Science, and Sigma Xi.