THE BELL SYSTEM
TECHNICAL JOURNAL

VOLUME XXXVIII SEPTEMBER 1959 NUMBER 5

Studies in Tropospheric Propagation Beyond the Horizon*

By A. B. CRAWFORD, D. C. HOGG and W. H. KUMMER

(Manuscript received April 17, 1959)

This paper describes an extended series of experiments in beyond-the-horizon propagation on a 171-mile overland path using 460 and 4110 mc. The following aspects of the propagation were investigated: the effect of antenna size on signal level and fading characteristics, wavelength dependence, seasonal and diurnal effects, a new form of diversity reception, the bandwidth capability of the medium. Many of the experiments were directed toward a better understanding of the mechanism of propagation.

TABLE OF CONTENTS

* This work was supported in part by Contract AF18(600)-572 with the U. S. Air Force, Air Research and Development Command.

I. INTRODUCTION

There has been a tremendous increase in interest in tropospheric propagation beyond the horizon since Bullington[1] first suggested its useful properties in 1950. Early studies at Bell Telephone Laboratories and elsewhere[2] led to the design and installation of large-scale military and commercial systems. To a great extent, the design of these systems was based on limited data, since time did not permit the implementation of long-term research programs specifically designed to study the mechanism of the propagation. Hand-in-hand with the interest in the systems themselves, several theories were proposed to explain the propagation. These can be classified into three general types: scattering by atmospheric turbulence (Booker-Gordon),[3] mode propagation (Carroll-Ring)[4] and reflection from atmospheric layers (Friis, Crawford and Hogg).[5]

This paper describes the results of a series of experiments conducted on a 171-mile overland path between Pharsalia, New York, and Crawford Hill, New Jersey, during the period from May 1955 to September 1958.

At the time these experiments were planned, some of the characteristics of beyond-the-horizon propagation were known, but other important properties had not been investigated thoroughly. Some of the aims of our experiment were: to study the effect of antenna size on signal level and fading characteristics; to determine the wavelength dependence of the propagation; to look for seasonal and diurnal effects; to investigate a new form of diversity reception which combines the outputs of twin-feed horns horizontally or vertically disposed at the focal point of a single paraboloid; to study the bandwidth capability of

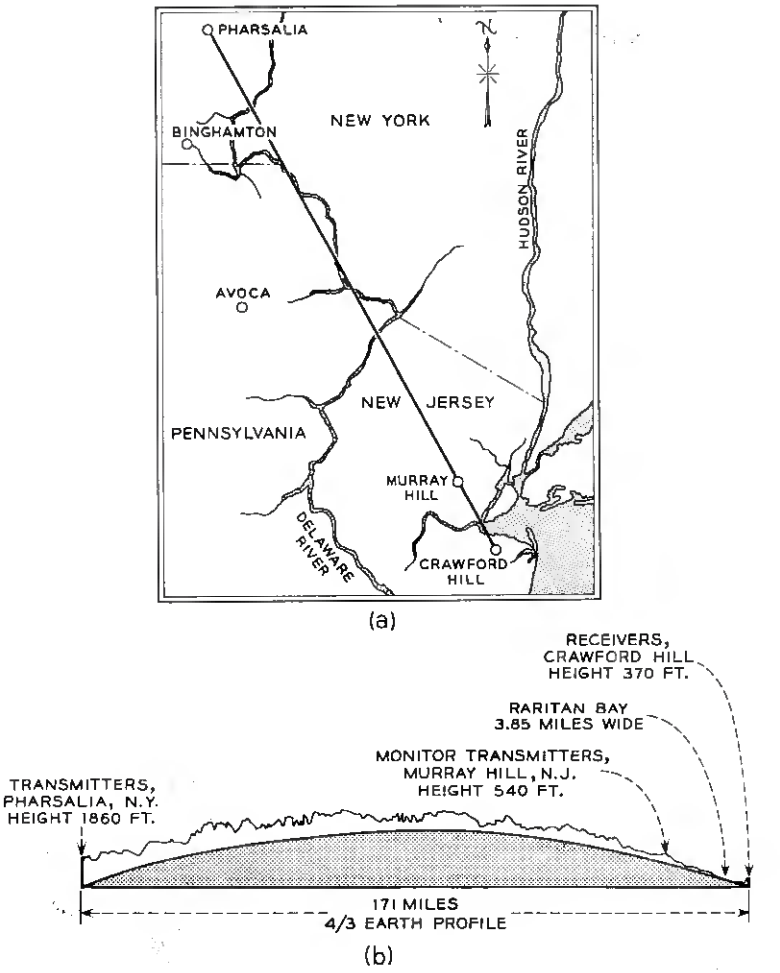


Fig. 1 — Propagation path.

the propagation; to learn as much as possible about the mechanism involved.

Previous experience in the short-wave and microwave fields had shown that a valuable research tool for studying propagation mechanisms is a narrow beam receiving antenna capable of being scanned through the incoming wave fronts. For the present experiment, a solid-surface aluminum paraboloid, 60 feet in diameter,* was constructed on

* This antenna was designed by H. T. Friis.[6]

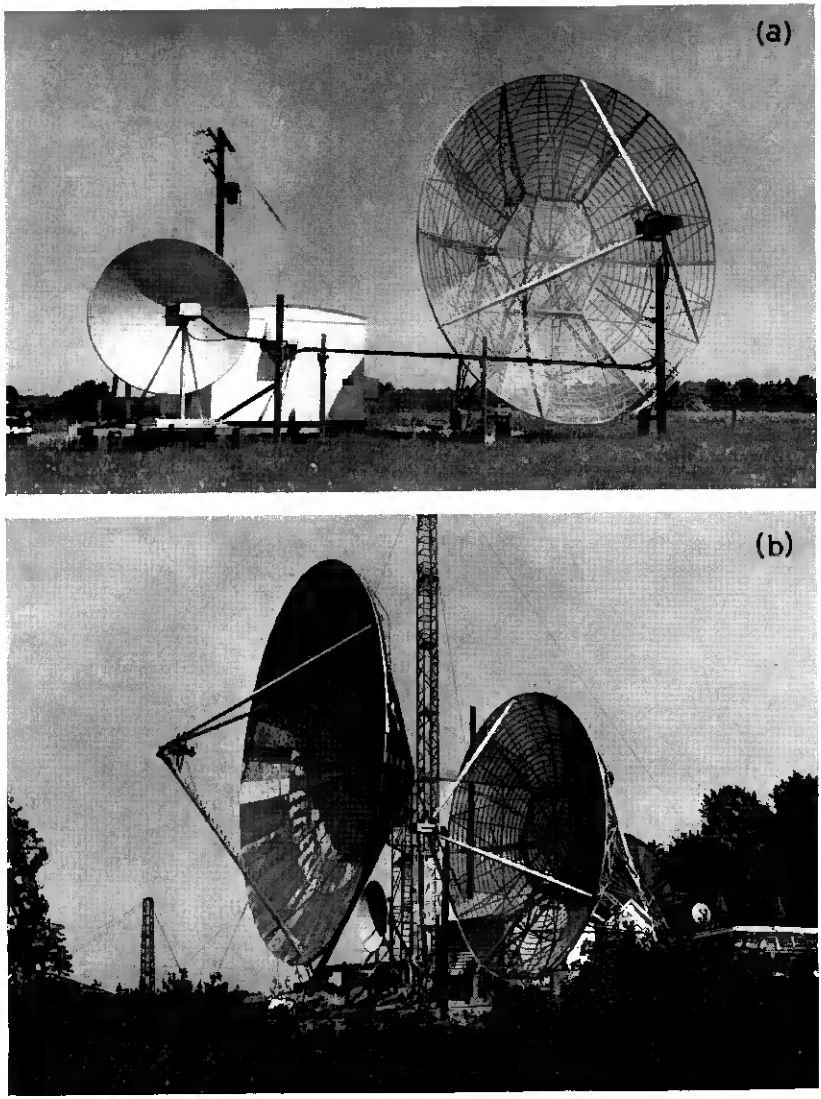


Fig. 2 — Terminal sites: (a) transmitter; (b) receiver.

Crawford Hill. Means were provided for scanning the antenna through an angle of approximately three degrees in azimuth and elevation at a rate of about four scans per minute.

The map in Fig. 1(a) shows the location of the transmission path for the experiment. The immediate foregrounds of the transmitting and

receiving sites are unobstructed, as shown in the profile map. This particular path was chosen to pass over the Murray Hill Laboratory, thereby providing a line-of-sight path with good clearance above ground between the Crawford Hill and Murray Hill locations [Fig. 1(b)]. The Murray Hill tower therefore was an excellent location for a monitor transmitter to be used in measuring the performance of the receiving antennas and to serve as a reference point for the angle of arrival measurements.

To investigate the effect of antenna size, two transmitting antennas, 10 feet and 28 feet in diameter, and three receiving antennas, 8 feet, 28 feet and 60 feet in diameter, were available (see Fig. 2). To study the propagation at two widely separated wavelengths, low-power transmitters designed for unattended operation at 4110 mc and 460 mc were installed at the Pharsalia site. Composite-feed horns consisting of a small 4110-mc horn within a 460-mc horn were available to permit operation at either of the two frequencies. Table I lists the frequencies, transmitter power, antenna gains and other data relating to the experiment.

In discussing the experimental results, comparisons are made with the theory of reflections from atmospheric layers.[5] This theory is based on the simple assumption that propagation beyond the horizon is due to reflections from a large number of randomly disposed layers in the volume of the atmosphere common to the beams of the transmitting and receiving antennas. The reflecting layers are formed by relatively steep gradients in the dielectric constant of the atmosphere; these gradients have been observed by refractometer measurements. While no claim is made that this theory is unique in explaining beyond-the-horizon propagation, it has been found to be in good agreement with experimental data.

References to some of the published work in the field are given in the paper but no attempt has been made to present a complete list of references. Ref. 2 contains many papers on the subject and references to many others. Refs. 8 and 9 give a list of references up to 1958.

The remainder of this paper is divided into seven sections, as follows:

Section II describes the design of the experiment and discusses the transmitting and receiving-recording apparatus.

Section III is concerned with the median received power and its dependence on antenna size and wavelength. It was found that, over a one-year period, the median signal level at 4110 mc received with the 60-foot antenna exceeded that received with the 8-foot antenna by only 6 db, compared with the difference of 17 db in the gains of the antennas. Comparing transmission at 460 mc and 4110 mc in a scaled experiment,

TABLE I — PARAMETERS FOR THE 171-MILE EXPERIMENTAL CIRCUIT

Frequency, mc	Transmitter Power	Free Space Path Loss, db	Antennas						Free Space Received Power		Estimated Beyond-the-Horizon Loss, db†	Estimated Received Power, dbm
			Transmitting			Receiving						
			Diameter, feet	Gain, db	Line loss, db	Diameter, feet	Gain, db	Line loss, db	Antenna Sizes, feet	Power, dbm*		
4110	20–30 watts 43 dbm	153.5	10	38.5	0.33	8	37.0	0.32	10–8	−36	73	−109
			28	47.5	0.44	28	47.5	0.47				
						60	54.5	1.6	28–60	−10.5		—
460	500–700 watts 57 dbm	134.5	10	20.5	0.79	8	—	—	10–28	−28.5	73	−101.5
			28	30.5	1.35	28	30.5	1.29				
						60	37.0	2.15	28–60	−13.5		—

* Including antenna line loss.
† From Ref. 7, Fig. 4.

it was found that, on the average, the median received power was proportional to the wavelength, as predicted by the theory.

Section IV gives the results of scanning or beam-swinging experiments using the 60-foot antenna at 4110 mc. The angular distribution of the received energy varied from day to day; it was found to be directly related to the ratio of the median power received on the 60-foot and 8-foot antennas. On the average, the apparent half-power beamwidth of the 60-foot antenna was about 1.0°, compared with the free space value of 0.3°.

Section V is a study of instantaneous fading rates. It is shown that the fading rate (at 4110 mc) is related to the size of the antennas and to the velocity of horizontal drift winds normal to the propagation path in the common volume of the antenna beams. In a scaled experiment, the ratio of the fading rates at 460 mc and 4110 mc was found to be significantly greater than the ratio of the frequencies.

Section VI discusses the twin-feed diversity experiments. At 4110 mc this type of diversity proved to be effective for both horizontally and vertically disposed feed horns. At 460 mc the diversity was effective at all times only for vertically placed feeds; with horizontal feed disposition, the effectiveness was a function of the rate of fading.

Section VII is a theoretical and experimental study of the characteristics of the short-term fading of the received signals; in particular, the number of one-way signal crossings and average duration of fade at a given signal level are treated for a single channel and for multichannel diversity systems.

Section VIII describes the results of frequency-sweep experiments intended to yield information on the bandwidth capability of the propagation. The transmission in a band of 15 mc centered at 4110 mc was characterized by the selective fading typical of multipath propagation. The bandwidth, defined as the frequency spacing of nulls in the transmission band, was about 5 mc on the average and was sensibly independent of antenna size, rate of fading and median signal level.

II. DESIGN OF THE EXPERIMENT

The experiments to be described have been carried out at Holmdel on beyond-the-horizon propagation using an experimental circuit of 171 miles between Pharsalia, New York, and Crawford Hill, New Jersey (Figs. 1 and 2). There are no obstructions in the transmitting and receiving foregrounds. The foreground profile at the receiving location is shown in Fig. 3.

The design of equipment for experiments in this type of propagation

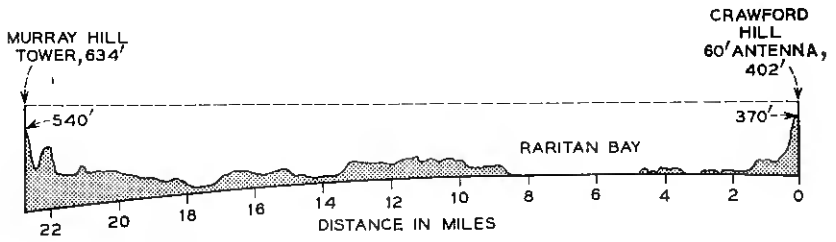


Fig. 3 — Foreground path profile at receiving end.

is dictated by the transmission path selected as well as by the experiments to be performed. Since propagation conditions vary considerably from hour to hour as well as from week to week and month to month, a particular experiment may run from several months to a year, with continuous recording of data being taken for several days each week. Several different equipment setups may be used in any one week. These are then repeated in succeeding weeks.

Equipment reliability and gain stability are important design considerations if a large amount of experimental data is to be obtained with minimum personnel. Flexibility of equipment setups and automatic data-taking also reduce the manpower needed and increase the time available for recording.

For our experiment, the transmitting equipment was designed for unattended operation with bimonthly maintenance periods; the receiving equipment was calibrated every four days.

2.1 *Receiver Design*

The receiver design is determined by the power available at the transmitter, the path loss and the antennas employed.

Table I summarizes the parameters used in this particular propagation path. The estimated beyond-the-horizon loss of 73 db was taken from Bullington[7] and gave the best estimate available at the beginning of our project. The signal will vary about 15 db around this value because of seasonal changes in transmission. From Table I it is seen that the average median level of received power will be -109 dbm at 4 kmc; during the worst period it will drop to -124 dbm.

It is known that the instantaneous signal is usually Rayleigh-distributed; it will drop 10 db below the median for 6.7 per cent of the time, 20 db below the median for 0.7 per cent of the time and 30 db below for 0.07 per cent of the time. The lowest expected signal level for 0.07 per cent of the time for the worst period would be -124 dbm $-$ 30 db $=$ -154 dbm.

TABLE II — REQUIRED RECEIVER BANDWIDTHS FOR VARIOUS SIGNAL-TO-NOISE RATIOS

Bandwidth, kc	Signal-to-Noise Ratio, db, for $P_R = -109$ dbm
150	1
25	9
6	15

Assuming an over-all noise figure of 12 db and a median received power of −109 dbm, we can calculate the bandwidth needed for various signal-to-noise ratios. The results are shown in Table II. Thus, from the table it is apparent that a 6-kc bandwidth is needed at low signal levels, since the median signal level can drop to −124 dbm. For that level, the signal-to-noise ratio would be 15 db + (−124 dbm + 109 dbm) = 0 db. The use of the 6-kc bandwidth means that the transmitter and receiver frequencies must be controlled to better than one part in 10^6.

The description of the equipment follows; the receivers will be discussed first, then the transmitters and finally the recording equipment.

2.2 *Receivers and Transmitters*

Two receivers are used; each has two RF heads, one for 4 kmc and one for 460 mc, with both feeding into the same IF portion of the receiver. Hence two channels of a single frequency or two different frequencies can be recorded simultaneously. The 4-kmc receiver will be discussed first; it is to be understood that the only difference between it and the 460 mc receiver is in the RF section and first beating oscillator.

In the 4-kmc receiver the RF signal is mixed with the first beating oscillator in a balanced converter. A matched pair of 1N23C and 1N23CR crystals are used, the IF cable to the first amplifier being adjusted for the best noise figure. The converter noise figure is 9 to 11 db, depending on the crystals selected.

To achieve the narrow bandwidths required, we used the conventional technique of quadruple detection. The details are indicated on Fig. 4. The gain of each IF stage (66 mc, 1.8 mc bandwidth; 3 mc, 150 kc bandwidth; 150 kc, 25 or 6 kc bandwidth) is adjusted so that its gain is equal to the signal-to-noise ratio improvement when that IF is added in the receiver. In other words, a broadband detector at the output of any of the IF stages will read the same amount of noise when the RF converter is connected to a matched load.

The 66-mc amplifier* has five stages with a grounded-grid WE417A input stage. The gain of the amplifier is 95 db, for low-level input. Automatic gain control is applied to two stages to increase the dynamic range from 28 db to 60 db. Compression is applied only to the 66 mc amplifier and is controlled by the output of the last detector; this means that the relative db calibration at the last detector is the same for any IF bandwidth selected, assuming that the other IF amplifiers are linear. Additional loss must be inserted at the RF end if the signal is high and the narrowest IF bandwidth is used.

Succeeding IF amplifiers are linear over the intended range of operation. The automatic gain control amplifier is a high-level 40-db linear amplifier with a low output impedance (5000 ohms); it is linear up to 0.7 volt input. Hence the dc output after the detector is at a high level (-100 volts for -71 dbm input to the receiver using the 25-kc bandwidth) and at low impedance. A portion of this output voltage is used as the AGC voltage, which is fed back to the 66-mc IF. Gain stability problems are considerably reduced by not using any dc amplification in the receiver.

Frequency stability is achieved by using crystal-controlled beating oscillators. The one with the most critical stability problem is the first beating oscillator, which consists of a modified Western Electric microwave generator.

The gain stability of the over-all system is ± 1 db for about one-half week; this is using the 50-db range on the receiver and the 6-kc bandwidth. The frequency stability for that same period is $\pm 1\frac{1}{2}$ kc. The frequency deviation is recorded on a recorder with a 0–10 kc scale.

The 460-mc receiver consists of a WE416B triode as an RF amplifier, a crystal-controlled beating oscillator and two IF stages. The IF output is connected through an IF attenuator to the rest of the receiver, as shown in Fig. 4.

There are two cw transmitters used in these propagation tests. Both are crystal-controlled and designed for unattended operation by remote control. The control circuits incorporate time sequence, with air pressure and safety interlocking devices. The transmitters will be described with the aid of Fig. 5.

The power output of the 4-kmc transmitter is obtained from a Sperry SAC41 three-stage klystron amplifier. The RF output is 20–30 watts with a drive of 50 mw.

As stated in the last section, the frequency stability of the transmitter must be better than $\pm 1\frac{1}{2}$ kc (four parts in 10^7), when the narrowest IF

* Designed by W. C. Jakes, Jr., of the Holmdel laboratory of Bell Telephone Laboratories.

amplifier of the receiver is used. To achieve this, the driver for the klystron amplifier is a modified TD-2 Western Electric microwave generator. The modifications include increased temperature stability of the crystal oscillator by the use of specially cut crystals and a different crystal oven. The complete generator is also temperature-controlled.

An Esterline-Angus recorder monitors the power delivered to the antenna. Calibrated RF attenuators, directional couplers and power meter are used to calibrate the E-A recorder.

The 460-mc transmitter uses a crystal-controlled oscillator with frequency multiplication and amplifying stages to obtain the required 500 watts output. The component parts of this transmitter are: Motorola Type TA110 15-watt crystal-controlled exciter; Radio Engineering Labs transmitter output amplifier and power amplifier.

Since the frequency is about one-tenth that of the 4 kmc transmitter,

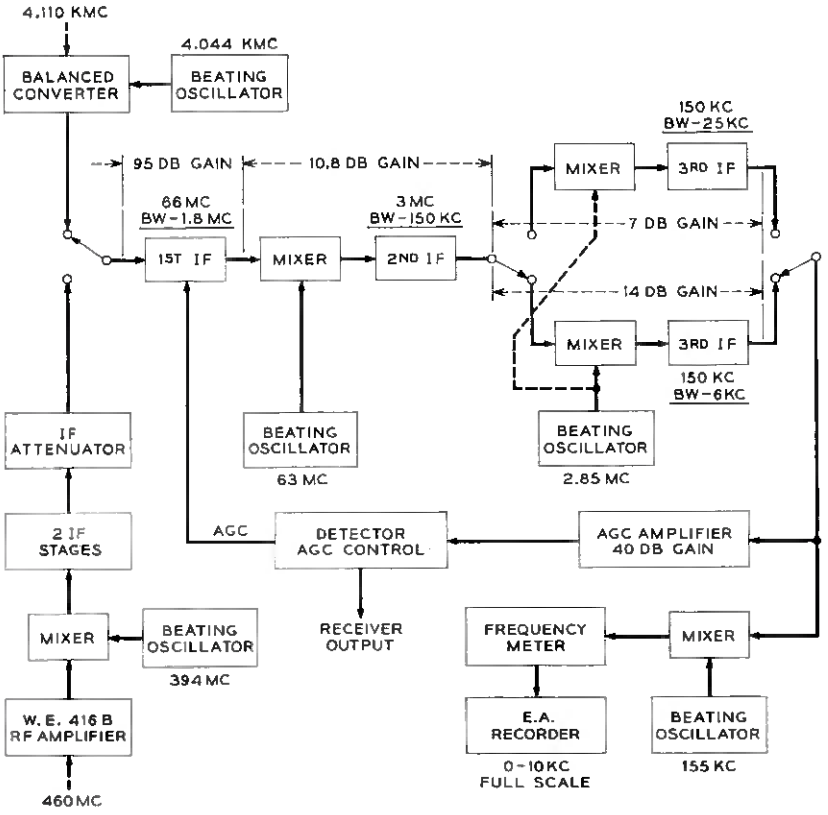


Fig. 4 — 460-mc and 4-kmc cw receivers.

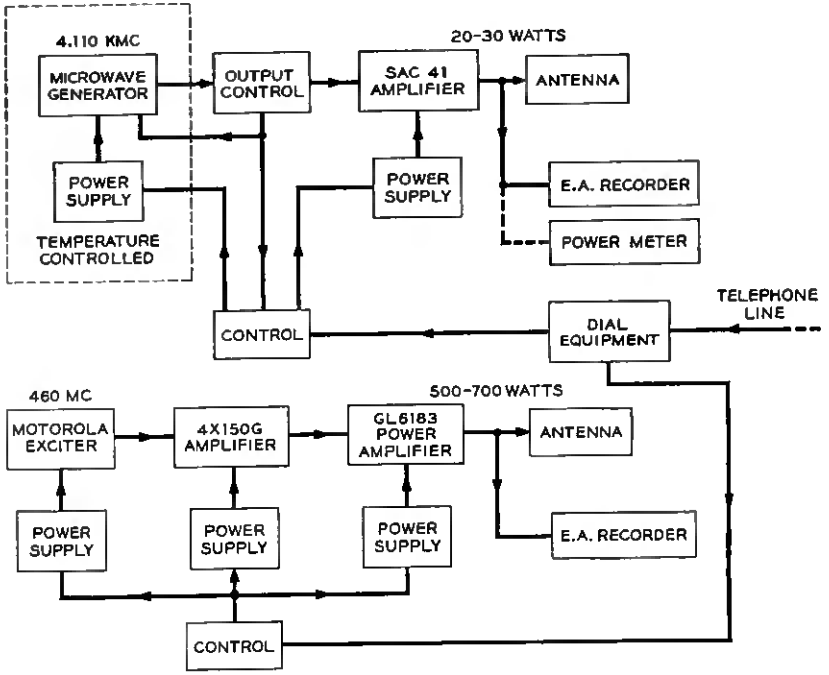


Fig. 5 — 460-mc and 4-kmc cw transmitters.

the frequency stability need only be one-tenth as good when measured on a percentage basis. However, the stability of the transmitter is only $\pm 2\frac{1}{2}$ kc. Since the transmitter power output is higher at 460 mc than at 4110 mc, the 6-kc receiver bandwidth is not required, and the stability of $\pm 2\frac{1}{2}$ kc is satisfactory when the 25-kc bandwidth is used.

2.3 *Recording Apparatus*

In propagation studies it is desirable to record the data so that the maximum information can be obtained without having to convert the original data into other presentations. Hence the presentation of the data depends on what use is to be made of them. Our program has consisted of the following analyses: statistical distribution of median signal levels and short-term fading for a long period, say a year (operation 60 hours a week); angle of arrival studies; diversity studies; and frequency sweep experiments.

Fig. 6 shows the setup for recording median signal levels with periodic sampling of the short-term fading. The output of the receiver is con-

nected to a Sanborn two-channel recorder. This recorder has a time constant (t.c.) of 0.015 second and has a recording speed of from 0.5 to 50 mm per second. The receiver output is also connected to a Doelcam dc amplifier which drives the Esterline-Angus (E-A) recorder through a 0 to 25 second t.c. circuit. The dc amplifier has a frequency response from 0 to 42 cps and is stable to 1 per cent full scale on a long-term basis. The Sanborn recorder is energized to take a sample of short-term fading every half-hour for an interval adjustable from one-half to five minutes. A marker pen on the E-A recorder is also energized during that period.

The recording setups for the angle of arrival, diversity and frequency-sweep experiments will be discussed in the sections where the experiments are described.

2.4 *Calibration Procedure*

The calibration procedure of the receivers consists of several steps. First, the RF input of the receiver is terminated in a matched load. Then each IF amplifier is adjusted to have the same noise output as the 66-mc IF amplifier of that receiver. This is done by means of a broadband crystal detector whose frequency response is constant within $\pm\frac{1}{4}$ db from 100 kc to 70 mc. The high-level 150-kc amplifier and high-level detector are adjusted to give -100 volts dc out with 0.7 volt ac in.

The matched load is removed from the input of the receiver and is replaced by a 4-kmc RF signal generator. This RF signal generator uses a microwave generator similar to the one used at the transmitter. Its output, monitored by a microwave power meter, is fed to the receiver input through calibrated directional couplers and RF attenuators. For an RF signal of -71 dbm (25 kc IF bandwidth), the AGC control is

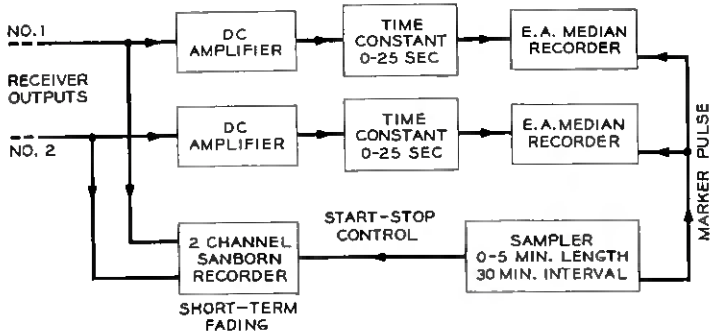


Fig. 6 — Recording setups for measuring median signal levels with periodic sampling of short-term fading.

adjusted to give −100 volts dc output. A calibration curve in steps of 5 db is taken; this completes the calibration.

To calibrate at 460 mc, a Hewlett-Packard Type 608A signal generator is used. This unit has a built-in power meter and calibrated attenuator.

III. MEDIAN RECEIVED POWER

In this section we shall discuss the median signal levels received on a tropospheric beyond-the-horizon system. In the first part our discussion will be limited to one frequency of transmission, 4 kmc; here we receive simultaneously on two antennas with different aperture areas. One of the interesting aspects of this type of propagation is that the power received by an antenna does not increase linearly with an increase in aperture area. The effect has been called "aperture-to-medium coupling loss" and "loss in antenna gain." The phenomenon has not been noticed at the lower frequencies primarily because it does not become appreciable until the aperture is of the order of a hundred wavelengths for a path of two hundred miles or so. Antennas of this aperture are physically realizable at present only in the microwave region.

In the second part of this section a scaled experiment to determine the wavelength dependence in this type of transmission is described; it is performed using simultaneous operation at two radio frequencies, 460 mc and 4 kmc, using almost equal antenna beamwidths. The antenna beamwidths are broad enough to make the "loss in gain" negligible; hence the atmosphere is the controlling factor in the transmission.

3.1 *Dependence on Antenna Size at 4 kmc*

In this study the received signal level was recorded for about 60 hours each weekend; the period of the experiment was from January 1956 to January 1957. The 4-kmc transmitter was used with the 10-foot antenna. Fig. 7 shows the common volume subtended by the antenna beams and the elevations involved. The results of this study show that the ratio of the median signal levels received by the 60-foot antenna and the 8-foot antenna are usually much smaller than would be expected on the basis of the antenna gains. The ratio of the gains of these two antennas is 17 db at 4 kmc. The ratio of the power levels received by the two antennas varies considerably from day to day and in some instances from hour to hour; on a yearly basis it is 5.7 db.

A strong seasonal dependence, as reported by Bean and Meany[10] and Bullington, has also been observed. The data indicate that the received signal level is approximately 15 db higher in summer than in winter.

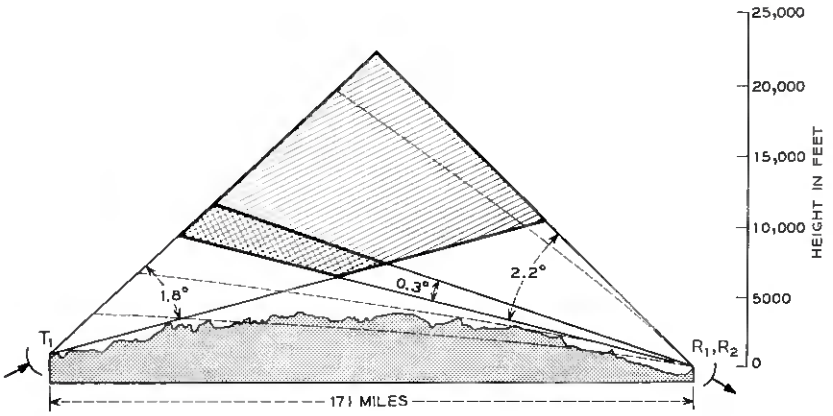


Fig. 7 — Common volumes in elevation plane for the 4-kmc circuit. The half-power beamwidths are shown for the 10-foot transmitting antenna (T_1) and the 60-foot (R_1) and 8-foot (R_2) receiving antennas. The dotted lines show the bending of the beams of the receiving antennas during unusual weather conditions at the receiving end (see Section 4.2).

This dependence is shown to be closely related to the refractive index of the atmosphere derived from data taken at ground level near the center of the path.

The fading range of the 15-minute medians is obtained for each weekend's recording. The fading range is taken as the ratio of the power levels at the 10 per cent and 90 per cent points of the distribution curve. The fading range of the 15-minute medians averaged over the year is 13 db for the signal on the 60-foot antenna and 9 db for the signal received on the 8-foot antenna.

3.1.1 *Method of Data Analysis — Samples of Recordings*

The received signal levels were recorded on Esterline-Angus recorders using a time constant of 25 seconds. The 15-minute medians were then tabulated and reduced to a percentage-of-time basis. The resultant distribution of these medians for the particular weekend period was plotted in percentage of time versus dbm on arithmetic probability paper. By this method the median received power level is established at the 50 per cent of time point for the given weekend. We will now discuss some samples of weekend recordings.

Fig. 8 shows a segment of the weekend recording for May 25–28, 1956. This record is typical of the field strengths encountered during the year. In this case the distributions for both antennas are fairly linear and

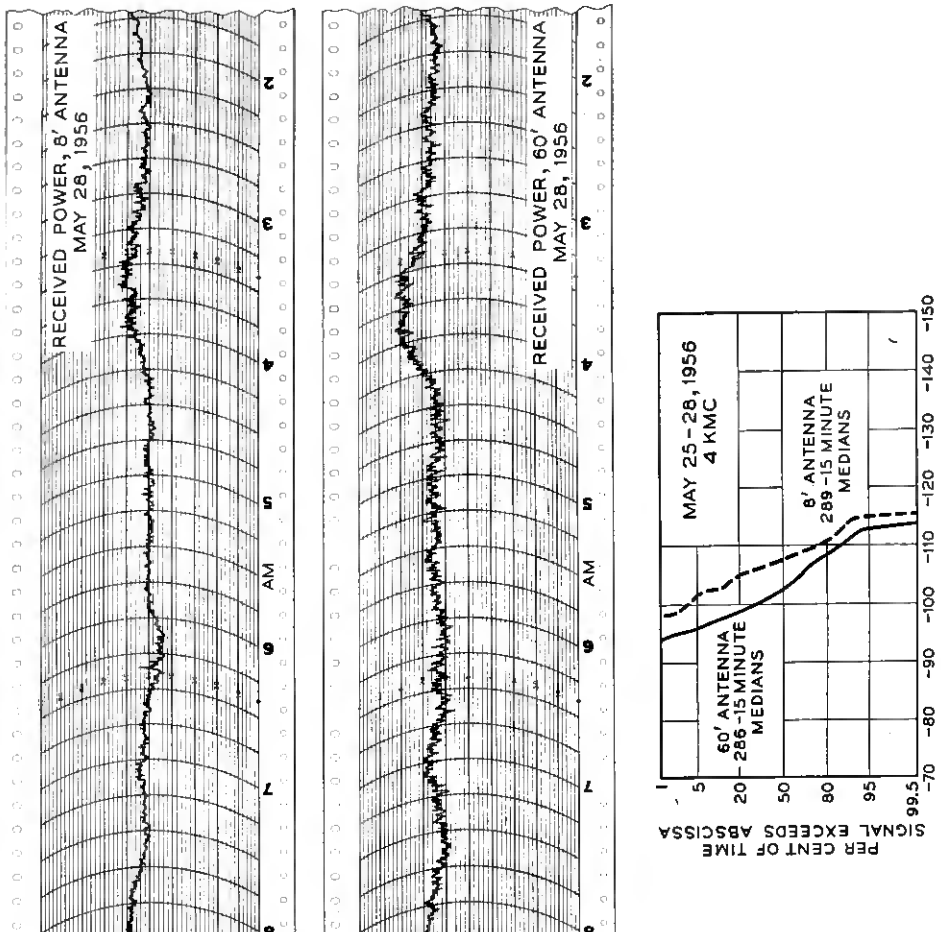


Fig. 8 — Segments of median signal level records and distribution curves for a typical weekend — not adjusted for antenna line losses (see Table I). The full scale on the records is 70 dbm. The 10 foot transmitting

indicate that the received power is distributed randomly when plotted in percentage of time versus decibels. The difference between the 50 per cent points of the two distributions is about 5.0 db, as indicated by the circles marked on the curves. When adjusted for antenna line losses, this difference becomes 6.3 db.

During the weekend of October 21–24, 1955, the difference in the medians of the two antenna was about 8.5 db, as shown in Fig. 9 (9.8 db when adjusted). Some of these distributions were far from linear, since there often was a departure from linearity in the curve at the higher signal level and low percentage-of-time regions. This condition was observed several times during the year; it is believed to be due to large stratified layers in the atmosphere.

The weekend of March 23–26, 1956, gave us our lowest medians of the year. A segment of the record is shown in Fig. 10. In this case the medians for the two antennas are almost equal; this is often found when the received power level is low.*

3.1.2 *Distribution of Weekend Medians for One Year*

The median signal levels of each weekend are plotted for the year from January 8, 1956, to January 13, 1957, in Fig. 11. The points represent the average power available at the output of the antenna feed horns. It may be seen from this curve that small ratios of the power received by the two antennas are most prevalent during the winter months.

In Fig. 12, the weekend (60-hour) medians of Fig. 11 have been plotted on a percentage-of-time basis for the year. The median values of these curves establish the performance of the antennas on a yearly basis. From the ratio of these medians we find that the power output of the 60-foot antenna is 5.7 db above that of the 8-foot antenna.

The median power output of the 8-foot antenna (on a yearly basis) was -109 dbm. Since the power received under free-space conditions for the 8-foot antenna would be -36 dbm, the transmission loss, relative to free space, was 73 db at 4110 mc for this 171-mile path.

3.1.3 *Comparison with Theory*

In the theory of reflection by layers, the transmission loss in radio propagation beyond the horizon for equal size transmitting and receiv-

* The numerous spikes in the record are reflections from aircraft flying through the path. These reflections are much more evident on the record obtained from the 8-foot antenna, since the 8-foot antenna gives a greater common volume than the 60-foot antenna (see Fig. 7).

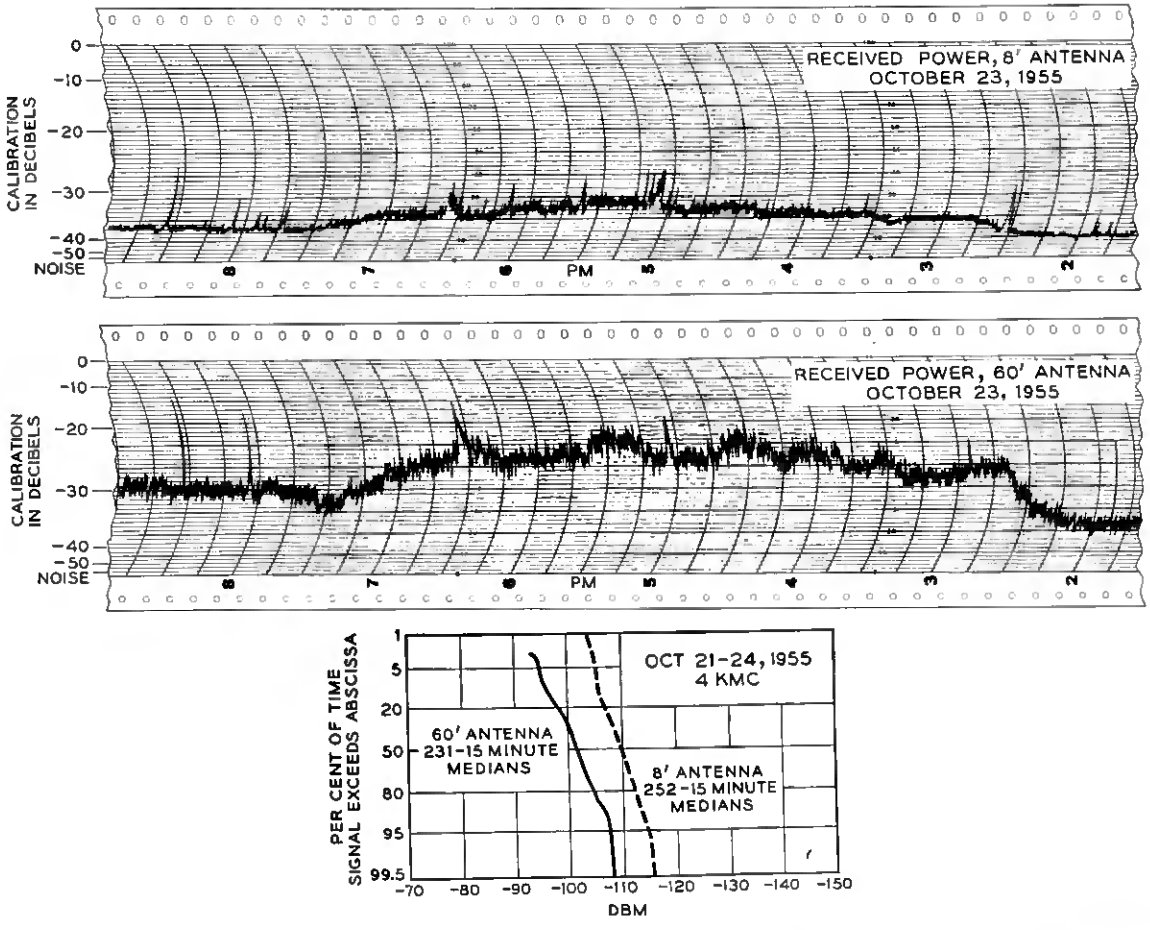


Fig. 9 — Segments of median signal level records and distribution curves for a high median signal weekend.

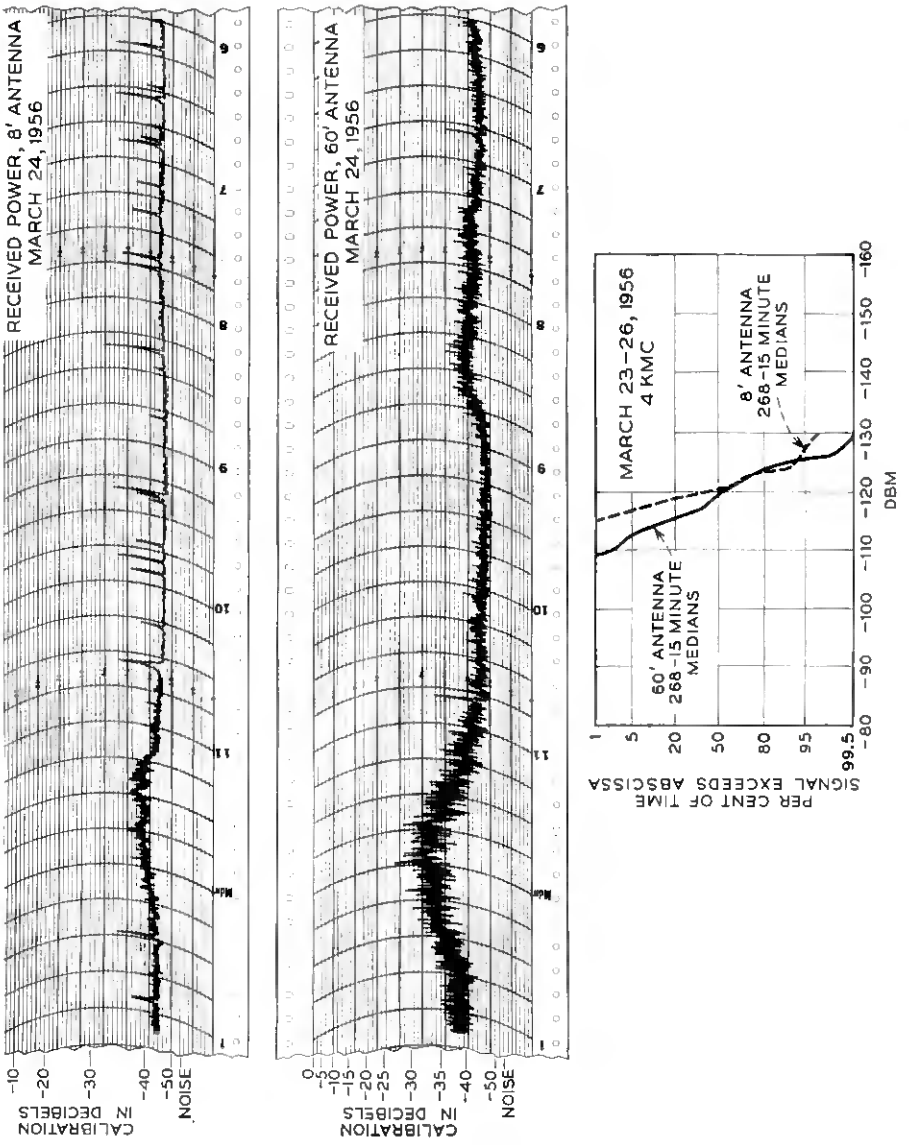


Fig. 10 — Segments of median signal level records and distribution curves for a low median signal weekend. The full scale on the records is −76 dbm.

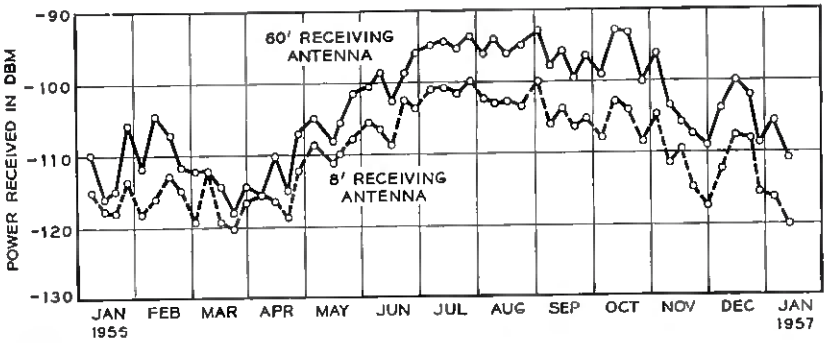


Fig. 11 — Weekend medians (60 hours) for one year of power received at 4110 mc. The 10-foot transmitting antenna was used with a power input of 43 dbm. Received power adjusted for antenna line losses. Period: January 1956 through January 1957.

ing antennas is given by three factors, as shown below in (1). The first term in brackets represents the power that would be received in free space, the second term involves the characteristics of the atmosphere, and the third term is a correction factor for narrow beam antennas:

$$P_R = \left[\frac{P_T\lambda^2}{4a^2\alpha^2}\right]\left[\frac{4M\lambda}{3\theta^4}\right]\left[\frac{\frac{\alpha}{\theta} f\left(\frac{\alpha}{\theta}\right)}{2+\frac{\alpha}{\theta}}\right], \tag{1}$$

where

P_R and P_T are the powers received and transmitted,
α, for actual antennas, may be taken as the half-power beamwidth,
θ is the angle between the lower edge of the idealized antenna pattern and a straight line joining the terminals and
$2a$ is the distance between the terminals.

These parameters are shown in Fig. 13 (Ref. 5, p. 636). Also,
λ is the wavelength,
M is a factor which includes the average size, number and strength of layers in the atmosphere, and

$$f\left(\frac{\alpha}{\theta}\right) = 1 + \frac{1}{\left(1+\frac{\alpha}{\theta}\right)^4} - \frac{1}{8}\left(\frac{2+\frac{\alpha}{\theta}}{1+\frac{\alpha}{\theta}}\right)^4. \tag{2}$$

We now compare the power received by the 60-foot and 8-foot antennas

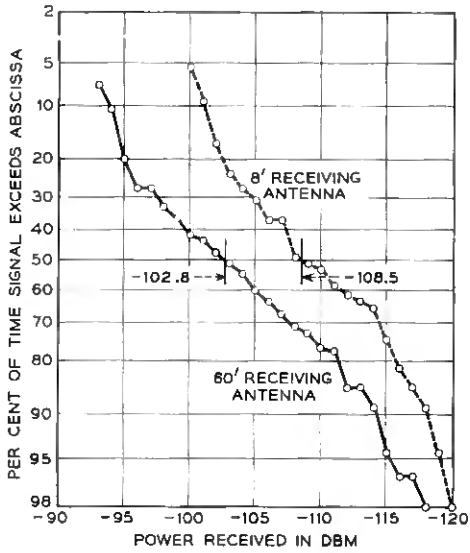


Fig. 12 — Yearly distribution of 60-hour weekend medians of power received at 4110 mc.

at 4 kmc using the theory. Since unequal beamwidths are involved, a generalization of (1) is necessary. (This is carried out in the Appendix.) The result is

$$\frac{P_{R_2}}{P_{R_1}} = \frac{\alpha_{R_1}{}^2(\alpha_T + 2\theta)\left[f\left(\frac{\alpha_T}{\theta}\right) + f\left(\frac{\alpha_{R_2}}{\theta}\right) - \left(\frac{\theta}{\theta + \alpha_{R_2}}\right)^4 f\left(\frac{\alpha_T - \alpha_{R_2}}{\theta + \alpha_{R_2}}\right)\right]}{\alpha_{R_2}\alpha_T(\alpha_{R_2} + 2\theta)\left[f\left(\frac{\alpha_T}{\theta}\right) + f\left(\frac{\alpha_{R_1}}{\theta}\right) - \left(\frac{\theta}{\theta + \alpha_T}\right)^4 f\left(\frac{\alpha_{R_1} - \alpha_T}{\theta + \alpha_T}\right)\right]}. \quad (3)$$

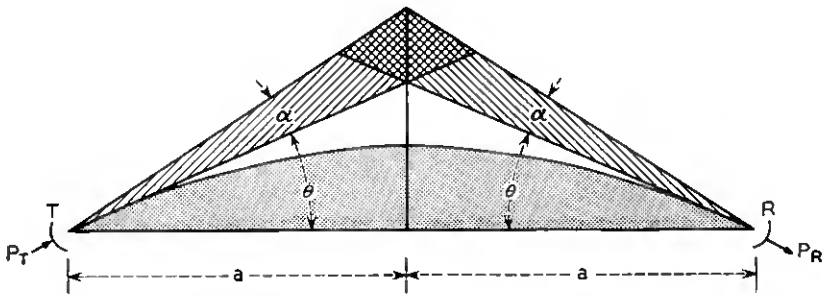


Fig. 13 — Parameters in a beyond-the-horizon circuit.

Substituting the appropriate parameters for this circuit, $\theta = 0.9°$ (4/3 earth), $\alpha_T = 1.8°$ (10-foot antenna), $\alpha_{R_1} = 2.2°$ (8-foot antenna), $\alpha_{R_2} = 0.33°$ (60-foot antenna), we get

$$\frac{P_{R_2}}{P_{R_1}} = 6.5 \text{ (8.1 db)}.$$

This compares favorably with our measurements of 5.7 db. This measured value is somewhat smaller than one might expect purely from the "loss in gain" of the 60-foot antenna, since it also includes the effects of local bending at the receiving site, as discussed in Section 4.2.

3.1.4 *Relationship between Propagation Data and Meteorological Conditions*

It is apparent from Fig. 11 that the received power has a strong seasonal dependence, being as much as 15 db higher in summer than in winter. In Fig. 14 the weekend values have been averaged over the month and replotted along with the average monthly temperature. The temperature, relative humidity and atmospheric pressure at ground level

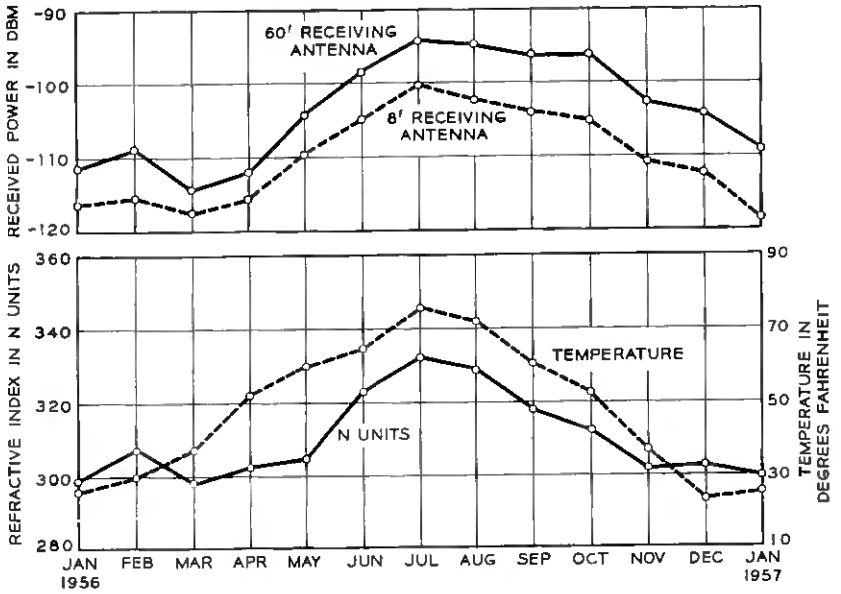


Fig. 14 — Monthly medians of received power at 4110 mc compared with surface temperature and refractive index from weather station at Avoca, Pa.

were obtained from the weather station at Avoca, Pennsylvania, which is near the middle of the propagation path. From these data the refractive index at the ground level was calculated and is also shown in Fig. 14. The formula[10] used for calculating the refractive index is:

$$N = (n - 1) \times 10^6 = \frac{77.6}{T}\left(p + \frac{4810\, e_s\, RH}{T}\right), \tag{4}$$

where

N = refractive index in "N" units,
n = refractive index of the atmosphere,
p = atmospheric pressure in millibars,
T = temperature in degrees Kelvin,
e_s = saturation vapor pressure in millibars for the temperature, T, and
RH = relative humidity.

The correlation between the received power level and the refractive index is believed to be due in part to the bending of the radio waves, as discussed in the theory (Ref. 5, p. 640). Thus, in the summer season when the refractive index is highest, more bending occurs and therefore the received power is higher.

3.2 *Wavelength Dependence*

This section describes a study in which measurements were made simultaneously at two widely separated wavelengths to investigate the wavelength dependence. The half-power beamwidths of the antennas are fairly broad at both frequencies, the vertical and horizontal beamwidths along with antenna sizes being shown in Fig. 15. While the beamwidths of the antennas used for the two wavelengths are not exactly equal, there is supporting evidence from our other experiments that the beams in all cases are sufficiently broad to insure that atmospheric effects are controlling over the effects of antenna size.

The method of analysis was the same as that for the experiment described in the previous section.

3.2.1 *Received Power at 460 and 4110 mc*

Fig. 16 shows the median power levels for both the 460-mc and 4110-mc signals. While the data are for a 21-month period (from April 1956 through December 1957) the data are plotted as a yearly period; the data from April through December 1956 were combined with data for

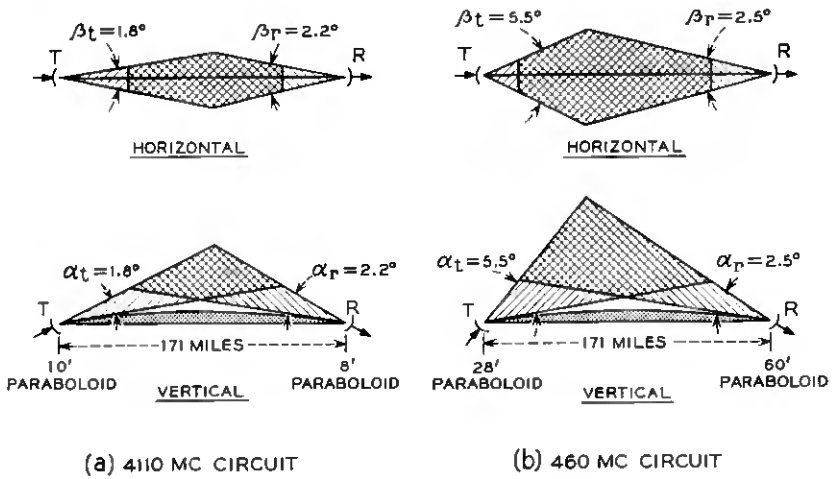


Fig. 15 — Antenna beamwidths and common volumes for wavelength dependence study: (a) 4110 mc; (b) 460 mc.

the corresponding month in 1957 and plotted as one month. The available data for the months of January through April 1957 were sparse. A seasonal variation in the received level similar to that noted in the previous section is evident in Fig. 16.

The weekend (60-hour) medians of received power were normalized to the power that would be received under free-space conditions for both wavelengths and cumulative distributions were made. These data are shown in Fig. 17. The median beyond-the-horizon transmission loss is 61.5 db at 460 mc and 69.8 db at 4110 mc. The long-term median received power relative to free-space transmission is 8.3 db greater at 460 mc than at 4110 mc.

3.2.2 *Variations in Wavelength Dependence — Comparison with Theory*

The wavelength dependence was found to vary considerably from day to day. In order to determine the variability of the wavelength dependence, the weekly median beyond-the-horizon transmission loss in excess of the free-space loss at 460 and 4110 mc was expressed as a ratio. A cumulative distribution curve of these ratios is shown in Fig. 18.* The ratio, T_{460}/T_{4110}, is the relative transmission coefficient at 460 and 4110 mc. It should be emphasized that, since the antenna beams used

* Bolgiano[11] has found a similar wavelength dependence using Lincoln Laboratory data.

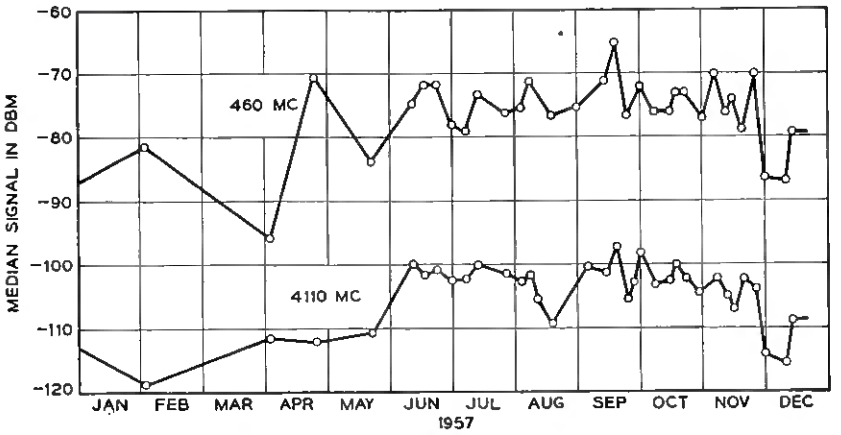


Fig. 16. — Median received signal levels for one year at 460 and 4110 mc. At 460 mc the 28-foot transmitting antenna was used with a power input of 57 dbm, signals were received on the 60-foot antenna. At 4110 mc, the 10-foot transmitting antenna was used with a power input of 43 dbm, signals were received on the 8-foot antenna. The data have not been adjusted for antenna line losses (see Table I). Period: April 1956 through December 1957.

were large, it is believed that this distribution is a true representation of the effect of the atmosphere on the propagation at the two widely

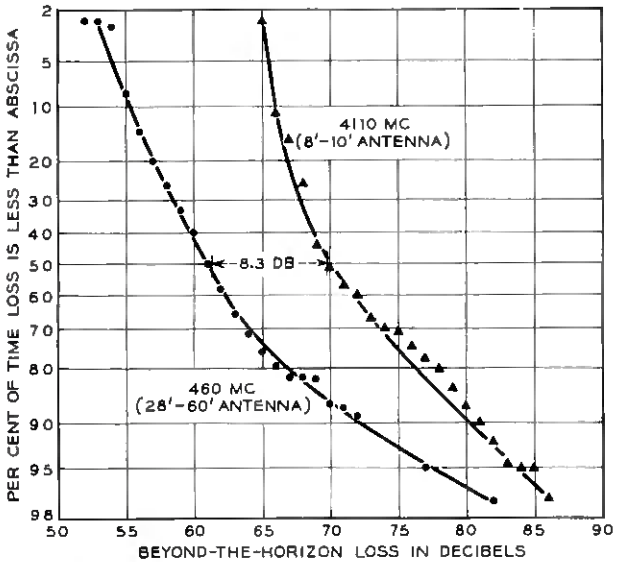


Fig. 17 — Distribution of the beyond-the-horizon transmission loss in excess of the free-space loss at 460 and 4110 mc. Period: April 1956 through December 1957.

separated wavelengths. Since the distributions were obtained using the 60-hour weekend medians instead of the 15-minute medians, an abrupt change of slope occurs in the extreme values.

With the aid of Fig. 18 we see that the median value of the relative transmission loss is 8.3 db. At the 95 per cent point on the distribution the 460-mc signal is about 19 db greater than the 4110-mc signal. At the 2 per cent point the received signal levels at both frequencies are about equal (all relative to free space).

Wavelength dependence has also been considered in the theory. In the last section we were concerned with the antenna sizes at one frequency; here we keep the beamwidths constant and vary the frequency. Using (1) one obtains

$$\frac{P_{R_1}/P_{R_2}\text{ (beyond horizon)}}{P_{R_1}/P_{R_2}\text{ (free space)}} = \left(\frac{\lambda_1}{\lambda_2}\right)^1,$$

when we have intermediate size layers (Case 3) (see Ref. 5, p. 637, Equation 17). For large layers (Case 1) this ratio becomes

$$\frac{P_{R_1}/P_{R_2}\text{ (beyond horizon)}}{P_{R_1}/P_{R_2}\text{ (free space)}} = \left(\frac{\lambda_1}{\lambda_2}\right)^2,$$

while for small layers (Case 2) we have

$$\frac{P_{R_1}/P_{R_2}\text{ (beyond horizon)}}{P_{R_1}/P_{R_2}\text{ (free space)}} = \left(\frac{\lambda_1}{\lambda_2}\right)^0.$$

For large layers in the atmosphere (Case 1), the wavelength dependence is 19.0 db for the frequencies used; if layers in the atmosphere are small (Case 2) the wavelength dependence is 0 db. For intermediate size layers (Case 3) we get 9.5 db, which compares favorably with the measured value of 8.3 db.

IV. CHARACTERISTICS OF THE SIGNALS RECEIVED ON A LARGE APERTURE ANTENNA

In order to learn more about the mechanism of the propagation, an experiment was carried out using the scanning feature of the 60-foot paraboloid antenna.[6] The experiment was designed to investigate the variation in received power as the beam of the antenna was scanned in azimuth and elevation. A further aim was to look for relationships between the scanning data and the median signal levels received simultaneously on the 60-foot antenna and on the 8-foot comparison antenna. A discussion of the aiming of narrow-beam antennas as it is affected by

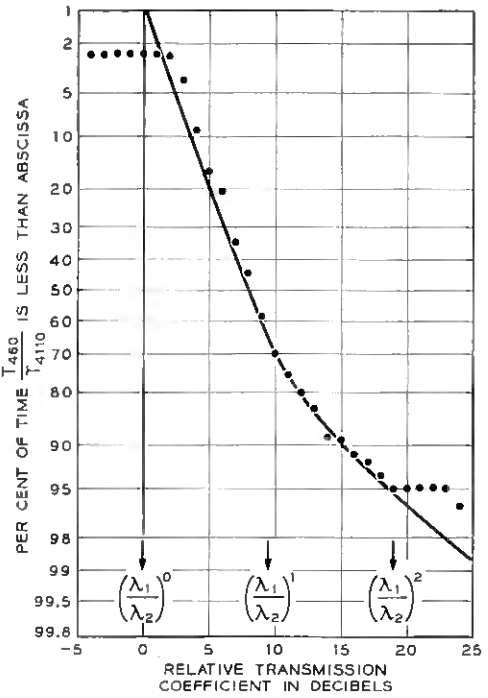


Fig. 18 — Disitrbution of the wavelength dependence of the relative transmission coefficient at 460 and 4110 mc. Period: April 1956 through December 1957.

foreground terrain and local atmospheric conditions concludes the section.

4.1 *Beam Swinging Experiment*

In this experiment the 10-foot transmitting antenna and the 4-kmc transmitter are used; Fig. 19 shows the setup at the receiving end. A dc voltage which is proportional to the angle of rotation of the 60-foot antenna is derived from potentiometers and fed to the X-axis of an oscillograph. The signal from the receiver is placed on the Y-axis and is recorded by using a Polaroid camera attached to the oscillograph. The total scan is 2.8° in azimuth and 3.0° in elevation. The period of a single scan is 15 seconds. Since the period of the short-term fading can be an order of magnitude smaller than the scan period, an integration method is used to average out the time variation of the received signal.*

* A rapid beam-swinging experiment has recently been performed by Waterman.[12]

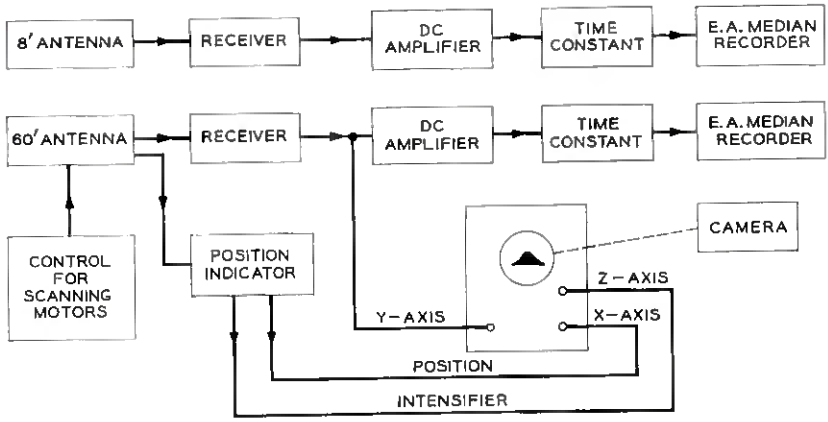


Fig. 19 — Recording setup for beam-swinging experiment.

This is accomplished by superposing 12 scans on a single photographic exposure. During this interval the type of fading as well as the median signal level remain fairly constant.

The median signal levels received simultaneously on the 60-foot antenna and on an 8-foot comparison antenna are recorded on Esterline-Angus recorders; the time constant of the recorder circuits is 25 seconds. The 60-foot antenna is aimed for maximum signal for recording the median level.

4.1.1 *Examples of Beam-Broadening*

Figs. 20 and 21 show sample scans taken with the 60-foot antenna at 4.11 kmc; these were chosen to illustrate interesting characteristics of the received signal. The angles, δ_{10} and γ_{10}, represent the width of the patterns in degrees at the 10-db points for horizontal and vertical scans respectively; they are obtained by drawing an envelope through the maxima of the photographed signal and measuring the width of the envelope at points 10 db below the maximum signal level. The zero on the angle scale represents the same antenna orientation for all the pictures; it is otherwise arbitrary.

Fig. 20(a) shows azimuth and elevation scans on the line-of-sight Murray Hill-Crawford Hill path. It will be noted that there is some asymmetry in the patterns of Fig. 20(a); this is because the feed horn used was one of an assembly of horns used for other experiments. Thus, the gain and beamwidth were not optimum.

Fig. 20(b) shows a rather unusual transmission phenomenon. It is

characterized by a very high median received signal level (about 35 db higher than that usually observed) and by very slow fading (on the order of one fade per minute). At this time the free-space beamwidth and gain of the large antenna were realized. These observations point to the existence of a large stratified layer in the region common to the beams of the transmitting and receiving antennas.

Fig. 20(c) has a pattern width slightly broader than the free-space

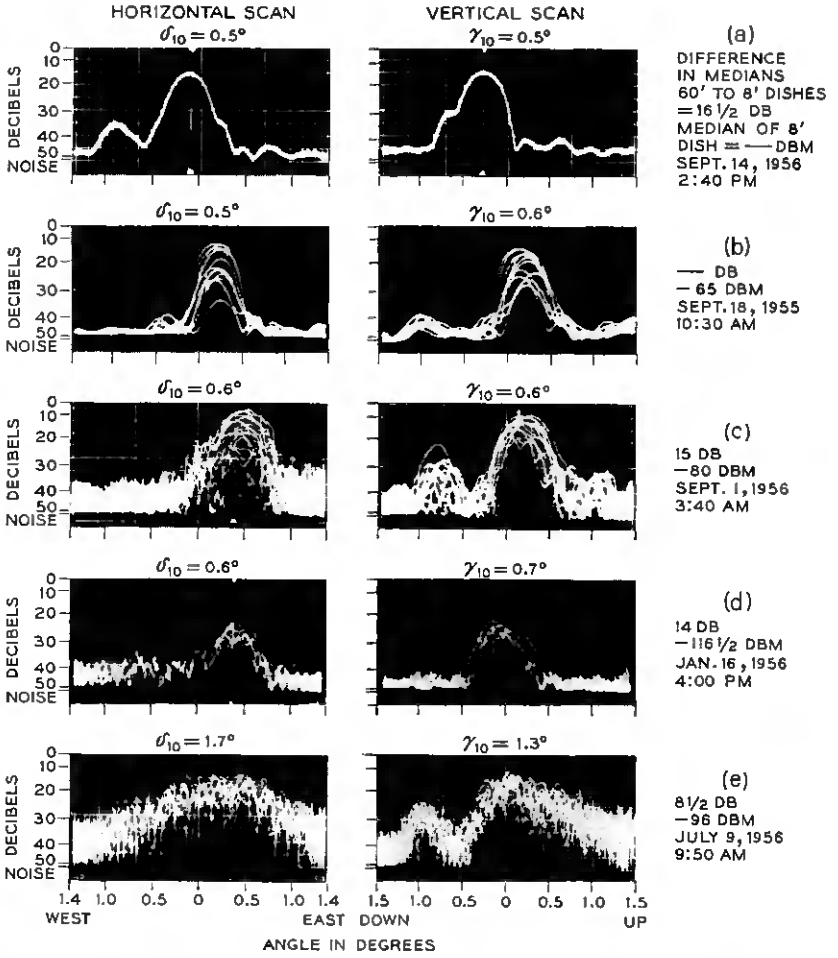


Fig. 20 — Sample scans of signals received on the 60-foot antenna at 4110 mc. The zero on the angle scale represents the same antenna orientation for all pictures; it is otherwise arbitrary.

beamwidth, but it is characterized by more rapid fading than in Fig. 20(b). Fig. 20(d) shows that it is possible to have very little beam-broadening (near free-space characteristic) for low median signal levels.

Fig. 20(e) is an example of the average beam-broadening observed on this circuit. Two features of the vertical scan are worth noting: the linear decrease in received power (in db) as the angle of elevation of the antenna is increased from 0° to 1.5° is typical; the lobe near 1° down is due to reflection from the foreground of the receiving site and will be discussed in more detail later.

Fig. 21(a) shows a very broad pattern; the received power is essentially independent of the azimuth setting of the antenna. We believe that this condition is due to numerous reflecting layers of small dimension in the atmosphere.

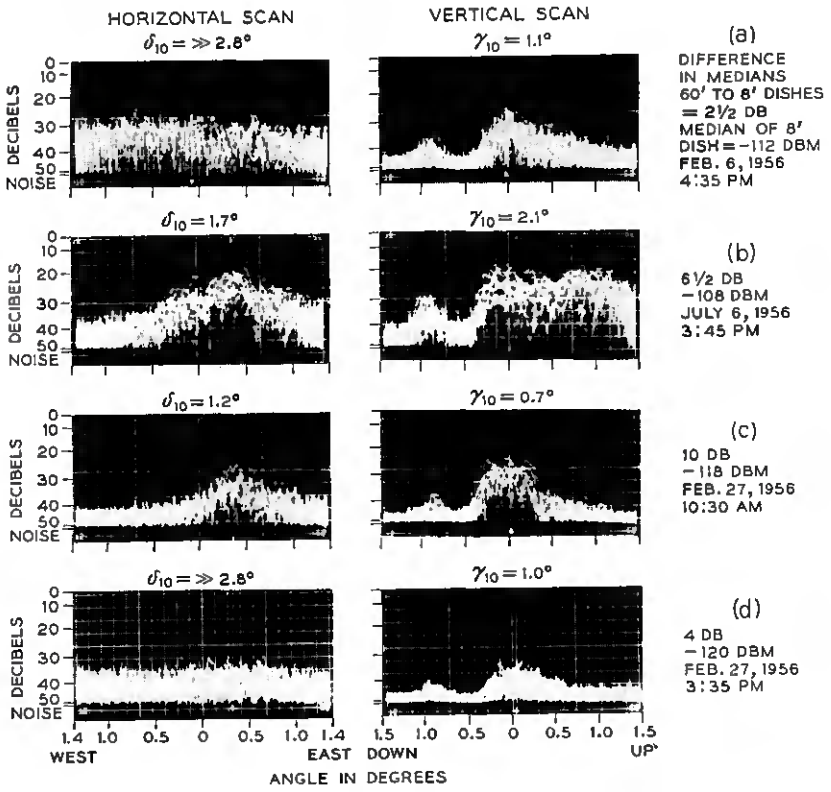


Fig. 21 — Additional sample scans of signals received on the 60-foot antenna at 4110 mc.

The vertical scan in Fig. 21(b) indicates the existence of two strongly reflecting regions in the atmosphere at heights corresponding to 3000 and 11,000 feet above the ground at midpath.

The rapidity with which atmospheric conditions can change is illustrated by comparing Figs. 21(c) and 21(d), which were taken on the same day five hours apart. The horizontal beam-broadening in Fig. 21(d) exceeds that in Fig. 21(c) by about three times, yet the median signal level is the same in the two cases. The weather map for that day indicates a front moving perpendicularly to the transmission path from the west.

The range and rate of fading at the edges of the pictures are greater than at the center. This is caused by two factors: the scanning motion of the antenna is sinusoidal, so that the antenna beam spends more time at the edges of the scan than in the middle, hence both the exposure time and the probability of a high signal are increased at the edges; and the fading rate off the great circle route is greater due to longer path delays.

4.1.2 *Average Scanning Patterns — Comparison with Theory*

To find the average beam-broadening, 30 sets of pictures, similar to those in Figs. 20 and 21 were taken over a period of several months; the envelopes of these were normalized and then averaged. In other words, the average envelope of received power versus angle was obtained. Fig. 22 shows the results of this analysis. The free-space patterns are shown for comparison.

The composite horizontal scan [Fig. 22(a)] shows that, on the average, most of the power directed toward the receiver is contained within an angle of one degree (measured at the 3-db points) in the horizontal plane. Thus the free-space pattern is broadened by a factor of three. There is definite asymmetry in this composite pattern; this is due to the fact that some of the horizontal scan patterns peaked between 0 and 0.5° west rather than at the dead-reckoning position (0.35° east). The cause of this anomalous condition is unknown at present. Large horizontal beam-broadening and rapid fading were present when the shift was evident [see Fig. 21(a)]. When the anomaly was present, a check on the orientation of the antenna was made using the Murray Hill transmitter, and it was determined that the shift was not due to deformation of the antenna itself.

Fig. 22(a) shows the measured azimuth position of Murray Hill with respect to the horizontal scan. Calculation of the initial course angles (dead reckoning) between Crawford Hill and Murray Hill and between

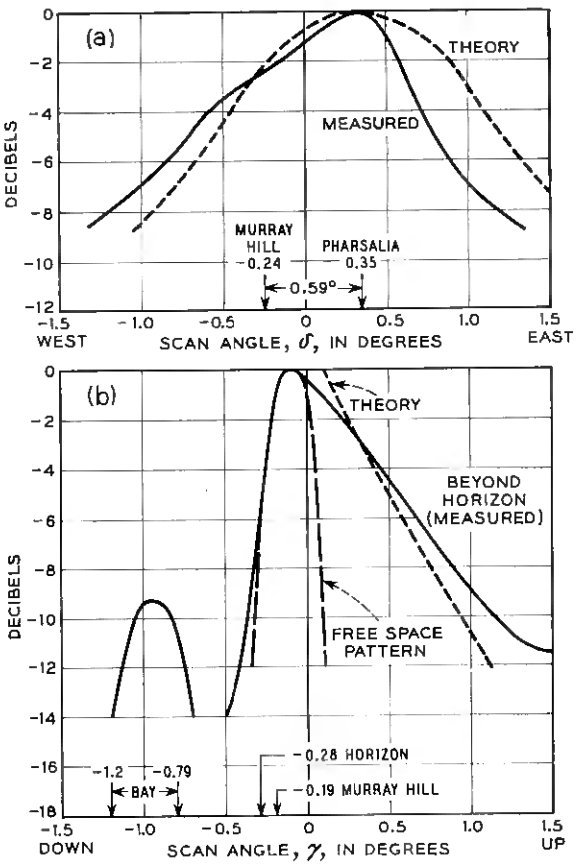


Fig. 22 — Average (a) horizontal and (b) vertical scanning patterns. The experimental curves represent a composite of 30 scans. The theoretical curves are calculated from the theory, using intermediate-size layers.

Crawford Hill and Pharsalia gives a difference in azimuth of 0.59°. This angle also is laid out on the figure; it will be noted that peak signal on the composite scan coincides with the calculated position of Pharsalia.

The composite vertical scan is shown in Fig. 22(b). Reflection from Raritan Bay (shown on the foreground profile in Fig. 3) accounts for the strong lobe which appears when the antenna beam is pointed below the horizon. Using this profile, the position of the bay and other information were calculated and superimposed on the angular scale. It is interesting to note that the drop-off in power as the antenna scans below Murray Hill closely follows the free-space pattern. The power directed

toward the receiver is contained within an angle of 0.6° (measured at the 3-db points) in the vertical plane. A comparison with the free space pattern shows that a three-to-one broadening occurs.* The average curve shows that the received power (in db) falls off rather linearly as the angle of elevation of the antenna is increased from the horizon to higher elevations. If one converts elevation angle to height at the center of the path, one finds that the power fall-off between 3500 feet above sea level (−0.1° on the scale) and 13,500 feet (+1.1°) is 10 db; thus the average rate of fall-off of received power with height is about 1 db per thousand feet.

Using Fig. 78 in the Appendix, one can calculate the beam-broadening of an antenna of given beamwidth. The free-space antenna pattern used in the theory is wedge-shaped, so that the theoretical curve for the vertical beam-broadened pattern rises very steeply for angles near the horizon. The theoretical patterns are shown in Fig. 22; the vertical pattern is normalized with the experimental one at the 3-db point.

The theoretical curves plotted are those for the intermediate size layers (theory, case 3). If one considers the atmosphere to be composed of several large horizontal layers, then one would expect the beamwidth at the receiver to approximate that of free space; this is consistent with the result shown in Fig. 20(b). On the other hand, if one considers several large layers which may be tilted with respect to the horizontal plane (theory, case 1) then a broadened pattern similar to that of Fig. 21(c) is obtained. Furthermore, if small layers (theory, case 2) which are uniformly distributed throughout the atmosphere are considered, a pattern broader than that of Fig. 22(a) is obtained. However, the theory in its present state does not account adequately for very broad patterns,† such as Fig. 21(b).

4.1.3 *Beam-Broadening and the Ratio of the Power Received on the 60-Foot and 8-Foot Antennas*

Since the beam-broadening of the 60-foot antenna varied so much from day to day, as shown by the data just discussed, it seemed desirable to try to obtain a quantitative relation between the beam-broadening and some other parameter such as median signal level. The following experimental procedure was adopted: the horizontal and vertical scans were taken and the 10-db widths of the patterns (δ_{10} and γ_{10}) were

* The beamwidth measured from the peak to the half-power point above the horizon is about 0.4°.

† The very broad patterns might be explained with the assumption of numerous isotropic reradiators.

measured from the photographs. Both before and after the scans, the median signal levels on the 60-foot antenna and on the 8-foot antenna were obtained from the Esterline-Angus recorders. From theoretical considerations, it was believed that the product of the 10-db widths in the two planes and the ratio of the power levels on the large and small antennas were the significant parameters.

In Fig. 23 observed values of the ratio of power on the two antennas are plotted versus measured inverse products of the 10-db pattern widths. The points fit the curve that is discussed in the next paragraph.

Consider antennas as used in a line-of-sight path with plane waves incident upon them. Then the ratio of the power received in free space by a paraboloid of a certain aperture relative to that received by an 8-foot paraboloid can be plotted as a function of the inverse product of the free-space horizontal and vertical beamwidths of that paraboloid as its aperture size is increased, as shown by the curve of Fig. 23. The end points of the curve correspond to the measured gain and beamwidth of the 60-foot and 8-foot antennas; the curve is normalized to give a gain of 0 db for the 8-foot antenna. The ratio of the gains of the 60-foot and the 8-foot is 16.5 db; the 10-db beamwidths are 0.5° and 2.8° respectively.

Consider now two antennas, a 60-foot and an 8-foot paraboloid, used

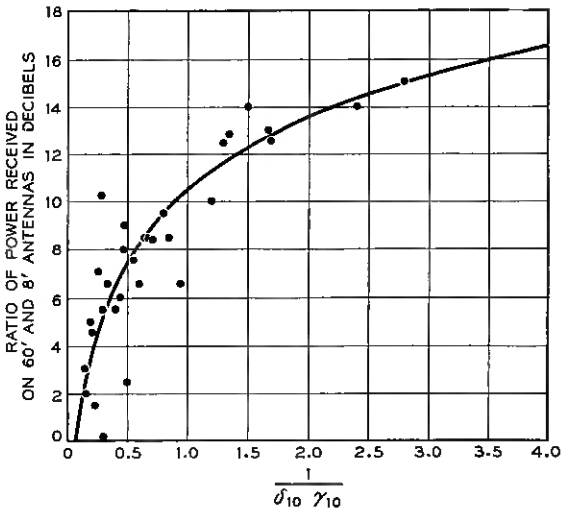


Fig. 23 — Ratio of median signal levels received simultaneously on 60-foot and 8-foot antennas as a function of the inverse product of the 10-db beamwidths; δ_{10} and γ_{10} are the beamwidths in degrees for the horizontal and vertical scans, respectively. Points are experimental; curve is theoretical.

in a beyond-the-horizon propagation test. This experiment is a duplicate of the one on the line of sight, except that antenna size is not a variable. The ratio of the power received on the two antennas is measured simultaneously with the horizontal and vertical beam patterns of the 60-foot antenna, and this ratio is plotted against the inverse product of the 10-db widths of the measured patterns as discussed previously.

Since the experimental points in Fig. 23 lie close to the theoretical curve, it is concluded that the beam pattern and received power of the 60-foot antenna in a beyond-the-horizon circuit change from day to day in a manner corresponding to an antenna of changing size in line-of-sight propagation. The ratio of the power received on the 60-foot and 8-foot antennas in the beyond-the-horizon path is less than that in free space, because the incident energy over the face of the 60-foot antenna has components which are not of equal amplitude nor of equal phase. It is believed that the smaller ratio results from an atmosphere which is made up of layers of small dimension. The 8-foot antenna was used as a reference in this experiment, since it is believed that the received field is correlated over its aperture most of the time.

There is some scatter of the experimental points in Fig. 23 because of the difficulty of measuring the angular widths of the beam patterns, especially if the width is large. If it is greater than 3°, an extrapolation has to be made.

From a year's recording of the power received on the 60-foot and 8-foot antennas (Section 3.1.2) we have found that the average difference in medians is 5.7 db. This corresponds to an average pattern width of 1.8° at the 10-db points compared with the free-space pattern of 0.5°.

4.2 *Antenna Aiming*

As can be seen from the previous section, a problem arises as to how to aim a high-gain antenna (3-db beamwidth $<1°$) in a beyond-the-horizon circuit. The previous section showed that, on the average, the maximum signal level is received when the antennas are aimed along the great circle between transmitter and receiver location, with each being aimed just above the horizon. However, local atmospheric refraction can exert considerable influence on the signals received on narrow beam antennas.

An example of what is believed to be the effect of local meteorological conditions is shown in Fig. 24 for June 8, 1956. In the early part of the evening the atmospheric conditions at the receiving site were quite stable. During the evening the humidity rose slowly, as plotted in the figure, and at about 11:00 p.m. the winds increased considerably. As

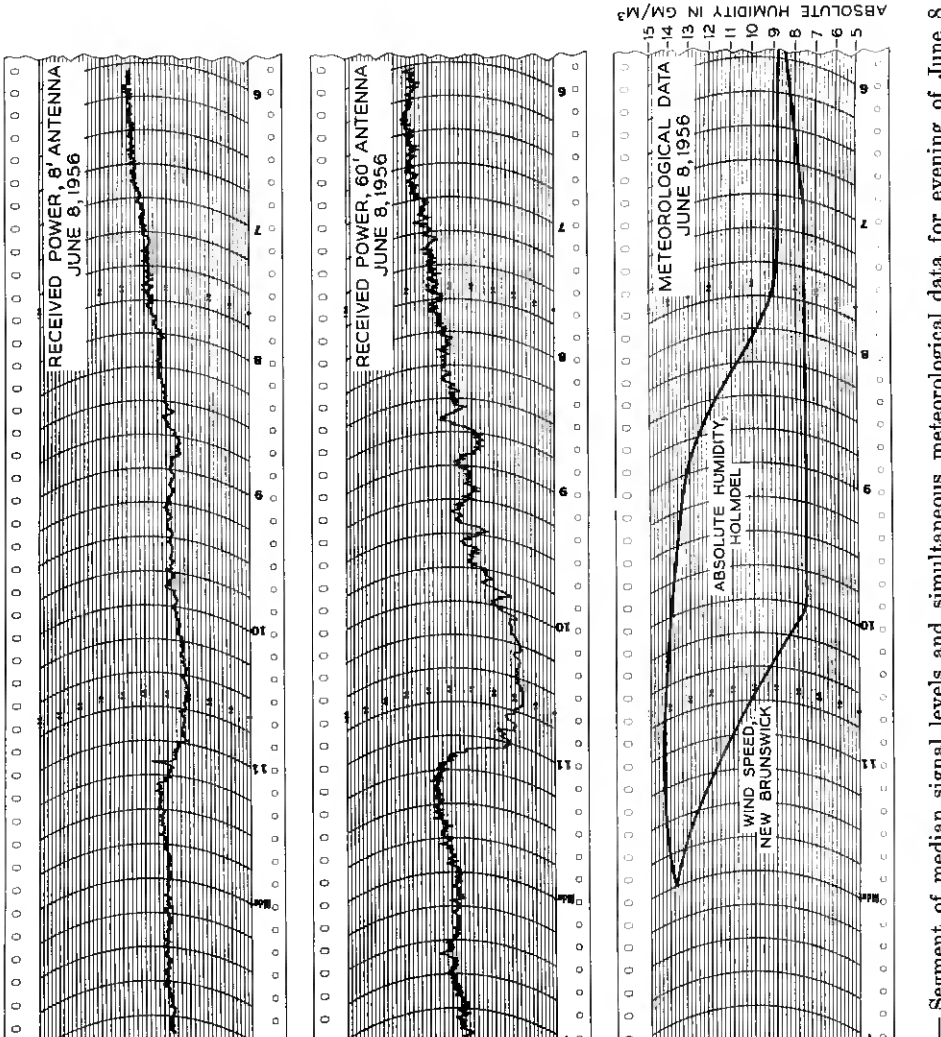


Fig. 24 — Segment of median signal levels and simultaneous meteorological data for evening of June 8, 1956. Full scale of records is −76 dbm. (For other details see caption for Fig. 8.)

observed between the hours of 6:00 p.m. and 11:00 p.m., the power level of the signal received on the 60-foot antenna at 4 kmc slowly decreased some 20 db, whereas the signal on the 8-foot antenna decreased only about 9 db; and then the signal levels of both antennas suddenly returned to their normal levels at 11:00 p.m. This effect is believed to have been caused by increased local bending of the radio waves during the period from 6:00 p.m. to 11:00 p.m. followed by a break-up in this stable condition by the winds at 11:00 p.m.

Fig. 7 illustrates the local refraction which is believed to have taken place. The narrow beam of the 60-foot antenna, as indicated by the broken lines in the figure, is bent some 0.3 degrees into Murray Hill in the foreground of the receiving site; thus the signal level on this antenna decreases considerably. On the other hand, the beam of the 8-foot antenna, although bent, is not completely shadowed and the decrease in signal level is not so pronounced.

Bending of this order of magnitude was frequently observed on the line-of-sight path between Crawford Hill and Murray Hill during the course of angle-of-arrival measurements. Similarly, in these measurements it was observed that the angle of arrival would return to its normal value when winds broke up a stable atmosphere.

This local bending may affect the median signal level over extended periods, as is shown by the data on March 9–12, 1956, and April 6–9, 1956 (Fig. 11). During these weekends the 60-foot antenna was oriented to receive the maximum signal level at the beginning of the period of recording. However, during the run the direction of the received energy changed, thus reducing the median signal level on the narrow beam antenna. The weekend median signal level received on the 60-foot antenna was about the same as that on the 8-foot antenna. The cumulative effect of the local bending undoubtedly decreases the ratio of the power received by the 60-foot and 8-foot antennas, hence the measured 5.7 db yearly difference includes both loss of gain and local bending effects of the 60-foot antenna.

The orientation of the antenna will depend on the beamwidth as well as on the foreground. If one assumes a flat foreground with unity reflection coefficient, then aiming at the horizon will be best for maximum median signal level. If the foreground is not flat, or does not have unity reflection coefficient, an aim slightly above the horizon is preferable. (We are considering a very narrow beamwidth antenna situated at ground level.) For large beamwidth antennas ($>2°$) the aim is not as critical; an aim toward the horizon is satisfactory.*

* Throughout the paper we have used the conventional common volume dia-

V. RATE OF FADING OF RECEIVED SIGNALS

Part of the program of study has been the observation and interpretation of the short-term fading of the received signal. In the first experiment to be described, a 4-kmc cw signal is transmitted; the received power is found to vary in noise-like fashion at a rate of a few cycles per second. The rate of fading varies considerably from day to day and, at times, from hour to hour, due to changes in the atmosphere. The rate is found to depend upon the beamwidths of the transmitting and receiving antennas. Data taken at various times throughout the year have been analyzed and a quantitative relation for fading rate has been developed which is consistent with the experimental results.

A further study is described using simultaneous transmission of two frequencies, 460 and 4110 mc. Here the wavelength dependence of the fading and its relation to the characteristics of the atmosphere is discussed. Antennas of comparable beamwidth are used.

5.1 *Rate of Fading at 4 kmc — Relation to Antenna Beamwidth*

The basic measuring circuit consists of a 10-foot paraboloid transmitting at 4110 mc and two receiving antennas — a 60-foot paraboloid and an 8-foot paraboloid. The outputs of the antennas are recorded separately and the simultaneous short-term fading is examined on a percentage-of-time basis to obtain cumulative probability distributions. Also, the median signal levels are drawn on the records and the fading rates (the number of one-way crossings of the median levels per second) are determined.

From these data one can obtain significant ratios in addition to the absolute fading rate: for instance, the ratio of the fading rate of the signal received on an 8-foot antenna from a 10-foot transmitting antenna, $N_{m(10\text{-}8)}$, to the fading rate observed simultaneously on a 60-foot antenna, $N_{m(10\text{-}60)}$. This ratio $K = N_{m(10\text{-}8)}/N_{m(10\text{-}60)}$ has turned out to be sensibly independent of the absolute fading rates.

On the assumption that propagation beyond the horizon is due to reflections from randomly positioned layers which drift through the atmosphere, a simple formula which predicts the fading ratio, K, for any given combination of antennas is developed. Values of K obtained experimentally by using several combinations of transmitting and receiving antennas check well with the formula.

grams which show idealized antenna beams with the lower *edge* of the beam along the horizon. It is clear from our vertical beam-swinging experiments that for an actual antenna pattern an aim along the horizon results in maximum received power.

From these investigations it has been possible to estimate the effective horizontal scattering angle of the atmosphere for the 171-mile path at 4 kmc. The scattering angle is a measure of the directional properties of the reflections from the layers in the atmosphere. From the data, this angle has turned out to be 0.8 degree, which is in good agreement with the average horizontal beam-broadening of 1 degree discussed in Section IV.

The absolute value of the fading rate for given transmitting and receiving antennas can be estimated with fair accuracy if the winds in the upper atmosphere are known in the region near the center of the path of propagation. Fading rates calculated from known wind data compare favorably with measured fading rates.

5.1.1 *Instantaneous Signal Level Recordings — Analysis*

The horizontal and vertical half-power beamwidths of the antennas used in the basic measuring circuit are shown in Fig. 25. The approximate volume of the atmosphere common to the beams of the 10-foot–8-foot antenna combination is shown by the shaded areas in the figure. The cross-hatched areas show the smaller volume utilized by the 10-foot–60-foot antenna combination. Approximate dimensions are scaled on the left of the sketch.

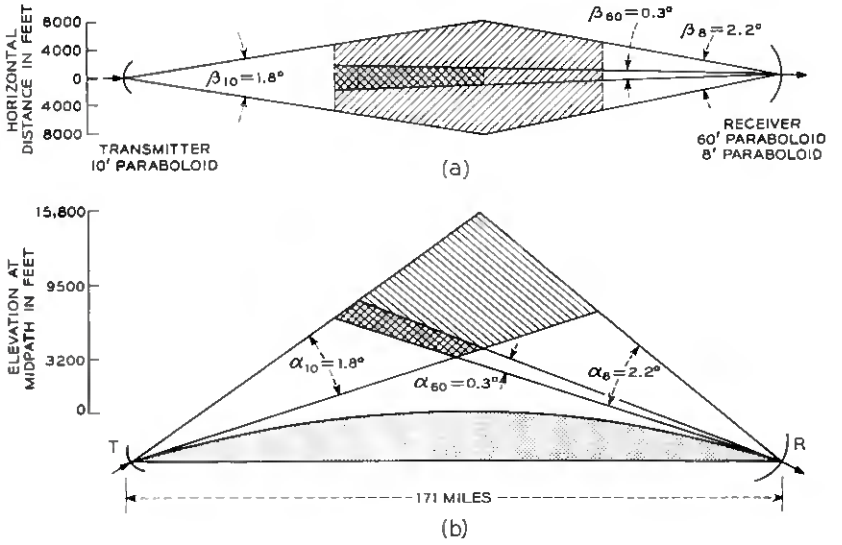


Fig. 25 — Antenna beamwidths and common volumes for fading rate studies at 4110 mc: (a) horizontal; (b) vertical.

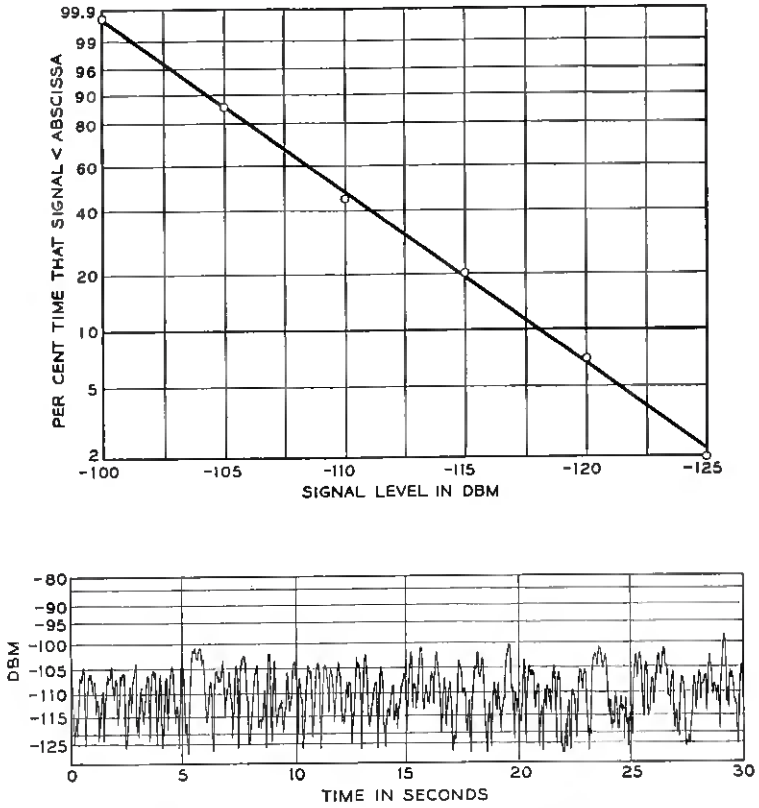


Fig. 26 — Sample and distribution of instantaneous signal levels — very rapid fading at 4110 mc.

The fluctuations in the received signal are recorded on a Sanborn recorder with a time constant of 0.015 second. It is believed that no fluctuations in the received signal were too rapid to be followed by this instrument and, therefore, that the records are a true representation of the received signal.

Sample records of the signal received on the 60-foot antenna, chosen to show the variety of fading rates encountered, are shown in Figs. 26 and 27, with the distribution for each record plotted above the sample. Analysis of Fig. 26, a sample of very rapid fading, indicates a Rayleigh distribution of the signal. Fairly slow fading is shown in Fig. 27 with its distribution curve; within the accuracy of measurement, the Rayleigh distribution is obtained. The Rayleigh distribution indicates that the

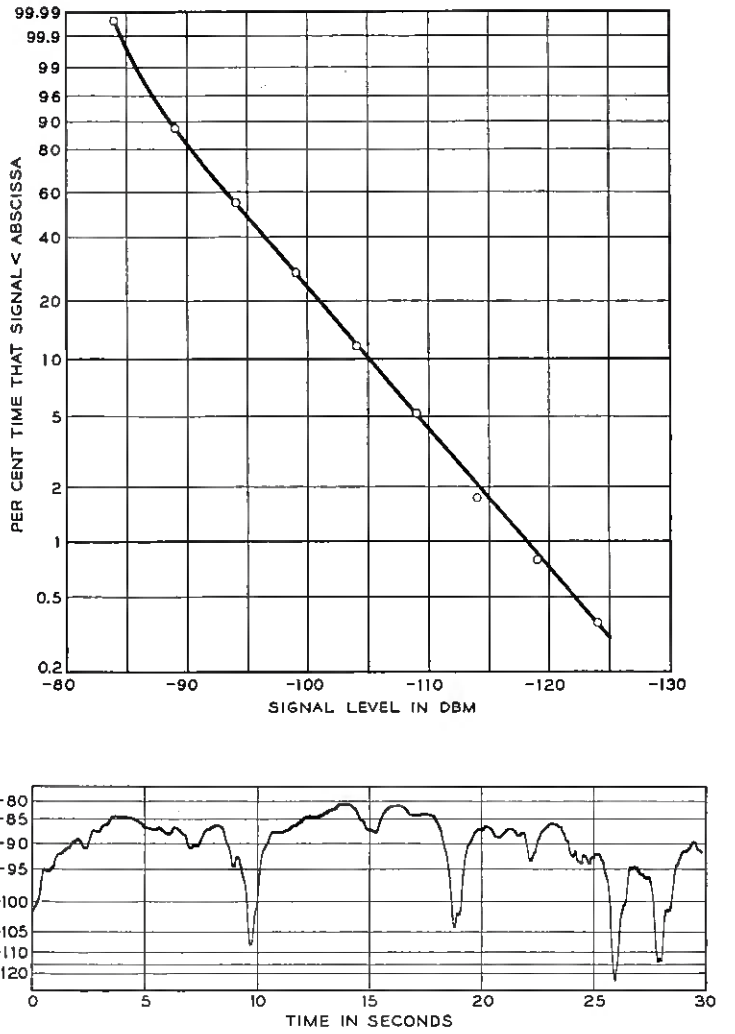


Fig. 27 — Sample and distribution of instantaneous signal levels — slow fading at 4110 mc.

received signal is due to propagation through the atmosphere via many paths of random relative phase.

As previously stated, the rate of fading varies considerably from day to day, and often within the hour. Samples of very rapid, rapid and slow fading signals recorded during one day are shown on Fig. 28. The output

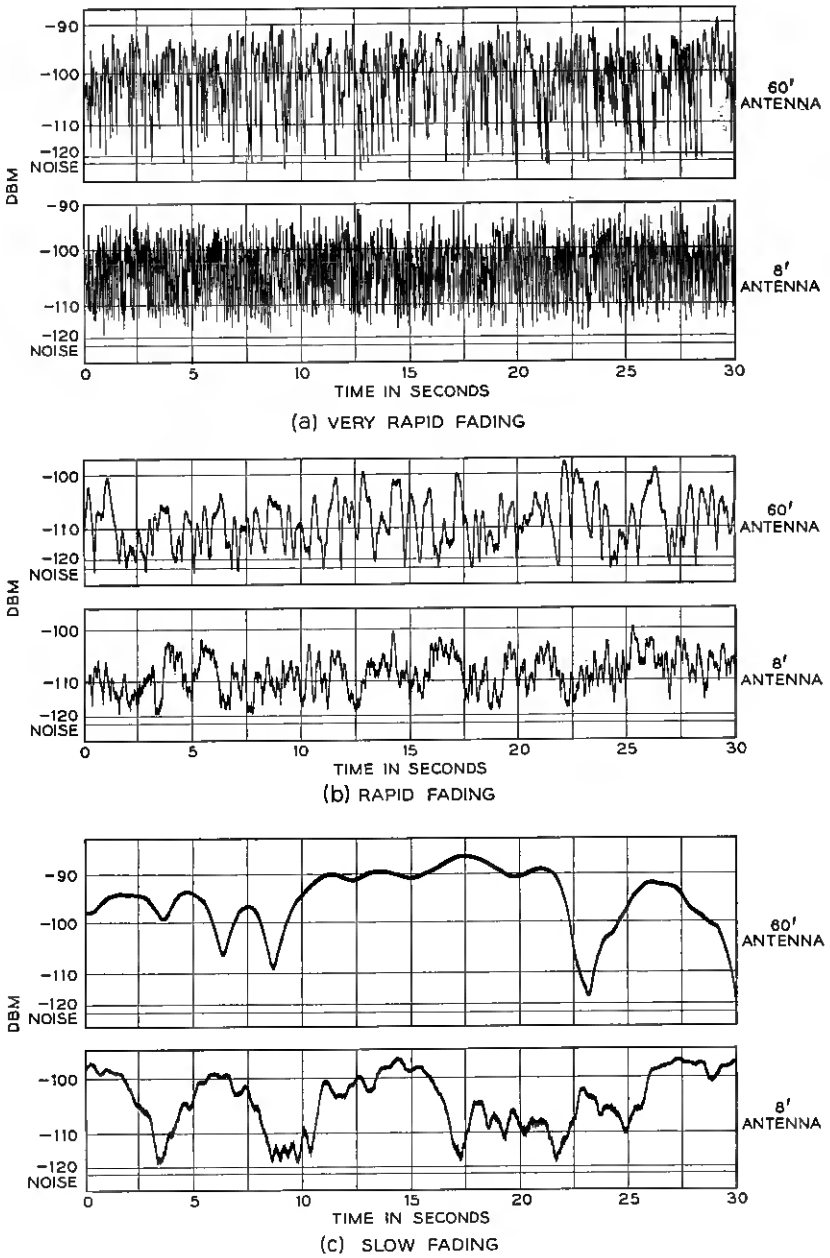


Fig. 28 — Samples of simultaneous recordings of instantaneous signal levels on the 60-foot and 8-foot antennas at 4110 mc.

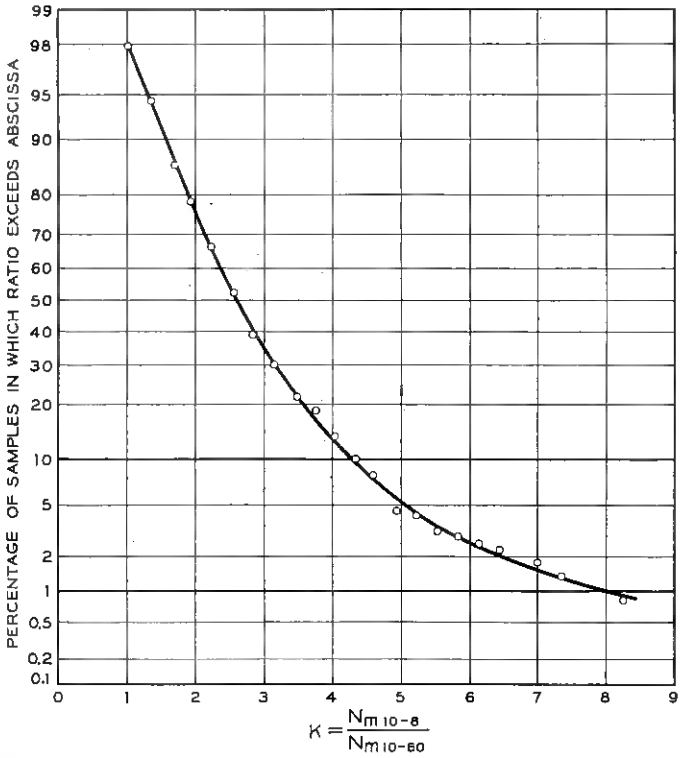


Fig. 29 — Statistical distribution of the ratio of fading rates of the 10-foot–8-foot antenna combination to the 10-foot–60-foot combination at 4110 mc, using 392 samples of 33 seconds in length. Periods: October 1955 through April 1956 and September through November 1957.

of the 60-foot antenna is shown on the top trace of each sample, the output of the 8-foot antenna on the bottom trace. In analyzing the data, the estimated median signal level is drawn on the record and the number of one-way crossings of the median levels is counted; from these, the fading rate and the ratio of the fading rates are derived.

The range of values over which the ratio of the fading rates varies may be seen on the plot of Fig. 29. These data were obtained from a study of 392 sample records taken during the period October 1955 to April 1956 and September to November 1957. The median of the ratios of the fading rate of the signal received on the 8-foot antenna to the fading rate of the signal received on the 60-foot antenna is 2.60. The ratio is well-defined, the spread of values being small.

Table III lists all the antenna combinations used during the measure-

TABLE III — MEASURED AND CALCULATED FADING RATE RATIOS FOR VARIOUS ANTENNA COMBINATIONS — 4 KMC

Transmitting Antenna, feet	Receiving Antennas, feet	K	Number of Samples	K = Ratio of Fading Rates	
				Measured	Calculated
10	8 60	$\frac{N_{m(10\text{-}8)}}{N_{m(10\text{-}60)}}$	392	2.60	—
10	8 28	$\frac{N_{m(10\text{-}8)}}{N_{m(10\text{-}28)}}$	98	1.61	1.47
10	28 60	$\frac{N_{m(10\text{-}28)}}{N_{m(10\text{-}60)}}$	181	1.62	1.67
28	8 60	$\frac{N_{m(28\text{-}8)}}{N_{m(28\text{-}60)}}$	61	1.85	1.86
28	8 28	$\frac{N_{m(28\text{-}8)}}{N_{m(28\text{-}28)}}$	36	1.18*	1.26
28	28 60	$\frac{N_{m(28\text{-}28)}}{N_{m(28\text{-}60)}}$	37	1.65*	1.47

* Limited data, not necessarily typical.

ments, the ratio of fading rates obtained using the various combinations and the number of sample records analyzed.

5.1.2 *Rate of Fading — Theory*

On the assumption that the short-term fading is due mainly to horizontal drift of many randomly positioned layers in the atmosphere which reflect the radio waves, several simple conclusions can be drawn regarding the relation of antenna size and winds to fading rate. The purpose here is to try to obtain a method for calculating the fading rate from known wind data and known antenna beamwidths.

Fig. 30(a) shows a horizontal plan view of a tropospheric circuit in which three rays have been drawn from the transmitting antenna T to reflecting layers A, B and C and thence to the receiving antenna R. The layers are assumed to drift with velocity V at angle δ to the line T-R; the wind component normal to the great circle route is therefore $V \sin \delta$. The length of the path via layer B in excess of the great circle route, TAR, is TBR $-$ TAR $= 2x^2/2a$, where x is the distance from layer A to layer B and $2a$ is the distance from T to R.* The phase difference between

* Disposition of the layers in the vertical plane causes path differences larger than this, but, since horizontal winds are much stronger than vertical winds, fading due to vertical motion is not considered here (see Section 6.1).

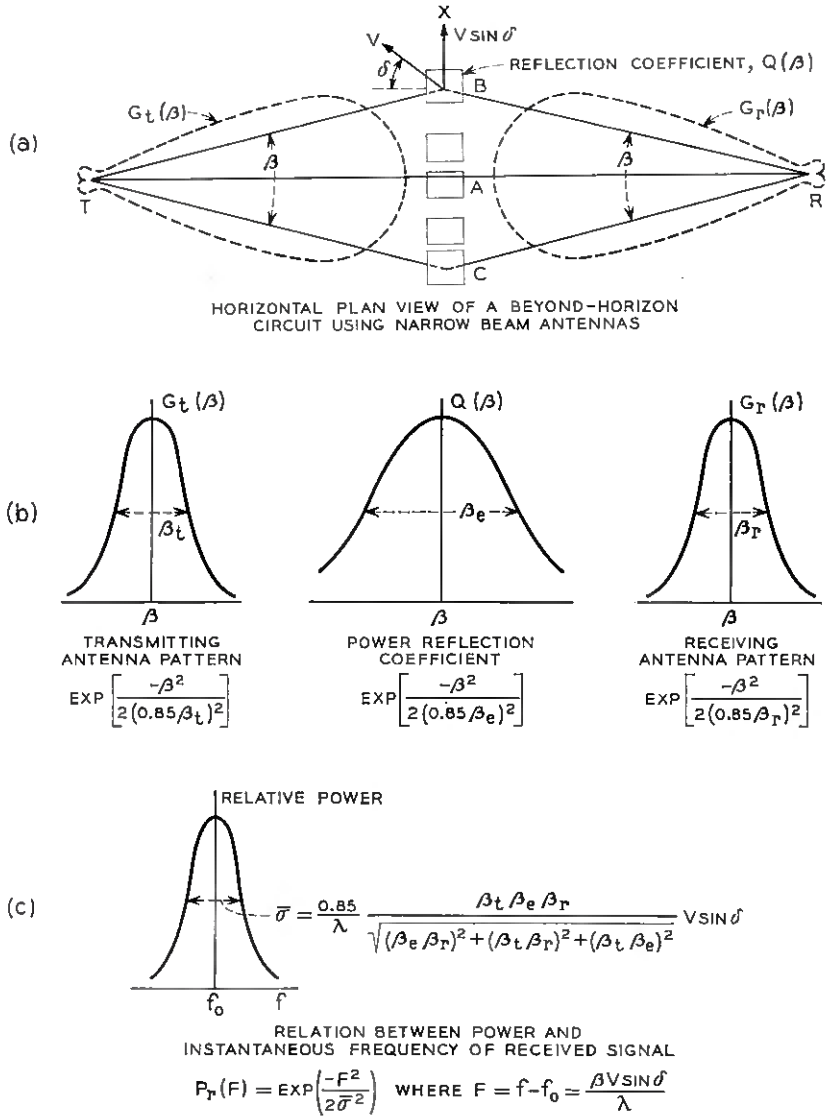


Fig. 30 — Factors affecting the instantaneous frequency of the received signal.

the two paths is $\varphi = [(2\pi/\lambda)\ (x^2/a)]$ and the frequency of fading is

$$F = f - f_0 = \frac{1}{2\pi}\frac{d\varphi}{dt} = \frac{2x}{\lambda a}\frac{dx}{dt} = \frac{\beta}{\lambda} V \sin\delta,$$

where

f_0 = transmitted frequency,
λ = transmitted wavelength,
f = instantaneous frequency, and
$\beta/2 = x/a$ = angle in the horizontal plane between great circle plane and the ray to a layer.

Thus the frequency of fading F is due to the Doppler shift caused by drift of B at velocity $V\sin\delta$ normal to the path. The rate of change of path length due to the wind component along the path is relatively small and is neglected.

The amount of power that reaches the receiver via layer B [Fig. 30(a)] depends on three factors. First, the radiation pattern of the transmitting antenna $G_t(\beta)$ controls the power transmitted in direction β, and therefore the amount of power incident on layer B. Similarly, the radiation pattern of the receiving antenna $G_r(\beta)$ determines the power received from direction β, and therefore influences the amount of power reflected by layer B that enters the receiver. Finally, the power reflection coefficient Q of the layer depends on the angle of incidence of the ray, and therefore on angle β; both experiment and theory* indicate that Q decreases as β increases. Since all three factors tend to decrease the power received from angle β as β increases, and since the frequency of fading, F, is related to β by $F = (\beta/\lambda)\ V\sin\delta$, one would expect the high-fading-frequency components to contribute only small amounts of power to the received signal. In fact, both the atmosphere and the antenna radiation patterns are low-pass filters as far as the frequencies of fading are concerned.

Consider first the effect of the reflection coefficient, Q, on the fading of the received signal. Fig. 30(a) shows numerous layers positioned at random along the x-axis; these are assumed to drift in the x-direction with velocity $V \sin \delta$. At first the transmitting and receiving antenna patterns will be assumed to be omnidirectional; thus, the only factor affecting the power received from a layer at a given angle β is the reflection coefficient Q. Since Q is assumed to be related to β in a gaussian manner, as shown in Fig. 30(b), the power reflection coefficient is

$$Q(\beta) = \exp\left[\frac{-\beta^2}{2(0.85\beta_e)^2}\right], \qquad Q(\beta = 0) = 1,$$

* See Section 4.1.

where β_e is the angle at which $Q = \frac{1}{2}$. β_e is of the order one degree for a 171-mile path at 4 kmc; a detailed discussion is given in Section 4.1.

The received signal is the sum of reflections from all of the randomly positioned layers and, since the reflections are randomly phased, the envelope of the received signal (the rectified signal) is Rayleigh-distributed. For the case of omnidirectional antennas, the distribution of received power versus frequency (the power spectrum) is directly related to the power reflection coefficient, Q. Thus the received power is distributed according to

$$P_r(F) = \exp\left(\frac{-F}{2\sigma_e^2}\right),$$

where $\sigma_e = (0.85\ \beta_e/\lambda)V\sin\delta$ is the standard deviation.* Note that $P_r(F)$ can be obtained directly from $Q(\beta)$ by substituting $F = (\beta/\lambda)V\sin\delta$ for β.

The antenna patterns are taken to be gaussian; that is,

$$G_t(\beta) = \exp\left[\frac{-\beta^2}{2(0.85\beta_t)^2}\right]$$

and

$$G_r(\beta) = \exp\left[\frac{-\beta^2}{2(0.85\beta_r)^2}\right];$$

β_t and β_r are the half-power beamwidths of the radiation patterns [Fig. 30(b)]. The power spectrum of the received signal now depends on $G_t(\beta)$ and $G_r(\beta)$ as well as on $Q(\beta)$. The power received from a layer at angle β is

$$P_r(\beta) = G_t(\beta)Q(\beta)G_r(\beta) = \exp\left[\frac{-F^2}{2}\left(\frac{1}{\sigma_t^2} + \frac{1}{\sigma_e^2} + \frac{1}{\sigma_r^2}\right)\right], \quad (5)$$

where

$$\sigma_{t,e,r} = \frac{0.85\beta_{t,e,r}}{\lambda} V\sin\delta. \quad (6)$$

If we let

$$\frac{1}{\bar{\sigma}^2} = \frac{1}{\sigma_t^2} + \frac{1}{\sigma_e^2} + \frac{1}{\sigma_r^2}, \quad (7)$$

* A formula similar to this was given by Rice[13] (see also Ref. 14). His discussion is based on the assumption of small scattering discontinuities which move in random fashion similar to the thermal motion of the molecules of a gas, and it involves angles in the vertical plane of the propagation path. In the present discussion, horizontal drift of relatively large discontinuities and angles in the horizontal plane are considered to be of major importance.

then

$$P_r(\beta) = \exp\left(\frac{-F^2}{2\bar{\sigma}^2}\right). \tag{8}$$

Substituting (6) in (7), one obtains

$$\bar{\sigma} = \frac{0.85}{\lambda}(V\sin\delta)\frac{\beta_t\beta_e\beta_r}{\sqrt{\beta_e^2\beta_r^2 + \beta_t^2\beta_r^2 + \beta_t^2\beta_e^2}}, \tag{9}$$

where $\bar{\sigma}$ is the standard deviation of the power spectrum that one would expect for the received signal in the circuit of Fig. 30(a). The antenna patterns and reflection coefficient shown in Fig. 30(b) result in the power spectrum of Fig. 30(c).

S. O. Rice has pointed out that fading in beyond-the-horizon circuits behaves like thermal noise that is passed through a filter. In both cases, the rectified signal is Rayleigh-distributed. Similarly, the width of the power spectrum of the signal (the standard deviation, $\bar{\sigma}$) in both cases is related to the fading rate of the envelope. If fading rate, N_m, of the envelope is defined as the number of one-way crossings of the median level per second (this is the definition used in analyzing the experimental data), then the theory of random noise gives $N_m = 1.5\ \bar{\sigma}$. Using the value in (9) for $\bar{\sigma}$, one obtains

$$\begin{aligned} N_m &= \frac{1.27}{\lambda}(V\sin\delta)\frac{\beta_t\beta_e\beta_r}{\sqrt{\beta_e^2\beta_r^2 + \beta_t^2\beta_r^2 + \beta_t^2\beta_e^2}} \\ &= \frac{1.27}{\lambda}(V\sin\delta)g(\beta_t, \beta_e, \beta_r) \end{aligned} \tag{10}$$

for the calculated fading rate, where

$$g(\beta_t, \beta_e, \beta_r) = \frac{\beta_t\beta_e\beta_r}{\sqrt{\beta_e^2\beta_r^2 + \beta_t^2\beta_r^2 + \beta_t^2\beta_e^2}}.$$

We are now able to compare fading rates and ratios of fading rates as calculated by means of (10) with the measured data. First consider the ratio $K = N_{m(10\text{-}8)}/N_{m(10\text{-}60)}$ of the fading rates of the 8-foot and 60-foot receiving antennas with a 10-foot transmitting antenna. Using (10), one obtains after simple reduction

$$\begin{aligned} K &= \frac{N_{m(10\text{-}8)}}{N_{m(10\text{-}60)}} = \frac{g(\beta_{10}, \beta_e, \beta_8)}{g(\beta_{10}, \beta_e, \beta_{60})} \\ &= \frac{\beta_8}{\beta_{60}}\sqrt{\frac{1 + (\beta_{60}/\beta_e)^2 + (\beta_{60}/\beta_{10})^2}{1 + (\beta_8/\beta_e)^2 + (\beta_8/\beta_{10})^2}}. \end{aligned} \tag{11}$$

Substituting the known antenna beamwidths, and the measured value of $K = 2.60$, one can solve for β_e, which turns out to be 0.8°. This value is in good agreement with the 1° obtained for β_e in the beam-swinging experiment. In Table III measured values of K for various antenna combinations are compared with values calculated, using (10) and $\beta_e = 0.8°$.

The good agreement between the measured and calculated values indicates that the term $g(\beta_t, \beta_e, \beta_r)$ in (10) is essentially correct. Therefore, the relationship between fading rate (at 4 kmc) and antenna beamwidth is essentially correct.

5.2 *Rate of Fading at 4 kmc — Relation to Wind*

So far, the discussion has been limited to drift of layers in a horizontal plane in the atmosphere. However, there are two effects with altitude which should be taken into account. First, the contributions to the received power by reflections from layers at high altitudes are less than those from low altitudes (it was shown in Section 4.1.2 that the received power decreases about 1 db per thousand feet of altitude). Also, the drift winds vary in both speed and direction with altitude. Both of these effects, in their relation to fading rate, can be taken into account approximately in the following manner.

The volume, V, common to the transmitting and receiving antenna beams is divided into horizontal segments 1000 feet thick, as shown in Fig. 31(a). Each segment is assumed to contain many randomly positioned layers of finite dimension. The power received from segment 1 is weighted by the factor unity, and the standard deviation of the power spectrum for signals reflected from segment 1 is

$$\sigma_1 = \frac{0.85}{\lambda} g(\beta_t, \beta_e, \beta_r) V_1 \sin \delta_1 ,$$

where V_1 and δ_1 are the speed and direction of the wind in segment 1 (as obtained from meteorological data). The power spectrum due to segment 1 is therefore

$$\exp\left(\frac{-F^2}{2\sigma_1^2}\right).$$

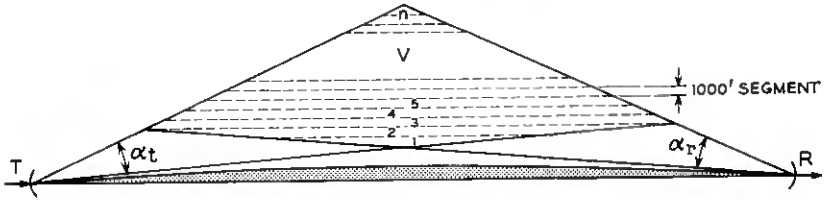
The power from segment 2 is down (on the average) 1 db from the power from segment 1; the power spectrum for segment 2 is therefore

$$0.8 \exp\left(\frac{-F^2}{2\sigma_2^2}\right),$$

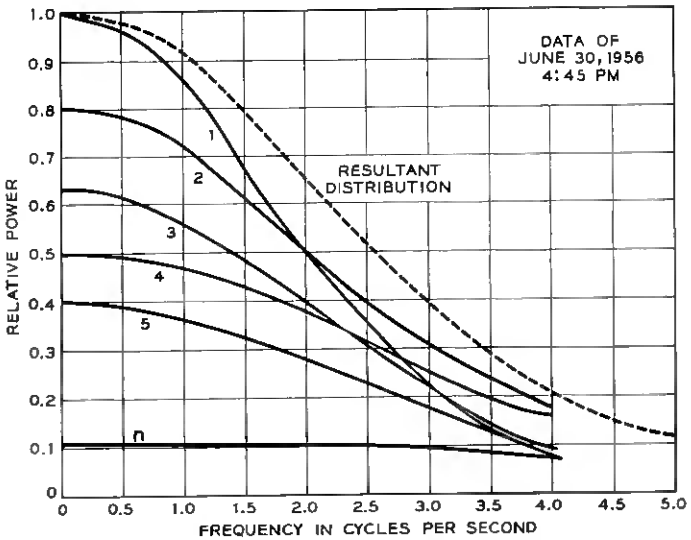
where

$$\sigma_2 = \frac{0.85}{\lambda} g(\beta_t, \beta_e, \beta_r) V_2 \sin \delta_2 ,$$

with V_2 and δ_2 being the speed and direction of the wind in segment 2; and so on to the nth segment. The resultant power spectrum of the received signal is obtained by summing the power spectra for the individual segments as shown in Fig. 31(b). This sum, or resultant, is considered gaussian-shaped for our purposes, although, strictly speaking, this is true only when the winds have the same speed and direction in



(a) VERTICAL PLAN VIEW OF A BEYOND-HORIZON CIRCUIT



(b) EXAMPLE OF A RESULTANT DISTRIBUTION. DISTRIBUTIONS 1,2---n ARE PLOTTED FROM WIND DATA FOR THE CORRESPONDING SEGMENTS IN (a)

Fig. 31 — Distribution of winds in the common volume region: (a) vertical plan view of a beyond-horizon circuit; (b) example of a resultant distribution, with distributions 1, 2, $\cdots$, n being plotted from wind data for the corresponding segments above.

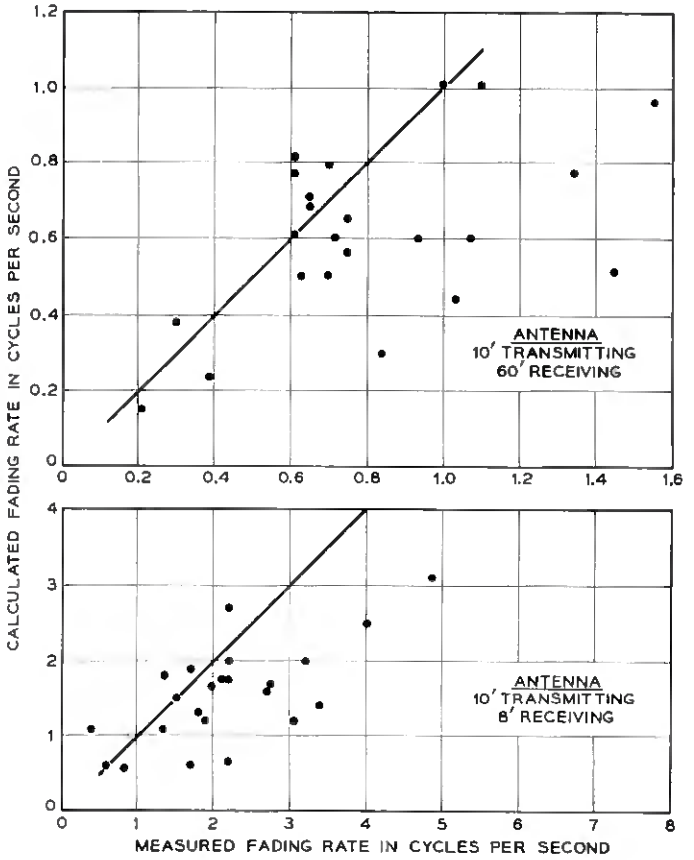


Fig. 32 — Comparison of measured and calculated fading rates. Samples were chosen from data at 4110 mc on 13 different days in the period January through August 1956.

all segments. The estimated fading rate is obtained from the width of the resultant curve. This estimated rate can be compared with fading rates measured from actual records of the received signal.*

In Fig. 32 measured fading rates are compared with fading rates calculated by the method just discussed. The speed and direction of the winds used in the calculations were obtained from the U. S. Weather Station at Binghamton, New York (see Fig. 1). The wind data and the

* If the wind speeds normal to the path are almost the same at all altitudes, the fading rate is given approximately by (10): $N_m = (1.27/\lambda)\ g(\beta_t, \beta_e, \beta_r)\ \overline{V \sin \delta}$, where $\overline{V \sin \delta}$ is the average normal wind speed.

signal records from which the fading rates were measured were taken at approximately the same time for each point shown.

It is apparent from Fig. 32 that the measured fading rate usually exceeds the calculated rate; this may be due to vertical winds or to turbulent winds, neither of which have been considered here. However, it is significant that no constants need to be adjusted in (10) to calculate fading rate of the correct order of magnitude. The rate of fading at 4 kmc is a strong function of horizontal drift wind.

5.3 *Wavelength Dependence of Fading Rates*

This section is concerned with a comparison of fading rates of signals received simultaneously at 4110 and 460 mc. Since it is known that the beamwidths of the antennas affect the fading rate, the present experiment employs two pairs of antennas, the beamwidths of the 4110-mc pair being of the same order as the beamwidths of the 460-mc pair. Furthermore, we have shown that antennas with very narrow beamwidths (less than about 1° at 4110 mc) reduce the rate of fading. Therefore, antennas with relatively large beamwidths were used to minimize such effects.

5.3.1 *Instantaneous Signal Level Recordings — Analysis*

In this study, the 10-foot transmitting and 8-foot receiving antennas are used at 4110 mc and the 28-foot (transmitting) and 60-foot (receiving) antennas are used at 460 mc. The setup is the same as that used in Section 3.2.

The rate of fading at both frequencies varies considerably from day to day and sometimes within the hour. Portions of samples of rapid, slow and very slow fading at 460 mc, and the corresponding 4110-mc signal, are shown on Fig. 33. The 460-mc signal is shown on the top trace of each sample, the 4110-mc signal on the bottom trace. The samples are normally two to three minutes long and are taken each half-hour during the recording period. It may be noted in Fig. 33 that the fading rate of the 460-mc signal, N_{m460}, changes considerably from sample to sample, whereas the fading rate of the 4110-mc signal, N_{m4110}, remains fairly constant; therefore, changes in the *ratio* of the rates of fading at the two frequencies are due predominantly to changes in the 460-mc fading rate.

The ratio, k, of the fading rates of the signals received simultaneously at the two frequencies ($k = N_{m4110}/N_{m460}$) varies widely, as shown in

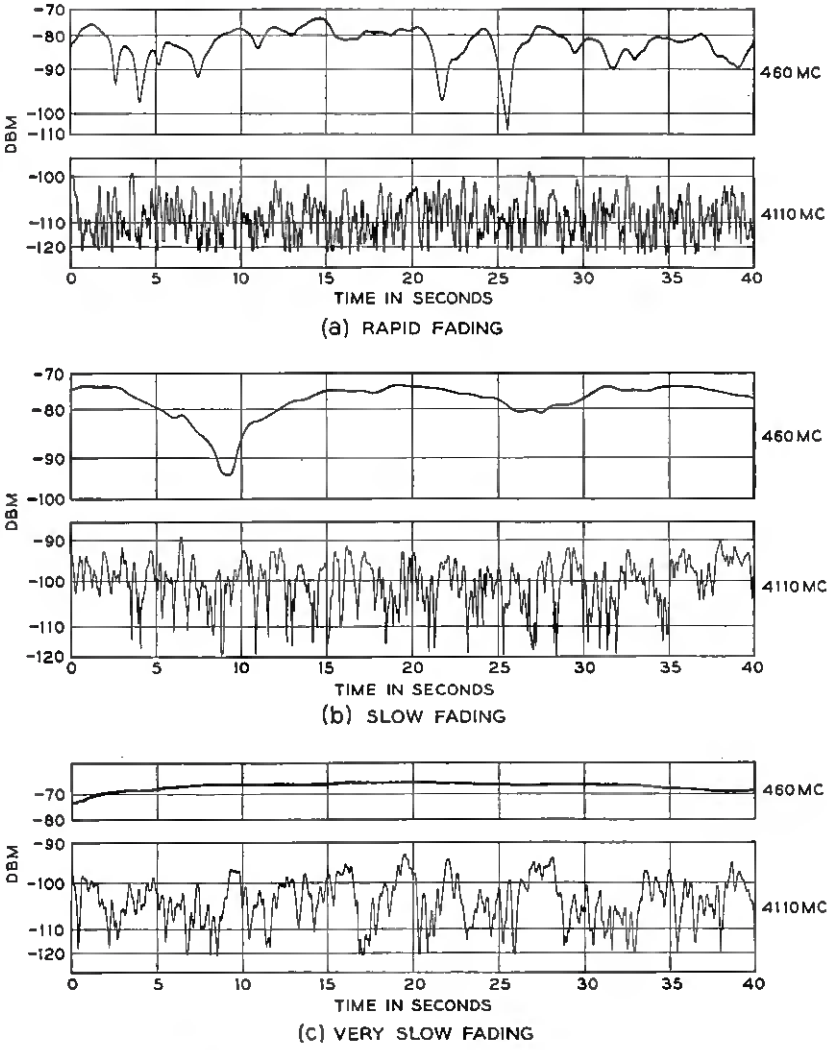


Fig. 33 — Segments of samples of simultaneous recordings of instantaneous signal levels at 460 and 4110 mc.

Fig. 34. The data are taken from a study of 387 samples and show that the median value of the ratio of the fading rate of the signals received at 4110 mc to the 460-mc rate is 12.7. It is believed that this ratio is significantly higher than the ratio of the radio frequencies due to the effect of layer size, as discussed in the following section.

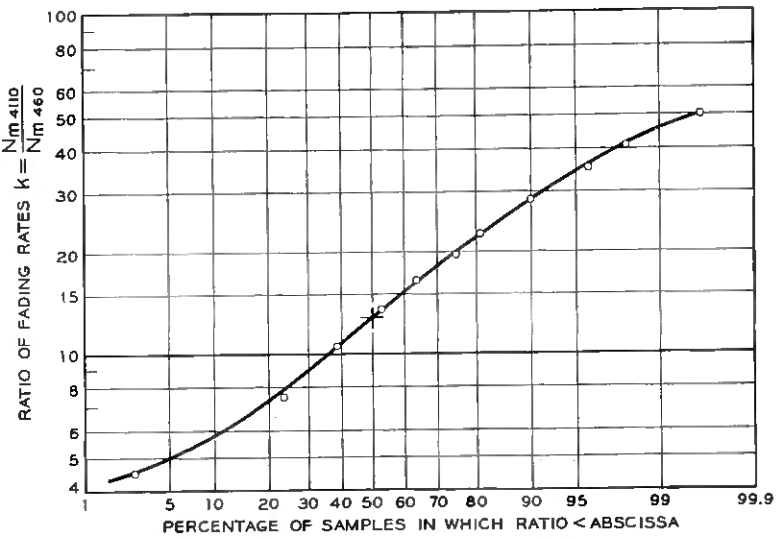


Fig. 34 — Statistical distribution of the ratio of fading rates recorded simultaneously at 4110 and 460 mc — 387 samples of 2 minutes duration. Period: March through December 1956.

5.3.2 *Ratio of Fading Rates and Layer Size*

A discussion of fading of 4-kmc signals was given in Section 5.1. There it was assumed that short-term fading was caused by drift of numerous randomly positioned layers formed by changes in the gradient of the dielectric constant of the atmosphere.

Consider now the effect of the size of the reflecting layers on the fading rates. It will be shown that small layers result in a ratio of fading rates equal to the ratio of the radio frequencies, whereas large periodic layers must be assumed to explain higher ratios of the fading rates.

Small Layers — The diagram in Fig. 30(a) is applicable if the horizontal dimensions of the reflecting layers are small compared to the horizontal dimension of the antenna beam at the operating wavelength. If this condition is fulfilled at wavelengths λ_1 and λ_2, (10) applies to both wavelengths, and for k one obtains the ratio of fading rates measured simultaneously at wavelengths λ_1 and λ_2 with antenna beamwidths β_{t_1}, β_{r_1} and β_{t_2}, β_{r_2}:

$$k = \frac{N_{m1}}{N_{m2}} = \frac{\lambda_2}{\lambda_1} \frac{g_1(\beta_{t_1}, \beta_{e_1}, \beta_{r_1})}{g_2(\beta_{t_2}, \beta_{e_2}, \beta_{r_2})}. \tag{12}$$

The factors g_1 and g_2 are very nearly equal if the beamwidths of the

antennas used are equal and if the reflecting layers are small compared to the Fresnel zone at both wavelengths ($\beta_{e_1} = \beta_{e_2}$). For this case, the ratio of fading rates is just proportional to the ratio of the operating wavelengths. Letting subscripts 1 and 2 refer to 4110 and 460 mc respectively, one obtains

$$k = \frac{N_{m4110}}{N_{m460}} = \frac{\lambda_2}{\lambda_1} = \frac{0.65m}{0.073m} = 9. \tag{13}$$

The experimental data of Fig. 33(a) are an example of the condition where the layers apparently are smaller than a Fresnel zone at both wavelengths, resulting in a value of $k \cong 9$.

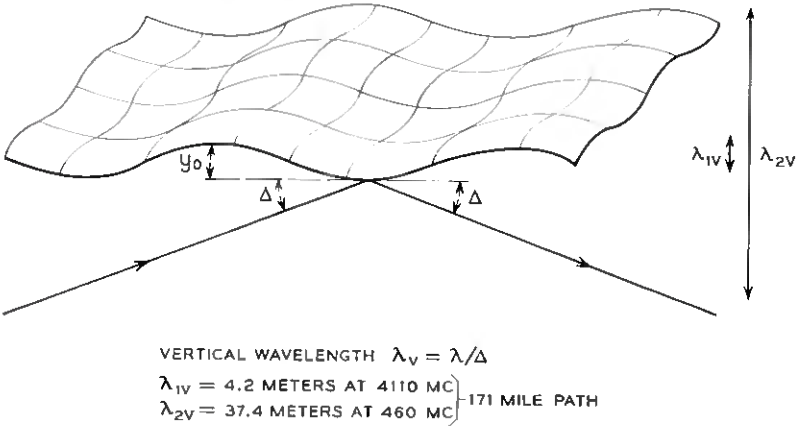


Fig. 35 — Idealized structure representing a single layer at 460 mc and several layers at 4110 mc.

Large Layers — Analysis of many samples makes it clear that the ratio of the fading rates, N_{m4110}/N_{m460}, often far exceeds the value given by (13); the ratio may reach values in excess of 50. This condition is due to a very low fading rate at 460 mc rather than a very high 4110-mc fading rate. When these high ratios exist, the atmosphere apparently acts much differently at 460 mc than at 4110 mc.

A layer such as that shown in Fig. 35 will act as a smooth reflecting surface at the lower frequency, and simultaneously will represent many reflecting surfaces at the higher frequency.* The layer is considered to be a smooth reflecting surface if the roughness as indicated by the spatial amplitude, y_0, in Fig. 35 is less than one-eighth of a vertical wavelength (Rayleigh's criterion). The vertical wavelengths $\lambda_v = \lambda/\Delta$ for a 171-

* A similar argument is used by Anderson and Smyth.[15]

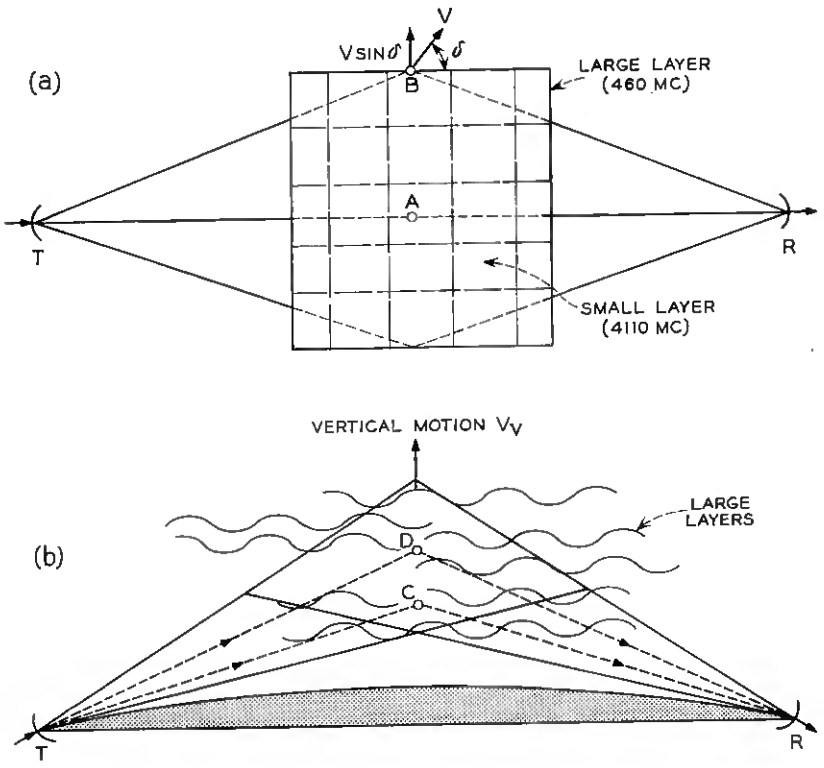


Fig. 36 — Large idealized periodic layers in the propagation path: (a) horizontal; (b) vertical.

mile path and frequencies 4110 and 460 mc are shown in the figure, Δ being the grazing angle.

Thus, if $y_0 < \lambda_{2v}/8 = 4.6$ meters, the layer appears large and smooth at 460 mc. On the other hand, this value for the roughness amplitude, y_0, exceeds an entire vertical wavelength at 4110 mc ($\lambda_{1v} = 4.2$ meters). Therefore, the loops and nodes of the idealized periodic layer will appear as individual layers at 4110 mc.*

The idealized layers which are assumed to exist in the propagation path are shown in Fig. 36. In the vertical plan view, several of the periodic layers are drawn in the volume common to the antenna beams, and in the horizontal plan view, Fig. 36(a), one of these is shown sub-

* Periodic layers such as these probably can be produced by changes in magnitude and direction of winds with height in the atmosphere; no doubt actual layers are more irregular and less systematic than the one shown in Fig. 35

divided by dotted lines to illustrate the periodic structure. Since the layers at 4110 mc are the loops and the nodes of the periodic structure, the situation at 4110 mc is much the same as that already discussed in connection with Fig. 30(a), and the fading rate at 4110 mc will have fairly high values depending on the horizontal wind, as given by (10). However, at 460 mc the horizontal drift wind has little effect on the fading rate if the layers are large and smooth at that frequency, since most of the power is propagated to the receiver along the great circle route TAR and very little along delayed paths such as TBR. Nevertheless, relative delays at 460 mc do exist in the vertical plane as shown by rays TCR and TDR in Fig. 36(b), and some fading will occur due to interference of reflections from the several periodic layers. But the fading rate at 460 mc will be relatively low, since the vertical motion of the layers V_v is usually very much less than the horizontal drift V. Large periodic layers therefore can give rise to high ratios N_{m4110}/N_{m460}.

5.3.3 *Fading Rate Ratios and Ratios of Median Received Powers — Diurnal Effects*

It has been observed that the ratio of the fading rates N_{m4110}/N_{m460} may change considerably during the course of a day, as shown in Fig. 37. Moreover, the relative median received power at the two frequencies, P_{4110}/P_{460}, is found to change simultaneously with the fading ratio; the out-of-phase relationship between fading ratio and received power ratio is evident in Fig. 37. The median received powers P_{4110} and P_{460} used in Fig. 37 both have the free-space transmission loss factored out; the ratio P_{4110}/P_{460} therefore is a measure of the relative transmission coefficient of the atmosphere at the two frequencies. When the fading ratio is low ($k \cong 9$) the ratio of received powers P_{4110}/P_{460} is between 0 and -5 db, whereas a high ratio of fading rates ($k \geqq 20$) occurs when the power ratio is between -10 and -15 db. The relationship between fading ratio and power ratio can be discussed qualitatively by means of the theory, as follows:

In Section 3.2.2 the median received power is discussed for three cases depending on the horizontal dimensions of the reflecting layers. The received power, relative to the free-space value, at two wavelengths λ_1 and λ_2 is:

Case 1. Large Layers (Dimension > Fresnel Zone):

$$\frac{P_R(\lambda_1)}{P_R(\lambda_2)} = \left(\frac{\lambda_1}{\lambda_2}\right)^2.$$

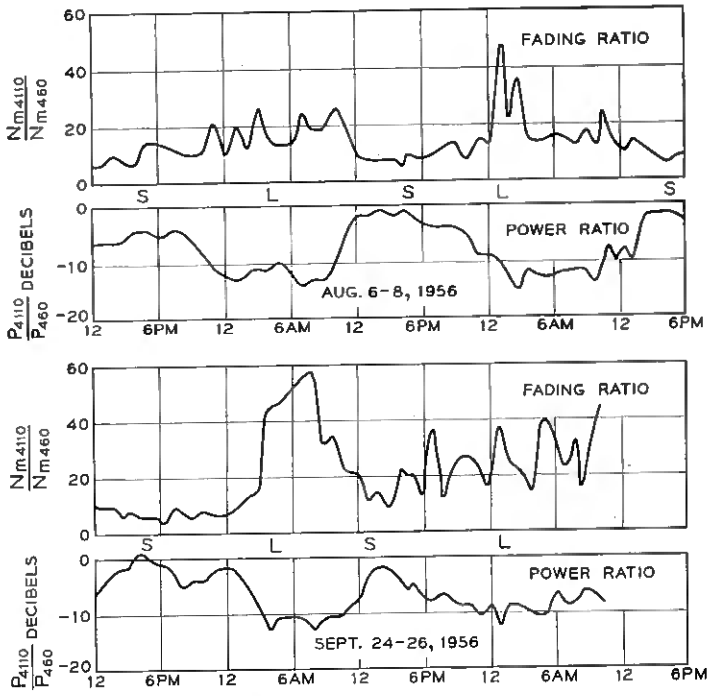


Fig. 37 — Correlation between ratio of median fading rates and ratio of median received powers (relative to free space) at 4110 and 460 mc.

Case 2. Small Layers (Dimension < Fresnel Zone):

$$\frac{P_R(\lambda_1)}{P_R(\lambda_2)} = \left(\frac{\lambda_1}{\lambda_2}\right)^0.$$

Case 3. Layers of Intermediate Size (Dimension ≈ Fresnel Zone):

$$\frac{P_R(\lambda_1)}{P_R(\lambda_2)} = \left(\frac{\lambda_1}{\lambda_2}\right)^1.$$

Consider Case 2. This formula indicates that both wavelengths, λ_1 and λ_2, are propagated equally well when the atmosphere contains small reflecting layers. In terms of Fig. 37, it means that small layers are present when $P_{4110}/P_{460} \approx 0$ db (various portions where $P_{4110}/P_{460} \approx 0$ are labeled s). During the same periods the ratio of fading rates is low ($k \approx 9$), as discussed previously. Thus, the observed ratio of received powers and the observed ratio of the fading rates are both consistent with the existence of small layers during the period labeled s.

If the formulas for Cases 3 and 1 are applied to the 4110- and 460-mc circuits, one obtains

Case 3: $\frac{P_{4110}}{P_{460}} = -9.5$ db for layers of intermediate size

Case 1: $\frac{P_{4110}}{P_{460}} = -19$ db for large layers.

Thus, the propagation at the lower frequency is favored more and more as the layer dimensions increase. It is believed that the observed power ratios of between -10 and -15 db, shown in the portions of Fig. 37 labeled L, are due to large layers at 460 mc and small layers at 4110 mc (Fig. 35). However, this case cannot be calculated with confidence by theory, since our knowledge of the sizes and refractive index gradients of the layers is insufficient. Nevertheless, the observed power ratios during periods L (Fig. 37) are consistent with the concept of large layers at 460 mc and with high ratios of fading rate as discussed previously.

Additional data have been plotted in the manner of Fig. 37; all of these show the out-of-phase relationship between N_{m4110}/N_{m460} and P_{4110}/P_{460}. The two plots shown in Fig. 37 were chosen to illustrate the diurnal effect; it will be noted that large layers exist in the early morning hours when the atmosphere is believed to be more stable.

Diurnal Variation. An analysis was made of the fading rate at 460 mc as a function of time of day; Fig. 38(a) shows these results. It indicates that the atmosphere is stable at night and during the morning hours.

The fading rate is low between the hours of 9:00 p.m. and 9:00 a.m. As heating by the sun occurs, the fading rate increases rapidly between 9:00 a.m. and 11:00 a.m., staying relatively high until 7:00 p.m. It drops rapidly between 7:00 p.m. and 9:00 p.m., as radiation cooling occurs.

The diurnal effect is not strong so far as the median signal level is concerned, particularly in the winter months; however, there is a trend which follows the fading rate variations; i.e., at higher fading rates the median signal level decreases. Fig. 38(b) shows this relation.

5.3.4 *Fading Ratio and Fading Rate*

An interesting relationship shows up if the measured ratios of fading rates $k = N_{m4110}/N_{m460}$ are plotted versus the rates themselves. Fig. 39 shows plots of 387 ratios obtained from fading rates measured on the 4110- and 460-mc circuits. It is evident that there is a sharp break in the data of Fig. 39(a), in which the ratios are plotted against the 460-mc fading rate. The ratio varies only slightly ($k \leqq 9$) as the fading rate

decreases from 1.4 cps to 0.2 cps, at which point the ratio begins to increase rapidly. This plot emphasizes that the high ratios N_{m4110}/N_{m460} are caused by very low values of N_{m460}. A similar effect is seen in Fig. 39(b), where the same measured data are plotted versus the fading rate at 4110 mc.

Consistent with all previous discussion, these data show that two rather well-defined ranges of 460-mc fading rate exist. One concludes that, when the periodic layers illustrated in Fig. 36 are present, the 460-mc fading rate is less than 0.2 cps [Fig. 39(a)] and the 4110-mc rate is less than about 2 cps [Fig. 39(b)], and the rates are higher than these values when the layers are broken up.

The rate of fading of 4110-mc signals, N_{m4110}, is related to the horizontal drift winds in the atmosphere. The fading rates calculated by

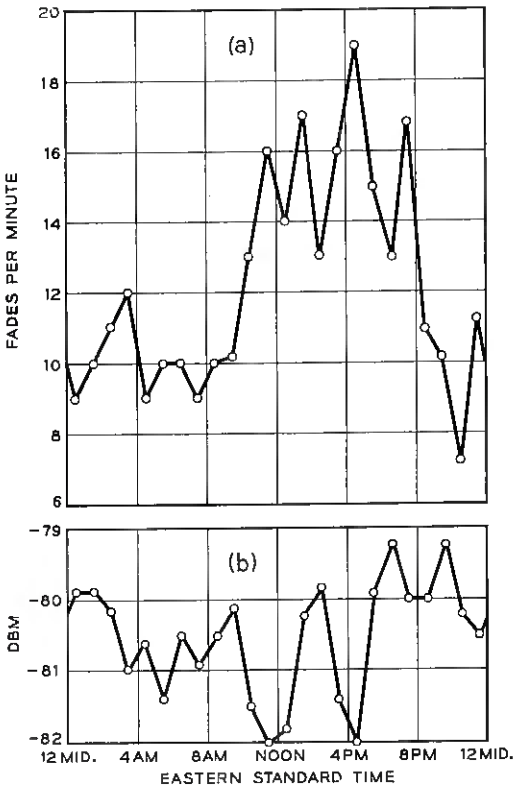


Fig. 38 — Diurnal variation in (a) median fading rate and (b) median signal level at 460 mc — 1680 samples. Period: September 1956 through April 1957.

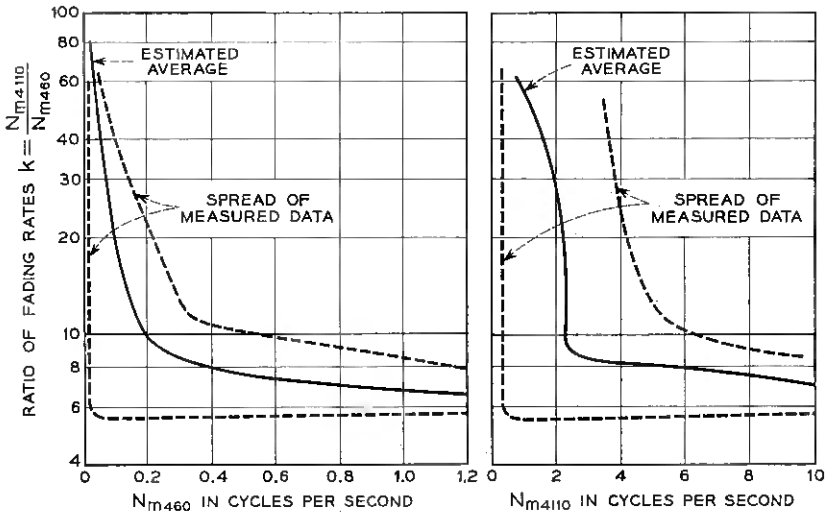


Fig. 39 — Ratio of median fading rates at 460 and 4110 mc as a function of absolute median fading rate — 387 samples.

(10), that is, $N_m = (1.27/\lambda)g(\beta_t, \beta_e, \beta_r)\ V\sin\delta$, using measured wind velocities (V,δ), were shown to agree fairly well with the actual fading rates of the received signal. Therefore, since $N_{m4110} = 2$ cps is the value at which the curve of Fig. 39(b) suddenly changes slope, one can write the equation

$$\frac{1.27}{\lambda}\, g(\beta_t, \beta_e, \beta_r)\ V\sin\delta = 2 \text{ cps},$$

and solve for the wind velocity corresponding to the breaking point of the curve. This results in $V \approx 15$ meters per second (30 knots) for the "critical" velocity below which large layers can exist. One concludes that formation of the large layers depends not only on the stability of the atmosphere as related to time of day but also on the wind conditions in the volume of the atmosphere that is common to the beams of the transmitting and receiving antennas.

VI. TWIN-FEED DIVERSITY STUDIES

In propagation beyond the horizon there are two principal types of fading, defined somewhat arbitrarily as "long-term" and "short-term" fading. The long-term fading is caused by changes in the refractive properties of the atmosphere over periods of from several hours to days

(Section III); short-term fading is due to multipath changes caused by movement of discontinuities (Section V). In short-term fading the received signal level follows, in most instances, a Rayleigh distribution.

A system is designed to operate for a given percentage of the time above a certain signal-to-noise level. To decrease the time the signal spends at low levels for a given system, diversity reception is used to obtain two or more uncorrelated signals at the receiver. The usual type is space diversity: two or more receiving antennas are spaced about 100 wavelengths perpendicular to the great circle between transmitter and receiver. The output from each antenna is amplified and combined with the outputs of the other antennas.[16] If the signals are uncorrelated, the percentage of time the combined output stays below a certain level will be less than the time the signal from either channel stays below that level.

The diversity systems to be described use one receiving antenna with two feeds.* Uncorrelated signals should be obtained from the two feeds because they are oriented to illuminate different parts of the atmosphere. Two types of feed positions are described: horizontal feeds placed side by side and vertical feeds placed one on top of the other.

In order to check the experimental results with theory, we first must determine whether or not the envelope of the signals is Rayleigh-distributed, since the subsequent calculations on the diversity systems are based on a Rayleigh distribution. The following few paragraphs discuss this aspect.

6.1 *Types of Distributions of Instantaneous Signal — Non-Rayleigh*

The distribution of the envelope of the instantaneous signal level usually follows a Rayleigh distribution. A sample is shown in Fig. 40. Other examples of the distributions may be found in Section 5.1.1.

An analysis of numerous short-term fading samples has shown that the envelope of the signals is Rayleigh-distributed when the fading rate is greater than about 8 fades per minute at 460 mc and 12 fades per minute at 4 kmc. In Fig. 41, distributions of fading rates taken over a considerable period of time are shown for both 4110 and 460 mc. From these data one concludes that the signals are not Rayleigh-distributed for 35 per cent of the time at 460 mc and 3 per cent of the time at 4 kmc.

The distributions which are not Rayleigh-distributed usually occur when the median fading rate is very slow. While the very slow fading rates occur at both 460 mc and 4.11 kmc, they occur more often at 460

* Suggested by H. T. Friis.

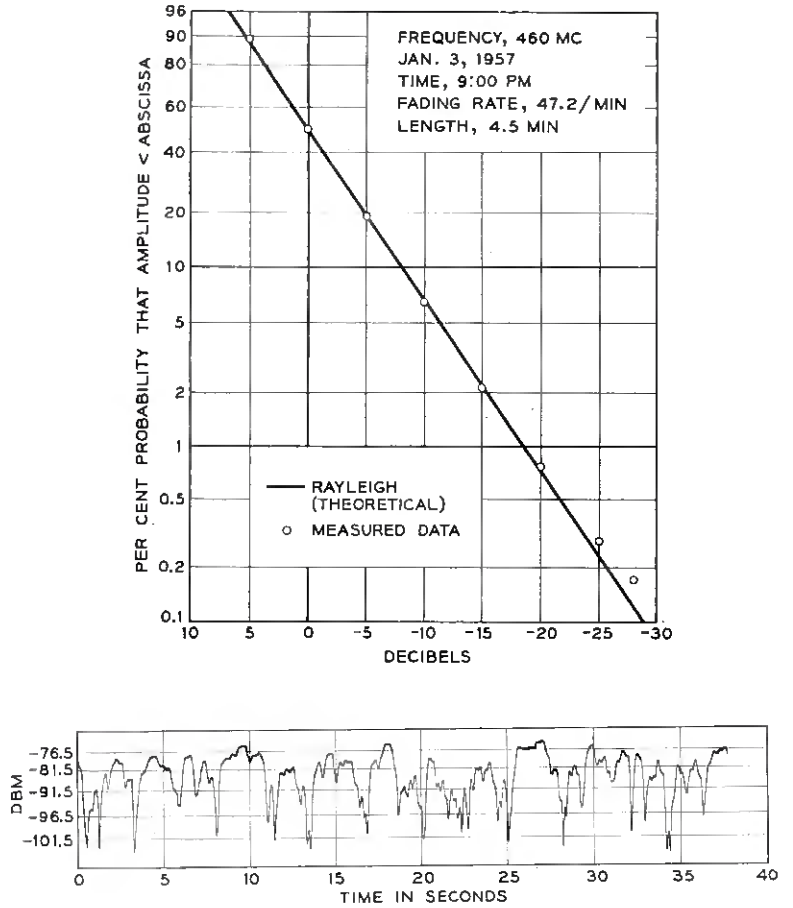


Fig. 40 — Segment of recording and distribution of instantaneous signal levels of Rayleigh-distributed signal.

mc; hence the examples cited are for 460 mc. During these periods of very slow fading the number of layers or discontinuities in the common volume is reduced to a few, and also the drift of these layers is slow. It is possible to fit these distributions to some theoretical models. Fig. 42 shows the distribution for the model of a constant vector and a Rayleigh-distributed vector;[17] the points are experimental. Figs. 43, 44 and 45 show the model for two, three and four equal-amplitude but randomly phased vectors, respectively. Some of the distributions obtained from other samples do not fit any simple theoretical model. One

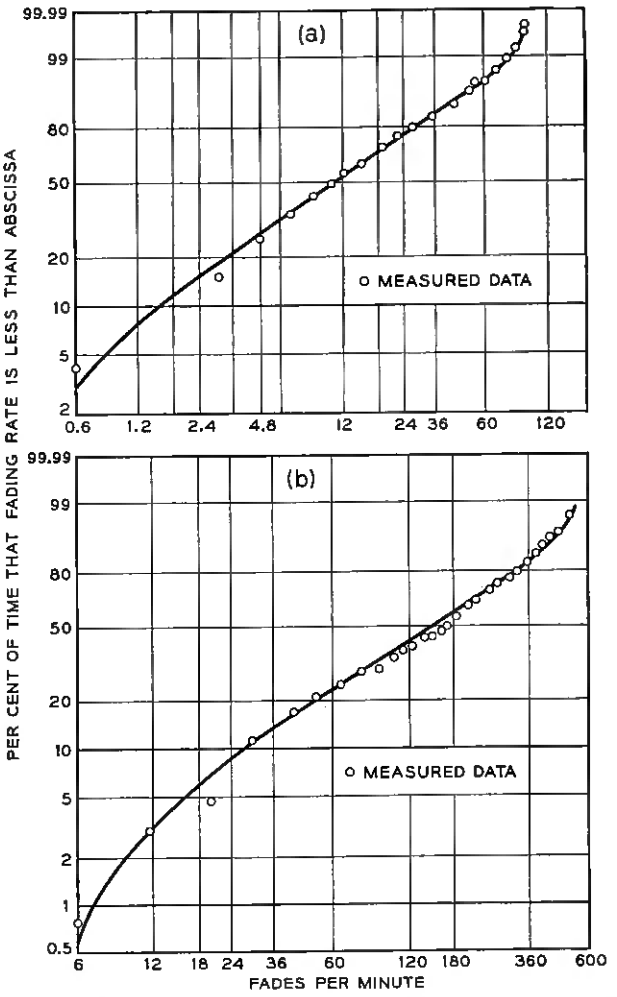


Fig. 41 — (a) Distribution of median fading rates at 460 mc — 1680 samples each 2 minutes long; 28-foot transmitting antenna, 60-foot receiving antenna. Period: September 1956 through April 1957. (b) Distribution of median fading rates at 4110 mc — 780 samples each 2 minutes long; 10-foot transmitting antenna, 8-foot receiving antenna. Period: October 1955 through April 1956.

would expect the lower fading rates to be due to a stable atmosphere composed of a few large layers. That there is a relationship between fading rate and horizontal wind speed has been shown in Section 5.2. These results do not indicate the correlation between vertical winds and fading rate, since the vertical component usually produces only a secondary effect.

During the periods of very slow fading it is sometimes possible to infer that the fading is caused by the vertical component of the wind. The 60-foot paraboloid is scanned in order to find the angular dependence of the scattered energy. (The rate of scan is four per minute; the transmitter frequency is 4.11 kmc.) Consider Fig. 46, which shows these scans. The horizontal patterns in Figs. 46(b) and 46(c), while changing in amplitude, do not move in position; the vertical scans have a tendency to change in position as well as in amplitude. If large layers tend to move vertically, one would expect the signals from two adjacent parts of the atmosphere to be correlated in the horizontal plane but not necessarily in the vertical plane. Fig. 46(d) shows an instance of some-

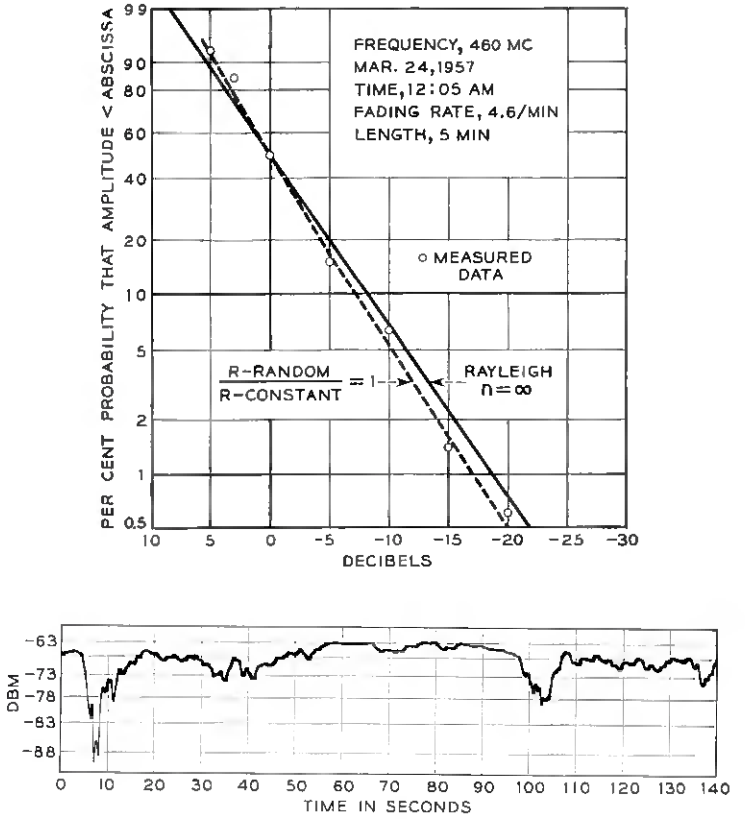


Fig. 42 — Segment of recording and distribution of instantaneous signal levels of the resultant of a constant vector plus a Rayleigh-distributed vector of equal amplitude. Points are experimental; curves are theoretical.

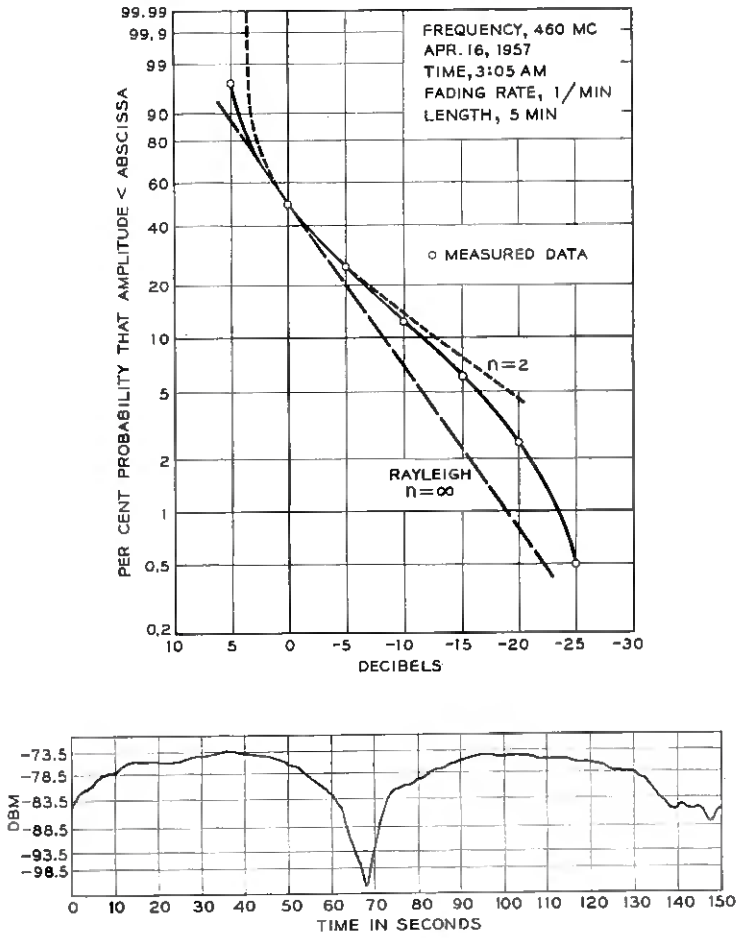


Fig. 43 — Segment of recording and distribution of instantaneous signal levels of the resultant of two unit vectors with random relative phase. Points are experimental.

what the opposite effect. However, here the fading is more rapid, indicating the effect of horizontal winds.

6.2 *Two-Channel Switch-Type Diversity*

We will first analyze the advantages obtainable in diversity systems. In this analysis we assume that the uncorrelated signals of two receivers

are combined by a switch-type combiner, which selects the higher of the instantaneous signals. The signals need not have the same median level.

The envelope of the instantaneous signal level will be Rayleigh-distributed in most cases, as was pointed out in the previous section. This probability distribution results from an addition of an infinite number of components of random phase and amplitude. The probability, P, that

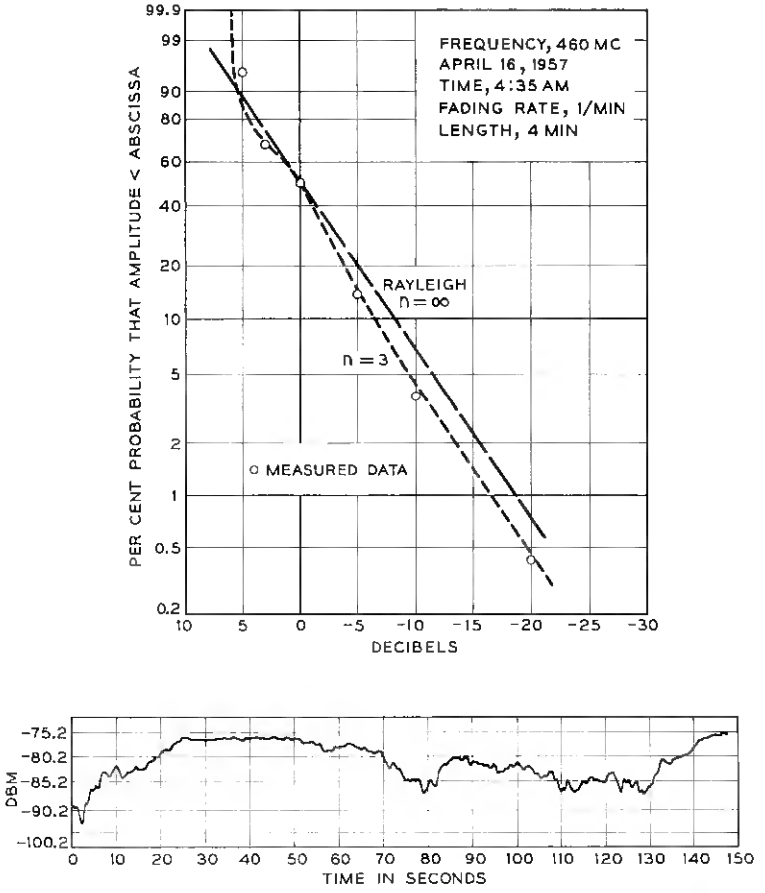


Fig. 44 — Segment of recording and distribution of instantaneous signal levels of the resultant of three unit vectors with random relative phase. Points are experimental; curves are theoretical.

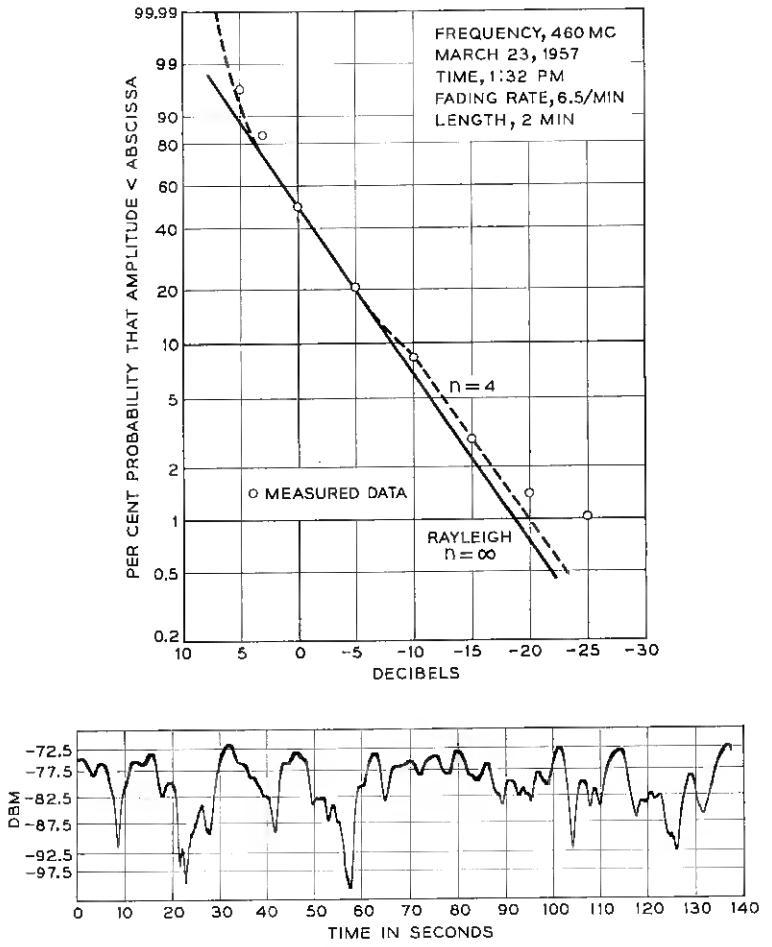


Fig. 45 — Segment of recording and distribution of instantaneous signal levels of the resultant of four unit vectors with random relative phase. Points are experimental; curves are theoretical.

the signal amplitude E_1 is less than a given value E is

$$P(E_1 < E) = 1 - \exp\left[-\left(\frac{E}{E_r}\right)^2\right], \tag{14}$$

where E_r is the rms value. For $P = 0.50$, E is equal to the median value E_m ; hence,

$$E_m^2 = (\ln 2)\, E_r^2; \tag{14a}$$

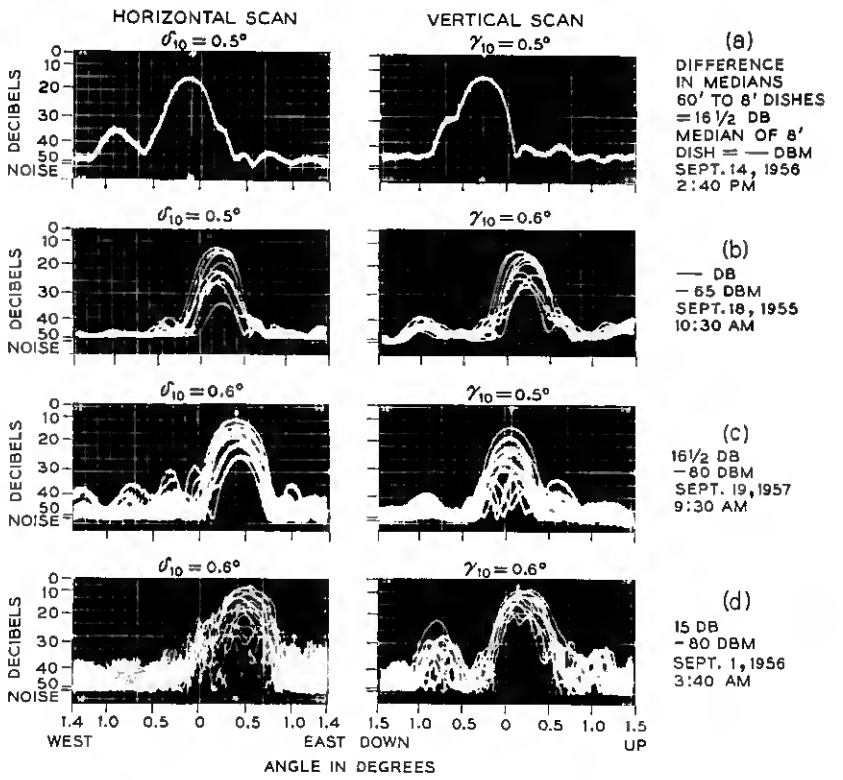


Fig. 46 — Scans of signals received on the 60-foot antenna at 4110 mc.

and

$$P(E_1 < E) = 1 - \exp\left[-\ln 2\left(\frac{E}{E_m}\right)^2\right]. \tag{15}$$

This equation, plotted in Fig. 47, is labeled "single channel." If we have two signals, E_1 and E_2, then

$$\begin{aligned} P(E_1 < E) &= 1 - \exp\left[-\ln 2\left(\frac{E}{E_{m_1}}\right)^2\right], \\ P(E_2 < E) &= 1 - \exp\left[-\ln 2\left(\frac{E}{E_{m_2}}\right)^2\right]. \end{aligned} \tag{16}$$

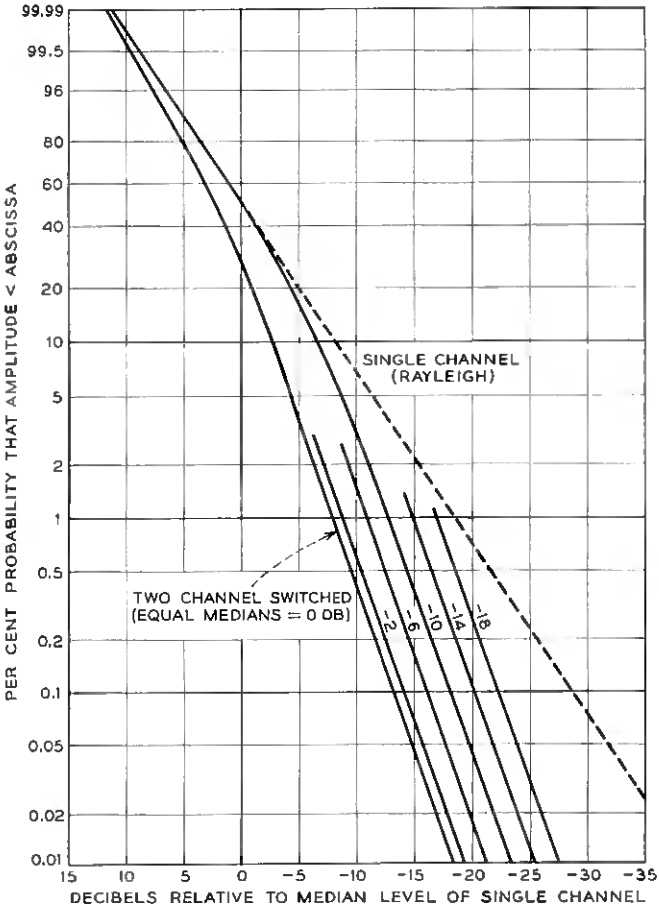


Fig. 47 — Distribution of amplitudes for two-channel switched diversity system with unequal median signal levels for uncorrelated Rayleigh-distributed channels. Curves are referred to the median of the single channel with the higher median.

The probability that both E_1 and E_2 are less than E is

$$\begin{aligned} P(E_1, E_2 < E) &= [P(E_1 < E)][P(E_2 < E)] \\ &= \left\{1 - \exp\left[-\ln 2 \left(\frac{E}{E_{m_1}}\right)^2\right]\right\} \\ &\quad \cdot \left\{1 - \exp\left[-\ln 2 \left(\frac{E}{E_{m_2}}\right)^2\right]\right\}. \end{aligned} \tag{17}$$

For equal medians $E_{m_1} = E_{m_2} = E_m$, (17) reduces to

$$P(E_1, E_2 < E) = \left\{1 - \exp\left[-\ln 2 \left(\frac{E}{E_m}\right)^2\right]\right\}^2. \tag{18}$$

The probability that both signals are below a pre-set level is the square of the probability of one signal channel being below that level.

For very small signal levels (E less than 0.3 E_m) we now compare the various systems — single-channel, equal median diversity and unequal median diversity — at the same percentage of time. We shall use as reference the median of the single-channel E_{m_1}.

i. Single-Channel:

$$P(E_1 < E) \approx \ln 2 \left(\frac{E}{E_{m_1}}\right)^2; \tag{19}$$

ii. Two-Channel, Equal Medians:

$$P(E_1, E_2 < E) \approx \left[\ln 2 \left(\frac{E}{E_{m_1}}\right)^2\right]^2; \tag{20}$$

iii. Two-Channel, Unequal Medians:

$$P'(E_1, E_2 < E') \approx \left[\ln 2 \frac{E'^2}{E'_{m_1} E_{m_2}}\right]^2. \tag{21}$$

For the same probability, $P = P'$, the degradation of the switched signal for an unequal median system over that of an equal median signal is (set $E'_{m_1} = E_{m_1}$):

$$\frac{E'^2}{E_{m_1} E_{m_2}} = \frac{E^2}{E_{m_1}{}^2}; \qquad \text{hence,} \quad \frac{E}{E'} = \sqrt{\frac{E_{m_1}}{E_{m_2}}}. \tag{22}$$

In this comparison we have set the median of the single channel for the equal median case (E_{m_1}) equal to the median of the single channel with the larger median (E'_{m_1}) for the unequal median case. At a given probability, the level for the two-channel switched diversity with unequal medians will be below that of the two-channel switched diversity with equal medians by one-half of the difference in decibels of the two medians. For a difference of 10 db in medians, the level of the switched channel with unequal medians will be 5 db below the level of the switched channel with equal medians. Fig. 47 shows the distributions for various ratios in medians.

6.3 *Twin-Feed Diversity — Horizontally Disposed Feeds*

In this type of diversity, the two horns are mounted side by side on one paraboloidal reflector so that each beam illuminates horizontally adjacent volumes in the common volume region (Fig. 48). The antenna is oriented so that each feed receives the same median signal level.

Successful diversity reception requires that the signals received on the two horns be uncorrelated. If the discontinuities in the common volume are so large that most of the energy is transmitted to the receiver along a great circle route, then the signals received on the separate horns will be correlated. An attempt to correct this condition by separation of horns will lead to degradation in the median signals received.

6.3.1 *Experimental Setup and Analysis of Data*

Fig. 49 shows the free-space antenna patterns at 460 mc using the 60-foot paraboloid. The separation of the beams is 2.33°, which corresponds to 3.5 miles at the midpath. The signal level received is 4.5 db down at the crossover point.

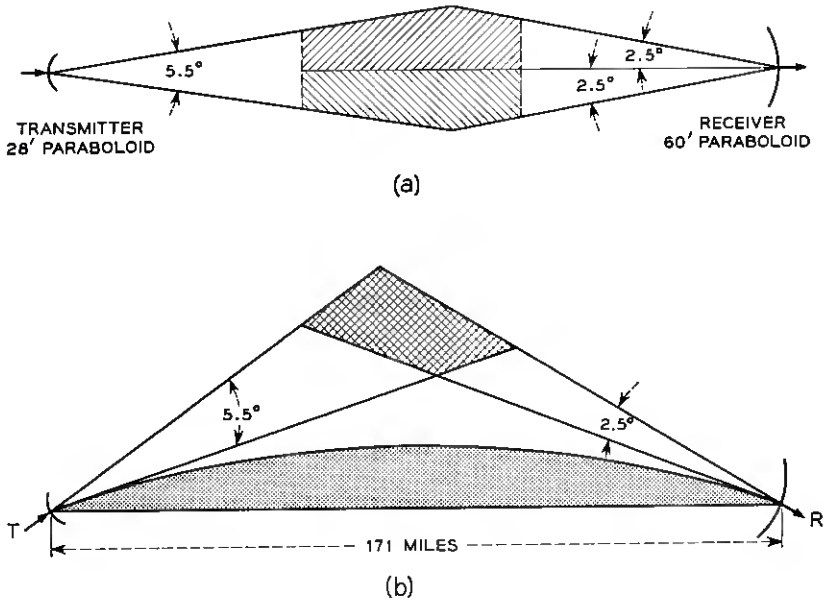


Fig. 48 — Common volume configuration for twin-feed diversity system with horizontally disposed feeds, at 460 mc, using 60-foot antenna: (a) horizontal; (b) vertical.

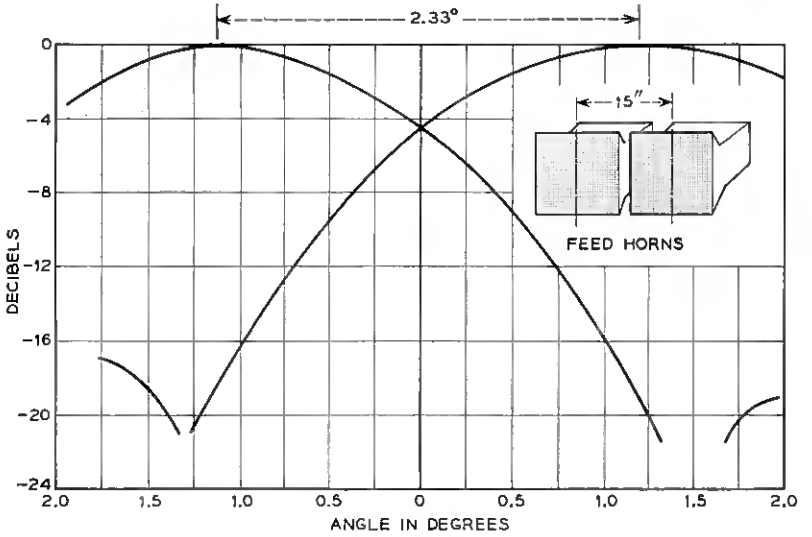


Fig. 49 — Free-space patterns for twin-feed diversity system with horizontally disposed feeds, at 460 mc, using 60-foot antenna.

The purpose of the experiment is to determine whether or not an advantage is obtained when the higher of the two signals is chosen instead of the signal from one channel. The degree of correlation between the two signals determines the diversity improvement of the system. The signal levels below the median value are chiefly of interest, since the low levels determine the operational reliability of the system.

Fig. 50 shows the setup wherein the signals are combined in the diversity combiner; the output of the combiner is put on one channel of the Sanborn recorder. One of the signals is put on the other channel. The median level of each channel is recorded on the E-A recorders. The combiner consists of two vacuum tube diodes whose plates are tied together, each cathode being connected to one of the two channels; the output is taken from the plates. Since the receiver output impedance is low compared to the diode resistance and the voltage into the combiner is high (0 to −100 volts dc), no appreciable error is introduced because of the diode resistance and emission current.

Since the nature of the received signals is such that the signal spends little time at levels well below the median, it becomes difficult to measure, using the Sanborn records, the time that the signal spends at these levels. This problem is especially acute in the diversity studies, since these levels are the ones of interest. More accurate results can be obtained by

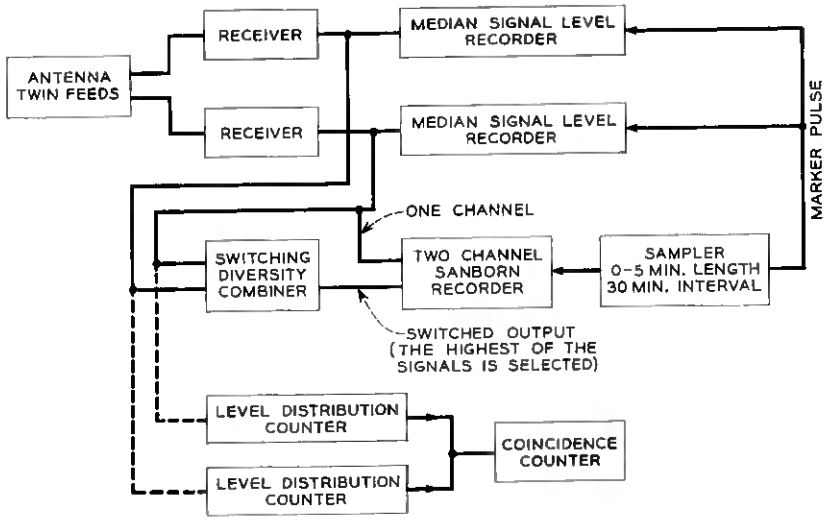


Fig. 50 — Recording setup for measurement of diversity properties of two-channel switched diversity system.

using level distribution counters which totalize the time the signal is below a certain set level. These counters are gated, with the gate opening when the signal drops below the preset level. The counting frequency is 2000 cps. The level can be set to an accuracy of 0.3 volts input for the gate to open or close. A clock is used as a time totalizer.

The counters can be so connected that one counts on one channel, the second counts on the same level on the other channel and the third counts when both channels are below the common preset level; the third counter registers coincidences, and hence gives a direct measure of the properties of the diversity system.

There is an important difference between using the Sanborn recorder with the diversity combiner and the level-distribution counters. The Sanborn recorder gives the diversity action at all levels of signal, while the counters only give information about one level. To get equivalent information a bank of counters would be needed.

For the recordings on the Sanborn to give quantitative results, a further condition is necessary. The combiner does not select the higher of the two signal levels but only the higher of the two receiver outputs, thus for quantitative results the calibration of the two receivers must be identical. By adjusting the gain and AGC level it is possible to get the calibrations to be within ± 1 db over the 50-db range.

For the horizontally spaced twin feeds the instantaneous signal levels are recorded on the Sanborn recorder and the following analysis of the data is made. A probability distribution is plotted for the single channel as well as the switched channel. Fig. 51 shows a sample distribution. Since the experimental points lie close to the theoretical curve, the signals are Rayleigh-distributed and uncorrelated in this case. We define diversity improvement as follows: at 10 db below the median of the single channel the probability is 6.70 per cent for the single channel and 0.45 per cent for the switched channel; hence the diversity improvement is 6.25 per cent. It is possible for the single-channel distribution to be Rayleigh, but for the switched channel not to follow the theoretical curve, due to correlation between the two signals.

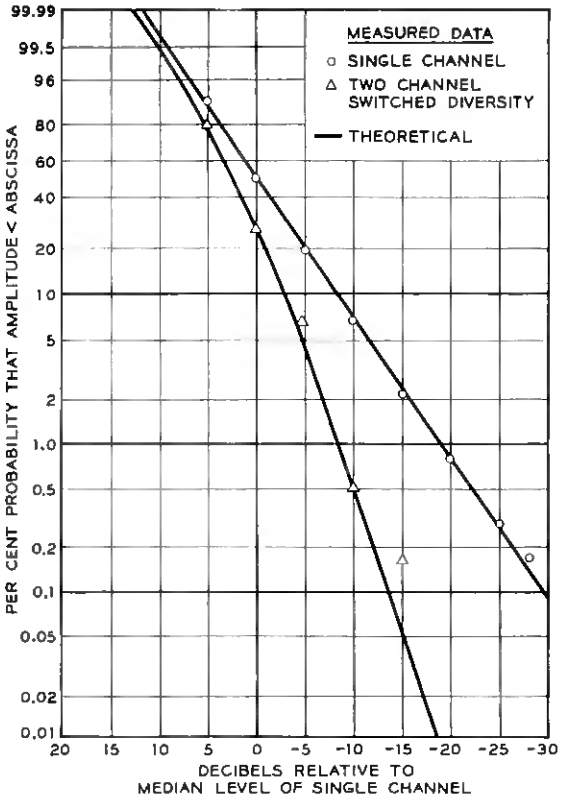


Fig. 51 — Distribution of instantaneous signal levels of a two-channel switched diversity system with equal median signals at 460 mc. Sample length, 4½ minutes; median fading rate of single channel, 47.2 cpm.

6.3.2 *Results of the Experiment*

The experiments at 4 kmc using the twin feeds on the 60-foot antenna as well as the 28-foot antenna show that the theoretical improvement was obtained in nearly all cases. The only time it did not occur was when the fading rate was extremely slow, indicating the presence of very large horizontal layers.

At 460 mc the theoretical diversity improvement was not obtained in many cases. Since the fading rate is an indication of the stability of the atmosphere, one might expect a relation between the fading rate and diversity improvement. To check this hypothesis, 460-mc records for several days during the year were selected. For each sample the probability distribution was plotted. In Fig. 52 the diversity improvement is plotted against the fading rate of a single channel. It shows that, when the fading is relatively fast (40 fades per minute or more) the theoretical improvement of 6.25 per cent is approached. As the fading becomes slower, the diversity improvement becomes less. The data give the same results for the three periods showing no seasonal dependence.

Since the individual signals are Rayleigh-distributed for fading rates greater than 8 per minute, it is possible to calculate a cross-correlation coefficient for two signals from the observed diversity improvement (as

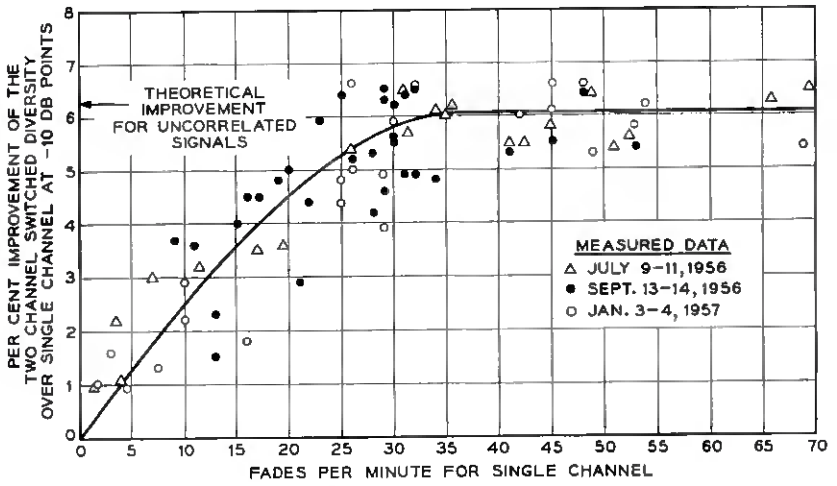


Fig. 52 — Improvement of twin-feed diversity system with horizontally disposed feeds as a function of the median fading rate, at 460 mc.

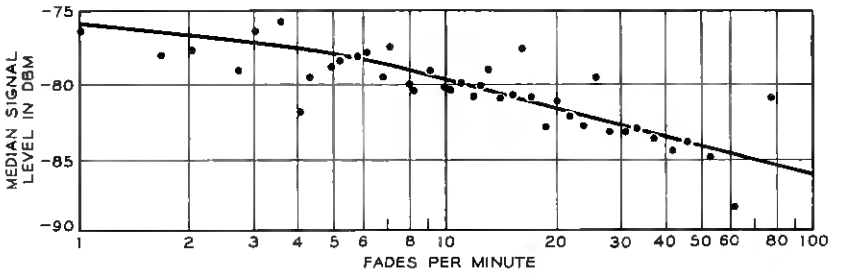


Fig. 53 — Median signal level as a function of the median fading rate. Each point represents the median signal level of 41 samples that had approximately the same fading rate at 460 mc — 1680 samples. Period: September 1956 through April 1957.

shown in Fig. 52) using S. O. Rice's theory.[18,19] The coefficient* of the envelope is indicated in this table:

Fades per Minute, Single Channel	0	20	25	30	35	70
Correlation coefficient, k^2	1.00	0.88	0.77	0.59	0.38	0.00

For two uncorrelated signals of equal median, the combined output is 0.45 per cent of the time below the level which is 10 db below the median of one of the channels. For a correlation coefficient of 0.6, the per cent of time increases to about 1 per cent, indicating that most of the diversity advantage is still obtained. For the system discussed, the correlation coefficient is less than 0.6 for fading greater than 30 cpm.

It has been observed that there is a seasonal dependence in the median signal level. During low-signal periods there is a tendency for the fading rate to be greater than during high-signal periods. If this trend is sufficiently strong, then this system will be as reliable as one which received uncorrelated signals at all signal levels, since a system has to be engineered only for low-signal levels.

An analysis was made to determine this effect. Fig. 53 shows a study of 1680 samples during the period of September 1956 to April 1957. This period was chosen because the median signal level is low during

* The correlation coefficient is calculated from the approximate expression:

$$P_j \cong [1 - k^2][(1 - e^{-u})^2 + k^2(u^2/2 - u^3/3 + u^4/8 - u^5/30)^2],$$

where

P_j = probability of the switched channel,
p = probability of a single channel,
$k^2 \approx$ correlation coefficient of the envelope,
$u \approx p/(1 - k^2)$.

This equation holds for $0 \leqq k^2 \leqq 0.9$ when $p < 0.1$.

the winter months. Each point represents the *median* signal level of 41 samples that had approximately the same fading rate. From 30 fades per minute (where diversity improvement begins to decrease) to 1 fade per minute, the increase in signal level is about 7 db. A trend is definitely apparent, but it is not significant enough to compensate for the increase in correlation between the signal channels.

One can conclude that the diversity is no longer effective for a fading rate less than 30 cpm. The percentage of time that the diversity is not fully effective can be obtained from Fig. 41(a); 80 per cent of the time the fading rate is less than 30 cpm. Horizontally spaced feed diversity at 460 mc is not useful from a systems viewpoint.

6.4 *Twin-Feed Diversity — Vertically Disposed Feeds*

In this section we discuss an experiment in which separate signals are obtained from two horns stacked vertically at the focus of the receiving paraboloid. The two feeds are placed one above the other; they are so aimed that the upper feed (in-line) receives the maximum signal level. Under these conditions, the lower feed (off-axis) receives a median signal level which is 10 db or so lower (Fig. 54). The diversity results from the fact that each receiving beam subtends a different volume in the atmosphere. The median signal level from the off-axis feed normally

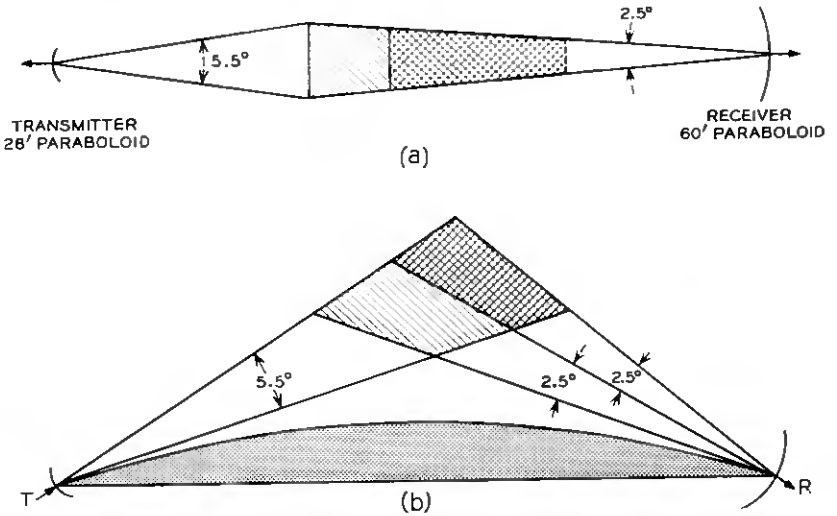


Fig. 54 — Common volume configuration for twin-feed diversity system with vertically disposed feeds, at 460 mc: (a) horizontal; (b) vertical.

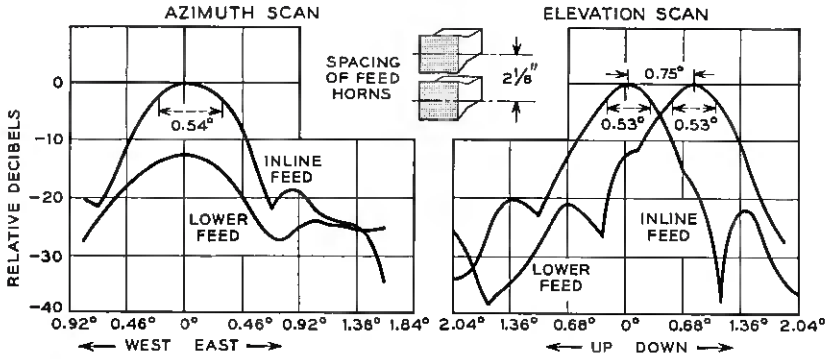


Fig. 55 — Free-space patterns for twin-feed diversity system with vertically disposed feeds using 28-foot antenna, at 4110 mc.

will be lower than that from the in-line feed because the power received from discontinuities at higher elevations is less (Section 4.1.2). The median signal level will vary between feeds depending on the strength of the atmospheric discontinuities.

6.4.1 *Experimental Setup and Analysis of Data*

Fig. 55 shows the antenna patterns taken on a line-of-sight path for twin feeds on the 28-foot antenna at 4 kmc. The signal is down 5.6 db at the crossover point. The level distribution counters are used to find the time that the signal spends below a preset level. The time when both channels are below the same level simultaneously is also measured on a coincidence counter. To check the theory the following is done:

i. The median signal level from each feed is recorded.
ii. The level at which the counters are set is recorded.
iii. The time the amplitude of the switched channel should spend below the level where the counter is set (if the signals are uncorrelated and Rayleigh-distributed) is computed using the curves of Fig. 47.
iv. The time the switched channel actually spends below the setting is read off the coincidence counter.

The percentage of time for the switched channel is computed with the aid of Fig. 47, using the medians of the two signals and the switched-channel level distribution counter setting; it is called "theoretical per cent of time." The percentage of time read off the coincidence counter is called "measured per cent of time." For uncorrelated signals, a plot of the measured versus the theoretical per cent of time should result in a 45° straight line.

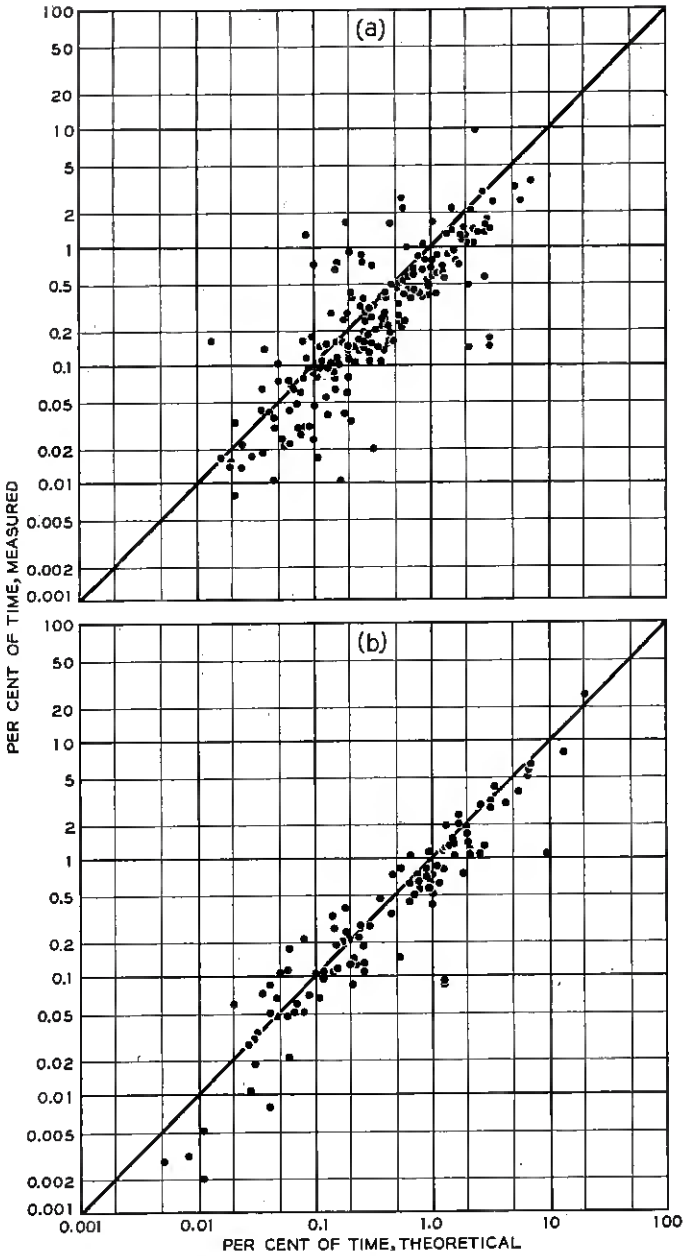


Fig. 56 — Effectiveness of twin-feed switch-diversity system with vertically disposed feeds, showing theoretical vs. measured percentage of time switched channel is below coincidence counter setting (counter set relative to median signal level of in-line single channel). Curves are theoretical for uncorrelated signals; points are experimental: (a) 460 mc; 28-foot–60-foot antennas; period: June through September 1957; (b) 4110 mc; 10-foot–28-foot antennas; period: April through June 1957.

6.4.2 *Results.*

Fig. 56 shows the results of a study of several months' data for both 460 and 4110 mc. The experimental points fall on both sides of the theoretical curve due to experimental inaccuracies. Several errors are introduced in recording the data, the most pronounced being that of estimating the median signal level.

The curves indicate that the predicted theoretical diversity improvement is realized in all cases; thus, the signals from the two vertically spaced feeds are not correlated.

It should be noted that the difference between the medians of the two feeds changes continuously as the structure of the atmosphere varies. Fig. 57 shows an extreme case. At 1:00 p.m. the difference between the medians on the feeds of the 28-foot antenna was about 8 db, with the in-line feed at the higher median signal level. By 4:00 p.m. the median of the off-axis feed had increased by 5 db, while the in-line feed

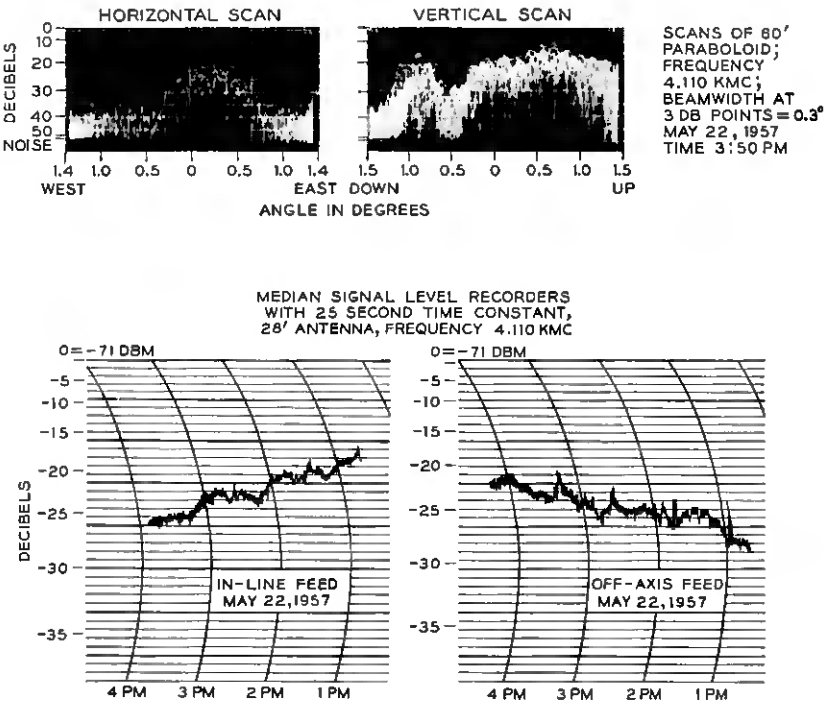


Fig. 57 — Segment of median signal recordings for twin-feed diversity system with vertically disposed feeds on the 28-foot antenna. Also shown are scans on the 60-foot antenna (at 4110 mc) taken at the same time.

median had dropped by 7 db, giving a difference of 4 db in favor of the off-axis feed. The lower part of Fig. 57 shows a scanning sequence using the 60-foot dish taken at 4:00 p.m. The vertical scan indicates that the maximum energy was coming from 0.6° up. The positioning of the feeds on the 28-foot antenna was such that the beam of the in-line feed was pointed 0.15° down and the off-axis feed was at 0.60° up. (The zero on the angle scale is arbitrary.*) This demonstrates why the off-axis feed was receiving the higher median signal in this case.

6.5 *Summary*

The twin-feed diversity studies give more evidence that the region in the common volume of two antenna beams used in a beyond-the-horizon system is composed of horizontal striations. Since the striations are essentially horizontal, uncorrelated signals will be received with the vertically placed feeds at both 460 mc and 4110 mc. These discontinuities are in the form of layers whose undulations are of such length and shape that they affect the propagation frequencies in different ways. Most of the time the layer size is large for 460 mc and small for 4 kmc. Under these conditions, diversity systems using horizontally disposed feed horns are effective at 4 kmc but not at 460 mc.

VII. CHARACTERISTICS OF SHORT-TERM FADING

In a beyond-the-horizon system using data transmission the length of fade at various levels is of importance in relation to the number of allowable errors.

Since rapid fluctuations in the level of received signal are always present, a problem exists in choosing the pulse width required for reliable transmission in systems using pulse-type modulation. For example, the simultaneous occurrence of a fade and the transmission of a pulse with a width less than the duration of the fade may result in the complete loss of that particular pulse. Accordingly, then, the choice of pulse width in these situations concerns the length of time the signal fades below a given level. If the length of fade is such that a sequence of pulses is lost, then knowledge of the length of fade will enable the system designer to devise the appropriate error-correcting code.

We will first outline the derivation for the number of crossings of the envelope of the instantaneous signal at any signal level. The results are given for the single channel, two-channel switch-type diversity and two- and four-channel combiner diversity — all for uncorrelated Rayleigh-

* See Section 4.1.1.

distributed signals with equal median signal levels. Results are also given for the average length of fade for these systems. A comparison is made between theory and experiment for the single channel and for the two-channel switch-type diversity. Finally, the system reliability is considered from the point of view of the length of fade.

7.1 *Number of One-Way Crossings and Average Length of Fade*

From the theory of random noise it is possible to find the number of crossings of the signal at a given level when the probability distribution of the signal is known. In our derivation, a Rayleigh distribution will be assumed for the envelope of the received signal. If we define $N(E)$ as the number of one-way crossings at a level E, then[20,21]

$$N(E) = (\sqrt{2\pi}\,\sigma E_r)(\text{probability density of } E). \tag{23}$$

This relation holds if the power spectrum is of gaussian form (σ is the standard deviation). With the aid of (14), we obtain for the single channel

$$\frac{dP_1}{dE} = \frac{2E}{E_r^2} \exp\left[-\left(\frac{E}{E_r}\right)^2\right]. \tag{24}$$

Let us write the number of one-way crossings $N(E)$ in terms of the median signal level E_m. Using (23), (24) and (14a), we obtain

$$N(E) = (8\pi \ln 2)^{1/2}\sigma \left(\frac{E}{E_m}\right) \exp\left[-\ln 2 \left(\frac{E}{E_m}\right)^2\right]. \tag{25}$$

At the median signal level, the number of one-way crossings will be

$$N(E_m) = (8\pi \ln 2)^{1/2}\sigma e^{-\ln 2} = \sqrt{2\pi \ln 2}\,\sigma. \tag{26}$$

Hence the number of one-way crossings at any level E with respect to the number of crossings at the median E_m is

$$n_1 = \frac{N(E)}{N(E_m)} = 2\left(\frac{E}{E_m}\right) \exp\left[-\ln 2 \left(\frac{E}{E_m}\right)^2\right]. \tag{27}$$

The average length of fade is the ratio of the probability distribution and the number of crossings.* For the single channel we have (14) divided by (27) [also using (14a)]:

$$\bar{t}_1 = \frac{1 - e^{-(E/E_r)^2}}{\frac{2}{\sqrt{\ln 2}}\left(\frac{E}{E_r}\right) e^{-(E/E_r)^2}}. \tag{28}$$

* Rice[22] has studied in detail the question of the distribution of the duration of fades.

TABLE IV — EQUATIONS FOR THE INSTANTANEOUS SIGNAL-LEVEL DISTRIBUTIONS, NUMBER OF ONE-WAY CROSSINGS AND AVERAGE LENGTH OF FADE FOR VARIOUS DIVERSITY SYSTEMS

Equations are based on uncorrelated Rayleigh-distributed channels with the same median. The equations are normalized to the variable at the median of the single channel. $E_m = \sqrt{\ln 2} E_r$, where E_m is median amplitude and E_r is rms amplitude.

Type of Combiner	Probability Distribution for Combined Channel	Number of One-Way Crossings	Average Length of Fade for $E/E_r < 0.3$
Single Channel	$1 - e^{-(E/E_r)^2}$	$\frac{2}{\sqrt{\ln 2}} \left(\frac{E}{E_r}\right) e^{-(E/E_r)^2}$	$\approx \frac{1}{2} \sqrt{\ln 2} \left(\frac{E}{E_r}\right)$
Two-Channel Switched	$[1 - e^{-(E/E_r)^2}]^2$	$\frac{4}{\sqrt{\ln 2}} \left(\frac{E}{E_r}\right) [e^{-(E/E_r)^2}][1 - e^{-(E/E_r)^2}]$	$\approx \frac{1}{4} \sqrt{\ln 2} \left(\frac{E}{E_r}\right)$
Two-Channel Combiner	$1 - e^{-(E/E_r)^2} \left[1 + \left(\frac{E}{E_r}\right)^2\right]$	$\frac{2}{\sqrt{\ln 2}} \left(\frac{E}{E_r}\right)^3 e^{-(E/E_r)^2}$	$\approx \frac{1}{4} \sqrt{\ln 2} \left(\frac{E}{E_r}\right)$
Four-Channel Combiner	$1 - e^{-(E/E_r)^2} \left[1 + \left(\frac{E}{E_r}\right)^2 + \frac{1}{2}\left(\frac{E}{E_r}\right)^4 + \frac{1}{6}\left(\frac{E}{E_r}\right)^6\right]$	$\frac{1}{3\sqrt{\ln 2}} \left(\frac{E}{E_r}\right)^7 e^{-(E/E_r)^2}$	$\approx \frac{1}{8} \sqrt{\ln 2} \left(\frac{E}{E_r}\right)$

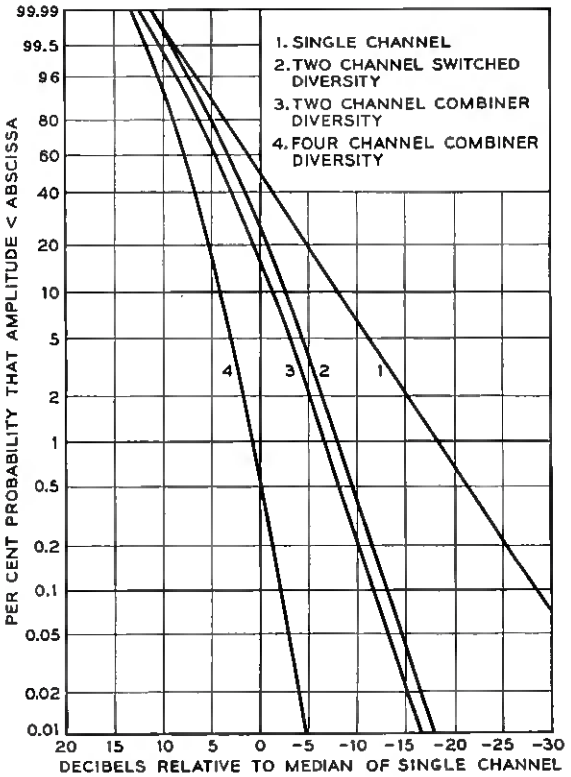


Fig. 58 — Rayleigh distributions of uncorrelated signals for various diversity systems. Curves are normalized to the median of a single channel. Results are for signal channels with the same median.

For small values of E/E_r ($E/E_r < 0.3$), (28) reduces to

$$\bar{l}_1 \approx \frac{1}{2}\sqrt{\ln 2}\left(\frac{E}{E_r}\right).$$

The equations for the probability distributions, number of crossings and average length of fade for the single channel, two-channel switch-type, two-channel and four-channel combiner diversity are listed in Table IV. The equations are plotted in Figs. 58, 59 and 60. The equations and curves are all for the combination of individual channels which have the same median signal level. (The case of unequal medians for the switch-type diversity was treated in Section VI.) In looking at the curves it must be remembered that the median level increases (as compared to a single channel) by 2.49 db for switched diversity, by 3.89 db for the

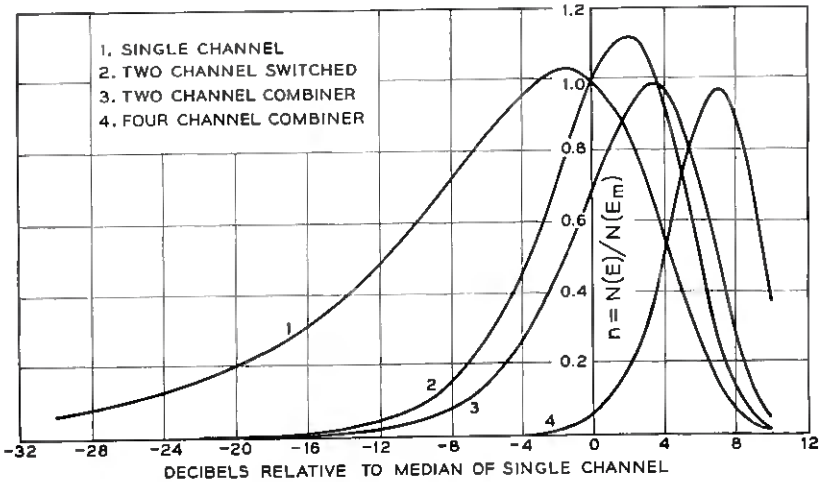


Fig. 59 — Number of one-way crossings, $N(E)$, of the instantaneous signal level for various diversity systems. Curves are for uncorrelated Rayleigh-distributed signal channels with the same median; curves are normalized to the number of crossings at the median of a single channel, $N(E_m)$.

two-channel diversity combiner and by 7.24 db for the quadruple diversity combiner.

This analysis applies for cw transmission. If FM or AM modulation gives a frequency diversity effect, then the results shown here should be modified to include it.

7.2 *Analysis of Experimental Results*

The measured number of crossings at various signal levels will now be compared with the theoretical values of a Rayleigh-distributed signal, given by (27).

Fig. 61 is a detailed analysis of a two-minute sample of the 460 mc signal with a fading rate of 1.2 cps; there is good agreement between the measured data and (27). The signal level as shown in Fig. 62 is, in fact, Rayleigh-distributed. Fig. 63 shows another comparison between data and theory. The agreement between the theoretical curve of (27) and the data points for the number of crossings is not good; a look at the corresponding signal-level distribution, Fig. 64, shows that the signal is not Rayleigh-distributed. In this sample the median fading rate is quite slow, hence a deviation from a Rayleigh distribution occurs, as discussed in Section 6.1. From the experimental signal-level distribution of Fig. 64 it is possible to calculate the corresponding distribution

of the number of crossings.[23] The results of this calculation are shown by the solid curve of Fig. 63; the agreement between the curve and the experimental points is good. It might be noted that the distribution of the number of crossings is very sensitive to a deviation from the Rayleigh distribution and offers a means of obtaining the distribution without the need for measuring the percentage of time the signals spends at various levels.

7.3 *Computed Average Duration of Fades*

The average length of time that the signal remains below a fixed threshold level in a system depends on two factors: the long-term

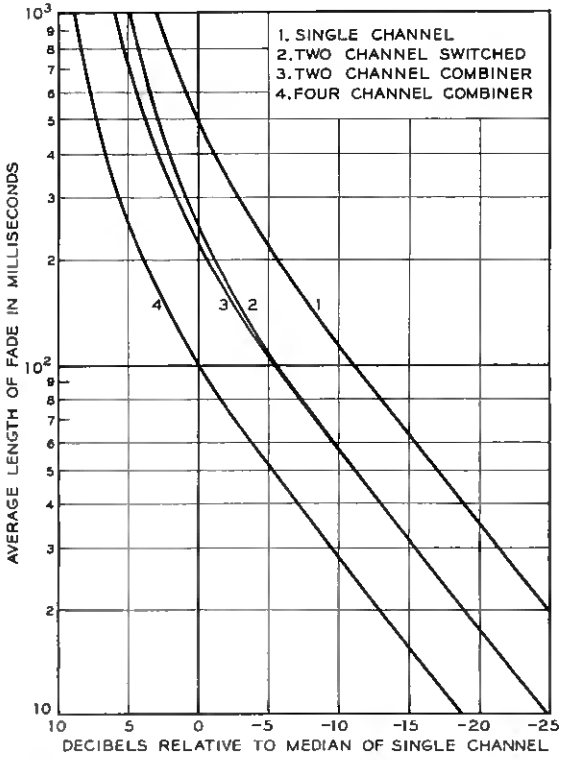


Fig. 60 — Average length of fade at a given level for various diversity systems. Curves are for uncorrelated Rayleigh-distributed signal channels with the same median; the curves are normalized for one crossing per second at the median of a single channel.

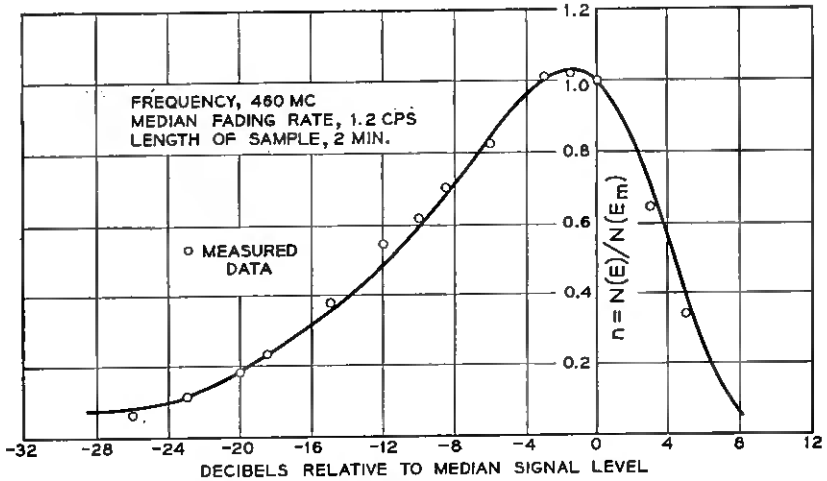


Fig. 61 — Number of one-way crossings of the instantaneous signal level for a single channel. The number of crossings is normalized to the number of crossings at the median level. Curve is theoretical [from (27)]; points are experimental.

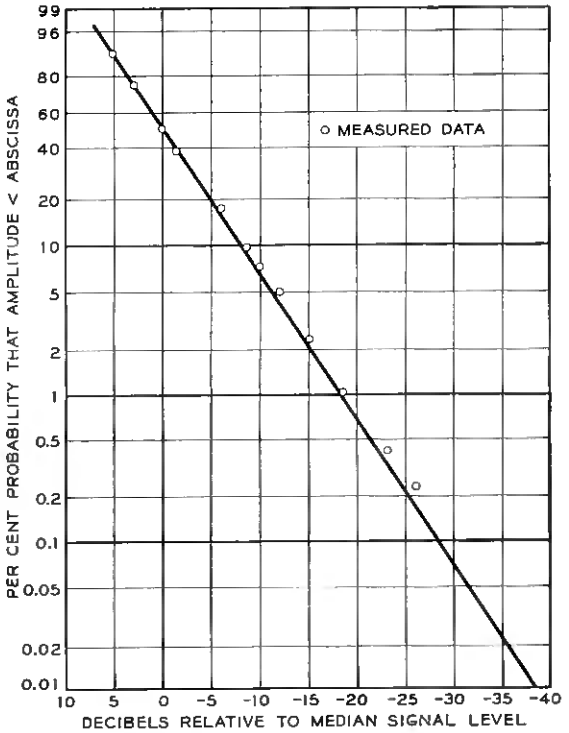


Fig. 62 — Distribution of instantaneous signal levels for the same sample analyzed in Fig. 61. Curve is Rayleigh distribution; points are experimental.

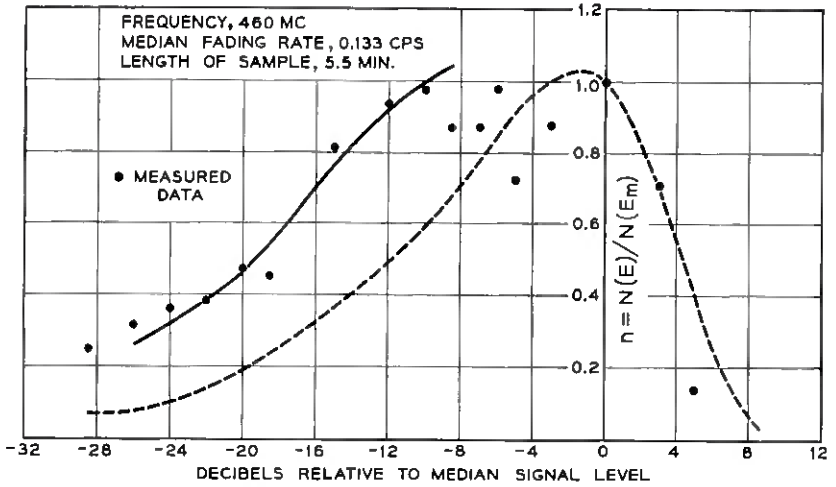


Fig. 63 — Number of one-way crossings of the instantaneous signal level for a single channel. The number of crossings is normalized to the number of crossings at the median level. Dashed curve gives the number of one-way crossings based on the Rayleigh distribution (27); solid curve is a calculation, for the number of crossings based on the measured distribution of Fig. 64. Frequency, 460 mc; median fading rate, 0.133 cps; length of sample, 5½ minutes.

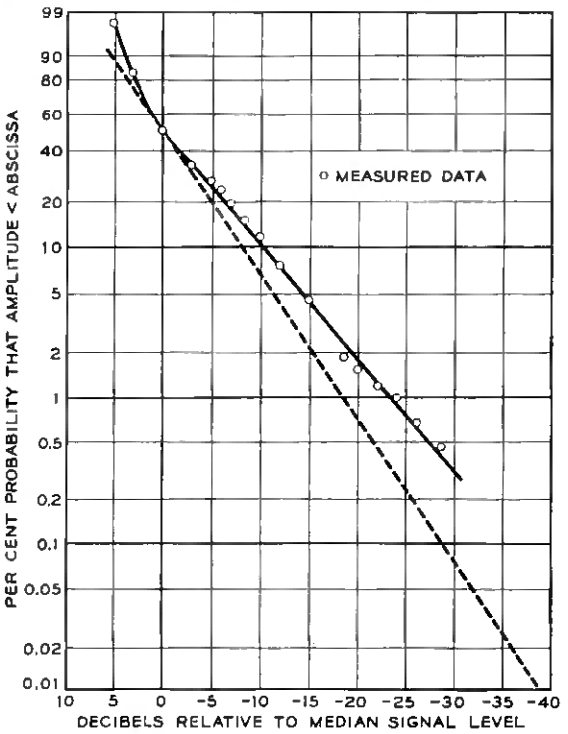


Fig. 64 — Distribution of instantaneous signal levels for the sample analyzed in Fig. 63. Dashed curve is Rayleigh distribution.

changes in the median level of the signal with respect to the fixed level and the rate of fading of the signal, assuming the median signal level to be fixed.

The following will show how the average length of fade at a given signal level can be computed using the distribution of fading rates (Fig. 41), the signal-level distribution (usually a Rayleigh distribution) and the number of times the signal crosses the particular level [from (27)].

In the following example it is arbitrarily assumed that, for a given system, the signal threshold during the worst period of transmission is at a level 18.5 db below the median (the 1 per cent point on the distribution curve). From the distribution of fading rates at 4 kmc (Fig. 41), the median fading rate of 150 cpm is obtained. The ratio of the number of crossings $N(E)$ at the threshold level (in this case -18.5 db) to the fading rate at the median $N(E_m)$ is obtained from Fig. 59. This value is 0.239; therefore

$$\frac{N(E)}{N(E_m)} = 0.239 \quad \text{and} \quad N(E) = 0.239\left(\frac{150}{60}\right) = 0.6 \text{ cps}$$

is the average number of crossings at the -18.5 db level.

The average length of fade is equal to the percentage of time the signal is below a given level, divided by the number of one-way crossings at this level. Since the time spent below the -18.5 db level is 1 per cent, $\bar{t} = 0.01/0.6 = 16.7$ milliseconds. Hence, 16.7 milliseconds is the average length of fade that will be exceeded 50 per cent of the time at the assumed threshold level of -18.5 db. A tabulation of the spread of average fade lengths to be expected using the 10 and 90 per cent values from the distributions of fading rate (Fig. 41) is given in Table V for both the 4110 and 460 mc signals.

From Table V, the average length of fade for a 4-kmc signal at the assumed threshold level of -18.5 db will be between 6.2 and 92 milli-

TABLE V — AVERAGE LENGTH OF FADE AT THE ASSUMED LEVEL OF 18.5 DB BELOW THE MEDIAN SIGNAL LEVEL (RAYLEIGH-DISTRIBUTED SIGNALS ARE ASSUMED)

Percentage of Time That Length of Fade Exceeds $\bar{t}$	4110 mc, $\bar{t}$ in Milliseconds	460 mc, $\bar{t}$ in Milliseconds
10	92	1590
50	16.7	233
90	6.2	61.2

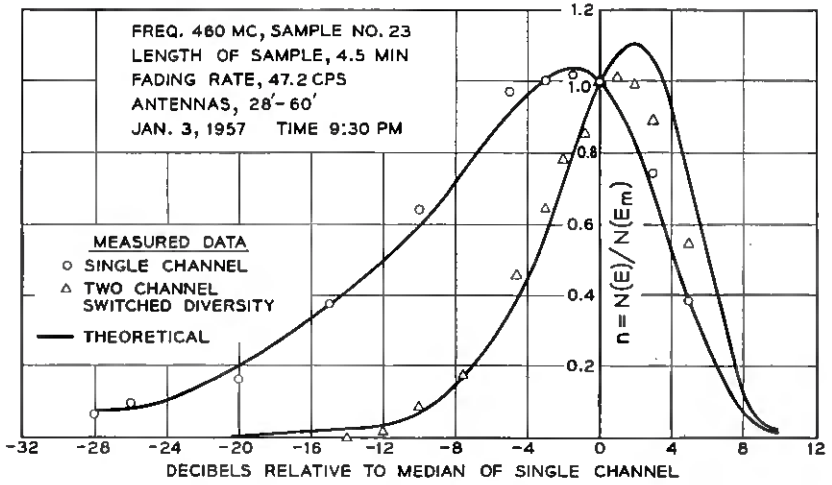


Fig. 65 — Number of one-way crossings at a given signal level for a two-channel switched diversity system. The number of crossings is normalized to the number at the median of the single channel. Curves are theoretical; points are experimental.

seconds for 80 per cent of the time. Similarly, a Rayleigh-distributed 460-mc signal at the same threshold level has average fade lengths between 61.2 and 1590 milliseconds for 80 per cent of the time.

7.4 *Twin-Feed Diversity — Average Length of Fade*

Fig. 65 shows the analysis of the number of crossings for a simultaneous recording of a switched channel and the single channel. (Fig. 51 gives the analysis of the probability distribution for this sample.) The agreement is very good between theory and experiment below the median level but does depart somewhat for the levels above the median. This is probably due to inaccuracies in receiver and recorder calibrations, since we depend on equal receiver calibration for quantitative data.

It is interesting to note that, while the number of crossings decreases very rapidly when more diversity channels are used, the average length of a fade stays fairly constant; this results from the fact that the time spent at low levels is also very small. This point is worth noting in the design of digital systems, since diversity will not reduce the average length of a drop-out as much as one might expect at first glance. We will discuss this aspect in the next section.

TABLE VI — CHARACTERISTICS OF DIVERSITY SYSTEMS FOR GIVEN VALUES OF POWER RELIABILITY

Diversity	Crossings*	Average Length of Fade, milliseconds†	Level Below Median, db‡	Reduction in Transmitter Power, db§
Power Reliability¶ 99.0 per cent				
None	0.23	42	18.5	0
Two-Channel Switched	0.148	73	8.0	10.5
Two-Channel Combiner	0.126	86	6.7	11.8
Four-Channel Combiner	0.028	116	−0.6	19.1
Power Reliability¶ 99.9 per cent				
None	0.0751	13	28.5	0
Two-Channel Switched	0.025	38	13.3	15.2
Two-Channel Combiner	0.023	44	11.8	16.7
Four-Channel Combiner	0.014	75	2.0	26.5

* Normalized to number of crossings at median of single channel.
† Normalized to one one-way crossing per second at median of single channel.
‡ With respect to median of single channel.
§ At the stated reliability.
¶ Defined as percentage of time the received power level stays above a given absolute power level.

7.5 *Reliability*

The usual definition of "reliability" is concerned with the percentage of time the received power stays above a given absolute level. Table VI illustrates the power reliability of a transmission system which employs diversity reception to reduce the power requirement on the transmitter. The received power at the output of the combined channel is maintained at the same absolute level as the order of diversity is changed. For example, for the 99 per cent reliability, the amount of transmitter power needed is 19 db less for quadruple combiner diversity than for a single channel. If we now look at the average length of fade, we note that it has increased about threefold. To keep the average length of fade constant when going from a single channel to the quadruple diversity combiner, a decrease of only 11.8 db in transmitter power could be realized (see Fig. 60). Table VI also shows the reliability factors when a 99.9 per cent power reliability is the design criterion.

VIII. BANDWIDTH IN TROPOSPHERIC PROPAGATION

In previous sections we have discussed the various characteristics of the received signals when cw signals are transmitted. In this section we consider the bandwidth characteristics of the propagating medium. To

make this study a frequency-sweep experiment* was performed using a 4.11-kmc transmitter sawtooth-modulated at a 1000 cps rate over a 20-mc band. The receiver was swept nonsynchronously over the same band at a 30-cps rate. The resultant pulses were displayed on an oscillograph and photographed at the rate of one frame every two seconds. The experiment used a 28-foot transmitting antenna and 8-, 28- and 60-foot receiving antennas.

Samples of sweep-frequency pictures are shown for various antenna combinations and transmission conditions. The bandwidths deduced from the experimental data are compared with a calculation based on the geometry of the common volumes.

An analysis of the sequences of frequency sweep photographs is made and the frequency auto-correlation function and distribution of bandwidths are computed. Photographs of signals received simultaneously from a twin-feed horizontal diversity system are also shown and discussed.

8.1 *Experimental Setup*

The transmitter consists of a klystron oscillator frequency-modulated by a 1000-cps sawtooth voltage applied to the repeller. This signal is amplified by a stagger-tuned klystron-amplifier. The output power is about 10 watts with a bandwidth of 10 mc at the 1-db points and 15 mc at the 3-db points. No ripples occur across the band.

The receiver consists of a triple detection set with two IF amplifiers whose center frequencies are 66 mc and 3 mc. The output of the second IF amplifier (150-kc bandwidth) is connected to the Y-axis of an oscilloscope (Fig. 66). The first beating oscillator is frequency-modulated at a 30- or 60-cps rate in a manner similar to the transmitter. These rates were chosen to be an order of magnitude higher than the fading rates usually encountered on this circuit.

With the aid of Fig. 67 we see that, whenever the difference between the transmitter frequency and the beating oscillator frequency is 66 mc, a pulse will appear on the oscilloscope connected to the last IF amplifier. The oscilloscope presentation is therefore amplitude versus frequency.

8.2 *Sweep-Frequency Photographs*

Samples of sweep-frequency photographs selected from runs taken on different days are shown in Fig. 68. Each run consisted of a series of 150

* It is possible to obtain equivalent information using pulse techniques but the interpretation of overlapping and distorted pulses is often difficult.[24]

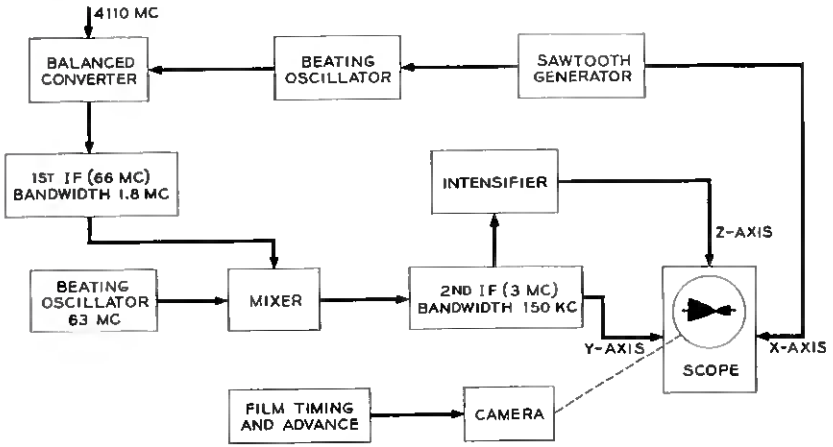


Fig. 66 — Block diagram of sweep-frequency receiver.

frames taken at $2\frac{1}{2}$ second intervals with a $\frac{1}{10}$ second exposure. In the first three sets, Fig. 68(a), (b) and (c), the antenna sizes (28-foot to 60-foot) remain fixed; these sets were taken on different days to show the variability of transmission conditions.

Fig. 68(a) shows the photographs for the 28-foot–60-foot combinations

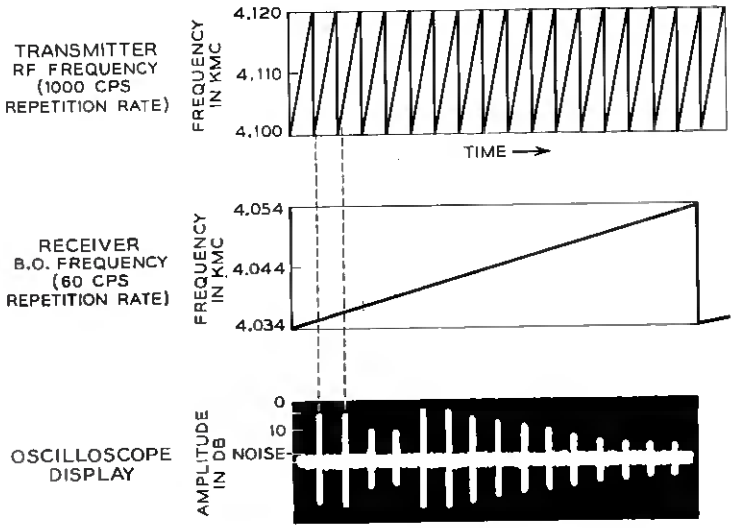


Fig. 67 — Combination of transmitter and receiver sweep-frequency waveforms.

on a day when the median signal level was high with a very slow fading rate (for a 4-kmc cw signal). Since the fading rate is slow, the sequence appears continuous. The signal goes to one deep fade for the samples shown. It should be noted that the amplitude fall-off at the extremities of the photographs is due to the limited bandwidth of the transmitter. In this case, the usable bandwidth of the medium exceeds 15 mc. The wide bandwidth suggests that rather large layers are primarily responsible for the propagation and that they are contained in a rather small height interval. In Fig. 68(b) we observe at times a narrower bandwidth (Photograph c) as well as a very wide one (D). The characteristic which is most noticeable is the variability from sample to sample. Had a continuous record been taken, continuous changes would have been noted. Fig. 68(c) shows, on the average, a broader bandwidth than the previous one. However, the narrowest bandwidth (A) is about 6 mc wide. (For the qualitative discussion we define the bandwidth as the frequency difference between two adjacent amplitude minima.)

Fig. 68(c) is for the 28-foot–60-foot antenna combination, as were all the preceding ones. This and the following figures were taken on the same day, with the receiving antenna as the variable. Fig. 68(d) shows the received signal using the 28-foot–28-foot antennas; these photographs were taken 7 minutes after those of Fig. 68(c). Here the narrowest bandwidth is about 3 mc (c). Fig. 68(e) shows the received signal on an 8-foot antenna a few minutes later. Due to the lower gain of the 8-foot antenna, the signal level is closer to the noise level of the system. However, one notes the narrowest bandwidth to be about 3 mc (c).

The photographs show that the medium introduces selective fading which bears a remarkable similarity to the fading which occurs on a line-of-sight path during anomalous propagation conditions.[25] This raises the question as to how many layer-like discontinuities are present at one time. The hypothesis that there are a limited number of regions of discontinuities does not necessarily contradict the fact that the received signal is Rayleigh-distributed as a function of time. In each region several layers can be present; these can change both in amplitude and position with time, therefore giving a Rayleigh distribution.

8.3 *Calculation of Bandwidth*

The bandwidth can be estimated by calculating the relative delay produced by the two limiting paths obtained from the common volume geometry. From the common volume concept it is apparent that the delays will be greatest in the vertical plane. Assuming the 3-db points of the antenna patterns to define the boundaries limiting the transmis-

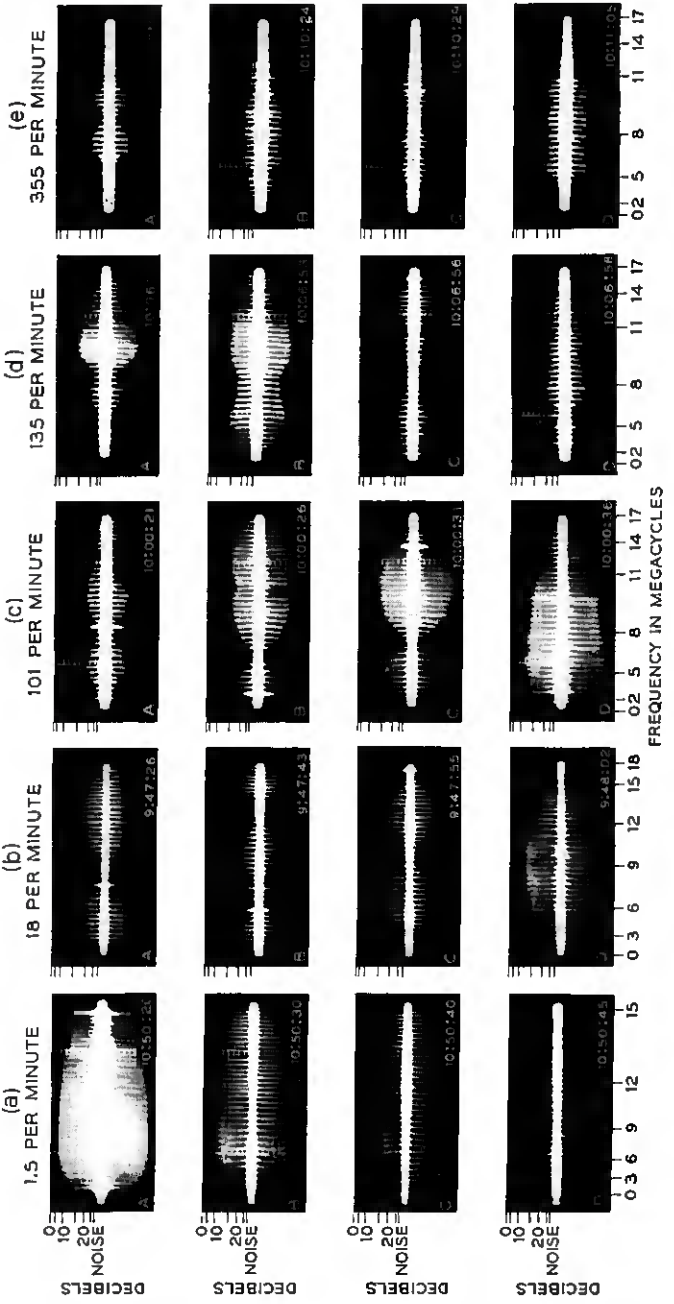


Fig. 68 — Sweep-frequency photographs: (a) September 30, 1957; 28-foot–60-foot antennas; cw median fading rate, 1.5 per minute. (b) October 15, 1957; 28-foot–60-foot antennas; cw median fading rate, 18 per minute. (c) November 8, 1957; 28-foot–60-foot antennas; cw median fading rate, 101 per minute. (d) November 8, 1957; 28-foot–28-foot antennas; cw median fading rate, 135 per minute. (e) November 8, 1957; 28-foot–8-foot antennas; cw median fading rate, 355 per minute.

sion, we calculate the maximum path difference, δ. With the aid of Fig. 13, we see that the path length difference δ will be

$$\delta = a\theta^2\left[\left(1 + \frac{\alpha_T}{\theta}\right)\left(1 + \frac{\alpha_R}{\theta}\right) - 1\right], \tag{29}$$

where α_T and α_R are the 3-db beamwidths in radians of the transmitting and receiving antennas and θ is the angle in radians between the lower edge of the beam at the horizon and the straight line joining the terminals. For our path, $2a = 171$ miles and $\theta = 0.9°$ (4/3 earth).

For the 28-foot–60-foot antenna combination $\alpha_T = 0.6°$ and $\alpha_R = 0.33°$. Using (29), we get $\delta = 142$ feet, which gives a corresponding bandwidth of 7.0 mc between adjacent amplitude minima. For the 28-foot–28-foot and 28-foot–8-foot, antenna combinations we note that the limiting factor in transmission is the atmosphere. From our beam-swinging experiment (Section 4.1.2), we found that the received power was contained on the average within an angle of 0.6° (half-power points) in the vertical plane. Hence, using $\alpha_T = \alpha_R = 0.6$, we obtain a maximum path difference of $\delta = 198$ feet; the corresponding bandwidth is 5.1 mc.

Depending on the structure of the atmosphere, these bandwidths will change. The actual bandwidth will be determined by the distribution and strength of the discontinuities in the common volume region. For example, if there were only one discontinuity, then one would expect a fairly large bandwidth. Nevertheless, the prediction of a bandwidth of the order of 5 to 7 mc is consistent with experimental data. In the next section a more quantitative analysis is made of the data.

8.4 *Statistical Distributions of Bandwidth*

While a qualitative presentation of bandwidth gives one a feeling for the capabilities of the propagating medium, a quantitative analysis should yield data which can form a basis for the design of communication systems.

The reduction of the photographic data consisted of the translation of the amplitude-versus-frequency function on each photograph into 30 discrete amplitude values; i.e., an amplitude value was obtained at $\frac{1}{2}$-mc intervals for the 15-mc band. The majority of the data reduction was accomplished using semiautomatic machines, which punched the coordinates for each discrete amplitude into IBM cards for computer processing.

A run of data for any antenna combination consisted of approximately 150 photographs. It should be noted that the median cw fading rate can

be of the order of magnitude of the rate of photographic sampling; hence one must insure that the number of photographs analyzed is sufficient to obtain a valid distribution. A test was therefore made to determine the amount of data needed for analysis. The cw fading rate for the run selected was 85 fades per minute; the rate of photographic sampling was about 25 frames per minute. For this analysis the same frequency was selected on each frame and a signal level distribution was plotted for the first 50, 100 and 160 frames. As the number of frames included in the distribution was increased, the level distribution curve approached the Rayleigh curve. While there was some deviation from the theoretical distribution, no significant change was introduced with the increase in sample length from 100 to 160 frames. One therefore concludes that 100 photographs are sufficient for analyzing the bandwidth. This test is perhaps more stringent than necessary when computing the frequency autocorrelation function, since each photographic frame contains 30 frequency points instead of only one.

There were two main analyses made on the data. One was the frequency autocorrelation function, the other was the amplitude-frequency distribution.

8.4.1 *Frequency Autocorrelation Function*

We define the frequency autocorrelation function of the envelope

$$\tau = \frac{\int_{-\infty}^{\infty} v_m(t)v_n(t)\ dt}{\int_{-\infty}^{\infty} v_m{}^2(t)\ dt}, \tag{30}$$

where v is the amplitude function of the envelope; m and n are different frequencies (see Fig. 69) and τ is a function of the frequency difference $m - n$.

Since the data are quantized in $\frac{1}{2}$-mc intervals and only a limited frequency band is available, the integrals are converted to these summations:

$$\tau = \frac{\sum_{N=1}^{M} v_m(t_N)v_n(t_N)}{\frac{1}{2}\sum_{N=1}^{M} [v_m{}^2(t_N) + v_n{}^2(t_N)]}. \tag{31}$$

The averaging is done in the denominator to smooth out transmitter amplitude fluctuation. This method works best for $m - n$ small. Due to the decrease in power output of the transmitter at the edges of the band,

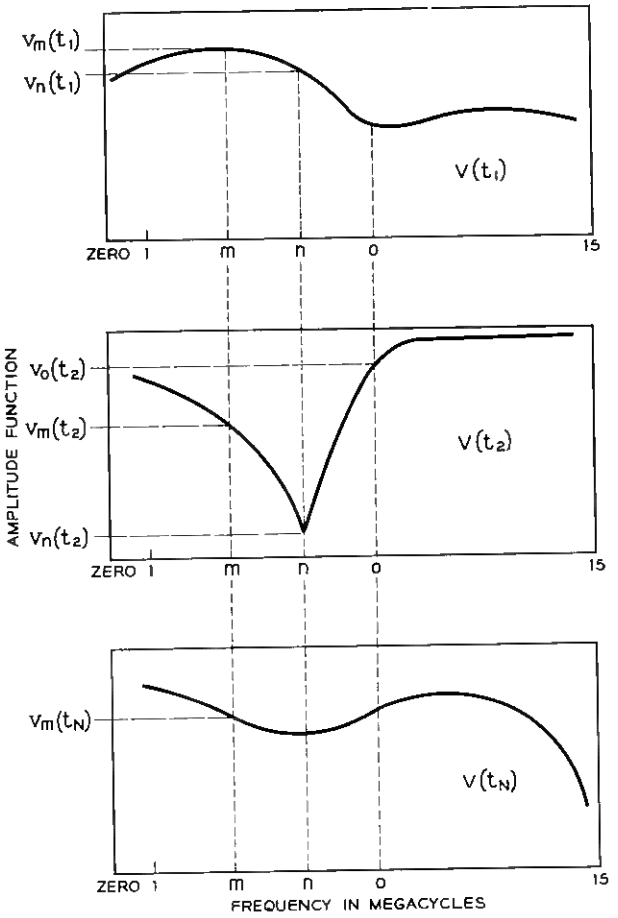


Fig. 69 — Sweep-frequency patterns; $v(t)$ is the amplitude function.

only the center 10 mc are used in the computation. The τ's for the same $m - n$ are calculated and averaged for each run. For $m - n = \frac{1}{2}$ mc (the smallest increment) one obtains 3000 values from a 150-frame sample when the center 10 mc are used in the computation; for $m - n = 10$ mc only 150 values are available.

From the frequency autocorrelation function, τ, of the envelope, we can compute the frequency autocorrelation coefficient, k. The equation relating the two is given[26] by

$$\tau = E(k) - \left(\frac{1 - k^2}{2}\right) K(k), \tag{32}$$

where $E(k)$ and $K(k)$ are the complete elliptic integrals of modulus k. Booker, Radcliffe and Shinn[27] have shown that k^2 is approximately the correlation coefficient of the envelope. It can be shown from (32) that as the frequency interval $m - n$ becomes large, $\tau \to \pi/4$.

In all, 10 runs were analyzed, using different combinations of receiving antennas. The runs were chosen to give a wide range of transmission conditions, and the cw fading rate and the median signal levels were the parameters measured before and after each frequency-sweep run.

Fig. 70 shows the autocorrelation function for the 28-foot–60-foot combination. Except for the data of September 30, 1957, the curves for the autocorrelation function do not change appreciably. The unusual propagation conditions on September 30, 1957, were discussed in Section 8.2 in conjunction with Fig. 68(a). We note that $\tau \to \pi/4$ as $(m - n) \to 10$ mc.

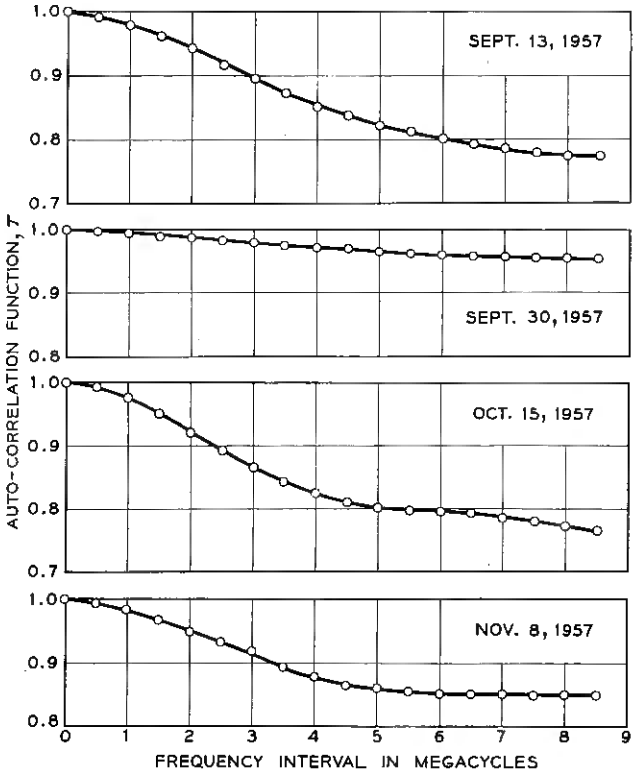


Fig. 70 — Frequency autocorrelation functions, 28-foot–60-foot antennas.

TABLE VII — PARAMETERS IN SWEEP-FREQUENCY RECEIVED DATA

frequency: 4110 mc
cw transmitting power: 43 dbm (10-ft transmitting antenna)
$k^2 \approx$ correlation coefficient of envelope of received signal

Antennas, feet	Date	cw Median Signal (−dbm)		cw Median Fading Rate, per Minute	Bandwidth in mc at Correlation Coefficient; $k^2 = 0.6$
		Before	After		
28–60	9-13-57	97.5	93	85	5.2
	9-30-57	58	65	1.5	≅16.0*
	10-15-57	96	95	18	4.5
	11-8-57	89	89	101	6.3
28–28	9-13-57	92	94	148	7.4
	9-30-57	68.5	62	1.25	≅12.5*
	10-15-57	99.5	102	18	5.6
	11-8-57	92.5	94	135	5.5
28–8	9-30-57	70	73.5	3	≅16.0*
	11-8-57	99	96	355	4.5

* $k^2 = 0.8$

Table VII summarizes the bandwidth for the various runs. A more meaningful definition is used than that employed in our qualitative analysis of bandwidth in Section 8.3. The bandwidth may be defined as the frequency separation for which the correlation coefficient $k^2 = 0.6$ ($\tau = 0.904$). [Since the bandwidth was extremely large on September 30, 1957, we have used the criterion $k^2 = 0.8$ ($\tau = 0.952$).] It is seen that the bandwidth varies from 4.5 to 6.3 mc (omitting the data of September 30, 1957) and does not appear to be correlated with cw median signal level, cw median fading rate or antenna size.

8.4.2 *Amplitude-Frequency Characteristics*

The analysis of the amplitude-frequency characteristics consisted of calculating the maximum ratio of signal amplitudes in a given frequency interval $m - n$. A cumulative distribution is plotted as a function of $m - n$. The distribution of slopes can be obtained by looking at the results for the smallest frequency interval ($\Delta f = \frac{1}{2}$ mc). For larger intervals the results do not give the slopes but only the ratio of amplitudes. [In frame 2 of Fig. 69 for $m - o$ the maximum amplitude ratio is $v_o(t_2)/v_n(t_2)$.]

Fig. 71 shows the amplitude-frequency distributions. These give the frequency separation for various amplitude ratios which are exceeded 10, 50 and 90 per cent of the time. Figs. 72 and 73 give the corresponding

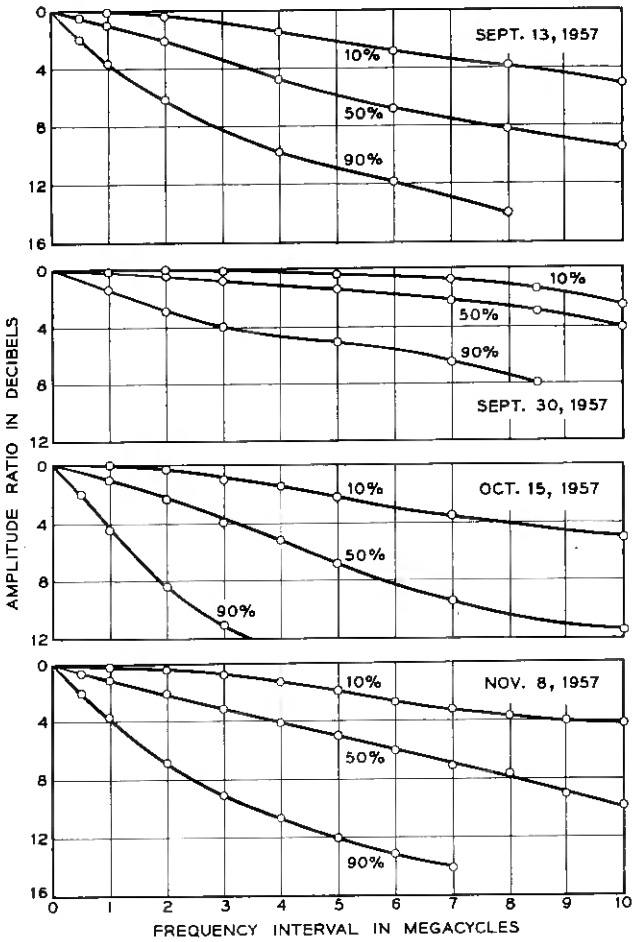


Fig. 71 — Bandwith exceeded 10, 50 and 90 per cent of the time — 28-foot–60-foot antennas.

distributions for the 28-foot–28-foot and 28-foot–8-foot combinations. While the curves differ somewhat, there are no significant differences except for the curve for September 30, 1957.

Fig. 74 gives the amplitude ratio distribution for the 28-foot–60-foot antenna combination on September 13, 1957. This curve gives the distribution for a day when the bandwidth was about average.

Except for the unusual propagation of September 30, 1957, the auto-correlation functions, as well as the bandwidth distributions, are sensibly independent of the median signal level, cw fading rate and antenna

sizes used in these tests. To obtain a wider bandwidth, one must use extremely narrow beam antennas on the order of 0.1°. In that case, one probably would expect a marked increase in bandwidth with antenna size.

8.5 *Synthesis of Sweep-Frequency Photographs*

One of the interesting questions in the analysis of the sweep-frequency records is the number of signal components present in the received signal

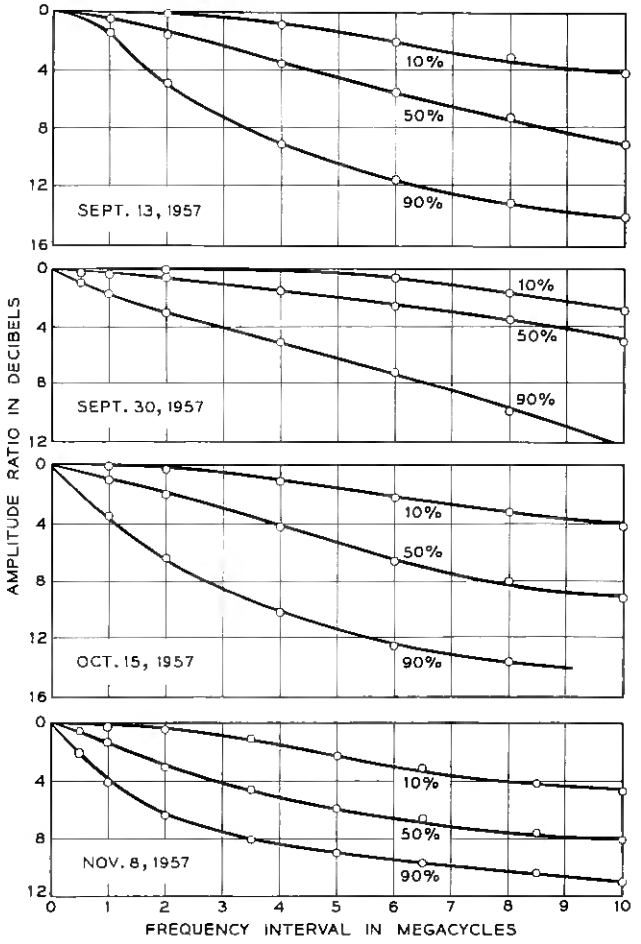


Fig. 72 — Bandwidth exceeded 10, 50 and 90 per cent of the time — 28-foot-28-foot antennas.

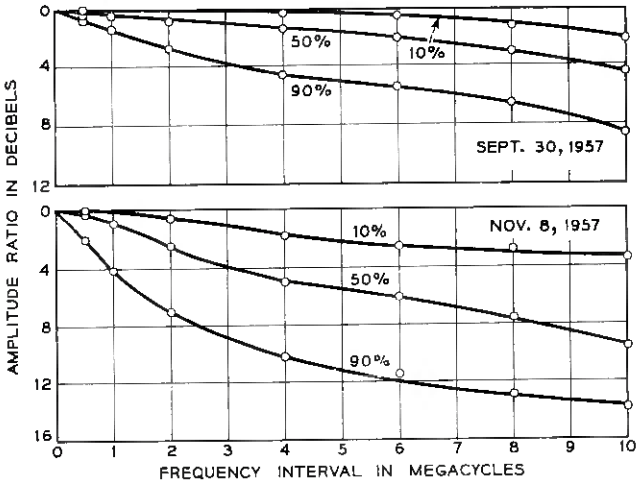


Fig. 73 — Bandwidth exceeded 10, 50 and 90 per cent of the time — 28-foot–8-foot antennas.

at any one instant. We indicated that the photographs bore a striking similarity to those obtained on a line-of-sight circuit under anomalous propagating conditions.[25] We have synthesized a sequence of seven photographs from the run of November 8, 1958. The analog computer was the one used previously by Crawford and Jakes,[25] which combines four components of arbitrary amplitude and phase. Fig. 75 shows the synthesis of one of the samples. It was possible to synthesize all seven photographs approximately with only four components. Due to the limited band of the frequency sweep, it is difficult to tag the components with small delays, but the four components give an adequate representation of those with large amplitudes and relatively short delay.

8.6 *Twin-Feed Diversity*

Twin-feed diversity experiments using cw transmissions have shown that signals from the two received channels are uncorrelated most of the time at 4 kmc. There has been speculation that a frequency effect gives an additional improvement in signal to noise for a single FM channel.[28] If the RF energy spectrum is distributed over several megacycles and if the dropout at low signal levels is only several hundred kilocycles wide, then only a small part of the signal information may be lost. This is in contrast to a cw channel, where the complete channel would be lost.

To examine the frequency correlation between two diversity channels, tests were made using the 28-foot transmitting antenna and the 28-foot receiving antenna with horizontally disposed twin feeds. Simultaneous recordings were made on the two sweep-frequency channels using a dual trace oscilloscope. Photographs were taken with a Polaroid camera, using a $\frac{1}{25}$-second exposure. Fig. 76(a) shows the calibration for the photographs. Several frames have been selected to illustrate the frequency diversity characteristics. Figs. 76(b), (c) and (d) were taken on a day when the bandwidth was rather broad, whereas the remaining ones were

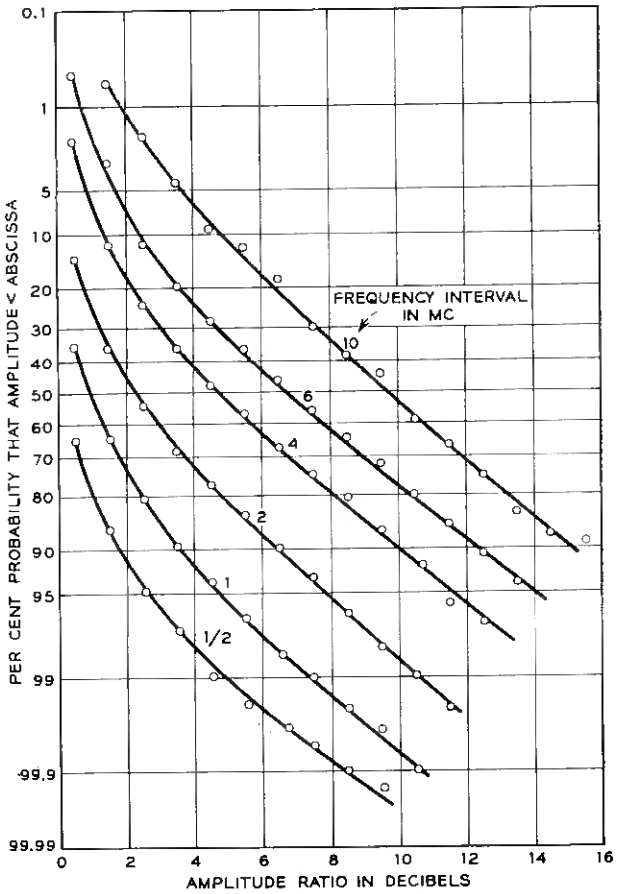
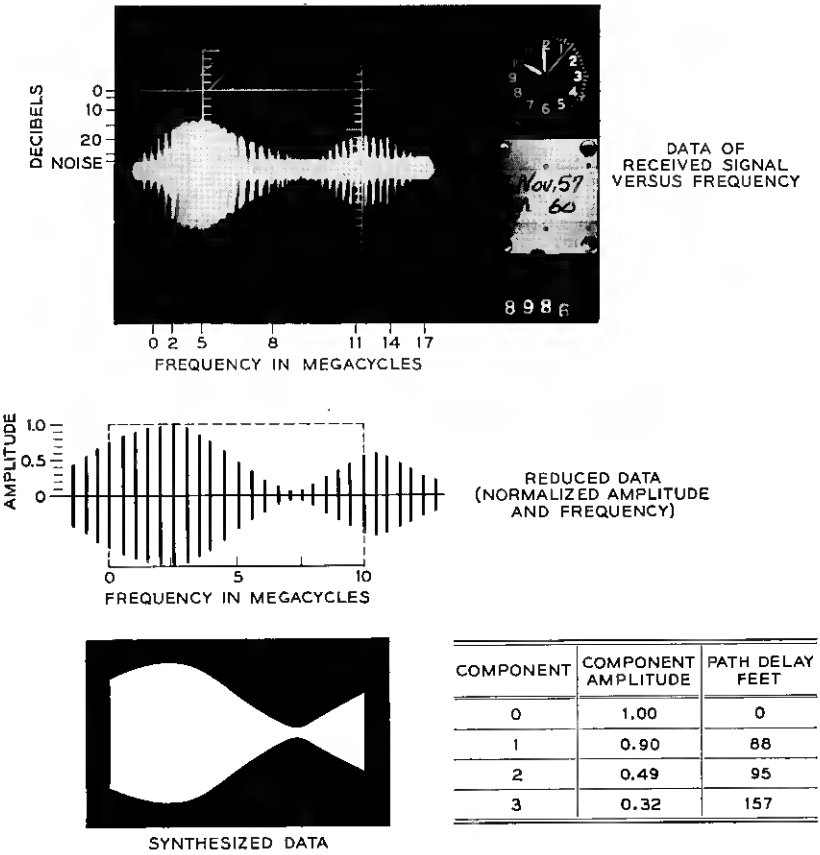


Fig. 74 — Distribution of amplitude ratios for a run where the bandwidth is about average.

recorded on a more representative day. In Figs. 76(b) and (f) the amplitude-frequency characteristics are very similar for the two feeds; on the remaining ones there is little similarity. In all the runs the signals on the two feeds were uncorrelated at any particular frequency. An examination of several hundred frames shows that a frequency effect does occur between the two diversity channels when the frequency separation is of the order of several megacycles at 4 kmc.

If the received channels are combined in a switch-type combiner, the percentage of time that the switched signal (at one frequency) spends below a certain level will decrease and the bandwidth will be broader.



COMPONENT	COMPONENT AMPLITUDE	PATH DELAY FEET
0	1.00	0
1	0.90	88
2	0.49	95
3	0.32	157

Fig. 75 — Synthesis of a sweep-frequency photograph.

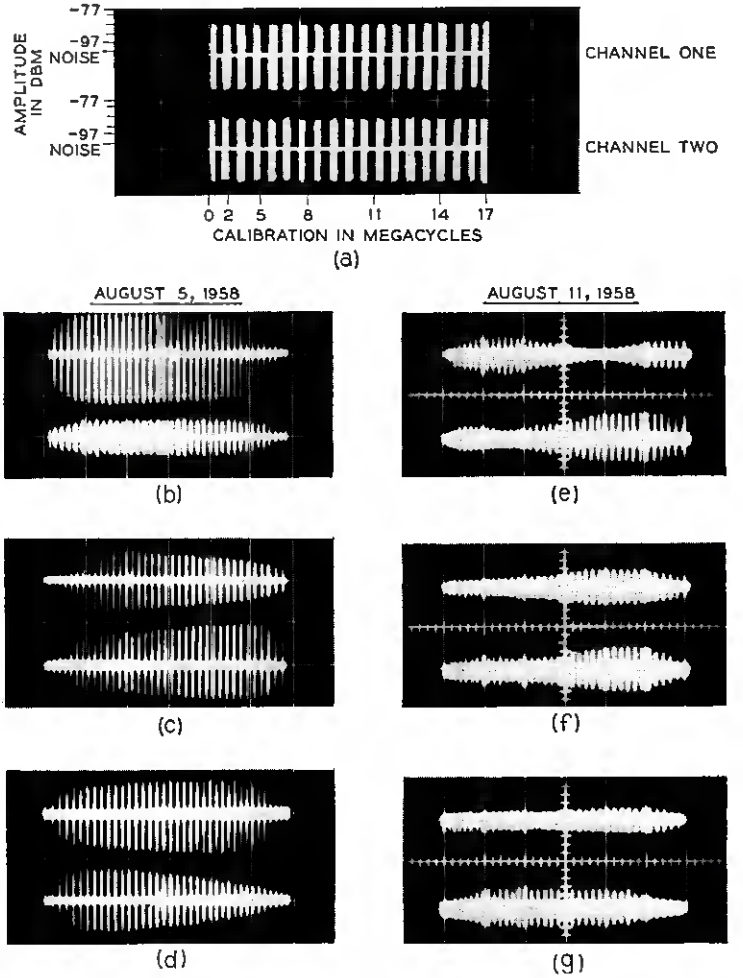


Fig. 76 — Sweep-frequency photographs of single channels in a twin-feed diversity system — horizontally disposed feeds, 4110 mc.

IX. CONCLUDING REMARKS

The results of the several experiments discussed in the foregoing sections are interrelated in various ways, but it is perhaps most significant that the results can be interpreted in terms of reflections from layers in the troposphere. In many cases, such as the dependence of the propagation on wavelength, antenna size and beam orientation, quantitative

agreement is obtained between experiment and the theory of reflection by layers of intermediate size. The 4110-mc data on bandwidth and on fading rate are consistent with the theory. Moreover, one concludes that delays in the vertical plane of propagation are most important in determining the bandwidth, whereas delays in the horizontal plane determine the fading rate. The propagation mechanism itself determines the bandwidth and fading rate unless antennas with extremely narrow beamwidths are employed. The 460-mc experiments on twin-feed diversity and those dealing with the dependence of received power and fading rate on wavelength indicate that the effective horizontal extent of the layers increases significantly as the frequency of propagation is decreased. Where possible, the influence of these concepts on the design of systems has been discussed.

X. ACKNOWLEDGMENTS

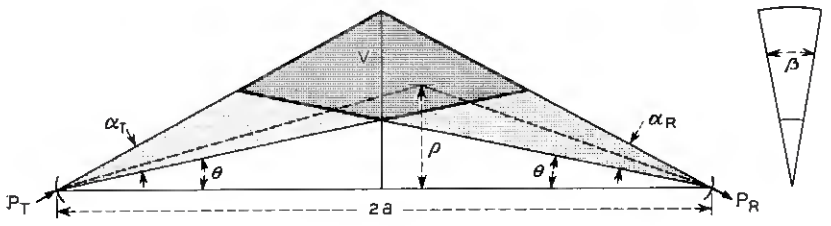
We wish to express appreciation to H. T. Friis for his active leadership and encouragement in this work. We also wish to thank R. P. Booth of the Military Systems Engineering Department of Bell Telephone Laboratories for his continued interest and for making available some of the equipment needed for the experiment. Several members of the Radio Research Department assisted in this program. R. A. Semplak, who was a full-time participant, built and maintained much of the apparatus and obtained and analyzed much of the data. L. R. Lowry, R. A. Desmond and J. H. Hammond participated in some phases of the experiment and analyzed some of the data. Others active in the program from time to time were W. E. Legg and H. A. Gorenflo. The cooperation of Orton J. Newton and his family, on whose property the transmitting site was located, was most valuable and greatly appreciated.

APPENDIX

Calculation of Received Power Using Unequal Transmitting and Receiving Antennas

The formula for received power using antennas with equal beamwidths is given by (1) in Section 3.1.3 and in the theory (Ref. 5, p. 635, Equations 5 and 7):

$$P_R = P_T M a A_T A_R \lambda^{-1} \int_V \rho^{-6}\, dV = P_T M a A_T A_R \lambda^{-1} \left[\frac{\beta}{6\theta^4 a^3} f\left(\frac{\alpha}{\theta}\right) \right]. \quad (33)$$



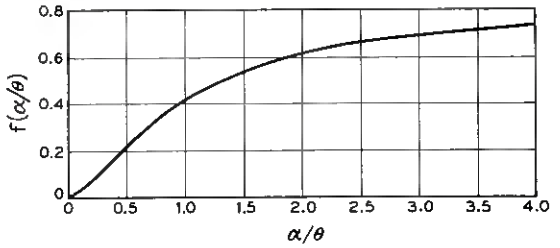


Fig. 77 — Power received beyond the horizon on antennas of equal size, $\alpha_R = \alpha_T$ for symmetrical volume $V = V(\alpha, \theta)$.

The symbols are defined in Section 3.1.3 and in Fig. 77, and A_T and A_R are the effective areas of the antennas. For antennas of equal size with beams of unity aspect ratio (see Ref. 5, pp. 635–636):

$$A_T = A_R = \frac{\lambda^2}{\alpha_T{}^2} = \frac{\lambda^2}{\alpha_R{}^2}.$$

The factor

$$\left[\frac{\beta}{6\theta^4 a^3} f\left(\frac{\alpha}{\theta}\right)\right]$$

depends upon the magnitude and location of the symmetrical volume, $V(\alpha,\theta)$; a plot of $f(\alpha/\theta)$ [see (2)] is shown in Fig. 77. For the idealized antenna beams used in the theory, the wedge angle β is given by

$$\beta = \frac{2\alpha}{\alpha + 2\theta}$$

where α is the 3-db beamwidth of the actual antennas.

Consider now the case of unequal transmitting and receiving beams, as shown in Fig. 78. The resulting common volume is asymmetrical about the center line OA, thus one cannot apply (33) directly. However, the asymmetrical volume can be expressed as a sum of symmetrical

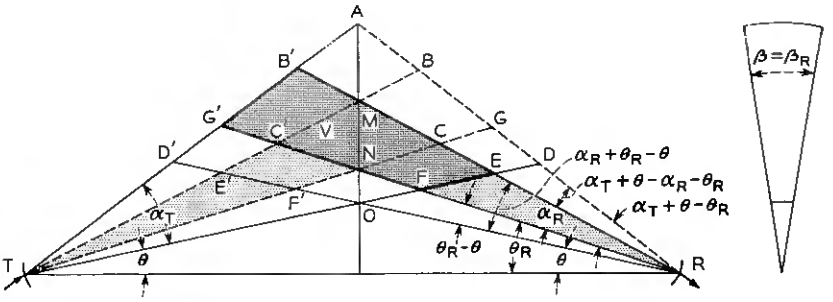


Fig. 78 — The asymmetrical volume B'EFG' expressed in terms of symmetrical volumes, valid when $\theta_R + \alpha_R \leqq \theta + \alpha_T$ and $\beta_R < \beta_T$ (unequal transmitting and receiving antennas). We require V = B'EFG' = ADOD' − ADEB' − G'D'OF. Now, ADEB' = ABMB' + BDEM and ADOD' − MEOE' = B'D'E'M + BDEM + ABMB'. Therefore, ADEB' = ½(ADOD' − MEOE' + ABMB'). Similarly, G'D'OF = ½(ADOD' − AGNG' + NFOF'). Therefore, V = ½(AGNG' + MEOE' − ABMB' − NFOF') = ½[$V_1(\alpha_T + \theta - \theta_R, \theta_R) + V_2(\alpha_R + \theta_R - \theta, \theta) - V_3(\alpha_T + \theta - \alpha_R - \theta_R, \theta_R + \alpha_R) - V_4(\theta_R - \theta, \theta)$].

volumes as shown in the figure. If both beams are oriented along the horizon, $\theta_R = \theta$ and the asymmetrical volume becomes

$$V = \tfrac{1}{2}[V_1(\alpha_T, \theta) + V_2(\alpha_R, \theta) - V_3(\alpha_T - \alpha_R, \theta + \alpha_R)]. \quad (34)$$

The received power is obtained by substituting the appropriate angles from the V_1, V_2, V_3 into the factor in (33) and summing. Thus

$$P_R = P_T M A_T A_R \lambda^{-1} \frac{\beta_R}{12a^2}\left[\frac{1}{\theta^4} f\left(\frac{\alpha_T}{\theta}\right) + \frac{1}{\theta^4} f\left(\frac{\alpha_R}{\theta}\right) - \frac{1}{(\theta + \alpha_R)^4} f\left(\frac{\alpha_T - \alpha_R}{\theta + \alpha_R}\right)\right], \quad (35)$$

which holds for $\beta_R < \beta_T$. For this case the effective areas of the idealized antennas with beams of unity aspect ratio are

$$A_T = \frac{\lambda^2}{\alpha_T{}^2} \quad \text{and} \quad A_R = \frac{2\lambda^2}{\alpha_R \beta_R (\alpha_R + 2\theta)},$$

which results in

$$P_R = P_T \frac{M}{6a^2} \frac{\lambda^3}{\alpha_T{}^2} \frac{1}{\alpha_R(\alpha_R + 2\theta)} \left(\frac{1}{\theta^4}\right)\left[f\left(\frac{\alpha_T}{\theta}\right) + f\left(\frac{\alpha_R}{\theta}\right) - \frac{\theta^4}{(\theta + \alpha_R)^4} f\left(\frac{\alpha_T - \alpha_R}{\theta + \alpha_R}\right)\right]. \quad (36)$$

Special treatment is required when the elevation and beam angles

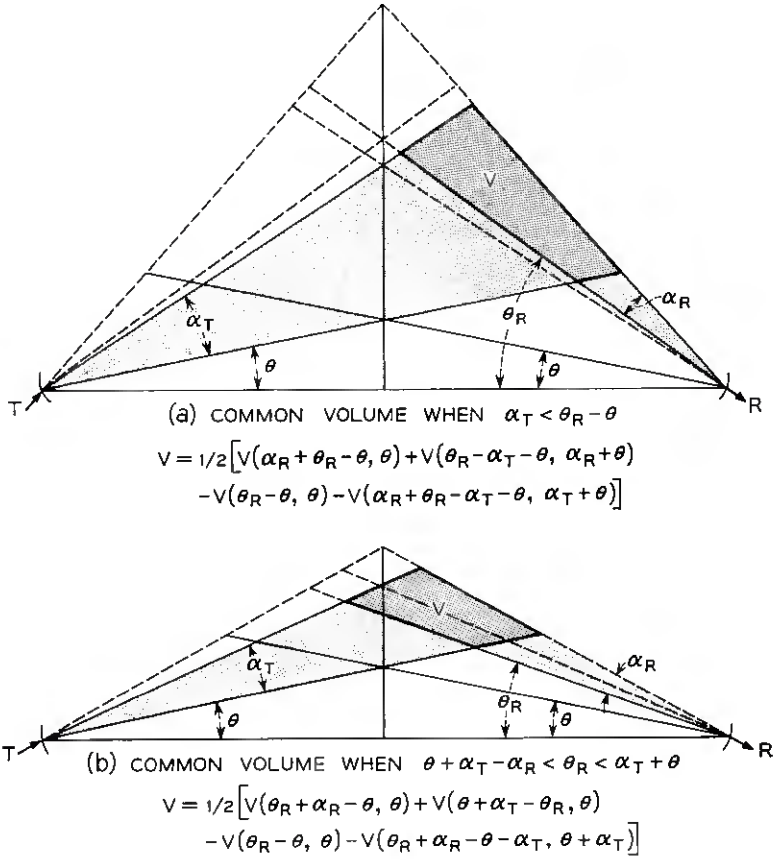


Fig. 79 — Asymmetrical volumes — special cases.

involved are such that the apex of the common volume in Fig. 78 is situated to the right of centerline OA; the resulting common volumes are shown in Fig. 79. Again, one can express these asymmetrical volumes in terms of symmetrical ones. The resulting expressions for the common volume and the range of validity are given in the figure; with these, (33) can be applied to obtain the received power.

REFERENCES

1. Bullington, K., Propagation of UHF and SHF Waves Beyond the Horizon, Proc. I.R.E., **38,** October 1950, p. 1221.
2. Proc. I.R.E., **43,** October 1955 (entire issue).
3. Booker, H. G. and Gordon, W. E., A Theory of Radio Scattering in the Troposphere, Proc. I.R.E., **38,** April 1950, p. 401.

4. Carroll, T. J. and Ring, R. M., Propagation of Short Radio Waves in a Normally Stratified Troposphere, Proc. I.R.E., **43,** October 1955, p. 1384.
5. Friis, H. T., Crawford, A. B. and Hogg, D. C., A Reflection Theory for Propagation Beyond the Horizon, B.S.T.J., **36,** May 1957, p. 627.
6. Crawford, A. B., Friis, H. T. and Jakes, W. C., Jr., A 60-Foot Diameter Parabolic Antenna for Propagation Studies, B.S.T.J., **35,** September 1956, p. 1199.
7. Bullington, K., Radio Transmission Beyond the Horizon in the 40 to 4000 mc Band, Proc. I.R.E., **41,** January 1953, p. 132.
8. Nupen, W. and Thuronyi, G., Checklist of References to Literature on Tropospheric Propagation of UHF, VHF and SHF Radio Waves, Bulletin 6001, National Bureau of Standards, Boulder, Colo., August 1958.
9. Ortwein, N. R., An Annotated Bibliography of Literature Pertinent to Tropospheric Scatter Propagation, 1945–1957, N.E.L. Report 858, August 1958.
10. Bean, B. R. and Meany, F. M., Some Applications of the Monthly Median Refractivity Gradient in Tropospheric Propagation, Proc. I.R.E., **43,** October 1955, p. 1419.
11. Bolgiano, R., Jr., Wavelength Dependence in Transhorizon Propagation, Proc. I.R.E., **47,** February 1959, p. 331.
12. Waterman, A. T., Jr., A Rapid Beam-Swinging Experiment in Transhorizon Propagation, Trans. I.R.E., **AP-6,** October 1958, p. 338.
13. Rice, S. O., Radio Field Strength Statistical Fluctuations Beyond the Horizon, Proc. I.R.E., **41,** February 1953, p. 274.
14. Norton, K. A., Rice, P. L., Janes, H. B. and Barsis, A. P., The Rate of Fading in Propagation Through a Turbulent Atmosphere, Proc. I.R.E., **43,** October 1955, p. 1341.
15. Anderson, L. J. and Smyth, J. B., The Effect of Uniform Layers on the Propagation of Radio Waves, Trans. I.R.E., **PGAP-2,** March 1952, p. 28.
16. Mack, L. C., Diversity Reception in UHF Long-Range Communications, Proc. I.R.E., **43,** October 1955, p. 1281.
17. Norton, K. A., Vogler, L. E., Mansfield, W. V. and Short, P. T., The Probability Distribution of the Amplitude of a Constant Vector Plus a Rayleigh-Distributed Vector, Proc. I.R.E., **43,** October 1955, p. 1354.
18. Rice, S. O., Mathematical Analysis of Random Noise, B.S.T.J., **24,** January 1945, p. 78.
19. Staras, H., Diversity Reception with Correlated Signals, J. Appl. Phys., **77,** January 1956, p. 93.
20. Rice, S. O., Statistical Properties of a Sine Wave Plus Random Noise, B.S.T.J., **27,** January 1948, p. 109.
21. Law, H. B., Lee, F. T., Looser, R. C. and Levett, F. A. W., An Improved Fading Machine, Proc. I.E.E., **104B,** March 1957, p. 117.
22. Rice, S. O., Distribution of the Duration of Fades in Radio Transmission—Gaussian Noise Model, B.S.T.J., **37,** May 1958, p. 581.
23. Rice, S. O., private communication.
24. Chisholm, J. H., Portmann, P. A., DeBettencourt, J. T. and Roche, J. F., Investigations of Angular Scattering and Multipath Properties of Tropospheric Propagation of Short Radio Waves Beyond the Horizon, Proc. I.R.E., **43,** October 1955, p. 1317.
25. Crawford, A. B. and Jakes, W. C., Jr., Selective Fading of Microwaves, B.S.T.J., **31,** January 1952, p. 68.
26. Lawson, J. L. and Uhlenbeck, G. E., *Threshold Signals*, McGraw-Hill Book Co., New York, 1939, pp. 62, 156.
27. Booker, H. G., Ratcliffe, J. A. and Shinn, D. H., The Diffraction from an Irregular Screen, Phil. Trans. Roy. Soc. (London), **242,** 1950, p. 579.
28. Willson, F. E., private communication.

Group Testing To Eliminate Efficiently All Defectives in a Binomial Sample

By MILTON SOBEL and PHYLLIS A. GROLL

(Manuscript received November 13, 1958)

In group-testing, a set of x units is taken from a total starting set of N units, and the x units ($1 \leqq x \leqq N$) are tested simultaneously as a group with one of two possible outcomes: either all x units are good or at least one defective unit is present (we don't know how many or which ones). Under this type of testing, the problem is to find the best integer x for the first test and to find a rule for choosing the best subsequent test-groups (which may depend on results already observed), in order to minimize the expected total number of group-tests required to classify each of the N units as good or defective. It is assumed that the N units can be treated like independent binomial chance variables with a common, known probability p of any one being defective; the case of unknown p and several generalizations of the problem are also considered.

I. SUMMARY

A finite number, N, of units are to be tested in groups. A "group-test" is a simultaneous test on x units (x to be chosen so that $1 \leqq x \leqq N$) with only two possible outcomes: "success," indicating that all x units are good; and "failure," indicating that at least one of the x units is defective (we don't know how many or which ones). The problem is to define a simple and efficient procedure (or an optimal procedure) for separating all the defective units from the good units — efficiency being defined in the sense of minimizing the expected number of group-tests required. Each unit is assumed to represent an independent observation from a binomial population with a common known *a priori* probability, q, of being good and $p = 1 - q$ of being defective. (The case of q unknown is briefly treated in Section X.)

A procedure (or decision rule), R_1, which describes a mode of action for any given value of q, is proposed and compared with several other procedures applicable to the same problem. The procedure R_1 is simple

in the sense that at any time, t, the experimenter must separate the units not yet proven to be good or defective into only two sets; units within either of these two sets need not be distinguishable. If it is given that the identification of units within the group being tested is economically impractical or impossible, then the procedure R_1 is conjectured to be optimal for all values of q.

Explicit instructions for carrying out R_1 are given for $N = 1(1)16$ for all q and for $N = 17(1)100$ for the particular values $q = 0.90, 0.95$ and 0.99. Exact formulae for the expected number of group-tests required under R_1 are given in Table IV B for all q and for $N = 1(1)12$; numerical results for $q = 0.90, 0.95$ and 0.99 are given in Tables V A, V B and V C for $N = 1(1)100$. Other numerical comparisons are made in Tables II A and II B. [Tables II through VIII appear at the end of this paper.]

Another procedure, R_2, which is simpler to compute and compares favorably with R_1, is defined in Appendix A in terms of information theory concepts.

Several different directions for generalization of the problem and corresponding generalizations of the procedure R_1 are considered in Section XI. Industrial applications are mentioned, in addition to the known application to blood testing.

II. INTRODUCTION

A problem which has hitherto been considered only in connection with blood-testing applications[1,2,3] can be shown to have industrial applications, and these have focused interest on a more general treatment of the problem. During World War II, a great saving was accomplished in the field of blood testing by pooling a fixed number of blood samples and testing the pooled sample for some particular disease. If the disease was not present, then several people were passed by a single test; if the disease was present, then there was enough blood remaining in each blood sample to test each one separately. The amount of time, money and effort saved by such a procedure depends on how rare the disease is in the population of people being tested. In this application, the total number of people to be tested was regarded as unknown and very large.

The goal of the problem treated here is the same — namely, to separate the defective units from the good units with a minimal (or approximately minimal) number of group-tests. This problem differs from the blood-testing problem in the following respects:

i. The population size N (number of people to be tested) is known at the outset.

ii. The number of units in each group-test (pooled blood sample) is not necessarily constant.

iii. If a group-test fails (the disease is present) we do not necessarily test each item separately.

In practice, the simplicity of the procedure deserves some consideration. The proposed procedure R_1 defined in Section III, after having been computed and described explicitly in advance of any experimentation, is in some sense no more complicated than the blood-testing procedure described above; this is explained in Section V.

Some typical industrial applications are:

1. It is desired to remove all "leakers" from a set of N devices. One chemical apparatus is available and the devices are tested by putting x of them (where $1 \leqq x \leqq N$) in a bell jar and testing whether any of the gas used in constructing the devices has leaked out into the bell jar. It is assumed that the presence of gas in the bell jar indicates only that there is at least one leaker and that the amount of gas gives no indication of the number of leakers. The *a priori* probability, q, of a unit being good is given by the records of similar units tested in the past.

2. Paper capacitors are tested at most n at a time, and each test indicates by the presence or absence of a current whether or not there is at least one defective present. For given n and q and given cost of unit manufacture, should the operator throw away a whole set of n units if it contains at least one defective? If not, how should he proceed to sort out the defective units to minimize the expected number of tests required? If the cost of a group-test and the cost of producing a unit are known, a related problem is to find a procedure which minimizes the total cost (including testing costs) of producing a good unit.

3. *Christmas tree lighting problem.* A batch of n light bulbs is electrically arranged in series and tested by applying a voltage across the whole batch or any subset thereof. If this is to be done on a routine basis, what procedure should be used to minimize the expected number of tests required to remove all the defective light bulbs, assuming the value of q is given?

4. A test indicates whether or not there is at least one *good* unit present in a batch of n, without indicating which ones or how many are good. Given q, what procedure should be used to remove the good units? This dual problem, which is useful in salvaging good components on a routine basis, is mathematically equivalent to those above, if the definitions of good and defective are interchanged.

A procedure R_1 is defined to solve the above problems, and is compared with several other procedures for the same problem. A procedure R_2,

based on maximizing the information in each group-test, is defined in Appendix A. Another procedure, R_3, which does not allow any recombination, is defined in Appendix B. Two "halving procedures," which can be carried out without knowing the true value of q, are defined in Appendix C. Procedures R_7 and R_6 are the best procedures that can be obtained by the methods of Dorfman[1] and Sterrett,[3] respectively, for small population sizes. For $N = 4$, 8 and 12 and various q values, Table II A gives a numerical comparison of the expected number of group-tests required for all these procedures.

Different directions of generalization, some of which are discussed in Section XI, are the following:

1. Two (or more) different kinds of units with (say) known probabilities q_1, q_2 of a unit being good are present, and the two different kinds can be put into the same test group.

2. Two (or more) experimenters may be working on a single set of N units by carrying out simultaneous, parallel group-tests and cooperating in such a way as to minimize the time required to accomplish the task.

3. The restriction is sometimes applied (particularly in blood-testing) that any one unit can be included in at most k group-tests; here the goal is to minimize the expected number of group-tests subject to this restriction. For $k = 2$ the proposed procedure is necessarily based on the method of Dorfman;[1] i.e., if a group-test fails, then the units therein are all tested individually.

4. Various generalizations appear if it is assumed that each test on x units gives three (or more) different possible results. For example, a test could indicate that either (a) all are good or (b) all are defective or (c) there are at least one good unit and at least one defective present.

5. A unit can be defective in either of two ways (e.g., electrical or mechanical) with the two *a priori* probabilities of being defective assumed to be independent but not necessarily equal. If there are two different tests corresponding to the two types of defectives, then, in addition to deciding the next test-group size, it may be necessary to decide which test to use next.

6. For positive continuous chance variables with a known distribution (like weight) the following problem is analagous. It is desired to separate N units into two groups according as the weight per unit is less than or greater than a constant (say, unity). Any number of units can be included in a single weighing. The problem is to accomplish the separation in a minimal number of weighings, assuming that the individual weights are independent observations from the common known distribution.

Many of these generalizations will be omitted from this paper and treated separately.

III. THE PROCEDURE R_1

The procedure R_1 is defined implicitly by a pair of recursion formulae and boundary conditions, but first we shall need some definitions and preliminary results. The units proven to be good and the units proven to be defective are never used in subsequent tests. Aside from such units, this procedure requires that at every stage the remaining units be separated into at most two sets. For one set of size $m \geqq 0$, which we call the *defective set*, it is known that it contains at least one defective unit; for the other set of size $n - m \geqq 0$, which we call the *binomial set*, our *a posteriori* knowledge is, so to speak, in the original binomial state; i.e., given the past history of testing, the units in the binomial set act like independent binomial chance variables with a common probability p of being defective. For the defective set, the conditional probability that Y, the number of defectives present, equals y is

$$\Pr\{Y = y \mid Y \geqq 1\} = \frac{\binom{m}{y} p^y q^{m-y}}{1 - q^m} \qquad (y = 1, 2, \cdots, m). \tag{1}$$

If X denotes the number of defectives present in a subset of size x randomly chosen from the defective set, then

$$\Pr\{X = 0 \mid Y \geqq 1\} = \sum_{y=1}^{m-x} \frac{\binom{m}{y} p^y q^{m-y}}{1 - q^m} \frac{\binom{m-y}{x}}{\binom{m}{x}} = \frac{q^x(1 - q^{m-x})}{1 - q^m}. \tag{2}$$

Before defining the procedure it is convenient to prove a lemma in a more general setting. Let $T(r_i)$ $(i = 1, 2, \cdots, t)$ denote a test on n units $(1 \leqq r_i \leqq n)$ such that there are only two mutually exclusive possible outcomes: a "failure," indicating that there are at reast r_i defectives present, and a "success," indicating that at most $r_i - 1$ of the units in the test are defective. In Lemma 1 we consider any integers r_i, r_0 with $1 \leqq r_i \leqq n (i = 0, 1, 2, \cdots, t)$, but the most important application is the case $r_0 = r_1 = \cdots = r_t = 1$. Let $\mathfrak{A}$ be any set of units, and let $\mathfrak{B}_i (i = 1, 2, \cdots, t)$ denote sets not necessarily disjoint from one another but such that each is disjoint from $\mathfrak{A}$; the case $t = 1$ is the one used for procedure R_1. At the outset, all units are independently and binomially distributed with a common probability p of being defective.

Lemma 1: If a test $T(r_i)$ on $\mathfrak{A} + \mathfrak{B}_i$ produces a failure for $(i = 1, 2, \cdots, t)$ and another test $T(r_0)$ on $\mathfrak{A}$ also produces a failure, then for $r_0 \geqq \max r_i$ the conditional distribution associated with all the units in the sets $\mathfrak{B}_i (i = 1, 2, \cdots, t)$, given both conditions above, is exactly the same as the original binomial distribution.

Proof: Let A and B_i denote the chance number of defectives present in $\mathfrak{A}$ and $\mathfrak{B}_i$, respectively. For the jth set $\mathfrak{B}_j$ the conditional probability P of interest is

$$P = \Pr\{B_j \leqq b \mid A + B_i \geqq r_i (i = 1, 2, \cdots, t), A \geqq r_0\}. \quad (3)$$

Since $r_0 \geqq \max r_i (i = 1, 2, \cdots, t)$, the condition $A \geqq r_0$ implies that $A + B_i \geqq r_i$, and hence

$$P = \Pr\{B_j \leqq b \mid A \geqq r_0\}. \quad (4)$$

Since $\mathfrak{B}_j$ and $\mathfrak{A}$ are disjoint, it follows that B_j and A are independent and hence, from (4),

$$P = \Pr\{B_j \leqq b\}, \quad (5)$$

which proves the lemma.

Let $G_1(m, n; q) = G_1(m, n)$ denote the expected number of group-tests remaining to be performed if the defective set is presently of size m, the binomial set is presently of size $n - m$, the *a priori* probability of a good unit is the known constant q and the procedure R_1 is used. For the special case $m = 0$ we use the symbol $H_1(n; q) = H_1(n)$. The values of m and n vary as the procedure is carried out; at the outset, $m = 0$ and $n = N$. It will also be convenient to refer to the *G-situation* or *G(m, n)-situation* if $m > 1$ and to the *H-situation* or *H(n)-situation* if $m = 0$.

Recursion Formulae Defining Procedure R_1

If x denotes the size of the very next group-test, then we write for any situation with $m = 0$

$$H_1(n) = 1 + \min_{1 \leqq x \leqq n} \{q^x H_1(n - x) + (1 - q^x) G_1(x, n)\}, \quad (6)$$

and, with the help of (2) and Lemma 1, we write for $n \geqq m \geqq 2$

$$G_1(m, n) = 1 + \min_{1 \leqq x \leqq m-1} \left\{ \left(\frac{q^x - q^m}{1 - q^m} \right) G_1(m - x, n - x) + \left(\frac{1 - q^x}{1 - q^m} \right) G_1(x, n) \right\}. \quad (7)$$

The *boundary conditions* state that for all q

$$H_1(0) = 0 \quad \text{and} \quad G_1(1, n) = H_1(n-1) \quad \text{for} \quad n = 1, 2, \cdots . \tag{8}$$

In (6) and (7) the constant 1 represents the very next group-test of size x and the expression in braces is the conditional expected number of additional group-tests given x. It follows from (6) and (8) that $H_1(1) = 1$ for all q.

Remark 1: To justify writing $G_1(x, n)$ in (7) we make use of Lemma 1 with $t = 2$, $r_2 = 0$, $r_1 = r_0 = 1$; we take $\mathfrak{A} + \mathfrak{B}_1$ as the defective set of size m and $\mathfrak{A}$ as a subset of size $x < m$. Then, by Lemma 1, if the subset $\mathfrak{A}$ of size x is shown to contain at least one defective, the *a posteriori* distribution associated with the $m - x$ units in B_1 is exactly binomial. These are then mixed or recombined with the $n - m$ "binomial units," giving a total of $n - x$ binomial units, and this justifies the expression $G_1(x, n)$ in (7).

Remark 2: These two recursion formulae, together with the boundary conditions, allow one to compute successively for any q the functions $G_1(2, 2)$, $H_1(2)$, $G_1(2, 3)$, $G_1(3, 3)$, $H_1(3)$, $G_1(2, 4)$, $G_1(3, 4)$, $G_1(4, 4)$, $H_1(4)$, $\cdots$ to any desired value of m and n.

Remark 3: The integer x which accomplishes the minimization in (6) and (7) for each situation characterized by the integers m and n is particularly important, since this is the size of the next test to be run according to the procedure R_1. These integers $x = x_H(n; q)$ and $x = x_G(m, n; q)$ implicitly define the procedure R_1. An illustration of how the procedure R_1 is to be carried out is given in Section IV.

Remark 4: If $m > 1$, then it is assumed in (7) that a subset of size x with $1 \leqq x < m$ will be taken from the defective set without mixing it with units from the binomial set. It follows from (6), (7) and (8) that any lack of optimality can only arise from this "no mixing" assumption. This assumption was used in the derivation of the algorithm (7) (see Remark 1 above). It will be noted in Section XIII that, when all the units are individually identified, then by dropping this assumption, an improvement to the procedure R_1 for high values of q can be found. A specific example of a modification and improvement of the procedure R_1, which drops the "no mixing" assumption at the expense of more complication will be more thoroughly discussed in a separate paper.

IV. ILLUSTRATION OF THE PROCEDURE R_1

Suppose we start with $N = 12$ units and it is given that $q = 0.98$. Referring to the column headed $H_1(12)$ in Fig. 3, we find that the first

test-group is of size $x = 12$; i.e., we start by testing all 12 units. If a success occurs, the experiment is over; if a failure occurs, then according to the column headed $G_1(12, n)$ of Fig. 4 the next test group is of size $x = 4$ chosen at random from the 12. Similarly, we continue along one of the sample paths shown in Fig. 1. The complete "tree" is not shown here, but testing continues in a similar manner and the specific details can be obtained from Figs. 3 and 4, which appear at the end of this paper.

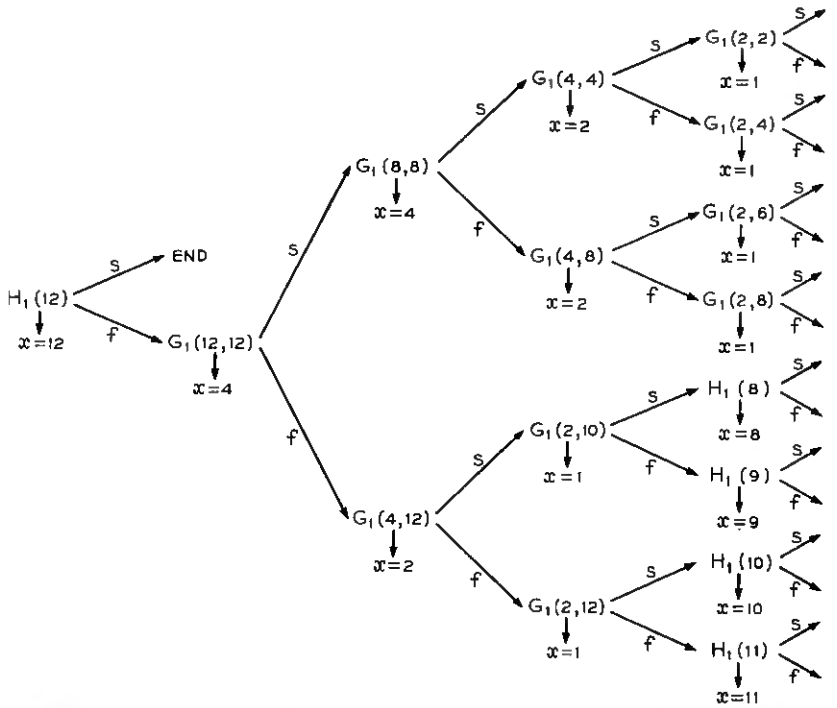


Fig. 1—Initial part of the tree for procedure R_1 for $q = 0.98$ and starting with an $H(12)$-situation.

It is obvious that the above procedure terminates in a finite number of steps. In fact, it can be shown for procedure R_1 (proof is omitted) that the maximum number $M(n)$ $[M(m,n)]$ for any H-situation $[G$-situation$]$ occurs when q is close to unity and the n unanalyzed units are all defective. It follows easily that

$$\begin{aligned} M(n) &= (n+1)\,[1+\alpha(n)] + 1 - 2^{1+\alpha(n)}, \\ M(m,n) &= \alpha(m) + n\,[1+\alpha(n-1)] + 1 - 2^{1+\alpha(n-1)}, \ (\mathrm{m} > 1) \end{aligned} \tag{10}$$

where $\alpha(z)$, for any positive integer z, is defined by

$$2^{\alpha(z)} \leqq z < 2^{1+\alpha(z)}. \tag{11}$$

For the example above, $\alpha(12) = 3$ and $M(12) = 37$; it is interesting to note that the constant term in $H_1(n)$ expressed in powers of q in Table IV B is also the length of the longest "chain" (that is, 37 for $n = 12$) in the "tree" used for the interval of q-values ending at unity. Although the maximum number is so large, the expected number of tests is only 2.07. This is explained partly by the fact that the procedure terminates after one group-test with probability 0.7847. A table of such probabilities for the number of tests T required when $q = 0.98$ and $N = 12$ is given below:

T	1	2	3	4	5	6	7	8	9	10–37
Probability	0.7847	0	0	0	0.0801	0.1124	0.0003	0.0016	0.0134	0.0075

If we assign the probability 0.0075 to $T = 10$, we obtain an estimate of $H_1(12)$ (namely, 2.070) which is a lower bound. The exact value, 2.073, can be obtained from the formula for $H_1(12)$ in Table IV B. Similarly, an estimate of the standard deviation is computed to be $\sigma \cong 2.1$, and this also is easily shown to be a lower bound.

It is interesting to note that, if the starting number N is exactly a power of 2 and q is large, the procedure R_1 starts off the same as a "halving" procedure. Such a procedure R_4 is defined in Appendix C for any N, and it has the property that it can be carried out without knowing the true value of q. To compare the means and standard deviations of R_1 and R_4, we consider the case $N = 6$, where $M = 14$ for both R_1 and R_4. For any $q > 0.844$, the expectation under R_1 can be put in the form

$$\begin{aligned} E(T;R_1) = {} & 1(q^6) + 4(3pq^5) + 5(3pq^5 + 2p^2q^4) + 6(2p^2q^4) \\ & + 7(7p^2q^4 + 2p^3q^3) + 8(4p^2q^4 + 5p^3q^3) + 9(8p^3q^3 + 3p^4q^2) \\ & + 10(5p^3q^3 + 5p^4q^2) + 11(3p^4q^2 + 2p^5q) \\ & + 12(3p^4q^2 + p^5q) + 13(p^4q^2 + 2p^5q) + 14(p^5q + p^6), \end{aligned} \tag{12}$$

where, for each term, the expression in parentheses is the probability that T takes on the value of the associated integer coefficient. For any q the corresponding expression under R_4 is

$$\begin{aligned} E(T;R_4) = {} & 1(q^6) + 4(3pq^5) + 5(3pq^5 + 2p^2q^4) + 6(p^2q^4 + p^3q^3) \\ & + 7(7p^2q^4) + 8(5p^2q^4 + 5p^3q^3) + 9(6p^3q^3 + 2p^4q^2) \\ & + 10(7p^3q^3 + 4p^4q^2) + 11(p^3q^3 + 4p^4q^2 + p^5q) \\ & + 12(4p^4q^2 + 2p^5q) + 13(p^4q^2 + 2p^5q) + 14(p^5q + p^6). \end{aligned} \tag{13}$$

The procedure R_1 is better than R_4 (at least for $q > 0.844$) since

$$E(T;R_4) - E(T;R_1) = 2p^2q^4 + 5p^3q^3 + 4p^4q^2 + p^5q \geqq 0. \tag{14}$$

[The fact that each term in (14) is positive indicates that R_1 is better than R_4, even if we know how many defectives are actually present among the N units.] Some numerical comparisons of means and standard deviations are given in Table I. The maximum difference for $N = 6$ between $E(T; R_4)$ and $E(T; R_1)$ occurs at $q = 0.844$, and is equal to 0.0379. For $N = 6$, the procedure R_1 appears to have a variance smaller than

TABLE I — COMPARISON OF PROCEDURES R_1 AND R_4 FOR $N = 6$

	$q = 0.85$	$q = 0.90$	$q = 0.95$	$q = 0.98$	$q = 0.99$
$E(T; R_1)$	3.8118	2.9434	2.0092	1.4133	1.2083
$E(T; R_4)$	3.8472	2.9605	2.0136	1.4141	1.2085
$\sigma(T; R_1)$	2.536	2.269	1.762	1.179	0.850
$\sigma(T; R_4)$	2.593	2.304	1.776	1.183	0.851

that of R_4 for all $q < 1$. A more complete comparison of $E(T; R_1)$ with that of several other procedures is given in Table II A.

V. THE SIMPLICITY OF R_1

It will be shown in this section that, for any given q and any situation $G(m, n)$, the appropriate x [i.e., the integer which accomplishes the minimization in (7)] does not depend on n. A somewhat simpler method of computing x is given and a new function of m alone is introduced to replace $G_1(m, n)$ in the definition of the procedure R_1. For any m and any pair of integers $(x, x + 1)$ both possible under R_1, there is always a single dividing point $q_d(x) = q_d(x, x + 1; m)$ that separates the interval for x from the interval for $x + 1$. (This property was observed for $m \leqq n \leqq 16$ and is treated as a conjecture for all m and n in Section VII.)

According to Remark 4 in Section III, the procedure R_1 for $m > 1$ is to "break down" the defective set. This "breaking down" is continued until a single unit is established to be defective and removed. Instead of randomizing the order of the units in the defective group again and again before each test group is selected, it will be convenient to assume,

without affecting the properties of the procedure R_1, that the order is randomized only once at the outset.† Units or groups of units removed later are then to be taken in that order. If the ith unit (in that order) is the first defective unit, then the "breaking down" mentioned above leads to an H-situation with $n - i$ binomial units, and the converse also holds true.‡

It is convenient to introduce $F_1(m, q) = F_1(m)$ defined as the expected number of group-tests required to "break down" a defective set of size m and for the first time reach an H-situation when q is given and the procedure R_1 is used. Then $F_1(m)$ clearly does not depend on n and the above argument permits us to write

$$G_1(m, n) = F_1(m) + \left(\frac{p}{1 - q^m}\right) \sum_{i=1}^{m} q^{i-1} H_1(n - i). \tag{15}$$

For algebraic simplicity we let

$$G_1^*(m, n) = \left(\frac{1 - q^m}{1 - q}\right) G_1(m, n) \text{ and } F_1^*(m) = \left(\frac{1 - q^m}{1 - q}\right) F_1(m). \tag{16}$$

Then (7) and (15) take on the simpler forms

$$G_1^*(m, n) = \sum_{i=1}^{m} q^{i-1} + \min_{1 \leq x \leq m-1} \{q^x G_1^*(m - x, n - x) + G_1^*(x, n)\}, \tag{17}$$

$$G_1^*(m, n) = F_1^*(m) + \sum_{i=1}^{m} q^{i-1} H_1(n - i). \tag{18}$$

Substituting (18) in (17), the three summations cancel and the result is

$$F_1^*(m) = \sum_{i=1}^{m} q^{i-1} + \min_{1 \leq x \leq m-1} \{q^x F_1^*(m - x) + F_1^*(x)\}, \tag{19}$$

which does not depend on n. The boundary condition, $F_1^*(1) = 0$ for all q, also does not depend on n. It is clear from this derivation that (19), which does not depend on n, must define the same integer values $x = x_0(m; q)$ as (17) or (7). This proves the following theorem.

† It should be pointed out that even this single randomization at the outset can be disregarded in carrying out the procedure R_1 if there is no doubt about the assumption of independent chance variables or if the units are already well-mixed in the process of delivery to the experimenter.

‡ It follows from the above that, for any procedure which "breaks down" the defective set in the above-mentioned manner (including a method of testing units from the defective set one at a time until a defective unit is found), the expected number of good units eliminated between a $G(m,n)$-situation and the next H-situation is $q/p - mq^m/(1 - q^m)$, and the number of defective units eliminated is always exactly one.

Theorem 1: For any G-situation with $n \geqq m > 1$ and any q, the size of the next test group, defined implicitly by (7), does not depend on n.

This result simplifies the explicit instructions needed to describe the procedure. Thus the two diagrams, Figs. 3 and 4, describe the procedure R_1 for all values of q and for any $N \leqq 16$.

Equations (15) and (16) can also be substituted in (6), yielding

$$H_1(n) = 1 + \min_{1 \leqq x \leqq n} \left\{ q^x H_1(n-x) + (1-q)\left[F_1^*(x) + \sum_{i=1}^{x} q^{i-1} H_1(n-i) \right] \right\}, \tag{20}$$

which, together with (19), gives a pair of "one-dimensional" recursion formulae for defining R_1 instead of the "two-dimensional" set, (6) and (7).

Remark 5: It should be pointed that, if one were to ask for a procedure that "breaks down" the defective set in as small an expected number of group tests as possible, then one would write (19) as one of the basic recursion formulae defining the procedure. This shows that R_1 "breaks down" the defective set and returns to an H-situation in a minimal number of tests.

VI. SOME PROPERTIES OF R_1 FOR q CLOSE TO UNITY

For any G-situation with $m > 1$, consider the effect of increasing q. It is easy to see by an induction argument that $G_1(m, n)$ is a strictly decreasing function of q. The function $H_1(n)$ is also strictly decreasing unless the value of q is such that the procedure R_1 tests all units one at a time, in which case $H_1(n)$ is constant. The "tree" remains the same in an interval as q increases, and changes only when it becomes more efficient under R_1 to increase the size of some test group in the "tree"; i.e., if we proceed down the "tree" along any path, the first change encountered, if any, will be an increase in some test group size. It therefore seems reasonable to expect (in both G- and H-situations) that the largest value of x assigned by R_1 (say, $x_{\max}$) occurs in an interval of q-values ending at unity. This unproved assertion that under R_1 large values of x are associated with large values of q is an immediate consequence of Conjectures 1 and 2 stated in Section 7.

For fixed $m > 1$, let the integers $\alpha(m)$ and $\beta(m)$

$$[\alpha(m) \geqq 1, \qquad 0 \leqq \beta(m) < 2^{\alpha(m)}]$$

be defined by

$$m = 2^{\alpha(m)} + \beta(m). \tag{20a}$$

Under the assumption that R_1 assigns $x_{\max}$ in an interval of q values ending at unity, it will be shown for any G-situation that in this interval the value of $x = x_{\max}$ is given by

$$x_{\max} = \begin{cases} 2^{\alpha(m)-1} & \text{for} \quad 2^{\alpha(m)} \leqq m \leqq 3\cdot 2^{\alpha(m)-1} \\ m - 2^{\alpha(m)} & \text{for} \quad 3\cdot 2^{\alpha(m-1)} \leqq m < 2^{\alpha(m)+1}. \end{cases} \tag{21}$$

As a corollary it then follows that, under R_1 (G-situation),

$$\frac{m}{3} \leqq x_{\max} \leqq \frac{m}{2}. \tag{22}$$

Also in the above-mentioned interval we have

$$F_1^*(m) = \alpha(m)\left(\frac{1-q^m}{1-q}\right) + q^{m-2\beta(m)}\left[\frac{1-q^{2\beta(m)}}{1-q}\right]. \tag{23}$$

Proof of (21), (22) and (23): In (19), since $x \leqq m-1$ and $m - x \leqq m - 1$, we can use the induction hypothesis (23) to obtain

$$\begin{aligned} F_1^*(m) - \sum_{i=1}^{m} q^{i-1} = \min_{1\leqq x\leqq m-1} \Bigg\{ & q^x \alpha(m-x)\left(\frac{1-q^{m-x}}{1-q}\right) \\ & + q^{m-2\beta(m-x)}\left(\frac{1-q^{2\beta(m-x)}}{1-q}\right) + \alpha(x)\left(\frac{1-q^x}{1-q}\right) \\ & + q^{x-2\beta(x)}\left(\frac{1-q^{2\beta(x)}}{1-q}\right)\Bigg\}. \end{aligned} \tag{24}$$

For q close to unity, the right side of (24) is equivalent to minimizing

$$Q(x) = [x\alpha(x) + 2\beta(x)] + [(m-x)\alpha(m-x) + 2\beta(m-x)]. \tag{25}$$

For the moment, let us utilize the symmetry of $Q(x)$ for $2 \leqq x \leqq m-2$ and limit our considerations to $x \leqq m/2$ and $m \geqq 4$.

Let $[m/2]$ denote the largest integer less than or equal to $m/2$. For $x = [m/2], [m/2] - 1, \cdots$, we consider $Q(x+1) - Q(x)$ and distinguish several cases according as

$$\begin{aligned} &\text{a.} \quad \alpha(x-1) = \alpha(x) = \alpha(m-x) \leqq \alpha(m-x+1); \\ &\text{b.} \quad \alpha(x-1) + 1 = \alpha(x) = \alpha(m-x) \leqq \alpha(m-x+1); \\ &\text{c.} \quad \alpha(x-1) = \alpha(x) = \alpha(m-x-1) \\ &\qquad\qquad = \alpha(m-x) - 1 = \alpha(m-x+1) - 1; \\ &\text{d.} \quad \alpha(x-1) + 1 = \alpha(x) = \alpha(m-x-1) \\ &\qquad\qquad = \alpha(m-x) - 1 = \alpha(m-x+1) - 1; \\ &\text{e.} \quad m \geqq 2x+1 \text{ and } \alpha(x) < \alpha(m-x-1). \end{aligned} \tag{26}$$

Using the fact that, for any integer $y > 1$,

$$y[\alpha(y) - \alpha(y-1)] + 2[\beta(y) - \beta(y-1)] = 2, \tag{27}$$

we obtain for all the cases in (26) the result,

$$Q(x+1) - Q(x) = \alpha(x) - \alpha(m-x) \leqq 0. \tag{28}$$

In particular, the value is zero only when (26a) holds and $\alpha(x) = \alpha(m-x)$. Hence, $Q(x)$ is a nonincreasing function of the integer x for $1 \leqq x \leqq m/2$ and is constant for $x_0 \leqq x \leqq m - x_0$, where x_0 is defined by

$$\alpha(x_0 - 1) + 1 = \alpha(x_0) = \alpha(m - x_0 - 1)$$

or (29)

$$\alpha(x_0) = \alpha(m - x_0 - 1) = \alpha(m - x_0) - 1,$$

whichever gives the maximum. These imply that

$$x_0 = 2^{\alpha(x_0)} \leqq m - x_0 - 1 < 2^{\alpha(x_0)+1} = 2x_0$$

or (30)

$$x_0 < m - x_0 = 2^{\alpha(x_0)+1} \leqq 2x_0 .$$

Both lead to the same results; namely,

$$\frac{m}{3} \leqq x_0 < \frac{m}{2}. \tag{31}$$

It follows that

$$\frac{m}{3} \leqq x_{\max} \leqq \frac{2m}{3}$$

and, from (29), for any integer x with $x_0 \leqq x \leqq m - x_0$, the values of $\alpha(x)$ and $\alpha(m-x)$ can differ by at most unity.

For any x with $x_0 \leqq x \leqq m - x_0$ we consider two cases according as $|\alpha(x) - \alpha(m-x)|$ is zero or unity. Let the expression in braces in (24) be denoted $C_1^*(x)$. For integers x with $x_0 \leqq x \leqq m - x_0$ the number of terms in $C_1^*(x)$ is constant and the expression $C_1^*(x)$ is to be minimized by making more powers of q larger.

Case 1. We can write

$$C_1^*(x) = \alpha(x)\left[\frac{1-q^m}{1-q}\right] + q^{m-2\beta(m-x)}\left[\frac{1-q^{2\beta(m-x)}}{1-q}\right] + q^{x-2\beta(x)}\left[\frac{1-q^{2\beta(x)}}{1-q}\right], \tag{32}$$

and the problem is to find the x which minimizes $C_1^*(x)$ for large values of q (i.e., to find $x_{\max}$). Since $\alpha(x) = \alpha(m - x)$, $\beta(m - x) - \beta(x) = m - 2x$ and, using the fact that $x_{\max} \geqq m/3$, we consider only integers $x \geqq m/3$ and obtain

$$m - x \geqq 2(m - 2x) = 2[\beta(m - x) - \beta(x)], \tag{33}$$

and hence

$$m - 2\beta(m - x) \geqq x - 2\beta(x). \tag{34}$$

It follows that the last term in (32) has lower powers of q than the previous term and that (32) is minimized by setting $\beta(x) = 0$, i.e., by taking $x_{\max}$ to be a power of 2. To complete the proof of (21) and (23), we note that

$$\begin{aligned} m = x + (m - x) &= 2^{\alpha(x)} + 2^{\alpha(m-x)} + \beta(m - x) \\ &= 2^{\alpha(x)+1} + \beta(m - x), \end{aligned} \tag{35}$$

so that $\alpha(m) = \alpha(x) + 1$, $\beta(m) = \beta(m - x)$ and, since $x = x_{\max}$ is a power of 2, we have

$$x_{\max} = 2^{\alpha(m)-1}. \tag{36}$$

Since, by (35), $\beta(m) = \beta(m - x) < \frac{1}{2}2^{\alpha(m)}$, then (36) holds for $2^{\alpha(m)} \leqq m < 3[2^{\alpha(m)-1}]$. Substituting these values of $\alpha(x)$, $\beta(x)$, $\alpha(m - x)$ and $\beta(m - x)$ into (32) and using (24) gives

$$F_1^*(m) = \alpha(m)\left(\frac{1 - q^m}{1 - q}\right) + q^{m-2\beta(m)}\left(\frac{1 - q^{2\beta(m)}}{1 - q}\right), \tag{37}$$

which completes the induction for Case 1. Conversely, if $2^{\alpha(m)} \leqq m < 3[2^{\alpha(m)-1}]$ then $x_0 = 2^{\alpha(m)-1}$ and, for any x with $x_0 \leqq x \leqq m - x_0$, we have $\alpha(x) = \alpha(m - x)$. Hence (32) is minimized for $x_{\max} = x_0 = 2^{\alpha(m)-1}$ by the above argument.

Case 2. Since we are now considering only possibilities outside Case 1, we must have $3[2^{\alpha(m)-1}] \leqq m \leqq 2^{\alpha(m)+1}$. Using (29), and the fact that $x_0 \leqq x \leqq m - x_0$, it follows that either $\alpha(x) = \alpha(m - x - 1) = \alpha(m - x) - 1$, so that $m - x$ is a power of 2 and $x < m/2$, or else $\alpha(m - x) = \alpha(x - 1) = \alpha(x) - 1$, so that x is a power of 2 and $x > m/2$. In the former alternative, we obtain

$$m = x + (m - x) = 3[2^{\alpha(x)}] + \beta(x),$$

$\alpha(m) = \alpha(x) + 1$ and $\beta(m) = 2^{\alpha(x)} + \beta(x) = x$. It follows that

$$x = x_{\max} = m - 2^{\alpha(m)}. \tag{38}$$

In the latter alternative, it can be concluded that $x = 2^{\alpha(m)} > m/2$. In the latter alternative (or in both alternatives), we now compare the values of $C_1^*(x)$, computed from (24), for the two arguments,

$$x_1 = 2^{\alpha(m)} > m/2 \qquad \text{and} \qquad x_2 = m - 2^{\alpha(m)} < m/2.$$

Using the general result that $\beta(m^*) < m^*/2$ for any positive integer m^* and the implied result that $\alpha(m - 2^{\alpha(m)}) = \alpha(m) - 1$, it is easy to show (the details are omitted) that $C_1^*(x_2) < C_1^*(x_1)$ for $0 < q < 1$. Hence, for either alternative, we obtain the same result, (38). As a corollary, (22) holds. In this case $C_1^*(x)$, after algebraic simplification, is given by

$$C_1^*(x) = [\alpha(m) - 1]\left(\frac{1 - q^m}{1 - q}\right) + q^{m-2\beta(m)}\left[\frac{1 - q^{2\beta(m)}}{1 - q}\right], \quad (39)$$

which, using (24), again gives the result (23). The fact that (23) holds for $m = 2$, 3 and 4 is easily shown and the details are omitted. This completes the proof of (23).

It is also possible to get expressions for $G_1^*(m, m)$, $G_1^*(m, n)$ and $H_1(n)$ for large values of q using (23) and the fact that, for any positive n and sufficiently large values of q (depending on n),

$$H_1(n) = 1 + (1 - q^n)G_1(n, n) = 1 + pG_1^*(n, n). \quad (40)$$

Then, by (18), we obtain for large values of q

$$G_1^*(m, m) = m - 1 + qF_1^*(m) + p \sum_{j=2}^{m} F_1^*(j), \quad (41)$$

$$\begin{aligned} G_1^*(m, n) &= n - 1 - (n - m - 1)q^m + F_1^*(m) \\ &\quad + p(1 - q^m) \sum_{j=2}^{n-m-1} F_1^*(j) + p \sum_{j=n-m}^{n-1} F_1^*(j), \end{aligned} \quad (42)$$

$$H_1(n) = q + np + pF_1^*(n) + p^2 \sum_{j=2}^{n-1} F_1^*(j), \quad (43)$$

where (42) is to be used only for $m < n$ and the summation from a to b is taken to be zero for $b < a$. Using these and (23), the last equation (i.e., the equation for the last q-interval which ends at unity) can be obtained independently and more simply for $F_1^*(m)$, $G_1^*(m, n)$ and $H_1(n)$.

It is clear from the results above that, as q approaches unity,

$$\lim H_1(n) = 1 \quad (43a)$$

and that, for q in an interval ending at unity, we have $x_H(n; q) = n$. Here $x_H(n; q)$ is the size of the next test group when the procedure R_1

is used for an $H(n)$-situation with *a priori* probability q of a unit being good. In an H-situation the probability that there are no defectives present approaches unity as q approaches unity. In a G-situation with $m > 1$, the probability that there is exactly one defective in the defective set and none in the binomial set approaches unity as q approaches unity. Defining $x_G(m;\ q)$ in a similar manner, it was shown above that $x_G(m; q) = x_{\max}$ [as given in (21)] in an interval ending at unity, and it follows from (23), (41) and (42) [or (23) and (15)] that, as q approaches unity,

$$\lim F_1(m) = \alpha(m) + 2\frac{\beta(m)}{m} \qquad (m > 1), \quad (43b)$$

$$\lim G_1(m, n) = \begin{cases} 1 + \lim F_1(m) = 1 + \alpha(m) + 2\dfrac{\beta(m)}{m} & (\text{for } n > m) \\ 1 - \dfrac{1}{m} + \alpha(m) + 2\dfrac{\beta(m)}{m} & (\text{for } n = m). \end{cases} \quad (43c)$$

It is interesting to note that the result (43c) depends on m but not on n.

VII. CONJECTURED PROPERTIES OF R_1

In this section we shall state some properties which appear to hold for procedure R_1 based on numerical calculations for $N \leqq 16$ but have not been proved for all N.

1. For any G-situation with fixed $m > 1$, if $x_G(m; q)$ denotes the size of the next test group under R_1, then $x_G(m; q)$ is a nondecreasing step function of q with step size unity. That is, for any pair $q \leqq q^+ \leqq 1$,

$$x_G(m; q) \leqq x_G(m; q^+) \quad (44)$$

and, for sufficiently small $\epsilon > 0$,

$$x_G(m; q + \epsilon) \leqq x_G(m; q) + 1. \quad (45)$$

Also for fixed q the value of $x_G(m + 1; q)$ is either the same or one greater than $x_G(m; q)$; i. e.

$$x_G(m; q) \leqq x_G(m + 1; q) \leqq x_G(m; q) + 1. \quad (46)$$

If the dividing point between x and $x + 1$ for any G-situation under R_1 [denoted by $q_G(x; m)$] is shown to exist (and be unique), then the three properties (44), (45) and (46) are equivalent to the two properties that, for any $m > 1$,

$$q_G(x; m) \geqq q_G(x; m + 1) \quad \text{for} \quad 1 \leqq x < x_{\max}, \tag{47}$$

$$q_G(x; m) < q_G(x + 1; m) \quad \text{for} \quad 1 \leqq x < x_{\max} - 1. \tag{48}$$

The assumption used in Section VI that the largest x-values are associated with the largest q-values is a simple consequence of (44).

2. For any H-situtation, we can define $x_H(n; q)$ similar to $x_G(m; q)$ and the property corresponding to (44) still holds: that, for $q \leqq q^+ \leqq 1$ and all positive integers n,

$$x_H(n; q) \leqq x_H(n; q^+). \tag{49}$$

We can also define $q_H(x; n)$ similar to $q_G(x; m)$ with the understanding that, if x does not appear under $H_1(n)$ in Fig. 3, then the interval for x is assumed to have length zero but the endpoints still exist, and, in fact, $q_H(x - 1, n)$ will then be equal to $q_H(x; n)$. Then (49) is equivalent to the property that for all $x \geqq 1$ and all positive integers n

$$q_H(x; n) \leqq q_H(x + 1; n). \tag{50}$$

3. If the experimental situation is such that it is impossible or economically impractical to identify or keep separate the individual units in any test group, then, after each test on a batch of x units, the disposition of the x units must be made on a batch basis. In such a situation it is conjectured that the procedure R_1 is the optimal procedure for all values of q.

4. There are various patterns existing in Table II B both within a column and across columns, none of which have been proved. For example, if $q = q_G(x; m)$ is the dividing point between x and $x + 1$, then the first entry in the appropriate column is $1 - q^{2x} - q^{2x+1}$, and the last entry, for $m = \infty$, is $1 - q^x - q^{x+1}$. In the first column, the entry can be written as $1 - q - q^2 + q^m$ for all m. In the second column of Table II B the pattern shown is to replace the highest power of q (say, q^h) by q^{h-2} or by $q^{h+2} + q^{h+3}$, depending on whether m is odd or even. Thus the pattern displays a cycle of 2. In the third column one can similarly find a pattern with a cycle of 3, starting with $m \geqq 9$. If a general rule for all these patterns were proved then it might be easier to find the dividing points for higher values of m. The conjecture in this case lies in the fact that these patterns exist and can be mathematically established.

VIII. CHARACTER OF R_1 FOR SMALL VALUES OF q

For the procedure R_1 it will now be shown that, when

$$q < q_0 = \tfrac{1}{2}(\sqrt{5} - 1) = 0.618$$

(to three decimal places), then, for both G- and H-situations with any

positive integers $m \leqq n$, the units are all tested one at a time. Of course, if we start with the H-situation and test units one at a time, then a G-situation never arises, but in the induction proof that follows it must first be shown for the G-situation and then for the H-situation. This property that units are tested one at a time was recently shown[4] to hold for the optimal procedure (without specifying what the optimal procedure is or whether it exists). Simple formulae for $H_1(n)$, $G_1(m, n)$ and $F_1(m)$ are obtained for $q \leqq q_0$.

Theorem 2: For procedure R_1 with $1 \leqq m \leqq n$ and $0 \leqq q < q_0$,

$$x_G(m; q) = x_H(n; q) = 1, \tag{51}$$

$$H_1(n) = n, \tag{52}$$

$$G_1(m, n) = n - \frac{pq^{m-1}}{1 - q^m}, \tag{53}$$

$$F_1(m) = \frac{q}{p} + \frac{1 - q^{m-1} - mq^m}{1 - q^m}. \tag{54}$$

[Remarks: The last term in (53) results from the possibility of saving one test if the defective set of size m contains exactly one defective unit that is discovered inferentially by showing that all the other $m - 1$ units are good. It is interesting to note that (54) can be obtained by summing the series

$$F_1(m) = \frac{p}{1 - q^m}(1 + 2q + 3q^2 + \cdots + (m - 1)q^{m-2} + (m - 1)q^{m-1}). \tag{55}$$

In the proof below, (54) is shown first and then (52); the proof of these contain the result (51). Then (53) follows from (15), (52) and (54)].

Proof: The proof of (54) is by induction. The result holds for $m = 1$, since $F(1) = 0$. Assuming (54) holds for arguments less than m, we can use (19) with (16) to obtain

$$F_1^*(m) = \frac{1 - q^m}{p} + \min_{1 \leqq x \leqq m-1} \left\{ q^x \left[\frac{1 - q^{m-x-1} - (m - x - 1)pq^{m-x}}{p^2} \right] + \frac{1 - q^{x-1} - (x - 1)pq^x}{p^2} \right\} = \left[\frac{1 - q^{m-1} - (m - 1)pq^m}{p^2} \right] + \frac{1 - q^m}{p} + \frac{1}{p} \min_{1 \leqq x \leqq m-1} \{x(q^m - q^x) - pq^{x-1}\}. \tag{56}$$

To prove part of (51) it is now shown that the minimum of the expression above in braces [say, $f_0(x)$] is attained at $x = 1$. Since this is obvious for $m = 2$, it is now assumed that $m \geqq 3$. Then

$$f_0(x) - f_0(1) = (x - 1)q^m + 1 - xq^x - pq^{x-1}, \tag{57}$$

and it suffices to show that, for $q \leqq q_0$ and $x = 2, 3, \cdots, m - 1$,

$$f_1(x) = 1 - xq^x - pq^{x-1} \geqq 0. \tag{58}$$

Similarly, it suffices to show that, for $q \leqq q_0$ and $x = 3, 4, \cdots, m - 1$,

$$f_2(x) = \frac{f_1(x) - f_1(2)}{q} = 1 + q - xq^{x-1} - pq^{x-2} \geqq 0. \tag{59}$$

More generally, if it suffices to show that, for $q \leqq q_0$ and $x = y + 1$, $y + 2, \cdots, m - 1$,

$$f_y(x) = 1 + (y - 1)q - xq^{x-y+1} - pq^{x-y} \geqq 0, \tag{60}$$

then it also suffices to show that, for $q \leqq q_0$ and $x = y + 2, y + 3, \cdots, m - 1$,

$$f_{y+1}(x) = \frac{f_y(x) - f_y(y + 1)}{q} = 1 + yq - xq^{x-y} - pq^{x-y-1} \geqq 0. \tag{61}$$

Setting $y = m - 3$ in (61), it suffices to show that, for $q \leqq q_0$ and $x = m - 1$ (we can now replace x by $m - 1$),

$$1 + (m - 4)q - (m - 2)q^2 = 1 - q - q^2 + (m - 3)pq \geqq 0. \tag{62}$$

Since q_0 is the root of $1 - q - q^2$ and $m \geqq 3$, the inequality (62) is proved and the minimum of (56) is attained at $x = 1$. Setting $x = 1$ in (56) gives the bracketed expression in (56) and proves the result (54).

To prove (52), we substitute (54) in (20) to obtain

$$\begin{aligned} H_1(n) &= 1 + \min_{1 \leqq x \leqq n} \left\{ q^x(n - x) - pq^{x-1} + p \sum_{i=1}^{x} iq^{i-1} \right. \\ &\qquad \left. + p \sum_{i=1}^{x} (n - i)q^{i-1} \right\} \\ &= n + 1 - \max_{1 \leqq x \leqq n} \{xq^x + pq^{x-1}\}, \end{aligned} \tag{63}$$

where the value of q^{x-1} for $q = 0$ and $x = 1$ is taken to be unity.

To prove the rest of (51), it is now shown that the maximum of the expression above in braces [say, $h(x)$] is attained at $x = 1$ for $q \leqq q_0$.

Then, for $q \leqq q_0$ and $x \geqq 1$,

$$\begin{aligned} h(x) - h(x+1) &= xpq^x - 2q^x + q^{x-1} \\ &= q^{x-1}[1 - q - q^2 + (x-1)pq] \geqq 0. \end{aligned} \tag{64}$$

Clearly, the maximum of $h(x)$ is attained at $x = 1$ not only for $0 < q \leqq q_0$ but also for $q = 0$. Setting $x = 1$ in (63) gives $H_1(n) = n$, and this completes the proof of (52). The fact that the minimum is attained only at $x = 1$ for $q < q_0$ in both (56) and (63) proves (51) and shows that, under R_1 with $q < q_0$, units are tested one at a time.

IX. CONSTRUCTION OF TABLES FOR R_1

Figs. 3 and 4 describe the procedure R_1 for $n = 2(1)16$ and $m \leqq n$ in the form of two diagrams that are easy to use in a practical situation. Tables III A and III B give the polynomials, the roots of which are the dividing points in Figs. 3 and 4. Table IV A gives the polynomial equation for $F_1^*(m)$ for $m = 2(1)16$. Tables IV B and IV C give the polynomial equations for $H_1(n)$ and $G_1^*(m, n)$, respectively, for $n = 2(1)12$ and $2 \leqq m \leqq n$. These can be obtained from (6), (7) and (8), or from (18), (19) and (20) and the boundary conditions, $H_1(0) = F_1^*(1) = 0$. For the sake of brevity, Table IV C has been reduced so that it gives the results only for $q \geqq 0.85$, and only for pairs m,n which can arise starting from an $H(n)$-situation with $n \leqq 12$.

Having computed $H_1(n)$ and $F_1^*(m)$ for $2 \leqq m \leqq n \leqq 12$, we can make the procedure R_1 explicit for $12 \leqq n \leqq 16$ by a different method, which will now be explained. Let $H_1(n \mid x)$ denote the value of $H_1(n)$ if (i.e., for those q-values for which) the next sample size is x; let $F_1^*(m \mid x)$ be defined similarly. Then (20) can be written as

$$\begin{aligned} H_1(n \mid x) - H_1(n-1) &= 1 + pF_1^*(x) \\ &\quad - \sum_{i=1}^{x-1} q^i[H_1(n-i) - H_1(n-i-1)]. \end{aligned} \tag{65}$$

Writing a similar equation for $H_1(n \mid y)$ for $y > x$ and subtracting gives

$$\begin{aligned} H_1(n \mid x) - H_1(n \mid y) &= -p[F_1^*(y) - F_1^*(x)] \\ &\quad + \sum_{j=x}^{y-1} q^j[H_1(n-j) - H_1(n-j-1)]. \end{aligned} \tag{66}$$

If $4 \leqq x < y \leqq 16$ and $12 < n \leqq 16$, then the right member of (66) involves H-function arguments only up to 12. In particular, for $y = x + 1$ we set the right member of (66) equal to zero and obtain a poly-

nomial whose root (between zero and one) is the dividing point, $q_H(x, n)$, between x and $x + 1$. Table II A shows that, for $n > 12$ and $x < 4$, the pattern for the dividing points is well stabilized; for example, the value of $q_H(1, n) = 0.618$ (to three decimal places) is shown to hold for all n in Section VIII. By considering various pairs x, y (most often of the type x, $x + 1$) in (66) it is possible to determine the procedure R_1 for the $H(n)$-situation for $12 < n \leqq 16$, without explicitly computing the formula for $H_1(n)$.

Similarly, we can do the same for the G-situation by using

$$F_1^*(m \mid x) - F_1^*(m \mid y) = F_1^*(y) - F_1^*(x) + q^y F_1^*(m - y) - q^x F_1^*(m - x), \tag{67}$$

but in this case (according to Section VI) we need only consider values $y = x + 1$ up to and including $m/2$.

Table II A gives a numerical comparison for $N = 4$, 8 and 12 of R_1 and several other procedures, two of which are based on the work of Dorfman[1] and Sterrett;[3] the others are defined in the Appendices to this paper. Table II B gives a brief numerical comparison of $H_1(n)$ and $H_3(n)$ (corresponding to procedure R_3 defined in Appendix B) for large values of n [viz., $n = 10(10)100$] and $q = 0.90$, 0.95 and 0.99; these entries were computed on the IBM-704.

Tables V A, V B and V C give the numerical values of $H_1(n)$ and $G_1(m, n)$ as well as the values of $x_G(m; q)$ and $x_H(n; q)$ for $q = 0.90$, 0.95 0.99 for $2 \leqq n \leqq 100$, and for appropriate values of m; these entries were computed on the IBM-704. For $q = 0.90$ the value of x is always at most 9 and hence, if we start with an H-situation, there is no need to consider values of $m > 9$; similarly, for $q = 0.95$ we disregard values of $m > 19$. For $q = 0.99$ we should consider all values of m up to and including $m = 100$ but many of these were omitted for the sake of brevity. It is interesting to note that $G_1(m, n)$ is strictly monotonic in the second argument (and hence also in the argument $n - m$) for fixed m, but it is curious and difficult to explain why it is not monotonic in the first argument for fixed n.

X. A SUGGESTED PROCEDURE FOR THE CASE OF UNKNOWN q

It is reasonable to expect that a knowledge of good procedures for the case of known q will suggest good procedures for the case of unknown q. From this point of view we consider modifications of the basic procedure R_1 that make it adaptable when q is unknown. It is suggested that after each test we form a new estimate of q and that the procedure R_1 be

used with the estimated value in place of the true value. At the outset we can start with an estimate based on past experience or we can start by testing one unit at a time. A thorough investigation of the relative merit of this procedure has not been carried out. Some discussion on the maximum likelihood method of estimating q is given below.

Let d and s denote the number of units proven defective and proven good, respectively, so that at any stage of experimentation we have

$$N = d + s + m + (n - m) = d + s + n, \tag{68}$$

where N is the total number of units at the outset, m is the size of the defective set (which is known to contain at least one defective) and $n - m$ is the size of the binomial set. The likelihood L of the observed result (68) is given by

$$L = \binom{N - n}{d} p^d q^{N-n-d}(1 - q^m). \tag{69}$$

Then it is easily shown that

$$\frac{d}{dq}(\log L) = -\frac{d}{dp}(\log L) = -\frac{1}{pq}\left[d - (N - n)p + \frac{mpq^m}{1 - q^m}\right]. \tag{70}$$

Setting the latter equal to zero, we find that, for $m \neq 0$, the maximum likelihood estimate $\hat{q}$ of q is the root of

$$(N - n + m)(1 - \hat{q}^m)(1 - \hat{q}) - d(1 - \hat{q}^m) - m(1 - \hat{q}) = 0 \tag{71}$$

or, equivalently,

$$s - d\sum_{i=1}^{m} \hat{q}^i - (m + s)\hat{q}^m = 0 \tag{72}$$

and, for $m = 0$, we have $\hat{q} = s/(d + s)$, the usual estimate. For $s = 0$ and $m + d \geqq 1$, we get $\hat{q} = 0$ and, for $s = 1$, it is easily seen, using the Descartes Rule of Signs, that (72) has exactly one root $\hat{q}$ (allowing multiplicities) in the unit interval and hence $\hat{q}$ is uniquely defined. The remaining case, $s = m = d = 0$, can only occur at the outset when there is no observations on which to base an estimate. It is interesting to note that the same result (71) or (72) can also be obtained by computing the conditional expected proportion of defectives among the N units, given the observed s, d, m and n, and setting it equal to $1 - q$. The equation thus obtained is the same as (71), and its root is $\hat{q}$.

The above method of getting an estimate is being suggested in connection with procedure R_1, but it can also be used in connection with

the procedures R_2 and R_4 (see Appendices) without any change. For procedures R_3 and R_5 we can have several defective sets and several binomial sets at any one time, and (71) then becomes

$$d + (1 - \hat{q}) \left[\sum_{i=1}^{I} \frac{m_i}{1 - \hat{q}^{m_i}} - \left(N - \sum_{j=1}^{J} n_j' \right) \right] = 0, \tag{73}$$

where $m_1, m_2, \cdots, m_I$ are the sizes of the defective sets, $n_1', n_2', \cdots, n_J'$ are the sizes of the binomial sets, and

$$N = d + s + \sum_{i=1}^{I} m_i + \sum_{j=1}^{J} n_j'. \tag{74}$$

If the number of tests carried out is large (and hence $N - n$ must be large), the maximum likelihood estimate is approximately normally distributed with expectation equal to the true value of q and variance given by

$$\sigma^2(\hat{q}) \cong \left[E \left(\frac{d \log L}{dq} \right)^2 \right]^{-1}, \tag{75}$$

where m and n in (69) are to be regarded as chance variables. Taking expectation first for fixed m and n and then with respect to m and n, gives

$$\sigma^2(\hat{q}) \cong \left[E \left(\frac{N - n}{pq} \right) + E \left(\frac{mq^{m-1}}{1 - q^m} \right)^2 \right]^{-1}. \tag{76}$$

Since $mpq^{m-1} < 1 - q^m$ for all m and all $q < 1$, and since the expectation of a square is nonnegative, for asymptotically large $N - n$

$$\frac{p^2 q}{p + q[N - E(n)]} \leqq \sigma^2(\hat{q}) \leqq \frac{pq}{N - E(n)}. \tag{77}$$

In particular, if we continue to test until a fixed proportion $\theta > 0$ of the N units are determined to be good or bad (i.e., until $n/N = 1 - \theta$, approximately) then, for asymptotically large N (so that $N\theta$ is also large), we obtain

$$\frac{p^2 q}{N\theta p + q} \leqq \sigma^2(\hat{q}) \leqq \frac{pq}{N\theta}. \tag{78}$$

For $\theta = 1$ and large N, the two bounds are essentially equal and the common value is the same as for ordinary binomial sampling. In general, at any stage of experimentation it appears to be conservative to estimate $\sigma^2(\hat{q})$ in the same way as for ordinary binomial sampling based on $N - n$ observations, using the value of n that is actually realized at that time.

In regard to the procedure, if the size of the very first test group is based on past experience, the question arises as to whether this past experience should also enter into the second, third and other early estimates of q. If it does not enter, then in the early tests we may find sudden jumps from testing very small numbers to testing very large numbers and *vice versa*, both of which are undesirable. This makes it useful to find a method to continue to use past experience until the estimate of q (without using past experience) is stabilized. In the absence of past experience, this same feature may make it desirable to test several units one at a time before starting to use any group-testing procedure.

XI. SOME GENERALIZATIONS OF R_1

Returning to the case of known probabilities q, we consider some generalizations of the same basic problem and, in each case, the appropriate generalization of the procedure R_1. The appropriate formulae will be given, but only a few simple computations will be carried out.

1. Two (or more) different kinds of units with known probabilities (say, $q_1 \leqq q_2$) of a good unit are present and both can be put into the same test group.

Let $H_{11}(n_1, n_2)$ denote the expected number of tests required under the proposed procedure R_{11} if there are n_i units of type i with $q = q_i (i = 1, 2)$ and the binomial chance variables associated with the units are mutually independent. Let $G_{11}(m_1, m_2; n_1, n_2)$ denote the expected number of tests required under R_{11} if there is a defective set containing m_1 units of type 1 and m_2 units of type 2 (known to contain at least one defective among the $m_1 + m_2$ units) and a binomial set containing $n_1 - m_1 \geqq 0$ of type 1 and $n_2 - m_2 \geqq 0$ of type 2. The recursion formulae corresponding to (6), (7) and (8) are

$$H_{11}(n_1, n_2) = 1 + \min \{q_1^x q_2^y H_{11}(n_1 - x, n_2 - y) + (1 - q_1^x q_2^y) G_{11}(x, y; n_1, n_2)\}, \tag{79}$$

where the minimum is over pairs (x, y) with $0 \leqq x \leqq n_1$, $0 \leqq y \leqq n_2$ and $x + y \geqq 1$, and

$$G_{11}(m_1, m_2; n_1, n_2) = 1 + \min \left\{ \left(\frac{q_1^x q_2^y - q_1^{m_1} q_2^{m_2}}{1 - q_1^{m_1} q_2^{m_2}} \right) \cdot G_{11}(m_1 - x, m_2 - y; n_1 - x, n_2 - y) + \left(\frac{1 - q_1^x q_2^y}{1 - q_1^{m_1} q_2^{m_2}} \right) G_{11}(x, y; n_1, n_2) \right\}, \tag{80}$$

where the minimum is over pairs (x, y) with $0 \leqq x \leqq m_1$, $0 \leqq y \leqq m_2$ and $1 \leqq x + y \leqq m_1 + m_2 - 1$. The boundary conditions state that, for all $q_1 \leqq q_2$,

$$G_{11}(1, 0; n_1, n_2) = H_{11}(n_1 - 1, n_2) \quad \text{for all} \quad n_1 \geqq 1, n_2 \geqq 0, \tag{81}$$

$$G_{11}(0, 1; n_1, n_2) = H_{11}(n_1, n_2 - 1) \quad \text{for all} \quad n_1 \geqq 0, n_2 \geqq 1, \tag{82}$$

$$G_{11}(m_1, 0; n_1, 0) = G_1(m_1, n_1; q_1); \qquad G_{11}(0, m_2; 0, n_2) = G_1(m_2, n_2; q_2), \tag{83}$$

$$H_{11}(n_1, 0) = H_1(n_1; q_1); \qquad H_{11}(0, n_2) = H_1(n_2; q_2), \tag{84}$$

where the right-hand member of each equality in (83) and (84) refers to the basic procedure R_1 defined by (6), (7) and (8).

It is clear that $H_{11}(1, 0) = H_{11}(0, 1) = 1$ and $H_{11}(0, 0) = 0$. It follows from (80) that, for $q_1 \leqq q_2$,

$$G_{11}(1, 1; 1, 1) = \frac{2 - q_2 - q_1q_2}{1 - q_1q_2}, \tag{85}$$

and the rule is to test first the unit of type 2. Using this result, we can compute

$$H_{11}(1, 1) = 1 + \min(1, 2 - q_2 - q_1q_2), \tag{86}$$

and the rule is to test either unit separately if $1 - q_2 - q_1q_2 > 0$ and to test both simultaneously if $1 - q_2 - q_1q_2 \leqq 0$. The latter inequality is a direct generalization of the inequality $1 - q - q^2 \leqq 0$, which played a prominent role in the basic procedure R_1. We state (without proof) that, if $q_1 \leqq q_2 < \frac{1}{2}(\sqrt{5} - 1) = 0.618$ (to three decimals), all testing is carried out one unit at a time.

2. Two (or more) experimenters may be working on a single set of N units by carrying out simultaneous, parallel group-tests and cooperating in such a way as to minimize the time required to accomplish the task.

It is clear that no saving can be effected in the expected total number of tests by having more than one experimenter. However, if the simultaneous tests are regarded as a stage, each of which lasts the same amount of time, then minimizing the expected number of stages is equivalent to minimizing the expected time required to accomplish the task. These remarks indicate that there may be some conflict in these two aims of reducing the expected time and the expected total number of tests. For this reason, it should be stated that our primary emphasis in this problem is to reduce the expected time.

Let m and m' denote the sizes of defective sets and let $n - (m + m') \geqq 0$ denote the size of the binomial set. Let $H_{12}(n)$ denote the expected number of stages required for $m = m' = 0$ by the proposed procedure R_{12}. Let $G_{12}(m, m', n)$ denote the expected number of stages required by R_{12} if we have two defective sets of size m, m' and one binomial set of size $n - m - m' \geqq 0$. Let $G_{12}(m, 0, n)$ and $G_{12}(0, m, n)$ be denoted by $G_{12}(m, n)$, so that $G_{12}(0, n) = H_{12}(n)$. The recursion formulae for R_{12} are

$$\begin{aligned} H_{12}(n) = 1 + \min_{\substack{x, y \geqq 1 \\ x+y \leqq n}} [&q^{x+y} H_{12}(n - x - y) + q^x(1 - q^y) G_{12}(y, n - x) \\ &+ q^y(1 - q^x) G_{12}(x, n - y) + (1 - q^x)(1 - q^y)\, G_{12}(x, y, n)], \end{aligned} \tag{87}$$

$$\begin{aligned} G_{12}(m, m', n) = 1 + \min_{\substack{1 \leqq x \leqq m-1 \\ 1 \leqq y \leqq m'-1}} \Bigg\{ &\left(\frac{q^x - q^m}{1 - q^m}\right)\left(\frac{q^y - q^{m'}}{1 - q^{m'}}\right) \\ &\cdot G_{12}(m - x, m' - y, n - x - y) \\ &+ \left(\frac{q^x - q^m}{1 - q^m}\right)\left(\frac{1 - q^y}{1 - q^{m'}}\right) G_{12}(m - x, y, n - x) \\ &+ \left(\frac{1 - q^x}{1 - q^m}\right)\left(\frac{q^y - q^{m'}}{1 - q^{m'}}\right) G_{12}(x, m' - y, n - y) \\ &+ \left(\frac{1 - q^x}{1 - q^m}\right)\left(\frac{1 - q^y}{1 - q^{m'}}\right) G_{12}(x, y, n) \Bigg\} \end{aligned} \tag{88}$$

and

$$G_{12}(m, n) = 1 + \min \left\{ \min_{\substack{x, y \geqq 1 \\ x+y \leqq m}} G_{12}'(x, y), \min_{\substack{1 \leqq x \leqq m-1 \\ 1 \leqq y \leqq n-m}} G_{12}''(x, y) \right\}, \tag{89}$$

where $G_{12}'(x, y)$ and $G_{12}''(x, y)$ are defined by

$$\begin{aligned} G_{12}'(x, y) = &\left(\frac{q^{x+y} - q^m}{1 - q^m}\right) G_{12}(m - x - y, n - x - y) \\ &+ \frac{q^x(1 - q^y)}{1 - q^m} G_{12}(y, n - x) \\ &+ \frac{q^y(1 - q^x)}{1 - q^m} G_{12}(x, n - y) \\ &+ \frac{(1 - q^x)(1 - q^y)}{1 - q^m} G_{12}(x, y, n) \end{aligned} \tag{90}$$

and

$$\begin{aligned} G_{12}''(x, y) = {} & \frac{q^y(q^x - q^m)}{1 - q^m} G_{12}(m - x, n - x - y) \\ & + \frac{q^y(1 - q^x)}{1 - q^m} G_{12}(x, n - y) \\ & + \frac{(1 - q^y)(q^x - q^m)}{1 - q^m} G_{12}(m - x, y, n - x) \\ & + \frac{(1 - q^y)(1 - q^x)}{1 - q^m} G_{12}(x, y, n). \end{aligned} \tag{91}$$

The boundary conditions state that, for all q,

$$G_{12}(m, 1, n) = G_{12}(1, m, n) = G_{12}(m, n - 1) \quad \text{for} \quad 0 \leqq m \leqq n - 1, \tag{92}$$

$$G_{12}(1, n) = H_{12}(n - 1) \quad \text{for} \quad n \geqq 1, \tag{93}$$

$$H_{12}(0) = 0. \tag{94}$$

It is easy to see that $H_{12}(1) = G_{12}(2, 2) = H_{12}(2) = 1$ and $G_{12}(3, 4) = G_{12}(4, 4) = 2$ for all q.

Remark 6: The extra complication in (89) insures that, for $n \geqq 2$, one experimenter will not be idle while another is carrying out a test.

Remark 7: It is conjectured that in (89) the possibility $x + y = m$ can be omitted, with the exception of the single case $m = n = 2$ (and $m' = 0$).

Remark 8: It is also conjectured that G_{12}'', which is needed for the cases $n > m = 2$, can be disregarded when $m > 2$; i.e., that G_{12}' always gives a smaller minimum for $m > 2$.

Remark 9: It is conjectured that, at any stage in which $m = m' \geqq 2$ or in which we have both $m = m' = 0$ and n even, the two test group sizes, x and y, will be equal. If either m or $m' = 0$, it is conjectured that the two test group sizes will differ by at most unity.

Further calculations yield

		Test Group Sizes x y	
$G_{12}(2, 3) = \dfrac{2 + q}{1 + q}$	for all q	1*D* 1*B*,	(95)
$G_{12}(3, 3) = \dfrac{2 + 2q + q^2}{1 + q + q^2}$	for all q	1*D* 1*D*,	(96)

$$G_{12}(2, 4) = \begin{cases} 2 & \text{for } 0 \leqq q < 0.682 \quad 1D\ 1B, \\ \dfrac{3 + q - q^3}{1 + q} & \text{for } 0.682 \leqq q \leqq 1.000\ 1D\ 2B, \end{cases} \tag{97}$$

$$G_{12}(2, 2, 4) = \frac{2 + 4q + q^2}{(1 + q)^2} \qquad \text{for all } q \qquad 1D\ 1D', \tag{98}$$

$$H_{12}(4) = \begin{cases} 2 & \text{for } 0 \leqq q < 0.691 \quad 1B\ 1B, \\ 3 - 3q^2 + 2q^3 - q^4 & \text{for } 0.691 \leqq q \leqq 1.000\ 2B\ 2B, \end{cases} \tag{99}$$

where $1B$ indicates that 1 unit is taken from the binomial set to form one of the two test-groups and D, D' denote different defective sets.

It is interesting to compare the above result for $H_{12}(4)$ for $q \geqq 0.691$ with the procedure $R_{12}{}^*$ of giving each experimenter two units to analyze independently of each other and without any mutual cooperation. Let T denote the total number of tests and S denote the number of stages required. Then, for $q \geqq 0.691$ (letting T_1, T_2 denote the number of tests in two independent experiments with $n = 2$ under R_1), it is easily shown that

$$E\{S \mid R_{12}{}^*\} = E\{\max(T_1, T_2) \mid R_1\} = 3 - q^2 - q^4, \tag{100}$$

$$E\{T \mid R_{12}\} = 2H_{12}(4) - 2q - 2q^2 + 2q^3 - 2q^4. \tag{101}$$

Hence we find that, for $q \geqq 0.691$,

$$E\{S \mid R_{12}\} - E\{S \mid R_{12}{}^*\} = H_{12}(4) - (3 - q^2 - q^4) = -2q^2(1 - q) \leqq 0, \tag{102}$$

$$E\{T \mid R_{12}\} - E\{T \mid R_{12}{}^*\} = 6 - 2q - 2q^2 + 2q^3 - 2q^4 - 2H_1(2) = 2q^3(1 - q) \geqq 0, \tag{103}$$

which illustrates the fact that R_{12} effects an improvement in the expected number of stages at the expense of a slight increase in the expected total number of tests.

3. In this generalization we apply the restriction that any one unit can be included in at most K group-tests. This is particularly appropriate in the blood testing application, where a single blood sample can be used in a small number K of blood tests and the patient does not want to be annoyed by having more than one blood sample taken.

In this problem there is again only one defective set but it is now denoted by a vector $\mathbf{m} = \{m_0, m_1, \cdots, m_{K-1}\}$, where $m_j \geqq 0$ is the num-

ber of units that have already been included in j group-tests. Similarly, the union of the binomial and defective sets is denoted by

$$\mathbf{n} = \{n_0, n_1, \cdots, n_{K-1}\}$$

and the binomial set alone is the difference $\mathbf{n} - \mathbf{m}$. The symbol for the size of the next test group will be $\mathbf{x} = \{x_0, x_1, \cdots, x_{K-1}\}$, where x_j is the number of units taken from n_j in an H-situation (from m_j in a G-situation). The symbols x, m, n will be used for the sum of the components in the vectors $\mathbf{x}$, $\mathbf{m}$, $\mathbf{n}$, respectively. Let $G_{13}{}^K(\mathbf{m}; \mathbf{n})$ denote the expected number of group-tests required to remove all defective units if the defective set is $\mathbf{m}$ and the binomial set is $\mathbf{n} - \mathbf{m}$. If $m = 0$, we denote this expectation by $H_{13}{}^K(\mathbf{n})$. For the special case in which $\mathbf{n}$ has all except one component (say, n_j) equal to zero, we will drop the zeros and write $H_{13}{}^K(n_j)$, with a scalar argument. The recursion formulae for this procedure $R_{13}{}^K$ are given by

$$H_{13}{}^K(\mathbf{m}) = 1 + \min_{\substack{\text{all } \mathbf{x} \text{ with} \\ 1 \leqq x \leqq n \text{ and} \\ 0 \leqq x_j \leqq n_j \\ (j = 0, 1, \cdots, K-1)}} \{q^x H_{13}{}^K(\mathbf{n} - \mathbf{x}) + (1 - q^x) G_{13}{}^K(\mathbf{x}; \mathbf{n})\}, \tag{104}$$

$$G_{13}{}^K(\mathbf{m}; \mathbf{n}) = 1 + \min_{\substack{\text{all } \mathbf{x} \text{ with} \\ 1 \leqq x \leqq m-1 \\ 0 \leqq x_j \leqq m_j \\ (j = 0, \cdots, K-1)}} \left\{ \left(\frac{q^x - q^m}{1 - q^m} \right) G_{13}{}^K(\mathbf{m} - \mathbf{x}; \mathbf{n} - \mathbf{x}) + \left(\frac{1 - q^x}{1 - q^m} \right) G_{13}{}^K(\mathbf{x}; \mathbf{n}) \right\}, \tag{105}$$

where, as usual, $m > 1$. The boundary conditions state that, for all q, we have $H_{13}{}^K(\mathbf{O}) = 0$ and, for $m = 1$, we can write $G_{13}{}^K(\mathbf{m}; \mathbf{n}) = H_{13}{}^K(\mathbf{n} - \mathbf{m})$. It is easy to see that $H_{13}{}^K(\mathbf{n}) = 1$ for $n = 1$ and all q. Some further computations for $n = 2$ and $n = 3$ give, for any K:

$H_{13}{}^K(2_j) =$

	q – interval	x – value	j – value
$H_1(2)$	$0 \leqq q \leqq 1.000$	(see R_1)	$0 \leqq j \leqq K - 2$
2	$0 \leqq q \leqq 1.000$	$(0, \cdots, 0, 0, 1)$	$j = K - 1$

(106)

$H_{13}{}^K(3_j) =$

$H_1(3)$	$0 \leqq q \leqq 1.000$	(see R_1)	$0 \leqq j \leqq K - 3$
3	$0 \leqq q \leqq 0.618$	$(0, \cdots, 0, 1, 0)$	$j = K - 2$
$4 - q - q^2$	$0.618 \leqq q \leqq 0.707$	$(0, \cdots, 0, 1, 0)$	
$4 - q^2 - 2q^3$	$0.707 \leqq q \leqq 1.000$	$(0, \cdots, 0, 3, 0)$	
3	$0 \leqq q \leqq 1.000$	$(0, \cdots, 0, 0, 1)$	$j = K - 1$.

(107)

If $K = 1$, all units are tested individually. If $K = 2$, then, after a set of size $m > 1$ is shown to be defective, all the units in that set are tested individually; this is the procedure recommended by Dorfman.[1] A more thorough investigation for particular values of $K \geqq 3$ will be given in a separate paper.

Under R_1 for any $H(n)$-situation the maximum number of additional tests $M^*(n)$ in which any particular unit will be included before experimentation is concluded (allowing a random reordering of units in the binomial and defective sets after every test) occurs when q is close to unity and the sample has all units defective. It is easily seen that, under R_1,

$$M^*(n) = M(n) - (n - 1) = (n + 1)\alpha(n) + 3 - 2^{1+\alpha(n)}, \quad (108)$$

where $\alpha(n)$ and $M(n)$ are defined by (10) and (11). Under ${R_{13}}^K$, let $j(\mathbf{n})$ denote the largest subscript associated with a nonzero component of $\mathbf{n}$. For $K \geqq M^*(n) + j(\mathbf{n})$, the restriction that no unit should be included in more than K group-tests does not affect the procedure R_1, and hence we have for procedure ${R_{13}}^K$

$$H_{13}{}^K(\mathbf{n}) = H_1(n) \quad \text{for} \quad 0 \leqq j(\mathbf{n}) \leqq K - M^*(n), \quad (109)$$

which generalizes some results in (106) and (107) and shows that ${R_{13}}^K$ is a generalization of R_1.

Since units in the last component of the defective or the binomial sets cannot be tested in groups, we can remove at any time for individual testing all units in the last component of the binomial set and all but one of the units in the last component of the defective set without affecting the expected number of tests under ${R_{13}}^K$. It is easy to show that this leads to the two reduction formulae

$$H_{13}{}^K(\mathbf{n}) = n_{K-1} + H_{13}{}^K(n_0, \cdots, n_{K-2}, 0), \quad (110)$$

$$\begin{aligned} G_{13}{}^K(\mathbf{m}; \mathbf{n}) = n_{K-1} &+ \left(\frac{1 - q^{-1+m_{K-1}}}{1 - q^m}\right) H_{13}{}^K(n_0, \cdots, n_{K-2}, 0) \\ &+ q^{-1+m_{K-1}} \left(\frac{1 - q^{+1-m_{K-1}}}{1 - q^m}\right) \\ &\cdot [G_{13}{}^K(m_0, \cdots, m_{K-2}, 1; n_0, \cdots n_{K-2}, 0) - 1], \end{aligned} \quad (111)$$

which are useful in computations and for checking.

It is conjectured that, for $m = 0$ or $m > 1$, the procedure ${R_{13}}^K$ can always be carried out by putting in the next test-group only units that have been included in the same number of group tests (i.e., units in the same subset); the only possible exception to this is that, in any H-situa-

tion, if q is sufficiently close to unity, then $R_{13}{}^{K}$ will call for a test of all the remaining units; i.e., $\mathbf{x} = \mathbf{n}$.

Under the above conjecture, it is possible to carry out a simplification as in Section V and show, in direct analogy with (19), that

$$F_{13}{}^{*K}(\mathbf{m}) = \sum_{i=1}^{m} q^{i-1} + \min_{1 \leqq x \leqq m_i - 1} \{q^x F_{13}{}^{*K}(\mathbf{m} - \mathbf{x}) + F_{13}{}^{*K}(\mathbf{x})\}, \quad (112)$$

where $\mathbf{x}$ has x_i in the $(i + 1)$th position and zeros elsewhere (so that $x = x_i$), and i is defined as the subscript associated with the first non-zero component of $\mathbf{m}$. The function $F_{13}{}^{K}(\mathbf{m})$ is defined as the expected number of group-tests required to reach the next H-situation, and then we define, as in Section V,

$$F_{13}{}^{*K}(\mathbf{m}) = \left(\frac{1 - q^m}{1 - q}\right) F_{13}{}^{K}(\mathbf{m}). \quad (113)$$

Hence under the above conjecture it is again seen that, for any G-situation with $m > 1$, the next test-group $\mathbf{x}$ depends on $\mathbf{m}$ but is independent of $\mathbf{n} - \mathbf{m}$.

It appears to be true (but has not been rigorously proved) that, in this case also, for $q < q_0 = 0.618$ (to three decimal places) all units are tested one at a time.

XII. AN ASYMPTOTIC FORMULA FOR $H_1(n)$

In this section we shall use results obtained by considering an information procedure R_2, which is defined in Appendix A. The procedure R_2 appears to be a best test in the sense that it maximizes the information in the very next test but does not take into account the exact finite number of units present and the possible ways of distributing them among subsequent tests. It is therefore intuitively reasonable to expect that the procedure R_1 tends toward R_2 in the H-situation as $n \to \infty$ and also in the G-situation as $m \to \infty$. A more rigorous proof of this assertion would be desirable. It should also be pointed out that there is considerable numerical evidence in Tables IIIA and IIIB of the above assertion, which explains the reason for putting opposite $m = \infty$ and $n = \infty$ in these tables the polynomial equations

$$1 - q^x - q^{x+1} = 0 \qquad (x = 1, 2, \cdots), \quad (114)$$

which are derived for procedure R_2 in Appendix A.

We shall now derive an asymptotic formula for $H_1(n)$ for large n based on the assumption that the above reasoning is correct. For large

values of n and fixed q, the expected number of tests required under procedure R_1 is approximately given by

$$H_1(n) \cong n \left(\frac{\text{expected number of tests needed to reach the next } H\text{-situation under } R_1}{\text{expected number of units analyzed between } H\text{-situations under } R_1} \right). \quad (115)$$

The ratio of n to the denominator in (115) is the approximate number of H-situations reached if we start with n units, and this is clearly to be multiplied by the expected number of tests required to proceed from one H-situation to the next. Let T and U denote the chance variables in the numerator and denominator, respectively, of (115). For a fixed q we find from the limiting procedure R_2 that, for an $H(n)$-situation with n large, we will, under R_1, "almost always" be using the same test-group size x, where x is that positive integer for which q^x is closer to one-half than is either q^{x-1} or q^{x+1}. Then, for this fixed integer x, which depends on the given q, we have

$$E\{U \mid R_1\} = xq^x + p \sum_{j=1}^{x} jq^{j-1} = \frac{1 - q^x}{1 - q}, \quad (116)$$

which is obtained by assuming a single randomization of the order of the units at the outset and considering the different possible positions of the first defective.

Since $F_1(x)$ is the expected number of tests required under R_1 to get from a $G(m, n)$-situation to the next $H(n)$-situation, we have

$$E\{T \mid R_1\} = q^x + (1 - q^x)[1 + F_1(x)] = 1 + pF_1^*(x), \quad (117)$$

where $F_1^*(x)$ is tabulated in Table IV A for $x = 2(1)16$ and all values of q. Hence, we obtain from (115), (116) and (117)

$$H_1(n) \cong \frac{np[1 + pF_1^*(x)]}{1 - q^x}, \quad (118)$$

where x is defined above in terms of q. This is the main result of this section; we now consider some special cases.

For values of q close to unity we can use (23) to replace $F_1^*(x)$ by an explicit expression. If we also replace $(1 - q^x)/(1 - q)$ by x for q close to unity, we obtain for q close to unity

$$H_1(n) \cong \frac{n}{x} \{1 + p[x\alpha(x) + 2\beta(x)]\}, \quad (119)$$

where $\alpha(x)$ and $\beta(x)$ are defined in (20a). If q approaches unity, x be-

comes large; if $\alpha(x) \geqq 2$, then $2\beta(x) < 2^{\alpha(x)+1} \leqq x\alpha(x)$ and, since $\alpha(x) \to \infty$, it follows that we can disregard $2\beta(x)$ in (119). For q close to unity and x large, the dividing points get closer and closer, and we obtain

$$0 \cong 1 - q^x - q^{x+1} \cong 1 - 2q^x, \tag{120}$$

so that $x \cong \left[\log_2 \left(\frac{1}{q}\right)\right]^{-1}$. Also, from the definition of $\alpha(x)$,

$$2^{\alpha(x)} \leqq x \leqq 2 \cdot 2^{\alpha(x)}, \tag{121}$$

so that

$$\log_2 \left(\frac{x}{2}\right) \leqq \alpha(x) \leqq \log_2 (x). \tag{122}$$

Using the upper value in (122) gives

$$H_1(n) \cong n \log_2 \left(\frac{1}{q}\right) + np \log_2 \left[\log_2 \left(\frac{1}{q}\right)\right]^{-1}. \tag{123}$$

It can be shown that the first term in (123) goes to zero faster than the second as q approaches unity, and hence we drop the first term and rewrite the second in the form

$$H_1(n) \cong -np \log_2 p. \tag{124}$$

In particular, if $p = 1/n$ so that $q = 1 - (1/n)$, we obtain from (124)

$$H_1(n) \cong \log_2 n. \tag{125}$$

In this case, we also have $x = n = 1/p$, and a better estimate is obtained by setting $q^x = e^{-1}$ in (23) and (118). Assuming that (23) is either equal to $F_1{}^*(x)$ or is a good approximation to it, we obtain

$$H_1(n) \cong \frac{n}{x}\left[\frac{1 + pF_1{}^*(x)}{1 - e^{-1}}\right] = \frac{n}{x}\left[1 + \alpha(x) + \frac{e^{2\beta(x)/x}}{e - 1}\right]. \tag{126}$$

For the case $n = N = 100$ and $q = 0.99$, we obtain 8.32 as the exact value from Table V C; $\log_2 100 = 6.64$, from (125); 7.79, from (119); and 8.20 from (126).

A rough lower bound on $H(n)$ for any procedure can be easily obtained from information theory. The total information in n units is

$$-n(p \log_2 p + q \log_2 q),$$

and this is to be equated with the product of the expected number of tests $H(n)$ and the average information obtained per group-test. Since

the maximum information per group-test (in bits) is unity, we obtain for any procedure

$$H(n) \geqq -n(p \log_2 p + q \log_2 q). \tag{127}$$

For $n = N = 100$ and $q = 0.99$, this gives 8.09 as a lower bound. Since a result better than 8.09 is impossible, the smallness of the difference $8.32 - 8.09 = 0.23$ is an indication of how far R_1 can possibly be from an optimal solution. However, it should not be inferred that the lower bound (127) can be reached for any value of q (except possibly for $p = q = \frac{1}{2}$) by any procedure. In fact, for $q < \frac{1}{2}$ and q decreasing towards zero, it has been shown[4] that an optimal procedure must have $H(n) = n$, whereas the right member of (127) approaches zero.

It has been pointed out to the authors by S. W. Roberts that a lower bound for $G(m, n)$ for any procedure is easily shown to be

$$\begin{aligned} G(m,n) &\geqq \sum_{i=1}^{m} \frac{pq^{i-1}}{1-q^m}\left[\log_2\left(\frac{pq^{i-1}}{1-q^m}\right)\right. \\ &\qquad \left. + (n-i)(p\log_2 p + q\log_2 q)\right] \\ &= \frac{1}{1-q^m}[q^m \log_2 q^m + (1-q^m)\log_2(1-q^m)] \\ &\qquad - \left(n + \frac{mq^m}{1-q^m}\right)(p\log_2 p + q\log_2 q). \end{aligned} \tag{128}$$

XIII. LACK OF OPTIMALITY OF PROCEDURE R_1

To illustrate the fact that R_1 is not optimal in the general case when units are identifiable and "mixing" of units from the binomial and defective sets is allowed, we shall describe a method of obtaining an improvement on R_1. It is sufficient to consider the case $N = 3$, but the case $N = 4$ is more typical, and we shall use the latter. Let R_0^* denote a procedure for $N = 4$, part of which is described by Fig. 2 and the remaining part of which is arbitrary. (We can therefore also regard R_0^* as a set of procedures, with the common part shown in Fig. 2.) Let a_1, a_2, b_1, b_2 denote individual units; it will be assumed that the a-units are distinguishable from the b-units. The part of Fig. 2 enclosed by dashed lines is different from R_1, since it includes mixing; the rest of the procedure agrees with R_1 for q close to unity. For q close to unity and $m = 2$, after the first two group-tests result in failure, we should act as if there was exactly one defective present until it is proved otherwise. Then, for q close to unity, the above procedure R_0^* terminates in one or two ad-

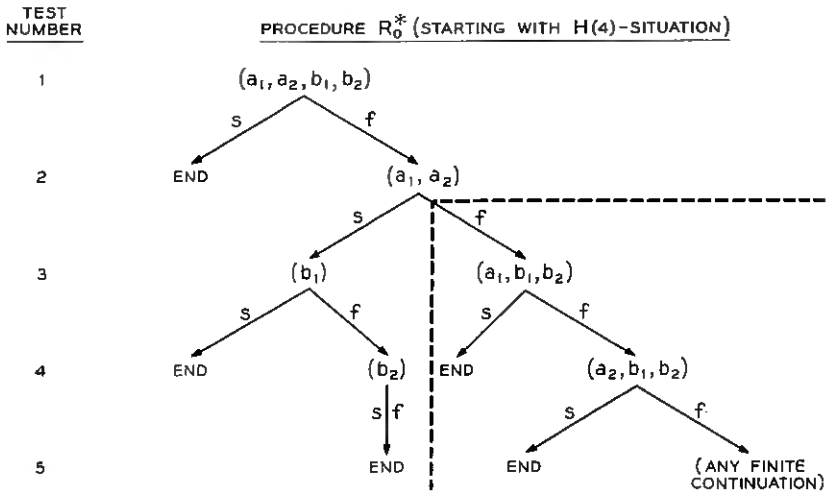


Fig. 2 — Procedure R_0^*.

ditional tests with probability close to $\frac{1}{2}$ for each. More precisely, we obtain for the conditional expected number of additional tests required under R_0^*, given that the first two tests result in a failure,

$$G(2,4 \mid R_0^*) = \frac{pq^3}{1-q^2} + \frac{2pq^3}{1-q^2} + \frac{p^2 f(q)}{1-q^2} = \frac{3q^3 + pf(q)}{1+q}, \quad (129)$$

where $f(q)$ is a polynomial in q. In comparison, we have under R_1 for $0.707 < q < 1.000$

$$\begin{aligned} G_1(2,4) &= 1 + \frac{pH_1(3)}{1-q^2} + \frac{pqH_1(2)}{1-q^2} \\ &= \frac{1}{1+q}(6 + 2q - 2q^2 - 2q^3). \end{aligned} \quad (130)$$

For q approaching unity, the value in (129) approaches $\frac{3}{2}$, while that in (130) approaches 2. This proves that any finite continuation in Fig. 2 will be better than R_1 for q sufficiently close to unity. In a particular procedure to be discussed in a separate paper, the dividing point for the $G(2, 4)$-situation between "no-mixing" and "mixing" is

$$q = (1 + \sqrt{33})/8 = 0.843$$

(to three decimal places). The maximum improvement over R_1 for $n = 4$ in the expected number of tests required for the H-situation is a decrease of 0.04. The price to be paid for this improvement will be an increase in the complexity of the procedure.

ACKNOWLEDGMENT

The authors wish to thank J. E. Clark of Bell Telephone Laboratories for bringing this problem to their attention. Grateful acknowledgment is also due to H. O. Pollak and E. N. Gilbert for suggesting the procedure R_3 and for supplying some numerical values for R_3 given in Table II B. Thanks are also due to S. W. Roberts, Jr., and R. B. Murphy of Bell Telephone Laboratories and I. R. Savage of the University of Minnesota for helpful comments made after reading the paper. We also wish to thank D. E. Carlson, Mrs. W. L. Mammel and D. E. Eastwood, for their help in setting up the program on the IBM-704 for the numerical computations pertaining to procedures R_1 and R_3.

APPENDIX A

The Information Procedure

Another procedure which was investigated is based on choosing that value of x which maximizes the "amount of information" that the next test will give. The amount of information in a test with two outcomes is $p \log_2(1/p) + q \log_2(1/q)$ if p is the probability of either outcome. Hence, equating the information in an H-situation obtained by taking x and $x + 1$ in the next test gives

$$q^x \log_2 q^x + (1 - q^x) \log_2 (1 - q^x) = q^{x+1} \log_2 q^{x+1} + (1 - q^{x+1}) \log_2 (1 - q^{x+1}), \tag{131}$$

whose root will be used as a dividing point between x and $x + 1$. It is easy to verify that, for any integer x and any integer $n \geqq x + 1$, the unique positive root of (131) is also the unique positive root of

$$1 - q^x - q^{x+1} = 0. \tag{132}$$

It is interesting to note that the solution implicit in (132) is easily seen to be equivalent to finding that positive integer x (for fixed, known q) for which q^x is closer to $\frac{1}{2}$ than is either q^{x-1} or q^{x+1}. In fact, by (132), the right endpoint of the interval for x is such that q^x and q^{x+1} are centered about $\frac{1}{2}$ and the left endpoint of the interval for x is such that q^x and q^{x-1} are centered about $\frac{1}{2}$. Similarly, for the G-situation with $m > 1$, we equate

$$\left(\frac{q^x - q^m}{1 - q^m}\right) \log_2 \left(\frac{q^x - q^m}{1 - q^m}\right) + \left(\frac{1 - q^x}{1 - q^m}\right) \log_2 \left(\frac{1 - q^x}{1 - q^m}\right) \tag{133}$$

with the same expression, except that x is replaced by $x + 1$, and find

that, for any $n \geqq m > 1$, the dividing point between x and $x + 1$ is the unique root in the interior of the unit interval (if it exists) of

$$1 - q^x - q^{x+1} + q^m = 0. \tag{134}$$

If we remove the root $q = 1$ in (134), the dividing point is the unique positive root (if it exists) of

$$1 + q + q^2 \cdots + q^{x-1} - q^{x+1} - q^{x+2} - \cdots - q^{m-1} = 0. \tag{135}$$

If the root does not exist for some $m > 1$, then $x + 1$ will never be used for that m. It should be noted that the left member of (134) is a strictly increasing function of x and, for $x \geqq (m - 1)/2$, $m > 1$ and any fixed q with $0 \leqq q < 1$, we have

$$1 - q^x - q^{x+1} + q^m \geqq (1 - q^{(m-1)/2})(1 - q^{(m+1)/2}) > 0. \tag{136}$$

It follows that the highest value of x for which a nondegenerate root exists is such that $x + 1 < (m + 1)/2$ and hence, under this procedure, we never take a test group of size greater than $m/2$. It is interesting to note that the dividing points for any G-situation do not depend on n.

These equations define a new procedure R_2, which we shall also call the *information procedure*. For this, let $F_2(m)$ denote the expected number of group-tests required to "break up" a defective set of size m, i.e., to reach an H-situation. Let $F_2^*(m) = (1 - q^m)p^{-1}F_2(m)$. Then we can write as in (20), for any n and for the appropriate interval where the next test group is of size x,

$$F_2^*(m \mid x) = \sum_{i=1}^{m} q^{i-1} + q^x F_2^*(m - x) + F_2^*(x) \qquad (m > 1), \tag{137}$$

$$H_2(n \mid x) = 1 + q^x H_2(n - x) + pF_2^*(x) + p\sum_{i=1}^{x} q^{i-1} H_2(n - i). \tag{138}$$

The boundary conditions state that $F_2^*(1) = H_2(0) = 0$ for all q. Those expressions for $F_2^*(m)$ which are used to generate expressions for $H_2(n)$ for $2 \leqq m \leqq n \leqq 12$ are given in Table VIII; the resulting expressions for $H_2(n)$ (with x-values) are given in Table VI. Table VII gives the dividing points for $n = 1(1)100$ and for $m = 1(1)16$, $20(5)100$ for procedure R_2.

It should be noted that the $F_2^*(m)$ as well as the $H_2(n)$-function are not all continuous. At the point of discontinuity the x corresponding to the smaller expectation should be used.

It is interesting to observe in the numerical comparisons of Table III A

that the procedure R_2 compares quite favorably with the procedure R_1. In addition, the fact that the dividing points are easier to compute makes it better for practical applications, since the dividing points for R_1 are only known exactly up to $n = 16$. It is also interesting to note that the limiting expressions in Table III A as $n \to \infty$ and in Table III B as $m \to \infty$ are the same as (132).

It is interesting to note that a succession of modifications $R_2^{(j)}(j = 1, 2, \ldots)$ of the information procedure, R_2, are possible such that $R_2^{(1)} = R_2$ and $R_2^{(j)} = R_1$ for $j \geqq M(m,n)$. Here $M(m,n)$, as defined in (10), is the maximum number of group-tests required if we start with a $G(m,n)$-situation [where $G(0,n)$ corresponds to $H(n)$]. Under the procedure $R_2^{(j)}$ we find and use that x which maximizes the ratio of the information expected from at most j group-tests to the conditional expected number of tests required given that we will stop after at most j group-tests. In the special case when there is no possibility of stopping before j tests, we can disregard the denominator and simply maximize the information. For the case $j = 1$ this is clearly equivalent to R_2. For $j \geqq M(m,n)$ the information expected from at most j tests is the same regardless of what x is used next and of what sample path is taken, since all units are then analyzed. Hence, the numerator above can be disregarded and the problem is to minimize the denominator or expected number of tests. Under the assumption of "no-mixing" of units from the binomial and defective sets, this gives the procedure R_1.

For any $H(n)$-situation with $n \geqq 4$ and $j \geqq 2$, these procedures appear to eliminate the possibility of taking $n - 1$ units in the next test-group. For example, if $n = 4, j = 2$ and $q > 0.618$, then we will want to compare $x = 2$ and $x = 3$. For $j = 1$, the dividing point between $x = 2$ and $x = 3$ is $q = 0.755$. Since neither $x = 2$ nor $x = 3$ can result in termination after one test, we can disregard the denominator and compare for $x = 2$ and $x = 3$ the information expected from two group-tests. After simplification, the difference between the results expected after $x = 2$ and $x = 3$ can be written as

$$p^2q[(1 + q) \log_2 (1 + q) - q \log_2 q] \geqq 0, \tag{139}$$

which shows that $x = 2$ is preferable to $x = 3$ for all $q > 0.618$. The same result holds for all $j \geqq 2$. Then we find that the dividing point between $x = 2$ and $x = 4$ for $j = 2$ is the nondegenerate root between zero and unity of

$$\begin{aligned}(2 - 2q - q^4 - q^5 - 2q^7)q \log_2 q - (2 + q^6) \\ (1 - q^2) \log_2 (1 + q) - q^4(1 - q^4) \log_2 p = 0,\end{aligned} \tag{140}$$

which is 0.789 to three decimal places. For $j \geqq 8 = M(0,4)$ the corresponding dividing point for $R_2^{(j)} = R_1$ is the root of $1 - q^2 - q^4 = 0$, or 0.786 to three decimal places. Curiously enough, the same result 0.786 is also the dividing point between $x = 2$ and $x = 4$ for $j = 1$.

For the special case $x = 1$, we state without proof that, for any $H(n)$-situation, the dividing point between $x = 1$ and $x = 2$ is again $\frac{1}{2}(\sqrt{5} - 1) = 0.618$ to three decimal places.

Formulae for the expected number of tests under $R_2^{(j)}$ for $1 < j < M(m,n)$ have not been derived in this paper.

APPENDIX B

Definition of Procedure R_3

It may happen in some problems that recombination is undesirable or impossible, or it may be that we are interested in finding out just how much is saved by allowing recombinations. Both are good reasons for considering a procedure R_3 that is similar to R_1 except that recombinations are not allowed. This simply means that any two operationally formed sets cannot be combined to form a new set from which subsequent test-groups are to be taken. In procedure R_1 the possibility of mixing proper subsets of two different sets was never used, and the same will be true for R_3. If both recombinations and mixing are not used, then, as the experiment continues, the operationally formed sets can only be broken down further into smaller and smaller sets, yielding a nested set of partitions; i.e., any two units separated at some stage remain separated in subsequent stages. It follows that any defective set present is not affected by the number or size or nature of other sets present. Hence, we define $G_3(m)$, with a single argument, as the conditional expected number of group-tests required to remove all the defectives from a set of size m which is known to have at least one defective.

The recursion formulae for R_3 are, for $n \geqq 1$ and $m > 1$,

$$H_3(n) = 1 + \min_{1 \leqq x \leqq n} \{q^x H_3(n - x) + (1 - q^x)[G_3(x) + H_3(n - x)]\}, \tag{141}$$

$$G_3(m) = 1 + \min_{1 \leqq x \leqq m-1} \left\{\left(\frac{q^x - q^m}{1 - q^m}\right) G_3(m - x) + \left(\frac{1 - q^x}{1 - q^m}\right)[G_3(x) + H_3(n - x)]\right\}, \tag{142}$$

with boundary conditions $H_3(0) = G_3(1) = 0$ for all q. If we let $G_3^*(m)$ denote $(1 - q^m)p^{-1}G_3(m)$ and simplify, we obtain

$$H_3(n) = 1 + \min_{1 \leqq x \leqq n} \{H_3(n - x) + pG_3^*(x)\}, \tag{143}$$

$$G_3^*(m) = \sum_{i=1}^{m} q^{i-1} + \min_{1 \leqq x \leqq m-1} \left\{ G_3^*(x) + q^x G_3^*(m - x) + H_3(n - x) \left(\sum_{i=1}^{x} q^{i-1} \right) \right\}, \tag{144}$$

with boundary conditions $H_3(0) = G_3^*(1) = 0$ for all q.

Numerical comparisons of the results for R_1 and R_3 are given in Tables II A and II B.

APPENDIX C

Definition of Two Halving Procedures

Two "halving" procedures R_4 and R_5 are defined below, the principal purpose being to compare the results on the expected number of tests required with comparable results for R_1 and R_3. The procedure R_4 allows recombinations exactly as in R_1, while R_5 is the same as R_4 except that recombinations are not allowed. Both R_4 and R_5 are of particular interest, since they can be carried out without knowing the true value of q.

The procedure R_4 is carried out like R_1 except that, if the defective set is of size $m > 1$, the next test group is a subset of size $m' = [m/2]$ (i.e., the largest integer contained in $m/2$) randomly selected from the defective set and, if $m = 0$ or 1, the entire binomial set is used in the next test-group. In particular, we start with all N units in the first test-group. The recursion formulae for R_4 are

$$H_4(n) = q^n + (1 - q^n)[1 + G_4(n, n)] = 1 + (1 - q^n)G_4(n, n), \tag{145}$$

$$G_4(m, n) = 1 + \left(\frac{1 - q^{m'}}{1 - q^m} \right) G_4(m', n) + q^{m'} \left(\frac{1 - q^{m-m'}}{1 - q^m} \right) G_4(m - m', n - m') \qquad (m > 1), \tag{146}$$

with the same boundary conditions as in R_1. If we let $F_4(m)$ denote the expected number of tests required to break up a defective set of size m, it can be shown as in the case of R_1 that

$$G_4(m, n) = F_4(m) + \left(\frac{p}{1 - q^m}\right) \sum_{i=1}^{m} q^{i-1} H_4(n - i). \qquad (147)$$

If we let

$$G_4^*(m,n) = \left(\frac{1 - q^m}{1 - q}\right) G_4(m,n) \quad \text{and} \quad F_4^*(m) = \left(\frac{1 - q^m}{1 - q}\right) F_4(m), \qquad (148)$$

the recursion formulae for R_4 reduce to

$$F_4^*(m) = \sum_{i=1}^{m} q^{i-1} + F_4^*(m') + q^{m'} F_4^*(m - m') \qquad (m > 1), \qquad (149)$$

$$\begin{aligned} H_4(n) &= 1 + pG_4^*(m, n) \\ &= 1 + p\{F_4^*(n) + \sum_{i=1}^{n} q^{i-1} H_4(n - i)\}, \end{aligned} \qquad (150)$$

with boundary conditions $F_4^*(1) = H_4(0) = 0$ for all q. For $n \leqq 5$, the results are the same as those for R_1, if we take q close to unity in the formulae for R_1. The results for $H_4(n)$ for $n = 6(1)12$ are, for all q:

$$\begin{aligned} H_4(6) &= 14 - 9q - 2q^2 - q^3 - q^5, \\ H_4(7) &= 17 - 11q - 3q^2 - q^3 + q^4 - 2q^5, \\ H_4(8) &= 21 - 14q - 4q^2 - q^3 + q^4 - 2q^5, \\ H_4(9) &= 25 - 18q - 4q^2 - q^3 + q^4 - 2q^5 + q^7 - q^9, \\ H_4(10) &= 29 - 22q - 4q^2 + q^4 - 3q^5 + q^7 - q^9, \\ H_4(11) &= 33 - 26q - 4q^2 + q^3 - 4q^5 + 2q^6 + q^7 - q^8 - q^9, \\ H_4(12) &= 37 - 29q - 4q^2 - 4q^5 + 2q^6 + q^7 - q^8 - q^9. \end{aligned} \qquad (151)$$

The recursion formulae for the halving procedure R_5, which continues to separate sets into smaller and smaller subdivision, are

$$H_5(n) = 1 + (1 - q^n) G_5(n), \qquad (152)$$

$$\begin{aligned} G_5(m) = 1 &+ \left(\frac{1 - q^{m'}}{1 - q^m}\right) [G_5(m') + H_4(m - m')] \\ &+ q^{m'} \left(\frac{1 - q^{m-m'}}{1 - q^m}\right) G_5(m - m') \quad (m > 1), \end{aligned} \qquad (153)$$

where m' is defined as above. Here it was not necessary to use a double argument with G because there is no recombination allowed. The bound-

ary condition is $G_5(1) = 0$ for all q. If we define $G_4^*(m)$ as in (153) and use (152), the recursion formulae reduce to

$$H_5(n) = 1 + pG_5^*(n), \tag{154}$$

$$G_5^*(m) = \sum_{i=1}^{m'} q^{i-1} + \sum_{i=1}^{m} q^{i-1} + G_5^*(m') + G_5^*(m - m') \qquad (m > 1), \tag{155}$$

with the boundary condition $G_5^*(1) = 0$ for all q. For $n \leqq 3$, the results are the same as for R_1 if we take q close to unity in the formulae for R_1. The results for $H_5(n)$ for $n = 4(1)12$ are:

$$\begin{aligned}
H_5(4) &= 7 - 2q - 3q^2 - q^4, \\
H_5(5) &= 9 - 3q - 3q^2 - q^3 - q^5, \\
H_5(6) &= 11 - 4q - 2q^2 - 3q^3 - q^6, \\
H_5(7) &= 13 - 4q - 4q^2 - 2q^3 - q^4 - q^7, \\
H_5(8) &= 15 - 4q - 6q^2 - 3q^4 - q^8, \\
H_5(9) &= 17 - 5q - 6q^2 - q^3 - 2q^4 - q^5 - q^9, \\
H_5(10) &= 19 - 6q - 6q^2 - 2q^3 - 3q^5 - q^{10}, \\
H_5(11) &= 21 - 7q - 5q^2 - 4q^3 - 2q^5 - q^6 - q^{11}, \\
H_5(12) &= 23 - 8q - 4q^2 - 6q^3 - 3q^6 - q^{12}.
\end{aligned} \tag{156}$$

APPENDIX D

Known Procedures

An attempt has been made to put the Dorfman procedure[1] and the Sterrett procedure[3] in the best form that is comparable with the other procedures treated here. For each of $N = 4$, 8 and 12, we have found the division into equal (or approximately equal) subsets such that the Dorfman plan of testing defective sets one at a time gives the smallest possible expected number of tests required. It should not be inferred that these results would be the same if a straightforward application of the tables published by Dorfman and Sterrett, respectively, were made, since their tables are only concerned with very large N. In the Dorfman plan we use a common test-group size for binomial sets and, for defective sets, the units are all tested one at a time. In the Sterrett plan, there is a common test group size for binomial sets at the outset and, for defective sets, the units are tested one at a time only until a defective unit is found. Then the remaining units, from that defective set only, are pooled and tested. This is continued until that particular defective set is completely analyzed before we start with other sets. We have also assumed that

logical inference would be used whenever possible in the Dorfman and Sterrett procedures.

For the Dorfman procedure R_7, if the common group size is c, then, for any binomial set of size n (where $n \leqq c$), we obtain

$$H_7(n) = q^n + (1 - q^n)\left(n + 1 - \frac{pq^{n-1}}{1 - q^n}\right) = n + 1 - q^{n-1} - (n - 1)q^n. \tag{157}$$

For example, for $N = 12$ and $q = 0.90$, we find that $c_1 = c_2 = c_3 = 4$ gives the best results and, using (157), with $c = 4$, we obtain

$$H_7(12) = 3H_7(4) = 3(5 - q^3 - 3q^4) = 6.908. \tag{158}$$

For the Sterrett procedure R_6, if the group size at the outset is c, then, for each binomial set of size n (where $n \leqq c$), we obtain

$$\begin{aligned} H_6(n) &= q^n + p[2 + H_6(n-1)] + qp[3 + H_6(n-2)] \\ &\quad + \cdots + q^{n-2}p[n + H_6(1)] + q^{n-1}pn \\ &= q^n + npq^{n-1} + p\sum_{i=2}^{n} iq^{i-2} + p\sum_{i=1}^{n-1} q^{i-1}H_6(n-i), \end{aligned} \tag{159}$$

with boundary condition $H_6(1) = 1$ for all q. It can be verified (we omit the details) that the solution of this system is given by

$$H_6(n) = (2n - 1) - (n - 1)q - \sum_{i=2}^{n} q^i. \tag{160}$$

For example, for $N = 12$ and $q = 0.90$, we find that $c_1 = c_2 = c_3 = 4$ gives the best results and, using (160), with $c = 4$, we obtain

$$H_6(12) = 3H_6(4) = 6.315. \tag{161}$$

APPENDIX E

Cost Considerations

In this Appendix we introduce another procedure R_8, which brings into play the cost of throwing away a good unit and balances it against the cost of conducting another group-test. It is interesting to note that R_8 was the solution given when the problem was first brought to the authors' attention in a practical application.

For procedure R_8 we divide all the N units into approximately equal subsets of size x', where x' is the nearest positive integer to the solution in x of

$$(1 - p)^x = \tfrac{1}{2}, \tag{162}$$

where p is the known *a priori* probability of a unit being defective. It is assumed here that $p \ll \frac{1}{2}$. Each subgroup is tested either simultaneously or in sequence, and good subgroups are removed. Assuming $x' > 1$, subsets shown to contain at least one defective are pooled. Since the same number of defectives has now been put into a pooled set of approximate size $N/2$, it follows that the probability of drawing a defective from the pooled set is approximately double the original *a priori* probability p. Then the pooled subset is again divided into approximately equal subsets of size x'', where x'' is the nearest positive integer to the solution of

$$(1 - 2p)^x = \tfrac{1}{2}. \tag{163}$$

The process is repeated (say) a total of t times. If tp gets larger than $\frac{1}{2}$, x is taken to be unity and the units would then be tested one at a time. However, it may be more economical to stop the procedure before tp reaches $\frac{1}{2}$ and scrap the pooled defective subgroup. The amount of saving may be substantial if the cost of manufacturing a unit, say c_0, is small compared to the cost of each group-test, say c_1.

Suppose, for example, that we start with $N = 8$ units and q is given to be 0.90. The approximate solution of (162) is $x' = 7$ but, since this leaves a subset of size 1, we make our first test on all 8 units. If the test is a success we are through; otherwise we look for a solution of (163) and find that $x'' = 3$. The 8 units are divided into subsets of size 3, 3 and 2, and each is tested. Good subsets are removed. We purposely avoid the next stage which requires testing one at a time. Hence, any of these three sets of size 3, 3, and 2 that proves to contain at least one defective is scrapped.

Let T denote the number of tests required in the above example and let D denote the total number of scrapped units (i.e., good units and defectives that are discarded). Then, for this procedure R_8 with $N = 8$ and $q = 0.90$, we obtain

$$E\{T \mid R_8\} = 4 - 3q^8 = 2.709, \tag{164}$$

$$E\{D \mid R_8\} = 8 - 2q^2 - 6q^3 = 2.006. \tag{165}$$

If we define the expected loss $E\{L \mid R_i\}$ for any procedure R_i by

$$E\{L \mid R_i\} = c_0 E\{D \mid R_i\} + c_1 E\{T \mid R_i\}, \tag{166}$$

we find for procedures R_1 and R_8, respectively,

$$E\{L \mid R_1\} = 0.800\, c_0 + 3.904\, c_1, \tag{167}$$

$$E\{L \mid R_8\} = 2.006\, c_0 + 2.709\, c_1. \tag{168}$$

A comparison of these two expressions shows that R_8 will be more economical in this case if the ratio

$$\frac{c_1}{c_0} \gtrless \frac{2.006 - 0.800}{3.904 - 2.709} = 1.009 = 1, \text{ approximately.} \qquad (169)$$

Hence, R_8 is more economical in this case if the total cost of a single test is greater than the total cost of manufacturing a single unit.

Similarly, it can be shown that it would be more economical to stop after the first test (and scrap all 8 units if there is at least one defective present) when the ratio of the two costs in (169) is greater than 1.492 (or approximately 1.5).

REFERENCES

1. Dorfman, R., The Detection of Defective Members of Large Populations, Ann. Math. Stat., **14,** 1943, p. 436.
2. Feller, W., *An Introduction to Probability Theory and Its Applications*, 2nd edition, John Wiley & Sons, New York, 1951, p. 189.
3. Sterrett, A., On the Detection of Defective Members of Large Populations, Ann. Math. Stat., **28,** 1957, p. 1033.
4. Ungar, P., The Cut-Off Point for Group Testing, to be published.

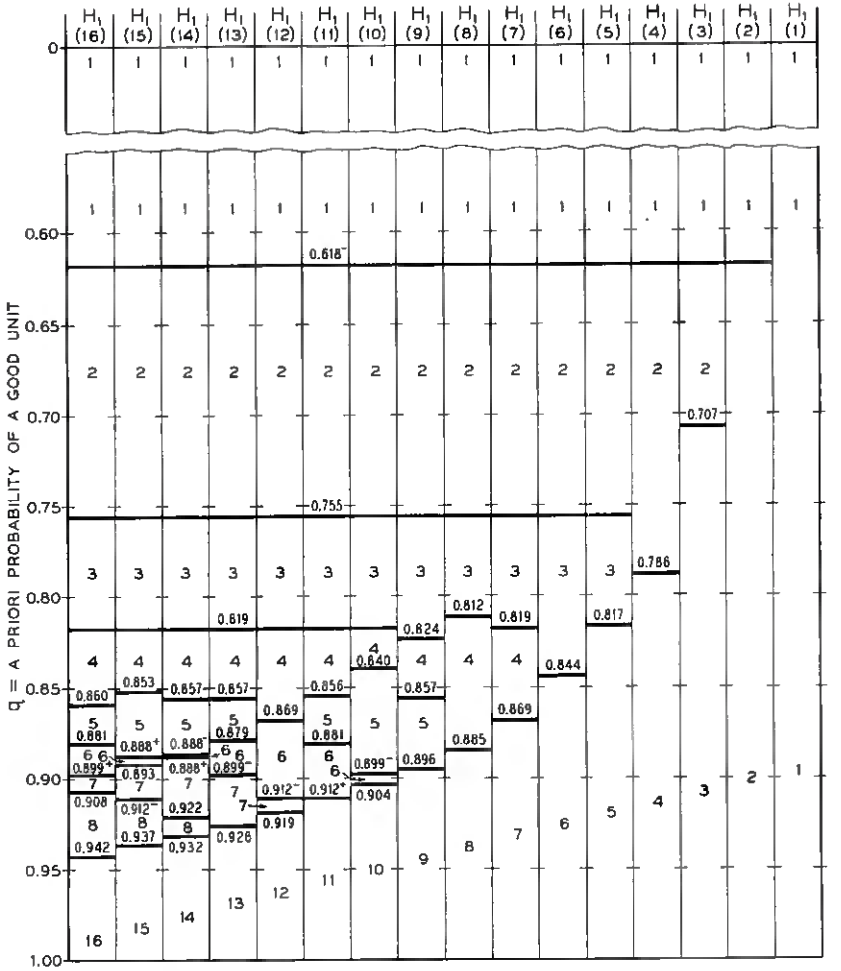


Fig. 3 — Diagram showing number of observations to be taken in any H-situation for $n = 1$ through 16 — for procedure R_1 .

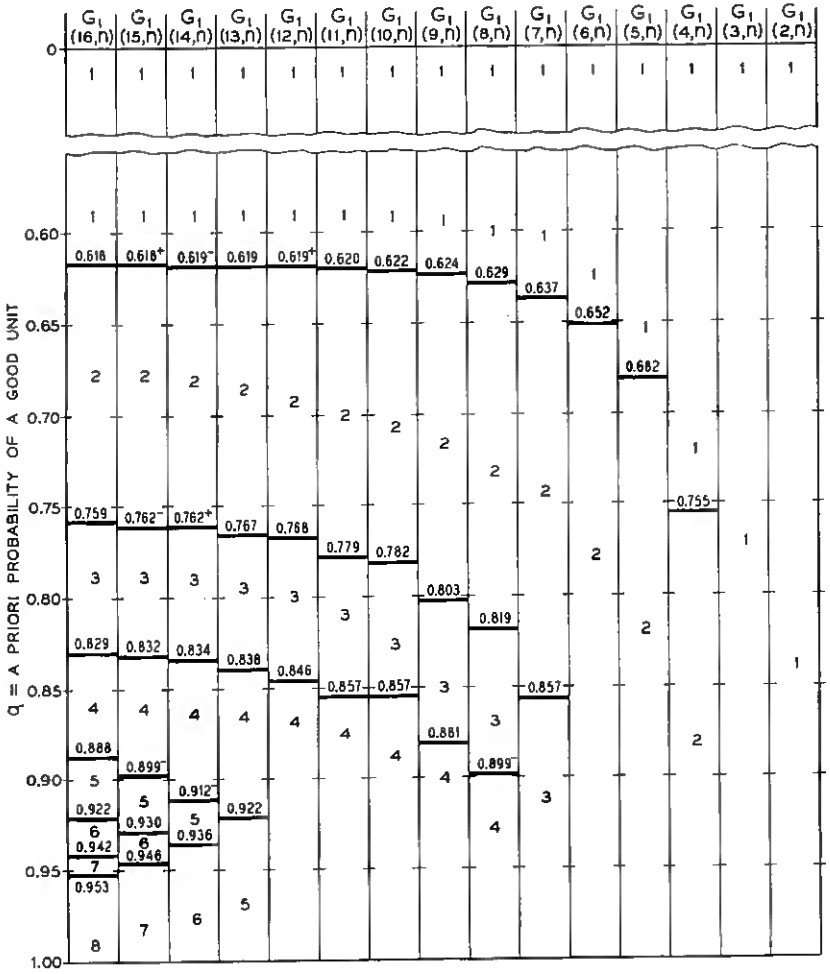


Fig. 4 — Diagram showing number of observations to be taken in any G-situation for $m = 2$ through 16 and any $n \geqq m$ — for procedure R_1 .

TABLE II A — COMPARISON OF THE EXPECTED NUMBER OF GROUP-TESTS REQUIRED FOR DIFFERENT PROCEDURES FOR $N = 4, 8, 12$ AND SELECTED VALUES OF q

Procedure	q							
	0.00	0.50	0.75	0.90	0.95	0.98	0.99	1.00
R_1 (Proposed Procedure)	4.000	4.000	3.332	2.051	1.538	1.218	1.110	1.000
	8.000	8.000	6.619	3.904	2.499	1.612	1.308	1.000
	12.000	12.000	9.905	5.790	3.594	2.073	1.543	1.000
R_2 (Information Procedure)	4.000	4.000	3.332	2.051	1.538	1.218	1.110	1.000
	8.000	8.000	6.663	4.141	2.500	1.612	1.308	1.000
	12.000	12.000	9.956	5.839	3.599	2.078	1.544	1.000
R_3 (Proposed without Recombinations)	4.000	4.000	3.375	2.105	1.576	1.236	1.119	1.000
	8.000	8.000	6.719	4.052	2.631	1.680	1.345	1.000
	12.000	12.000	10.062	6.103	3.851	2.209	1.617	1.000
R_4 (Halving with Recombinations)	4.000	4.000	3.332	2.051	1.538	1.218	1.110	1.000
	21.000	12.875	7.670	3.906	2.500	1.612	1.308	1.000
	37.000	21.408	12.365	6.021	3.620	2.078	1.544	1.000
R_5 (Halving without Recombinations)	7.000	5.188	3.496	2.114	1.578	1.236	1.119	1.000
	15.000	11.309	7.576	4.141	2.678	1.700	1.355	1.000
	23.000	17.203	11.653	6.309	3.900	2.229	1.627	1.000
R_6 (Sterrett)	4.000	4.000	3.375	2.105	1.576	1.236	1.119	1.000
	8.000	8.000	6.750	4.210	3.151	1.807	1.412	1.000
	12.000	12.000	10.125	6.315	4.333	2.973	1.852	1.000
R_7 (Dorfman)	4.000	4.000	3.375	2.303	1.699	1.292	1.148	1.000
	8.000	8.000	6.750	4.605	3.398	2.176	1.609	1.000
	12.000	12.000	10.125	6.908	5.097	3.334	2.354	1.000

TABLE II B — COMPARISON OF $H_1(n)$ AND $H_3(n)$ FOR THREE VALUES OF q AND LARGER n-VALUES

n	$q = 0.90$		$q = 0.95$		$q = 0.99$	
	$H_1(n)$	$H_3(n)$	$H_1(n)$	$H_3(n)$	$H_1(n)$	$H_3(n)$
10	4.872	5.101	3.039	3.242	1.425	1.481
20	9.572	10.155	5.940	6.456	2.051	2.221
30	14.301	15.209	8.791	9.626	2.738	3.057
40	19.024	20.260	11.671	12.798	3.478	3.943
50	23.750	25.361	14.555	16.009	4.243	4.853
60	28.475	30.415	17.438	19.246	5.026	5.792
70	33.200	35.469	20.316	22.415	5.830	6.754
80	37.925	40.520	23.197	25.591	6.647	7.717
90	42.650	45.621	26.078	28.781	7.477	8.683
100	47.375	50.675	28.959	32.019	8.230	9.687

TABLE III A — TO BE USED WITH PROCEDURE R_1 IN ANY $H(n)$ SITUATION WITH $n \leqq 16$ OR n VERY LARGE

Each polynomial shown with its unique real root in the unit interval $0 < q < 1$; these roots determining the region where the next test group is of size x.

The exponential symbols $+$, $-$ indicate only the relative magnitude of two different roots that are equal to three decimal places (i.e., $a^- < a^+$).

n	The unique real root in the unit interval of the polynomial shown is the dividing point between:																			
	$x = 1$ and $x = 2$		$x = 2$ and $x = 3$		$x = 3$ and $x = 4$		$x = 4$ and $x = 5$		$x = 5$ and $x = 6$		$x = 6$ and $x = 7$		$x = 7$ and $x = 8$		(maximum possible $x < n$) and $x = n$	
2	$1-q-q^2$	0.618													(see first column)	
3	$1-q-q^2$	0.618	(see last col.)												$1-2q^2$	0.707
4	$1-q-q^2$	0.618													$1-q^2-q^4$	0.786
5	$1-q-q^2$	0.618	$1-q^2-q^3$	0.755											$1-q+q^3-2q^5$	0.817
6	$1-q-q^2$	0.618	$1-q^2-q^3$	0.755											$1-q+q^2-q^4-q^6$	0.844
7	$1-q-q^2$	0.618	$1-q^2-q^3$	0.755	$1-q^3-q^4$	0.819									$1-q^4-q^5$	0.869
8	$1-q-q^2$	0.618	$1-q^2-q^3$	0.755	$1-q-q^5$	0.812									$1-q^3+q^5-2q^7$	0.885
9	$1-q-q^2$	0.618	$1-q^2-q^3$	0.755	$1-q-q^9$	0.824	$1-q^4-q^5$	0.857							$1-q^3+q^7-2q^9$	0.896
10	$1-q-q^2$	0.618	$1-q^2-q^3$	0.755	$1-q^3-q^4$	0.819	$1-q^2-q^7$	0.840	$1-q^6-q^7$	0.899$^-$					$1-q^2+q^6-2q^{10}$	0.904
11	$1-q-q^2$	0.618	$1-q^2-q^3$	0.755	$1-q^3-q^4$	0.819	$1-q+q^4-q^6-q^9$	0.856	$1-q^5-q^6$	0.881					$1-q^2+q^5-q^9-q^{11}$	0.912$^+$
12	$1-q-q^2$	0.618	$1-q^2-q^3$	0.755	$1-q^3-q^4$	0.819	$1-q^4-q^8$ ($x = 4$ and $x = 6$)*	0.869			$1-q^7-q^8$	0.912$^-$			$1-q+q^4-q^{10}-q^{12}$	0.919
13	$1-q-q^2$	0.618	$1-q^2-q^3$	0.755	$1-q^3-q^4$	0.819	$1-q^4-q^5$	0.857	$1-q+q^6-q^7-q^{11}$	0.879	$1-q^6-q^7$	0.899$^-$			$1-q+q^3-q^9-q^{12}$	0.926
14	$1-q-q^2$	0.618	$1-q^2-q^3$	0.755	$1-q^3-q^4$	0.819	$1-q^4-q^5$	0.857	$1-q+q^5-q^6+q^{10}-2q^{12}$	0.888$^-$	$1-q+q^2-q^3+q^4-q^5-q^{11}$	0.888$^+$	$1-q^5-q^8$	0.922	$1-q^8-q^{12}$	0.932
15	$1-q-q^2$	0.618	$1-q^2-q^3$	0.755	$1-q^3-q^4$	0.819	$1-q-q^{12}$	0.853	$1-q+q^2-q^3+q^4-q^5-q^{11}$	0.888$^+$	$1-q+q^7-q^8+q^{12}-2q^{14}$	0.893	$1-q^7-q^8$	0.912$^-$	$1-q^8-q^{14}$	0.937
16	$1-q-q^2$	0.618	$1-q^2-q^3$	0.755	$1-q^3-q^4$	0.819	$1-q-q^{13}$	0.860	$1-q^5-q^6$	0.881	$1-q+q^5-q^7+q^{13}-2q^{15}$	0.899$^+$	$1-q^7-q^{14}-q^{15}$	0.908	$1-q^8-q^{16}$	0.942
∞	$1-q-q^2$	0.618	$1-q^2-q^3$	0.755	$1-q^3-q^4$	0.819	$1-q^4-q^5$	0.857	$1-q^5-q^6$	0.881	$1-q^6-q^7$	0.899$^-$	$1-q^7-q^8$	0.912		1.000

* The interval for $x = 5$ vanishes for $n = 12$.

TABLE III B — TO BE USED WITH PROCEDURE R_1 IN ANY $G(m, n)$ SITUATION WITH $2 \leqq m \leqq n \leqq 16$ OR m VERY LARGE

Each polynomial shown with its unique real root in the unit interval $0 < q < 1$; these roots determining the region where the next test group (taken from the defective set) is of size x.

The exponential symbols $+$, $-$ indicate only the relative magnitude of two or three different roots that are equal to three decimal places (i.e., $a^- < a^+$ or $a^- < a < a^+$).

m	The unique real root in the unit interval of the polynomial shown is the dividing point between:													
	$x = 1$ and $x = 2$		$x = 2$ and $x = 3$		$x = 3$ and $x = 4$		$x = 4$ and $x = 5$		$x = 5$ and $x = 6$		$x = 6$ and $x = 7$		$x = 7$ and $x = 8$	
2														
3														
4	$1-q^2-q^3$	0.755												
5	$1-q^2-q^3-q^4$	0.682												
6	$1-q^2-q^3-q^4-q^5$	0.652												
7	$1-q^2-q^3-q^4-q^5-q^6$	0.637	$1-q^4-q^5$	0.857										
8	$1-q^2-q^3-\cdots-q^7$	0.629	$1-q^3-q^4$	0.819	$1-q^6-q^7$	0.899^-								
9	$1-q^2-q^3-\cdots-q^8$	0.624	$1-q^3-q^6-q^7$	0.803	$1-q^5-q^6$	0.881								
10	$1-q^2-q^3-\cdots-q^9$	0.622	$1-q^3-q^5-q^6$	0.782	$1-q^4-q^5$	0.857								
11	$1-q^2-q^3-\cdots-q^{10}$	0.620	$1-q^3-q^5-q^8-q^9$	0.779	$1-q^4-q^9-q^{10}$	0.857								
12	$1-q^2-q^3-\cdots-q^{11}$	0.619^+	$1-q^3-q^5-q^7-q^8$	0.768	$1-q^4-q^8-q^9$	0.846								
13	$1-q^2-q^3-\cdots-q^{12}$	0.619	$1-q^3-q^5-q^7-q^{10}-q^{11}$	0.767	$1-q^4-q^8-q^{11}-q^{12}$	0.838	$1-q^6-q^9$	0.922						
14	$1-q^2-q^3-\cdots-q^{13}$	0.619^-	$1-q^3-q^5-q^7-q^9-q^{10}$	0.762^+	$1-q^4-q^7-q^8$	0.834	$1-q^7-q^8$	0.912^-	$1-q^{10}-q^{11}$	0.936				
15	$1-q^2-q^3-\cdots-q^{14}$	0.618^+	$1-q^3-q^5-q^7-q^9-q^{12}-q^{13}$	0.762^-	$1-q^4-q^7-q^{11}-q^{12}$	0.832	$1-q^6-q^7$	0.899^-	$1-q^9-q^{10}$	0.930	$1-q^{12}-q^{13}$	0.946		
16	$1-q^2-q^3-\cdots-q^{15}$	0.618	$1-q^3-q^5-q^7-q^9-q^{11}-q^{12}$	0.759	$1-q^4-q^7-q^{11}-q^{14}-q^{15}$	0.829	$1-q^6-q^{11}-q^{12}$	0.888^+	$1-q^8-q^9$	0.922	$1-q^{11}-q^{12}$	0.941	$1-q^{14}-q^{15}$	0.953
∞	$1-q-q^2$	0.618^-	$1-q^2-q^3$	0.755	$1-q^3-q^4$	0.819	$1-q^4-q^5$	0.857	$1-q^5-q^6$	0.881	$1-q^6-q^7$	0.899	$1-q^7-q^8$	0.912^-

TABLE IV A — FORMULAE FOR $F_1^*(m)$ FOR PROCEDURE R_1 AND VALUES OF THE NEXT TEST-GROUP SIZE x FOR $m = 2(1)16$

	q-interval	x	1	q	q^2	q^3	q^4	q^5	q^6	q^7	q^8	q^9	q^{10}	q^{11}	q^{12}	q^{13}	q^{14}	q^{15}	
$F_1^*(2)$	0.000 to 1.000	1	1	1															
$F_1^*(3)$	0.000 to 1.000	1	1	2	2														
$F_1^*(4)$	0.000 to 0.755	1	1	2	3	3													†
	0.755 to 1.000	2	2	2	2	2													
$F_1^*(5)$	0.000 to 0.682	1	1	2	3	4	4												†
	0.682 to 1.000	2	2	2	2	3	3												
$F_1^*(6)$	0.000 to 0.652	1	1	2	3	4	5	5											†
	0.652 to 0.755	2	2	2	2	3	4	4											†
	0.755 to 1.000	2	2	2	3	3	3	3											
$F_1^*(7)$	0.000 to 0.637	1	1	2	3	4	5	6	6										†
	0.637 to 0.682	2	2	2	2	3	4	5	5										†
	0.682 to 0.857	2	2	2	3	3	3	4	4										†
	0.857 to 1.000	3	2	3	3	3	3	3	3										
$F_1^*(8)$	0.000 to 0.629	1	1	2	3	4	5	6	7	7									†
	0.629 to 0.652	2	2	2	2	3	4	5	6	6									†
	0.652 to 0.755	2	2	2	3	3	3	4	5	5									†
	0.755 to 0.819	2	2	2	3	3	4	4	4	4									†
	0.819 to 0.899⁻	3	2	3	3	3	3	3	4	4									
	0.899⁻ to 1.000	4	3	3	3	3	3	3	3	3									
$F_1^*(9)$	0.000 to 0.624	1	1	2	3	4	5	6	7	8	8								†
	0.624 to 0.637	2	2	2	2	3	4	5	6	7	7								†
	0.637 to 0.682	2	2	2	3	3	3	4	5	6	6								†
	0.682 to 0.803	2	2	2	3	3	4	4	4	5	5								†
	0.803 to 0.881	3	2	3	3	3	3	4	4	4	4								†
	0.881 to 1.000	4	3	3	3	3	3	3	3	4	4								
$F_1^*(10)$	0.000 to 0.622	1	1	2	3	4	5	6	7	8	9	9							†
	0.622 to 0.629	2	2	2	2	3	4	5	6	7	8	8							†
	0.629 to 0.652	2	2	2	3	3	3	4	5	6	7	7							†
	0.652 to 0.755	2	2	2	3	3	4	4	4	5	6	6							†
	0.755 to 0.782	2	2	2	3	3	4	4	5	5	5	5							†
	0.782 to 0.857	3	2	3	3	3	3	4	4	4	5	5							†
	0.857 to 1.000	4	3	3	3	3	3	3	4	4	4	4							
$F_1^*(11)$	0.000 to 0.620	1	1	2	3	4	5	6	7	8	9	10	10						†
	0.620 to 0.624	2	2	2	2	3	4	5	6	7	8	9	9						†
	0.624 to 0.637	2	2	2	3	3	3	4	5	6	7	8	8						†
	0.637 to 0.682	2	2	2	3	3	4	4	4	5	6	7	7						†
	0.682 to 0.779	2	2	2	3	3	4	4	5	5	5	6	6						†
	0.779 to 0.819	3	2	3	3	3	3	4	4	5	5	5	5						†
	0.819 to 0.857	3	2	3	3	3	4	4	4	4	4	5	5						†
	0.857 to 1.000	4	3	3	3	3	3	4	4	4	4	4	4						
$F_1^*(12)$	0.000 to 0.619	1	1	2	3	4	5	6	7	8	9	10	11	11					†
	0.619 to 0.622	2	2	2	2	3	4	5	6	7	8	9	10	10					†
	0.622 to 0.629	2	2	2	3	3	3	4	5	6	7	8	9	9					†
	0.629 to 0.652	2	2	2	3	3	4	4	4	5	6	7	8	8					†
	0.652 to 0.755	2	2	2	3	3	4	4	5	5	5	6	7	7					†
	0.755 to 0.768	2	2	2	3	3	4	4	5	5	6	6	6	6					†
	0.768 to 0.803	3	2	3	3	3	3	4	4	5	5	5	6	6					†
	0.803 to 0.846⁺	3	2	3	3	3	4	4	4	5	5	5	5	5					†
	0.846⁺ to 0.899⁻	4	3	3	3	3	3	4	4	4	4	4	5	5					
	0.899⁻ to 1.000	4	3	3	3	3	4	4	4	4	4	4	4	4					

† These equations are not needed to compute $H_1(n)$-formulae if the experiment *starts* with a "pure binomial" set of any size N (i.e., if $m = 0$ at the outset).

TABLE IV A — *Continued*

	q-interval	x	1	q	q^2	q^3	q^4	q^5	q^6	q^7	q^8	q^9	q^{10}	q^{11}	q^{12}	q^{13}	q^{14}	q^{15}	
	0.000 to 0.619	1	1	2	3	4	5	6	7	8	9	10	11	12	12				†
	0.619 to 0.620	2	2	2	3	4	5	6	7	8	9	10	11	11					†
	0.620 to 0.624	2	2	2	3	3	3	4	5	6	7	8	9	10	10				†
	0.624 to 0.637	2	2	2	3	3	4	4	4	5	6	7	8	9	9				†
	0.637 to 0.682	2	2	2	3	3	4	4	5	5	5	6	7	8	8				†
F_1*(13)	0.682 to 0.767	2	2	2	3	3	4	4	5	5	6	6	6	7	7				†
	0.767 to 0.782	3	2	3	3	3	3	4	4	5	5	6	6	6	6				†
	0.782 to 0.834	3	2	3	3	3	4	4	4	4	5	5	5	6	6				†
	0.834 to 0.881	4	3	3	3	3	3	4	4	4	4	5	5	5	5				†
	0.881 to 0.922	4	3	3	3	3	4	4	4	4	4	4	4	5	5				†
	0.922 to 1.000	5	3	3	3	4	4	4	4	4	4	4	4	4	4				
	0.000 to 0.619^-	1	1	2	3	4	5	6	7	8	9	10	11	12	13	13			†
	0.619^- to 0.619^+	2	2	2	3	4	5	6	7	8	9	10	11	12	12				†
	0.619^+ to 0.622	2	2	2	3	3	4	5	6	7	8	9	10	11	11				†
	0.622 to 0.629	2	2	2	3	3	4	4	4	5	6	7	8	9	10	10			†
	0.629 to 0.652	2	2	2	3	3	4	4	5	5	5	6	7	8	9	9			†
	0.652 to 0.755	2	2	2	3	3	4	4	5	5	6	6	6	7	8	8			†
	0.755 to 0.762^+	2	2	2	3	3	4	4	5	5	6	6	7	7	7	7			†
F_1*(14)	0.762^+ to 0.779	3	2	3	3	3	3	4	4	5	5	6	6	6	7	7			†
	0.779 to 0.819	3	2	3	3	3	4	4	4	4	5	5	6	6	6	6			†
	0.819 to 0.834	3	2	3	3	3	4	4	4	5	5	5	5	5	6	6			†
	0.834 to 0.857	4	3	3	3	3	3	4	4	4	4	5	5	5	6	6			†
	0.857 to 0.912^-	4	3	3	3	3	4	4	4	4	4	4	5	5	5	5			†
	0.912^- to 0.936	5	3	3	3	4	4	4	4	4	4	4	4	4	5	5			
	0.936 to 1.000	6	3	3	4	4	4	4	4	4	4	4	4	4	4	4			
	0.000 to 0.618^+	1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	14		†
	0.618^+ to 0.619	2	2	2	2	3	4	5	6	7	8	9	10	11	12	13	13		†
	0.619 to 0.620	2	2	2	3	3	3	4	5	6	7	8	9	10	11	12	12		†
	0.620 to 0.624	2	2	2	3	3	4	4	4	5	6	7	8	9	10	11	11		†
	0.624 to 0.637	2	2	2	3	3	4	4	5	5	5	6	7	8	9	10	10		†
	0.637 to 0.682	2	2	2	3	3	4	4	5	5	6	6	6	7	8	9	9		†
	0.682 to 0.762	2	2	2	3	3	4	4	5	5	6	6	7	7	7	8	8		†
F_1*(15)	0.762 to 0.768	3	2	3	3	3	3	4	4	5	5	6	6	7	7	7	7		†
	0.768 to 0.803	3	2	3	3	3	4	4	4	4	5	5	6	6	6	7	7		†
	0.803 to 0.832	3	2	3	3	3	4	4	4	5	5	5	5	6	6	6	6		†
	0.832 to 0.857	4	3	3	3	3	3	4	4	4	5	5	5	5	5	6	6		†
	0.857 to 0.899^-	4	3	3	3	3	4	4	4	4	4	5	5	5	5	5	5		†
	0.899^- to 0.930	5	3	3	3	4	4	4	4	4	4	4	4	5	5	5	5		†
	0.930 to 0.946	6	3	3	4	4	4	4	4	4	4	4	4	4	4	5	5		
	0.946 to 1.000	7	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4		
	0.000 to 0.618^-	1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	15	†
	0.618^- to 0.619^-	2	2	2	2	3	4	5	6	7	8	9	10	11	12	13	14	14	†
	0.619^- to 0.619^+	2	2	2	3	3	3	4	5	6	7	8	9	10	11	12	13	13	†
	0.619^+ to 0.622	2	2	2	3	3	4	4	4	5	6	7	8	9	10	11	12	12	†
	0.622 to 0.629	2	2	2	3	3	4	4	5	5	5	6	7	8	9	10	11	11	†
	0.629 to 0.652	2	2	2	3	3	4	4	5	5	6	6	6	7	8	9	10	10	†
	0.652 to 0.755	2	2	2	3	3	4	4	5	5	6	6	7	7	7	8	9	9	†
	0.755 to 0.759	2	2	2	3	3	4	4	5	5	6	6	7	7	8	8	8	8	†
F_1*(16)	0.759 to 0.767	3	2	3	3	3	3	4	4	5	5	6	6	7	7	7	8	8	†
	0.767 to 0.782	3	2	3	3	3	4	4	4	4	5	5	6	6	7	7	7	7	†
	0.782 to 0.829	3	2	3	3	3	4	4	4	5	5	5	5	6	6	6	7	7	†
	0.829 to 0.846^+	4	3	3	3	3	3	4	4	4	5	5	5	5	6	6	6	6	†
	0.846^+ to 0.888^+	4	3	3	3	3	4	4	4	4	4	5	5	5	5	5	6	6	†
	0.888^+ to 0.922	5	3	3	3	4	4	4	4	4	4	4	5	5	5	5	5	5	†
	0.922 to 0.941	6	3	3	4	4	4	4	4	4	4	4	4	4	5	5	5	5	†
	0.941 to 0.953	7	3	4	4	4	4	4	4	4	4	4	4	4	4	4	5	5	
	0.953 to 1.000	8	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	

The exponential symbols +, − indicate only the relative magnitude of two different roots that are equal to three decimal places (i.e., $a^- < a^+$).

TABLE IV B — FORMULAE FOR $H_1(n)$ AND VALUES OF THE NEXT SAMPLE SIZE x FOR ANY H-SITUATION EXPRESSED AS FUNCTIONS OF q FOR $n = 2(1)12$ WHEN THE PROCEDURE R_1 IS USED.

The integer shown below q^y opposite $H_1(n)$ is the coefficient of q^y in the polynomial formula for $H_1(n)$.

	q-interval	x	1	q	q^2	q^3	q^4	q^5	q^6	q^7	q^8	q^9	q^{10}	q^{11}	q^{12}
$H_1(2)$	0.000 to 0.618	1	2												
	0.618 to 1.000	2	3	−1	−1										
$H_1(3)$	0.000 to 0.618	1	3												
	0.618 to 0.707	2	5	−3	−1	1									
	0.707 to 1.000	3	5	−2	−1	−1									
$H_1(4)$	0.000 to 0.618	1	4												
	0.618 to 0.707	2	7	−5	0	1	−1								
	0.707 to 0.786	2	7	−4	−1	−1	1								
	0.786 to 1.000	4	8	−4	−2	−1									
$H_1(5)$	0.000 to 0.618	1	5												
	0.618 to 0.707	2	9	−7	1	0	−1	1							
	0.707 to 0.755	2	9	−6	0	−1	1	−1							
	0.755 to 0.786	3	9	−5	−1	−2	1								
	0.786 to 0.817	3	10	−6	−2	−1	0	1							
	0.817 to 1.000	5	11	−7	−2	0	0	−1							
$H_1(6)$	0.000 to 0.618	1	6												
	0.618 to 0.707	2	11	−9	2	−1	0	1	−1						
	0.707 to 0.755	2	11	−8	1	−2	1	−1	1						
	0.755 to 0.786	3	11	−6	−2	−2	2	0	−1						
	0.786 to 0.817	3	12	−7	−3	−1	1	1	−1						
	0.817 to 0.844	3	13	−9	−2	0	0	−1	1						
	0.844 to 1.000	6	14	−10	−1	0	−1	−1							
$H_1(7)$	0.000 to 0.618	1	7												
	0.618 to 0.707	2	13	−11	3	−2	1	0	−1	1					
	0.707 to 0.755	2	13	−10	2	−3	2	−1	1	−1					
	0.755 to 0.786	3	13	−7	−3	−2	2	0	−1	1					
	0.786 to 0.817	3	14	−8	−4	0	1	0	−1						
	0.817 to 0.819	3	15	−10	−3	1	0	−2	1						
	0.819 to 0.844	4	16	−11	−3	0	0	−1	1						
	0.844 to 0.869	4	17	−13	−1	−1	−1	0	0	1					
	0.869 to 1.000	7	17	−12	−1	−1	−1	−1							
$H_1(8)$	0.000 to 0.618	1	8												
	0.618 to 0.707	2	15	−13	4	−3	2	−1	0	1	−1				
	0.707 to 0.755	2	15	−12	3	−4	3	−2	1	−1	1				
	0.755 to 0.786	3	15	−8	−4	−2	3	0	−2	1					
	0.786 to 0.812	3	16	−9	−5	0	1	0	−1	0	1				
	0.812 to 0.817	4	17	−10	−5	0	1	0	−1						
	0.817 to 0.819	4	18	−12	−4	1	0	−2	1						
	0.819 to 0.844	4	19	−14	−3	0	1	−1							
	0.844 to 0.869	4	20	−16	−1	−1	0	0	−1	1					
	0.869 to 0.885	4	20	−15	−2	−1	0	−1	0	0	1				
	0.885 to 0.899⁻	8	20	−14	−2	−1	−1	−1	1	0	−1				
	0.899⁻ to 1.000	8	21	−15	−2	−1	−1	−1							

TABLE IV B — *Continued*

	q-interval	x	1	q	q^2	q^3	q^4	q^5	q^6	q^7	q^8	q^9	q^{10}	q^{11}	q^{12}
	0.000 to 0.618	1	9												
	0.618 to 0.707	2	17	−15	5	−4	3	−2	1	0	−1	1			
	0.707 to 0.755	2	17	−14	4	−5	4	−3	2	−1	1	−1			
	0.755 to 0.786	3	17	−9	−5	−2	4	−1	−2	2	0	−1			
	0.786 to 0.812	3	18	−10	−6	0	2	−1	−1	1	1	−1			
	0.812 to 0.817	3	19	−12	−5	0	2	−1	−1	1					
	0.817 to 0.819	3	20	−14	−4	1	1	−3	1	1					
$H_1(9)$	0.819 to 0.824	3	21	−16	−3	0	2	−2	0	1					
	0.824 to 0.844	4	22	−17	−3	0	2	−2	0	1	0	−1			
	0.844 to 0.857	4	23	−19	−1	−1	1	−1	−1	2	0	−1			
	0.857 to 0.869	5	23	−19	−1	0	0	−1	−1	1					
	0.869 to 0.885	5	23	−18	−2	0	0	−2	0	0	1				
	0.885 to 0.896	5	23	−17	−3	0	−1	−1	1	−1	−1	2			
	0.896 to 0.899⁻	9	24	−17	−3	−1	−1	−1	1	0	−1				
	0.899⁻ to 1.000	9	25	−19	−2	−1	−1	−1	0	1	0	−1			
	0.000 to 0.618	1	10												
	0.618 to 0.707	2	19	−17	6	−5	4	−3	2	−1	0	1	−1		
	0.707 to 0.755	2	19	−16	5	−6	5	−4	3	−2	1	−1	1		
	0.755 to 0.786	3	19	−10	−6	−2	5	−2	−2	2	0	−1	1		
	0.786 to 0.812	3	20	−11	−7	0	3	−2	0	1	0	−1			
	0.812 to 0.817	3	21	−13	−6	0	3	−2	0	1	−1				
	0.817 to 0.819	3	22	−15	−5	1	2	−4	2	1	−1				
	0.819 to 0.824	4	24	−18	−4	0	2	−2	0	1					
$H_1(10)$	0.824 to 0.840	4	25	−20	−3	0	2	−2	0	1	0	−1	1		
	0.840 to 0.844	5	25	−20	−3	1	2	−3	0	1	0	−1			
	0.844 to 0.857	5	26	−22	−1	0	1	−2	−1	2	0	−1			
	0.857 to 0.869	5	26	−22	−1	1	−1	−1	−1	1	1	0	−1		
	0.869 to 0.885	5	26	−21	−2	1	−1	−2	0	0	2	0	−1		
	0.885 to 0.896	5	26	−20	−3	1	−2	−1	1	−1	0	2	−1		
	0.896 to 0.899⁻	5	27	−21	−3	0	−1	−1	1	0	−1	0	1		
	0.899⁻ to 0.904	6	28	−23	−1	−1	−1	−1	0	1	−1	−1	2		
	0.904 to 1.000	10	29	−23	−2	−1	−1	−1	1	1	−1	−1			
	0.000 to 0.618	1	11												
	0.618 to 0.707	2	21	−19	7	−6	5	−4	3	−2	1	0	−1	1	
	0.707 to 0.755	2	21	−18	6	−7	6	−5	4	−3	2	−1	1	−1	
	0.755 to 0.786	3	21	−11	−7	−2	6	−3	−2	3	0	−2	1		
	0.786 to 0.812	3	22	−12	−8	0	4	−3	0	1	0	−1	0	1	
	0.812 to 0.817	3	23	−14	−7	1	3	−3	0	1	−1				
	0.817 to 0.819	3	24	−16	−6	2	2	−5	2	1	−1				
	0.819 to 0.824	4	27	−21	−4	0	3	−2	−1	1					
	0.824 to 0.840	4	28	−23	−3	0	3	−2	−1	1	0	−1	1		
$H_1(11)$	0.840 to 0.844	4	28	−23	−3	1	2	−3	0	1	0	−1	0	1	
	0.844 to 0.856	4	29	−25	−1	0	1	−2	−1	2	0	−1	0	1	
	0.856 to 0.857	5	29	−25	−1	1	0	−2	−1	3	0	−2			
	0.857 to 0.869	5	29	−25	−1	2	−2	−1	−1	2	1	−1	−1		
	0.869 to 0.881	5	29	−24	−2	2	−2	−2	0	1	2	−1	−1		
	0.881 to 0.885	6	29	−24	−1	1	−2	−2	0	0	2	0	−1		
	0.885 to 0.896	6	29	−23	−2	1	−3	−1	1	−1	0	2	−1		
	0.896 to 0.899⁻	6	30	−24	−2	0	−2	−1	1	0	−1	0	1		
	0.899⁻ to 0.904	6	31	−26	0	−2	−1	−1	0	1	−1	0	2	−1	
	0.904 to 0.912	6	32	−27	−1	−1	−1	−1	1	0	−1	0	0	1	
	0.912 to 1.000	11	33	−27	−2	−1	−1	0	1	0	−1	−1			

TABLE IV B — *Continued*

	q-interval	$\varkappa$	1	q	q^2	q^3	q^4	q^5	q^6	q^7	q^8	q^9	q^{10}	q^{11}	q^{12}
	0.000 to 0.618	1	12												
	0.618 to 0.707	2	23	−21	8	−7	6	−5	4	−3	2	−1	0	1	−1
	0.707 to 0.755	2	23	−20	7	−8	7	−6	5	−4	3	−2	1	−1	1
	0.755 to 0.786	3	23	−12	−8	−2	7	−4	−2	4	−1	−2	2	0	−1
	0.786 to 0.812	3	24	−13	−9	0	5	−4	0	2	−1	−1	1	1	−1
	0.812 to 0.817	3	25	−15	−8	1	3	−3	0	2	−2	0	1		
	0.817 to 0.819	3	26	−17	−7	2	2	−5	2	2	−2	0	1		
	0.819 to 0.824	4	30	−24	−4	0	4	−3	−1	1	1	0	−1		
	0.824 to 0.840	4	31	−26	−3	0	4	−3	−1	1	1	−1			
	0.840 to 0.844	4	31	−26	−3	1	3	−4	0	1	1	−1	−1	1	
$H_1(12)$	0.844 to 0.856	4	32	−28	−1	0	2	−3	−1	2	1	−1	−1	1	
	0.856 to 0.857	4	32	−28	−1	1	0	−2	−1	3	0	−2	0	0	1
	0.857 to 0.869	4	32	−28	−1	2	−2	−1	−1	2	1	−1	−1	0	1
	0.869 to 0.881	6	32	−27	−1	2	−3	−2	0	1	2	−1	−1		
	0.881 to 0.885	6	32	−27	0	0	−2	−2	0	0	3	0	−2		
	0.885 to 0.896	6	32	−26	−1	0	−3	−1	1	−1	1	2	−2		
	0.896 to 0.899^-	6	33	−27	−1	−1	−2	−1	1						
	0.899^- to 0.904	6	34	−29	1	−3	−1	−1	0	1	0	0	1	−1	
	0.904 to 0.912^-	6	35	−30	0	−2	−1	−1	1	0	0	0	−1	1	
	0.912^- to 0.912^+	7	35	−29	−1	−2	−1	−1	1	0	−1	0	0	1	
	0.912^+ to 0.919	7	36	−30	−2	−1	−1	0	0	0	−1	−1	1	0	1
	0.919 to 1.000	12	37	−31	−2	−1	0	0	0	0	−1	−1			

The exponential symbols +, − indicate only the relative magnitude of two different roots that are equal to three decimal places (i.e., $a^- < a^+$).

TABLE IV C — FORMULAE FOR $G_1^*(m, n)$ AND THE VALUES OF THE NEXT SAMPLE SIZE x FOR CERTAIN G-SITUATIONS EXPRESSED AS FUNCTIONS OF q FOR $n = 2(1)12$ AND $q \geqq 0.850$ WHEN PROCEDURE R_1 IS USED

The integer shown below q^y opposite $G_1^*(m,n)$ is the coefficient of q^y in the polynomial formula for $G_1^*(m,n)$.

n	m	q-interval	x	$G_1^*(m,n)$											
				1	q	q^2	q^3	q^4	q^5	q^6	q^7	q^8	q^9	q^{10}	q^{11}
2	2	0.850 to 1.000	1	2	1										
3	2	0.850 to 1.000	1	4	1	−1									
	3	0.850 to 1.000	1	4	2	1									
4	2	0.850 to 1.000	1	6	2	−2	−2								
	4	0.850 to 1.000	2	7	3	1									
5	2	0.850 to 1.000	1	9	2	−4	−2	−1							
	3	0.850 to 1.000	1	9	3	1	−3	−2							
	5	0.850 to 1.000	2	10	3	1	1	1							
6	2	0.850 to 1.000	1	12	2	−6	−2	−1	−1						
	3	0.850 to 1.000	1	12	3	1	−4	−2	−2						
	6	0.850 to 1.000	2	13	3	2	2	1							
7	2	0.850 to 1.000	1	15	2	−8	−2	−1	−1	−1					
	3	0.850 to 1.000	1	15	3	2	−6	−3	−2	−1					
	4	0.850 to 1.000	2	16	3	2	1	−5	−3	−2					
	7	0.850 to 0.857	2	16	3	3	2	1	1	1					
		0.857 to 1.000	3	16	4	3	2	1							
8	2	0.850 to 0.869	1	18	2	−11	−2	−1	−1	−1	1				
		0.869 to 1.000	1	18	3	−11	−2	−1	−2	−1					
	3	0.850 to 0.869	1	18	3	2	−9	−3	−1	−1					
		0.869 to 1.000	1	18	4	2	−9	−3	−2	−1	−1				
	4	0.850 to 0.869	2	19	3	2	1	−7	−3	−2					
		0.869 to 1.000	2	19	4	2	1	−7	−4	−2	−1				
	8	0.850 to 0.869	3	19	4	3	2	1	1	1	2				
		0.869 to 0.899⁻	3	19	5	3	2	1	0	1	1				
		0.899⁻ to 1.000	4	20	5	3	2	1							
9	2	0.850 to 0.869	1	21	2	−14	−2	−1	−1	−1	1	1			
		0.869 to 0.885	1	21	3	−14	−2	−1	−2	−1	0	1			
		0.885 to 0.899⁻	1	21	4	−14	−2	−2	−2	0	0	−1			
		0.899⁻ to 1.000	1	22	3	−14	−2	−2	−2	−1					
	3	0.850 to 0.869	1	21	3	2	−12	−2	−1	−2	0	1			
		0.869 to 0.885	1	21	4	2	−12	−2	−2	−2	−1	1			
		0.885 to 0.899⁻	1	21	5	2	−12	−3	−2	−1	−1	−1			
		0.899⁻ to 1.000	1	22	4	2	−12	−3	−2	−2	−1				
	4	0.850 to 0.869	2	22	3	2	1	−9	−3	−2					
		0.869 to 0.885	2	22	4	2	1	−9	−4	−2	−1				
		0.885 to 0.899⁻	2	22	5	2	1	−10	−4	−1	−1	−2			
		0.899⁻ to 1.000	2	23	4	2	1	−10	−4	−2	−1	−1			

TABLE IV C — *Continued*

n	m	q-interval	x	$G_1^*(m,n)$											
				1	q	q^2	q^3	q^4	q^5	q^6	q^7	q^8	q^9	q^{10}	q^{11}
9	5	0.850 to 0.869	2	22	3	2	2	2	−7	−4	−1				
		0.869 to 0.885	2	22	4	2	2	2	−8	−4	−2				
		0.885 to 0.899⁻	2	22	5	2	2	1	−8	−3	−2	−2			
		0.899⁻ to 1.000	2	23	4	2	2	1	−8	−4	−2	−1			
	9	0.850 to 0.869	3	22	4	3	2	2	2	1	2	2			
		0.869 to 0.881	3	22	5	3	2	2	1	1	1	2			
		0.881 to 0.885	4	23	5	3	2	2	0	0	1	2			
		0.885 to 0.899⁻	4	23	6	3	2	1	0	1	1				
		0.899⁻ to 1.000	4	24	5	3	2	1	0	0	1	1			
10	2	0.850 to 0.857	1	24	2	−17	−2	0	−1	−1	1	1	−1		
		0.857 to 0.869	1	24	2	−17	−1	−1	−1	−1	0	1			
		0.869 to 0.885	1	24	3	−17	−2	−1	−2	−1	0	1	1		
		0.885 to 0.896	1	24	4	−17	−2	−2	−2	0	0	−1	1		
		0.896 to 0.899⁻	1	25	4	−17	−3	−2	−2	0	1	−1	−1		
		0.899⁻ to 1.000	1	26	3	−17	−3	−2	−2	−1	1	0	−1		
	3	0.850 to 0.857	1	24	3	2	−15	−1	−2	−2	1	1			
		0.857 to 0.869	1	24	3	2	−14	−2	−2	−2	0	1	1		
		0.869 to 0.885	1	24	4	2	−14	−2	−3	−2	−1	1	1		
		0.885 to 0.896	1	24	5	2	−14	−3	−3	−1	−1	−1	1		
		0.896 to 0.899⁻	1	25	5	2	−15	−3	−3	−1	0	−1	−1		
		0.899⁻ to 1.000	1	26	4	2	−15	−3	−3	−2	0	0	−1		
	4	0.850 to 0.857	2	25	3	2	1	−11	−3	−2					
		0.857 to 0.869	2	25	3	2	2	−12	−3	−2	−1	0	1		
		0.869 to 0.885	2	25	4	2	2	−12	−4	−2	−2	0	1		
		0.885 to 0.896	2	25	5	2	2	−13	−4	−1	−2	−2	1		
		0.896 to 0.899⁻	2	26	5	2	1	−13	−4	−1	−1	−2	−1		
		0.899⁻ to 1.000	2	27	4	2	1	−13	−4	−2	−1	−1	−1		
	5	0.850 to 0.857	2	25	3	2	2	3	−10	−4	0	0	−1		
		0.857 to 0.869	2	25	3	2	3	2	−10	−4	−1				
		0.869 to 0.885	2	25	4	2	3	2	−11	−4	−2				
		0.885 to 0.896	2	25	5	2	3	1	−11	−3	−2	−2			
		0.896 to 0.899⁻	2	26	5	2	2	1	−11	−3	−1	−2	−2		
		0.899⁻ to 1.000	2	27	4	2	2	1	−11	−4	−1	−1	−2		
	6	0.850 to 0.857	2	25	3	3	2	3	1	−8	−2	−1	−1		
		0.857 to 0.869	2	25	3	3	3	2	1	−8	−3	−1			
		0.869 to 0.885	2	25	4	3	3	2	0	−8	−4	−1			
		0.885 to 0.896	2	25	5	3	3	1	0	−7	−4	−3			
		0.896 to 0.899⁻	2	26	5	3	2	1	0	−7	−3	−3	−2		
		0.899⁻ to 1.000	2	27	4	3	2	1	0	−8	−3	−2	−2		
	10	0.850 to 0.857	3	25	4	3	2	3	2	1	3	3	2		
		0.857 to 0.869	4	26	4	3	3	2	1	1	2	2	2		
		0.869 to 0.885	4	26	5	3	3	2	0	1	1	2	2		
		0.885 to 0.896	4	26	6	3	3	1	0	2	1	0	2		
		0.896 to 0.899⁻	4	27	6	3	2	1	0	2	2				
		0.899⁻ to 1.000	4	28	5	3	2	1	0	1	2	1			
11	2	0.850 to 0.857	1	27	2	−20	−1	0	−1	−2	1	2	−1	−1	
		0.857 to 0.869	1	27	2	−20	0	−1	−1	−2	0	2	0	−1	
		0.869 to 0.885	1	27	3	−20	−1	−1	−2	−2	0	2	1	−1	
		0.885 to 0.896	1	27	4	−20	−2	−2	−2	0	0	−1	1	1	
		0.896 to 0.899⁻	1	28	4	−20	−3	−2	−2	0	1	−1	−1	1	
		0.899⁻ to 0.904	1	29	3	−20	−3	−2	−2	−1	1	0	−1	1	
		0.904 to 1.000	1	30	3	−21	−3	−2	−2	0	1	0	−1	−1	

TABLE IV C — *Continued*

n	m	q-interval	x	$G_3^*(m,n)$											
				1	q	q^2	q^3	q^4	q^5	q^6	q^7	q^8	q^9	q^{10}	q^{11}
11		0.850 to 0.857	1	27	3	2	−17	−1	−2	−2	1	1	0	−1	
		0.857 to 0.869	1	27	3	2	−16	−2	−2	−2	0	1	1	−1	
		0.869 to 0.885	1	27	4	2	−16	−3	−3	−2	−1	2	1		
		0.885 to 0.896	1	27	5	2	−16	−4	−3	−1	−1	0	1		
	3	0.896 to 0.899⁻	1	28	5	2	−17	−4	−3	−1	0	0	−1		
		0.899⁻ to 0.904	1	29	4	3	−18	−4	−3	−2	0	0	−1	1	
		0.904 to 1.000	1	30	4	2	−18	−4	−3	−1	0	0	−1	−1	
		0.850 to 0.857	2	28	3	2	2	−14	−3	−3	0	1			
		0.857 to 0.869	2	28	3	2	3	−15	−3	−3	−1	1	1		
		0.869 to 0.885	2	28	4	2	3	−15	−4	−3	−2	1	1		
	4	0.885 to 0.896	2	28	5	2	3	−16	−4	−2	−2	−1	1		
		0.896 to 0.899⁻	2	29	5	2	2	−16	−4	−2	−1	−1	−1		
		0.899⁻ to 0.904	2	30	4	3	1	−16	−4	−3	−1	−1	−1	1	
		0.904 to 1.000	2	31	4	2	1	−16	−4	−2	−1	−1	−1	−1	
		0.850 to 0.857	2	28	3	2	3	3	−13	−4	0	0	−1		
		0.857 to 0.869	2	28	3	2	4	2	−13	−4	−1				
		0.869 to 0.885	2	28	4	2	4	2	−14	−4	−2				
	5	0.885 to 0.896	2	28	5	2	4	1	−14	−3	−2	−2			
		0.896 to 0.899⁻	2	29	5	2	3	1	−14	−3	−1	−1	−2	−2	
		0.899⁻ to 0.904	2	30	4	2	3	1	−14	−4	−1	−1	−2		
		0.904 to 1.000	2	31	4	1	3	1	−14	−3	−1	−1	−2	−2	
		0.850 to 0.857	2	28	3	3	3	3	1	−11	−2	0	−1	−1	
		0.857 to 0.869	2	28	3	3	4	2	1	−11	−3	0	0	−1	
		0.869 to 0.885	2	28	4	3	4	2	0	−11	−4	0	0	−1	
	6	0.885 to 0.896	2	28	5	3	4	1	0	−10	−4	−2	0	−1	
		0.896 to 0.899⁻	2	29	5	3	3	1	0	−10	−3	−2	−2	−1	
		0.899⁻ to 0.904	2	30	4	4	2	1	0	−11	−3	−2	2		
		0.904 to 1.000	2	31	4	3	2	1	0	−10	−3	−2	−2	−2	
		0.850 to 0.857	3	28	4	3	3	4	2	1	3	3	2	2	
		0.857 to 0.869	4	29	4	3	4	2	2	1	2	3	2	1	
		0.869 to 0.885	4	29	5	3	4	2	1	1	1	3	2	1	
	11	0.885 to 0.896	4	29	6	3	4	1	1	2	1	1	2	1	
		0.896 to 0.899⁻	4	30	6	3	3	1	1	2	2	1	0	1	
		0.899⁻ to 0.904	4	31	5	4	2	1	1	2	2	1	0	2	
		0.904 to 1.000	4	32	5	3	2	1	1	2	2	1			
12		0.850 to 0.856	1	30	2	−23	−1	1	−1	−3	1	2	−1	−1	1
		0.856 to 0.857	1	30	2	−23	0	0	−1	−3	2	2	2	−2	−1
		0.857 to 0.869	1	30	2	−23	1	−1	−2	−2	1	2	0	−1	−1
		0.869 to 0.881	1	30	3	−23	0	−1	−3	−2	1	2	1	−1	−1
	2	0.881 to 0.885	1	30	3	−22	−1	−1	−3	−2	0	2	2	−1	−1
		0.885 to 0.896	1	30	4	−22	−2	−2	−3	0	0	−1	2	1	−1
		0.896 to 0.899⁻	1	31	4	−23	−3	−2	−2	0	1	−1	−1	1	1
		0.899⁻ to 0.904	1	32	3	−23	−3	−2	−2	−1	1	0	−1	1	1
		0.904 to 0.912⁻	1	33	3	−24	−3	−2	−2	0	1	0	−1	−1	1
		0.912⁻ to 1.000	1	34	3	−25	−3	−2	−1	0	1	0	−2	−1	
		0.850 to 0.856	1	30	3	2	−20	0	−2	−2	0	1	1	−1	
		0.856 to 0.857	1	30	3	2	−19	−1	−2	−2	1	1	0	−1	−1
		0.857 to 0.869	1	30	3	2	−18	−2	−2	−2	0	1	1	−1	−1
		0.869 to 0.881	1	30	4	2	−18	−3	−3	−2	−1	2	1	0	−1
	3	0.881 to 0.885	1	30	4	3	−19	−3	−3	−2	−2	2	2	0	−1
		0.885 to 0.896	1	30	5	3	−19	−5	−3	−1	−1	0	1	0	1
		0.896 to 0.899⁻	1	31	5	3	−20	−5	−3	−1	0	0	−1	0	1
		0.899⁻ to 0.904	1	32	4	4	−22	−4	−3	−2	0	0	0	1	
		0.904 to 0.912⁻	1	33	4	3	−22	−4	−3	−1	0	0	0	−1	
		0.912⁻ to 1.000	1	34	4	2	−22	−4	−2	−1	0	0	−1	−1	−1

TABLE IV C — *Continued*

n	m	q-interval	x	$G_1^*(m,n)$											
				1	q	q^2	q^3	q^4	q^5	q^6	q^7	q^8	q^9	q^{10}	q^{11}
12		0.850 to 0.856	2	31	3	2	2	−16	−3	−3	0	1			
		0.856 to 0.857	2	31	3	2	3	−17	−3	−3	1	1	−1	0	−1
		0.857 to 0.869	2	31	3	2	4	−18	−3	−3	0	1	0	0	−1
		0.869 to 0.881	2	31	4	2	4	−18	−5	−3	−1	1	1		
	4	0.881 to 0.885	2	31	4	3	3	−18	−5	−3	−2	1	2		
		0.885 to 0.896	2	31	5	3	3	−19	−5	−2	−2	−1	2		
		0.896 to 0.899⁻	2	32	5	3	2	−19	−5	−2	−1	−1			
		0.899⁻ to 0.904	2	33	4	4	1	−19	−5	−3	−1	−1	0	1	
		0.904 to 0.912⁻	2	34	4	3	1	−19	−5	−2	−1	−1	0	−1	
		0.912⁻ to 1.000	2	35	4	2	1	−19	−4	−2	−1	−1	−1	−1	−1
		0.850 to 0.856	2	31	3	3	3	4	1	−14	−2	0	−1	−1	1
		0.856 to 0.857	2	31	3	3	4	3	1	−14	−1	0	−2	−1	
		0.857 to 0.869	2	31	3	3	5	2	1	−14	−2	0	−1	−1	
		0.869 to 0.881	2	31	4	3	5	2	0	−14	−3	0	−1	−1	
		0.881 to 0.885	2	31	4	4	4	2	0	−14	−4	0	0	−1	
	6	0.885 to 0.896	2	31	5	4	4	1	0	−13	−4	−2	0	−1	
		0.896 to 0.899⁻	2	32	5	4	3	1	0	−13	−3	−2	−2	−1	
		0.899⁻ to 0.904	2	33	4	5	2	1	0	−14	−3	−2	−2		
		0.904 to 0.912⁻	2	34	4	4	2	1	0	−13	−3	−2	−2	−2	
		0.912⁻ to 1.000	2	35	4	3	2	1	1	−13	−3	−2	−3	−2	−1
		0.850 to 0.856	2	31	3	3	3	4	2	1	−9	−2	−1	−1	
		0.856 to 0.857	2	31	3	3	4	3	2	1	−8	−2	−2	−1	−1
		0.857 to 0.869	3	31	4	3	5	2	1	0	−9	−2	−1	−1	−1
		0.869 to 0.881	3	31	5	3	5	2	0	0	−10	−2	−1	−1	−1
	7	0.881 to 0.885	3	31	5	4	4	2	0	0	−11	−2	0	−1	−1
		0.885 to 0.896	3	31	6	4	4	1	0	1	−11	−4	0	−1	−1
		0.896 to 0.899⁻	3	32	6	4	3	1	0	1	−10	−4	−2	−1	−1
		0.899⁻ to 0.904	3	33	5	5	2	1	0	0	−10	−4	−2	0	−1
		0.904 to 0.912⁻	3	34	5	4	2	1	0	1	−10	−4	−2	−2	−1
		0.912⁻ to 1.000	3	35	5	3	2	1	1	1	−10	−4	−3	−2	−2
		0.850 to 0.856	4	32	4	3	3	4	2	1	3	3	2	2	3
		0.856 to 0.857	4	32	4	3	4	3	2	1	4	3	1	2	2
		0.857 to 0.869	4	32	4	3	5	2	2	1	3	3	2	2	2
		0.869 to 0.881	4	32	5	3	5	2	1	1	2	3	2	2	2
	12	0.881 to 0.885	4	32	5	4	4	2	1	1	1	3	3	2	2
		0.885 to 0.896	4	32	6	4	4	1	1	2	1	1	3	2	2
		0.896 to 0.899⁻	4	33	6	4	3	1	1	2	2	1	1	2	2
		0.899⁻ to 0.904	4	34	5	5	2	2	1	1	2	1	1	2	1
		0.904 to 0.912⁻	4	35	5	4	2	2	1	2	2	1	1	0	1
		0.912⁻ to 1.000	4	36	5	3	2	2	2	2	2	1			

The exponential symbols +, − indicate only the relative magnitude of two different roots that are equal to three decimal places (i.e., $a^- < a^+$).

TABLE V A — EXPECTED NUMBER OF TESTS REQUIRED AND SIZE x OF THE NEXT SAMPLE TO BE TAKEN FOR ANY $G_1(m,n)$ AND ANY $H_1(n)$ SITUATION WHEN PROCEDURE R_1 IS USED AND $q = 0.90$

n	$H_1(n)$	x	$G_1(m,n)$							
			$m = 2$ $x = 1$	$m = 3$ $x = 1$	$m = 4$ $x = 2$	$m = 5$ $x = 2$	$m = 6$ $x = 2$	$m = 7$ $x = 3$	$m = 8$ $x = 4$	$m = 9$ $x = 4$
1	1.000	1								
2	1.290	2	1.526							
3	1.661	3	2.153	2.439						
4	2.051	4	2.485	2.971	3.056					
5	2.490	5	2.866	3.325	3.547	3.637				
6	2.943	6	3.282	3.727	3.925	4.115	4.148			
7	3.414	7	3.728	4.157	4.343	4.512	4.621	4.628		
8	3.904	8	4.191	4.613	4.784	4.943	5.033	5.100	5.100	
9	4.395	9	4.672	5.085	5.250	5.396	5.476	5.528	5.575	5.543
10	4.872	6	5.163	5.570	5.728	5.868	5.938	5.980	6.014	6.021
11	5.327	6	5.646	6.056	6.210	6.346	6.411	6.445	6.470	6.466
12	5.790	6	6.111	6.528	6.687	6.821	6.883	6.914	6.933	6.922
13	6.261	7	6.570	6.993	7.157	7.294	7.356	7.385	7.401	7.385
14	6.732	7	7.037	7.456	7.623	7.765	7.829	7.857	7.872	7.853
15	7.213	7	7.509	7.925	8.089	8.232	8.300	8.330	8.345	8.324
16	7.695	7	7.985	8.400	8.561	8.702	8.770	8.802	8.819	8.798
17	8.161	6	8.467	8.878	9.038	9.176	9.242	9.275	9.293	9.273
18	8.629	6	8.940	9.354	9.513	9.651	9.715	9.746	9.764	9.746
19	9.100	7	9.407	9.825	9.986	10.124	10.188	10.218	10.235	10.217
20	9.572	7	9.877	10.294	10.458	10.596	10.660	10.690	10.706	10.687
21	10.044	7	10.348	10.764	10.927	11.068	11.133	11.163	11.179	11.159
22	10.520	7	10.820	11.236	11.398	11.538	11.605	11.635	11.651	11.632
23	10.996	6	11.295	11.709	11.871	12.010	12.076	12.108	12.124	12.105
24	11.466	6	11.770	12.184	12.345	12.484	12.549	12.580	12.597	12.578
25	11.937	7	12.243	12.658	12.819	12.957	13.022	13.052	13.069	13.051
26	12.408	7	12.714	13.130	13.292	13.430	13.495	13.525	13.541	13.522
27	12.881	7	13.185	13.601	13.763	13.902	13.967	13.997	14.014	13.994
28	13.353	7	13.657	14.073	14.235	14.374	14.440	14.470	14.486	14.467
29	13.827	7	14.129	14.545	14.707	14.846	14.912	14.943	14.959	14.939
30	14.301	7	14.603	15.018	15.179	15.319	15.384	15.415	15.431	15.412
31	14.773	7	15.077	15.491	15.653	15.792	15.857	15.888	15.904	15.885
32	15.244	7	15.549	15.965	16.126	16.264	16.329	16.360	16.377	16.358
33	15.717	7	16.021	16.437	16.598	16.737	16.802	16.833	16.849	16.830
34	16.189	7	16.493	16.909	17.071	17.210	17.275	17.305	17.321	17.302
35	16.661	7	16.965	17.381	17.543	17.682	17.747	17.778	17.794	17.775
36	17.135	7	17.438	17.853	18.015	18.154	18.219	18.250	18.266	18.247
37	17.608	7	17.910	18.326	18.487	18.626	18.692	18.723	18.739	18.720
38	18.080	7	18.384	18.799	18.960	19.099	19.164	19.195	19.212	19.192
39	18.552	7	18.856	19.272	19.433	19.572	19.637	19.668	19.684	19.665
40	19.024	7	19.328	19.744	19.906	20.044	20.110	20.140	20.157	20.137
41	19.497	7	19.801	20.216	20.378	20.517	20.582	20.613	20.629	20.610
42	19.969	7	20.273	20.688	20.850	20.989	21.055	21.085	21.102	21.082
43	20.442	7	20.745	21.161	21.323	21.462	21.527	21.558	21.574	21.555
44	20.915	7	21.218	21.634	21.795	21.934	21.999	22.030	22.047	22.027
45	21.387	7	21.691	22.106	22.268	22.407	22.472	22.503	22.519	22.500
46	21.860	7	22.164	22.579	22.740	22.879	22.945	22.975	22.992	22.972
47	22.332	7	22.636	23.051	23.213	23.352	23.417	23.448	23.464	23.445
48	22.805	7	23.108	23.524	23.685	23.825	23.890	23.920	23.937	23.917
49	23.277	7	23.581	23.996	24.158	24.297	24.362	24.393	24.409	24.390
50	23.750	7	24.053	24.469	24.630	24.769	24.835	24.865	24.882	24.862

TABLE V A — *Continued*

n	$H_1(n)$	x	$G_1(m,n)$							
			$m=2$ $x=1$	$m=3$ $x=1$	$m=4$ $x=2$	$m=5$ $x=2$	$m=6$ $x=2$	$m=7$ $x=3$	$m=8$ $x=4$	$m=9$ $x=4$
51	24.222	7	24.526	24.941	25.103	25.242	25.307	25.338	25.354	25.335
52	24.695	7	24.998	25.414	25.575	25.714	25.780	25.810	25.827	25.807
53	25.167	7	25.471	25.886	26.048	26.187	26.252	26.283	26.299	26.280
54	25.640	7	25.943	26.359	26.521	26.660	26.725	26.755	26.772	26.752
55	26.112	7	26.416	26.831	26.993	27.132	27.197	27.228	27.244	27.225
56	26.585	7	26.888	27.304	27.465	27.605	27.670	27.700	27.717	27.697
57	27.057	7	27.361	27.776	27.938	28.077	28.142	28.173	28.189	28.170
58	27.530	7	27.833	28.249	28.410	28.550	28.615	28.645	28.662	28.643
59	28.002	7	28.306	28.721	28.883	29.022	29.087	29.118	29.134	29.115
60	28.475	7	28.779	29.194	29.356	29.495	29.560	29.590	29.607	29.588
61	28.947	7	29.251	29.666	29.828	29.967	30.032	30.063	30.079	30.060
62	29.420	7	29.723	30.139	30.301	30.440	30.505	30.535	30.552	30.533
63	29.892	7	30.196	30.611	30.773	30.912	30.977	31.008	31.024	31.005
64	30.365	7	30.668	31.084	31.246	31.385	31.450	31.480	31.497	31.478
65	30.837	7	31.141	31.556	31.718	31.857	31.922	31.953	31.969	31.950
66	31.310	7	31.614	32.029	32.191	32.330	32.395	32.425	32.442	32.423
67	31.782	7	32.086	32.502	32.663	32.802	32.867	32.898	32.914	32.895
68	32.255	7	32.559	32.974	33.136	33.275	33.340	33.370	33.387	33.368
69	32.727	7	33.031	33.446	33.608	33.747	33.812	33.843	33.859	33.840
70	33.200	7	33.504	33.919	34.081	34.220	34.285	34.316	34.332	34.313
71	33.672	7	33.976	34.392	34.553	34.692	34.757	34.788	34.804	34.785
72	34.145	7	34.449	34.864	35.026	35.165	35.230	35.261	35.277	35.258
73	34.617	7	34.921	35.337	35.498	35.637	35.702	35.733	35.749	35.730
74	35.090	7	35.394	35.809	35.971	36.110	36.175	36.206	36.222	36.203
75	35.562	7	35.866	36.282	36.443	36.582	36.647	36.678	36.694	36.675
76	36.035	7	36.339	36.754	36.916	37.055	37.120	37.151	37.167	37.148
77	36.507	7	36.811	37.227	37.388	37.527	37.592	37.623	37.639	37.620
78	36.980	7	37.284	37.699	37.861	38.000	38.065	38.096	38.112	38.093
79	37.453	7	37.756	38.172	38.333	38.472	38.537	38.568	38.584	38.565
80	37.925	7	38.229	38.644	38.806	38.945	39.010	39.041	39.057	39.038
81	38.398	7	38.701	39.117	39.278	39.417	39.482	39.513	39.530	39.510
82	38.870	7	39.174	39.589	39.751	39.890	39.955	39.986	40.002	39.983
83	39.343	7	39.646	40.062	40.223	40.362	40.427	40.458	40.475	40.455
84	39.815	7	40.119	40.534	40.696	40.835	40.900	40.931	40.947	40.928
85	40.288	7	40.591	41.007	41.168	41.307	41.373	41.403	41.420	41.400
86	40.760	7	41.064	41.479	41.641	41.780	41.845	41.876	41.892	41.873
87	41.233	7	41.536	41.952	42.113	42.252	42.318	42.348	42.365	42.345
88	41.705	7	42.009	42.424	42.586	42.725	42.790	42.821	42.837	42.818
89	42.178	7	42.481	42.897	43.058	43.197	43.263	43.293	43.310	43.290
90	42.650	7	42.954	43.369	43.531	43.670	43.735	43.766	43.782	43.763
91	43.123	7	43.426	43.842	44.003	44.142	44.208	44.238	44.255	44.235
92	43.595	7	43.899	44.314	44.476	44.615	44.680	44.711	44.727	44.708
93	44.068	7	44.371	44.787	44.948	45.087	45.153	45.183	45.200	45.180
94	44.540	7	44.844	45.259	45.421	45.560	45.625	45.656	45.672	45.653
95	45.013	7	45.316	45.732	45.893	46.032	46.098	46.128	46.145	46.125
96	45.485	7	45.789	46.204	46.366	46.505	46.570	46.601	46.617	46.598
97	45.958	7	46.261	46.677	46.838	46.978	47.043	47.073	47.090	47.070
98	46.430	7	46.734	47.149	47.311	47.450	47.515	47.546	47.562	47.543
99	46.903	7	47.206	47.622	47.784	47.923	47.988	48.018	48.035	48.016
100	47.375	7	47.679	48.094	48.256	48.395	48.460	48.491	48.507	48.488

TABLE V B—EXPECTED NUMBER OF TESTS REQUIRED AND SIZE x OF THE NEXT SAMPLE TO BE TAKEN FOR ANY $G(m,n)$ AND ANY $H_1(n)$ SITUATION WHEN PROCEDURE R_1 IS USED AND $q = 0.95$

n	$H_1(n)$	x	$G_1(m,n)$																
			$m=2$ $x=1$	$m=3$ $x=1$	$m=4$ $x=2$	$m=5$ $x=2$	$m=6$ $x=2$	$m=7$ $x=3$	$m=8$ $x=4$	$m=9$ $x=4$	$m=10$ $x=4$	$m=11$ $x=4$	$m=12$ $x=4$	$m=13$ $x=5$	$m=14$ $x=6$	$m=15$ $x=7$	$m=16$ $x=7$	$m=17$ $x=8$	$m=18$ $x=8$	$m=19$ $x=8$
1	1.000	1																	
2	1.148	2	1.513																
3	1.340	3	2.076	2.385															
4	1.538	4	2.246	2.818	2.898														
5	1.771	5	2.441	2.998	3.268	3.409													
6	2.009	6	2.657	3.206	3.462	3.749	3.810												
7	2.252	7	2.893	3.430	3.679	3.953	4.130	4.150											
8	2.499	8	3.134	3.668	3.908	4.175	4.341	4.457	4.455										
9	2.767	9	3.379	3.911	4.148	4.408	4.568	4.675	4.753	4.780									
10	3.039	10	3.637	4.164	4.398	4.655	4.808	4.908	4.978	5.074	5.082								
11	3.315	11	3.907	4.427	4.656	4.909	5.059	5.153	5.218	5.306	5.373	5.368							
12	3.594	12	4.180	4.699	4.923	5.171	5.318	5.409	5.467	5.550	5.610	5.657	5.643						
13	3.878	13	4.458	4.975	5.196	5.440	5.583	5.671	5.726	5.804	5.859	5.899	5.932	5.913					
14	4.166	14	4.739	5.254	5.474	5.716	5.855	5.939	5.991	6.066	6.116	6.152	6.178	6.201	6.179				
15	4.458	15	5.025	5.538	5.756	5.997	6.133	6.214	6.262	6.334	6.382	6.413	6.435	6.451	6.467	6.443			
16	4.753	16	5.316	5.826	6.042	6.281	6.416	6.494	6.540	6.608	6.653	6.681	6.699	6.712	6.721	6.732	6.704		
17	5.051	17	5.609	6.118	6.332	6.569	6.702	6.779	6.822	6.888	6.930	6.955	6.971	6.979	6.985	6.990	6.993	6.961	
18	5.348	18	5.906	6.414	6.626	6.861	6.992	7.067	7.109	7.173	7.211	7.234	7.247	7.253	7.256	7.257	7.255	7.251	7.214
19	5.648	19	6.203	6.710	6.922	7.155	7.285	7.358	7.398	7.461	7.498	7.518	7.528	7.532	7.532	7.530	7.524	7.516	7.504	7.465
20	5.940	11	6.502	7.009	7.219	7.452	7.580	7.652	7.691	7.752	7.787	7.806	7.814	7.815	7.813	7.809	7.800	7.788	7.772	7.757
21	6.220	12	6.798	7.305	7.516	7.748	7.876	7.947	7.985	8.044	8.078	8.096	8.102	8.101	8.097	8.090	8.079	8.065	8.046	8.026
22	6.499	12	7.083	7.595	7.808	8.041	8.169	8.240	8.277	8.336	8.369	8.386	8.391	8.389	8.382	8.374	8.361	8.345	8.323	8.301
23	6.780	12	7.363	7.879	8.095	8.330	8.460	8.531	8.568	8.627	8.659	8.675	8.679	8.677	8.669	8.659	8.644	8.626	8.603	8.578
24	7.064	13	7.643	8.159	8.378	8.616	8.747	8.820	8.858	8.916	8.949	8.964	8.968	8.964	8.956	8.945	8.929	8.909	8.884	8.858
25	7.348	13	7.926	8.440	8.659	8.899	9.033	9.107	9.146	9.205	9.238	9.253	9.257	9.252	9.243	9.232	9.215	9.194	9.167	9.139
26	7.633	13	8.210	8.723	8.941	9.181	9.316	9.392	9.432	9.492	9.526	9.542	9.545	9.541	9.531	9.519	9.501	9.480	9.451	9.423
27	7.921	14	8.494	9.008	9.225	9.464	9.598	9.676	9.718	9.779	9.813	9.829	9.833	9.829	9.819	9.806	9.788	9.766	9.737	9.707
28	8.210	14	8.781	9.293	9.510	9.748	9.882	9.959	10.002	10.064	10.100	10.117	10.121	10.117	10.107	10.094	10.076	10.053	10.024	9.993
29	8.500	14	9.069	9.581	9.796	10.034	10.167	10.243	10.286	10.349	10.386	10.403	10.408	10.405	10.395	10.382	10.364	10.341	10.311	10.280
30	8.791	14	9.358	9.870	10.085	10.321	10.454	10.530	10.571	10.634	10.671	10.690	10.695	10.692	10.683	10.671	10.652	10.629	10.599	10.567

TABLE V B — *Continued*

n	$H_1(n)$	x	$G_1(m,n)$																		
			$m=2$ $x=1$	$m=3$ $x=1$	$m=4$ $x=2$	$m=5$ $x=2$	$m=6$ $x=2$	$m=7$ $x=3$	$m=8$ $x=4$	$m=9$ $x=4$	$m=10$ $x=4$	$m=11$ $x=4$	$m=12$ $x=4$	$m=13$ $x=5$	$m=14$ $x=6$	$m=15$ $x=7$	$m=16$ $x=7$	$m=17$ $x=8$	$m=18$ $x=8$	$m=19$ $x=8$
31	9.084	15	9.649	10.160	10.374	10.610	10.742	10.817	10.858	10.920	10.957	10.976	10.982	10.980	10.971	10.959	10.940	10.917	10.887	10.855	
32	9.377	15	9.941	10.451	10.665	10.901	11.032	11.106	11.146	11.208	11.244	11.263	11.269	11.267	11.259	11.248	11.229	11.206	11.176	11.144	
33	9.669	15	10.234	10.743	10.957	11.192	11.323	11.396	11.436	11.497	11.532	11.550	11.557	11.555	11.547	11.536	11.518	11.495	11.465	11.433	
34	9.956	13	10.527	11.036	11.249	11.484	11.614	11.687	11.726	11.787	11.822	11.839	11.845	11.843	11.835	11.825	11.807	11.785	11.755	11.723	
35	10.240	13	10.816	11.327	11.540	11.775	11.905	11.978	12.017	12.077	12.111	12.128	12.133	12.131	12.123	12.112	12.095	12.073	12.044	12.012	
36	10.524	13	11.102	11.614	11.829	12.064	12.195	12.268	12.306	12.366	12.400	12.417	12.422	12.419	12.411	12.400	12.383	12.361	12.332	12.300	
37	10.810	13	11.386	11.899	12.116	12.352	12.483	12.556	12.595	12.655	12.689	12.706	12.710	12.707	12.698	12.687	12.670	12.648	12.620	12.588	
38	11.096	13	11.671	12.184	12.401	12.639	12.771	12.845	12.884	12.944	12.978	12.994	12.999	12.995	12.986	12.975	12.957	12.935	12.907	12.876	
39	11.383	13	11.957	12.469	12.686	12.924	13.057	13.132	13.172	13.232	13.266	13.282	13.287	13.283	13.274	13.262	13.245	13.223	13.194	13.163	
40	11.671	14	12.243	12.756	12.972	13.209	13.343	13.418	13.459	13.519	13.554	13.570	13.575	13.571	13.562	13.550	13.532	13.510	13.481	13.450	
41	11.959	14	12.531	13.043	13.258	13.496	13.629	13.704	13.745	13.807	13.842	13.858	13.863	13.859	13.850	13.838	13.820	13.798	13.768	13.737	
42	12.247	14	12.819	13.330	13.546	13.783	13.915	13.990	14.032	14.094	14.129	14.146	14.151	14.147	14.138	14.126	14.108	14.085	14.056	14.024	
43	12.536	14	13.107	13.618	13.833	14.070	14.203	14.277	14.318	14.380	14.416	14.434	14.439	14.435	14.426	14.414	14.396	14.373	14.344	14.312	
44	12.826	14	13.395	13.907	14.122	14.358	14.490	14.565	14.605	14.667	14.703	14.721	14.726	14.723	14.714	14.702	14.684	14.661	14.632	14.600	
45	13.116	14	13.685	14.196	14.411	14.647	14.779	14.853	14.893	14.955	14.990	15.008	15.014	15.011	15.003	14.991	14.972	14.950	14.920	14.888	
46	13.406	14	13.975	14.485	14.700	14.936	15.068	15.142	15.182	15.243	15.278	15.296	15.301	15.299	15.291	15.279	15.261	15.238	15.208	15.177	
47	13.696	13	14.265	14.776	14.990	15.226	15.357	15.431	15.471	15.532	15.566	15.584	15.589	15.587	15.579	15.567	15.549	15.527	15.497	15.465	
48	13.982	13	14.555	15.066	15.280	15.515	15.647	15.720	15.760	15.821	15.855	15.872	15.878	15.875	15.867	15.855	15.838	15.815	15.786	15.754	
49	14.268	13	14.843	15.354	15.569	15.805	15.936	16.009	16.049	16.109	16.144	16.161	16.166	16.163	16.155	16.143	16.126	16.104	16.074	16.042	
50	14.555	13	15.129	15.641	15.857	16.093	16.224	16.298	16.337	16.398	16.432	16.449	16.454	16.451	16.443	16.431	16.413	16.392	16.362	16.331	
51	14.842	13	15.416	15.928	16.144	16.381	16.512	16.586	16.626	16.686	16.721	16.738	16.742	16.739	16.730	16.719	16.701	16.679	16.650	16.619	
52	15.130	13	15.702	16.215	16.430	16.668	16.800	16.874	16.914	16.974	17.009	17.026	17.031	17.027	17.018	17.007	16.989	16.967	16.937	16.906	
53	15.418	14	15.990	16.502	16.717	16.954	17.087	17.162	17.202	17.262	17.297	17.314	17.319	17.315	17.306	17.295	17.277	17.254	17.225	17.194	
54	15.706	14	16.277	16.789	17.005	17.242	17.374	17.449	17.489	17.550	17.585	17.602	17.607	17.603	17.595	17.583	17.565	17.542	17.513	17.481	
55	15.994	14	16.565	17.077	17.292	17.529	17.661	17.736	17.777	17.838	17.873	17.890	17.895	17.891	17.883	17.871	17.853	17.830	17.801	17.769	
56	16.282	14	16.853	17.365	17.580	17.817	17.949	18.024	18.064	18.125	18.161	18.178	18.183	18.179	18.171	18.159	18.141	18.118	18.089	18.057	
57	16.570	14	17.142	17.653	17.868	18.105	18.237	18.311	18.352	18.413	18.448	18.466	18.471	18.467	18.459	18.447	18.429	18.406	18.377	18.345	
58	16.859	14	17.430	17.941	18.156	18.393	18.525	18.599	18.639	18.701	18.736	18.753	18.759	18.755	18.747	18.735	18.717	18.694	18.665	18.633	
59	17.149	14	17.719	18.230	18.445	18.681	18.813	18.887	18.927	18.989	19.024	19.041	19.046	19.043	19.035	19.023	19.005	18.983	18.953	18.921	
60	17.438	14	18.008	18.519	18.734	18.970	19.102	19.176	19.216	19.277	19.312	19.329	19.334	19.331	19.323	19.311	19.293	19.271	19.241	19.210	

61	17.726	13	18.297	18.808	19.023	19.259	19.390	19.464	19.504	19.565	19.600	19.617	19.622	19.619	19.611	19.599	19.582	19.559	19.530	19.498
62	18.013	13	18.585	19.097	19.311	19.547	19.679	19.753	19.793	19.854	19.888	19.905	19.911	19.907	19.899	19.888	19.870	19.847	19.818	19.786
63	18.300	13	18.873	19.385	19.600	19.836	19.967	20.041	20.081	20.142	20.177	20.194	20.199	20.196	20.187	20.175	20.158	20.136	20.106	20.074
64	18.588	13	19.160	19.672	19.887	20.124	20.256	20.330	20.369	20.430	20.465	20.482	20.487	20.484	20.475	20.463	20.446	20.423	20.394	20.363
65	18.875	13	19.448	19.959	20.175	20.412	20.544	20.618	20.658	20.718	20.753	20.770	20.775	20.772	20.763	20.751	20.734	20.711	20.682	20.651
66	19.163	14	19.735	20.247	20.462	20.699	20.831	20.906	20.946	21.006	21.041	21.058	21.063	21.060	21.051	21.039	21.021	20.999	20.970	20.938
67	19.451	14	20.023	20.535	20.750	20.987	21.119	21.194	21.234	21.295	21.329	21.346	21.351	21.348	21.339	21.327	21.310	21.287	21.258	21.226
68	19.740	14	20.311	20.823	21.038	21.275	21.407	21.481	21.522	21.582	21.617	21.634	21.639	21.636	21.627	21.615	21.598	21.575	21.546	21.514
69	20.028	14	20.599	21.111	21.326	21.563	21.695	21.769	21.809	21.870	21.905	21.922	21.927	21.924	21.915	21.904	21.886	21.863	21.834	21.802
70	20.316	14	20.887	21.399	21.614	21.851	21.982	22.057	22.097	22.158	22.193	22.210	22.215	22.212	22.204	22.192	22.174	22.151	22.122	22.090
71	20.604	14	21.175	21.687	21.902	22.139	22.270	22.345	22.385	22.446	22.481	22.498	22.503	22.500	22.492	22.480	22.462	22.439	22.410	22.378
72	20.892	14	21.464	21.975	22.190	22.427	22.559	22.633	22.673	22.734	22.769	22.786	22.791	22.788	22.780	22.768	22.750	22.727	22.698	22.666
73	21.181	14	21.752	22.263	22.478	22.715	22.847	22.921	22.961	23.022	23.057	23.074	23.079	23.076	23.068	23.056	23.038	23.016	22.986	22.954
74	21.470	13	22.040	22.552	22.767	23.003	23.135	23.209	23.249	23.310	23.345	23.362	23.367	23.364	23.356	23.344	23.326	23.304	23.274	23.243
75	21.757	13	22.329	22.840	23.055	23.292	23.423	23.498	23.538	23.598	23.633	23.650	23.655	23.652	23.644	23.632	23.614	23.592	23.562	23.531
76	22.045	13	22.617	23.129	23.344	23.580	23.712	23.786	23.826	23.887	23.921	23.939	23.944	23.940	23.932	23.920	23.902	23.880	23.850	23.819
77	22.333	13	22.905	23.417	23.632	23.868	24.000	24.074	24.114	24.175	24.210	24.227	24.232	24.229	24.220	24.208	24.190	24.168	24.139	24.107
78	22.621	13	23.193	23.704	23.920	24.156	24.288	24.362	24.402	24.463	24.498	24.515	24.520	24.517	24.508	24.496	24.478	24.456	24.427	24.395
79	22.909	14	23.480	23.992	24.207	24.444	24.576	24.650	24.690	24.751	24.786	24.803	24.808	24.805	24.796	24.784	24.766	24.744	24.715	24.683
80	23.197	14	23.768	24.280	24.495	24.732	24.864	24.938	24.978	25.039	25.074	25.091	25.096	25.093	25.084	25.072	25.054	25.032	25.003	24.971
81	23.485	14	24.056	24.568	24.783	25.020	25.152	25.226	25.266	25.327	25.362	25.379	25.384	25.381	25.372	25.360	25.343	25.320	25.291	25.259
82	23.773	14	24.345	24.856	25.071	25.308	25.440	25.514	25.554	25.615	25.650	25.667	25.672	25.669	25.660	25.649	25.631	25.608	25.579	25.547
83	24.061	14	24.633	25.144	25.359	25.596	25.728	25.802	25.842	25.903	25.938	25.955	25.960	25.957	25.948	25.937	25.919	25.896	25.867	25.835
84	24.349	14	24.921	25.432	25.647	25.884	26.016	26.090	26.130	26.191	26.226	26.243	26.248	26.245	26.237	26.225	26.207	26.184	26.155	26.123
85	24.637	14	25.209	25.720	25.935	26.172	26.304	26.378	26.418	26.479	26.514	26.531	26.536	26.533	26.525	26.513	26.495	26.473	26.443	26.411
86	24.926	14	25.497	26.008	26.224	26.460	26.592	26.666	26.706	26.767	26.802	26.819	26.824	26.821	26.813	26.801	26.783	26.761	26.731	26.699
87	25.214	14	25.785	26.297	26.512	26.748	26.880	26.954	26.994	27.055	27.090	27.107	27.112	27.109	27.101	27.089	27.071	27.049	27.019	26.988
88	25.502	13	26.073	26.585	26.800	27.036	27.168	27.242	27.283	27.343	27.378	27.395	27.400	27.397	27.389	27.377	27.359	27.337	27.307	27.276
89	25.790	13	26.362	26.873	27.088	27.325	27.456	27.531	27.571	27.632	27.666	27.684	27.689	27.685	27.677	27.665	27.647	27.625	27.595	27.564
90	26.078	13	26.650	27.161	27.376	27.613	27.745	27.819	27.859	27.920	27.955	27.972	27.977	27.974	27.965	27.953	27.935	27.913	27.883	27.852
91	26.366	14	26.938	27.449	27.664	27.901	28.033	28.107	28.147	28.208	28.243	28.260	28.265	28.262	28.253	28.241	28.223	28.201	28.171	28.140
92	26.654	14	27.226	27.737	27.952	28.189	28.321	28.395	28.435	28.496	28.531	28.548	28.553	28.550	28.541	28.529	28.511	28.489	28.460	28.428
93	26.942	14	27.514	28.025	28.240	28.477	28.609	28.683	28.723	28.784	28.819	28.836	28.841	28.838	28.829	28.817	28.800	28.777	28.748	28.716
94	27.230	14	27.802	28.313	28.528	28.765	28.897	28.971	29.011	29.072	29.107	29.124	29.129	29.126	29.117	29.105	29.088	29.065	29.036	29.004
95	27.518	14	28.090	28.601	28.816	29.053	29.185	29.259	29.299	29.360	29.395	29.412	29.417	29.414	29.405	29.394	29.376	29.353	29.324	29.292
96	27.806	14	28.378	28.889	29.104	29.341	29.473	29.547	29.587	29.648	29.683	29.700	29.705	29.702	29.693	29.682	29.664	29.641	29.612	29.580
97	28.094	14	28.666	29.177	29.393	29.629	29.761	29.835	29.875	29.936	29.971	29.988	29.993	29.990	29.982	29.970	29.952	29.929	29.900	29.868
98	28.382	14	28.954	29.465	29.681	29.917	30.049	30.123	30.163	30.224	30.259	30.276	30.281	30.278	30.270	30.258	30.240	30.218	30.188	30.156
99	28.670	14	29.242	29.754	29.969	30.205	30.337	30.411	30.451	30.512	30.547	30.564	30.569	30.566	30.558	30.546	30.528	50.506	30.476	30.445
100	28.959	14	29.530	30.042	30.257	30.493	30.625	30.699	30.739	30.800	30.835	30.852	30.857	30.854	30.846	30.834	30.816	30.794	30.764	30.733

TABLE V C — EXPECTED NUMBER OF TESTS REQUIRED AND SIZE x OF THE NEXT SAMPLE TO BE TAKEN FOR ANY $G_1(m,n)$ AND ANY $H_1(n)$-SITUATION WHEN PROCEDURE R_1 IS USED AND $q=0.99$

			$G_1(m,n)$																	
n	$H_1(n)$	x	$m=2$ $x=1$	$m=3$ $x=1$	$m=4$ $x=2$	$m=5$ $x=2$	$m=6$ $x=2$	$m=7$ $x=3$	$m=8$ $x=4$	$m=9$ $x=4$	$m=10$ $x=4$	$m=20$ $x=8$	$m=30$ $x=14$	$m=40$ $x=16$	$m=50$ $x=18$	$m=60$ $x=28$	$m=70$ $x=32$	$m=80$ $x=32$	$m=90$ $x=32$	$m=100$ $x=35$
1	1.000	1	—																	
2	1.030	2	1.503																	
3	1.070	3	2.015	2.343																
4	1.110	4	2.050	2.697	2.779															
5	1.159	5	2.090	2.733	3.053	3.240														
6	1.208	6	2.134	2.776	3.092	3.468	3.560													
7	1.258	7	2.184	2.823	3.137	3.510	3.757	3.798												
8	1.308	8	2.233	2.872	3.184	3.556	3.800	3.974	3.986											
9	1.366	9	2.283	2.922	3.234	3.603	3.847	4.018	4.145	4.237										
10	1.425	10	2.337	2.974	3.286	3.655	3.896	4.066	4.191	4.385	4.447									
11	1.484	11	2.396	3.030	3.340	3.708	3.949	4.117	4.241	4.433	4.585									
12	1.543	12	2.455	3.089	3.397	3.763	4.003	4.171	4.293	4.483	4.634									
13	1.603	13	2.514	3.148	3.455	3.821	4.059	4.226	4.347	4.536	4.685									
14	1.662	14	2.573	3.207	3.515	3.879	4.117	4.282	4.402	4.591	4.739									
15	1.722	15	2.632	3.266	3.574	3.939	4.176	4.340	4.459	4.647	4.794									
16	1.782	16	2.692	3.326	3.633	3.998	4.235	4.399	4.517	4.704	4.851									
17	1.849	17	2.752	3.386	3.693	4.057	4.294	4.458	4.577	4.762	4.908									
18	1.916	18	2.815	3.448	3.754	4.119	4.355	4.519	4.637	4.822	4.967									
19	1.984	19	2.883	3.513	3.818	4.181	4.417	4.580	4.698	4.884	5.028									
20	2.051	20	2.950	3.580	3.883	4.246	4.481	4.643	4.761	4.946	5.090	5.773								
21	2.119	21	3.018	3.647	3.951	4.312	4.546	4.708	4.824	5.009	5.153	5.871								
22	2.187	22	3.085	3.715	4.018	4.379	4.612	4.773	4.889	5.073	5.216	5.928								
23	2.255	23	3.153	3.783	4.086	4.447	4.680	4.839	4.955	5.138	5.281	5.986								
24	2.324	24	3.221	3.851	4.154	4.515	4.747	4.907	5.021	5.204	5.346	6.046								
25	2.392	25	3.290	3.919	4.222	4.583	4.815	4.975	5.089	5.271	5.412	6.107								
26	2.461	26	3.358	3.987	4.290	4.651	4.883	5.043	5.157	5.339	5.479	6.169								
27	2.530	27	3.427	4.056	4.359	4.719	4.952	5.111	5.225	5.407	5.547	6.232								
28	2.599	28	3.495	4.125	4.427	4.788	5.020	5.179	5.293	5.475	5.615	6.296								
29	2.668	29	3.564	4.194	4.496	4.856	5.089	5.248	5.362	5.543	5.684	6.360								
30	2.738	30	3.634	4.263	4.565	4.925	5.157	5.317	5.431	5.612	5.752	6.426	6.676							

31	2.807	31	3.703	4.332	4.634	4.994	5.227	5.386	5.499	5.681	5.821	6.492	6.763							
32	2.877	32	3.773	4.401	4.704	5.064	5.296	5.455	5.568	5.749	5.890	6.558	6.824							
33	2.952	33	3.842	4.471	4.773	5.133	5.365	5.524	5.638	5.819	5.959	6.625	6.887							
34	3.027	34	3.915	4.543	4.844	5.204	5.436	5.594	5.708	5.888	6.028	6.692	6.950							
35	3.101	35	3.989	4.616	4.917	5.276	5.507	5.665	5.778	5.959	6.099	6.761	7.014							
36	3.176	36	4.064	4.690	4.990	5.348	5.579	5.737	5.850	6.030	6.170	6.830	7.080							
37	3.252	37	4.139	4.765	5.065	5.422	5.652	5.810	5.922	6.102	6.241	6.900	7.146							
38	3.327	38	4.214	4.840	5.140	5.497	5.726	5.883	5.995	6.175	6.314	6.970	7.213							
39	3.402	39	4.289	4.915	5.215	5.572	5.801	5.958	6.069	6.248	6.387	7.041	7.281							
40	3.478	40	4.365	4.991	5.290	5.647	5.876	6.033	6.143	6.322	6.460	7.112	7.349	7.486						
41	3.554	41	4.440	5.066	5.366	5.723	5.952	6.108	6.218	6.396	6.534	7.183	7.418	7.570						
42	3.630	42	4.516	5.142	5.441	5.798	6.027	6.183	6.294	6.472	6.609	7.255	7.487	7.635						
43	3.706	43	4.592	5.218	5.517	5.874	6.103	6.259	6.369	6.547	6.684	7.328	7.557	7.701						
44	3.782	44	4.668	5.294	5.593	5.950	6.178	6.334	6.445	6.622	6.760	7.400	7.628	7.767						
45	3.859	45	4.744	5.370	5.669	6.025	6.254	6.410	6.520	6.698	6.835	7.474	7.699	7.835						
46	3.935	46	4.821	5.446	5.745	6.102	6.330	6.486	6.596	6.774	6.911	7.547	7.770	7.903						
47	4.012	47	4.897	5.522	5.821	6.178	6.406	6.562	6.672	6.850	6.987	7.621	7.842	7.972						
48	4.089	48	4.974	5.599	5.898	6.254	6.483	6.638	6.748	6.926	7.063	7.695	7.915	8.042						
49	4.166	49	5.051	5.676	5.975	6.331	6.559	6.715	6.825	7.002	7.139	7.770	7.987	8.112						
50	4.243	50	5.128	5.753	6.052	6.408	6.636	6.791	6.901	7.079	7.215	7.845	8.060	8.183	8.211					
51	4.321	51	5.205	5.830	6.129	6.485	6.713	6.868	6.978	7.155	7.292	7.920	8.134	8.254	8.293					
52	4.398	52	5.282	5.907	6.206	6.562	6.790	6.945	7.055	7.232	7.369	7.996	8.207	8.325	8.361					
53	4.476	53	5.360	5.985	6.283	6.639	6.867	7.022	7.132	7.309	7.446	8.072	8.281	8.397	8.430					
54	4.554	54	5.438	6.062	6.361	6.717	6.945	7.100	7.209	7.386	7.523	8.148	8.356	8.470	8.499					
55	4.632	55	5.515	6.140	6.438	6.794	7.022	7.177	7.287	7.464	7.600	8.225	8.430	8.543	8.569					
56	4.711	56	5.594	6.218	6.516	6.872	7.100	7.255	7.364	7.541	7.677	8.301	8.505	8.616	8.639					
57	4.789	57	5.672	6.296	6.594	6.950	7.178	7.333	7.442	7.619	7.755	8.378	8.581	8.690	8.710					
58	4.868	58	5.750	6.375	6.763	7.028	7.256	7.411	7.520	7.697	7.833	8.455	8.656	8.763	8.782					
59	4.947	59	5.829	6.453	6.751	7.107	7.334	7.489	7.598	7.775	7.910	8.532	8.732	8.838	8.854					
60	5.026	60	5.908	6.532	6.830	7.185	7.413	7.567	7.676	7.853	7.989	8.609	8.808	8.912	8.927	8.891				

TABLE V C — *Continued*

n	$H_1(n)$	x	$G_1(m,n)$																
			$m=2$ $x=1$	$m=3$ $x=1$	$m=4$ $x=2$	$m=5$ $x=2$	$m=6$ $x=2$	$m=7$ $x=3$	$m=8$ $x=4$	$m=9$ $x=4$	$m=10$ $x=4$	$m=20$ $x=8$	$m=30$ $x=14$	$m=40$ $x=16$	$m=50$ $x=18$	$m=60$ $x=28$	$m=70$ $x=32$	$m=80$ $x=32$	$m=90$ $x=32$	$m=100$ $x=35$
61	5.105	61	5.987	6.611	6.909	7.264	7.491	7.646	7.755	7.931	8.067	8.687	8.884	8.987	8.999	8.972				
62	5.185	62	6.066	6.690	6.988	7.343	7.570	7.724	7.833	8.010	8.145	8.765	8.961	9.062	9.073	9.042				
63	5.265	63	6.145	6.769	7.067	7.422	7.649	7.803	7.912	8.088	8.224	8.842	9.038	9.137	9.146	9.113				
64	5.345	64	6.225	6.849	7.146	7.501	7.728	7.883	7.991	8.167	8.303	8.920	9.116	9.213	9.220	9.184				
65	5.425	65	6.305	6.929	7.226	7.581	7.808	7.962	8.070	8.247	8.382	8.999	9.193	9.289	9.294	9.256				
66	5.506	66	6.385	7.009	7.306	7.661	7.887	8.041	8.150	8.326	8.461	9.077	9.270	9.365	9.369	9.328				
67	5.587	67	6.466	7.089	7.386	7.741	7.967	8.121	8.230	8.406	8.541	9.156	9.348	9.442	9.444	9.401				
68	5.667	68	6.546	7.170	7.467	7.821	8.048	8.201	8.310	8.485	8.620	9.234	9.426	9.518	9.520	9.475				
69	5.748	69	6.627	7.250	7.547	7.901	8.128	8.282	8.390	8.565	8.700	9.313	9.504	9.595	9.595	9.548				
70	5.830	70	6.708	7.331	7.628	7.982	8.209	8.362	8.470	8.646	8.781	9.393	9.583	9.673	9.671	9.622	9.560			
71	5.911	71	6.789	7.412	7.709	8.063	8.289	8.443	8.551	8.726	8.861	9.472	9.661	9.750	9.747	9.697	9.641			
72	5.992	72	6.870	7.493	7.790	8.144	8.370	8.524	8.632	8.807	8.941	9.551	9.740	9.828	9.824	9.772	9.714			
73	6.074	73	6.952	7.575	7.871	8.225	8.451	8.605	8.713	8.888	9.022	9.631	9.819	9.906	9.900	9.847	9.786			
74	6.155	74	7.033	7.656	7.953	8.307	8.533	8.686	8.794	8.969	9.103	9.711	9.898	9.984	9.977	9.922	9.859			
75	6.237	75	7.115	7.737	8.034	8.388	8.614	8.767	8.875	9.050	9.184	9.791	9.977	10.063	10.054	9.998	9.933			
76	6.319	76	7.196	7.819	8.115	8.469	8.695	8.848	8.956	9.131	9.265	9.872	10.057	10.142	10.132	10.074	10.007			
77	6.401	77	7.278	7.901	8.197	8.551	8.777	8.930	9.037	9.212	9.347	9.952	10.136	10.220	10.209	10.150	10.081			
78	6.483	78	7.360	7.983	8.279	8.633	8.858	9.011	9.119	9.294	9.428	10.033	10.216	10.299	10.287	10.227	10.156			
78	6.565	79	7.442	8.064	8.361	8.714	8.940	9.093	9.201	9.375	9.510	10.113	10.296	10.378	10.365	10.304	10.231			
80	6.647	80	7.524	8.146	8.443	8.796	9.022	9.175	9.282	9.457	9.591	10.194	10.376	10.458	10.444	10.381	10.307	10.221		
81	6.729	81	7.606	8.229	8.525	8.878	9.104	9.257	9.364	9.539	9.673	10.275	10.456	10.537	10.522	10.458	10.383	10.302		
82	6.812	82	7.688	8.311	8.607	8.960	9.186	9.339	9.446	9.621	9.755	10.357	10.537	10.617	10.601	10.536	10.459	10.376		
83	6.895	83	7.771	8.393	8.689	9.043	9.268	9.421	9.528	9.703	9.837	10.438	10.617	10.697	10.680	10.613	10.535	10.450		
84	6.977	84	7.853	8.476	8.772	9.125	9.351	9.503	9.611	9.785	9.919	10.520	10.698	10.777	10.759	10.691	10.612	10.525		
85	7.060	85	7.936	8.559	8.854	9.208	9.433	9.586	9.693	9.867	10.001	10.601	10.779	10.857	10.839	10.770	10.689	10.600		
86	7.143	86	8.019	8.641	8.937	9.290	9.516	9.668	9.775	9.950	10.084	10.683	10.860	10.937	10.918	10.848	10.766	10.676		
87	7.227	87	8.102	8.724	9.020	9.373	9.599	9.751	9.858	10.032	10.166	10.765	10.941	11.017	10.998	10.927	10.844	10.752		
88	7.310	88	8.185	8.807	9.103	9.456	9.681	9.834	9.941	10.115	10.249	10.847	11.022	11.098	11.078	11.006	10.922	10.828		
89	7.393	89	8.268	8.890	9.186	9.539	9.764	9.917	10.024	10.198	10.331	10.929	11.104	11.179	11.158	11.085	10.999	10.904		
90	7.477	90	8.352	8.974	9.269	9.622	9.848	10.000	10.107	10.281	10.414	11.011	11.186	11.260	11.238	11.164	11.078	10.981	10.881	

91	7.561	91	8.435	9.057	9.353	9.706	9.931	10.083	10.190	10.364	10.497	11.094	11.267	11.341	11.318	11.244	11.156	11.058	10.962	
92	7.644	92	8.519	9.141	9.436	9.789	10.014	10.166	10.273	10.447	10.580	11.176	11.349	11.422	11.399	11.323	11.235	11.135	11.037	
93	7.728	93	8.603	9.225	9.520	9.873	10.098	10.250	10.356	10.530	10.664	11.259	11.431	11.503	11.479	11.403	11.313	11.213	11.113	
94	7.813	94	8.687	9.308	9.604	9.956	10.181	10.333	10.440	10.614	10.747	11.342	11.514	11.585	11.560	11.483	11.393	11.291	11.189	
95	7.897	95	8.771	9.392	9.688	10.040	10.265	10.417	10.524	10.697	10.831	11.425	11.596	11.666	11.641	11.564	11.472	11.369	11.266	
96	7.981	96	8.855	9.477	9.772	10.124	10.349	10.501	10.607	01.781	10.914	11.508	11.678	11.748	11.722	11.644	11.551	11.447	11.342	
97	8.066	97	8.939	9.561	9.856	10.208	10.433	10.585	10.691	10.865	10.998	11.591	11.761	11.830	11.803	11.725	11.631	11.526	11.420	
98	8.151	98	9.024	9.645	9.940	10.293	10.517	10.669	10.775	10.949	11.082	11.674	11.844	11.912	11.885	11.805	11.711	11.605	11.497	
99	8.235	99	9.108	9.730	10.025	10.377	10.602	10.753	10.860	11.033	11.166	11.758	11.927	11.994	11.966	11.886	11.791	11.684	11.575	
100	8.320	100	9.193	9.815	10.109	10.462	10.686	10.838	10.944	11.118	11.250	11.841	12.010	12.077	12.048	11.967	11.871	11.763	11.653	11.547

VALUES OF THE NEXT TEST-GROUP SIZE x UNDER PROCEDURE R_1 FOR $q = 0.99$ AND $m = 2(1)100$

	1	2	3	4	5	6	7	8	9	10
0+		1	1	2	2	2	3	4	4	4
10+	4	4	5	6	7	8	8	8	8	8
20+	8	8	8	8	9	10	11	12	13	14
30+	15	16	16	16	16	16	16	16	16	16
40+	16	16	16	16	16	16	16	16	17	18
50+	19	20	21	22	23	24	25	26	27	28
60+	29	30	31	32	32	32	32	32	32	32
70+	32	32	32	32	32	32	32	32	32	32
80+	32	32	32	32	32	32	32	32	32	32
90+	32	32	32	32	32	32	33	34	35	35

TABLE VI— EXPECTED NUMBER OF TESTS REQUIRED FOR PROCEDURE R_2

The integer below q^v opposite $H_2(n)$ is the coefficient of q^v in the polynomial formula for $H_2(n)$.

	q-interval	x	1	q	q^2	q^3	q^4	q^5	q^6	q^7	q^8	q^9	q^{10}	q^{11}	q^{12}
$H_2(2)$	0.000 to 0.618	1	2												
	0.618 to 1.000	2	3	−1	−1										
$H_2(3)$	0.000 to 0.618	1	3												
	0.618 to 0.755	2	5	−3	−1	1									
	0.755 to 1.000	3	5	−2	−1	−1									
$H_2(4)$	0.000 to 0.618	1	4												
	0.618 to 0.755	2	7	−5	0	1	−1								
	0.755 to 0.819	3	7	−3	−2	−2	2								
	0.819 to 1.000	4	8	−4	−2	−1									
$H_2(5)$	0.000 to 0.618	1	5												
	0.618 to 0.755	2	9	−7	1	0	−1	1							
	0.755 to 0.819	3	9	−4	−3	−2	3	−1							
	0.819 to 0.857	4	11	−7	−2	−1	−1	2							
	0.857 to 1.000	5	11	−7	−2	0	0	−1							
$H_2(6)$	0.000 to 0.618	1	6												
	0.618 to 0.755	2	11	−9	2	−1	0	1	−1						
	0.755 to 0.819	3	11	−5	−4	−2	4	−1	−1						
	0.819 to 0.857	4	14	−10	−2	−1	−1	3	−1						
	0.857 to 0.881	5	14	−10	−2	1	−1	−3	3						
	0.881 to 0.885	6	14	−10	−1	0	−1	−1							
	0.885 to 1.000	6	14	−9	−2	−1	0	−1							
$H_2(7)$	0.000 to 0.618	1	7												
	0.618 to 0.755	2	13	−11	3	−2	1	0	−1	1					
	0.755 to 0.819	3	13	−6	−5	−2	5	−2	−2	2					
	0.819 to 0.857	4	17	−13	−2	−1	−1	4	−1	−1					
	0.857 to 0.881	5	17	−13	−2	2	−2	−4	5	−1					
	0.881 to 0.885	6	17	−13	0	−1	−1	−1	−2	3					
	0.885 to 0.899⁻	6	17	−11	−3	−2	2	−2	−2	3					
	0.899⁻ to 1.000	7	17	−11	−3	−1	1	−2							
$H_2(8)$	0.000 to 0.618	1	8												
	0.618 to 0.755	2	15	−13	4	−3	2	−1	0	1	−1				
	0.755 to 0.819	3	15	−7	−6	−2	6	−3	−2	3	−1				
	0.819 to 9.857	4	20	−16	−2	−1	0	4	−2	−1					
	0.857 to 0.881	5	20	−16	−2	3	−3	−5	7	−1	−1				
	0.881 to 0.885	6	20	−16	1	−2	−1	−1	−3	5	−1				
	0.885 to 0.899⁻	6	20	−13	−4	−3	4	−3	−3	5	−1				
	0.899⁻ to 0.912	7	20	−13	−4	−1	1	−2	0	−2	3				
	0.912 to 0.932	8	20	−13	−4	−1	1	−2	1	0	−1				
	0.932 to 1.000	8	20	−14	−4	−1	1	−2							
$H_2(9)$	0.000 to 0.618	1	9												
	0.618 to 0.755	2	17	−15	5	−4	3	−2	1	0	−1	1			
	0.755 to 0.819	3	17	−8	−7	−2	7	−4	−2	4	−1	−1			
	0.819 to 0.857	4	23	−19	−2	−1	1	3	−2	−1	−1	2			
	0.857 to 0.881	5	23	−19	−2	4	−4	−5	8	−2	−1				
	0.881 to 0.890	6	23	−19	2	−3	−1	−1	−4	7	−1	−1			
	0.890 to 0.899⁻	6	23	−15	−5	−4	6	−4	−4	7	−1	−1			
	0.899⁻ to 0.912	7	23	−15	−5	−1	1	−2	0	−3	5	−1			
	0.912 to 0.922	8	23	−15	−5	−1	1	−2	2	−1	−4	4			
	0.922 to 0.938	9	24	−16	−5	−1	1	−2	1	0	−1				
	0.938 to 1.000	9	25	−18	−4	−1	1	−2	0	1	0	−1			

TABLE VI — *Continued*

	q-interval	x	1	q	q^2	q^3	q^4	q^5	q^6	q^7	q^8	q^9	q^{10}	q^{11}	q^{12}
	0.000 to 0.618	1	10												
	0.618 to 0.755	2	19	−17	6	−5	4	−3	2	−1	0	1	−1		
	0.755 to 0.819	3	19	−9	−8	−2	8	−5	−2	5	−2	−2	2		
	0.819 to 0.857	4	26	−22	−2	−1	2	2	−2	−1	−1	3	−1		
	0.857 to 0.881	5	26	−22	−2	5	−5	−5	8	−2	0	0	−1		
	0.881 to 0.890	6	26	−22	3	−4	−1	−1	−4	8	−2	−1			
$H_2(10)$	0.890 to 0.899⁻	6	26	−17	−6	−5	8	−5	−4	8	−2	−1			
	0.899⁻ to 0.912	7	26	−17	−6	−1	1	−2	0	−4	7	−1	−1		
	0.912 to 0.922	8	26	−17	−6	−1	1	−2	3	−2	−6	7	−1		
	0.922 to 0.930	9	28	−20	−5	−1	1	−2	1	1	−2	−3	4		
	0.930 to 0.938	10	28	−20	−5	−1	1	−1	1	−1	−1				
	0.938 to 0.960	10	29	−22	−4	−1	1	−1	0	0	0	−1			
	0.960 to 1.000	10	29	−22	−4	0	1	−3	0	1	0	−1			
	0.000 to 0.618	1	11												
	0.618 to 0.755	2	21	−19	7	−6	5	−4	3	−2	1	0	−1	1	
	0.755 to 0.819	3	21	−10	−9	−2	9	−6	−2	6	−3	−2	3	−1	
	0.819 to 0.857	4	29	−25	−2	−1	3	1	−2	−1	−1	4	−1	−1	
	0.857 to 0.881	5	29	−25	−2	6	−6	−5	8	−2	1	−1	−3	3	
	0.881 to 0.890	6	29	−25	4	−5	−1	−1	−4	8	−2	0	0	−1	
	0.890 to 0.899⁻	6	29	−19	−7	−6	10	−6	−4	8	−2	0	0	−1	
$H_2(11)$	0.899⁻ to 0.912	7	29	−19	−7	−1	1	−2	0	−4	8	−2	−1		
	0.912 to 0.922	8	29	−19	−7	−1	1	−2	4	−3	−8	10	−1	−1	
	0.922 to 0.930	9	32	−24	−5	−1	1	−2	1	2	−3	−5	7	−1	
	0.930 to 0.936	10	32	−24	−5	−1	1	0	0	−2	1	−1	3	4	
	0.936 to 0.938	11	32	−24	−5	0	1	−2	1	−1	−1				
	0.938 to 0.960	11	33	−26	−4	0	1	−2	0	0	0	−1			
	0.960 to 1.000	11	33	−26	−4	1	0	−4	2	1	−1	−1			
	0.000 to 0.618	1	12												
	0.618 to 0.755	2	23	−21	8	−7	6	−5	4	−3	2	−1	0	1	−1
	0.755 to 0.819	3	23	−11	−10	−2	10	−7	−2	7	−4	−2	4	−1	−1
	0.819 to 0.857	4	32	−28	−2	−1	4	0	−2	−1	0	4	−2	−1	
	0.857 to 0.881	5	32	−28	−2	7	−7	−5	8	−2	2	−2	−4	5	−1
	0.881 to 0.890	6	32	−28	5	−6	−1	−1	−4	8	−1	0	−1	−1	
	0.890 to 0.899⁻	6	32	−21	−8	−7	12	−7	−4	9	−2	−1	0	−1	
	0.899⁻ to 0.912	7	32	−21	−8	−1	1	−2	0	−4	8	−2	0	0	−1
$H_2(12)$	0.912 to 0.922	8	32	−21	−8	−1	1	−2	5	−4	−9	12	−2	−1	
	0.922 to 0.930	9	36	−28	−5	−1	1	−2	1	3	−4	−7	10	−1	−1
	0.930 to 0.936	10	36	−28	−5	−1	1	1	−1	−3	3	−2	−5	7	−1
	0.936 to 0.938	11	36	−28	−5	1	0	−3	3	−2	−2	2	−1	−3	4
	0.938 to 0.941	11	37	−30	−4	1	0	−3	2	−1	−1	1	−1	−3	4
	0.941 to 0.960	12	37	−30	−4	1	0	−3	2	−1	0	0	−1		
	0.960 to 0.972	12	37	−30	−4	2	−1	−5	4	0	−1	0	−1		
	0.972 to 1.000	12	37	−29	−4	0	0	−4	2	1	−1	−1			

The exponential symbols +, − indicate only the relative magnitude of two different roots that are equal to three decimal places (i.e., $a^- < a^+$).

TABLE VII — THE DIVIDING POINTS BETWEEN x AND $x + 1$ FOR THE INFORMATION PROCEDURE R_2

G-Situation

x	m									
	4	6	7	8	9	11	12	13	14	16
1	0.7549	0.6518	0.6369	0.6289	0.6245$^-$	0.6204	0.6195$^-$	0.6189	0.6186	0.6182
2		0.8899	0.8376	0.8087	0.7913	0.7728	0.7677	0.7642	0.7617	0.7586
3				0.9378	0.9016	0.8631	0.8524	0.8446	0.8388	0.8313
4						0.9340	0.9160	0.9031	0.8935$^-$	0.8806
5							0.9723	0.9528	0.9384	0.9193
6									0.9796	0.9529
7										0.9844

G-Situation

m										x
10	20	30	40	50	60	70	80	90	100	
0.6210	0.6181	0.6180	0.6180	0.6180	0.6180	0.6180	0.6180	0.6180	0.6180	1
0.7802	0.7560	0.7549	0.7549	0.7549	0.7549	0.7549	0.7549	0.7549	0.7549	2
0.8786	0.8241	0.8198	0.8193	0.8192	0.8192	0.8192	0.8192	0.8192	0.8192	3
0.9601	0.8677	0.8587	0.8571	0.8568	0.8567	0.8567	0.8567	0.8567	0.8567	4
	0.8999	0.8854	0.8823	0.8816	0.8814	0.8813	0.8813	0.8813	0.8813	5
	0.9262	0.9056	0.9008	0.8993	0.8989	0.8987	0.8987	0.8987	0.8987	6
	0.9490	0.9218	0.9150	0.9129	0.9121	0.9118	0.9117	0.9116	0.9116	7
	0.9699	0.9355$^-$	0.9267	0.9236	0.9225$^-$	0.9220	0.9218	0.9217	0.9216	8
	0.9900	0.9474	0.9364	0.9325$^+$	0.9309	0.9302	0.9299	0.9297	0.9296	9
		0.9582	0.9449	0.9400	0.9380	0.9370	0.9365$^+$	0.9363	0.9362	10
		0.9682	0.9524	0.9466	0.9440	0.9428	0.9422	0.9418	0.9417	11
		0.9776	0.9593	0.9524	0.9493	0.9478	0.9470	0.9466	0.9463	12
		0.9867	0.9655$^+$	0.9576	0.9540	0.9522	0.9512	0.9507	0.9504	13
		0.9956	0.9714	0.9624	0.9582	0.9561	0.9549	0.9543	0.9539	14
46	0.9981		0.9770	0.9668	0.9620	0.9596	0.9583	0.9575$^-$	0.9570	15
45	0.9982		0.9823	0.9709	0.9656	0.9628	0.9613	0.9604	0.9598	16
44	0.9973		0.9874	0.9747	0.9689	0.9658	0.9640	0.9630	0.9623	17
43	0.9964		0.9925$^-$	0.9784	0.9719	0.9685$^-$	0.9665$^+$	0.9654	0.9646	18
42	0.9955$^+$		0.9975$^-$	0.9820	0.9748	0.9710	0.9689	0.9676	0.9667	19
41	0.9946	0.9989		0.9854	0.9776	0.9734	0.9711	0.9696	0.9687	20
40	0.9937	0.9978		0.9887	0.9802	0.9757	0.9731	0.9715$^-$	0.9705$^-$	21
39	0.9928	0.9967		0.9920	0.9827	0.9779	0.9750$^+$	0.9733	0.9721	22
38	0.9919	0.9956		0.9952	0.9852	0.9799	0.9769	0.9750$^-$	0.9737	23
37	0.9909	0.9944		0.9984	0.9876	0.9819	0.9786	0.9765$^+$	0.9752	24
36	0.9899	0.9933	0.9986		0.9899	0.9838	0.9803	0.9780	0.9766	25
35	0.9889	0.9921	0.9972		0.9922	0.9856	0.9819	0.9795$^-$	0.9779	26
34	0.9879	0.9910	0.9957		0.9944	0.9874	0.9834	0.9809	0.9792	27
33	0.9869	0.9898	0.9943		0.9967	0.9892	0.9849	0.9822	0.9804	28
32	0.9858	0.9886	0.9928		0.9989	0.9909	0.9863	0.9835$^-$	0.9816	29
31	0.9847	0.9873	0.9914	0.9981		0.9926	0.9877	0.9847	0.9827	30
30	0.9836	0.9860	0.9899	0.9962		0.9943	0.9891	0.9859	0.9838	31
29	0.9824	0.9847	0.9883	0.9943		0.9959	0.9904	0.9871	0.9848	32
28	0.9812	0.9834	0.9868	0.9924		0.9976	0.9918	0.9882	0.9858	33
27	0.9800	0.9820	0.9852	0.9904		0.9992	0.9931	0.9893	0.9868	34
26	0.9786	0.9805$^+$	0.9835$^+$	0.9884	0.9974		0.9943	0.9904	0.9878	35
25	0.9773	0.9790	0.9818	0.9864	0.9947		0.9956	0.9914	0.9887	36
24	0.9758	0.9775$^-$	0.9800	0.9843	0.9920		0.9969	0.9925$^-$	0.9896	37
23	0.9743	0.9758	0.9782	0.9822	0.9893		0.9981	0.9935$^+$	0.9905$^+$	38
22	0.9727	0.9741	0.9763	0.9800	0.9866		0.9994	0.9945$^+$	0.9914	39
21	0.9709	0.9722	0.9742	0.9776	0.9837	0.9961		0.9955$^+$	0.9922	40
20	0.9691	0.9702	0.9721	0.9752	0.9808	0.9921		0.9965$^+$	0.9931	41
19	0.9671	0.9681	0.9698	0.9727	0.9778	0.9881		0.9975$^+$	0.9939	42
18	0.9649	0.9659	0.9674	0.9700	0.9746	0.9840		0.9985$^+$	0.9948	43
17	0.9626	0.9634	0.9648	0.9671	0.9713	0.9798		0.9995$^+$	0.9956	44
16	0.9600	0.9608	0.9619	0.9640	0.9678	0.9754	0.9935$^-$		0.9964	45
15	0.9572	0.9578	0.9588	0.9606	0.9640	0.9708	0.9869		0.9972	46
14	0.9541	0.9546	0.9554	0.9570	0.9599	0.9659	0.9802		0.9980	47
13	0.9505$^-$	0.9509	0.9516	0.9529	0.9555$^-$	0.9607	0.9733		0.9988	48
12	0.9464	0.9467	0.9473	0.9484	0.9506	0.9551	0.9660		0.9996	49
11	0.9417	0.9420	0.9424	0.9433	0.9451	0.9489	0.9582	0.9872		
10	0.9362	0.9364	0.9367	0.9374	0.9388	0.9419	0.9498	0.9742		
9	0.9297	0.9298	0.9300	0.9305$^-$	0.9315$^+$	0.9340	0.9404	0.9608		
8	0.9216	0.9217	0.9218	0.9222	0.9229	0.9248	0.9298	0.9463		
7	0.9116	0.9116	0.9117	0.9119	0.9124	0.9137	0.9174	0.9303		
6	0.8987	0.8987	0.8987	0.8988	0.8991	0.8999	0.9024	0.9120	0.9645$^-$	
5	0.8813	0.8813	0.8813	0.8813	0.8814	0.8818	0.8834	0.8898	0.9276	
4	0.8567	0.8567	0.8567	0.8567	0.8567	0.8569	0.8576	0.8612	0.8862	
3	0.8192	0.8192	0.8192	0.8192	0.8192	0.8192	0.8194	0.8209	0.8346	
2	0.7549	0.7549	0.7549	0.7549	0.7549	0.7549	0.7549	0.7552	0.7599	
1	0.6180	0.6180	0.6180	0.6180	0.6180	0.6180	0.6180	0.6180	0.6184	0.6823
x	95	85	75	65	55	45	35	25	15	5
	m									

TABLE VII — *Continued*

H-Situation

x	q	x	q
1	0.6180	51	0.9866
2	0.7549	52	0.9869
3	0.8192	53	0.9871
4	0.8567	54	0.9874
5	0.8813	55	0.9876
6	0.8987	56	0.9878
7	0.9116	57	0.9880
8	0.9216	58	0.9882
9	0.9296	59	0.9884
10	0.9361	60	0.9886
11	0.9415^{-}	61	0.9888
12	0.9460	62	0.9890
13	0.9499	63	0.9891
14	0.9533	64	0.9893
15	0.9563	65	0.9895^{-}
16	0.9588	66	0.9896
17	0.9612	67	0.9898
18	0.9632	68	0.9899
19	0.9651	69	0.9901
20	0.9667	70	0.9902
21	0.9683	71	0.9904
22	0.9697	72	0.9905^{-}
23	0.9709	73	0.9906
24	0.9721	74	0.9907
25	0.9732	75	0.9909
26	0.9742	76	0.9910
27	0.9751	77	0.9911
28	0.9760	78	0.9912
29	0.9768	79	0.9913
30	0.9775^{+}	80	0.9914
31	0.9782	81	0.9915^{+}
32	0.9789	82	0.9916
33	0.9795^{+}	83	0.9917
34	0.9801	84	0.9918
35	0.9807	85	0.9919
36	0.9812	86	0.9920
37	0.9817	87	0.9921
38	0.9822	88	0.9922
39	0.9826	89	0.9923
40	0.9830	90	0.9924
41	0.9834	91	0.9925^{-}
42	0.9838	92	0.9925^{+}
43	0.9842	93	0.9926
44	0.9845^{+}	94	0.9927
45	0.9849	95	0.9928
46	0.9852	96	0.9928
47	0.9855^{+}	97	0.9929
48	0.9858	98	0.9930
49	0.9861	99	0.9931
50	0.9864	100	0.9931

The exponents $^{+}$ and $^{-}$ are to be used for rounding in the usual manner.

TABLE VIII — POLYNOMIAL EXPRESSIONS FOR $F_2^*(m)$ FOR PROCEDURE R_2

m	Coefficient of												q-interval	Next Test Group Size
	1	q	q^2	q^3	q^4	q^5	q^6	q^7	q^8	q^9	q^{10}	q^{11}		
2	1	1											$0.000 \leqq q \leqq 1.000$	1
3	1	2	2										$0.000 \leqq q \leqq 1.000$	1
4	2	2	2	2									$0.755 \leqq q \leqq 1.000$	2
5	2	2	2	3	3								$0.682 \leqq q \leqq 1.000$	2
6	2	2	3	3	3	3							$0.652 \leqq q \leqq 0.890$	2
	2	3	3	2	3	3							$0.890 \leqq q \leqq 1.000$	3
7	2	3	3	3	3	3	3						$0.838 \leqq q \leqq 1.000$	3
8	2	3	3	3	3	3	4	4					$0.809 \leqq q \leqq 0.938$	3
	3	3	3	3	3	3	3	3					$0.938 \leqq q \leqq 1.000$	4
9	3	3	3	3	3	3	3	4	4				$0.902 \leqq q \leqq 1.000$	4
10	3	3	3	3	3	4	4	3	4	4			$0.879 \leqq q \leqq 0.960$	4
	3	3	3	4	4	3	3	3	4	4			$0.960 \leqq q \leqq 1.000$	5
11	3	3	3	4	4	3	4	4	3	4	4		$0.934 \leqq q \leqq 1.000$	5
12	3	3	3	4	4	3	4	4	4	4	4	4	$0.916 \leqq q \leqq 0.972$	5
	3	4	4	3	4	4	3	4	4	3	4	4	$0.972 \leqq q \leqq 1.000$	6

A Network for Combining Radio Systems at 4, 6 and 11 kmc

By EARL T. HARKLESS

(Manuscript received May 11, 1958)

Development of the broadband horn reflector antenna has permitted the simultaneous radiation and reception of radio signals on different frequencies in the three common-carrier bands in which the Bell System has developed radio relay systems. A necessary adjunct to the antenna is a network to combine or separate the common carrier bands and also to combine or separate the two polarizations of any one band. The particular form of the network that is described was designed to meet strict system requirements on impedance match, insertion loss and cross-coupling between ports.

Microwave radio relay systems have been developed for Bell Telephone System toll transmission in the common carrier frequency bands of 3700 to 4200 mc, 5925 to 6425 mc and 10,700 to 11,700 mc. There will be many routes where two, or all three, of these radio systems will be in use simultaneously. On such routes, a common antenna[1] at the top of the station tower will transmit or receive all three frequency bands. The broadband antennas will operate with both polarizations of the radio waves, and the feed line will be a 3-inch circular copper waveguide (WC 281). The 3-inch waveguide is capable of supporting three modes in the 4-kmc frequency range and 22 modes in the 11-kmc frequency range. With the possibility of so many propagating modes, small discontinuities in the internal dimensions of the waveguide may cause undesirable mode conversion with associated transmission loss or reflection of signal energy. In particular, any bend made in the circular waveguide must have very small curvature to minimize coupling between the two cross-polarized dominant modes at 11,000 mc.

The cross-coupling requirements mean that the circular waveguide must run nearly straight down the side of the tower to a network located at the base of the tower. After the signals are in dominant-mode waveguide, bends and twists can be utilized to bring the signal into the repeater stations. The systems combining network is required to distribute

the six signals (two polarizations at each of three frequency bands, 4, 6 and 11 kmc) to the proper ports with small insertion loss (approximately $\frac{1}{2}$ db attainable), with small reflections (30 db return loss) and with small delay distortion (echoes 60 db below the signal).

The basic configuration of this systems combining network was suggested by Miller,[2] and an experimental network was reported on by Dawson.[3] A series of directional couplers is used to successively distribute the different polarizations and frequency bands. The circular waveguide from the antenna is tapered down to a dominant-mode square waveguide in the 3700- to 4200-mc band. The directional couplers are constructed by running rectangular dominant-mode waveguide parallel to the two-polarization square waveguide, with the narrow face of the rectangular waveguide adjacent to the two-polarization waveguide (see Fig. 1). Longitudinal slots cut in the common wall will couple to only

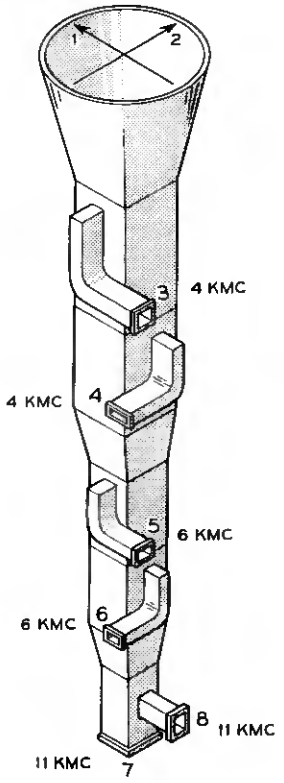


Fig. 1 — Designation of dominant-mode terminal pairs for systems combining network.

one polarization of the 4-kmc signal in the two-polarization waveguide, and a sufficient number of slots will be used to extract all the energy in this polarization. (The frequency-selection means will be discussed later.) A similar directional coupler rotated by 90 degrees is connected below the first coupler and couples to the other polarization of 4 kmc. All the 4-kmc energy in both polarizations has now been extracted, so the two-polarization waveguide is tapered down to square dominant-mode size at 6 kmc. Two more directional couplers with longitudinal slots are connected to extract the two polarizations at 6 kmc. The two-polarization waveguide is again tapered down in size, and a network to combine both polarizations at 11 kmc is connected.

All the directional couplers are reciprocal networks, so the structure can be used to combine the frequencies and polarizations into a single waveguide or to separate the signals from a single waveguide into their respective ports — or various combinations of combining and separating (transmitting and receiving) can be achieved.

The two directional couplers extracting the 4 kmc energy must, of course, not produce appreciable mode conversion of 4-, 6- or 11-kmc energy. Any size waveguide chosen as the two-polarization waveguide for the 4 kmc couplers will propagate several higher-order modes with cutoffs below 11,700 mc. It has been found that the most troublesome regions for higher-order mode couplings are in the vicinity of the mode cutoffs. For this reason, it is desirable to have all the mode cutoffs as far away as possible from the frequency bands 3700 to 4200 mc, 5925 to 6425 mc and 10,700 to 11,700 mc. This selection of mode-cutoff patterns leads to a very restricted choice of waveguide sizes, and results in dominant mode cutoff for the two-polarization waveguide (3300 mc) being close to the lower edge (3700 mc) of the 4-kmc band.

Fig. 2 graphically illustrates the relationship between the mode cutoff pattern and the common-carrier frequency bands. It is clear that the waveguide size or mode cutoff pattern cannot be shifted very much in either direction. The use of longitudinal slots (to separate polarizations) and waveguide operating close to cutoff (to avoid couplings to higher-

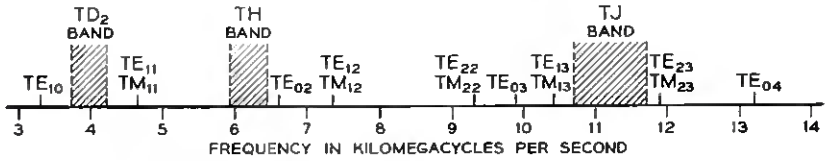


Fig. 2 — Distribution of mode cutoffs with relation to common carrier bands for square waveguide 1.790 inches on a side.

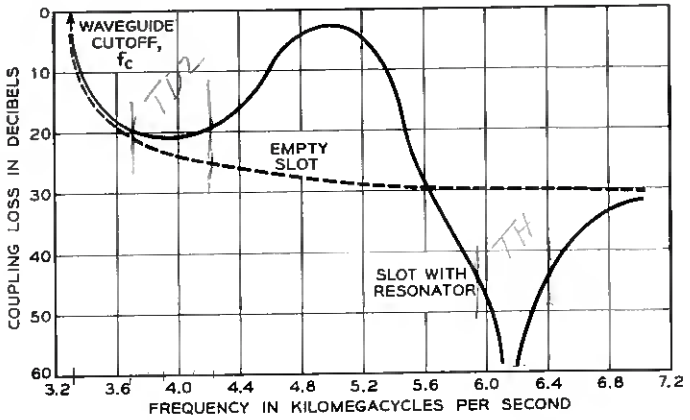


Fig. 3 — Transfer of energy between guides for single slot.

order modes) results in the coupling-versus-frequency characteristic shown in Fig. 3.

Using procedures described by Bethe[4] and Surdin,[5] we have equations for the amplitude of the wave coupled through a small slot in the side wall of a waveguide and the power carried by waves of unit amplitude in each waveguide:

$$A_1 = j\frac{2\pi}{\lambda_0 S_1} M H_{1z} H_{2z},$$

$$S_1 = \frac{a_1 b_1 \lambda_{g1}}{2\lambda_0},$$

$$S_2 = \frac{a_2 b_2 \lambda_{g2}}{2\lambda_0},$$

where

A_1 is the amplitude of the coupled wave,
$S_1 A_1^2$ is the power carried by a wave in waveguide 1,
S_2 is the power carried by a unit-amplitude wave in waveguide 2,
λ_0 is free-space wavelength,
λ_g is guide wavelength,
H_z is the magnetic field at the coupling slot for a wave of "unit" amplitude,
a and b are waveguide width and height and
M is the magnetic polarizability of the slot in the z direction.

The coefficient of coupling, c, is the square root of the ratio of the power excited in waveguide 1 to the exciting power in waveguide 2:

$$c = \frac{A_1\sqrt{S_1}}{\sqrt{S_2}} = M\frac{\pi}{2}\sqrt{\frac{\lambda_{g1}\lambda_{g2}}{a_1^3 b_1 a_2^3 b_2}}. \tag{1}$$

This expression shows that energy coupled between waveguides by a single slot decreases as the frequency is increased across the 4-kmc band.

For an array of coupling apertures, Miller[2] has shown that the amplitude of the wave coupled between two waveguides is given by

$$E_2 = \frac{j}{\sqrt{\left(\frac{B_1 - B_2}{2c}\right)^2 + 1}} \sin\left[nc\sqrt{\left(\frac{B_1 - B_2}{2c}\right)^2 + 1}\right], \tag{2}$$

where n is the number of apertures and B_1 and B_2 are the phase velocities in the coupled guides.

Assuming that $B_1 = B_2$, it is seen that the number of coupling slots, n, can be adjusted to give complete power transfer ($E_2 = 1$) for a particular c. However, c is a function of frequency and, as the frequency is increased, the number of slots will be insufficient to obtain complete power transfer, while at lower frequencies the number of coupling slots will be too large to obtain complete power transfer.

For the 4-kmc coupler, the two-polarization waveguide can be chosen to be square waveguide 1.790 inches on a side. The dominant mode cuts off at 3300 mc, and there are no modes with cutoffs in the 4-, 6- or 11-kmc bands. A coupling interval of four wavelengths was chosen to obtain complete power transfer from the square waveguide to a parallel rectangular waveguide ($n = 16$). The coupling obtained in each quarter wavelength interval must be equal[2] to $(\pi/2)/16 = 0.09830$. A slot 1.215 inches long in the sidewall of a rectangular waveguide 1.752 by 0.872 inches that couples to a 1.790-inch square waveguide will yield a coupling coefficient $c = 0.09830$ at 3870 mc. This calculation assumes that we have adjusted B_1 to be equal to B_2. (The method for doing this will be discussed later.) For 16 such slots, the amplitude of the wave emerging from port 3 of Fig. 4 when we excite port 1 is given by

$$E_1 = \cos nc\sqrt{\left(\frac{B_1 - B_2}{2c}\right)^2 + 1} - j\frac{B_1 - B_2}{2c}\frac{\sin\left[nc\sqrt{\left(\frac{B_1 - B_2}{2c}\right)^2 + 1}\right]}{\sqrt{\left(\frac{B_1 - B_2}{2c}\right)^2 + 1}}. \tag{3}$$

For the slot 1.215 inches long, we find at 3700 mc:

$$\begin{aligned} c &= 0.1203, \\ E_2 &= 0.940, \\ E_1 &= 0.342. \end{aligned}$$

At 4200 mc:

$$\begin{aligned} c &= 0.0755, \\ E_2 &= 0.940, \\ E_1 &= 0.355. \end{aligned}$$

Translating E_1 into uncoupled energy, we get the curve of Fig. 5, which shows that the insertion loss straight through the square waveguide (port 1 to port 3 of Fig. 4) is only about 10 db at the band edges.

These −10-db signals pass through the second coupler for the other polarization of 4 kmc with small loss and are completely reflected by the taper to 6-kmc dominant-mode two-polarization waveguide. The reflected signal suffers another 10-db loss in returning through the first coupler for 4 kmc, and emerges as a reflected signal that is only 20 db down. This is far from the objective of 30- to 35-db return loss. The above analysis leads to abandonment of the simple slot as a suitable element in the construction of directional couplers for systems combining networks.

A method has been found to modify the shape of the longitudinal

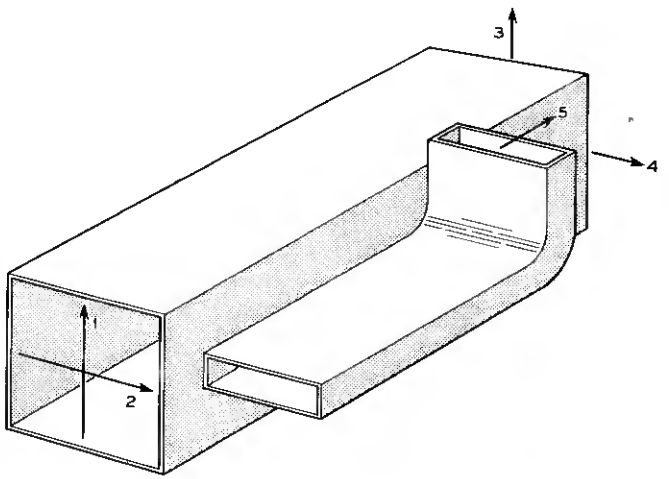


Fig. 4 — Accessible terminal pairs for directional couplers used in systems combining networks.

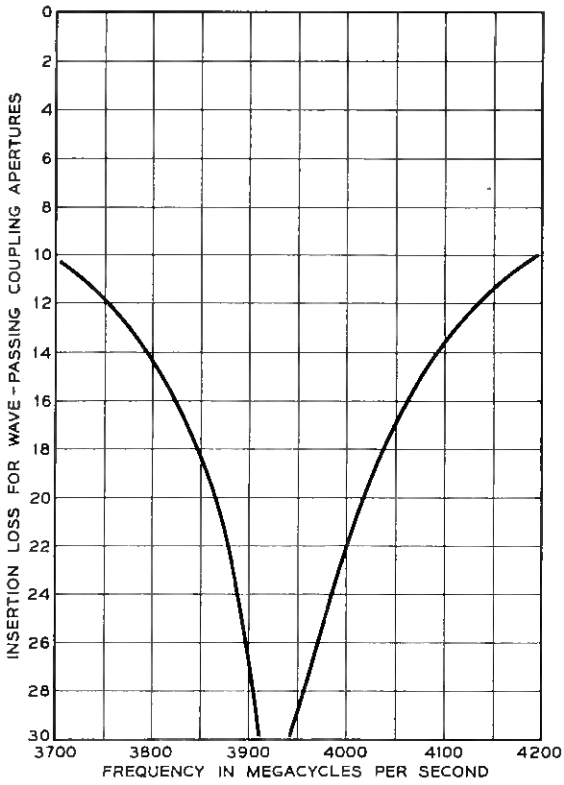


Fig. 5 — Insertion loss for wave passing coupling apertures.

slot to flatten the coupling-versus-frequency characteristic and reduce the amount of uncoupled energy. This structure is shown on Fig. 6. Placing the short piece of wire in the slot results in the coupling-loss-versus-frequency characteristic shown in Fig. 3. The wire is made approximately a quarter-wavelength long at 6 kmc, and results in a rejection peak at this frequency. The trial of a network element consisting of a wire resonator in the slot was suggested by viewing the coupling slot by itself as a short section of waveguide extending through the coupling plate. A resonant element placed in this approximately rectangular section of very short waveguide should be able to stop transmission at a particular desired frequency.

The frequency of the rejection peak is controlled by the total length of the wire while the width of the rejection peak is controlled by the location of the point at which the wire is bent and attached to the side

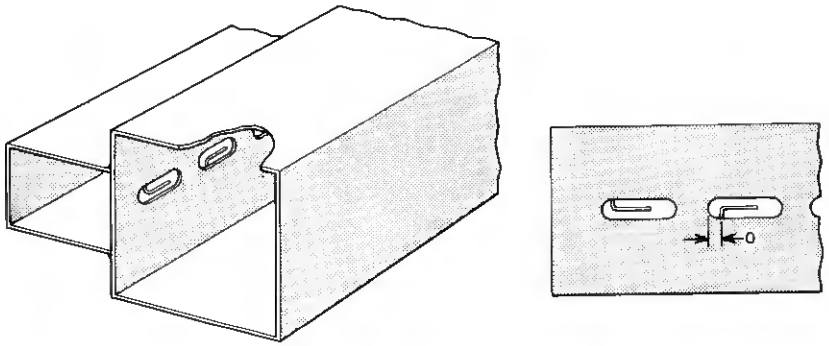
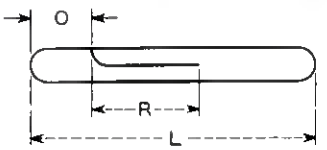


Fig. 6 — Directional coupler structure with modified slots.

of the slot (dimension O). Moving the point of attachment of the wire toward the center of the slot widens the rejection peak, but the left end of the curve, f_c, is determined by the waveguide size. It is seen that, as the width of the rejection peak is varied (wire offset changed), the slope of the coupling across the 4-kmc band can be changed. The width of the rejection peak is adjusted to produce a minimum of coupling near the center of the 4-kmc band. For each frequency in the lower part of the band, there will then be a frequency with the same coupling amplitude in the upper part of the band. With a series of such coupling elements, correct coupling amplitude for complete power transfer can be obtained at two frequencies in the 4-kmc band. Fig. 7 gives some data on the relation between resonator offset and coupling characteristics. The uncoupled energy at the two ends and the center of the 4-kmc band can be reduced to about 20 db, resulting in a return loss limitation of about 40 db due to the uncoupled energy.

The phase velocities in both the square waveguide and the rectangular waveguide are modified by the presence of the coupling slots. The amplitude and phase of the wave scattered in the forward direction in each waveguide can be calculated from (1). This scattered wave is added to the incident wave to obtain the phase shift caused by each coupling aperture. In the Appendix it is shown that an aperture-loaded rectangular 1.752- by 0.872-inch waveguide will produce the desired phase-velocity match with a square 1.790-inch waveguide near the center of the 4-kmc band.

The measured characteristics of a slot with resonator are shown in Fig. 8. The dimensions of the slot, resonator and offset have been experimentally adjusted to give the same coupling at the two edges of the



CURVE	DIMENSIONS		
	O	R	L
A	0.193	0.396	0.976
B	0.195	0.396	0.978
C	0.199	0.396	0.982

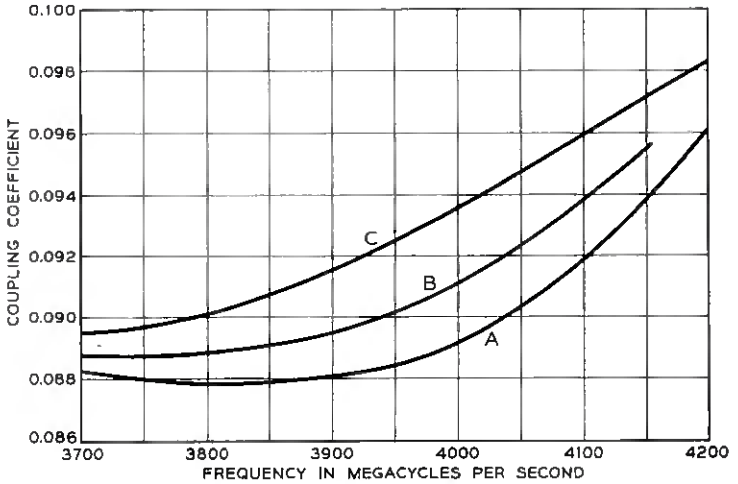


Fig. 7 — Effect of changing offset on coupling in 4-kmc band.

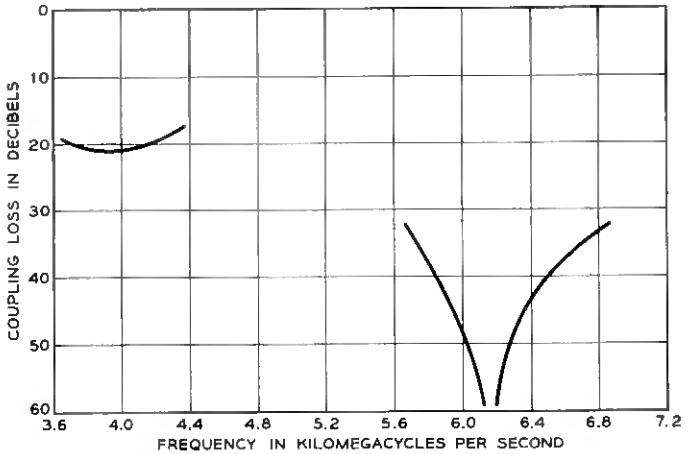


Fig. 8 — Measured response of single slot with resonator.

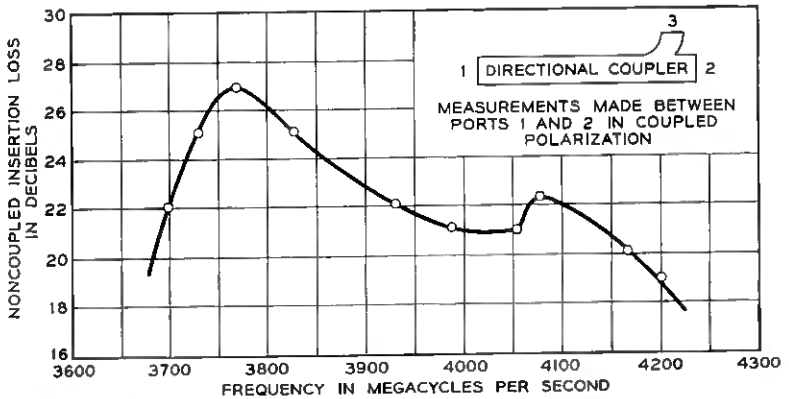


Fig. 9 — Straight-through loss for TD-2 band coupler.

4-kmc band, and a combination of 16 such slots has been made to give the coupling required for complete power transfer ($\sin nc = \pi/2$) at two points in the band.

The effect of the slot loading on the phase velocities in the square and rectangular waveguides is such that the phase velocities in the two waveguides can be exactly matched at only one frequency (approximately midband). At the two frequencies where the coupling is correct for complete power transfer, the phase velocity mismatch will limit the power transfer, so that the straight-through loss from port 3 to port 1 in Fig. 4 is only about 26 db. Fig. 9 shows the straight-through loss for a coupler with 16 coupling elements of the type shown in Fig. 8. The variation in coupling loss with frequency and the mismatch of phase velocity in the two waveguides both contribute to produce a straight-through loss of 19.5 db at the band edges and 21 db at the band center. The straight-through loss is now high enough to enable the realization of 30- to 35-db return loss on production models when other small reflections are included.

The resonant wires not only flatten the coupling characteristic in the 4-kmc band; they also reduce the coupling coefficients of the slots over the 6- and 11-kmc bands to such a small amount that very little energy (a few tenths of a db) is lost when these frequencies pass through the 4-kmc couplers. The only remaining deficiency of the network now is the return-loss performance at 11 kmc.

The slots in the 4-kmc coupler were spaced 1.440 inches apart, which is $\frac{5}{4}$ guide wavelengths at 10,740 mc, so the reflections should cancel at this frequency. However, the return-loss characteristic reached almost 23 db near 11 kmc. This was increased to over 30 db by putting in another

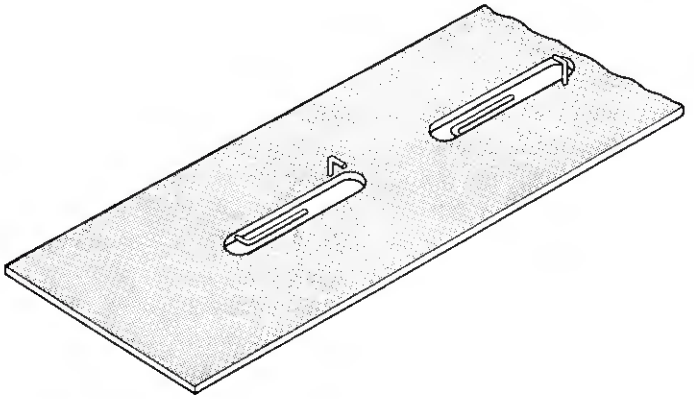


Fig. 10 — Coupling plate with added set of 11-kmc resonators.

set of wire resonators tuned to 11 kmc. There was insufficient room to place both the 6-kmc and 11-kmc resonators entirely within the slot, so the 11-kmc resonator was soldered in perpendicular to the plane of the coupling plate and it projected into the rectangular guide as shown in Fig. 10.

A similar design was produced for the 6-kmc directional coupler to be used in the systems combining network. In the 6-kmc version, it was found necessary to place two resonant wires in each slot, as shown in Fig. 11, in order to obtain adequate suppression of the 11-kmc energy. There was ample room for this, since the slot lengths in the 6-kmc directional coupler are considerably greater than a half wavelength at 11

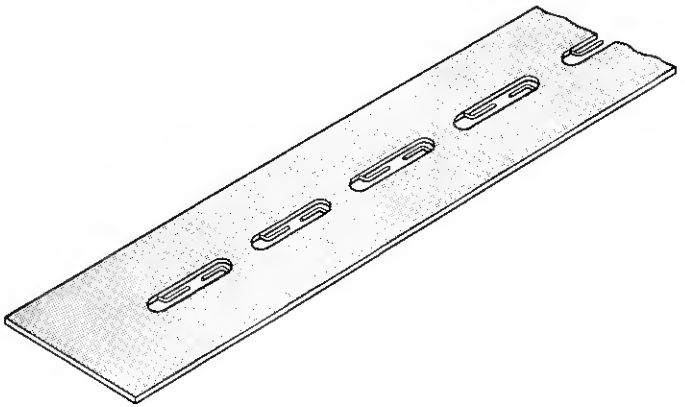


Fig. 11 — Coupling plate for 6-kmc coupler.

kmc. The energy transfer in the 6-kmc band was complete enough, so the straight-through loss exceeded 22 db over the band and the return loss due to this factor was 44 db.

Separation of the 11-kmc polarizations is accomplished by means of a polarization separation network previously described by Ohm.[6]

The complete systems combining network for both polarizations of all three common carrier frequency bands has eight accessible ports and 36 independent elements in the scattering matrix, which defines network performance. Fig. 1 illustrates the numbering of the ports, with port 1 and port 2 being the two dominant-mode polarizations in the square or circular waveguide. Table I summarizes the measurements made of the magnitude of the most important scattering elements in the three frequency bands of interest. The loss in the various desired signal paths is seen to be about $\frac{1}{2}$ db, and the impedance match at all ports is seen to exceed 30-db return loss. The cross-couplings in the various undesired signal paths are seen to range from 18 db to higher losses.

A photograph of a laboratory model of the systems combining network

TABLE I — MAGNITUDE OF SCATTERING COEFFICIENTS FOR COMPLETE SYSTEMS COMBINING NETWORK (WORST PERFORMANCE AT ANY POINT IN THE BAND)

Element (see Fig. 1)	Measurement	Decibels Loss at		
		4 kmc	6 kmc	11 kmc
S11	Return Loss	33	30	30
S33	Return Loss	35	—	—
S55	Return Loss	*	38	—
S77	Return Loss	*	*	30
S88	Return Loss	*	*	30
S13	Insertion Loss	0.4	20	28
S14	Insertion Loss	40	50	40
S15	Insertion Loss	*	0.6	18
S16	Insertion Loss	*	33	40
S17	Insertion Loss	*	*	0.5
S18	Insertion Loss	*	*	40
S25	Insertion Loss	*	0.5	18
S34	Insertion Loss	50	—	—
S35	Insertion Loss	*	45	—
S36	Insertion Loss	*	55	—
S37	Insertion Loss	*	*	40
S38	Insertion Loss	*	*	30
S56	Insertion Loss	*	50	—
S57	Insertion Loss	*	*	35
S58	Insertion Loss	*	*	36
S78	Insertion Loss	*	*	50

* means the guide is beyond cutoff for this frequency.
— means the coefficient is of no interest at this frequency.

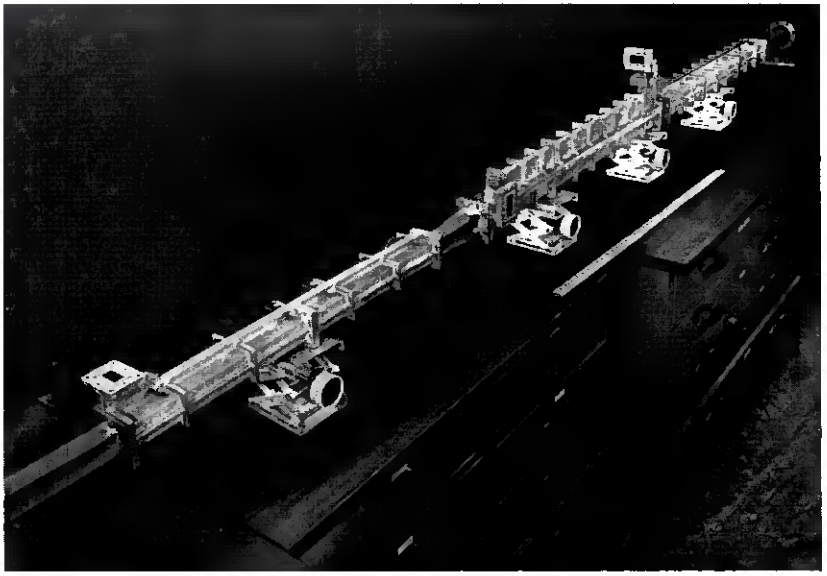


Fig. 12 — Laboratory model of systems combining network.

for separating polarizations of the 4-kmc and 6-kmc bands is shown in Fig. 12. The total length for combining both polarizations of all three frequency bands is 15 feet $10\frac{1}{2}$ inches.

There will be many repeater installations where all three frequency bands in both polarizations will not be used. In these cases, less than a full complement of directional couplers can be used. It is expected that there will be a large number of stations at which the horn reflector antennas will be fed with one polarization of 4 kmc and two polarizations of 6 kmc. The systems combining networks then need consist of only a circular-to-square taper, one 4-kmc directional coupler, a taper to 6-kmc dominant-mode square waveguide, one 6-kmc directional coupler and a transition from square to rectangular guide for the remaining 6-kmc polarization. The over-all length for this version of the systems combining network is less than 9 feet. In the case where only one polarization of each band is used, the total length would be 9 feet 1 inch.

CONCLUSION

A method has been developed for designing a systems combining network which separates by polarization and frequency the signals received

by a common antenna for the 4000-mc TD-2, the 6000-mc TH and the 11,000-mc TJ radio relay systems. The design results in a theoretical reflected-signal limitation of about 40 db, while production models realize well over 30 db return loss. Insertion loss in the desired signal paths is about $\frac{1}{2}$ db, while couplings between various other ports range from 18 db to over 50 db. The structure is mechanically simple and can be fabricated to the required tolerances by conventional machining techniques. The same principles can be utilized for other frequency bands. However, new designs must be based on experimentally determined characteristics for the slots and resonators.

ACKNOWLEDGMENTS

I would like to acknowledge the alert assistance of R. Vincent and C. N. Tanga in adjusting the performance of the directional couplers and in gathering much of the data presented here.

APPENDIX

The amplitude of the wave coupled between two guides has been computed by means of (1) for the coupling coefficient

$$c = M \frac{\pi}{2} \sqrt{\frac{\lambda_{g1}\lambda_{g2}}{a_1^3 b_1 a_2^3 b_2}}. \tag{1}$$

The amplitude of the wave scattered forward in guide 1 due to a wave incident in guide 1 can be calculated from the equation

$$c_{11} = M \frac{\pi}{2} \sqrt{\frac{\lambda_{g1}\lambda_{g1}}{a_1^3 b_1 a_1^3 b_1}} = M \frac{\pi}{2} \frac{\lambda_{g1}}{a_1^3 b_1},$$

and, of course, for guide 2,

$$c_{22} = M \frac{\pi}{2} \frac{\lambda_{g2}}{a_2^3 b_2}.$$

With c in coupling per unit length, the corrected phase shift per unit length is given by

$$\beta_{11}' = c_{11} + \beta_{11},$$

where β_{11}' is the corrected phase shift per unit length, and β_{11} is the unloaded guide phase shift per unit length. At 3900 mc, we have set $c = 0.0682$ per inch. Then, with $a_1 = b_1 = 1.790$ inches, $a_2 = 1.752$ inches and $b_2 = 0.872$ inch, we have

$$\beta_{11}' = 1.158 + 0.0682 \sqrt{\left(\frac{1.752}{1.790}\right)^3 \left(\frac{0.872}{1.790}\right)\left(\frac{0.5227}{0.5508}\right)}$$

$$= 1.158 + 0.045$$

$$= 1.203 \text{ radians per inch,}$$

$$\beta_{22}' = 1.099 + 0.0682 \sqrt{\left(\frac{1.790}{1.752}\right)^3 \left(\frac{1.790}{0.872}\right)\left(\frac{0.5508}{0.5227}\right)}$$

$$= 1.099 + 0.104$$

$$= 1.203 \text{ radians per inch.}$$

The 1.790-inch square guide and the 1.752- by 0.872-inch rectangular guide will now have the same phase velocity at 3900 mc, when they are loaded by slots that have a coupling coefficient equal to 0.0682 per inch.

REFERENCES

1. Friis, R. W. and May, A. S., A New Broad-Band Microwave Antenna System, Comm. & Elect., No. 35, March 1958, p. 97.
2. Miller, S. E., Coupled Wave Theory and Waveguide Applications, B.S.T.J., **33,** May 1954, p. 661.
3. Dawson, R. W., An Experimental Dual Polarization Antenna Feed for Three Radio Relay Bands, B.S.T.J., **36,** March 1957, p. 391.
4. Bethe, H. A., Theory of Diffraction by Small Holes, Phys. Rev., **66,** October 1944, p. 163.
5. Surdin, M., Directive Couplers in Wave Guides, J.I.E.E., **93,** Part IIIA, 1946, p. 725.
6. Ohm, E. A., A Broadband Microwave Circulator, Trans. I.R.E., **MTT-4,** October 1956, p. 215.

Synthesis of Active *RC* Networks

By B. K. KINARIWALA

(Manuscript received May 28, 1959)

A basic theorem is derived for RC networks containing active elements. It is shown that no more than one active element, embedded in a passive RC network, is needed to realize any driving-point function. Sufficiency of only one active element is shown by developing a synthesis method. A synthesis technique for n-port passive RC networks is developed in order to establish the sufficiency proof of the basic theorem. A more practical method of realizing driving-point functions, using active RC networks, termed the "cascade" method, is also presented. This method is applied to the design of a tenth-order Tchebycheff parameter filter.

I. INTRODUCTION

Passive filters using only resistive and capacitive elements are attractive for reasons of size, cost and reliability. Their use has been limited because the network complexity of *RC* filters (due to restrictions on impedance functions realizable with *R*'s and *C*'s only) is greater than that of equivalent *RLC* filters. This defect can be overcome by using active elements in addition to passive *RC* networks. Active *RC* networks are particularly attractive for achieving exacting performance at low frequencies, where it is not practical to use either inductors or crystals.

The practicality of active networks is a direct result of the availability of precision resistors and capacitors of small size having small drifts and low temperature coefficients, as well as the development of reliable junction transistors. In fact, it is not unusual to find that the active element's drift with time and temperature is no worse than that of a passive element.

There are several techniques available for synthesis of transfer functions by active *RC* networks.[1,2,3,4,5] The active element used is either a stabilized high-gain feedback amplifier or a negative-impedance converter. With the feedback amplifier, one *RC* network is used to produce the zeros of the transfer function and another *RC* network produces the poles of the function. The number of passive elements required in this

case tends to be rather large. The use of a converter usually leads to bridged–T, twin–T or parallel ladder networks when there are finite frequencies of infinite attenuation. There are several modifications of these basic synthesis methods which yield simpler networks. However, the emphasis in all of these methods is on the realization of transfer functions.

This paper presents some theoretical and practical results in the synthesis of active *RC* driving-point functions. Obviously, driving-point impedances can be used to yield any desired transfer characteristic. A theoretical investigation is first undertaken to show that no more than one active element (embedded in an *RC* network) is required to yield any desired driving-point function. The proof of this basic theorem is presented in Section II. Sufficiency of only one active element is shown by an actual synthesis method. This method involves realization of n-port passive *RC* networks. The results used in Section II for this purpose are drawn from material presented in the Appendices.

A more practical method of synthesizing driving-point functions is presented in Section III. This is termed the "cascade" method, since it involves cascading of a passive *RC* network with another passive *RC* network through a negative impedance converter. Network functions of rank 2 lead to particularly simple structures. Such functions are discussed in Section IV. The results obtained in this section are applied to the design of a tenth-order Tchebycheff parameter filter in Section V. The concluding section summarizes the results and discusses the outstanding problems as well as suggests some new avenues of approach.

II. BASIC THEOREM

In this section it will be proved that any driving-point immittance function* can be realized by a transformerless *RC* structure in which is embedded only one active element. The active element may be represented as a "controlled" source or as an amplifier. In any case, the active element is a two-port transducer. These two ports may be considered as external ports which are connected through an active element.

The remaining network is now purely passive. Thus, a driving-point immittance can be represented as a three-port passive *RC* network with two of the ports connected through an active element (Fig. 1). The proof of the basic theorem is then accomplished in two parts:

* In this paper, an immittance function is assumed expressible as a ratio of two polynomials in the complex frequency variable, $p = \sigma + j\omega$. The only restriction on these polynomials, except where other restrictions are specifically introduced, is that they have real coefficients.

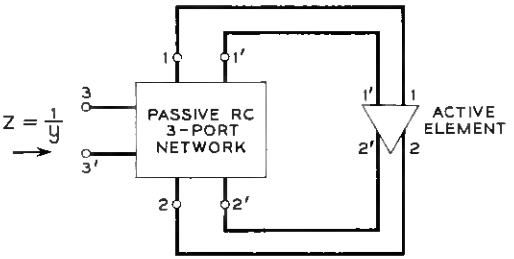


Fig. 1 — Realization of driving-point function with one active element.

i. Determination of the three-port *RC* matrix from the specified immittance function.

ii. Realization of the three-port *RC* network without transformers.

2.1 *Determination of Three-Port RC Matrix*

A driving-point admittance function $y(p)$, where $p = \sigma + j\omega$, is the complex frequency variable, is to be realized by the structure shown in Fig. 1. The admittance function is chosen because it leads to a three-port short-circuit admittance matrix, Y, which is convenient to use in the realization of the desired network.

In the figure, y is the admittance as seen from the terminals 3-3′, μ is the gain of the active element and T_{12} is the current transfer function of the passive *RC* network between terminals 1-1′ and 2-2′. Then, by Blackman's equation,[6]

$$y = y_{33}\left(\frac{1 - \mu T_{12}{}^{0}}{1 - \mu T_{12}{}^{\infty}}\right), \tag{1}$$

where y_{33} is the admittance seen at 3-3′ when the active element is placed in a reference condition of zero forward transmission, $\mu T_{12}{}^{0}$ is the loop-current transmission when zero admittance is introduced between terminals 3-3′ and $\mu T_{12}{}^{\infty}$ is the loop current transmission when infinite admittance is introduced between 3-3′.

For convenience and without loss of generality, it is assumed that the amplifier has zero input and output impedances. The amplifier is thus a current-controlled voltage source (Fig. 2). Thus, y_{33} is the short-circuit admittance at port 3-3′ and $\mu T_{12}{}^{\infty}$ is the loop current transmission when all ports are short-circuited, where

$$\mu T_{12}{}^{\infty} = -R_m y_{12}\,, \tag{2}$$

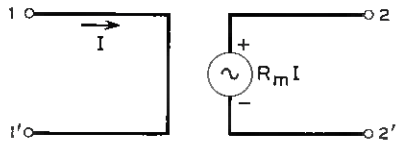


Fig. 2 — Current-controlled voltage source.

and μT_{12}^{0}, the loop current transmission when port 3-3′ is open-circuited, may be written as

$$\mu T_{12}^{0} = R_m \left[y_{13} \left(\frac{y_{23}}{y_{33}} \right) - y_{12} \right]. \tag{3}$$

Substituting (2) and (3) in (1),

$$y = y_{33} \left[\frac{1 - R_m \left(\frac{y_{13} y_{23}}{y_{33}} - y_{12} \right)}{1 + R_m y_{12}} \right]. \tag{4}$$

Let y be the admittance given as a ratio of two polynomials. Express the numerator and denominator as differences of two polynomials:

$$y = \frac{N}{D} = \frac{N_1 - N_2}{D_1 - D_2}, \tag{5}$$

The polynomials N_1 and D_1 must have only negative real roots. Further, N_1 and D_1 are so chosen that N_1/D_1 is an RC driving-point admittance. The reasons for these requirements will become clear later. It is also desired that the degree of N_2 does not exceed that of N_1 by more than one and that the degree of D_2 does not exceed that of D_1 by more than one. In Appendix A it is shown that, given N and D, it is always possible to find the polynomials N_1, N_2, D_1 and D_2 satisfying these requirements.

It is now possible to rewrite (5),

$$y = \frac{N_1}{D_1} \left(\frac{1 - \frac{N_2}{N_1}}{1 - \frac{D_2}{D_1}} \right) \tag{6}$$

and, comparing with (4),

$$y_{33} = \frac{N_1}{D_1},$$

$$-R_m y_{12} = \frac{D_2}{D_1},$$

$$R_m \left(\frac{y_{13} y_{23}}{y_{33}} - y_{12} \right) = \frac{N_2}{N_1} \tag{7}$$

or

$$-R_m y_{13} y_{23} = \frac{N_1 D_2 - N_2 D_1}{D_1^2}.$$

The previously mentioned requirements on the polynomials N_1, D_1, etc. are seen to be necessary if the identifications in (7) are to be made. The first equation in (7) gives the requirements on N_1/D_1, the second equation gives those on the degrees of D_2 and D_1, and so on. Since these requirements are already met, one obtains from (7) some of the elements of the short-circuit admittance matrix Y of the three-port *RC* network; y_{33} is completely determined, whereas y_{12} and the product $y_{13}y_{23}$ are determined within a constant multiplier.

Equation (7) gives all the constraints on the elements of the Y matrix imposed by the admittance function. Other elements of the matrix can be chosen arbitrarily. These remaining elements of the Y matrix, R_m and the identification of y_{13} and y_{23}, must be so determined that Y is realizable without transformers. In order to determine these quantities, it is necessary to have a knowledge of the conditions under which the network can be realized without transformers.

2.2 *Realization of Three-Port RC Networks*

Since synthesis of three-port networks is a special case of general n-port synthesis, a synthesis technique for n-port *RC* networks is developed below. The results obtained are then applied to the three-port case.

2.2.1 *Synthesis of n-Port RC Networks*

The synthesis method consists of generating an auxiliary $(m + n)$-port* matrix representing a purely resistive network. When this network is terminated at its m-ports in appropriate capacitors, the desired n-port *RC* network is obtained (Fig. 3). It is assumed that the terminations are unit capacitors. If desired, the values of these capacitors are changed by a simple scaling of impedances at the appropriate ports.

* m represents the number of capacitors in the *RC* network.

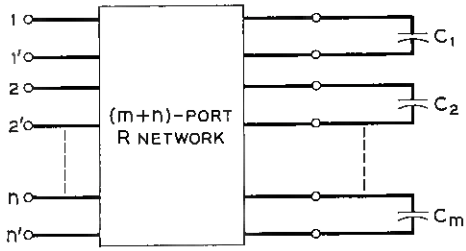


Fig. 3 — $(m + n)$-port resistance network terminated in m capacitors.

The short-circuit admittance matrix Y of the RC network may be expressed as

$$Y = Kp + K_\infty - \sum_\nu K_\nu \left(\frac{1}{p + \sigma_\nu}\right), \tag{8a}$$

where K, K_∞ and K_ν are $(n \times n)$ matrices of residues. It is shown in Appendix B that it is preferable, though not necessary, to have $K \equiv 0$; i.e., the matrix Y has no pole at $p = \infty$. Furthermore, it is always possible to determine Y from the given driving-point admittance such that there is no pole at $p = \infty$ (see Appendix A). Consequently, it is convenient to consider

$$Y = K_\infty - \sum_\nu K_\nu \left(\frac{1}{p + \sigma_\nu}\right). \tag{8b}$$

Let the short-circuit admittance matrix for the purely resistive $(m + n)$-port be

$$G = \begin{matrix} \\ n \\ m \end{matrix} \begin{matrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \\ n \quad m \end{matrix}, \tag{9}$$

where A_{ij} are matrices of appropriate order. It is desired to determine G such that, when the m-ports are terminated in unit capacitors, the matrix Y is obtained. It is shown in Appendix B that

$$\begin{aligned} A_{11} &= K_\infty, \\ A_{12} &= M, \qquad A_{21} = \bar{M} = \text{transpose of } M \\ A_{22} &= S. \end{aligned} \tag{10}$$

where

$$M = [M_1, M_2, \quad], \tag{11}$$

and M_ν is defined by

$$K_\nu = M_\nu \bar{M}_\nu ; \tag{12}$$

M_ν has n rows and m_ν columns. Thus, M has n rows and $m = \sum_\nu m_\nu$ columns. It is seen from (9) that m determines the number of capacitors. S is a diagonal matrix defined by

$$S = \begin{bmatrix} \sigma_1 I_1 & 0 & 0 & 0 & & & 0 \\ 0 & \sigma_2 I_2 & 0 & 0 & & & 0 \\ 0 & 0 & \sigma_3 I_3 & 0 & & & 0 \\ 0 & 0 & 0 & & & & 0 \\ & & & & & & 0 \\ & & & & & & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \end{bmatrix}, \tag{13}$$

where I_j is a unit matrix of order m_j.

If it is desirable to obtain a network with a minimum number of capacitors, it is necessary to obtain a minimum m. This is accomplished by having each m_ν a minimum number. A minimum m_ν is obtained if

$$M_\nu = [M_{\nu 1}, M_{\nu 2}, M_{\nu 3}, \cdots], \tag{14}$$

where $M_{\nu i}$ are determined by expressing the quadratic form of K_ν as sum of squares,

$$\bar{X} K_\nu X = \sum_i (X M_{\nu i})^2. \tag{15}$$

Each $M_{\nu i}$ is a column vector. It is seen that, if K_ν is of order n and rank δ_ν, then M_ν has n rows and $m_\nu = \delta_\nu$ columns. The matrix M is then obtained by using (11). The minimum number m of capacitors is seen to be the sum of the ranks of matrices of residues in the finite poles.

The $(m + n)$-port G matrix is thus determined by (9) and (10). The $(m + n)$-port resistance network must be terminated at its m-ports in unit capacitors to obtain the desired n-port RC network.

If the matrix G can now be realized without ideal transformers, then terminating the resistance network in capacitors at appropriate ports will yield the RC network without transformers. Synthesis of resistance networks is considered below.

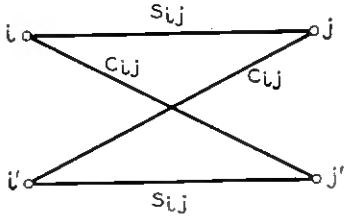


Fig. 4 — A typical unit in the realization of n-port resistance network.

2.2.2 *Synthesis of n-Port R Networks*

It is generally known and easy to show that a resistance network need have no internal nodes.* It follows that an n-port R network can always be realized as a completely connected graph having only the $2n$ external nodes. Necessary and sufficient conditions for the general n-port R network are not yet obtained. However, conditions for a restricted class of networks are obtained here.

Consider a symmetrical balanced structure constructed by introducing symmetrical lattices between each set of node-pairs (Fig. 4). If all node-pairs are short-circuited and a one-volt signal is introduced between one pair of terminals (say i-i'), then from the balanced structure it is seen that the current drawn by the network is determined solely by structures directly connected to terminals i-i'. Short-circuit admittance seen at i-i' is the sum of the short-circuit admittances of the lattices connected directly to i-i'. Furthermore, the short-circuit transfer admittances between terminals i-i' and j-j', for all j, are determined only by the lattice structure between each set of node-pairs. Thus,

$$g_{ii} = \sum_{j \neq i} A_{ij}, \tag{16}$$

where A_{ij} is the short-circuit admittance of the lattice between node-pairs i-i' and j-j', and g_{ii} is the short-circuit admittance seen at i-i'. Also,

$$g_{ij} = B_{ij},$$

where B_{ij} is the short-circuit transfer admittance of the lattice between node-pairs i-i' and j-j', and g_{ij} is the short-circuit transfer admittance between terminals i-i' and j-j'.

Since for any lattice $A_{ij} \geqq |\, B_{ij} \,|$, (16) may be written as

$$g_{ii} \geqq \sum_{j \neq i} |\, g_{ij} \,|. \tag{18}$$

* One way to show this is by repeated application of the results of Rosen.[7]

The conditions of (18) are shown to be necessary for the structure of Fig. 4. If any additional branches were placed between terminals i and i', the conditions of (18) would still be valid.

Sufficiency of the above conditions is proved by realization of the network as follows. If S_{ij} and C_{ij} are series and cross-arm admittances respectively of the i–j lattice,

$$\left.\begin{aligned} A_{ij} &= \tfrac{1}{2}\,(S_{ij} + C_{ij}) \\ g_{ij} &= B_{ij} = \tfrac{1}{2}\,(-S_{ij} + C_{ij}) \end{aligned}\right\}. \tag{19}$$

Let $A_{ij} = |\,B_{ij}\,| = |\,g_{ij}\,|$. Then,

$$\left.\begin{aligned} C_{ij} &= |\,g_{ij}\,| + g_{ij} \\ S_{ij} &= |\,g_{ij}\,| - g_{ij} \end{aligned}\right\}. \tag{20}$$

It is seen that C_{ij} and S_{ij} are non-negative, and therefore realizable. Any excess admittance determined by

$$g_{ii} - \sum_{j \neq i} |\,g_{ij}\,|$$

is then introduced between terminals i and i'. This proves the sufficiency of conditions (18). A matrix G that satisfies the conditions of (18) will be called a "dominant diagonal" or D matrix.

Necessary and sufficient conditions for the existence of balanced R network are that the short-circuit admittance matrix be a D matrix.

2.2.3 *Three-Port RC Networks*

In the case of three-ports for active networks under consideration, it will be shown that it is possible to find a matrix G which is a D matrix. It is also desirable to obtain the three-port RC network with a minimum number of capacitors.

If the case $n = 3$ is considered and the minimum number of capacitors is desired, then $m_\nu = 1$ and K_ν should be expressible as

$$K_\nu = M_\nu \bar{M}_\nu\,, \tag{21}$$

where M_ν is a matrix with three rows and one column. If $K_\nu = \| k_{ij}^{(\nu)} \|$ it is shown in Appendix C that, from (7),

$$k_{13}^{(\nu)} k_{23}^{(\nu)} - k_{12}^{(\nu)} k_{33}^{(\nu)} = 0, \qquad \text{for all } \nu. \tag{22}$$

Furthermore, since y_{11} and y_{22} are not specified, they may be chosen so that the K_ν, for all ν, are positive semidefinite matrices of rank 1. Hence,

the representation in (21) is possible with

$$K_\nu = \begin{bmatrix} \dfrac{k_{13}^{(\nu)2}}{k_{33}^{(\nu)}}, & k_{12}^{(\nu)}, & k_{13}^{(\nu)} \\ k_{12}^{(\nu)}, & \dfrac{k_{23}^{(\nu)2}}{k_{33}^{(\nu)}}, & k_{23}^{(\nu)} \\ k_{13}^{(\nu)}, & k_{23}^{(\nu)} & k_{33}^{(\nu)} \end{bmatrix}, \tag{23}$$

$$\bar{M}_\nu = \pm \frac{1}{\sqrt{k_{33}^{(\nu)}}} [k_{13}^{(\nu)}, k_{23}^{(\nu)}, k_{33}^{(\nu)}],$$

and relation (22) is assumed to be valid.

The elements of interest in G are

$$G = \left[\begin{array}{ccc|ccc} k_{11}^{\infty} & k_{12}^{\infty} & k_{13}^{\infty} & & & \\ k_{21}^{\infty} & k_{22}^{\infty} & k_{23}^{\infty} & & & \\ k_{31}^{\infty} & k_{32}^{\infty} & k_{33}^{\infty} & \pm\sqrt{k_{33}^{(1)}} & \pm\sqrt{k_{33}^{(2)}} & \sqrt{k_{33}^{(m)}} \\ & & & \sigma_1 & 0 & 0 \\ & & & 0 & \sigma_2 & 0 \\ \hline & & & 0 & 0 & \sigma_m \end{array}\right].$$

The first two rows of G can be made to satisfy the dominant diagonal conditions, since the diagonal terms may be chosen arbitrarily. The rows having σ_ν for diagonal terms can also be made to satisfy the same conditions by multiplying appropriate row and column by $\sqrt{C_\nu}$. This effectively changes the termination from unit capacitor to one having value C_ν. To make the third row satisfy the same conditions, R_m must be chosen large enough, thereby reducing the magnitudes of the residues of y_{13} and y_{23} [see (7)]. Further, the freedom in the choice of P_1 and Q_1 [see (5)] can be used to make the diagonal term in the third row greater than the sum of magnitudes of the remaining terms.

It is seen that the three-port RC matrices arising from the active networks under discussion can be realized without transformers.

A much simpler way of proving the basic theorem becomes obvious at this point. From discussion of n-port R networks it is seen that n-port RC networks are realizable if the matrices of residues in various poles and $Y(0)$ are all D matrices. For example, the Y matrix may be written as

$$Y = K_0 + \sum_\nu \frac{p}{p + \sigma_\nu} K_\nu', \tag{25}$$

where $K_0 = Y(0)$ and $K_\nu' = K_\nu/\sigma_\nu$. If K_0 is a D matrix, it is realized by resistors. Each remaining term $p(K_\nu')/p + \sigma_\nu$ is realized in the same manner as the R networks, except that each branch would be a resistor and a capacitor in series. The networks are all connected in parallel to obtain the n-port RC network.

From (7), it is seen that y_{11} and y_{22} are unconstrained and hence can be chosen such that the matrices K_0 and K_ν' have the diagonal terms in the first and second rows greater than the sum of the magnitudes of the off-diagonal terms of the corresponding rows. The third rows can be made to satisfy these conditions by an appropriate choice of R_m. A large enough value of R_m can always be chosen to make the residues of y_{13} and y_{23} sufficiently small [see (7)].

Thus, there are two alternative ways of proving the basic theorem. Comparing the two as synthesis methods, the second way is much simpler, but uses a rather large number of resistors and capacitors. The first method uses a minimum number of capacitors but still a large number of resistors. It also offers a choice as to the number of capacitors desired. However, it is not a practical method because of the large number of resistors required. The first method could turn out to be practical if it were possible to realize the R networks with much fewer resistors.

A much more practical way of designing active RC impedances is next undertaken.

III. CASCADE SYNTHESIS

In active networks, negative impedances are admissible and can be incorporated into a realization technique with relative ease. A possible way of realizing a negative impedance is with the help of a negative-impedance converter.[8] A negative-impedance converter is an active two-port which presents at either of its ports the negative of the impedance connected at the other port. Consider an impedance function

$$Z(p) = \frac{N}{D} = \frac{N_1 - N_2}{D_1 - D_2}, \tag{26}$$

where N and D are polynomials in p, whose degrees do not differ by more than one.

If it is possible to break up N and D into N_1, N_2 and D_1, D_2 as in (26) such that N_1, N_2, D_1 and D_2 all have only real and negative roots, then it might be possible to realize $Z(p)$ as a two-terminal-pair RC network terminated by either a negative resistance or a negative RC impedance. (Refer to Fig. 5.)

It is shown in Appendix A that any polynomial having real coefficients

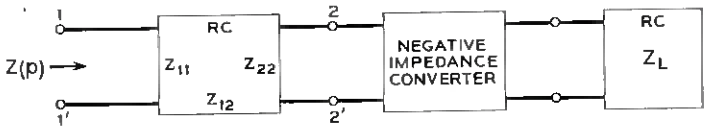


Fig. 5 — Realization of driving-point function by cascade method.

can be easily expressed as a difference of two polynomials each of which has only negative real roots. It is easy to show that an RC network terminated in a negative resistance cannot produce complex poles of the driving-point impedance function. The poles of $Z(p)$ are the zeros of $(z_{22} - R)$, where z_{22} is the open-circuit impedance of the RC two-port and $-R$ is the termination at end 2. It should be obvious that any resistance (positive or negative) added to an RC impedance merely shifts the zeros along the σ-axis without moving them off the axis. If the termination is $-Z_L$, a negative RC impedance, then the poles of $Z(p)$ are the zeros of $(z_{22} - Z_L)$. It is seen from Appendix A that a difference of two RC impedances can indeed produce the desired complex zeros.

3.1 *Realization Problem*

Consider an impedance function

$$Z(p) = \frac{N}{D} = \frac{\dfrac{N}{B}}{\dfrac{D}{B}}, \tag{27}$$

where B is a polynomial having only negative real roots. Then, as per Appendix A,

$$Z(p) = \frac{\dfrac{P_1}{P_2} - \dfrac{P_3}{P_4}}{\dfrac{Q_1}{P_2} - \dfrac{Q_3}{P_4}}, \qquad \text{where } P_2P_4 = B \tag{28a}$$

$$= \frac{P_1}{Q_1}\,\frac{\dfrac{P_3}{P_1} - \dfrac{P_4}{P_2}}{\dfrac{Q_3}{Q_1} - \dfrac{P_4}{P_2}}. \tag{28b}$$

For an RC two-port terminated in a negative impedance,

$$Z(p) = z_{11}\frac{\frac{1}{y_{22}} - Z_L}{z_{22} - Z_L}, \tag{29}$$

where z_{11} , z_{22} are the open-circuit impedances of the *RC* two-port, y_{22} is its short-circuit admittance and $-Z_L$ is the terminating impedance. Comparing (28) and (29):

$$\begin{aligned} z_{11} &= \frac{P_1}{Q_1}, \\ z_{22} &= \frac{Q_3}{Q_1}, \\ \frac{1}{y_{22}} &= \frac{P_3}{P_1}, \\ Z_L &= \frac{P_4}{P_2}. \end{aligned} \tag{30}$$

Without loss of generality, we will assume that $Z(p)$ has only complex zeros and poles. It should be noted that P_1 , P_3 , Q_1 , Q_3 have only negative real roots. As shown in Appendix A, Z_L is necessarily an *RC* impedance function and $-Z_L$ is realized by means of a negative-impedance converter. It remains to be shown that z_{11} , z_{22} , y_{22} represent a two-port *RC* network. From (30),

$$\frac{1}{y_{22}} = \frac{P_3}{P_1} = \frac{z_{11}z_{22} - z_{12}^2}{z_{11}}$$

and

$$z_{12} = \frac{\sqrt{P_1Q_3 - P_3Q_1}}{Q_1}. \tag{31}$$

Since z_{12} must be a rational function $(P_1Q_3 - P_3Q_1)$ must be a perfect square.

Then, to sum up, the necessary and sufficient conditions for physical realizability are that:

i. z_{11} and z_{22} represent *RC* driving-point impedances;

ii. $(P_1Q_3 - P_3Q_1)$ be a perfect square;

iii. residue condition be satisfied at all the poles.*

* For *RC* networks, the statement includes the point $p = \infty$, where the functions may be nonzero.

The condition iii is always satisfied. The residue condition is

$$k_{11}k_{22} - k_{12}^2 \geqq 0, \tag{32}$$

where k_{ij} represents the residue of z_{ij} in a given pole. The residues of z_{11} and z_{22} may be written as

$$\begin{aligned} k_{11}^{\,i} &= \frac{P_1(-\sigma_i)}{\dot{Q}_1(-\sigma_i)}, \\ k_{22}^{\,i} &= \frac{Q_2(-\sigma_i)}{\dot{Q}_1(-\sigma_i)}, \end{aligned} \tag{33}$$

where σ_i is any root of Q_1 and $\dot{Q}_1$ is the derivative of Q_1 with respect to p. The residue of z_{12} is

$$k_{12}^{\,i} = \frac{\sqrt{P_1(-\sigma_i)Q_2(-\sigma_i) - Q_1(-\sigma_i)P_2(-\sigma_i)}}{\dot{Q}_1(-\sigma_i)}. \tag{34}$$

But

$$Q_1(-\sigma_i) \equiv 0,$$

and so

$$k_{11}^{\,i}k_{22}^{\,i} - (k_{12}^{\,i})^2 \equiv 0. \tag{35}$$

Thus, the residue condition is always satisfied with an equal sign at all the poles. From (30) and (31) it is seen that the requirements at infinity are also met.

In reference to condition i, consider the location of the roots of P_1, P_2, etc. as shown in Fig. 6. As already pointed out, P_4/P_2 is an RC im-

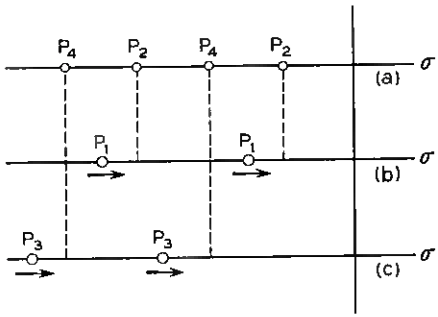


Fig. 6 — Adjustment of the roots of P_3 and P_1 so that P_3/P_1 represents an RC impedance function: (a) roots of P_2 and P_4; (b) roots of P_1; (c) roots of P_3.

pedance function. Hence it will have alternating roots in the correct order. This is shown as (a) in the figure. P_3/P_1 may or may not have the desired alternation. This can be corrected by adding the same constant to P_1/P_2 and P_3/P_4 [see (28a)]. Such addition of a constant will move the roots of P_1 closer to those of P_2 and move the roots of P_3 closer to those of P_4. This can always be done such that P_3/P_1 will have the desired alternation. In a similar manner, Q_3/Q_1 is corrected such that P_1/Q_1 also has the required form. In the case of an impedance Z of rank 2, the above corrections can generally be done so as to make $(P_1Q_3 - P_3Q_1)$ a perfect square. If and when this is not possible surplus factors might have to be added to $Z(p)$ to obtain a rational z_{12}.

It is not at all necessary to make all the ratios $(P_1/Q_1, Q_3/Q_1, P_3/P_1)$ have the desired alternation. It is obvious that, if $z_{11}(P_1/Q_1)$ and $z_{22}(Q_3/Q_1)$ are *RC* impedances and z_{12} is a rational function satisfying the residue conditions, $1/y_{22}$ or P_3/P_1 will necessarily have the proper form. Furthermore, if z_{12} is a ratio of two polynomials, the residues in its poles (real and negative) will be real and $(k_{12}{}^i)^2$ will always be positive. If z_{22} is an *RC* impedance function, its residues are necessarily positive. From (35), it is seen that z_{11} will automatically have positive residues; i.e., it must be an *RC* impedance. Thus, in order to realize the *RC* two-port one need only make z_{11} or z_{22} an *RC* impedance and $(P_1Q_3 - P_3Q_1)$ a complete square.

It is shown in Appendix D that, if Z has only complex poles, then z_{22} is always an *RC* impedance and remains so under the addition and subtraction of a constant in the denominator.* Thus, in such a case all that is needed is to make z_{12} a rational function. An illustrative example is considered to show how z_{12} can be made rational by the addition and subtraction of an appropriate constant in the denominator of Z.

Example:

$$Z(p) = \frac{p^2 + p + 1}{p^2 + p + 2} = \frac{\dfrac{p^2 + p + 1}{(p+1)(p+2)}}{\dfrac{p^2 + p + 2}{(p+1)(p+2)}}$$

$$= \frac{\dfrac{p+2}{p+1} - \dfrac{3}{p+2}}{\dfrac{p+3}{p+1} - \dfrac{4}{p+2}}$$

* This proof was developed by J. M. Sipress.

$$= \frac{p+2}{p+3} \frac{\dfrac{3}{p+2} - \dfrac{p+2}{p+1}}{\dfrac{4}{p+3} - \dfrac{p+2}{p+1}}.$$

In this case, $P_1Q_3 - P_3Q_1$ is to be made a perfect square. It is seen that Q_3/Q_1 has the proper alternation. Now, add a constant K to

$$(p+3)/(p+1)$$

and $4/(p+2)$. Then,

$$\begin{aligned} P_1Q_3 - P_3Q_1 &= (p+2)[4 + K(p+2)] - 3[K(p+1) + p + 3] \\ &= Kp^2 + (K+1)p + (K-1). \end{aligned}$$

For this to be a complete square,

$$3K^2 - 6K - 1 = 0$$

and

$$K = 1 + \frac{2}{\sqrt{3}}.$$

The open-circuit impedances of the two-port are

$$z_{11} = \frac{p+2}{(K+1)(p+1)+2},$$

$$z_{22} = \frac{K(p+2)+4}{(K+1)(p+1)+2},$$

$$z_{12} = \sqrt{K}\,\frac{p + \dfrac{K+1}{2K}}{(K+1)(p+1)+2}.$$

It is seen that these satisfy all the required conditions. It should be noted that adding the constant K to and subtracting it from the numerator alone instead of the denominator alone of $Z(p)$ will not yield a rational z_{12}. This is an obvious consequence of the fact that the introduction of K in the numerator alone is an attempt to make z_{12} nonvanishing at infinity, whereas z_{22} vanishes at infinity.

Realization of the RC two-port can be easily accomplished when there is only one term in the partial-fraction expansion of each z_{ik}.[9] Only one ideal transformer is present in such a case, and this is removed by means of impedance-scaling of the relevant part of the network. When several

poles are present in the impedances z_{ik}, such a realization procedure requires several ideal transformers, which cannot be eliminated in the above manner.

In such cases, one can make z_{11} and z_{12} satisfy the conditions of Fialkow and Gerst[10] by introducing one ideal transformer. The network is then realized without transformers and appropriately scaled to remove the ideal transformer that was introduced above. Such a realization would not give the desired z_{22}. To obtain this desired z_{22}, positive or negative elements are then added in series at the load end of the network. Negative elements are then realized together with the load through the negative-impedance converter.

3.2 *Surplus Factors*

As seen above, it is possible in simple cases to make $(P_1Q_3 - P_3Q_1)$ a perfect square by the use of an appropriate constant. However, this is not always possible. It will now be shown how $(P_1Q_3 - P_3Q_1)$ can always be made a perfect square by introducing surplus factors in the original expression for $Z(p)$. A new formulation is necessary so as to have z_{12} present explicitly in the expressions. Such a formulation can be obtained by writing $Z(p)$ in the following manner:

$$Z(p) = z_{11} - \frac{{z_{12}}^2}{z_{22} - Z_L} = \frac{N}{D}$$

$$= \frac{P_1}{Q_1} - \frac{\dfrac{g^2}{{Q_1}^2}}{\dfrac{Q_3}{Q_1} - \dfrac{P_4}{P_2}}$$

[from (30) and letting $z_{12} = g/Q_1$]

$$= \frac{P_1}{Q_1} - \frac{\dfrac{g^2}{{Q_1}^2}}{\dfrac{-D}{Q_1P_2}}. \tag{36}$$

This equation follows because $(-D = P_2Q_3 - Q_1P_4)$, as can be seen from (28a). Then,

$$Z(p) = \frac{\dfrac{DP_1}{D_1} + \dfrac{g^2P_2}{Q_1}}{D}$$

and

$$N = \frac{DP_1 + g^2P_2}{Q_1}$$

or

$$g^2P_2 = NQ_1 - DP_1 . \tag{37}$$

This equation must be subject to the constraints that $(P_1/Q_1)(z_{11})$ be an RC impedance. A further constraint requires Q_1/P_2 to be also an RC impedance. This is seen to be satisfied in (28a) since roots of P_2 and P_4 alternate. From a different viewpoint, it is seen from (36) that $-D/Q_1P_2$ must give a difference of two RC impedances having Q_1 and P_2 for their denominators. So the roots of Q_1 and P_2 must alternate. The sign of (Q_3/Q_1) is positive, and so the nearest singularity must be that of P_2. Therefore, Q_1/P_2 must be an RC impedance. If (37) is satisfied and P_1/Q_1 and Q_1/P_2 are RC impedances, then z_{11}, z_{22} and Z_L will all be RC impedances. If g is also a polynomial, then the impedance Z is physically realizable. The conditions for physical realizability can then be written as

$$g^2P_2 = NQ_1 - DP_1$$

and (38)

$$\left.\begin{array}{l} \dfrac{P_1}{Q_1} \\ \\ \dfrac{Q_1}{P_2} \end{array}\right\} \text{ are } RC \text{ impedances.}$$

It is important to note that (37) contains N and D, which are already known. Choice of P_2 is arbitrary. Freedom in the choice of P_4 is not explicit in (37). However, this freedom can be assigned to Q_1, since $-D/Q_1P_2$ gives z_{22} and Z_L. Thus, Q_1 and P_2 can be considered to be arbitrary instead of P_2 and P_4. Further, one can consider that P_1 and Q_1 are arbitrary in (37). This follows since, if (38) is satisfied, it does not matter whether P_2 and Q_1 are chosen independently and P_1 is determined from (37) or if P_1 and Q_1 are chosen and P_2 determined from (37).

Assume for the moment that it is somehow possible to find P_1, Q_1, P_2 which satisfy (38). It will be shown that it is possible to make g^2 a complete square by introducing surplus factors. Once this point is clarified, it will be shown how P_1, Q_1 and P_2 can always be chosen to satisfy (38), N and D being still assumed to have only complex conjugate roots.

This restriction will be removed later without altering the arguments presented here. For convenience in discussion, $Z(p)$ is chosen to be of rank $n = 2$. Similar results are obtained for any n.

Consider $Z(p)$ of rank $n = 2$ (i.e., N and D of degree 2). From (28), P_1, Q_1 and P_2 are of degree 1. These are determined such that (38) is satisfied. Now g^2 may or may not be a perfect square. If g^2 is a perfect square, the problem is solved and the network can be realized. If g^2 is not a perfect square, surplus factors M are introduced in $Z(p)$, making $N' = NM$ and $D' = DM$ of fourth degree. New polynomials P_1', Q_1' and P_2' of second degree are required [see (28)] and are somehow determined. Then (37) may be written as

$$g^2 P_2' = M(NQ_1' - DP_1'). \tag{39}$$

In place of (39), consider

$$(NQ_1' - DP_1') = P_2' R. \tag{40}$$

If (40) is satisfied subject to the constraints of (38), where R is any remainder polynomial, one merely multiplies both sides of (40), by $M = R$ to give (39). Thus, it is seen how g^2 can be made a complete square by the introduction of an appropriate surplus factor M. All that remains to be shown is that (40) can be satisfied subject to the constraints of (38). This is no different from the problem of finding P_1, Q_1 and P_2 satisfying (38), which was assumed somehow possible. It will now be shown that it *is* possible to find polynomials P_1, Q_1 and P_2 such that

$$(NQ_1 - DP_1) = P_2 R, \tag{40a}$$

subject to the constraints

$$\left.\begin{array}{l} \dfrac{P_1}{Q_1} \\ \\ \dfrac{Q_1}{P_2} \end{array}\right\} \text{ are } RC \text{ impedances.} \tag{40b}$$

As discussed previously, P_1, Q_1 of appropriate degrees are chosen arbitrarily such that P_1/Q_1 is an RC impedance. The roots of P_1 and Q_1 of second degree are shown in Fig. 7. No singularity is chosen at the origin for reasons that will soon become clear. Symbol I refers to the function NQ_1 and II refers to $(-DP_1)$. The superscripts of I and II represent the corresponding signs of the values of the functions in the designated regions on the real axis. The shaded areas in the figure show regions where

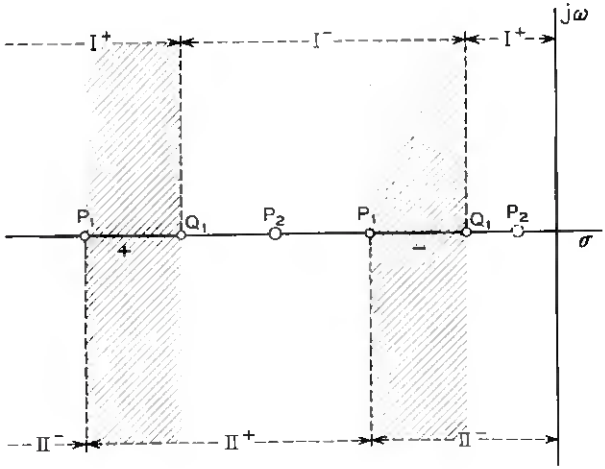


Fig. 7 — Determination of roots of P_2 when roots of P_1 and Q_1 are known.

both functions have the same sign. The two shaded areas shown are of opposite signs. Therefore, there must be at least one zero of the function $(I + II)$ in the region between these two areas. Such a zero in this region is assigned to P_2. Another zero of P_2 is required to the right of all the four roots so as to satisfy (40b). This is easily accomplished by making $z_{11} = Z$ at some point to the right of the other roots. This gives the other root of P_2 shown dotted in the figure. Thus (40a) and (40b) are completely satisfied.

The relevant steps involved for $Z(p)$ of rank n are:

1. Choose P_1 and Q_1 each of degree n, so that P_1/Q_1 is an RC impedance.

2. Evaluate $Z(p)$ at some point on the negative real axis closer to the origin than the root of Q_1 nearest to the origin.

3. Make $Z(p) = z_{11} = P_1/Q_1$ at this point by merely multiplying z_{11} by an appropriate constant.

4. Determine P_2R by finding $(NQ_1 - DP_1)$, and find the roots of P_2R.

5. Assign the appropriate roots to P_2 from steps 3 and 4.

6. Determine R from steps 4 and 5 and find $g^2 = R^2$, $N' = NR$ and $D' = DR$.

7. Express $(-D'/Q_1P_2)$ in partial-fraction form to obtain

$$(Q_3/Q_1 - P_4/P_2).$$

8. All the polynomials are determined and the relevant impedances are obtained from (30) and noting that $z_{12} = g/Q_1$.

Note that z_{22} (i.e., Q_3/Q_1) vanishes at infinity, whereas z_{11} and z_{12} do not. A sufficiently large constant is added to z_{22} and Z_L such that the two-port impedances satisfy the residue condition at infinity.

Example:

$$Z(p) = \frac{N}{D} = \frac{p^2 + p + 1}{p^2 + p + 2}.$$

1. $z_{11}(p) = K_0 \dfrac{(p+3)(p+5)}{(p+2)(p+4)} = K_0 \dfrac{(p^2 + 8p + 15)}{(p^2 + 6p + 8)}$.

2. $Z(-1) = \frac{1}{2}$.

3. $z_{11}(-1) = K_0 \frac{8}{3} = Z(-1) = \frac{1}{2}$,

$$K_0 = \tfrac{3}{16},$$

$$P_1 = 3(p^2 + 8p + 15),$$

$$Q_1 = 16(p^2 + 6p + 8).$$

4. $NQ_1 - DP_1 = P_2R = 13p^4 + 85p^3 + 165p^2 + 131p + 38$

$$\cong (p+1)(p+3.89)(13p^2 + 21.38p + 9.76).$$

5. $P_2 = (p+1)(p+3.89)$.

6. $R = 13p^2 + 21.38p + 9.76$,

$$g^2 = R^2,$$

$$N' = NR,$$

$$D' = DR = (p^2 + p + 2)(13p^2 + 21.38p + 9.76).$$

7. $$\frac{-D}{Q_1P_2} = -\frac{1}{16}\left(13 + \frac{0.318}{p+1} + \frac{2716.72}{p+3.89}\right) + \frac{1}{16}\left(\frac{20.104}{p+2} + \frac{2804.81}{p+4}\right).$$

8. $$z_{11} = \frac{1}{16}\left(3 + \frac{4.5}{p+2} + \frac{1.5}{p+4}\right),$$

$$z_{12} = \frac{1}{16}\left(13 + \frac{9.5}{p+2} + \frac{66.12}{p+4}\right),$$

$$z_{22} = \frac{1}{16}\left(K_1 + \frac{20.1}{p+2} + \frac{2804.81}{p+4}\right),$$

$$Z_L = \frac{1}{16}\left(13 + K_1 + \frac{0.318}{p+1} + \frac{2716.72}{p+3.89}\right),$$

where K_1 is chosen large enough so that $K_1 \geqq \frac{160}{3}$. The complete network is then realized.

3.3 *Elimination of Redundant Elements*

Introduction of surplus factors as proposed in the last section will introduce redundant elements in the realization. A possible formulation which may avoid the redundant elements is briefly presented here. If the roots of g are given by $p = p_i$, then, from (36),

$$Z(p_i) = z_{11}(p_i)$$

and (41)

$$\frac{d}{dp} Z \Big|_{p=p_i} = \frac{d}{dp} z_{11} \Big|_{p=p_i}, \qquad \text{for all } i.$$

Equation (41) assures that g is a polynomial. The frequencies p_i may be arbitrarily selected and (41) may then be solved. Then P_2 can be determined from (37) such that it satisfies the last restriction of (38). There is considerable freedom in the solution of (41), and this freedom may be used to satisfy the restrictions on P_2.

The difficulty encountered in the solution of (41) stems from the nature of the simultaneous equations involved. These equations are nonlinear and cumbersome to solve. It is not possible to state whether the method outlined above has a solution or not. It merely indicates a possible direction for further work on this subject.

3.4 *Restrictions on Impedance Functions*

It has been assumed that the driving-point impedances have only complex conjugate poles and zeros. The only use made of this assumption was the consequent positive sign of the impedance function on the negative σ-axis. This fact ensured that, for N/P_2P_4 and D/P_2P_4, the residues in the negative real poles (arbitrarily chosen) had alternating signs in the alternate poles. Thus, the method is applicable to the realization of any driving-point impedance (having poles and zeros anywhere in the complex plane) as long as the impedance function is positive on some interval of the negative σ-axis. If no such interval exists, then the negative of the desired impedance is realized and a second converter is used to obtain the required impedance function.

Functions that are negative on the entire negative real axis fall in two classes: (a) total number of poles and zeros on the positive σ-axis including the origin is even, with the function having a negative multiplier and (b) total number of poles and zeros on the positive σ-axis including the origin is odd, with a positive multiplier. In both these cases, there are no odd-order zeros or poles on the negative σ-axis excluding the origin. A second converter is needed in both the above cases, with the possible exception of the case when there is an odd-order pole at the origin with no other odd-order zeros or poles on the whole σ-axis. In such a case, a simple pole at the origin is removed if the residue in this simple pole is positive, and the remainder function is checked to determine if it is positive or negative on the negative real axis. If it is positive, the realization is carried out on the remainder function and a capacitor corresponding to the removed pole is added in series. If the remainder function turns out to be negative on the negative real axis, a second converter seems to be needed.

In some cases, a pole at the origin can be handled in the following manner:

$$\begin{aligned} Z = \frac{N}{pD} &= \frac{1}{p}\left[\frac{\dfrac{P_1}{P_2} - \dfrac{P_3}{P_4}}{\dfrac{Q_1}{P_2} - \dfrac{Q_3}{P_4}}\right] \\ &= \frac{1}{p}\left[\frac{\dfrac{P_1}{P_2} - \dfrac{P_3}{P_4}}{\left(\dfrac{K}{p} + \dfrac{Q_1}{P_2}\right) - \left(\dfrac{K}{p} + \dfrac{Q_3}{P_4}\right)}\right] \\ &= \frac{1}{p}\left[\frac{\dfrac{P_1}{P_2} - \dfrac{P_3}{P_4}}{\dfrac{Q_1''}{pP_2} - \dfrac{Q_3''}{pP_4}}\right] \\ &= \left[\frac{\dfrac{P_1}{P_2} - \dfrac{P_3}{P_4}}{\dfrac{Q_1''}{P_2} - \dfrac{Q_3''}{P_4}}\right]. \end{aligned} \tag{42}$$

Equation (42) is now used to determine the value of K for which z_{12} is a rational function. In the next section, such a development is applied to the synthesis of functions of rank 2.

The only other restriction on the impedance function is that the de-

grees of the numerator and denominator polynomials do not differ by more than one. This ensures that the degrees of the polynomials P_1, P_2, Q_1, Q_3, etc., lead to physical expressions for z_{ik}. It should be noted that the impedance function is not required to be positive-real, nor are the poles and zeros required to be in the left-half plane.

IV. NETWORK FUNCTIONS OF RANK 2

The cascade synthesis method will be applied to the realization of transfer functions of rank 2 (i.e., the numerator and denominator polynomials of second degree). Functions of rank 2 are by far the most important functions from sensitivity considerations. Functions with a single pair of complex poles are less sensitive to changes in the converter constant than are functions with more than one pair of complex poles.[3] Consequently, it is preferable to realize a function of rank n in sections of rank 2 and cascading these sections through isolation sections such as emitter followers.

Synthesis of driving-point functions can be adapted in several ways to the realization of transfer functions. Consider the realization of open-circuit voltage transfer function. Only two particularly simple but extremely important cases will be discussed here.

Case 1

In some instances, the input impedance may be chosen to be

$$Z(p) = \frac{N}{D} = \frac{E_0}{I_0} \tag{43}$$

and synthesized so that the resulting network has a shunt resistor R across the input terminals. Conversion from current source to voltage source $E_i = RI_0$ yields the desired transfer function [Fig. 8(a)] within a constant gain multiplier

$$\frac{E_0}{E_i} = \frac{1}{R}\frac{N}{D}. \tag{44}$$

Case 2

In other instances, $Z(p)$ may be chosen such that

$$Z(p) = \frac{E_0}{I_0} = \frac{N}{pD} \tag{45}$$

and synthesized so that the resulting network has a shunt capacitor, C, across the input terminals. A source conversion $E_i = I_0/pC$ yields the

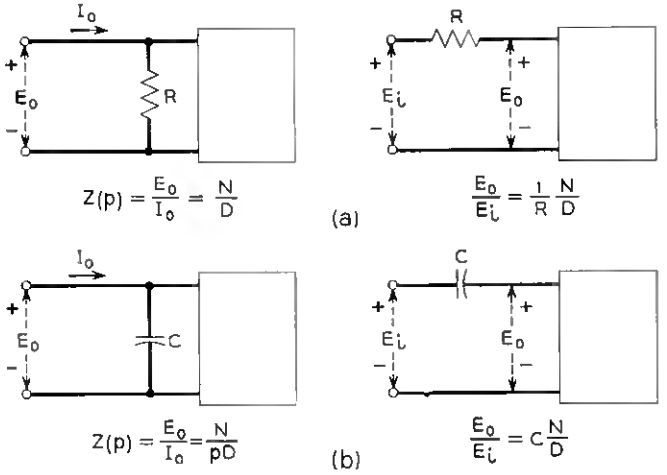


Fig. 8 — Adaptation of driving-point synthesis to voltage transfer function synthesis: (a) case 1; (b) case 2.

desired transfer function [Fig. 8(b)]:

$$\frac{E_0}{E_i} = C\,\frac{N}{D}. \tag{46}$$

The above two cases play an important role in the realization of transfer functions of rank 2. Consider the realization of open-circuit voltage transfer function of the form

$$T(p) = K_0\,\frac{p^2 + ap + b}{p^2 + cp + d}. \tag{47}$$

It is assumed that $T(p)$ has complex conjugate poles and zeros in the left-half plane. Similar developments can be carried out for other cases, but they are not discussed here.

It is shown in Appendix E that d must be greater than b if the rationalization of z_{12} is to be accomplished through the addition and subtraction of a constant, K, in the denominator of (28a). In this case, the driving-point impedance, $Z(p)$, may be synthesized so that the resulting network has a shunt resistor across its input terminals. Consequently, the transfer function of (47) may be realized as discussed above if d is greater than b. The structure is shown in Fig. 9(a).

If d is less than b, the addition and subtraction of a term of the form K/p in the denominator of (28a) will permit the rationalization of z_{12},

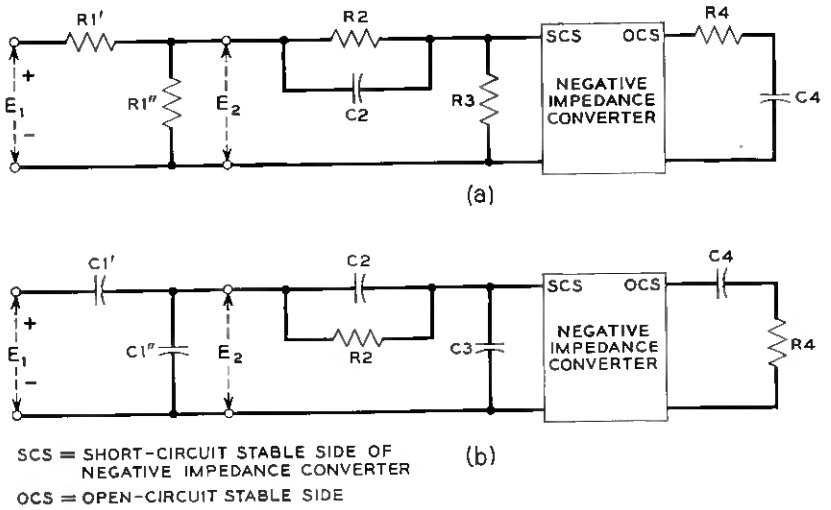


Fig. 9 — Realization of transfer functions of rank 2: (a) $d > b$; (b) $d < b$.

and the driving-point impedance $Z(p) = [T(p)/p]$ may be synthesized so that the resulting network has a shunt capacitor across its input terminals. The proper source conversion yields the desired transfer function for d less than b. The structure for this case is shown in Fig. 9(b).

V. DESIGN EXAMPLE

The results obtained in Section IV for the realization of transfer functions of rank 2 by the cascade method will be applied to the design of a practical filter. Bandpass filters with high selectivity are of great importance in frequency-multiplexing schemes. A typical bandpass filter frequently encountered in such systems is selected as a design example. The over-all transfer function is represented as a product of functions of rank 2, and each of these is realized as discussed in Section IV. Each of the sections is a simple ladder structure regardless of the positions of the zeros and poles.

The specifications on the filter are as follows: the pass band extends from 12.33 to 15.25 kc, and in this range the transmission characteristic must be flat to within ±0.1 db; the stop band is symmetrical; and, starting at 11.50 kc, a minimum of 50 db of rejection must be provided.

The approximation problem is solved through the utilization of a Tchebycheff parameter equal-ripple chraacteristic. The transmission

function $T(p)$ requires five pairs of complex poles, four pairs of imaginary zeros and a simple zero at the origin and infinity:

$$T(p) = K \frac{p(p^2 + \omega_{01}^2)(p^2 + \omega_{02}^2)(p^2 + \omega_{03}^2)(p^2 + \omega_{04}^2)}{(p^2 + 2\alpha_1 p + \omega_{n1}^2)(p^2 + 2\alpha_2 p + \omega_{n2}^2)(p^2 + 2\alpha_3 p + \omega_{n3}^2) \cdot (p^2 + 2\alpha_4 p + \omega_{n4}^2)(p^2 + 2\alpha_5 p + \omega_{n5}^2)},$$

where

$$\begin{aligned} \omega_{01} &= 2\pi \times 10349.7 \text{ rps}, \\ \omega_{02} &= 2\pi \times 11388.7 \text{ rps}, \\ \omega_{03} &= 2\pi \times 16510.4 \text{ rps}, \\ \omega_{04} &= 2\pi \times 18167.9 \text{ rps}, \end{aligned}$$

$$\begin{aligned} \alpha_1 &= 2\pi \times 143.6 \text{ rps}, & \omega_{n1} &= 2\pi \times 12273.5 \text{ rps}, \\ \alpha_2 &= 2\pi \times 492.1 \text{ rps}, & \omega_{n2} &= 2\pi \times 12695.6 \text{ rps}, \\ \alpha_3 &= 2\pi \times 767.1 \text{ rps}, & \omega_{n3} &= 2\pi \times 13691.0 \text{ rps}, \\ \alpha_4 &= 2\pi \times 573.2 \text{ rps}, & \omega_{n4} &= 2\pi \times 14788.6 \text{ rps}, \\ \alpha_5 &= 2\pi \times 179.3 \text{ rps}, & \omega_{n5} &= 2\pi \times 15318.1 \text{ rps}. \end{aligned}$$

This transmission characteristic is illustrated in Fig. 10 and the pole-zero configuration is shown in Fig. 11. The transmission function is expressed as a product of second order expressions:

$$T(p) = K \left(\frac{p^2 + \omega_{04}^2}{p^2 + 2\alpha_1 p + \omega_{n1}^2}\right)\left(\frac{p^2 + \omega_{03}^2}{p^2 + 2\alpha_2 p + \omega_{n2}^2}\right) \cdot \left(\frac{p}{p^2 + 2\alpha_3 p + \omega_{n3}^2}\right)\left(\frac{p^2 + \omega_{02}^2}{p^2 + 2\alpha_4 p + \omega_{n4}^2}\right) \cdot \left(\frac{p^2 + \omega_{01}^2}{p^2 + 2\alpha_5 p + \omega_{n5}^2}\right).$$

Each factor is selected so as to obtain a reasonably low sensitivity to changes in element values of the network section. Each of the second-order expressions is then realized in a separate section as a voltage transfer ratio. The last two factors belong to Case 1 and are realized to yield structures of the form shown in Fig. 9(a). The first two factors belong to Case 2 and are realized to yield structures of the form shown in Fig. 9(b). The middle factor has a zero at the origin and would either require the use of surplus factors, or perhaps two converters, as discussed above. However, the Q of the poles of this factor is rather low. Consequently, for the sake of convenience, this factor is realized with a simple series RLC circuit with a ferrite coil. Design of one section is considered below as an illustration.

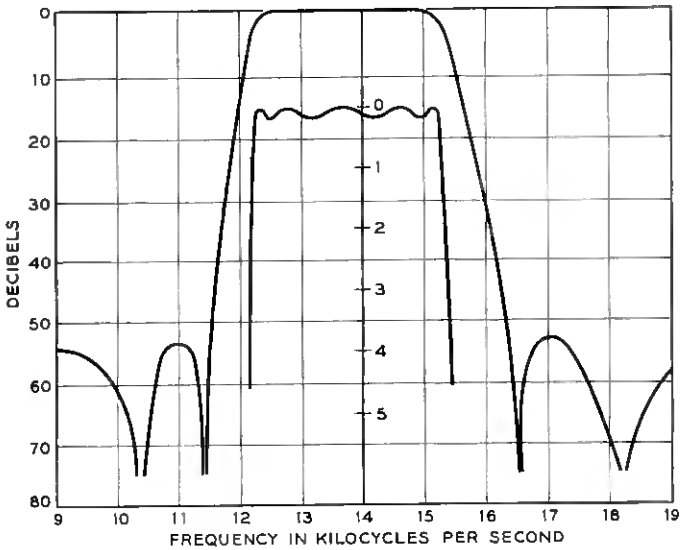


Fig. 10 — Design example — filter characteristic.

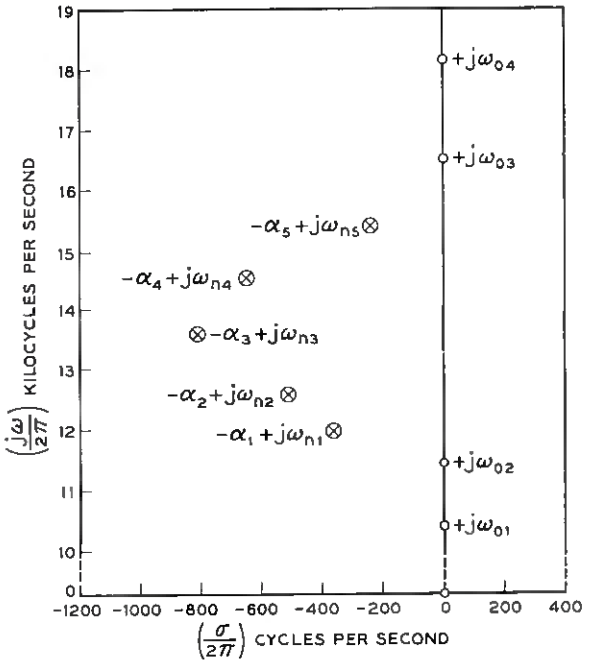


Fig. 11 — Design example — pole-zero configuration (not drawn to scale; pole and zero positions normalized with respect to 2π).

Consider the factor

$$\frac{p^2 + {\omega_{01}}^2}{p^2 + 2\alpha_5 p + {\omega_{n5}}^2}$$

normalized with respect to frequency with a scale factor of 10^5:

$$Z(p) = \frac{p^2 + 0.423}{p^2 + 0.023p + 0.926}.$$

Choose

$$\sigma_1 = 0,$$
$$\sigma_2 = \sqrt{0.423} = 0.650.$$

Then

$$P_2 = p,$$
$$P_4 = p + 0.650,$$
$$P_1 = p + 0.650,$$
$$P_3 = 2 \times 0.650 = 1.30,$$
$$Q_1' = (1 + K)p + \frac{0.926}{0.650},$$
$$Q_3' = Kp + \frac{(1 + K)0.423 - 0.023 \times 0.65 + 0.926}{0.650}.$$

The value of K for $(P_1Q_3' - Q_1'P_3)$ to be a complete square is found to be $K = 1.455$. Impedances z_{11}, z_{12}, z_{22} and Z_L are now obtained and the network realized. In order to ensure that there are no transformers present, it is sufficient to make at $p = \infty$

$$z_{22}(\infty) = z_{12}(\infty) = z_{11}(\infty).$$

The above requirement for no transformers is obvious in view of the structure of Fig. 9(a). This requirement is satisfied very simply by the appropriate scaling of the impedances involved. Consider

$$Z = z_{11} - \frac{K^* {z_{12}}^2}{K^* z_{22} - K^* Z_L}.$$

Here a scale factor K^* is introduced to satisfy the above condition of no transformers.

The new impedances are z_{11}, $\sqrt{K^*}\, z_{12}$, $K^* z_{22}$ and $K^* Z_L$. It should be observed that $K^* = 1/K$. The network is now realized to yield the de-

sired impedance function. The transfer function is obtained by the source conversion; either the whole or part of the shunt resistor may be used for this purpose. The actual values of the elements are obtained by introducing appropriately the frequency scale factor of 10^5. An impedance scale factor of 5×10^3 is then introduced to obtain element values which are practical. The element values of all the four active sections are given in Tables I and II. Table III gives the values for the series *RLC* circuit.

It is indeed possible to prepare tables which directly yield the element values in terms of the coefficients of the numerator and denominator polynomials. Such tables have been prepared and used to design filter sections with a considerable saving in time.

TABLE I — ELEMENT VALUES IN OHMS AND MICROMICROFARADS FOR TWO SECTIONS OF THE DESIGN EXAMPLE — $d > b$

$T(p)$	$K_1 \frac{p^2 + \omega_{01}^2}{p^2 + 2\alpha_5 p + \omega_{n5}^2}$	$K_2 \frac{p^2 + \omega_{02}^2}{p^2 + 2\alpha_4 p + \omega_{n4}^2}$
$R1'$	36755	4166
$R1''$	2661	2528
$R2$	17314	17020
$R3$	11412	8442
$R4$	3410	2790
$C2$	2227	2466
$C4$	4453	4934

TABLE II — ELEMENT VALUES IN OHMS AND MICROMICROFARADS FOR TWO SECTIONS OF THE DESIGN EXAMPLE — $d < b$

$T(p)$	$K_3 \frac{p^2 + \omega_{03}^2}{p^2 + 2\alpha_2 p + \omega_{n2}^2}$	$K_4 \frac{p^2 + \omega_{04}^2}{p^2 + 2\alpha_1 p + \omega_{n1}^2}$
$R2$	7350	6601
$R4$	3650	3275
$C1'$	1769	247.4
$C1''$	2915	3416
$C3$	873	797
$C2$	433.5	525
$C4$	2613	2644

TABLE III — ELEMENT VALUES IN OHMS, MILLIHENRIES AND MICROMICROFARADS

$T(p)$	$K_5 \frac{p}{p^2 + 2\alpha_3 p + \omega_{n3}^2}$
R	375
L	52.25
C	2438.20

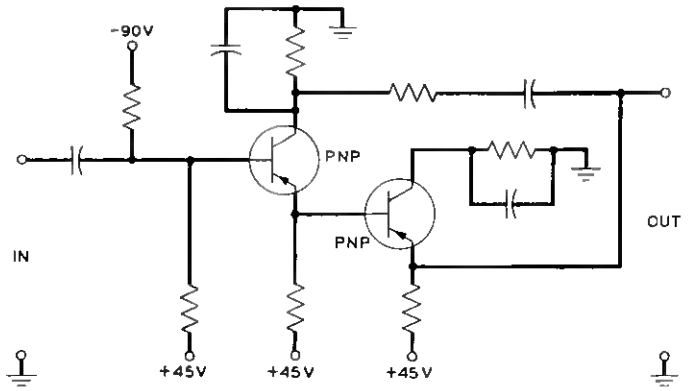


Fig. 12 — Emitter follower.

The above filter sections are then cascaded through isolation sections, which are emitter followers. The negative-impedance converter used in the experimental model is of the type proposed by Larky.[11] Circuit designs of the emitter follower and the converter are shown in Figs. 12 and 13. Note that converters are short-circuit stable at one port and open-circuit stable at the other port. Care must be taken to connect them properly so that the system is stable.

The experimental model is constructed on five 2 × 3 inch cards, with one filter section and an isolation network on each. The complete filter

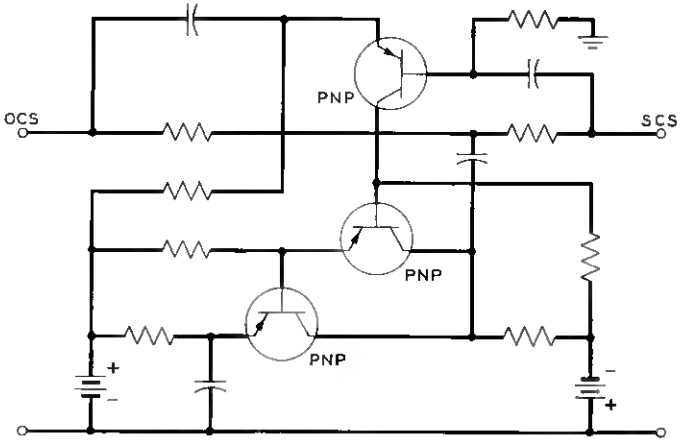


Fig. 13 — Negative impedance converter.

can be placed in a $5\frac{1}{2} \times 4 \times 2\frac{3}{4}$ inch can. Pictures of the filter are shown in Figs. 14 and 15.

The passband transmission characteristic of the experimental model has the five ripples and is flat to within the specified ± 0.1 db. The characteristic stays flat to within ± 0.15 db over a 10°C temperature variation.

VI. CONCLUSION

It is shown that any driving-point function can be synthesized using passive RC networks and one active element. The active element is assumed to be a current-controlled voltage source. A good approximation to such sources is practically possible. The synthesis technique involves synthesis of n-port RC networks and n-port R networks. Sufficient condition for synthesis of n-port R networks without transformers is shown to be the dominant diagonal matrix. Synthesis of n-port RC networks is considered in the Appendices. It is shown that this problem can be reduced to the synthesis of $(m + n)$-port R network.

The above method of synthesis of driving-point functions leaves much to be desired. The major objection is the balanced structure and the large number of passive elements resulting from the synthesis of $(m + n)$-port R network. It should be obvious that the realization of the R net-

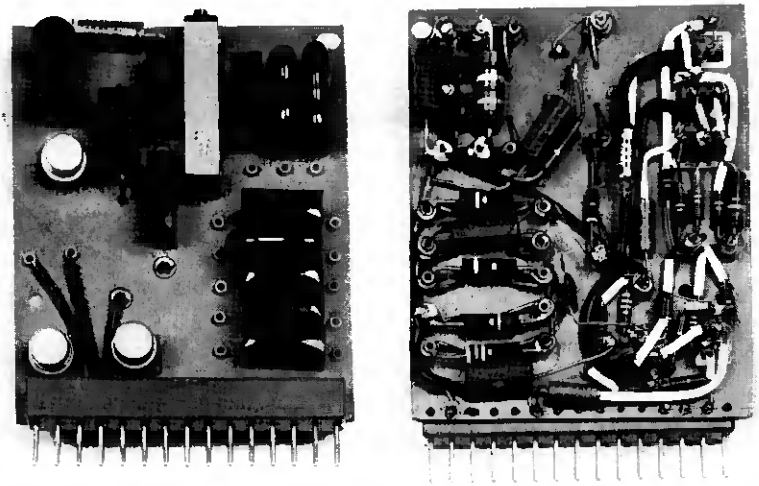


Fig. 14 — Front and rear views of one filter section.

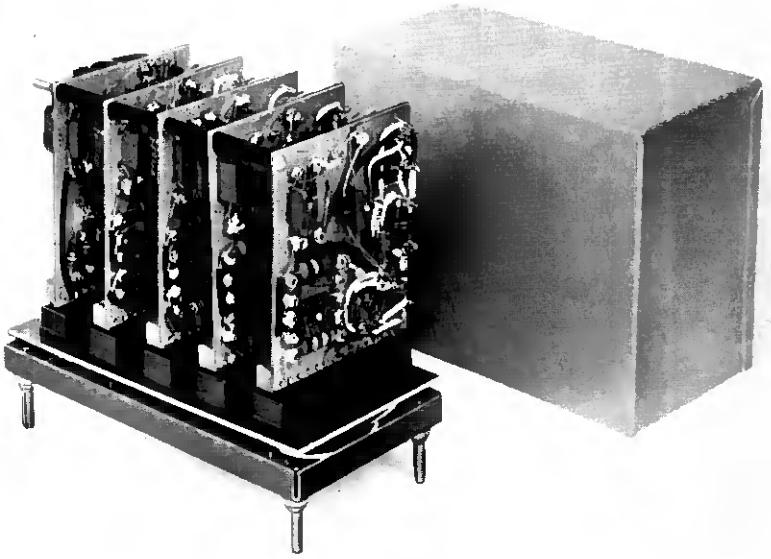


Fig. 15 — Complete bandpass filter.

work with a minimum number of resistors would render this method quite practical.

A more practical synthesis method for driving-point functions is considered next. The cascade method uses the negative-impedance converter as an active element. This method is adapted to the synthesis of voltage transfer functions. Particularly simple structures result when functions of rank 2 are to be realized. A rather elaborate filter is designed using these ideas. An experimental model of such a filter has been built and tested and the results are reported.

Further work on active networks is indicated in several directions. Theoretical investigation may be undertaken to determine the minimum number of active elements required to realize an arbitrary two-port matrix. The use of active elements with *RLC* networks may yield networks with fewer elements. Possibility of using several active elements scattered through the network to yield more stable characteristics may bear some investigation. It is indeed not farfetched to envisage a day when the use of active elements in network design will be viewed with as much equanimity as the use of passive elements is viewed today.

VII. ACKNOWLEDGMENTS

Reduction to practice of the cascade method was carried out by R. P. Abraham and J. M. Sipress. The filter was built and tested by E. P. Greener. Computing work was done by Mrs. H. D. Reinecke.

APPENDIX A

Given $F = N/D$, it will be shown that this may be expressed as

$$F = \frac{N_1 - N_2}{D_1 - D_2}, \tag{48}$$

where N_1/D_1 is an RC admittance and the degrees of N_2 and D_2 are equal to the degrees of N_1 and D_1 respectively.

Choose N_1 and D_1 of the same degree such that N_1/D_1 is an RC admittance. Then

$$\begin{aligned} N_2 &= N_1 - N, \\ D_2 &= D_1 - D. \end{aligned} \tag{49}$$

Degrees of N_2 and D_2 can always be made equal to the degrees of N_1 and D_1, respectively, by a proper choice of the degrees of N_1 and D_1. Thus, the polynomials N_1, D_1, N_2, D_2 are all of the same degree, say n_1. This fact ensures that N_1/D_1, N_2/N_1 and D_2/D_1 have the desirable property (see Appendix B) of having no poles at infinity. Obviously, the degree n_1 should not be less than n, the higher of the degrees of N and D.

It will now be shown that, given a polynomial N, it can always be written as a difference of two polynomials having only negative real roots. This development can in many cases be used to obtain the desired result above. The development is of greater interest in the cascade synthesis method.

Form a function

$$F_1(p) = \frac{N}{\prod_{\nu=1}^{n} (p + \sigma_\nu)}, \tag{50}$$

where σ_ν are real positive numbers and n is either equal to or exceeds by one the degree of N. The residue of the function F_1 in any of the poles is seen to be real. So,

$$F_1(p) = K_\infty + \sum_i \frac{|K_i|}{p + \sigma_i} - \sum_j \frac{|K_j|}{p + \sigma_j}. \tag{51}$$

In (51), the part with positive residues represents a passive *RC* impedance, and so the zeros of this part must be negative real. The part with negative residues is simply an *RC* impedance with a negative sign and so it must also have negative real zeros. Then

$$F_1(p) = \frac{P_1}{P_2} - \frac{P_3}{P_4}, \tag{52a}$$

where all the polynomials P_1, P_2, P_3, P_4 have negative real roots. Further,

$$F_1(p) = \frac{P_1P_4 - P_2P_3}{P_2P_4} = \frac{P_1P_4 - P_2P_3}{\prod_{\nu=1}^{n} (p + \sigma_\nu)} \tag{52b}$$

and

$$N = P_1P_4 - P_2P_3 = N_1 - N_2, \tag{53}$$

where N_1 and N_2 have only negative real roots, as was to be shown.

In a similar manner, the polynomial D can be expressed as $D = D_1 - D_2$, where D_1 and D_2 have only negative real roots. With a proper choice of σ_ν, it may be possible to make N_1/D_1 an *RC* immittance function. No attempt has been made to determine under what conditions the above is possible, since such a general statement does not seem to be of great importance. The difficulty involved in determining the proper σ_ν for a specific problem would govern whether one should apply this method or not.

In (50), if all σ_ν are selected on intervals of the negative real axis where $N(p)$ has the same sign, then the residues of F_1 will have alternating signs at alternate poles. This is easily observed by considering a polynomial N having only complex conjugate roots as shown in Fig. 16.

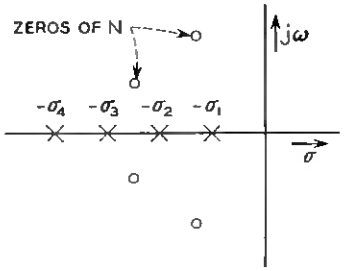


Fig. 16 — Residues with alternating signs in alternate poles.

Along the σ-axis the sign of the function $F_1(p)$ changes on passing a pole but not between the poles. Hence, the residues in these poles have alternating signs at alternate poles. Moreover, if the σ_r are selected on intervals of the negative σ-axis where $N(p)$ has a positive sign, then the residue of F_1 in the pole closest to the origin, say σ_1, has a positive sign and P_4/P_2 is an RC impedance function.

The results of interest in the cascade method are:

i. From (50) and (51) it is seen that a difference of two RC impedance functions can produce any desired complex zeros.

ii. With a proper choice of σ_r, it is observed that P_4/P_2 is an RC impedance function.

APPENDIX B

n-Port RC Networks

Assume that it is possible to find a minimum number, m, of capacitors in the desired RC n-port network. The network can then be modified to a $(m + n)$-port terminated in its m-ports in capacitors (Fig. 3). The $(m + n)$-port network is purely resistive and can be investigated to determine if ideal transformers can be avoided. This will not give a complete class of networks without ideal transformers; it will give only a subclass of networks with a minimum number of capacitors and no transformers. In order to obtain the complete class of networks m must be made greater than the minimum number. It will be seen later how the value of m can be chosen to give the desired number of capacitors.

The problem may be stated as follows:

A short-circuit admittance matrix Y of order n representing a n-port RC network is given. It is desired to find a conductance matrix G representing a resistive network such that, when this network is terminated at its m-ports in certain capacitors, the n-port RC network Y is obtained.

Let

$$G = \begin{matrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{matrix} n \\ m \end{matrix} \\ \begin{matrix} n & m \end{matrix} \end{matrix}, \tag{54}$$

where the A_{ij}'s are matrices. When the m-ports are properly terminated in capacitors, the resulting $(n \times n)$ matrix is required to be Y. For convenience, the terminations are assumed to be unit capacitors. Any transformers that appear at these m-ports are eliminated by a simple scaling of these capacitors. Then,

$$Y = A_{11} - A_{12}[A_{22} + Ip]^{-1} \tilde{A}_{12}, \tag{55}$$

where

I = unit matrix,
$\bar{A}$ = transpose of A,
$\bar{A}_{12} = A_{21}$, since G is symmetrical.

The short-circuit admittance matrix Y of the RC network may be expressed as

$$Y = Kp + K_\infty - \sum_\nu K_\nu \left(\frac{1}{p + \sigma_\nu}\right), \tag{56a}$$

where K, K_∞ and K_ν are $(n \times n)$ matrices of residues. For a passive network, these matrices are positive semidefinite. At $p = 0$, $Y(0)$ must also be positive semidefinite for physical realizability. It is desired to identify the expression of (55) with that of (56a). The only way to do this is by making $A_{11} = Kp + K_\infty$. This, of course, does not keep G purely resistive. An alternative approach is to let $Y = Kp + Y_1$, and then identify Y_1 with the expression in (55). Then, the network corresponding to Kp is added in parallel to the network corresponding to Y_1. As will be seen below, the networks obtained are balanced structures and no transformers are needed when the structures are connected in parallel. Even though Y in (56a) can always be realized in this manner, it is not at all desirable, since the realization of Kp requires a large number of additional capacitors. This whole difficulty can be avoided by considering the open-circuit impedance matrix Z and carrying out a development analogous to the one below. No pole at infinity is present in Z and the above difficulty does not arise. The other possible way of avoiding the additional capacitors is to make $K \equiv 0$. It is seen in Appendix A that it is indeed possible to do so when Y is obtained for the active network problem under consideration. Furthermore, since sufficient conditions for realizability of resistance networks without transformers are obtained in terms of the short-circuit admittance matrix G, it is desirable to make the development here in terms of Y. Assuming that K is made equal to zero, Y may be expressed as

$$Y = K_\infty - \sum_\nu K_\nu \left(\frac{1}{p + \sigma_\nu}\right), \tag{56b}$$

where K_∞, K_ν and $Y(0)$ are positive semidefinite.

It is now necessary to put (56b) in such a form that it is readily identified with (55). Let

$$K_\nu = M_\nu \bar{M}_\nu, \tag{57}$$

where M_ν has n rows and m_ν columns. It is then possible to write

$$K_\nu\left(\frac{1}{p+\sigma_\nu}\right) = M_\nu\left[\frac{1}{p+\sigma_\nu}\right]_{m_\nu}\tilde{M}_\nu\,, \tag{58}$$

where

$$\left[\frac{1}{p+\sigma_\nu}\right]_{m_\nu}$$

is a diagonal matrix of order m_ν and elements

$$\left(\frac{1}{p+\sigma_\nu}\right).$$

Let

$$\left[\frac{1}{p+\sigma_\nu}\right]_{m_\nu} = D_{m_\nu}\,. \tag{59}$$

Then

$$\sum_\nu K_\nu\left(\frac{1}{p+\sigma_\nu}\right) = M_\nu D_{m_\nu}\tilde{M}_\nu \tag{60}$$

and

$$\sum_\nu M_\nu D_{m_\nu} M_\nu$$

$$= [M_1, M_2, \cdots]\begin{bmatrix} D_{m_1} & 0 & 0 & 0 & 0 \\ 0 & D_{m_2} & 0 & 0 & 0 \\ 0 & 0 & D_{m_3} & 0 & 0 \\ 0 & 0 & 0 & & 0 \\ 0 & 0 & 0 & 0 & \end{bmatrix}\begin{bmatrix} M_1 \\ M_2 \\ \\ \\ \end{bmatrix} = MH\tilde{M}, \tag{61}$$

where

$$M = [M_1, M_2, \cdots] \tag{62}$$

and

$$H = \begin{bmatrix} D_{m_1} & 0 & 0 & 0 & 0 \\ 0 & D_{m_2} & 0 & 0 & 0 \\ 0 & 0 & D_{m_3} & 0 & 0 \\ 0 & 0 & & & \end{bmatrix}. \tag{63}$$

The columns in M and the order of H are both given by

$$m = \sum_{\nu} m_{\nu} . \tag{64}$$

It is seen from (54) that m also determines the number of capacitors. Thus, m can be arbitrarily chosen by use of (57) and (64) to be any value greater than or equal to the minimum value discussed below. It is, however, necessary to determine G in terms of the known quantities in (56b). Equation (56b) can be rewritten using (61):

$$Y = K_{\infty} - MH\tilde{M}. \tag{65}$$

Comparing it with (55),

$$\begin{aligned} A_{11} &= K_{\infty}, \\ A_{12} &= M, \\ [A_{22} + Ip]^{-1} &= H. \end{aligned} \tag{66}$$

Since H is a diagonal matrix, its inverse is given simply by inverting each term of the matrix, each term being of the form $[1/(p + \sigma)]$:

$$H^{-1} = [Ip + S], \tag{67}$$

where I is a unit matrix of order m and

$$S = \begin{bmatrix} \sigma_1 I_1 & 0 & 0 & 0 & 0 \\ 0 & \sigma_2 I_2 & 0 & 0 & 0 \\ 0 & 0 & \sigma_3 I_3 & 0 & 0 \\ 0 & 0 & 0 & & 0 \\ 0 & 0 & 0 & 0 & \end{bmatrix}, \tag{68}$$

where I_j is a unit matrix of order m_j. Thus,

$$\left.\begin{aligned} A_{11} &= K_{\infty} \\ A_{12} &= M \\ A_{22} &= S \end{aligned}\right\}. \tag{69}$$

If it is desirable to obtain a network with a minimum number of capacitors, it is necessary to obtain a minimum m. This is accomplished by having each m_{ν} a minimum number. Consider the quadratic form of a matrix K_{ν} expressed as sum of the squares

$$\bar{X}\, K_{\nu} X = \sum_{i} [X\, M_{\nu i}]^2,$$

where each $M_{\nu i}$ is a column vector. Then

$$\begin{aligned} \bar{X}K_\nu X &= \sum_i \bar{X}\, M_{\nu i}\bar{M}_{\nu i}X \\ &= \bar{X}\, M_\nu \bar{M}_\nu X, \end{aligned} \tag{70}$$

where

$$M_\nu = [M_{\nu 1}\,,\ M_{\nu 2}\,,\ M_{\nu 3}\,,\ \cdots]. \tag{71}$$

If K_ν is of order n and rank δ_ν, then there are δ_ν independent terms in the quadratic form. The matrix M_ν has n rows and δ_ν columns. This δ_ν is the minimum number of columns in M_ν, since the quadratic form in (70) must have at least δ_ν terms. It is thus necessary to use (70) instead of (57) to obtain a minimum

$$m = \sum_\nu \delta_\nu = \sum_\nu m_\nu\,,$$

since δ_ν is the minimum value of m_ν.

It is easy to show that G, whose submatrices are determined by (69), represents a physically realizable resistance network. Consider the quadratic form of matrix G,

$$\begin{aligned} Q &= [\bar{x},\bar{y}]\begin{bmatrix} A_{11} & A_{12} \\ \bar{A}_{12} & A_{22} \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} \\ &= \bar{x}A_{11}x + \bar{x}A_{12}y + \bar{y}\bar{A}_{12}x + \bar{y}A_{22}y \\ &= \bar{x}K_\infty x + \bar{x}My + \bar{y}\bar{M}x + \bar{y}Sy \\ &= \bar{x}K_\infty x - \bar{x}MS^{-1}\bar{M}x + [\bar{x}M + S\bar{y}]S^{-1}[\bar{M}x + Sy] \\ &= \bar{x}[K_\infty - S^{-1}M\bar{M}]x + [\bar{x}M + S\bar{y}]S^{-1}[\bar{M}x + Sy] \\ &= \bar{x}[Y(0)]x + \bar{z}S^{-1}z, \end{aligned}$$

where $z = \bar{M}x + Sy$ is a column matrix. From (55) and (69), $Y(0)$ is a positive semidefinite matrix and S^{-1} is a diagonal matrix with positive elements. Therefore, $Q > 0$ for all x and z. Hence the quadratic form is positive definite and the matrix G is physically realizable.

It has been shown that, given an n-port RC matrix Y, it is possible to find a matrix G for an auxiliary $(n + m)$-port resistance network. This $(m + n)$-port R network must be terminated at its m-ports in unit capacitors to obtain the desired n-port RC network.

APPENDIX C

It will be shown that, if K_ν is a third-order matrix some of whose elements are specified by (7), then y_{11} and y_{22} may be chosen so that K_ν is positive semidefinite of rank 1.*

From (7) it is clear that

$$R_m(y_{13}y_{23} - y_{12}y_{33}) = \frac{P_2}{Q_1}. \tag{72}$$

Since Q_1 has simple zeros, then $y_{13}y_{23} - y_{12}y_{33}$ must have simple poles. Substituting

$$y_{ij} = k_{ij}^{\infty} - \sum_{\nu} \frac{k_{ij}^{(\nu)}}{p + \sigma_\nu} \tag{73}$$

in the left-hand side of (72) and setting the coefficients of $1/(p + \sigma_\nu)^2$ equal to zero gives

$$k_{13}^{(\nu)}k_{23}^{(\nu)} - k_{12}^{(\nu)}k_{33}^{(\nu)} = 0 \qquad \text{for all } \nu. \tag{74}$$

Thus, y_{11} and y_{22} are driving-point admittances with poles σ_ν but residues as yet unspecified. It is well known that there exists an orthogonal transformation which transforms a positive semidefinite matrix of rank 1 into the diagonal form $[k, 0, 0]$ with $k > 0$. Since the eigenvalues of a matrix A are roots of the equation

$$\lambda^3 - \lambda^2 \operatorname{tr} A + \tfrac{1}{2}\lambda[(\operatorname{tr} A)^2 - \operatorname{tr} A^2] - \det A = 0, \tag{75}$$

where tr means the sum of the diagonal elements, and since tr A, tr A^2 and det A are invariant under orthogonal transformations, then for a matrix of rank 1

$$\operatorname{tr} A = k, \qquad \operatorname{tr} A^2 = k^2 \qquad \text{and} \quad \det A = 0. \tag{76}$$

Then (75) reduces to

$$\lambda^3 - \lambda^2 k = 0$$

with roots $(k, 0, 0)$. Hence $k_{11}^{(\nu)}$ and $k_{22}^{(\nu)}$ must be chosen so that they are positive and (76) is satisfied for K_ν. This can be done provided (74) is satisfied and

$$k_{11}^{(\nu)} = \frac{[k_{13}^{(\nu)}]^2}{k_{33}^{(\nu)}},$$

* The proof is due to Mrs. B. A. Morrison.

and

$$k_{22}^{(\nu)} = \frac{[k_{23}^{(\nu)}]^2}{k_{33}^{(\nu)}}.$$

To determine M_ν, observe that, under the orthogonal transformation $X = TY$, where

$$X = \{x_i\}, \qquad Y = \{y_i\}, \qquad T = \| t_{ij} \| \quad \text{and} \quad T^{-1} = \bar{T},$$

the quadratic form for K_ν becomes

$$\bar{X}K_\nu X = \bar{Y}\bar{T}K_\nu TY = ky_1^2 = k\left[\sum_{j=1}^{3} t_{j1}X_j\right]^2 = \left[\pm\sqrt{k}\sum_{j=1}^{3} t_{j1}X_j\right]^2.$$

Then,

$$M_\nu = \pm\{\sqrt{k}\, t_{j1}\}.$$

Since

$$\bar{T}K_\nu T = \underline{K}_\nu,$$

where $\underline{K}_\nu$ is the diagonal matrix $[k, 0, 0]$, then

$$\bar{T}K_\nu = \underline{K}_\nu\bar{T}$$

and t_{j1} is seen to be the solution to the following set of equations:

$$\sum_{j=1}^{3} t_{j1}[k_{jm}^{(\nu)} - k\delta_{jm}] = 0 \qquad \text{for } m = 1, 2, 3.$$

The result obtained is

$$\bar{M}_\nu = \pm\frac{1}{\sqrt{k_{33}^{(\nu)}}}[k_{13}^{(\nu)}, k_{23}^{(\nu)}, k_{33}^{(\nu)}].$$

APPENDIX D

It will be demonstrated that Q_3/Q_1 is always an RC driving-point impedance provided $D(p)$ contains only complex zeros. Further, Q_3'/Q_1'

as well as Q_3''/Q_1'' are also *RC* driving-point impedance functions. Polynomials Q, Q_3 and Q_1 are defined in (27) and (28), and $Q_1' = Q_1 + KP_2$ and $Q_3' = Q_3 + KP_4$ [P_2 and P_4 defined in (28)]. Polynomials Q_1'', and Q_3'' are defined in (42), and are given by $Q_1'' = Q_1 p + KP_2$, $Q_3'' = Q_3 p + KP_4$.

It follows from the definitions above and from (28) that

$$Q_1' P_4 - Q_3' P_2 = D(p), \tag{77}$$

where Q_1'/P_2, Q_3'/P_4 and P_4/P_2 characterize *RC* driving-point impedance functions. Also, the order of Q_1' and Q_3' equals the order of P_2 and P_4, respectively. Consequently, the roots of $D(p)$ are given by the zeros of an expression of the form

$$1 - R \frac{\prod (p + d_{2i}') \prod (p + f_{1j})}{\prod (p + d_{1k}') \prod (p + f_{2l})} = 0, \tag{78}$$

where

$$Q_1' = k_a' \prod (p + d_{1k}'),$$

$$Q_3' = k_b' \prod (p + d_{2i}'),$$

$$P_2 = k_c \prod (p + f_{1j}),$$

$$P_4 = k_d \prod (p + f_{2l}).$$

Consider the root locus of (78) for $R > 0$. For some specific value of R, the expression yields the roots, and only the roots, of $D(p)$. Consider the constraints that this places on the open-loop poles and zeros of the locus, i.e., the roots of $Q_1' P_4$ and $Q_3' P_2$, respectively, along the negative real axis.

The open-loop root closest to the origin must be a zero of the locus, i.e., a root of P_2, since Q_1'/P_2, Q_3'/P_4 and P_4/P_2 are all *RC*. The next, or second, root is an open-loop pole of the locus. The third root must also be an open-loop pole of the locus, since, if it were an open-loop zero, then it would follow that for all positive values of R there must be a zero of (77) on the negative real axis. This is not permissible, since there must exist a positive value of R for which (78) must yield only the roots of $D(p)$ which are complex. Thus, the second and third roots must be open-loop poles of the locus. The only way this can occur is if one is a root of Q_1' and the other is a root of P_4.

A similar line of reasoning is carried out until all the roots are used. Hence, if the degree of P_2 equals that of P_4, the only possible configuration for the open-loop poles and zeros of the locus is that shown in Fig. 17(a); if the degree of P_2 is one greater than that of P_4, the configuration is as shown in Fig. 17(b). Consequently, the roots of Q_1' and Q_3' must alternate along the negative real axis, with the one closest to the origin being a root of Q_1'. Therefore, Q_3'/Q_1' characterizes an RC driving-point impedance function if the roots of $D(p)$ are all complex.

Thus, Q_3/Q_1 is simply the special case of Q_3'/Q_1' for $K = 0$ in the definitions above. Configurations of the roots are similar to those obtained above and Q_3/Q_1 is an RC impedance function.

In the case of Q_3''/Q_1'', the only difference is that now

$$Q_1'' P_4 - Q_3'' P_2 = pD(p). \tag{79}$$

An expression identical to that of (78) gives the roots of $pD(p)$ which are all complex except for the one at origin. The root-locus is again considered and the configuration of the open-loop zeros and poles is determined. The configuration is shown in Fig. 17(c). The roots of Q_3'' and Q_1'' alternate along the negative real axis with the one closest to the origin being a root of Q_1''; Q_3''/Q_1'' characterizes an RC driving-point impedance function.

APPENDIX E

Conditions are obtained under which it is possible to rationalize z_{12} without use of surplus factors for network functions of rank 2. An additional condition to be satisfied is that the resulting two-port RC network be a ladder structure. Consider a transfer function

$$T(p) = K_0 \frac{p^2 + ap + b}{p^2 + cp + d}, \tag{80}$$

and choose the impedance functions as

$$Z(p) = \frac{p^2 + ap + b}{p^2 + cp + d} \tag{81}$$

or

$$Z(p) = \frac{1}{p} \frac{p^2 + ap + b}{p^2 + cp + d}. \tag{82}$$

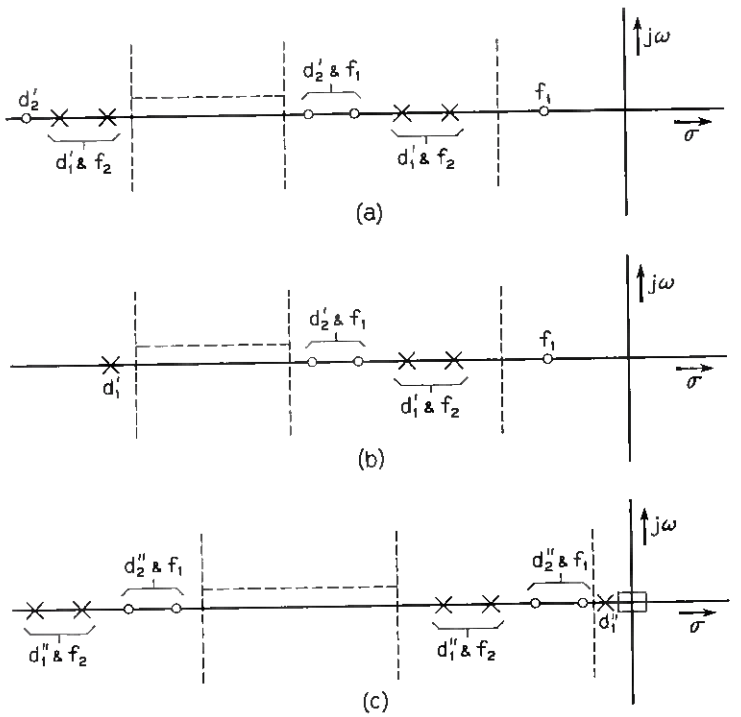


Fig. 17 — Configurations for open-loop poles and zeros of the root locus (a) when P_2 and P_4 are of same degree; (b) when degree of P_2 is one greater than degree of P_4; (c) when polynomials Q_3'' and Q_1'' are used to determine the configuration.

Equation (81) is to be considered when an introduction of constant K is ufficient to rationalize z_{12}, and (82) must be used when K/p is required.

When (81) is used,

$$P_2 = p + \sigma_1 ,$$

$$P_4 = p + \sigma_2 \qquad \sigma_2 > \sigma_1 > 0,$$

$$P_1 = p + \frac{\sigma_1\sigma_2 - a\sigma_1 + b}{\sigma_2 - \sigma_1},$$

$$P_3 = \frac{\sigma_2^{\,2} - a\sigma_2 + b}{\sigma_2 - \sigma_1}.$$

When the constant K is appropriately introduced,

$$Q_1' = (K+1)p + \frac{(K+1)\sigma_1\sigma_2 - K\sigma_1^2 - c\sigma_1 + d}{\sigma_2 - \sigma_1},$$

$$Q_3' = Kp + \frac{(K+1)\sigma_2^2 - K\sigma_1\sigma_2 - c\sigma_2 + d}{\sigma_2 - \sigma_1}.$$

It is required that $P_1Q_3' - Q_1'P_3 = 0$ have a double root at $p = \eta \leqq 0$, where η is assumed to be nonpositive so as to obtain a ladder network when the two-port is synthesized.

Thus,

$$P_1Q_3' - Q_1'P_3 = Kp^2 + p\,(x + Ka) + (y + Kb) = 0, \quad (83)$$

where

$$x = \frac{d - b + \sigma_2(a - c)}{\sigma_2 - \sigma_1}$$

and (84)

$$y = \frac{\sigma_2(b - d) - cb + da}{\sigma_2 - \sigma_1}.$$

For $\eta \leqq 0$,

$$(x + Ka) \geqq 0. \quad (85)$$

For (83) to have a double root,

$$(x + Ka)^2 = 4K\,(y + Kb),$$

which yields

$$K = \frac{(2a - 4y) \pm \sqrt{(2a - 4y)^2 + 4x^2(4b - a^2)}}{2(4b - a^2)}. \quad (86)$$

From (86) it is seen that K is positive if $(4b - a^2)$ is positive. It will be assumed that all the zeros and poles in (80) are complex conjugate, so that introduction of K keeps z_{22} an RC impedance, and $(4b - a^2) > 0$. Equation (85) can be satisfied by a proper choice of σ_2, provided that $d > b$.

If $d < b$, (82) is used, and the polynomials Q_1'' and Q_3'' are obtained by introducing K/p, as in (42):

$$Q_1'' = p^2 + p\left(K + \frac{\sigma_1\sigma_2 - c\sigma_1 + d}{\sigma_2 - \sigma_1}\right) + K\sigma_1,$$

$$Q_3'' = p\left(K + \frac{\sigma_2^2 - c\sigma_2 + d}{\sigma_2 - \sigma_1}\right) + K\sigma_2$$

and

$$P_1Q_3'' - Q_1''P_3 = p^2(K + y) + p(Ka + x) + Kb = 0, \quad (87)$$

where

$$x = \frac{da - bc + \sigma_2(b - d)}{\sigma_2 - \sigma_1},$$

$$y = \frac{d - b + \sigma_2(a - c)}{\sigma_2 - \sigma_1}. \quad (88)$$

Equation (87) has a double root if

$$(Ka + x)^2 = 4Kb(K + y) \quad (89)$$

or

$$K = \frac{(2ax - 4by) \pm \sqrt{(2ax - 4by)^2 + 4(4b - a^2)x^2}}{2(4b - a^2)}. \quad (90)$$

Under the same assumptions as before, $4b - a^2 > 0$, and so $K > 0$. For $\eta \leqq 0$, it is necessary that

$$\frac{Ka + x}{K + y} \geqq 0, \quad (91)$$

but $K + y > 0$ [see (89)], so (91) implies that $(Ka + x) > 0$. It is seen from (88) that $(Ka + x)$ can be made to be positive if $b > d$.

The impedances z_{11}, z_{12}, z_{22} and Z_L can be determined using the polynomials determined above. The two-port can be realized by using zero-shifting techniques. Only one ideal transformer appears, and it is removed by appropriate impedance scaling. Source conversion is then performed to obtain the desired transfer function to within a constant multiplier. The structures obtained are shown in Fig. 9(a) for $d > b$, and in Fig. 9(b) for $b > d$.

The above development can also be carried out when the zeros are not complex conjugate. In this case, however, one must consider separately the cases when $(4b - a^2) > 0$ and $(4b - a^2) < 0$. Results can also be obtained when some of the coefficients are nonpositive. The case of com-

plex zeros and poles in the left-half plane is considered above, since this is the more important case practically.

REFERENCES

1. Dietzold, R. L., U. S. Patent 2,549,065, April 17, 1951.
2. Bangert, J. T., The Transistor as a Network Element, B.S.T.J., **33,** March 1954, p. 329.
3. Linvill, J. G., *RC* Active Filters, Proc. I.R.E., **42,** March 1954, p. 555.
4. Horowitz, I. M., Synthesis of Active *RC* Transfer Functions, Research Report R-507-56, PIB-437, Microwave Research Inst., Polytechnic Inst. of Brooklyn, November 1956.
5. Yanagisawa, T., *RC* Active Networks Using Current-Inversion-Type Negative-Impedance Converters, Trans. I.R.E., **CT-4,** September 1957, p. 140.
6. Bode, H. W., *Network Analysis and Feedback Amplifier Design,* D. Van Nostrand Co., New York, 1956, p. 67.
7. Rosen, A., A New Network Theorem, J. I. E. E., **62,** 1924, p. 916.
8. Linvill, J. G., Transistor Negative-Impedance Converters, Proc. I.R.E., **41,** June 1953, p. 725.
9. Guillemin, E. A., *Communication Networks,* Vol. 2, John Wiley and Sons, New York, 1951, p. 218.
10. Fialkow, A. and Gerst, I., The Transfer Function of General Two Terminal Pair *RC* Networks, Quart. Appl. Math.. **10,** July 1952, p. 113.
11. Larky, A. I., Negative-Impedance Converters, Trans. I.R.E., **CT-4,** September 1957, p. 124.

Mode Conversion at the Junction of Helix Waveguide and Copper Pipe

By J. W. LECHLEIDER

(Manuscript received April 23, 1959)

Scattering coefficients for a junction of helix waveguide and copper pipe are calculated. Fields exterior to the helix in the helix waveguide are neglected, but no other approximations are made.

I. INTRODUCTION

The circular electric wave in circular waveguide has an attenuation which theoretically decreases monotonically with increasing frequency. This fact suggests its use as the information carrier in a high-frequency waveguide communication system. However, as the frequency in any circular waveguide is increased, a greater number of the normal modes become propagating modes. Mode conversion at variations from ideal circular cylinder geometry (such as in bends) thus leads to the presence of energy in propagating modes other than the circular electric mode. This eventually leads to rather severe variation of attenuation with frequency for the circular electric mode.

To reduce this variation of attenuation with frequency, mode filter sections, which attenuate unwanted modes much more rapidly than the circular electric mode is attenuated, may be inserted into the line. Short sections of helix waveguide may be used for this purpose.[1] This technique has proven to be effective in smoothing out the attenuation-frequency characteristic of the circular electric mode.

Using mode filters to effect equalization as discussed above suggests the problem of the calculation of mode conversion at the joint between the mode filter and the propagating guide. The problem of calculating the mode conversion at the intersection of the helix waveguide mode filter sections and solid copper wall waveguide has been attacked by others[2] by the use of approximate methods. These methods generally neglect the reflected modes at the junction of the two guides. In this paper, the reflected modes are included, but any currents excited in a shoulder that exists at the junction are neglected. In Section II of this

paper the representations of the fields associated with the natural modes of propagation in helix waveguide and copper pipe are given. Useful properties of these representations are also given. In Section III the equations relating mode voltages and currents imposed by boundary conditions at the helix-waveguide-copper-pipe junction are derived. In Section IV these equations are solved.

The author is indebted to H. G. Unger and J. A. Young of Bell Telephone Laboratories for their patient assistance in the preparation of this paper. The criticism of Z. Szekely is also appreciated.

GLOSSARY OF SYMBOLS USED

$\mathbf{E}_t$ = transverse electric field,
e_1, e_2 = scale factors associated with the cylindrical coordinates u_1 and u_2,
h = propagation constant of a mode
$\mathbf{H}_t$ = transverse magnetic field,
$\mathbf{i}_1$, $\mathbf{i}_2$ = transverse unit vectors associated with u_1 and u_2,
I = current associated with a mode,
j = $(-1)^{1/2}$,
k^2 = $\omega^2\mu\epsilon$,
l = a scattering coefficient for a helix waveguide and copper pipe junction,
T = a solution of the scalar wave equation,
u_1, u_2 = generalized cylindrical coordinates,
V = voltage associated with a mode,
Y = wave admittance of a mode,
$\mathbf{z}_0$ = unit vector pointed in axial direction of a cylindrical coordinate system,

$$\nabla_t = \frac{\mathbf{i}_1}{e_1}\frac{\partial}{\partial u_1} + \frac{\mathbf{i}_2}{e_2}\frac{\partial}{\partial u_2},$$

ϱ = $\mathbf{i}_1 u_1 + \mathbf{i}_2 u_2$,
χ^2 = a separation constant for the scalar wave equation,
ω = angular temporal frequency,
μ = permeability,
ϵ = permittivity,
δ_{nm} = Kronecker delta.

II. PRELIMINARIES

Before delving into the formulation of the problem at hand, it is necessary to make a few preliminary remarks. In the problem considered

herein, it will be necessary to represent the fields associated with the normal modes of propagation in helix waveguide and in copper pipe. The copper pipe is assumed to have perfectly conducting walls and, hence, its normal modes are transverse magnetic and transverse electric modes.

The transverse electric and magnetic fields associated with a transverse magnetic mode in the copper pipe may be represented as

$$\mathbf{E}_{t(n)} = V_{(n)}(z)\nabla_t T_{(n)}(\varrho), \tag{1}$$

$$\mathbf{H}_{t(n)} = I_{(n)}(z)\nabla_t \times [z_0 T_{(n)}(\varrho)]. \tag{2}$$

In (1) and (2) the temporal dependence of the fields ($e^{j\omega t}$) is understood, as it is in all equations in this paper. The dependence of the fields upon the distance along the waveguide axis is (z) absorbed in the voltage $V_{(n)}(z)$ and current $I_{(n)}(z)$. The function $T_{(n)}(\varrho)$ vanishes for ϱ on the guide wall and is a solution of the scalar wave equation:

$$\nabla_t^2 T_{(n)}(\varrho) = \frac{1}{e_1 e_2}\left[\frac{\partial}{\partial u_1}\left(\frac{e_2}{e_1}\frac{\partial T_{(n)}}{\partial u_1}\right) + \frac{\partial}{\partial u_2}\left(\frac{e_1}{e_2}\frac{\partial T_{(n)}}{\partial u_2}\right)\right] = -\chi_{(n)}^2 T_{(n)}. \tag{3}$$

In this equation, $\chi_{(n)}^2$ is a separation constant related to the propagation constant of the mode being discussed by

$$\chi_{(n)}^2 = k^2 - h_{(n)}^2. \tag{4}$$

The voltage and the current associated with the mode are related via the wave admittance for the mode

$$I_{(n)} = \pm Y_{(n)} V_{(n)} = \pm \frac{\omega\epsilon}{h_{(n)}} V_{(n)}, \tag{5}$$

where the sign depends upon the direction in which the wave is moving.

For transverse electric modes in the copper pipe, the transverse electric and magnetic fields can be expressed in the form

$$\mathbf{E}_{t[n]} = V_{[n]}(z)\nabla_t \times \mathbf{z}_0 T_{[n]}(\varrho), \tag{6}$$

$$\mathbf{H}_{t[n]} = I_{[n]}(z)\nabla_t T_{[n]}(\varrho). \tag{7}$$

The scalar function $T_{[n]}(\varrho)$ is a solution of (3) with separation constant $\chi_{[n]}^2$, and the normal derivative of $T_{[n]}(\varrho)$ at the waveguide wall vanishes. Relations similar to (4) and (5) can also be written for transverse electric modes:

$$\chi_{[n]}^2 = k^2 - h_{[n]}^2, \tag{8}$$

$$I_{[n]} = \pm Y_{[n]} V_{[n]} = \pm \frac{h_{[n]}}{\omega\mu} V_{[n]}. \tag{9}$$

The transverse vector functions defining the transverse electric fields associated with the transverse electric and transverse magnetic modes in copper pipe form a complete orthogonal set over the internal cross section of the copper pipe. As such, they may be used to represent any transverse electric field in the copper pipe. The same statement can be made of the associated modal transverse magnetic fields and the transverse magnetic field in the copper pipe. The functions $T_{(n)}(\varrho)$ and $T_{[n]}(\varrho)$ are assumed to be normalized, so that the following orthonormality conditions hold:

$$\begin{aligned}
\chi_{(n)}^{2} \iint T_{(n)}T_{(m)}\, ds &= \chi_{[n]}^{2} \iint T_{[n]}T_{[m]}\, ds \\
&= \iint (\nabla_t T_{(n)})(\nabla_t T_{(m)})\, ds \\
&= \iint (\nabla_t \times \mathbf{z}_0 T_{(n)})(\nabla_t \times \mathbf{z}_0 T_{(m)})\, ds \\
&= \iint (\nabla_t T_{[n]})(\nabla_t T_{[m]})\, ds \\
&= \iint (\nabla_t \times \mathbf{z}_0 T_{[n]})(\nabla_t \times \mathbf{z}_0 T_{[m]})\, ds \\
&= \delta_{nm}\,,
\end{aligned} \tag{10}$$

where the integration is over the internal cross section of the copper pipe.

In addition to the relations listed in (10), the following relations will be found useful in this paper:

$$\begin{aligned}
\iint (\nabla_t T_{(n)})(\nabla_t \times \mathbf{z}_0 T_{[m]})\, ds &= \iint (\nabla_t T_{[n]})(\nabla_t \times T_{[m]})\, ds \\
&= \iint (\nabla_t T_{[m]})(\nabla_t \times \mathbf{z}_0 T_{(n)})\, ds = 0\,,
\end{aligned} \tag{11}$$

which hold for all m and n.

For helix waveguide, two scalar functions are necessary for the specification of the transverse fields since all modes but the circular electric modes are hybrid.

The transverse electric and magnetic fields in helix waveguide may

be written in the form

$$\mathbf{E}_{tn} = V_n(\nabla_t T_n + d_n \nabla_t \times \mathbf{z}_0 T'_n), \tag{12}$$

$$\mathbf{H}_{tn} = I_n\left(-\nabla_t \times \mathbf{z}_0 T_n + d_n \frac{h_n^2}{k^2} \nabla_t T'_n\right), \tag{13}$$

which are similar to Unger's representation.[2]

The functions T_n and T_n' are both solutions of (3) with separation constant χ_n^2; they are related to each other by means of the boundary and continuity requirements on the electromagnetic field in the helix waveguide. The constant d_n in (12) and (13) is determined by the fact that the electric field component tangential to the helix wire must vanish at the helix. For a helix of zero pitch,

$$d_n = \left.\frac{\dfrac{\partial T_n}{\epsilon_2 \partial u_2}}{\dfrac{\partial T_n'}{\epsilon_1 \partial u_1}}\right| \text{ at helix.} \tag{14}$$

Expressions similar to (4) and (5) can also be written for the helix waveguide modes:

$$\chi_n^2 = k^2 - h_n^2, \tag{15}$$

$$I_n = \pm Y_n V_n = \pm \frac{\omega \epsilon_0}{h_n} V_n . \tag{16}$$

Equation (16) is written for the region interior to the helix, which is assumed to be empty (i.e., $\epsilon = \epsilon_0$ and $\mu = \mu_0$).

Finally, for any two solutions of (3), which are designated here as T_1 and T_2,

$$\begin{aligned} \iint \nabla_t T_1 \cdot \nabla_t T_2 \, ds &= \iint (\nabla_t \times \mathbf{z}_0 T_1)(\nabla_t \times \mathbf{z}_0 T_2 \, ds \\ &= \chi_2^2 \iint T_1 T_2 \, ds + \oint T_1 \frac{\partial T_2}{e_1 \partial u_1} e_2 \, du_2 \\ &= \chi_1^2 \iint T_1 T_2 \, ds + \oint T_2 \frac{\partial T_1}{e_1 \partial u_1} e_2 \, du_2 , \end{aligned} \tag{17}$$

and hence

$$\iint T_1 T_2 \, ds = \oint \frac{\left(T_1 \dfrac{\partial T_2}{e_1 \partial u_1} - T_2 \dfrac{\partial T_1}{e_1 \partial u_1}\right) e_2 \, du_2}{\chi_1^2 - \chi_2^2} \tag{18}$$

and

$$\iint (\nabla_t \times \mathbf{z}_0 T_1)(\nabla_t T_2)\, ds = \oint T_2 \frac{\partial T_1}{\partial u_2}\, du_2 = -\oint T_1 \frac{\partial T_2}{\partial u_2}\, du_2 . \quad (19)$$

In (17), (18) and (19) the line integrals are to be evaluated along the waveguide wall, or along the boundary enclosing the domain of integration of the integrals over the area.

III. THE BOUNDARY PROBLEM — HELIX WAVEGUIDE MODE INCIDENT

The problem under consideration is the calculation of the scattering properties of a junction between helix waveguide and copper pipe. The region interior to the helix in the helix waveguide and the interior of the copper pipe are assumed to be empty. Furthermore, the helix diameter and inside diameter of the copper pipe are assumed to be equal.

To formulate the problem under consideration analytically, assume that a single helix waveguide mode is incident upon the junction from the helix waveguide. This will give rise to reflected helix waveguide modes and transmitted transverse electric and transverse magnetic modes in the copper pipe. The requirement of continuity of the transverse electric field across the plane of the junction indicates that the total transverse electric field due to the incident and reflected helix waveguide modes must equal the total transverse electric field due to the transmitted transverse electric and transverse magnetic modes at the plane of the junction. This equality is represented analytically by

$$V_i(\nabla_t T_i + d_i \nabla_t \times \mathbf{z}_0 T_i') + \sum_n V_n(\nabla_t T_n + d_n \nabla_t \times \mathbf{z}_0 T_n') = \sum_{(n)} V_{(n)} \nabla_t T_{(n)} + \sum_{[n]} V_{[n]} \nabla_t \times \mathbf{z}_0 T_{[n]} . \quad (20)$$

Equation (20) is true in the area interior to the helix of the helix waveguide. In (20) the first term on the left is due to the incident mode, the summation on the left is due to the reflected helix waveguide modes, the first summation on the right is due to the transmitted transverse magnetic modes and the second summation is due to the transmitted transverse electric modes.

In a similar fashion, continuity of the transverse magnetic field across the plane of the junction is expressed as

$$I_i \left(-\nabla_t \times \mathbf{z}_0 T_i + d_i \frac{h_i^2}{k^2} \nabla_t T_i'\right) + \sum_n I_n \left(-\nabla_t \times \mathbf{z}_0 T_n + d_n \frac{h_n^2}{k^2} \nabla_t T_n'\right) = -\sum_{(n)} I_{(n)} \nabla_t \times \mathbf{z}_0 T_{(n)} + \sum_{[n]} I_{[n]} \nabla_t T_{[n]} . \quad (21)$$

Equations (20) and (21), in addition to the specification of the boundary conditions on the helix waveguide field in the region exterior to the helix, would be sufficient to determine the transmitted and reflected mode voltages and currents. In this paper it is assumed that the fields exterior to the helix in the helix waveguide are unimportant. This makes (20) and (21) sufficient for the determination of the transmitted and reflected mode voltages and currents.

Equations (20) and (21) can be converted to linear algebraic equations in the unknown voltages and currents by using some of the relationships presented in the preliminaries. For example, if (20) is multiplied scalarly on both sides by $\nabla_t T_{(p)}$ and the result is integrated over the internal cross section of the copper pipe, there results

$$\frac{V_i\chi_i^2 \oint T_i \dfrac{\partial T_{(p)}}{e_1 \partial u_1} e_2\, du_2}{\chi_i^2 - \chi_{(p)}^2} + \sum_n \frac{V_n\chi_n^2 \oint T_n \dfrac{\partial T_{(p)}}{e_1 \partial u_1} e_2\, du_2}{\chi_n^2 - \chi_{(p)}^2} = V_{(p)}\,, \tag{22}$$

where (10), (11), (17), (18) and (19) have been used. The integrals in (22) are along the waveguide wall, which is assumed to be a surface of the form $u_1 =$ a constant.

In a similar fashion, the following equations can be derived:

$$\frac{I_i\chi_i^2 \oint T_i \dfrac{\partial T_{(p)}}{e_1 \partial u_1} e_2\, du_2}{\chi_i^2 - \chi_{(p)}^2} + \sum_n \frac{I_n\chi_n^2 \oint T_n \dfrac{\partial T_{(p)}}{e_1 \partial u_1} e_2\, du_2}{\chi_n^2 - \chi_{(p)}^2} = I_{(p)}\,, \tag{23}$$

$$\frac{V_i\chi_i^2 \oint \dfrac{\partial T_i}{\partial u_2} T_{[p]}\, du_2}{\chi_{[p]}^2 - \chi_i^2} + \sum_n \frac{V_n\chi_n^2}{\chi_{[p]}^2 - \chi_n^2} \oint \frac{\partial T_n}{\partial u_2} T_{[p]}\, du_2 = V_{[p]}\,, \tag{24}$$

$$\frac{I_i\chi_i^2 \oint \dfrac{\partial T_i}{\partial u_2} T_{[p]}\, du_2}{\chi_{[p]}^2 - \chi_i^2} + \sum_n \frac{I_n\chi_n^2 \oint \dfrac{\partial T_n}{\partial u_2} T_{[p]}\, du_2}{\chi_{[p]}^2 - \chi_n^2} = \frac{k^2 I_{[p]}}{h_{[p]}^2}\,. \tag{25}$$

The voltages and currents of the transmitted transverse electric and transverse magnetic modes can be eliminated from (22) through (25) by use of (5) and (9). Multiplying (22) on both sides by $Y_{(p)}$ and subtracting (23) from the result yields, with application of (16),

$$\frac{V_i\chi_i^2 \oint T_i}{h_i(h_{(p)} + h_i)} \frac{\partial T_{(p)}}{e_1 \partial u_1} e_2\, du^2 - \sum_n \frac{V_n\chi_n^2 \oint T_n \dfrac{\partial T_{(p)}}{e_1 \partial u_1} e_2\, du_2}{h_n(h_{(p)} - h_n)} = 0. \tag{26}$$

The incident helix waveguide mode is treated as a forward-moving wave in deriving (26); the reflected helix waveguide modes are then back-

ward-moving waves and the transmitted copper pipe modes are forward-moving waves.

Consideration of the nature of the T functions can now be used to simplify the set of equations represented by (26). In the coordinate systems of practical interest here, e.g., a circular cylindrical geometry, the T functions for all modes separate, so that

$$T_n = T_n^{(1)}(u_1)T_n^{(2)}(u_2), \tag{27}$$

$$T_{(p)} = T_{(p)}^{(1)}(u_1)T_{(p)}^{(2)}(u_2). \tag{28}$$

The integrals in (26) will vanish unless

$$T_n^{(2)}(u_2) = T_{(p)}^{(2)}(u_2), \tag{29}$$

so that only modes for which (29) is satisfied will appear in (26). Furthermore, the integrals in (26) assume the form

$$\oint T_n \frac{\partial T_{(p)}}{e_1 \partial u_1} e_2 \, du_2 = T_n^{(1)} \frac{\partial T_{(p)}^{(1)}}{\partial u_1} K, \tag{30}$$

where K is a constant independent of n. Using this fact reduces (26) to the form

$$\frac{V_i T_i^{(1)} \chi_i^2}{h_i(h_{(p)} + h_i)} - \sum_n \frac{V_n T_n^{(1)} \chi_n^2}{h_n(h_{(p)} - h_n)} = 0. \tag{31}$$

Applying reasoning similar to that used in deriving (31) to (24) and (25) yields the following set of equations:

$$\frac{V_i T_i^{(1)} \chi_i^2}{h_i(h_i + h_{(p)})} - \sum_n \frac{V_n T_n^{(1)} \chi_n^2}{h_n(h_{(p)} - h_n)} = 0. \tag{32}$$

Equations (31) and (32) must be solved for the unknown reflected mode voltages V_n.

IV. SOLUTION OF THE EQUATIONS

The simultaneous set of equations indicated by (31) and (32) may be solved by successively eliminating the V_n's. For example, multiply the rth equation of the set (31) by

$$\frac{h_{(r)} - h_a}{h_{(p)} - h_a} \tag{33}$$

and subtract the result from the pth equation. The result obtained, after

some manipulation, is

$$\sum_n \frac{V_n T_n^{(1)} \chi_n^2 (h_a - h_n)}{h_n (h_{(r)} - h_n)(h_{(p)} - h_n)} = \frac{V_i T_i^{(1)} \chi_i^2 (h_i + h_a)}{h_i (h_{(p)} + h_i)(h_{(r)} + h_i)}. \quad (34)$$

Similarly, multiplication of the rth equation of the set (31) by

$$\frac{(h_{(r)} - h_a)}{(h_{[p]} - h_a)}$$

and subtraction of the result from the pth equation of the set indicated by (32) ultimately leads to

$$\sum_n \frac{V_n T_n^{(1)} \chi_n^2 (h_a - h_n)}{h_n (h_{(r)} - h_n)(h_{[p]} - h_n)} = \frac{V_i T_i^{(1)} \chi_i^2 (h_i + h_a)}{h_i (h_{[p]} + h_i)(h_{(r)} + h_i)}. \quad (35)$$

The set of equations indicated by (34) and (35) consists of one less equation and one less unknown (V_a) than the set indicated by (31) and (32). The process of elimination of successive V_n's can proceed in this fashion until only one of the V_n's is left, and the system of equations indicated by (31) and (32) is solved. The final result is

$$V_n = \frac{h_n V_i T_i^{(1)} \chi_i^2}{h_i T_n^{(1)} \chi_n^2} \prod_{\substack{m \\ m \neq n}} \left(\frac{h_m + h_i}{h_m - h_n} \right) \prod_{(m)} \left(\frac{h_{(m)} - h_n}{h_{(m)} + h_i} \right) \prod_{[m]} \left(\frac{h_{[m]} - h_n}{h_{[m]} + h_i} \right). \quad (36)$$

Inspection of (36) reveals the fact that no helix modes which have a propagation constant equal to the propagation constant of a copper pipe mode will be reflected from a helix-waveguide-copper-pipe junction. This result is especially important in the case of the circular electric modes in helix waveguide, since these modes have the same propagation constant in helix waveguide and copper pipe. This is quite reasonable, since the boundary conditions on the circular electric modes are essentially the same in helix waveguide and copper pipe.

Having formally obtained the reflected mode voltages, the transmitted mode voltages can be calculated by inserting the expressions for the V_n's as given by (36) into (22) and (24). Thus, inserting (36) into (22) yields the expression for the transmitted transverse magnetic mode voltages. The result after some manipulation is

$$V_{(p)} = V_i \chi_i^2 \oint T_i \frac{\partial T_{(p)}}{e_1 \partial u_1} e_2 \, du_2 \left[\frac{1}{(h_{(p)}^2 - h_i^2)} + \sum_n \frac{1}{(h_{(p)}^2 - h_n^2)} \cdot \prod_{\substack{m \\ m \neq n}} \left(\frac{h_m + h_i}{h_m - h_n} \right) \prod_{(m)} \left(\frac{h_{(m)} - h_n}{h_{(m)} + h_i} \right) \prod_{[m]} \left(\frac{h_{[m]} - h_n}{h_{[m]} + h_i} \right) \right]. \quad (37)$$

In the Appendix it is shown that this can be simplified to

$$V_{(p)} = \frac{V_i {\chi_i}^2 \oint T_i \frac{\partial T_{(p)}}{e_1 \partial u_1} e_2 \, du_2}{h_i(h_{(p)}^2 - h_i^2)} \prod_m \left(\frac{h_m + h_i}{h_m + h_{(p)}}\right) \cdot \prod_{\substack{(m) \\ (m)\neq(p)}} \left(\frac{h_{(m)} + h_{(p)}}{h_{(m)} + h_i}\right) \prod_{[m]} \left(\frac{h_{[m]} + h_{(p)}}{h_{[m]} + h_i}\right). \tag{38}$$

Similar manipulations yield the following expression for the transmitted transverse electric mode voltages:

$$V_{[p]} = \frac{V_i {\chi_i}^2 \oint \frac{\partial T_i}{\partial u_2} T_{[p]} \, du_2}{h_i(h_i^2 - h_{[p]}^2)} \prod_m \left(\frac{h_m + h_i}{h_m + h_{[p]}}\right) \cdot \prod_{(m)} \left(\frac{h_{(m)} + h_{[p]}}{h_{(m)} + h_i}\right) \prod_{\substack{[m] \\ [m]\neq[p]}} \left(\frac{h_{[m]} + h_{[p]}}{h_{[m]} + h_i}\right). \tag{39}$$

Equations (36), (38) and (39) are the formal solution to the problem of scattering at a junction of helix waveguide and copper pipe for a helix waveguide mode incident.

The most important results for practical application are the expressions for the transmitted voltages as given by (38) and (39). The quantities on the right in (38) and (39) are the voltage transmission coefficients of the junction multiplied by the incident mode voltages. These coefficients can be converted to scattering coefficients for travelling waves, as calculated by Unger,[2] by multiplying each by the square root of the appropriate wave admittance ratios. Thus, the scattering coefficients are

$$l_{(p)i} = \frac{{\chi_i}^2}{(h_i h_{(p)})^{1/2}} \frac{\oint T_i \frac{\partial T_{(p)}}{e_1 \partial u_1} e_2 \, du_2}{(h_{(p)}^2 - h_i^2)} \prod_{\substack{m \\ m\neq n}} \left(\frac{h_m + h_i}{h_m + h_{(p)}}\right) \cdot \prod_{\substack{(m) \\ (m)\neq(p)}} \left(\frac{h_{(m)} + h_{(p)}}{h_{(m)} + h_i}\right) \prod_{[m]} \left(\frac{h_{[m]} + h_{[p]}}{h_{[m]} + h_i}\right) \tag{40}$$

and

$$l_{[p]i} = \left(\frac{h_{[p]}}{h_i}\right)^{1/2} \frac{{\chi_i}^2}{k} \frac{\oint \frac{\partial T_i}{\partial u_2} T_{[p]} \, du_2}{(h_i^2 - h_{[p]}^2)} \prod_m \left(\frac{h_m + h_i}{h_m + h_{[p]}}\right) \cdot \prod_{(m)} \left(\frac{h_{(m)} + h_{[p]}}{h_{(m)} + h_i}\right) \prod_{\substack{[m] \\ [m]\neq[p]}} \left(\frac{h_{[m]} + h_{[p]}}{h_{[m]} + h_i}\right). \tag{41}$$

Furthermore, as pointed out by Unger,[2] proper normalization of the T functions for the helix waveguide insures that these coefficients will be the same whether they indicate conversion from an incident helix waveguide mode to a transmitted copper mode or conversion from an incident copper pipe mode to a transmitted helix waveguide mode.

V. CONCLUSIONS

Circular electric modes are transmitted through a junction of helix waveguide with a zero-pitch helix and copper pipe without any mode conversion or reflection at the junction. Neglecting the fields exterior to the helix in helix waveguide permits solution of the infinite set of equations which arises from requiring continuity of the transverse fields at the junction between helix waveguide and copper pipe.

APPENDIX

Reduction of the Formula for the Transmitted Voltages

The summation appearing in (37) is

$$S = \sum_n \frac{1}{(h_{(p)}^2 - h_n^2)} \prod_{\substack{m \\ m \neq n}} \left(\frac{h_m + h_i}{h_m - h_n}\right) \prod_{(m)} \left(\frac{h_{(m)} - h_n}{h_{(m)} + h_i}\right) \cdot \prod_{[m]} \left(\frac{h_{[m]} - h_n}{h_{[m]} + h_i}\right). \tag{42}$$

Since all the h's appearing in (42) are for forward-moving waves, they will all lie in the fourth quadrant of the complex h plane ($e^{-ihz} = e^{-i(-i\alpha+\beta)z}$). Thus, S can be seen to be $2\pi i$ times the sum of the fourth quadrant residues of the function

$$f(h) = \frac{1}{2\pi i(h_{(p)}^2 - h^2)} \prod_m \left(\frac{h_m + h_i}{h_m - h}\right) \prod_{(m)} \left(\frac{h_{(m)} - h}{h_{(m)} + h_i}\right) \cdot \prod_{[m]} \left(\frac{h_{[m]} - h}{h_{[m]} + h_i}\right) \frac{1}{(h + h_i)}. \tag{43}$$

Thus, S is simply a contour integral in the complex h plane:

$$S = \frac{1}{2\pi i} \oint_{c_1} \frac{1}{(h_{(p)}^2 - h^2)} \prod_m \left(\frac{h_m + h_i}{h_m - h}\right) \prod_{(m)} \left(\frac{h_{(m)} - h}{h_{(m)} + h_i}\right) \cdot \prod_{[m]} \left(\frac{h_{[m]} - h}{h_{[m]} + h_i}\right) \frac{dh}{(h + h_i)}, \tag{44}$$

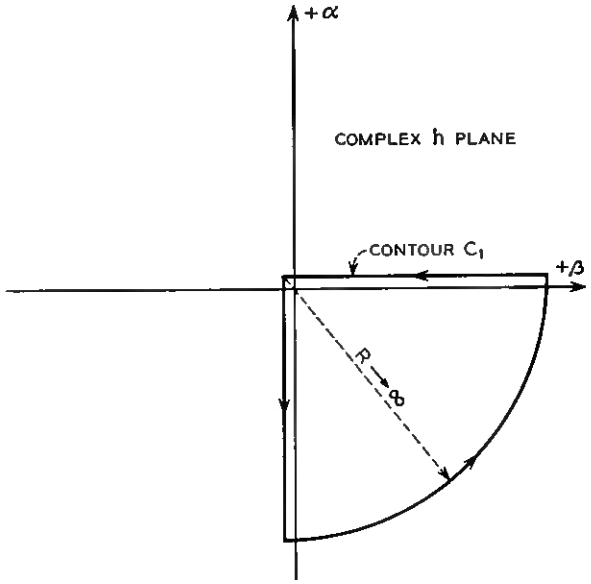


FIG. 1 — Contour for (44).

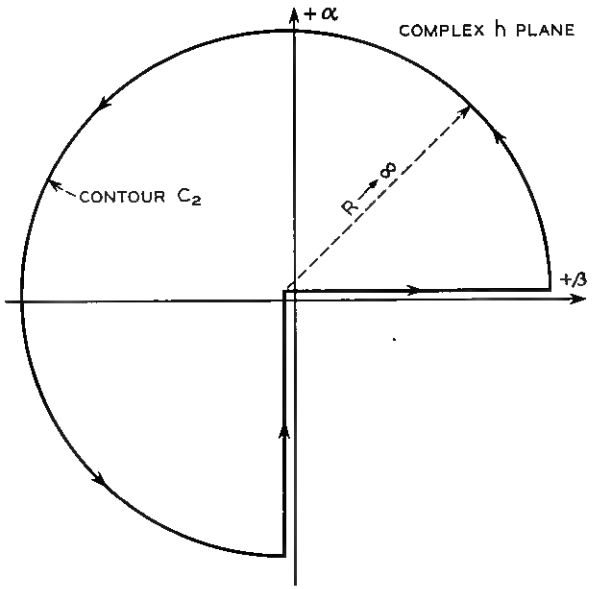


Fig. 2 — Contour for (45).

where the contour is that shown in Fig. 1. Since the contribution to the integral along the circumference of the circle is zero, it can be seen that the value of the integral in (44) is unchanged if the contour is chosen to be that of the other three quarters of the circle enclosing the h plane. Thus, the integral in (44) may be written in the form

$$S = -\frac{1}{2\pi i}\oint_{c_2}\frac{1}{(h_{(p)}^2 - h^2)}\prod_{m}\left(\frac{h_m + h_i}{h_m - h}\right)\prod_{(m)}\left(\frac{h_{(m)} - h}{h_{(m)} + h_i}\right) \cdot\prod_{[m]}\left(\frac{h_{[m]} - h}{h_{[m]} + h_i}\right)\frac{dh}{(h + h_i)}, \tag{45}$$

where the contour is now that shown in Fig. 2.

The integrand in (45) has just two poles in the first, second and third quadrants of the complex h plane, namely, the poles at

$$h = -h_{(p)} \qquad \text{and} \qquad h = -h_i\,,$$

so that S is just $2\pi i$ times the sum of the residues of the integrand at $h = -h_{(p)}$ and $h = -h_i$. Thus,

$$S = -\frac{1}{(h_{(p)}^2 - h_i^2)} \cdot\left[1 - \prod_{m}\left(\frac{h_m + h_i}{h_m + h_{(p)}}\right)\prod_{\substack{(m)\\(m)\neq(p)}}\left(\frac{h_{(m)} + h_{(p)}}{h_{(m)} + h_i}\right)\prod_{[m]}\left(\frac{h_{[m]} + h_{[p]}}{h_{[m]} + h_i}\right)\right]. \tag{46}$$

Consideration of (46) in (37) yields (38).

REFERENCES

1. Miller, S. E., Waveguide as a Communications Medium, B.S.T.J., **33**, November 1954, p. 1209.
2. Unger, H. G., Helix Waveguide Theory and Applications, B.S.T.J., **37**, November 1958, p. 1599.

Recent Monographs of Bell System Technical Papers Not Published in This Journal*

Abbott, H. H.

New Small Crossbar Telephone System for Private Branch Exchanges, Monograph 3252.

Al-Madfai, S., see Frisch, H. L.

Anderson, P. W.

Random-Phase Approximation in the Theory of Superconductivity, Monograph 3264.

Ashkin, A.

Dynamics of Electron Beams from Magnetically Shielded Guns, Monograph 3220.

Ashkin, A.

Parametric Amplification of Space Charge Waves, Monograph 3221.

Ashkin, A., Bridgers, T. J., Louisell, W. H. and Quate, C. F.

Parametric Electron Beam Amplifiers, Monograph 3265.

Batdorf, R. L. and Smits, F. M.

Diffusion of Impurities into Evaporating Silicon, Monograph 3248.

Bond, W. L., see Geller, S.

Boyle, W. S. and Rodgers, K.

Performance Characteristics of a New Low-Temperature Bolometer, Monograph 3222.

* Copies of these monographs may be obtained on request to the Publication Department, Bell Telephone Laboratories, Inc., 463 West Street, New York 14, N. Y. The numbers of the monographs should be given in all requests.

Contributors to This Issue

Arthur B. Crawford, B.S.E.E., 1928, Ohio State University; Bell Telephone Laboratories, 1928—. Mr. Crawford has been engaged in radio research since joining Bell Laboratories. He has worked on measuring techniques, propagation and antenna studies in the ultra short wave and microwave fields. Fellow I.R.E.; member Sigma Xi, Tau Beta Pi, Eta Kappa Nu, Pi Mu Epsilon.

Earl T. Harkless, B.S. in E.E., 1947, and M.S., 1949, Case Institute of Technology; Bell Telephone Laboratories, 1949—. He has been engaged primarily in the design of waveguide filters and networks for use in telephone communications systems. Member I.R.E., Tau Beta Pi, Sigma Xi.

David C. Hogg, B.S., 1949, University of Western Ontario; M.S., 1950, and Ph.D., 1953, McGill University; Bell Telephone Laboratories, 1953—. Mr. Hogg has been engaged in studies of artificial dielectrics for microwaves, antenna problems and over-the-horizon and millimeter-wave propagation. Senior member I.R.E.; member, Sigma Xi.

Phyllis A. Groll, A.B., 1957, Cedar Crest College; Bell Telephone Laboratories, 1957—. As a member of the reliability and statistics group Miss Groll has been engaged in statistical research involving the use of high-speed computers.

B. K. Kinariwala, B.S., 1951, Benares University (India); M.S., 1954, and Ph.D., 1957, University of California; Bell Telephone Laboratories, 1957—. He has been engaged in studies of circuit theory problems including active and time-varying networks. Member I.R.E., Sigma Xi.

W. H. Kummer, B.S., 1946, M.S., 1947, and Ph.D., 1954, University of California; Bell Telephone Laboratories, 1953–59. He was engaged primarily in studies in the field of microwaves, particularly on waveguides, antennas and beyond-the-horizon propagation. Member I.R.E., A.I.E.E., Tau Beta Pi, Eta Kappa Nu, Sigma Xi, Phi Beta Kappa.

J. W. Lechleider, B.S.M.E., 1954, Cooper Union; M.E.E., 1957, Brooklyn Polytechnic Institute; Bell Telephone Laboratories, 1955—. He has been primarily engaged in fundamental cable development, in studies of crosstalk, high-frequency propagation (especially in the field of helix waveguide development) and heat transfer. Member Sigma Xi, Pi Tau Sigma.

Milton Sobel, B.S., 1940, College of the City of New York; M.A., 1946 and Ph.D., 1951, Columbia University; Wayne University, assistant professor of mathematics, 1950–52; Columbia University, visiting lecturer, department of mathematical statistics, 1952; Cornell University, 1952–54; Bell Telephone Laboratories, 1954—. He is engaged in mathematical research on various statistical problems, including life-testing and reliability, and is a consultant on many Laboratories projects. Adjunct associate professor and senior research scientist, Institute of Mathematical Sciences, New York University. Fellow Institute Mathematical Statistics; member American Statistical Association, Sigma Xi.