A Method for Predicting Interchannel Modulation due to Multipath Propagation in FM and PM Tropospheric Radio Systems*

By C. D. BEACH and J. M. TRECKER

This paper describes a method for predicting the magnitude of interchannel modulation due to multipath propagation on angle-modulated tropospheric scatter radio systems. Values of signal-to-intermodulation ratio, S/I, are calculated for various pairs of signal reflections in the troposphere, taking into account the base bandwidth and frequency deviation of the system, the antenna patterns, the path geometry, and climatic conditions during the "worst month" of propagation. The lowest value of S/I (worst intermodulation) computed for such pairs of signal reflections is then corrected empirically to account for multiple reflections. The result represents the median value of S/I expected during the worst month of transmission on a specified path.

The method yields results that, when compared to measured results from four widely different paths normalized to worst-month conditions, have a standard error of estimate of about 2.6 db.

TABLE OF CONTENTS

I. INTRODUCTION

Various theories have been proposed to explain the mechanism by which energy radiated from a transmitting antenna is reflected or re-fracted in the atmosphere and ultimately arrives at a receiving antenna beyond the horizon.[1,2,3] Apart from the merits of any particular theory of propagation, the geometry of a typical over-the-horizon, or "tropospheric scatter," path suggests an inherent multipath transmission problem. As illustrated by Fig. 1, components of signal reflected to the receiving antenna from successively higher points in the atmosphere are delayed by increasing amounts behind the earliest arriving component.

When such time-delayed RF signals — called echoes henceforth — are demodulated by an FM or PM receiver, they are converted into energy at undesired regions of the baseband. In a frequency-division multiplexed, multi-channel voice system, this undesired energy is commonly called "interchannel modulation," or simply "intermodulation." It adds to other distortion energy generated by equipment nonlinearities and to the ever-present white noise, to produce the total background of noise faced by the desired speech signal. Whether or not "path intermodulation" noise is objectionable in a given system depends on its magnitude relative to these other components of noise.

The general theory of interchannel modulation due to transmission echoes in angle-modulated systems has been treated by Bennett, Curtis,
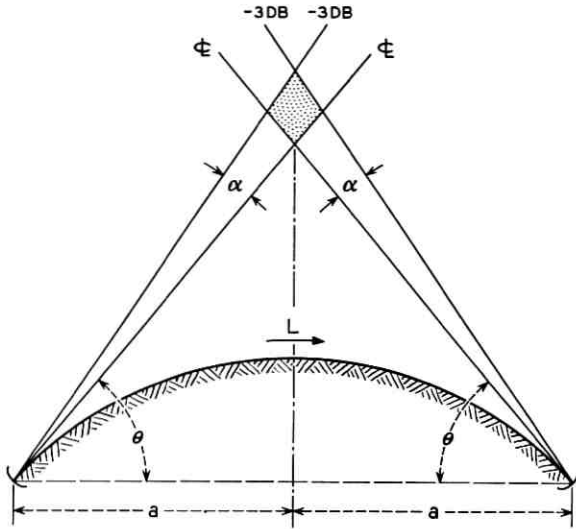
Fig. 1 — Symmetrical tropospheric scatter path.

and Rice.[4] The objective of the present paper is to outline a method for predicting the magnitude of intermodulation due to echoes in tropospheric scatter radio systems, making use of the intermodulation theory developed in Ref. 4 and the reflection theory of tropospheric propagation given in Ref. 3.

The first step in the method, following specification of a suitable model, involves calculation of the relative amplitudes of signal echoes received from various angles of elevation above the horizon. The corresponding time delays of these echoes are then calculated, relative to the signal arriving via the antenna beam axes, or "centerline" path.

Next, signal-to-intermodulation ratios are determined for selected combinations of echo amplitude and corresponding time delay, radio system base bandwidth, and frequency deviation.

These initial, primarily analytical, steps provide an indication of the magnitude of intermodulation which would be generated by single echoes in an angle-modulated system without diversity. The results postulate a so-called "worst echo"; that is, an echo having a combination of amplitude and time delay that produces more interchannel interference than any other combination.

In calculating the values of echo amplitude and time delay, the meteorological conditions of the path are assumed to be the median values during the worst month of propagation. Thus, the calculated in-

termodulation represents the median value for a single worst echo during the worst month of transmission on the given path.

From this point, the method becomes empirical in accounting for variations in the calculated median value of intermodulation as propagation conditions change, and in accounting for the effects of multiple echoes and diversity reception. Satisfactory analytical methods of handling these factors have been elusive, to say the least. However, experimental results have provided encouraging means of estimating correction factors for their effects, and these corrections are discussed.

II. MODEL OF TROPOSPHERIC SCATTER PATH

The typical geometry of a tropospheric scatter path is illustrated in Fig. 1, showing two narrow antenna beams intersecting to form a "common volume." The lower edges of the pattern represent the beam centerlines; each is pointed toward the horizon on the azimuth of the other antenna. The upper edges of the patterns are arbitrarily chosen to correspond to some value of antenna gain — e.g., 3 db — less than the centerline maximum. Throughout this paper, the antenna pattern above the centerline is assumed to be the same as half the free-space pattern.

2.1 Reflection Theory

The reflection theory of beyond-the-horizon transmission proposed by Friis, Crawford, and Hogg[3] has been found to yield estimates of received power which are in good agreement with reported experimental data. The expression for received power proposed by this theory will be adopted in the present model for use later in deriving an expression for echo amplitude.

Assuming identical transmitting and receiving antennas, and "intermediate-size" reflecting layers,* the received power in a system with symmetrical geometry as in Fig. 1 is given by:

$$P_r = P_t \frac{M}{3} \frac{\lambda^3}{\alpha^3} \frac{1}{\theta^5 \alpha^2} \frac{1}{\left(2 + \frac{\alpha}{\theta}\right)} f\left(\frac{\alpha}{\theta}\right) \tag{1}$$

where

P_t = transmitted power, in watts,

M = a constant of the path proportional to the number of reflecting

---

* Ref. 3 proposes somewhat different expressions for the cases of large, intermediate, and small reflecting layers, where the dimensions are defined in terms of wavelength and Fresnel zones. All three may be present on a given path at various times, but intermediate-sized layers are probably most prevalent.

layers and the squares of path length, reflecting layer length, and change in gradient of refractive index at the bottom of the common volume,

λ = wavelength, meters,

α = antenna beamwidth to the half-power points, radians,

θ = angle between a chord joining the ends of the path and the axis of the antenna beam, radians,

a = half the length of the chord between the ends of the path, in meters. For practical paths, this can also be taken as half the ground distance between terminals,

and

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

Equation (1) indicates that the received power is critically dependent on the angle θ, and it will become apparent later that path intermodulation also depends heavily on this parameter. Therefore, it is of interest to consider those factors which influence the value of θ on a path of given length. These factors — atmospheric refractivity, unequal elevation of antenna sites, and nonhorizontal antenna take-off angles — will now be discussed.

2.2 Effect of Refractivity on Angle θ

As is customary in many radio problems where the transmitted signal is refracted by the atmosphere, the lines of propagation in Fig. 1 have been drawn straight over an earth having an effective radius greater than the true earth's radius. The symmetrical path is then defined completely by the path length L (or alternatively, the chord length $2a$) and the angle θ between the chord and the earth-tangent ray at either end of the path.

The value of θ which gives the correct altitude of intersection for horizontally directed antenna beams, allowing for refractive bending, is

$$\theta = \frac{L}{2R} = \frac{L}{2R_oK} \tag{2}$$

where R is the effective earth's radius, R_o is the true earth's radius (3960 statute miles), and K is a constant.

The value $K = \frac{4}{3}$ is used in many radio applications. It has been found in the present work that better results are obtained if the value of K is

Fig. 2 — Minimum monthly mean value of N. (Taken from Ref. 6)

determined for the specific path under study. We need, then, to determine the relation between K and refractivity.

A method for determining the approximate value of K for a given path is developed in Appendix A. The initial step in the procedure requires the selection of appropriate values of minimum monthly mean refractivity at sea level, N_o, for the geographical coordinates of the stations at each end of the path. Fig. 2 gives such data, collected from 306 weather stations over a period of years.

Next, values of refractivity, N_s, are calculated for the altitude at each station. This is done using equation (22), repeated here for convenience as:

$$N_s = N_o \exp\left(-0.0322h\right) \tag{3}$$

where N_o and N_s are in "N-units," and h is the station altitude, in thousands of feet.

Finally, the values of K corresponding to the surface refractivity at each station are read from Fig. 3, and a single average for the path is estimated from the values at the end points. Unless the path is very long
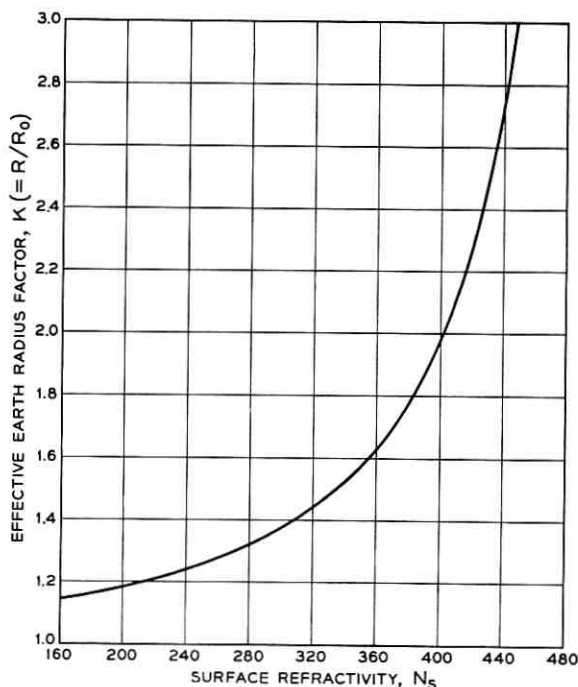

Fig. 3 — Effective earth-radius factor vs refractivity at surface.

or there is a large difference between the altitudes of the two stations, the values of K determined for the two stations are likely to be almost the same.

2.3 *Effect of a Difference in Antenna Site Altitude*

The angle θ for a symmetrical path with no obstructions can be determined by consideration of refractivity alone, as outlined above. When the altitudes of the antenna sites above sea level are significantly different, or when the antenna centerlines are directed above or below the horizontal, corrections are necessary to determine the angles of the beam centerlines above the chord. The correction factors may be (and generally will be) different at the two ends of the path. The effect of a difference in altitude of antenna sites will be considered first.

Fig. 4(a) illustrates the geometry of an unsymmetrical path. The asymmetry here is due to the high altitude above sea level of the antenna at the left. For paths of practical interest, the difference in altitude, Δ, between antenna sites is not likely to exceed a few thousand feet. (The

(a)

$$\theta_1 = \theta + \beta + \theta_{t_1}$$
$$(\theta_{t_1} \text{ IS NEGATIVE})$$

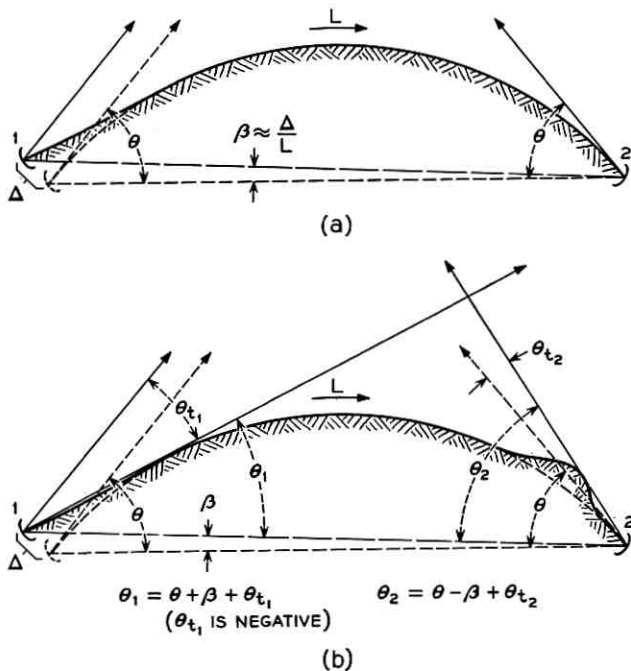$$\theta_2 = \theta - \beta + \theta_{t_2}$$

(b)

Fig. 4 — (a) Antenna 1 at altitude Δ above Antenna 2; (b) negative take-off angle at raised site 1, positive take-off angle at site 2.

greatest difference in existing systems known to the authors is about 8000 feet, on a path in the Arctic.) Thus, since Δ is very small relative to the path length, L, the correction angle β to account for the elevation difference is given with sufficient accuracy by

$$\beta \approx \frac{\Delta}{L} \qquad (4)$$

2.4 Effect of Nonhorizontal Antenna Take-Off Angle

When an antenna site is elevated and there is no foreground obstruction the beam centerline axis is generally depressed from the horizontal and aimed at the optical horizon to increase the received signal. Conversely, when there is an obstruction between an antenna and the normal horizon it may be necessary to elevate the antenna axis above the horizontal.

In the case of Fig. 4(a), for example, the beam from the elevated antenna at the left would normally be depressed by an angle θ_{t1} in aiming at the horizon. This is illustrated in Fig. 4(b). Also in Fig. 4(b), an obstruction has been added in front of the antenna at the right, requiring that its beam axis be elevated by angle θ_{t2}. θ_t at either end will be considered positive when measured upward from the horizontal.

By combining the effects of antenna altitude difference and nonhorizontal take-off angles, the following general expressions are obtained for the angles θ_1 and θ_2 between the chord joining the antenna sites and the centerlines of the respective antenna beams:

$$\theta_1 = \theta + \beta + \theta_{t1} \qquad (5)$$

$$\theta_2 = \theta - \beta + \theta_{t2} \qquad (6)$$

where θ is defined by (2), β is defined by (4), and θ_1 is at the end with greater altitude. The factor β generally can be ignored if the antenna sites are within about a thousand feet of the same altitude above sea level.

2.5 Concept of "Echoes"

The conventional diagram for a tropospheric scatter path, as illustrated in Fig. 1, shows well-defined antenna beams projecting from each terminal. This is a convenient model for many purposes, particularly for determining received signal power. It can be misleading when one is concerned about interference effects, however. We will therefore discard the notion of a bounded antenna beam, and turn instead to the geometry illustrated in Fig. 5.

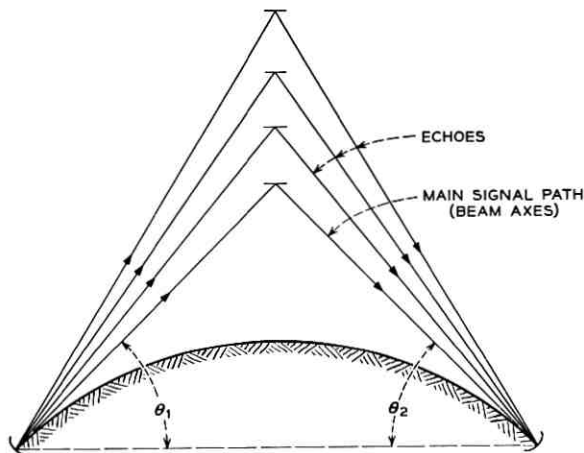In this illustration, the earliest-received — and generally the strongest

Fig. 5 — Illustration of tropospheric transmission in terms of "main signal" and "echo" paths.

— signal component travels along the axes of the antenna beams and is reflected at their intersection in the troposphere. This will be called the "main signal." Other components of signal are assumed to be smaller in amplitude because they are somewhat off the maximum-gain centerline of the antenna beam, and also because the reflected power decreases with increasing altitude of the reflection point. These reflections will be called "echoes" because of their delay behind the main signal.

Although there must be an infinitely large number of echoes from a continuously illuminated common volume, some sections of the volume probably produce higher power density than others as seen by the receiving antenna. It is helpful in this argument to consider the total received power as some equivalent result of a finite number of signal paths.

In the remaining analytical sections of this paper, it is assumed that reflections occur independently, at random, and only in the plane of the antenna beam centerlines. Among the entire ensemble of such echoes, one is to be designated as the "worst" echo — the echo calculated to produce most distortion in the output of the system. Other echoes — including those on either side of the beam centerlines — will contribute to the total distortion, but their effects are included later as part of the empirical correction for multiple echoes.

2.6 *Summary of Model*

The results up to this point can be summarized as follows: A path structure has been proposed which replaces the conventional 3-db an-

tenna beamwidth by a series of signal paths. The shortest of these paths is assumed to follow the antenna beam centerlines, and is designated the main signal path; the rest are echo paths. Critical parameters in the model are the angles θ_1 and θ_2 between the antenna beam centerlines and the chord joining the two antenna sites. The received power, given by (1), depends critically on these angles. When the path geometry is symmetrical, (which would be the case with both antennas at sea level and no foreground obstructions), the angles depend only on path length and atmospheric refractivity (2). When there is a significant difference in the altitudes of the antenna sites or non-horizontal antenna take-off angles, θ_1 and θ_2 must be modified accordingly (5, 6).

The foregoing model will be used in deriving expressions for echo amplitude, echo delay, and finally, intermodulation.

III. ECHO AMPLITUDE

The magnitude of an echo relative to the main signal over a tropospheric scatter path is derived in Appendix B. The derivation makes use of (1) to express the power received in vanishingly small beamwidths α and α_e, representing main signal and echo paths at angles θ and θ_e, respectively (See Fig. 6). The "reflection loss," r_e, of the received echo power relative to the received main signal power is given by (28); thus:

$$r_e \text{ (in db)} = 10 \log_{10} \left(\frac{\theta_e}{\theta} \right)^7 = 10 \log_{10} \rho^7. \tag{7}$$

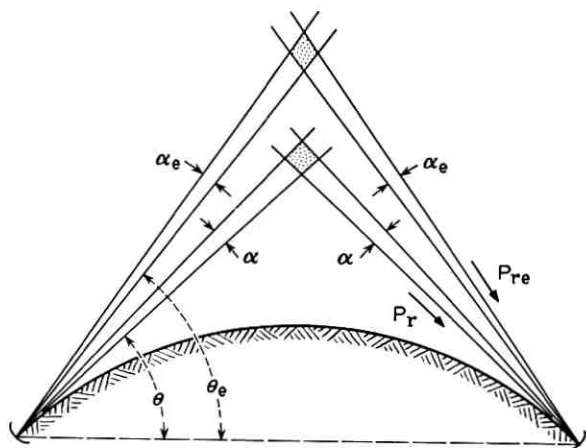It can be shown that (7) is valid for the unsymmetrical path of Fig. 7

Fig. 6 — Geometry used to calculate ratio of main signal power to echo power in a symmetrical path (7).
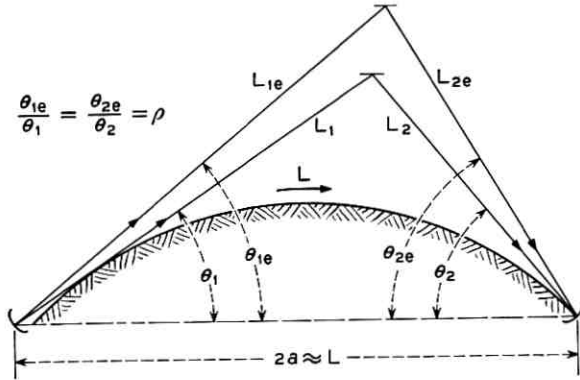
Fig. 7 — General unsymmetrical path.

as well, where the main signal and echo angles are related by:

$$\frac{\theta_{2e}}{\theta_2} = \frac{\theta_{1e}}{\theta_1} = \rho \tag{8}$$

Equation (7) also holds for the "small layer" version of the reflection theory. For the "large layer" theory, which may be more applicable under high-signal conditions, the reflection loss r_e can be shown to vary as $(\rho)^9$.

The derivation of (7) assumed that the antenna gains in the directions of the echo path were equal to the gains along the main signal path. In many cases, particularly those involving half-power beamwidths less than about one degree, the difference in antenna gain between the angles of the echo and main signal paths will add a significant factor to the difference between the received echo and main signal powers. Thus, if the combined loss due to the decrease in the radiation patterns of the two antennas at the angles θ_{1e} and θ_{2e} is given by r_a, in db, the resultant ratio of main signal to echo power is:

$$r \text{ (in db)} = r_e + r_a. \tag{9}$$

The reflection loss, r_e, is related to ρ by (7). The antenna loss, r_a, can be determined with sufficient accuracy from antenna patterns such as those illustrated in Fig. 8. The beam axes correspond to angles θ_1 and θ_2. Thus, the antenna losses at the echo angles θ_{1e} and θ_{2e} can be read from the appropriate antenna patterns at angles from the axes given by

$$\theta_{1e} - \theta_1 = (\rho - 1)\theta_1 \tag{10}$$
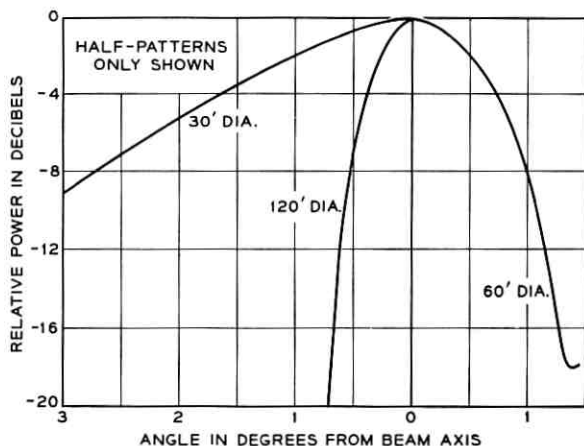
at one end of the path and

Fig. 8 — Typical free-space patterns of parabolic antennas (750 mc).

$$\theta_{2e} - \theta_2 = (\rho - 1)\theta_2 \qquad (11)$$

at the other.

IV. ECHO TIME DELAY

The time delay of an echo traveling the path θ_{1e}, θ_{2e} relative to the main signal traveling the path θ_1, θ_2 (Fig. 7) is:

$$T = 2.68 \, L \, \theta_1\theta_2(\rho^2 - 1) \text{ microseconds.} \qquad (12)$$

Here the path length, L, is in statute miles and ρ is as defined by (8). The derivation of (12) is given in Appendix C.

V. INTERMODULATION, SINGLE ECHO

A method for calculating intermodulation distortion caused by a single echo in the transmission medium of an angle-modulated system is given in Ref. 4. The method assumes a baseband loaded with random noise simulating many frequency-division multiplexed voice channels. The signal-to-intermodulation ratio, S/I, in a narrow frequency slot any-where in the baseband output can be computed for given values of signal-to-echo power ratio (r), echo delay (T), base bandwidth $(f_b)^*$ and rms frequency deviation (σ). The technique is applicable for signal-to-echo power ratios greater than about 5 or 6 db.

---

* The theory presented in Ref. 4 assumes that the baseband extends from zero to a top frequency f_b. In practical systems, the lower edge of the band starts at some nonzero frequency, say 12 kc. For systems of many channels, the assumption of a band extending to zero introduces negligible error.

Figure 5.7 in Ref. 4 shows a set of curves from which one can determine the S/I ratio in the top channel of an FM system without pre-emphasis. Fig. 9 is a similar set of curves computed for a pre-emphasized FM system. The preemphasis characteristic, Fig. 10, is typical of those used on many tropospheric scatter systems. The curves of Fig. 9 are probably also sufficiently accurate for phase modulation.

Fig. 9 is used as follows:

Given values of base bandwidth, frequency deviation, and an assumed value of time delay, compute the rms deviation ratio, σ/f_b, and the product of top baseband frequency and echo delay, f_bT. Note the value of the contour line on Fig. 9 at the intersection of the lines for the computed values of σ/f_b and f_bT. The value of the contour at the intersection
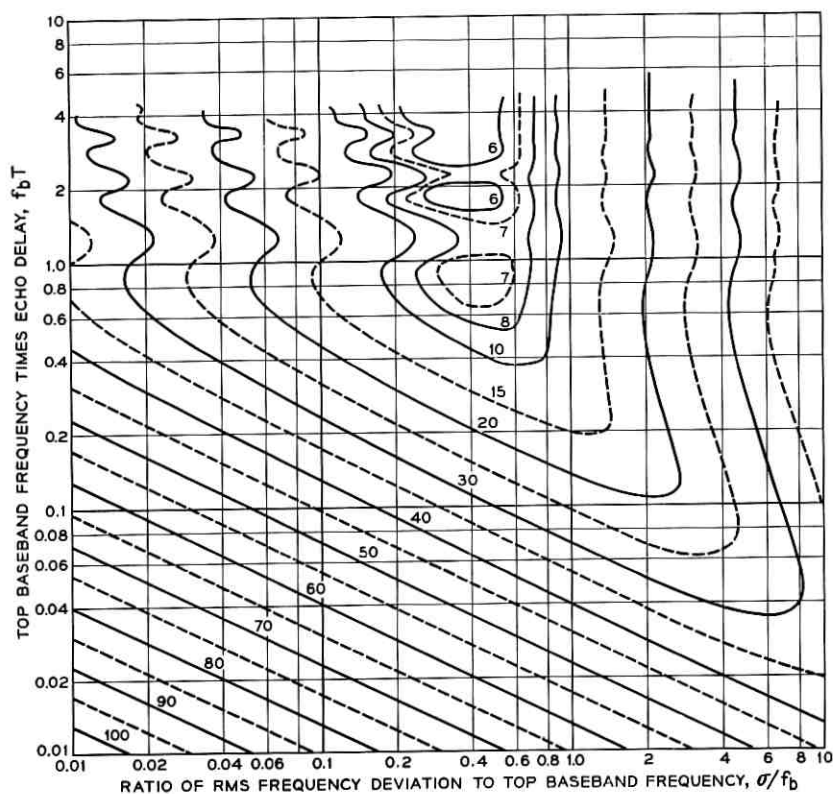

Fig. 9 — Contours of constant interference. (Value in db by which S/I ratio exceeds signal-to-echo power ratio in multichannel pre-emphasized FM system. Value is for top channel only.)
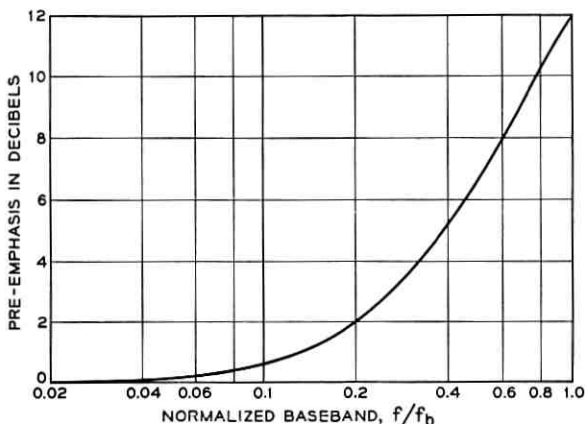
Fig. 10 — Pre-emphasis typical of tropospheric scatter systems.

represents the amount by which the S/I ratio *exceeds* the received signal-to-echo power ratio, r, in db. Thus, the values of the contour and the signal-to-echo power ratio must be added to arrive at the S/I caused by the echo.

The manner in which S/I varies as a function of echo time delay can be determined from Fig. 9, for a system of given parameters and a single echo of known magnitude. A family of curves of S/I vs. time delay can be developed, where each curve in the family applies to an echo of fixed amplitude. Fig. 11 shows such a family. The particular case chosen for illustration is that of a one-megacycle baseband, pre-emphasized and frequency modulated with an rms deviation ratio of one. Each line of this family of curves corresponds to a cross-section of the contours of Fig. 9 on the vertical line $\sigma/f_b = 1.0$. To the cross-section from Fig. 9 has been added, in the case of each curve on Fig. 11, the indicated signal-to-echo power ratio r.

Note that the family of curves in Fig. 11 applies generally to any system having the selected base bandwidth and deviation. It has not been tied to any particular path. The problem now is to find the combination of echo amplitude and delay time for a selected path which produces the most distortion in the system. To illustrate the method of doing this, the following path parameters have been arbitrarily chosen:

Path length, $L = 200$ miles

Refractivity factor, $K = \frac{4}{3}$

Antennas, both ends = 60-foot parabolas, with patterns as in Fig. 8

A symmetrical path
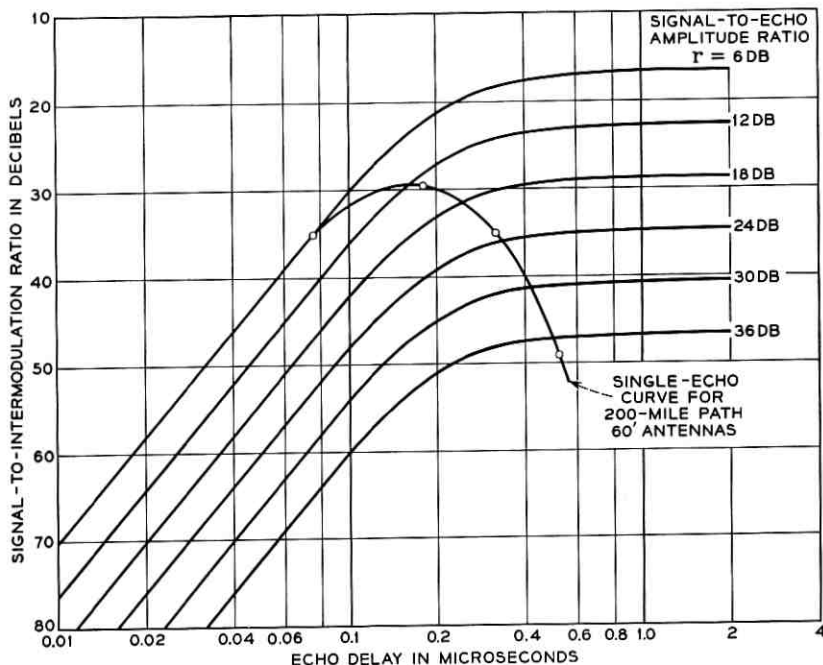
Fig. 11 — S/I vs echo delay (single echo) for pre-emphasized FM; $f_b = 1.0$ mc, $\sigma/f_b = 1.0$.

Equation (2) is used to calculate $\theta_1 = \theta_2 = 0.019$ radians. This is the angle between the chord separating stations 200 miles apart, and the antenna beam axis for a horizontal take-off angle.

Next, choose an arbitrary value of reflection loss, r_e. Suppose 5 db is chosen, corresponding to the power ratio 3.16. This is the ratio of the *reflection* loss along some echo path at angle θ_e, to the reflection loss along the main signal path at angle θ_1 (or θ_2). To this value of reflection loss must be added the reduction in antenna gain at the angle of the echo.

To find the reduction in antenna gain at the angle θ_e corresponding to 5 db reflection loss, first compute ρ from (7). Then (10) or (11) can be used to compute the angular difference $(\theta_e - \theta_1)$ or $(\theta_e - \theta_2)$. The difference in this example turns out to be $0.19°$. At this angle from the beam axis of a 60-foot antenna, Fig. 8 shows that the antenna loss, r_a, for *two* antennas at the angle of the echo path is about one db. The total loss

along the echo path, relative to the main signal path, is then r_e (5 db) plus r_a (1 db), or 6 db.

The time delay of this echo, given by (12), is about 0.075 microseconds.

Turning now to Fig. 11, the S/I ratio corresponding to the echo delay $T = 0.075$ microseconds and the echo amplitude ratio $r = 6$ db is about 35 db.

Similar computations are now performed for other arbitrarily-selected values of reflection loss, say 10, 15, and 20 db. Each case is associated with a separate echo path, and the total echo loss and delay along that path is computed. The results of such calculations are summarized in Table I, together with the S/I ratios that each echo individually would

TABLE I — EXAMPLE OF SINGLE-ECHO S/I CALCULATION

Reflection Loss. r_e (db)	Two-Ant. Loss, r_a (db)	Total Loss. r (db)	Delay T (μsec)	$\dfrac{S/I}{(\text{db})}$
5	1	6	0.075	35
10	3	13	0.18	30
15	8	23	0.32	35
20	18	38	0.52	50

produce if it were the only contributor. These values of S/I are also plotted on Fig. 11, resulting in the overlay curve showing how S/I varies with echo delay and amplitude for the specific path used in the example.

The important point to note in Fig. 11 is that there is a peak in the S/I intersection curve, occurring at time delays near 0.16 microsecond for the 200 mile example path and the assumed baseband and deviation. This tendency toward a "worst echo" effect has been found to be characteristic of all paths studied. The location of the peak on the time delay scale varies depending on the path length, base bandwidth, deviation, and other parameters involved. In particular, it has been observed that, as path length increases, the peak of the intersection curve for wide frequency deviations tends to shift to the region of echo amplitude r less than 5 or 6 db. It is in this region that the intermodulation theory loses accuracy. This has not appeared to be a serious limitation in work to date, however.

VI. LIMITATIONS OF THE ANALYTICAL METHOD

The analytical method as outlined predicts the intermodulation in the output of a receiver due to a single echo, when the received power is assumed given by (1). There are three significant limitations to the method at this point:

(*i*) It predicts a single fixed value of intermodulation, whereas measurements show that there are both long-term changes in the average value as well as more rapid variations about an average during a short interval.

(*ii*) The analysis predicts the effects of single echoes only; practically, there must be a continuum of echoes spread over something vaguely defined as a "common volume," and the total intermodulation generated by these echoes is bound to be greater than that of a single "worst" echo.

(*iii*) The theory of intermodulation due to transmission echoes as presented in Ref. 4 is valid only for echo amplitudes smaller than about 5 or 6 db relative to the main signal. On tropospheric paths, theory suggests that the strongest echoes are normally associated with the shortest time delays. Having short time delays, they would not contribute greatly to interchannel modulation, although they would be responsible for signal fading and the associated increases in fluctuation noise during fades. This is not the entire story, however. While strong echoes may *normally* be associated with short time delays, there are surely some meteorological conditions which produce strong reflecting or scattering regions at abnormally high altitudes. Occasionally, then, an echo might be comparable in magnitude to the main signal *and* be delayed enough to cause severe distortion. It should be emphasized that this apparently is not the average situation, which is of primary concern in this paper.

Efforts to develop satisfactory analytical techniques to remove the limitations noted above have thus far been unsuccessful. Empirical corrections have therefore been developed to account for them, based on the comparison of calculated values and results of tests presented in the following section.

VII. COMPARISON OF CALCULATED AND MEASURED INTERMODULATION

Measurements of intermodulation have been made on four tropospheric scatter paths with diverse climates, path lengths, and siting conditions. Significant parameters of the four paths are given in Table II.

The four systems were Gaussian-noise loaded except for a slot near the top of the baseband. Median values of interference power were measured in this slot at the receiver during approximately $1\frac{1}{2}$ to 3-minute samples. The median of each sample was adjusted to account for fluctuation noise and equipment intermodulation, leaving the value of intermodulation noise chargeable to the path. A more detailed account of test procedures and methods of data analysis is given in Ref. 7.

Table III shows a comparison of the raw measured non-diversity

TABLE II — PARAMETERS OF FOUR TEST PATHS

	Path 1		Path 2		Path 3		Path 4	
	Trans.	Rec.	Trans.	Rec.	Trans.	Rec.	Trans.	Rec.
Location	Caribbean (over water)		Arctic (over ice)		Arctic (over water)		Arctic (over water)	
Length, Statute Miles	185		194		440		340	
Radio Freq., mc (Nominal)	725		900		900		800	
Ant. Dia., ft.	60	60	30	60	120	120	120	120
Site Elev., ft. (Above Sea Level)	200	10	8000	1100	70	1100	700	1100
Eff. Earth Radius Factor, K	1.55		1.26		1.37		1.38	
Ant. Takeoff $\angle$, θ_t, radians	0	≈ 0	-0.0035	-0.004	-0.010	-0.0026	-0.0005	-0.0014
$\angle$'s Between Ant. Axes and Chord, θ_1 and θ_2, Radians	0.0151	0.0151	0.0227	0.0087	0.0375	0.0309	0.0295	0.0308

results from the four test paths and the results calculated for the single "worst" echo. The same data are presented as a scatter plot in Fig. 12. With the exception of the Path 2 data, each point represents an average of the median values from at least 5 to as many as 20 test samples, for one combination of base bandwidth and frequency deviation. Only 3 samples were available for each combination from the Path 2 tests.

VIII. ADJUSTMENT OF MEASURED INTERMODULATION TO WORST-MONTH
      PROPAGATION CONDITIONS

Results from the four tropospheric scatter test paths have shown that there is a decided correlation between average intermodulation and average path loss (or received power). As path loss increases, intermodu-

TABLE III — COMPARISON OF SINGLE-ECHO CALCULATIONS AND RAW MEASURED S/I, DB

Top Baseband Freq., kc	RMS Dev., kc*	Path 1		Path 2		Path 3		Path 4	
		Calc.	Meas.	Calc.	Meas.	Calc.	Meas.	Calc.	Meas.
108	62	—	—	—	—	51	38	—	—
	125	—	—	68	48	45	35	51	47
	250	—	—	62	43	40	31	45	43
	500	—	—	57	37	36	30	41	38
	800	—	—	54	33	36	32	—	—
	1000	—	—	—	—	—	—	40	33
300	62	—	—	—	—	43	32	—	—
	125	—	—	59	43	37	28	42	41
	250	55	48	53	38	32	27	37	37
	500	49	41	48	33	27	27	32	32
	800	—	—	45	30	26	27	—	—
	1000	45	36	—	—	—	—	30	28
	2000	43	34	—	—	—	—	—	—
552	62	—	—	—	—	37	31	—	—
	125	—	—	53	36	31	27	37	35
	250	50	42	48	33	26	24	32	32
	500	44	36	43	29	22	24	27	27
	800	—	—	40	27	21	23	—	—
	1000	40	29	—	—	—	—	25	23
	2000	37	28	—	—	—	—	—	—
1052	125	—	—	49	30	—	—	—	—
	250	44	37	43	27	—	—	33	27
	500	39	32	37	23	—	—	27	25
	800	—	—	34	21	—	—	22	22
	1000	35	26	—	—	—	—	—	—
	2000	32	24	—	—	—	—	20	18

* The deviations used on Path 3 were slightly smaller than listed here. In each case, however, calculated values of S/I are given for the actual deviations used in tests.

lation also increases. While the relationship between path loss and intermodulation may not be highly obvious from the various equations presented to this point, the fact that intermodulation does vary and that a correlation with path loss shows up in measured data leads to the following questions:

(*i*) To what value of path loss does the computed value of intermodulation correspond? In other words, what is a useful point of reference for the observed dependence?

(*ii*) What is the rate of variation of intermodulation with path loss?

The point of reference for the relation between intermodulation and path loss has already been suggested in the earlier discussion of refractivity. It was indicated there that in calculating the effective earth radius
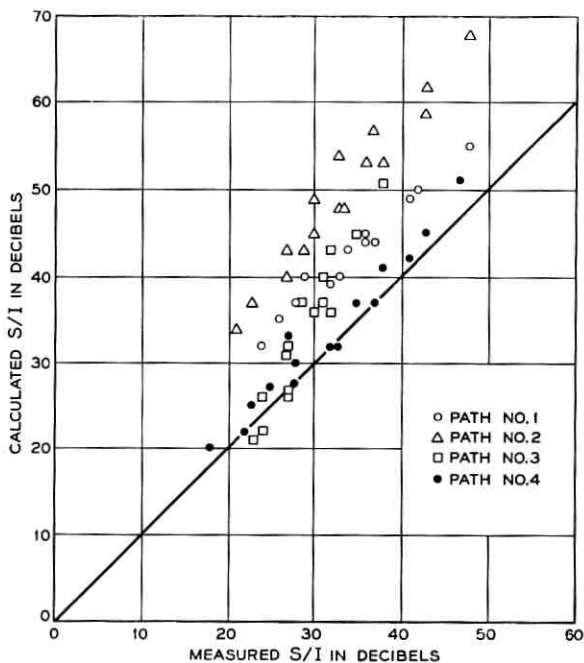
Fig. 12 — Calculated single-echo S/I vs average measured (unadjusted) S/I.

factor, K, the minimum monthly mean value of refractivity indicated in Fig. 2 should be used. Minimum refractivity corresponds to minimum bending of the radio waves, maximum values of antenna centerline angles θ_1 and θ_2, maximum time delay for an echo of given amplitude, and therefore maximum intermodulation. Correspondingly, (1) indicates that minimum refractivity (maximum θ) tends to minimize the average received power. Thus, the observed direction of variation of intermodulation and path loss is borne out by our model. It may be deduced that the value of intermodulation computed on the basis of minimum monthly mean refractivity corresponds directly to what is commonly called "worst-month" path loss — the monthly median path loss during the worst, month-long period of propagation during the year. (There should be some probability attached to this value, but it is usually approximated without this refinement.) Estimates of worst-month loss on a new path are generally derived from data from other paths, with subsequent adjustment based on measurements over the new path.

It can be noted that the choice of worst-month conditions as the point of reference is somewhat arbitrary. We might have chosen instead to use

an average annual value of refractivity, compute an average annual value of intermodulation, and relate the latter to the average annual path loss.

Apart from the point of reference chosen, there remains the second question posed above: what is the rate of variation of intermodulation with path loss, for propagation conditions other than the reference condition? If data were available on the time variations of refractivity for any path of interest, it might be possible to compute an approximate relationship. Unfortunately, such detailed refractivity data are not generally available. As a substitute, we are forced to use empirical adjustments based on results of measurements of intermodulation and received power.

Fig. 13 is typical of the data which have been obtained. Shown are two plots of intermodulation (in terms of S/I ratio) versus received signal measured on Path 3 — a 440-mile over-water path in the Arctic. Each point represents the median values of the two variables during a test sample of about 90 seconds. The total test period covered about ten days, during which time a wide range of propagation conditions was encountered. The rather wide scatter of points is an indication that factors other than *average* path loss can affect the value of intermodulation during any specific test sample. For example, a localized region of high reflecting ability could produce a strong echo and unusually high intermodulation.

The data of Fig. 13 apply for a base bandwidth of 300 kc, rms frequency deviations of 125 and 250 kc, and no receiver diversity. The slope of the curve drawn through the 125 kc deviation scatter plot is slightly greater than one; that is, intermodulation and path loss varied
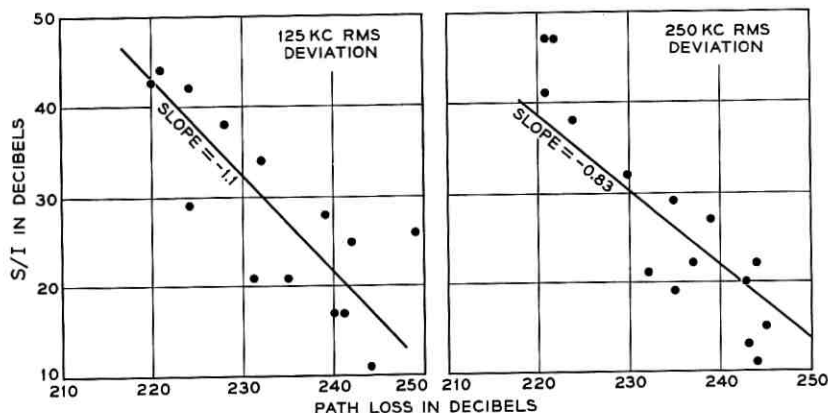

Fig. 13 — S/I vs path loss — 440-mi arctic path, 300-kc baseband, no diversity.

approximately db for db. The slope of the 250 kc plot is slightly more than 0.8. The least squares method was used to calculate these slopes.

Similar scatter diagrams plotted for other combinations of these transmission parameters and the other three paths indicate that the change in intermodulation ranged from about 0.4 db to 1.2 db per db of path loss change. The lowest values were noted on Path 1 — a 185-mile path between Florida and Cuba. There was no apparent pattern to the slopes of the scatter plots as a function of baseband or deviation on any of the four paths tested, except that the slope was generally lowest for the highest deviation ratio. There is a technical justification for the latter effect contained in the contour curves of Fig. 9. Note that a variation in time delay of an echo of given amplitude represents a change in the ordinate $f_b T$ on these curves. It is clear that for a given change of time delay (say, corresponding to a change in $f_b T$ from 0.1 to 0.2), a system operating at a high deviation ratio (toward the right on the σ/f_b scale) will in general exhibit a smaller change in S/I than will a system operating at a low deviation ratio. Except for this trend, however, the effects of different path lengths, antenna sizes, take-off angles and climates confound any attempt to find a precise dependence of S/I on path loss.

The data referred to above, together with similar data from three other paths, indicate that an adjustment to the computed value of S/I of about 0.7 db per db of path loss change is reasonable when one is trying to predict the probable average value of S/I for a path at times other than the worst month. If one is trying to compare a computed worst-month value of S/I with results of measurements on a specific path, it will be more accurate to determine the average slope of scatter plots similar to those of Fig. 13, and use that slope to place the computed and measured results on a comparable path loss basis. Thus, the correction to be added to the median measured S/I during a test period to adjust it to worst-month conditions can be expressed as:

$$D = S(TPL - WPL) \tag{13}$$

where S is the average slope, TPL is median path loss during the test period, and WPL is the worst-month median path loss.

During the test periods for which the results in Table III and Fig. 12 apply, the median path losses on Paths 1, 3 and 4 were four to eight db less than worst-month estimates. During the tests on Path 2, which were conducted over a period of 10 days, the median path loss was about 8 db greater than the worst-month median estimate. Thus, adjustments to the measured data are necessary to convert them to propagation conditions comparable to those assumed in calculations. This was done as follows:

For each path, the average slope was determined of the least-squares lines drawn through scatter plots of S/I versus path loss (similar to those illustrated in Fig. 13). The average slope was then used to adjust the measured S/I results to the worst-month path loss condition, using (13). The estimated worst-month median path losses, test period median path losses, and the slopes used for the S/I adjustments for each of the four test paths are given in Table IV.

TABLE IV

	Path 1	Path 2	Path 3	Path 4
Worst Month Median Path Loss (WPL)	197 db	213 db	243 db	240 db
Test Period Median Path Loss (TPL)	193 db	221 db	235 db	236 db
Avg. Slope of S/I vs Path Loss	0.5	0.7	0.9	0.6
Total Adjustment to Measured S/I Data	−2 db	+6 db	−7 db	−2 db

When the adjustments indicated in the last line of Table IV are made to the measured data plotted in Fig. 12, the result is as given in Fig. 14. With the exception of the Path 3 data, the adjusted points exhibit a noticeably narrower spread than before the path loss adjustments were made. However, the group of points as a whole still differs significantly from the calculated results for single echoes.

IX. ADJUSTMENT OF CALCULATED INTERMODULATION TO ACCOUNT FOR MULTIPLE ECHOES

It has been emphasized that a calculated value of intermodulation represents only the contribution of the single worst echo among echoes which have at least 5 or 6 db lower amplitude than the first-arriving signal element. In fact, however, there are probably very many echoes present at any given time on a tropospheric scatter path. The combined effects of a multiplicity of echoes is to produce more intermodulation than the maximum attributed to any single echo. Thus, it seems qualitatively reasonable to attribute the remaining differences between calculated and measured results in Fig. 14 to the effects of multiple echoes.

The results for Paths 1 and 2 in Fig. 14 appear to be closely intermingled, and they differ from the calculated values for single echoes by an average of about 10 db. The Path 4 results form a group somewhat closer to calculated values; their average difference is only about 4 db. As already noted, the differences for Path 3 vary more widely. This is thought
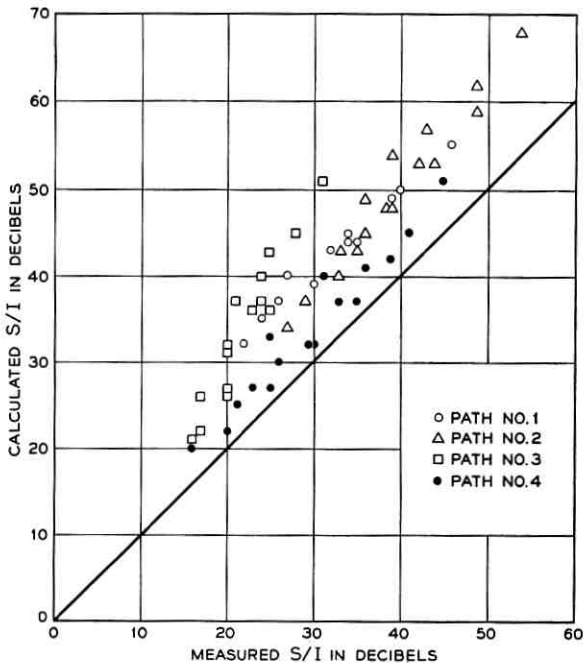
Fig. 14 — Calculated single-echo S/I vs average measured S/I normalized to worst-month path loss conditions.

to be at least partially the result of a propagation anomaly observed on this path, the effects of which will be discussed shortly. There are also a number of other recognized sources of possible errors in both the calculated and measured data, and these will also be discussed. Since there is no way at this time to resolve the quantitative effects of these errors, we are left with the undesirable but rather inescapable alternative of trying to estimate an adjustment for multiple echoes from the scatter of points as they stand in Fig. 14.

Fig. 15 shows the result of subtracting nine db from each point in Fig. 14. This value gives a little extra weight to the data from Paths 1, 2, and 3 for the following reasons: (a) there is relatively good agreement among the adjusted data for these three paths, and (b) up to the time of this writing, there has been considerable uncertainty as to the path loss data which should be assumed for Path 4.

While 9 db appears to be a reasonable average adjustment, it should not be expected that the intermodulation measured during any single test sample would be 9 db more than the calculated figure. The number
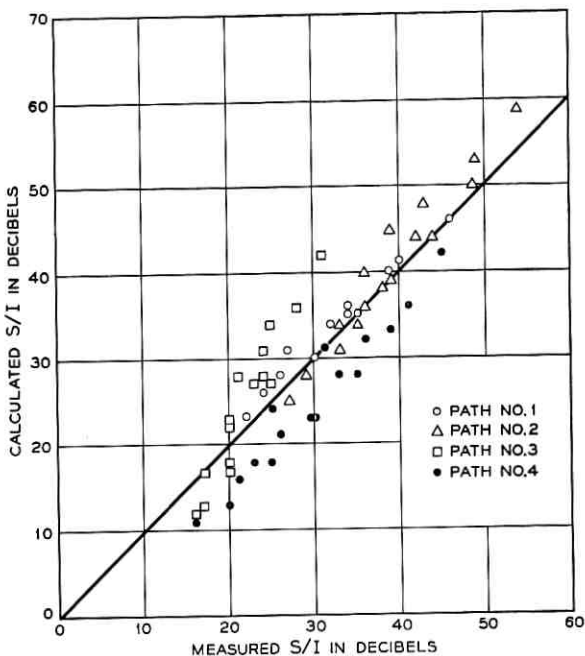
Fig. 15 — Calculated S/I adjusted for multiple echoes vs average measured S/I normalized to worst-month propagation.

of significant echoes is bound to vary from test sample to test sample as atmospheric conditions change. We are aware, too, of various sources of error in calculated or measured results. These errors are discussed in a later section, and it will be apparent that some of them can have an effect on the values assumed for path loss and multiple echo adjustments.

With the caution that some change in the number can be expected as better understanding and more and better data become available, it is felt that nine db is a reasonable number to account for the effects of multiple echoes in calculations of path intermodulation.

X. ANOMALY ON PATH 3

Fig. 15 shows that there is a wider variation of differences between measured and calculated results for Path 3 than for the other paths. This might be explained by the fact that the extreme length of the path makes it possible for strong reflections from the tropopause to cause more interference, at times, than echoes in the troposphere alone. The altitude of the tropopause over Path 3 was reported at 25,000 to 35,000 feet from

day to day during the test period. When the tropopause is in the lower half of this range, according to the method of calculation described earlier in this paper, reflections could cause distortion that is significant with respect to the maximum from a single echo in the troposphere. The effect is greatest for low FM deviations, as indicated by the calculated results in Table V.

Fig. 16 shows the same comparison as Fig. 15 except that on Path 3, a reflection is assumed to occur at 27,000 feet with no more reflection loss than the main signal. Thus, the echo at the receiver would be reduced below the main signal only by the antenna losses due to the elevated angles of the path to the tropopause. In plotting Fig. 16, the interference from this echo was added to the single echo maximum before making the 9 db correction for multiple echoes.

To make the calculated S/I for Path 3 correspond even better with the measured values, it would be necessary to weigh the effect of reflections from the tropopause more heavily than the effects of the tropospheric reflections. One way to increase the effect of the tropopause reflections would be to assume its altitude to be 25,000 or 26,000 feet. Another possibility is to assume that the tropopause provides a better reflecting surface than do the layers in the troposphere, so that, except for the antenna loss, the echo from the tropopause would be stronger than the main signal.

Still another possibility is contained in the following argument. In

TABLE V — EFFECT OF TROPOPAUSE ON S/I RATIOS, PATH 3

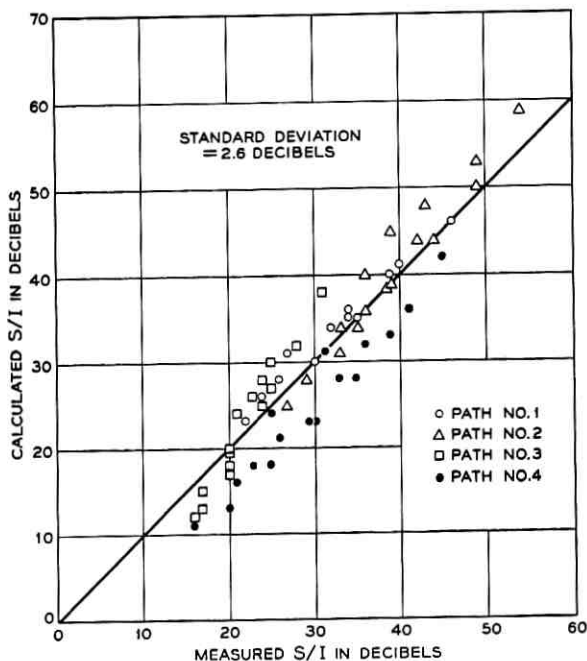| Top Baseband Frequency, kc	RMS Dev., kc	Calculated S/I, db				Measured S/I	
		Tropo-sphere Single Echo	Tropo-pause Echo, 27,000'	Sum of First Two Echoes	9-db Corr., Multiple Echoes	Raw	Adj. to Worst Month
108	62	51	49	47	38	38	31
	125	45	44	41	32	35	28
	250	40	41	37	28	31	24
	500	36	44	35	26	30	23
	700	36	50	36	27	32	25
300	62	43	42	39	30	32	25
	125	37	36	33	24	28	21
	250	32	33	29	20	27	20
	500	27	36	27	18	27	20
	800	26	41	26	17	27	20
552	62	37	38	34	25	31	24
	125	31	33	29	20	27	20
	250	26	29	24	15	24	17
	500	22	31	22	13	24	17
	800	21	36	21	12	23	16

Fig. 16 — Calculated vs measured S/I when tropopause reflection is added to calculated results for path 3.

applying the model of the tropospheric scatter path described in this article, the power in the lower half of the antenna beams has been neglected as though it were blocked by the horizon. Suppose that power in the lower half of the antenna beam grazing the surface of the sea were reflected (and spread somewhat by the convexity of the surface) with negligible loss. This power, with almost no antenna loss, may be reflected again from the tropopause and cause high distortion because of its long delay and low loss. This argument is particularly suited to Path 3 where the antennas at one end are located on a cliff 1100 feet above the sea.

There seems little point in speculating further about the reasons for the relatively peculiar behavior of the results for this path. While each of the arguments above has some plausibility, the fact remains that information is not now available to prove any of them.

XI. SOURCES OF ERRORS IN CALCULATED AND MEASURED DATA

The foregoing comparison of calculated and measured results indicates that the analytical-empirical method of predicting intermodulation has a standard error of estimate of about 2.6 db relative to measured results.

Part of this error is due to the limited accuracy of the empirical adjustments for path loss and multiple echoes. There are in addition a number of other possible sources of error in both the calculated and measured data, and these will now be discussed.

11.1 Path Loss Data

Estimates of worst-month path loss are difficult to make, and errors of 2 or 3 db are not unlikely. In fact, different methods of estimating path loss show that much discrepancy.

Determination of the median path loss from tests is also subject to errors. Knowledge of the values of transmitter output power, received signal power, RF plumbing losses, and antenna gains is required, and inevitably some error creeps into measurements of these parameters. Any error in such measurements should be approximately constant (assuming negligible changes between test periods) for a given path. Hence, it should not affect the slope of an S/I versus path loss scatter plot for that path. However, a constant error would affect the value of the adjustment used for multiple echoes.

11.2 Antenna Orientation and Pattern

Ideal free-space antenna patterns and centerlines directed at the horizon are assumed in calculating intermodulation by the method outlined in this paper. In practice, local siting conditions may have some effect on an antenna pattern and modify the assumed amplitude relationship between echoes at different angles from the horizon. The errors due to these factors should be a constant for any one path, but when the results for several paths are brought together as in Figs. 13 to 16, they will contribute to the spread of data.

The antenna patterns assumed in all the calculated results in this paper were those shown in Fig. 8. These patterns — determined from scale-model measurements — apply at a frequency of about 750 mc. As Table II indicates, the frequencies on the test paths varied from about 725 mc to 900 mc. In view of the uncertainty noted above relative to effects of local siting conditions, it has not appeared worthwhile to adjust the nominal patterns for 750 mc to patterns for specific test frequencies; however, it is worth pointing out this factor as a possible contributor to the spread of data.

11.3 Extrapolation of S/I vs Time Delay Curves

In the example used to illustrate the calculation of single-echo intermodulation, the peak of the S/I versus time delay curve (Fig. 11) was

quite pronounced. This example assumed a path length of 200 miles and an rms deviation ratio of unity. For longer paths and higher deviation ratios, it is sometimes found that the highest value of S/I which can be calculated falls on the 6-db echo amplitude curve. This indicates that the peak of the curve may be further to the left of the 6-db contour, or in the region where the intermodulation theory of Ref. 4 is not valid. Although we have taken the liberty of extrapolating into this region to obtain certain of the values used in Figs. 13 to 16, it is clear that this is another source of possible error.

It is worth pointing out that in those cases where such extrapolation has been necessary to arrive at an estimate of intermodulation, the resulting value of S/I is poor enough to render the system rather useless. In a sense, then, the fact that there may be an error due to the extrapolation is academic.

XII. EFFECT OF DIVERSITY ON INTERMODULATION

The calculated results have been compared here to the S/I ratios measured without diversity. No approach has been found to date which permits analytical prediction of the diversity improvement; however, experimental results give some indication of the advantage to be expected.

Measurements of S/I were made with dual and quadruple diversity reception as well as without diversity on Paths 2, 3, and 4, and with dual diversity and nondiversity on Path 1. Generally, the difference in S/I between nondiversity and dual, and between dual diversity and quadruple, was 2 to 5 db. On the three Arctic paths, though, it is questionable whether the full advantage of diversity was being realized because the combiners were not working as well as expected, and there was some unexpected interference from testing equipment.

Dual diversity results for Path 1 showed an improvement in intermodulation over nondiversity during high deviations, but the nondiversity results were about equal to dual diversity results for low deviations.

To generalize, on the basis of measured data, one might expect to realize 3 or 4 db improvement in median S/I ratio with dual diversity over nondiversity reception, and another 3 or 4 db improvement for quadruple diversity over dual diversity.

XIII. SUMMARY OF ANALYTICAL-EMPIRICAL METHOD

The relatively good agreement between calculated and measured results indicates that the approach described in this paper is some progress

in the right direction. Further measurements and more accurate methods of measurement may show the way to refinements in the analysis which will improve its accuracy and minimize the present reliance on empirical adjustments.

There is given below a step-by-step summary of the method suggested for predicting the average intermodulation due to the propagation medium in a tropospheric scatter system. While the sequence is rather long and involved, many of the steps can be formulated in graphical or chart form for general use in a variety of cases. In fact, a computer program has been developed to accomplish steps 2 through 11, using as inputs the parameters listed in step 1.

The steps involved in the method are as follows:

1. Determine these parameters of the system and the path:
   a) Path length, statute miles
   b) Antenna beam pattern for the appropriate antenna size and operating frequency
   c) Antenna site elevation, feet, at each end
   d) Antenna take-off angles, θ_{t1} and θ_{t2}
   e) Minimum monthly median values of radio refractivity at sea level, N_o, in the areas of the two antenna terminals. (From Fig. 2.)

2. Calculate the surface values of refractivity, N_s, using (3) and the values of N_o found in step 1(e).

3. From Fig. 3, determine the values of the earth radius factor K corresponding to the values of N_s found in step 2. Estimate an average value of K for the path.

4. Compute the base angle, θ, defined by Fig. 1 and (2).

5. Compute the correction angle β, (4), to account for the difference in antenna site altitudes.

6. Compute the total angles θ_1 and θ_2 between the chord joining the antenna sites and the centerlines of the respective antenna beams, using (5) and (6).

7. Choose arbitrarily four or five values of reflection loss, r_e, spaced through the range of approximately 5 to 30 db.

8. Compute the value of ρ corresponding to each value of reflection loss selected in step 7, using (7).

9. Determine the additional loss of the echo path due to the antenna patterns, for each echo selected in step 7. Use (10) and (11) to determine the echo angles, and the antenna patterns of Fig. 8 to determine the reduction of gain at those angles.

10. Calculate the total loss for each echo, using (9).

11. Compute the time delay, T, for each echo, using (12).

12. For given base bandwidth and frequency deviation, use Fig. 9 to determine the value of S/I for each echo. (For a straight-FM system, use Figure 5.7 in Ref. 4.)

13. Plot S/I versus time delay for each echo, and determine the minimum value of S/I. Alternatively, the minimum can usually be estimated with sufficient accuracy from inspection of the calculated S/I values for the several echoes.

14. To account for multiple echoes, subtract 9 db from the minimum value of S/I found in Step 13.

The result of step 14 is the predicted value of the monthly mean S/I ratio during worst-month propagation conditions. The result applies to the top channel of a pre-emphasized FM system operating without diversity. For dual or quadruple diversity systems, add 3 db or 6 db, respectively. For other than worst-month propagation, add 0.7 db to the S/I ratio for each one-db decrease in path loss. (The correction factor may be slightly higher or lower than 0.7 db for a specific path, and should of course be modified appropriately if the S/I versus path loss function has been measured.)

The geometry of some paths may present unusual propagation characteristics calling for special treatment beyond the steps outlined above. The case of the 440-mile path discussed earlier is an example; on this path, it appeared that reflections from the tropopause may have contributed significantly to the intermodulation in the system. The implication of this case is that propagation anomalies can be expected on paths several hundred miles long, where the common volume extends to very high altitudes.

Effect of Refractivity on Effective Earth Radius

We shall outline here a method for determining the approximate value of the effective earth's radius factor, K, applicable to a selected tropospheric scatter path.

K is expressed with sufficient accuracy for our purpose by[5]

$$K = \frac{R}{R_o} = \frac{1}{1 + \frac{dn}{dh} R_v} \tag{14}$$

where dn/dh is the vertical gradient of refractive index, n. In turn, define refractivity:

$$N = (n - 1)10^6. \tag{15}$$

Within a kilometer of the earth's surface,

$$\frac{dn}{dh} \approx \frac{\Delta N}{h} (10^{-6}) \tag{16}$$

where

$$\Delta N = N_h - N_s. \tag{17}$$

N_s is the refractivity at the surface, and N_h is the value at altitude h. When $h = 1$ km, it has been found from extensive data[5] that

$$\Delta N_{1\,\mathrm{km}} = -7.32 \exp (0.005577\, N_s). \tag{18}$$

One useful model of the atmosphere assumes an exponential decrease in N with altitude. Then:

$$N = N_s \exp (-Ch) \tag{19}$$

where h is altitude in kilometers above the ground. To fit the data quoted above, at $h = 1$ km,

$$C = \ln \frac{N_s}{N_s + \Delta N_{1\,\mathrm{km}}}. \tag{20}$$

Considering dn/dh to be constant with altitude (a good approximation only for the first 4000 or 5000 feet, but used here throughout because it introduces no appreciable error in calculating K for even the longest

path considered) and approximating it by the change in refractivity for the first 1000 feet only,

$$\frac{\Delta N}{1000 \text{ ft.}} = \frac{-N_s}{1000} [1 - \exp(-0.305 \ C)]. \tag{21}$$

Substituting (16), (18), (20) and (21) in (14), K is found to be a function of N_s. This is plotted in Fig. 3.

Since the quantity to be calculated later is intermodulation during worst-month conditions, the K of interest is the value for the worst month. Maximum intermodulation generally corresponds to the minimum value of K, since it will produce least bending and hence longest echo delays. Minimum K corresponds to minimum N_s. N_s may be calculated from N_o, the sea level value of N, by

$$N_s = N_o \exp(-0.0322 \ h) \tag{22}$$

where h here is the station altitude in thousands of feet. Minimum values of monthly mean N_o are given in Fig. 2, based on data collected from 306 weather stations over a period of several years.[6]

Calculation of Echo Amplitude Relative to Main Signal Amplitude

The technique involved here to calculate the magnitude of transmission echoes in a tropospheric scatter path makes use of (1) for received power. It will be convenient and helpful in this derivation to refer also to Fig. 6, which shows symmetrical main signal and echo paths having equal beamwidths α and α_e, respectively, and directed at angles θ and θ_e, respectively.

Equation (1) can be used to calculate the relative power received from the two paths, assuming the antenna gains in the directions θ and θ_e to be equal. (The effect of differences in antenna gain at θ and θ_e is discussed in the main text.) Expressing the ratio of the power of the main signal to the power of the echo by r', we have, from equation (1),

$$r' = \frac{P_r}{P_{re}} = \frac{\theta_e^{5}\left(2 + \dfrac{\alpha_e}{\theta_e}\right) f\left(\dfrac{\alpha}{\theta}\right)}{\theta^5\left(2 + \dfrac{\alpha}{\theta}\right) f\left(\dfrac{\alpha_e}{\theta_e}\right)}. \tag{23}$$

Setting $\alpha_e = \alpha$ and letting $\alpha \rightarrow 0$, the limit, r_e, will be the ratio between the power of a signal component traveling the centerline path and an echo traveling some longer path defined by θ_e.

Let $f(\alpha/\theta)$ be written as

$$g(\theta,\alpha) = \left[1 + \frac{\theta^4}{(\theta + \alpha)^4} - \frac{(2\theta + \alpha)^4}{8(\theta + \alpha)^4} \right]. \tag{24}$$

Then

$$\lim_{\alpha \to 0} r' = r_e = \frac{\theta_e^5}{\theta^5} \lim_{\alpha \to 0} \frac{g(\theta,\alpha)}{g(\theta_e,\alpha)} . \tag{25}$$

The above limit is undefined, and so is

$$\lim_{\alpha \to 0} \frac{g'(\theta,\alpha)}{g'(\theta_e, \alpha)} , \tag{26}$$

where $g' = dg/d\alpha$. Using L' Hospital's rule with the second derivative,

$$\lim_{\alpha \to 0} \frac{g''(\theta,\alpha)}{g''(\theta_e, \alpha)} = \left(\frac{\theta_e}{\theta} \right)^2 \tag{27}$$

or,

$$r_e = \left(\frac{\theta_e}{\theta} \right)^7 = (\rho)^7. \tag{28}$$

r_e will be called the "reflection loss" of an echo relative to the main signal.

APPENDIX C

Calculation of Echo Time Delay

The time delay of an echo which travels the path defined by θ_{1e}, θ_{2e} instead of the path θ_1, θ_2 (Fig. 7) is

$$T = \frac{\Delta L}{3(10)^8} \text{ seconds,} \tag{29}$$

where

$$\Delta L = (L_{1e} + L_{2e}) - (L_1 + L_2) \text{ meters.} \tag{30}$$

By the law of sines,

$$\begin{aligned}
\Delta L &= \frac{2a (\sin \theta_{1e} + \sin \theta_{2e})}{\sin (\theta_{1e} + \theta_{2e})} - \frac{2a (\sin \theta_1 + \sin \theta_2)}{\sin (\theta_1 + \theta_2)} \\
&= \frac{2a (\sin \theta_{1e} + \sin \theta_{2e})}{\sin \theta_{1e} \cos \theta_{2e} + \sin \theta_{2e} \cos \theta_{1e}} \\
&\quad - \frac{2a (\sin \theta_1 + \sin \theta_2)}{\sin \theta_1 \cos \theta_2 + \sin \theta_2 \cos \theta_1} .
\end{aligned} \tag{31}$$

For angles θ less than 0.1 radian* the following approximations are valid, applied to the above equation:

$$\cos \theta = 1 - \frac{\theta^2}{2}$$

$$\sin \theta = \theta$$

$$\frac{1}{1 - \frac{\theta_1 \theta_2}{2}} = 1 + \frac{\theta_1 \theta_2}{2}$$

$$2a = L.$$

Then

$$\Delta L = \frac{L}{2} (\theta_{1e} \theta_{2e} - \theta_1 \theta_2) \text{ meters.} \qquad (32)$$

Finally, for L in statute miles,

$$T = 2.68L \, \theta_1 \theta_2 (\rho^2 - 1) \text{ microseconds} \qquad (33)$$

where

$$\rho = \frac{\theta_{1e}}{\theta_1} = \frac{\theta_{2e}}{\theta_2}.$$

REFERENCES

1. Booker, H. G., and Gordon, W. E., A Theory of Radio Scattering in the Troposphere, Proc. I.R.E., **38**, April, 1950, p. 401.
2. Villars, F., and Weisskopf, V. F., Scattering of EM Waves by Turbulent Atmospheric Fluctuations, Phys. Rev., **94**, April, 1954, p. 232.
3. 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.
4. Bennett, W. R., Curtis, H. E., and Rice, S. O., Interchannel Interference in FM and PM Systems Under Noise Loading Conditions, B.S.T.J., **34**, May, 1955, p. 601.
5. Bean, B. R., and Thayer, G. D., Models of the Atmospheric Refractive Index, Proc. I.R.E., **47**, May, 1959, p. 740.
6. Bean, B. R., and Horn, J. D., Radio-Refractive-Index Climate Near the Ground, Journal of Research of the Nat. Bur. of Standards D. Radio Propagation, Vol. **63D**, Nov.–Dec., 1959.
7. Clutts, C. E., Kennedy, R. N., and Trecker, J. M., Results of Bandwidth Tests on the 185-Mile Florida-Cuba Tropospheric Scatter Radio System, I.R.E. Trans. on Comm. Systems, Vol. **CS-9**, Dec., 1961, p. 434.

* Corresponds to path lengths less than about 800 miles. Paths of practical interest are appreciably shorter, and angles θ correspondingly smaller.

A Wideband Transistor IF Amplifier for Space and Terrestrial Repeaters Using Grounded-Base Transformer-Coupled Stages

By W. F. BODTMANN and C. L. RUTHROFF

A wideband transistor IF amplifier is described. A measured model has a gain of 41 db and is flat to ±0.1 db from 50 to 100 mc. Data on noise figure, effect of temperature changes and power output are presented. An electronically variable attenuator is described which is suitable for use in such an amplifier as part of the AGC circuit.

I. INTRODUCTION

The general objectives of the work discussed in this paper are as follows:

(i) To design a wideband IF amplifier suitable for use on heavy-route, long-haul microwave radio relay systems.[1,2]

(ii) To design a wideband IF amplifier with low power consumption and good temperature stability for use in microwave radio relay systems[3,4] which may be battery powered and which must operate at outdoor ambient temperatures.

In accordance with these objectives it was decided to design an amplifier to be substantially flat over a 40-mc band centered at 70 mc. Transistors were chosen as the active elements because of the low bias power required, and the common-base circuit configuration was selected because it appeared likely to result in a stable band shape with respect to temperature.

A part of the initial philosophy was to start with a stable wideband amplifier. The desired frequency selective circuits could then be added as linear lumped circuits whose gain and phase properties are in good control. Such circuits could include any desired gain and phase equalization. This approach results in an IF amplifier with an excellent delay characteristic.

Fig. 1 — (a) Typical interstage circuit, (b) equivalent circuit for the interstage, (c) equivalent circuit for computation of input impedance, (d) equivalent interstage circuit including losses.

II. THE IF CIRCUIT[5,6]

The common-base configuration was chosen for several reasons. The transistors available have alpha cutoff frequencies in the region 500–700 mc. If these transistors were used in the common-emitter configuration in this application, the IF band would extend well beyond the beta cutoff frequency and considerable equalization would be necessary to shape the IF characteristic. Furthermore, changes in ambient temperature result in a shift in both alpha and beta cutoff frequencies. The alpha cutoff is well beyond the IF band but the beta cutoff is not. Hence, for operation over a reasonably large temperature range, the common-base stage has better stability with respect to band shape and midband gain.

Simple broadband transformers[7] are available which can be used in the common-base configuration and result in an interstage with at most one tuning element per stage.

The interstage circuit is shown in Fig. 1(a). L is the stray inductance, C_T the total interstage capacitance, L_b is an inductance added in the base of transistor Q_2, T_1 is a wideband transformer and R_1 has been added to help shape the interstage characteristic. This circuit has been redrawn in Fig. 1(b), where the assumptions have been made that the output of transistor Q_1 is a constant current and T_1 is an ideal transformer with an impedance ratio of 4:1. Transistor Q_2 has been replaced by its input impedance, which is considered to be $R_{IN} + j\omega L_{IN}$ where both R_{IN} and L_{IN} are functions of frequency. The blocking capacitors and bias circuits have been omitted. From Fig. 1(b) we see that the output capacitance of transistor Q_1, the input impedance of transistor Q_2, stray inductance L, and resistance R_1 form a peaking circuit. The equivalent circuit used to compute the input impedance of Q_2 is shown in Fig. 1(c). It has been assumed that $Z_c(1 - \alpha) \gg Z_L$ and $Z_c \gg Z_L + Z_b$. In the Appendix it is shown that the frequency response of Fig. 1(b) is

$$\left|\frac{I_0}{I_2}\right|^2 \approx \frac{(R_T - \mu)^2 + (\mu\gamma + X_L - X_c)^2}{X_c^2} \tag{1}$$

where $R_T = R_1 + r_e + r_b'$
$\mu = (r_b' + \gamma\omega L_b)/(1 + \gamma^2)$
$\gamma = f/f_\alpha$
r_e = emitter resistance
r_b' = high-frequency base resistance.

The high-frequency response of this circuit can be controlled by adjusting R_1 or L_b. Amplifiers have been built using both schemes. The effect of changing R_1 is to change the Q of the circuit. The effect of adding

a suitable inductance, L_b, in the base is a little more complicated but results in a better frequency response.

An inductance in series with the base of a transistor changes the transistor input impedance by decreasing the resistive part and increasing the reactive part. (See Appendix.) Thus, instead of adjusting R_1, a variable inductance in series with the base can be used to tune the interstage. Even though this method of tuning adds inductance to the emitter circuit it does not decrease the bandwidth of that portion of the response which is flat to a few tenths of a decibel. In fact if R_1 and L_b are properly chosen, the amplifier will be compensated over a wide frequency range. This is demonstrated in Figs. 2, 3, 4. Fig. 2(a) shows the measured responses of the peaking circuit with and without added
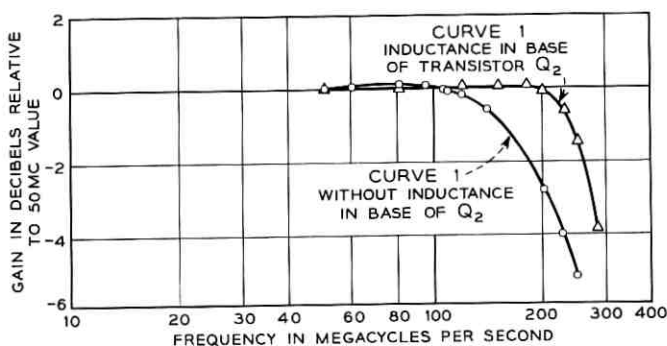
(a)
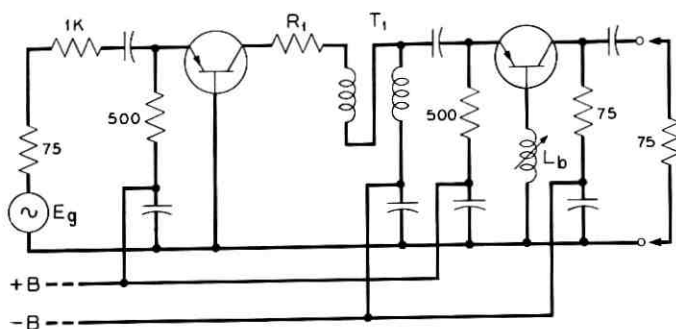
T_1 = BROADBAND TYPE TRANSFORMERS, Z = 4 TO 1
ALL CAPACITORS 0.001 μF
TRANSISTORS ARE W.E. 2N1195

(b)

Fig. 2 — (a) Measured responses of series peaking circuit, (b) circuit used to measure interstage series peaking response.
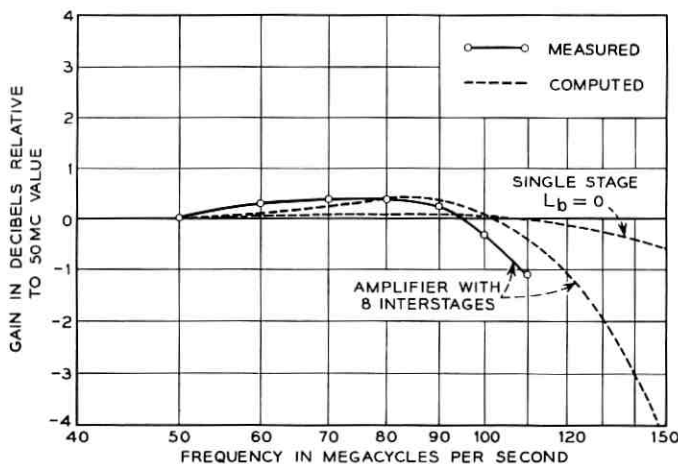
Fig. 3 — Amplitude-frequency response with no added inductance in base of transistor.

inductance in the base circuit. The circuit measured is shown in Fig. 2(b). The computed curves are identical to the measured ones when the following parameters are used in (1):

$$f_\alpha = 600 \text{ mc} \qquad C_T = 4\mu\mu f \qquad V_{CB} = -10v$$

$$L = 0.02\mu h \qquad L_b = 0.05\mu h \qquad I_E = 10 \text{ ma}$$

$$r_b' = 50 \text{ ohms} \qquad r_e = 3 \text{ ohms}$$

$$R_T = 115 \text{ ohms when } L_b = 0.05\mu h$$

$$R_T = 110 \text{ ohms when } L_b = 0.$$

Fig. 3 shows the computed response of a single stage and computed and measured responses for eight identical interstages, with no added inductance in the base lead. The transistor parameters are as listed above. $L_b = 0$ and $R_1 = 54$ ohms. For the computed responses α has been assumed equal to one and the transformers ideal. Since these assumptions are not exactly true, the measured response can be expected to be poorer than the computed response. The midband insertion gain between 75-ohm impedances is 41 db.

Fig. 4 shows the computed responses of a single interstage with two bias conditions and the measured response of eight interstages with inductance added in the base lead. The bias conditions for curve 1 are $V_{CB} = -10v$ and $I_E = 10$ ma; for curve 2, $V_{CB} = -6v$ and $I_E = 6$ ma. The bias conditions for the measured response are the same as for curve
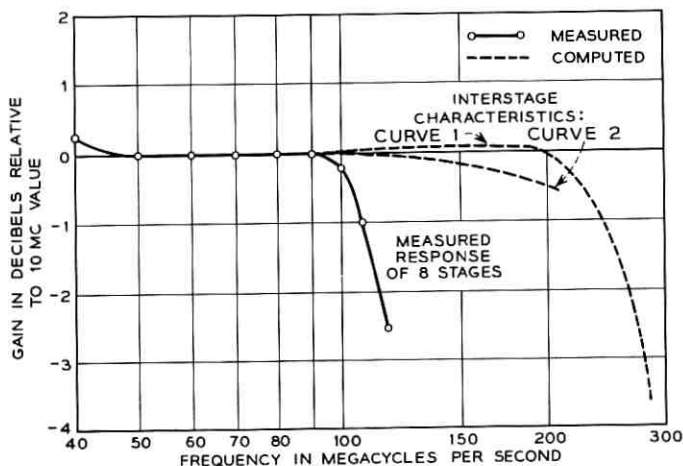
Fig. 4 — Amplitude-frequency response with inductance added in base of transistor.

2 in all except the last two stages which are discussed later. As suggested by Fig. 4, reducing the bias power results in some sacrifice in bandwidth. The final selection of bias conditions is a compromise between a flat response over the desired bandwidth and low power consumption for the amplifier. For the computation of curve 2 the circuit parameters are:

$$f_\alpha = 400 \text{ mc} \qquad C_T = 4.5\mu\mu f \qquad R_1 = 59 \text{ ohms}$$

$$r_b' = 50 \text{ ohms} \qquad L_b = 0.05\mu h$$

$$r_e = 4.3 \text{ ohms} \qquad L = 0.02\mu h$$

The measured midband insertion gain between 75-ohm impedances is 41 db for eight interstages. The dc power consumption is 0.7 watt.

III. INPUT AND OUTPUT CIRCUITS

A complete schematic diagram of the amplifier with a matched input circuit and two output circuits, is shown in Fig. 5. The input circuit shown is similar to the interstages in that a resistor is added in series with the emitter to provide the proper input resistance of 75 ohms, and a capacitor is shunted across the input to peak the circuit at higher frequencies. The return loss of the amplifier input as a function of frequency is shown in Fig. 6.
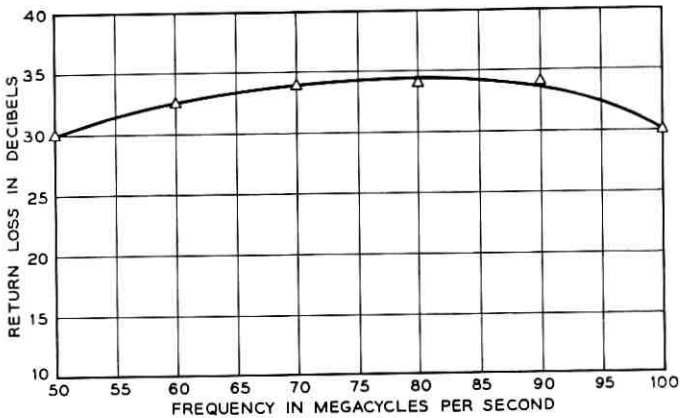
Fig. 5 — IF amplifier schematic.

Fig. 6 — Return loss vs frequency for input circuit of IF amplifier of Fig. 5 measured at 75 ohms.

The input conditions for operation with a low noise figure are discussed in Section VIII.

Two output circuits are shown in Fig. 5. One output circuit consists of a broadband transformer and a shunt peaking circuit. The other has a double-tuned transformer and is shown as the alternate output circuit in Fig. 5. The return losses of the two output circuits, measured at 75 ohms, are given in Fig. 7. The double-tuned transformer adds a measured variation of about one-tenth of a db to the frequency response. All data discussed in this paper were taken with the output circuit having the broadband transformer.

The AGC detectors shown in Fig. 5 have been designed so as to have no effect upon the frequency response. AGC problems are discussed in Section IX.

IV. POWER OUTPUT AND FREQUENCY RESPONSE

The power output of the amplifier depends upon the bias conditions in the output transistors, which in turn depend upon the maximum ambient temperature expected. The last two transistors in Fig. 5 have increased bias currents to allow for increased power output. The bias is arranged so that the rated junction temperature is not exceeded at the high end of the ambient temperature range, 130°F.

The compression characteristic of this amplifier for a single frequency sine wave input at 70 mc is given in Fig. 8. In addition to this information it is desirable to know the frequency response as a function of output
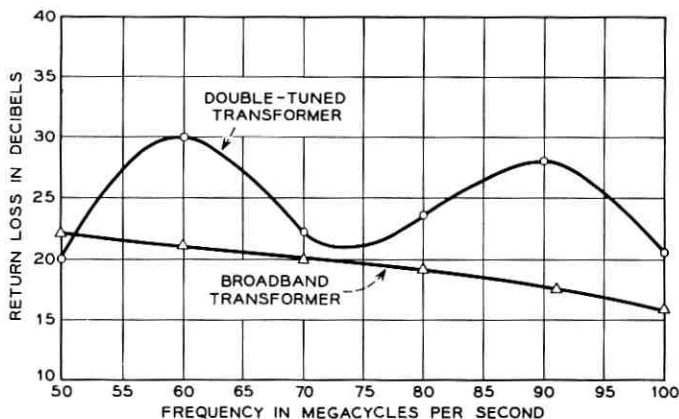
Fig. 7 — Return loss vs frequency for output circuits of IF amplifier of Fig. 5 measured at 75 ohms.

level. This is shown in Fig. 9. Fig. 9(a) shows the response for an output power of +6 dbm and a negligible compression. Fig. 9(b) shows responses for several other levels where it is to be noted that while the details of the responses have been preserved, the level differences between the curves have not.

V. TEMPERATURE-FREQUENCY RESPONSE

The frequency response as a function of temperature has been obtained for the amplifier that does not have added inductance in the base circuit. The data are presented in Fig. 10, plotted relative to the gain at 50 mc. The 50-mc gain decreases at high temperatures and increases at low temperatures. The total change in 50-mc gain in the range −40°F to +130°F is under 2 db.

VI. TRANSFORMERS

The interstage transformers used in this amplifier have been described elsewhere[7] and only a brief description is given here. These transformers are of a transmission line type and consist of a short bifilar winding on a small nickel-zinc ferrite toroidal core. The winding is formed by twisting a pair of wires and then winding the pair on the core. Fig. 11 is a schematic of the transformer and a plot of the measured frequency response into resistive terminations. The turns ratio is 2:1.

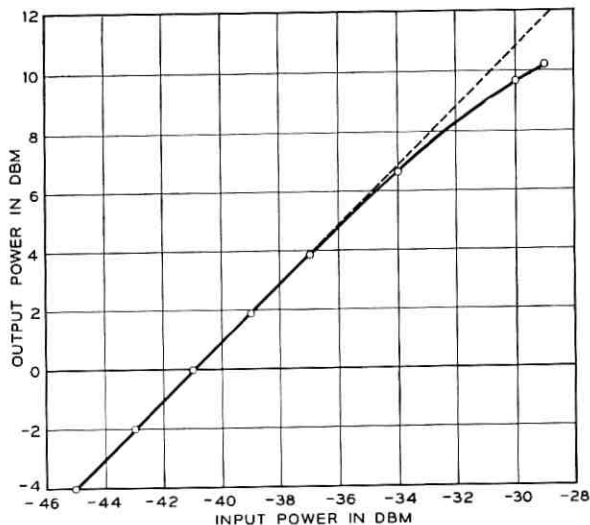This type of transformer may be regarded as an ideal transformer in

Fig. 8 — Compression characteristic for amplifier of Fig. 5.

series with a transmission line whose electrical length equals that of the bifilar winding. The analysis in this paper has assumed only the ideal transformer, which is reasonable as long as the winding is short compared to a wavelength at the highest frequency of interest.

The permeability and Q of this ferrite has been optimized in this frequency range to give minimum loss. With this ferrite and a winding 1.6 inches long, the transformer has the performance shown in Fig. 11.

VII. GAIN PER STAGE

All of the current gain is obtained from the interstage transformer so the maximum current gain per stage using a transformer with an impedance ratio of 4:1 would be 6 db. There are three sources of current loss in each interstage which reduce this gain. As seen in Fig. 1(d), they are:

(i) The transformer inductance L_T and resistance R_T which shunts the emitter-base circuit of Q_2.

(ii) The emitter bias resistor R_2 which shunts the input circuit of Q_2. This loss can be reduced at the expense of greater bias power.

(iii) The current gain of $Q_2 = \alpha < 1$.

The total loss from these sources is 0.8 db so that the actual measured gain per stage is 5.2 db.
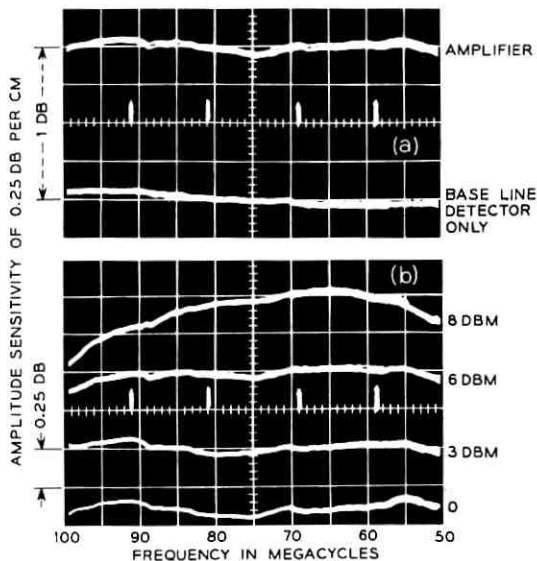
Fig. 9 — (a) Frequency response of amplifier and base line at output power level of +6 dbm and insertion gain of 41 db, (b) frequency response of IF amplifier as a function of the output level in dbm.

VIII. NOISE FIGURES

The amplifier of Fig. 5 has its input impedance adjusted to 75 ohms as indicated in Fig. 6. With this input circuit it is suitable for use as a main IF amplifier. For use as a preamplifier, however, the input circuit can be modified to improve the noise figure. Such a circuit is shown in Fig. 12. Note that the turns ratio of the interstage transformer is 8:3. The measurements reported below were made on the amplifier of Fig. 5 with the input circuit of Fig. 12 and using a diode noise generator. The bias on the first stage was $I_E = 3$ ma, $V_{CB} = -7$v.

Some amplifier noise figure measurements with W.E. 2N1195 transistors in the input circuit are shown in Fig. 13. The curve is the noise figure for the best of 5 transistors, and the amplifier noise figure is plotted versus generator resistance. Also shown are the amplifier noise figures for the 5 transistors with a generator resistance of 70 ohms. These measurements were made in a 0.7-mc band centered at 71 mc. At 50 mc the noise figure is typically 0.5 db lower and at 100 mc, 0.75 db higher.

It should be emphasized that these are amplifier noise figures and not transistor noise figures. In such a broadband amplifier, care must be taken to minimize the noise contribution of the second transistor. This
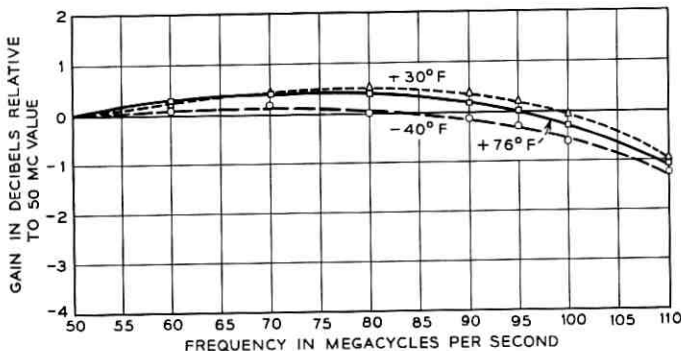
Fig. 10 — Frequency response vs temperature of IF amplifier.

accounts for the higher turns ratio of the interstage transformer. Use of the transformer described in Section VI results in an increase of 0.5 db in noise figure.

The noise contribution of the second transistor depends upon the available gain of the preceding circuit,[8] and the available gain depends in part upon the generator impedance and the transformer turns ratio. It is to be expected, then, that for narrower bandwidths the noise contribution of the second transistor can be further reduced, since higher turns ratios can be used with the corresponding increase in available gain of the first stage.

For the broadband common-base amplifier the best noise figures are obtained with transistors having low optimum source resistances. This comes about because the available gain increases as the optimum source resistance decreases, and the greater the available gain, the lower the noise contribution of the second transistor. It follows that for best results a low-noise transistor having a low optimum source resistance should be used.

The effect of the input circuit of Fig. 12 on the frequency response of the amplifier is quite small. With this input circuit the response was measured using a 75-ohm source impedance and the response equivalent to that of Fig. 9 was obtained by slight readjustment of the base lead inductance in the second stage.

IX. AUTOMATIC GAIN CONTROL CIRCUIT

Wideband transistor IF amplifiers present special problems with respect to AGC in that it is not always feasible to merely adjust the bias to obtain gain control. The method presented here consists of an elec-
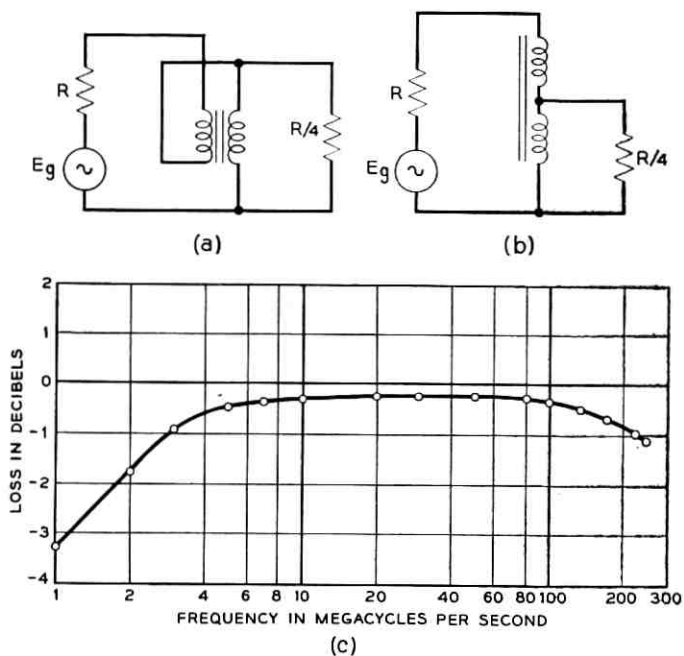
Fig. 11 — Transmission-line type transformer with impedance transformation of 4 to 1. (a) Transmission line form, (b) low-frequency equivalent, (c) amplitude vs frequency of transformer measured between 75 ohms and 18.75 ohms.
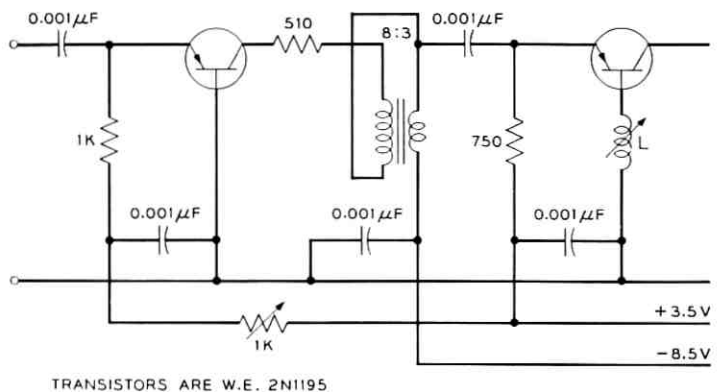

TRANSISTORS ARE W.E. 2N1195

Fig. 12 — Input circuit for low noise figure.

Fig. 13 — Noise figure measurements.

Fig. 14 — AGC attenuator.
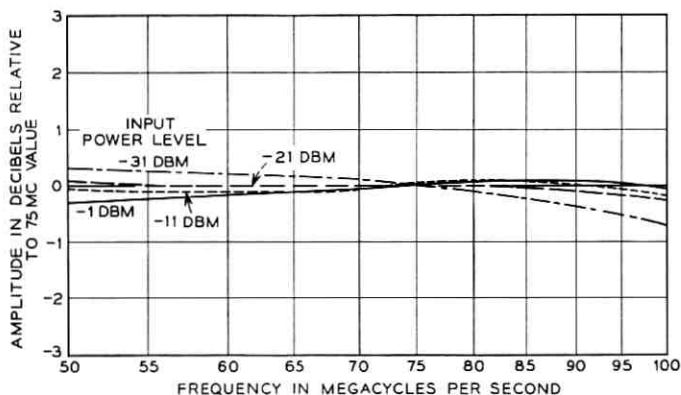
Fig. 15 — Frequency response of variable IF attenuator of Fig. 14. Output level held constant at −37 dbm.

tronically variable attenuator which can be inserted in the amplifier at any suitable point. The attenuator is described schematically in Fig. 14 and is basically a T-pad formed by three diodes. The three diodes are supplied with current at point A through the 750-ohm resistance. Transistor Q_3 controls the current through diode 3 which in turn controls the current through diodes 1 and 2. The resistance of the diodes is a function of the current through them so the attenuation of the T-pad can be controlled by adjusting the current in diode 3. This is done by applying the AGC detector output to the emitter of transistor Q_3.

If Q_3 is cut off, the current at point A splits between diodes 1 and 2, making them low resistances; diode 3 becomes a high resistance. This is the minimum-loss condition. When Q_3 conducts heavily, all of the current

Fig. 16 — Wideband IF amplifier.

at point A goes to diode 3, lowering its resistance and cutting off diodes 1 and 2. This is the maximum-loss condition. The reference diodes (TI-650) provide a 4-volt bias that assures that diodes 1 and 2 can be cut off and also provides collector voltage for Q_3.
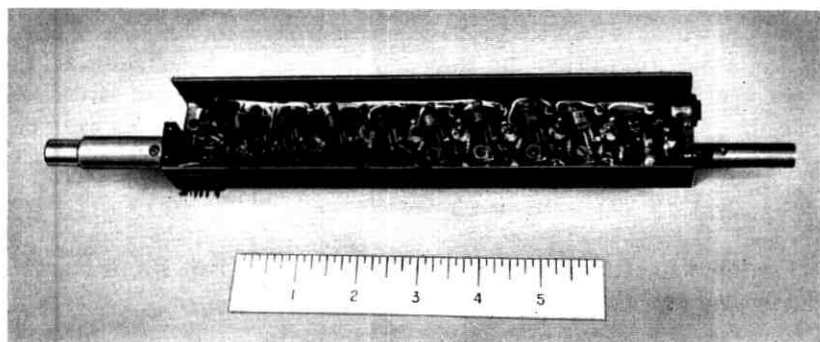
The frequency response of the attenuator is given in Fig. 15. These curves show the variation in response relative to the 75-mc value as the power input to transistor Q_1 (Fig. 14) is varied from -1 dbm to -31 dbm. For inputs greater than $+3$ dbm the response changes shape rather than slope.

The minimum measured insertion loss of the circuit of Fig. 14 is 6 db. The power consumption of the attenuator is 0.35 watt.

X. MECHANICAL ARRANGEMENT

The photograph of Fig. 16 shows the mechanical arrangement of one section of the amplifier which has 41 db of gain. The transistors are mounted inside the chassis to minimize stray inductances and to prevent accidental grounding of the collector, which is connected to the transistor header.

XI. DISCUSSION

It has been demonstrated that high quality wideband IF amplifiers can be made using transistors. Such amplifiers are capable of operation over large temperature ranges and small input power. For instance, an amplifier with a flat bandwidth of 50 mc, 90 db of gain, $+12$ dbm of output power, 34 db AGC range, a noise figure of 6 db, and capable of operation over a temperature range of $-40°F$ to $+130°F$ is feasible. Such an amplifier would require about 2 watts of bias power and would contain 23 transistors and 6 diodes.

It is possible, of course, to trade one attribute for another. For instance, a flat band extending to nearly 200 mc can be achieved at the expense of increased bias power. If a narrower bandwidth is desired, then the gain per stage can be increased. This can be done, for example, by using two transformers per stage or using another type of transformer.

APPENDIX

The input impedance of the equivalent circuit of Fig. 1(c) is approximately

$$Z_{IN} \approx r_e + Z_b(1 - \alpha) \tag{1}$$

where

$$Z_c(1 - \alpha) \gg Z_L$$

$$Z_c \gg Z_L + Z_b \, .$$

Now let

$$\alpha = \frac{\alpha_0}{1 + j\gamma} \quad \text{and} \quad \alpha_0 \approx 1$$

$$\gamma = f/f_\alpha$$

$$Z_b = r_b' + jX_b$$

$$X_b = \omega L_b = \gamma \omega_\alpha L_b = \gamma X_{b\alpha}$$

After these substitutions (1) becomes

$$Z_{IN} \approx \frac{(r_e + r_b')(1 + \gamma^2) - (r_b' + \gamma^2 X_{b\alpha}) + j(\gamma r_b' + \gamma^3 X_{b\alpha})}{1 + \gamma^2} . \quad (2)$$

Thus, by adding inductance L_b in the base circuit the input impedance is changed by the addition of the real term

$[-\gamma^2 X_{b\alpha}/(1 + \gamma^2)]$ and the reactive term $j[X_{b\alpha}\gamma^3/(1 + \gamma^2)]$.

When the input impedance of the transistor becomes part of the peaking circuit of Fig. 1(b), it becomes part of the resonant circuit formed by C_T, R_1, L, and Z_{IN}. The total inductance of this circuit will increase and the total resistance will decrease with increasing frequency due to the variation of Z_{IN}, thus changing the shape of the resonance curve. The greatest effects will be observed at the higher frequencies.

The peaking circuit to be analyzed is shown in Fig. 1(b). The total interstage capacitance is C_T, stray inductance L, and inductance added at base L_b. R_1 is used to adjust circuit Q. The output current of transistor Q_1 is I_0. The transformer has been assumed ideal.

I_2 is the emitter current of Q_2. Solving this circuit we get

$$\frac{I_0}{I_2} = \frac{R_1 + R_{IN} + j(X_{L_{IN}} - X_c)}{-jX_c} \quad (3)$$

and substituting from (1) and (2)

$$\frac{I_0}{I_2} = \frac{R_1 + r_e + r_b' - \dfrac{r_b' + \gamma X_b}{1 + \gamma^2} + j\left[\dfrac{(r_b' + \gamma X_b)\gamma}{1 + \gamma^2} + X_L - X_c\right]}{-jX_c}$$

$$= \frac{R_T - \mu + j(\mu\gamma + X_L - X_c)}{-jX_c} . \quad (4)$$

$$\left| \frac{I_0}{I_2} \right|^2 = \frac{(R_T - \mu)^2 + (\mu\gamma + X_L - X_c)^2}{X^2{}_c}, \qquad (5)$$

where

$$R_T = R_1 + r_e + r_b{}'$$

$$\mu = \frac{r_b{}' + \gamma\omega L_b}{1 + \gamma^2}$$

$$\gamma = f/f_\alpha .$$

REFERENCES

1. McDavitt, M. B., 6000-Megacycle-Per-Second Radio Relay System for Broad-Band Long-Haul Service in the Bell System, A.I.E.E. Transactions, Part I, No. 34, January, 1958.
2. Curtis, H. E., Collins, T. R. D., and Jamison, B. C., Interstitial Channels for Doubling TD-2 Radio System Capacity, B.S.T.J., 39, November, 1960, p. 1505.
3. Ruthroff, C. L., and Tillotson, L. C., An Experimental 'Short Hop' Microwave System, Bell Laboratories Record, June, 1960, p. 202
4. Gammie, J., and Hathaway, S. D., The TJ Radio System, B.S.T.J., 39, July, 1960, p. 821.
5. Linvill, J. G., and Schimpf, L. G., The Design of Tetrode Transistor Amplifiers, B.S.T.J., 35, July, 1956, p. 813.
6. Saari, V. R., Transistor 70-mc IF Amplifier, 1958 Transistor and Solid State Circuits Conference, Philadelphia, Pennsylvania.
7. Ruthroff, C. L., Some Broadband Transformers, Proc. I.R.E., 47, August, 1959, p. 1337.
8. Friis, H. T., Noise Figures of Radio Receivers, Proc. I.R.E., 32, July, 1944, p. 419.

Monitoring the Percussive Welding Process for Attaching Wires to Terminals

By J. C. COYNE

(Manuscript received September 5, 1962)

At present, no satisfactory nondestructive test is known which can be practically applied in the shop to detect a weak percussive-welded connection. Also, the breaking strengths within a population of percussive welds are distributed in such a way that large sample sizes are needed to determine, with reasonable confidence, if the population contains an excessive number of weak welds.

Monitoring the duration of the welding arc and the approach speed of the wire, for each weld as it is made, has been found to provide an effective control of the process. An analysis of the process yields necessary conditions which the arc duration and wire speed must satisfy. Test data are presented which confirm the necessity of these conditions, and show that when they are satisfied, percussive-welded connections can meet quality objectives.

I. INTRODUCTION

The low-voltage percussive welding process as applied to wired connections has been described and its merits pointed out in previous papers.[1-6] Therefore, only a brief review is included here.

Percussive welding is a form of capacitor discharge arc welding. The parts being welded, in this case a wire and a terminal, are themselves the only necessary electrodes (see Fig. 1). Also, the only advance preparation of the electrodes needed is to cut a tip on the wire end. The weld is made by propelling the wire across a 50-volt gap toward the terminal. When the two are nearly touching, an arc is established between the wire tip and the terminal. Initially, the power surge in the arc develops sufficient heat to melt back the wire tip faster than the electrodes are closing. But, with time, the power decays and melting slows down until eventually the electrodes impact, extinguishing the arc and completing the weld. A welding voltage of 50 volts is commonly referred to as "low voltage" as opposed to "high-voltage percussive welding" used to attach the fixed contacts to the wires of the wire spring relay.
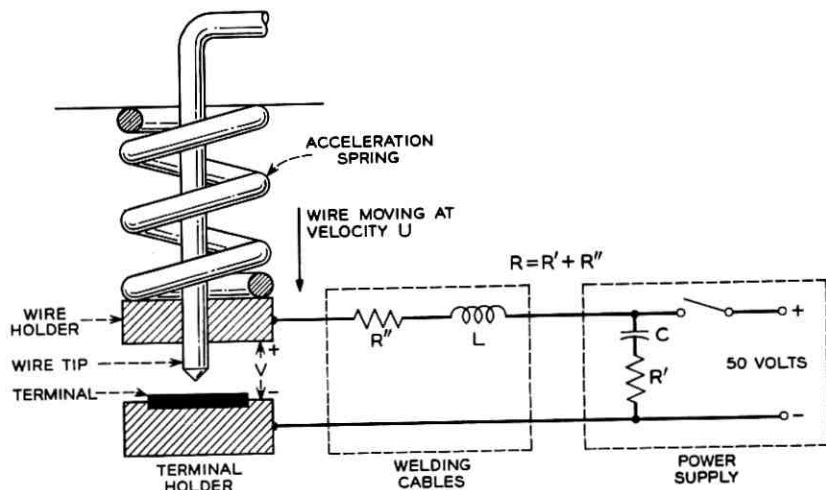
Fig. 1 — Welding circuit.

The welded joint made by this process is usually as strong as the wire itself. Nevertheless, it appears to be characteristic of the process to produce also, if adequate controls are not present, a small scattering of weaker welds, some ranging nearly to zero strength. Naturally, such a distribution of weld strengths is unacceptable from a connection reliability standpoint. Moreover, the usual statistical sampling techniques for assuring satisfactory quality in production have limited value here, as shall be seen later. For this reason, other more effective measures were sought.

Extensive investigations were first conducted to find a nondestructive screening test which might be applied to each completed weld. However, at the present time no satisfactory test has been devised which can be practically applied in the shop. More recently, attention has centered on control of the process by monitoring certain important welding parameters as each weld is made. The analysis and tests pertaining to monitoring are the subjects of this paper.

In the preliminary investigations, the time histories of wire velocity, current, and voltage were examined for aberrations which always coincided with the occurrence of a bad weld. Although this study revealed no easily measurable, consistent predictor of bad welds, it was observed that welds were consistently good if the time duration of the arc was within a certain range. Conversely, the percentage of defective welds increased for either longer or shorter times. Also, this optimum range

shifted, compressed, or expanded depending on the welding circuit constants, the wire velocity, and the wire tip shape.

Further testing and analysis were done to explore the effectiveness as well as the practical possibilities of monitoring arc duration. Two factors seemed to be in favor of this. First, the arc duration is a sensitive indicator of most possible changes in welding conditions. For instance, variations in the electrical characteristics of the welding circuit, wire tip geometry, or approach speed of the wire will all affect the duration of the arc. Second, it is a relatively simple measurement to make.

The areas of fractures for welds occurring at either very short or very long arc durations were observed to have certain distinguishing characteristics. In the case of very short arc durations, the welded area is usually noncircular and has less area than the cross-sectional area of the wire, as shown in Fig. 2(a). This results from the fact that the chisel-shaped tip on the wire does not become completely melted back during the short time duration of the arc. For the same reason, the crater burned into the terminal is usually undersized. The loss of weld strength in this case depends on the reduction of weld area.

In the case of very long arc durations, the smooth break along the interface of the two metals indicates an absence of fusion and suggests that the molten metal film over the electrodes had prematurely solidified. Fig. 2(c) shows an extreme example of this. This type of weld is invariably quite weak. As the arc duration is shortened, a corresponding increase in the average area of fusion in the weld is observed, but with considerable variability. A comparable improvement in the average weld strength also occurs.

If the arc duration is further shortened until it is in the optimum

(a)   (b)   (c)
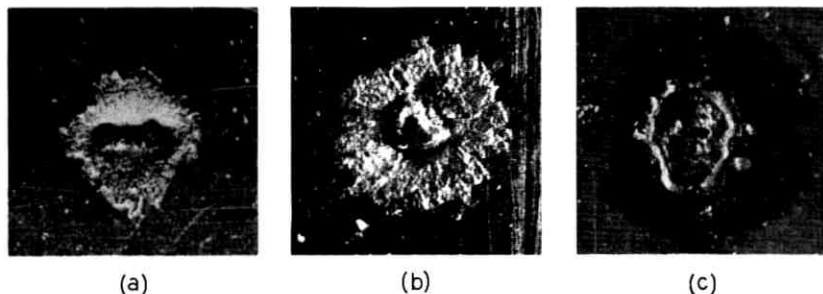
Fig. 2 — Appearance of weld breaks: (a) Excessively short arc durations do not provide sufficient melting energy and result in a weak weld due to the reduced area of fusion. (b) Arc durations within the optimum range result in a strong weld. (c) Excessively long arc durations allow excessive heat conduction losses and result in a weak weld due to premature solidification.

range (assuming that certain other conditions are also satisfied), then a full-area weld of wire to terminal will result as shown in Fig. 2(b), and the weld strength will equal or exceed the wire strength. The other conditions referred to above include: (i) No serious geometrical disparities in the surfaces of the arcing electrodes which would prevent a uniform full-area mating when contact is made. (ii) Adequate melting to remove the entire wire tip. (iii) No spreading or other disturbance of the arc due to surrounding fixture metal at electrode potential being too close. (iv) Development of sufficient force between electrodes at contact to effect good mating and fusion.

The foregoing observations suggest that premature solidification is a major cause of partial-area fusion. Excessive heat losses will occur at long arc durations, since the arc power decays at an approximately exponential rate, while the rate of heat conduction losses decreases inversely as the square root of time. Consequently, heat conduction losses grow with respect to arc power until eventually heat is being conducted away faster than it is being generated. Although radiation and convection losses are also present, these are small by comparison.

A second major cause for partial fusion in the weld zone is that the contours of the arcing surfaces can interfere with full area mating. This is particularly true at very short arc durations for reasons previously mentioned. However, to varying degrees, there will always be some differences in contour, because the two electrode shapes are vastly different at the start of the process. A proper choice of wire tip shape will minimize the disparity in the contours of the arcing surfaces. Any small differences that might still be present will not be harmful if the electrodes impact with sufficient force.

II. ANALYSIS OF WELDING PROCESS

2.1 General

Having seen in a descriptive way why either extreme of arc duration is undesirable, the welding process will be examined analytically in order to derive, for the general case, maximum and minimum monitoring limits of arc duration.

The major independent variables of the process to be used in this derivation are: (i) the wire and terminal metals; (ii) the welding voltage, V_o, which will be taken as 50 volts; (iii) the welding circuit effective resistance, R, and inductance, L, which will be treated as "lumped" constants; (iv) the power supply capacitance, C, also a lumped constant; (v) the wire velocity, U, which will be taken as the

average wire velocity during the time interval of the arc; (vi) the shape
of the wire tip, especially its length, l, and volume, v.

The more important dependent variables to be used in the derivation
are: (i) the welding current, I; (ii) the arc voltage, V; (iii) power ex-
pended in the arc, P; (iv) heat conduction losses, Q; (v) the burnback
of the wire tip, b; and (vi) the arc duration, t_a.

2.2 Some Relationships and Approximations Useful in the Derivation

The first step in the derivation will be to establish useful relationships
for arc voltage, current, power and heat losses. Wherever possible,
approximations will be made to simplify the formulation.

The arc voltage (V) will be approximated by a straight line function,
as shown in Figure 3:

$$V = V_1 - \frac{(V_1 - V_2)t}{t_a} \tag{1}$$

where V_1 and V_2 are the voltages across the arc just after arc initiation
and just prior to arc extinction respectively. For copper wire and nickel-
silver terminal metals, V_2 has been found to be approximately a 10-
volt constant, independent of welding voltage and current, wire velocity,
tip geometry, or wire gauge. Investigation of a few other metals also
showed no appreciable variation from this 10-volt constant.

V_1 showed a slight dependence on the parameters previously men-
tioned, but in each case it was noted that an increase in V_1 corresponded
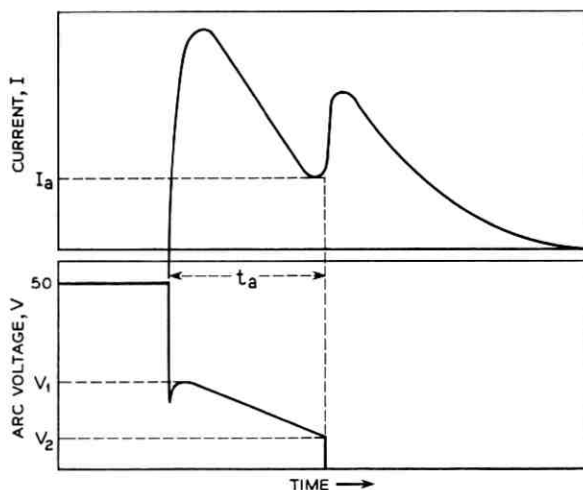to conditions producing a faster melting rate along the wire relative to

Fig. 3 — Typical traces of current and arc voltage.

the wire speed. For instance, V_1 would be increased if either the wire speed decreased or a longer and thinner tip were cut on the wire. In other words, it appears that the arc voltage is mainly a function of the instantaneous separation of the welding electrodes.

Although V_1 varied in the manner described, the magnitudes of the variations were small — in the order of 4 volts even when the peak welding current ranged from 300 to 900 amperes and the wire velocity ranged from 50 to 100 inches per second. A reasonable value for V_1, which will be used in the remainder of the analysis, was found to be 20 volts.

The differential equation describing the current flow for the welding circuit of Fig. 1 is:

$$L \frac{d^2I}{dt^2} + R \frac{dI}{dt} + \frac{I}{C} + \frac{dV}{dt} = 0 \tag{2}$$

with initial conditions at zero current and V_0 volts across C. Substituting (1) into (2), the welding current may be determined. The general solution for welding current consists of two parts: the standard "homogeneous solution" for a capacitor discharge through an RL series circuit, and an additive "particular solution" arising from the time dependent arc voltage function. The damping ratio of the welding circuit will be defined in the usual manner as the dimensionless ratio:

$$D = \frac{R}{2} \sqrt{\frac{C}{L}} \tag{3}$$

and confined to a range of one to infinity. For D equal to one (critical damping), the welding current becomes:

$$I = \left\{ \frac{4(V_o - V_1)\tau}{R} - \frac{(V_1 - V_2)}{R} \frac{(2\tau + 1)}{\tau_a} \right\} \epsilon^{-2\tau} + \frac{V_1 - V_2}{R\tau_a} \tag{4}$$

where $\tau = t/RC$ and $\tau_a = t_a/RC$.

For D approaching infinity (zero inductance) the welding current becomes:

$$I = \left\{ \frac{(V_o - V_1)}{R} - \frac{(V_1 - V_2)}{R\tau_a} \right\} \epsilon^{-\tau} + \frac{V_1 - V_2}{R\tau_a}. \tag{5}$$

For intermediate values of D, the current magnitude will be between (4) and (5).

When τ equals τ_a, either equation gives the current at arc extinction (I_a). If τ_a is sufficiently large, say greater than 4.0, then I_a from both (4) and (5) becomes

$$I_a \cong \frac{V_1 - V_2}{R\tau_a} \tag{6}$$

since in either case the exponential term becomes small by comparison.

If τ_a is not sufficiently large to use this approximation, then the contribution to I_a made by the exponential term should be included. Account is taken of this contribution by rewriting (6) with N substituted for $V_1 - V_2$:

$$I_a = \frac{N}{R\tau_a} = \frac{NC}{t_a} \qquad (7)$$

where N is a number larger than $V_1 - V_2$ and dependent on τ_a. N may be calculated by equating (7) to (4) or (5), with $\tau = \tau_a$. Some representative values are shown in Table I, calculated for $V_o = 50$, $V_1 = 20$ and $V_2 = 10$ volts.

The instantaneous power, P, expended in the arc is the product of V, given by (1) and I, given by (4) or (5). At arc extinction, using (7) for I, and with $t = t_a$,

$$P_a = \frac{NCV_2}{t_a}. \qquad (8)$$

The electrical energy, E_H , dissipated as heat in the arc, as a function of time, is obtained from a time integral of the power. To simplify the formulation without any great loss of accuracy, it will be assumed that the arc voltage remains constant (equal to V_1) and therefore the second terms of both (4) and (5) vanish. Then for $D = 1$

$$P \cong 4V_1 \frac{(V_o - V_1)\tau\epsilon^{-2\tau}}{R} \qquad (9)$$

and for $D \to \infty$

$$P \cong V_1 \frac{(V_o - V_1)\epsilon^{-\tau}}{R} . \qquad (10)$$

Using this approximation, the errors in arc voltage and current are of opposite sign, and therefore these partially cancel in the product. The remaining error is minimum at arc initiation and maximum at arc extinction. Table II has been prepared to show the relative magnitude of

TABLE I — VALUES OF N IN (7)

τ_a	$D = 1$	$D \to \infty$
2.0	17.87	16.77
2.5	14.64	15.33
3.0	12.50	13.97
3.5	11.27	12.85
4.0	10.62	12.03
4.5	10.02	11.39

Table II — Arc Power Versus τ

τ	(5) times (1)	(10) (approximation)
0	$600/R$	$600/R$
1	$218/R$	$221/R$
2	$92/R$	$81/R$
$3 = \tau_a$	$47/R$	$30/R$

the error made by using the approximation. The calculations have been made for $\tau_a = 3.0$ and for the same values of V_o, V_1, and V_2 used previously. Since the contribution to the total arc energy is many times greater at arc initiation than at arc extinction, it is evident that the percentage error is further reduced by the integration.

Therefore, for $D = 1$

$$E_H \cong CV_1(V_o - V_1)(1 - \epsilon^{-2\tau} - 2\tau\epsilon^{-2\tau}) \tag{11}$$

and for $D \to \infty$

$$E_H \cong CV_1 (V_o - V_1) (1 - \epsilon^{-\tau}). \tag{12}$$

The rate of heat conduction losses (dQ/dt) at any time after arc initiation can be taken as

$$\frac{dQ}{dt} = \frac{TkA}{\sqrt{\pi at}} = \frac{K}{\sqrt{t}} \tag{13}$$

where A is heat conducting area; e.g., wire area,

k is thermal conductivity,

a is thermal diffusivity, and

T is molten metal temperature.

Substituting, for instance, the appropriate constants for 24-gauge copper wire, K becomes approximately 7.0 watt $(\sec)^{\frac{1}{2}}$. This number is based on using an estimated average value for T. There is evidence that some very small amount of copper vaporizes during arcing.[3, 6] Consequently, the molten copper must range between its melting temperature and its boiling temperature.

The total heat conduction loss between arc initiation and time t is

$$Q = 2K \sqrt{t}. \tag{14}$$

2.3 Maximum Limit on Arc Duration

A maximum limit on arc duration, $t_{a(max)}$, may now be specified in order to prevent premature solidification. It is required that the power

dissipated in the arc just prior to arc extinction always exceed heat conduction losses. Assuming half of the arc energy is expended at each electrode and that radiation and convection losses are negligible,

$$\frac{P_a}{2} > \frac{dQ}{dt} \, . \tag{15}$$

Implicitly, it is postulated here that if the above inequality is satisfied at arc extinction, then it must necessarily be satisfied for all earlier times.

Substituting (8) and (13) into (15) and solving for t_a

$$t_a < \left(\frac{NCV_2}{2K}\right)^2 = t_{a\,(max)} \, . \tag{16}$$

Equation (16) is our first monitoring inequality and it is plotted in Fig. 4 for different RC products, and for the two extreme damping conditions, versus capacitance.

This formulation of $t_{a(max)}$ includes an inherent safety margin with regard to monitoring the process since the time to accomplish solidification is not taken into account.

2.4 Minimum Limits on Arc Duration

Burnback (b) is the length of wire melted back during arcing. With negligible error, it can be assumed that the electrodes are touching at arc initiation. (A potential of 50 volts will not normally initiate an arc across the air gap until it has closed to approximately 0.00005 inch.)[4] Consequently, the distance advanced by the wire during arcing must be equal to the average wire velocity, U, multiplied by the arc duration. This product will be very nearly equal to the total melted wire length, differing only by the thin layer of molten copper on the wire and by the crater depression in the terminal. These are both relatively small and of opposite sign. Therefore,

$$b = Ut_a \, . \tag{17}$$

In order to melt the entire wire tip length, l,

$$b > l + \text{safety margin.} \tag{18}$$

If the wire velocity is measured by the time, t_v, for the wire to traverse contacts separated by l, then (18) reduces to the monitoring inequality:
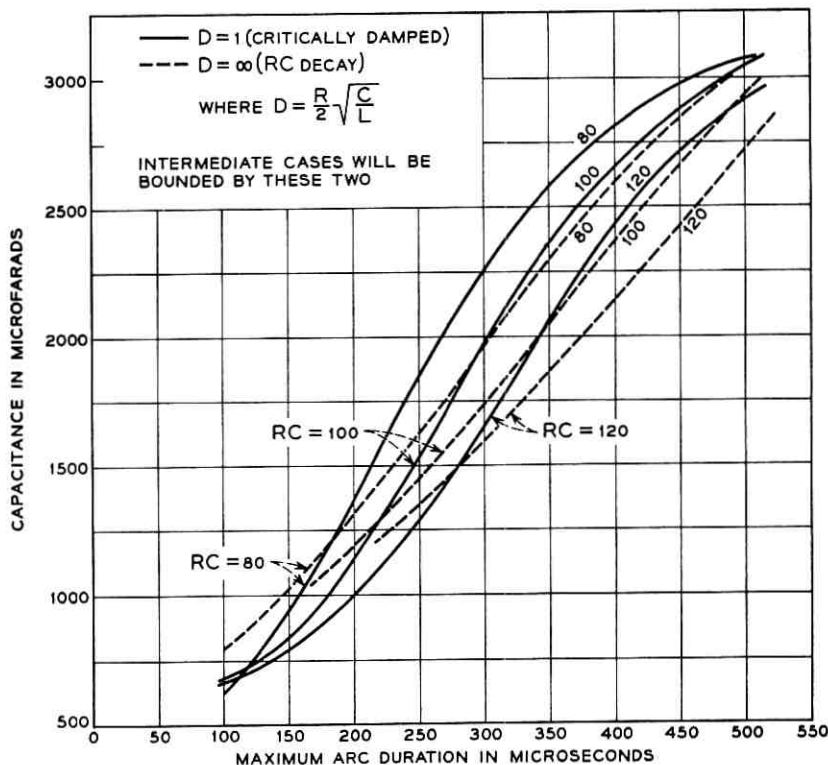
$$t_a > t_v \, . \tag{19}$$

Fig. 4 — Computed curves for $t_{a(\max)}$: 24-gauge copper wire, nickel-silver terminal, welding voltage 50v, $V_1 = 20$v, $V_2 = 10$v, $K = 7$ watt(sec.)$^{\frac{1}{2}}$.

Preliminary tests indicated the need for another minimum limit on arc duration, $t_{a(\min)}$, independent of t_v, since a scattering of weak welds sometimes occurred at slightly longer arc durations than predicted by (18). In general, $t_{a(\min)}$ must be at least long enough to provide for sufficient melting energy; however, other reasons may make a longer time necessary. This is discussed further in Section V. On this basis,

$$t_a > t_{a(\min)} \tag{20}$$

where $t_{a(\min)}$ is determined empirically. For the tests reported in Section IV, $t_{a(\min)}$ was found to be 200 microseconds.

2.5 Necessary Power Supply Capacitance

The melting energy (E_M) necessary to remove the entire wire tip, assuming the molten metal is at its melting temperature, is

$$E_M = (T_M c + \lambda)\rho v. \tag{21}$$

where T_m is the metal melting temperature, c is the specific heat, λ is the latent heat of fusion, ρ is the metal density and v is the wire tip volume.

The necessary power supply capacitance, C, can be calculated from an energy balance

$$\tfrac{1}{2} E_H - Q = E_M . \tag{22}$$

Substituting in (12), (14), and (21) and evaluating for average damping conditions, for copper wire and for $t_a = t_{a(\max)} = 3RC$,

$$C \cong 600\alpha v \tag{23}$$

where α is a safety factor at $t_{a(\max)}$. In general, α must include allowances for the reduction in input energy resulting from shorter arcing times as well as the greater demand for energy resulting from the variational increases in tip volume. The safety factor at $t_{a(\min)}$, (β), will be, of course, smaller than α and may be evaluated from the following approximate relationship

$$\beta \cong \frac{1 - \exp\ (-t_{a\ (\min)}/RC)}{1 - \exp\ (-t_{a\ (\max)}/RC)}\ \alpha. \tag{24}$$

An independent measurement of the melted wire volume versus capacitance agreed closely with the volume predicted by (23).

2.6 Monitoring Inequalities

To recapitulate, the inequalities in (16), (19), and (20) restrict the allowable range of arc durations for percussive welding

$$\begin{Bmatrix} t_v \\ t_{a(\min)} \end{Bmatrix} < t_a < t_{a(\max)} \tag{25}$$

where $t_{a(\min)}$ and $t_{a(\max)}$ are fixed limits evaluated either from the constants of the process or by empirical methods, and t_a and t_v are measurements made for each weld. The measurements of t_a and t_v can be made in a number of ways. Fig. 5 shows one method, in which contacts separated by the minimum burnback distance provide step voltages which are indicative of t_v , and a coil around one welding lead provides induced voltages at the beginning and the end of the arc which are indicative of t_a . These voltage indications are fed to a timer and alarm — any suitable apparatus capable of determining whether the monitoring inequalities in (25) are satisfied. Various other measurement methods might be em-
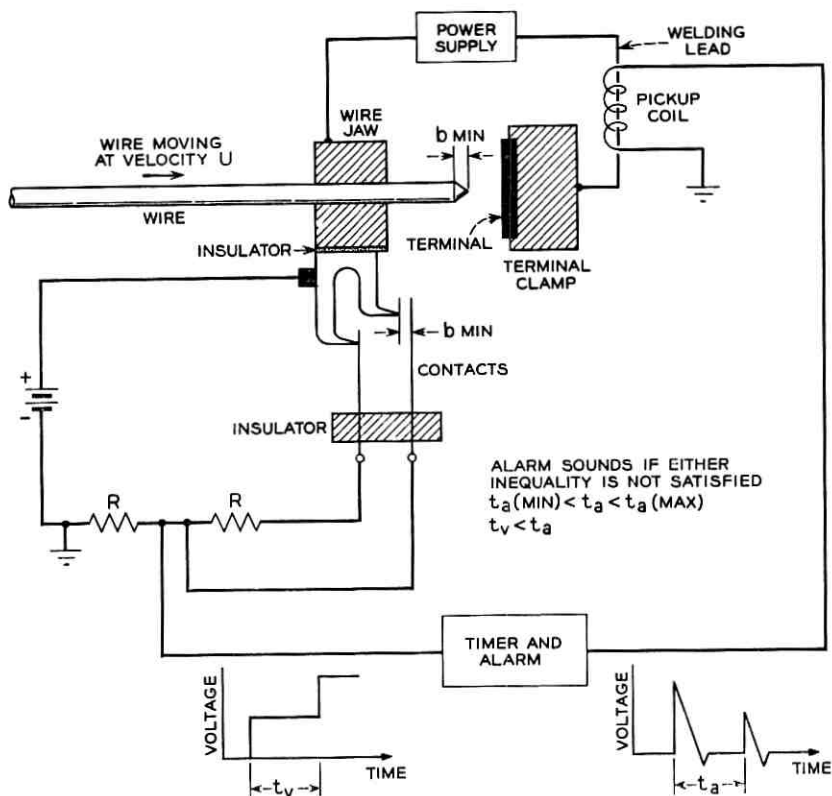
Fig. 5 — Wiring schematic of monitoring circuit.

ployed, such as a light beam and photocells in the case of t_v and the arc voltage change in the case of t_a.

III. OTHER CONSIDERATIONS

The described monitoring method naturally does not preclude the possibility of a defective weld. Other causes still exist.

3.1 Bad Wire Tip

Inconsistently or improperly formed tips are a potential hazard. Usually, monitoring will detect this condition since arc duration is sensitive to variations in either tip length or volume. Nevertheless, certain tip variations — for instance, variations in symmetry — are undetected by the measurement of arc duration.

3.2 Arc Dispersion

The wire as it approaches the terminal must "see" the terminal to the exclusion of any other metal at terminal potential. Otherwise the arc will spread and be weakened, resulting in poor welds despite monitoring.

3.3 Mechanical Effects at Closure

The most critical time in the weld process is, of course, when the two weld metals come together and fuse. The possible influences on weld strength at this time are manifold; the following are representative. At closure, the arcing surfaces should match reasonably well in order to achieve a uniform fusion zone. Also, any large disparity in contour will cause a time lag between initial electrode contact (arc extinguishes) and final full-area contact. The time lag is increased if the wire is appreciably decelerated. During such a time lag, premature solidification might be a problem in those areas that mate last. The shape of the original wire tip, especially its symmetry, plays an important role in determining how well the electrode surfaces will mate, since arcing generally is concentrated between the nearest points on the two electrodes.

With regard to impact forces at the weld interface, it has been found that up to a point these forces improve both connection strength and fatigue life. This might be attributed to various effects, some of which could be, (i) overcoming small contour differences on the arcing surfaces, (ii) heading, or increasing the area near the weld zone, (iii) strain hardening the copper above the weld, (iv) promoting better mixture in the fusion zone.

IV. DESCRIPTION OF TESTS

4.1 General

Welds were made for test under controlled laboratory conditions and monitored in accordance with the analysis. In order to test only the effects of the parameters in the monitoring inequalities, all other welding conditions were optimized to the best of our knowledge.

Wire tips were cut with a hand-held pair of diagonal wire cutters. Despite its seeming crudeness, this tip has been found to be sufficiently reproducible and has given good results in the past. The tip length on the wire averaged 0.019 inch with approximately half the volume of an equal length of wire. The wire was 24-gauge tinned copper.

The terminal was held flush with or slightly projecting from the ter-

minal clamp in order to minimize arc dispersion. The terminal was 0.025 inch x 0.062 inch nickel-silver.

The welding power supply contained 2300 microfarads capacitance. This value produces an adequate energy margin. No delay line or other type of network was used to supplement the capacitor bank.

The wire was accelerated with an adjustable tension spring and was permitted to fully impact into the terminal so that the weld interface experienced the complete impact forces. Weld strengths were measured by the "combined test" whereby the wire is bent over parallel to the terminal and pulled until either the wire or the weld fails. This test has proven to be more sensitive and more severe than a straight tension test. The break was classified as a "wire break" if the failure occurred in a portion of the wire completely above the terminal surface. Actually, these "wire breaks" occurred randomly along a span of wire of about three inches, so that usually there was very little difficulty in distinguishing a "wire break" from a "weld break." Furthermore, if a break occurred close to the weld-affected zone, and its classification was doubtful, it was usually taken to be a "weld break" since the probability is small that a "wire break" would occur randomly this close to the weld.

4.2 *Phase I*

The testing was divided into two phases. In Phase I, the burnback inequality at all times was satisfied, whereas the arc duration was varied in order to demonstrate that (16) and (20) are necessary conditions for optimizing the welding process. The arc duration was varied by means of changing the wire velocity. At least 600 welds were made and tested within each 50-microsecond interval between 150 and 500 microseconds, making a total of 4730 welds. The testing was not extended below 150 microseconds because of practical limitations of high wire velocity.

The following steps, based on the analysis, show the calculations made to determine the monitoring limits applicable to this first phase of testing.

(a) The welding circuit constants were determined by substituting measurements from an oscilloscope current trace into the theoretical current expression. Accordingly, the effective total series circuit resistance was calculated to be 0.040 ohm, the inductance 0.75 microhenry and the power supply capacitance 2300 microfarads. The capacitance figure agrees closely with an independent measured value. From (3) the damping ratio, D, becomes 1.1. The RC product becomes 92 microseconds.

(b) From Fig. 4, using the critical damping curves, $t_{a(max)}$ is 320 microseconds.

(c) In accordance with previous tests at the same welding conditions, $t_{a(\min)}$ was taken at 200 microseconds.

(d) The average wire tip volume was measured to be approximately 3×10^{-6} cubic inches. Using (23), the safety factor (α) at 320 microseconds is calculated to be 1.25 for a capacitance of 2300 microfarads. The safety factor is reduced to about 1.15 at 200 microseconds using (24).

4.3 *Phase II*

In Phase II, the burnback was varied in order to demonstrate the necessity of adequate burnback (18). The burnback was varied from 0.012 to 0.022 inch to span the average tip length of 0.019 inch. At least 200 welds were made and tested within each 0.002-inch increment, making a total of 1126 welds. The test was planned to be within the 200 to 350-microsecond interval of arc duration at a constant wire velocity of 63 inches per second. The only welding parameter independently varied was circuit resistance. This method was chosen not only to facilitate a sensitive control of burnback but also to avoid the possibility of exceeding $t_{a(\max)}$, since it was desired in this phase to test only the effect of reducing burnback.

The computations of monitoring limits for Phase II follow. Since resistance was varied in this test, no one value of $t_{a(\max)}$ applies throughout. But to indicate a typical case, at the longest range of arc durations in this test (320–350 microseconds), the incremental resistance was approximately 0.060 ohm, making the total resistance 0.1 ohm. The RC product then is 230 microseconds and $t_{a(\max)}$ is calculated from (16) and Table I to be approximately 600 microseconds ($N = 15$; $\tau_a = 2.5$).

From (24), the safety factor for energy at 320 microseconds is approximately 1.0, which means that this is the minimum allowable arc duration on a "sufficient energy basis." The burnback at 320 microseconds and 63 inches per second is 0.020 inch, in good agreement with the average tip length (0.019 inch) and giving us a check on the calculations. Consequently, in this test we can expect weld quality to degrade for arc durations less than 320 microseconds.

V. DISCUSSION OF TEST RESULTS

5.1 *General*

The test results of Phase I, shown in Fig. 6, clearly demonstrate that, for the conditions of this test, the welding process is optimized when the arc duration is between 200 and 300 microseconds. Within this range, 99.2 per cent of the samples failed in the wire and 99.85 per cent of the samples exceeded 90 per cent of the wire strength.
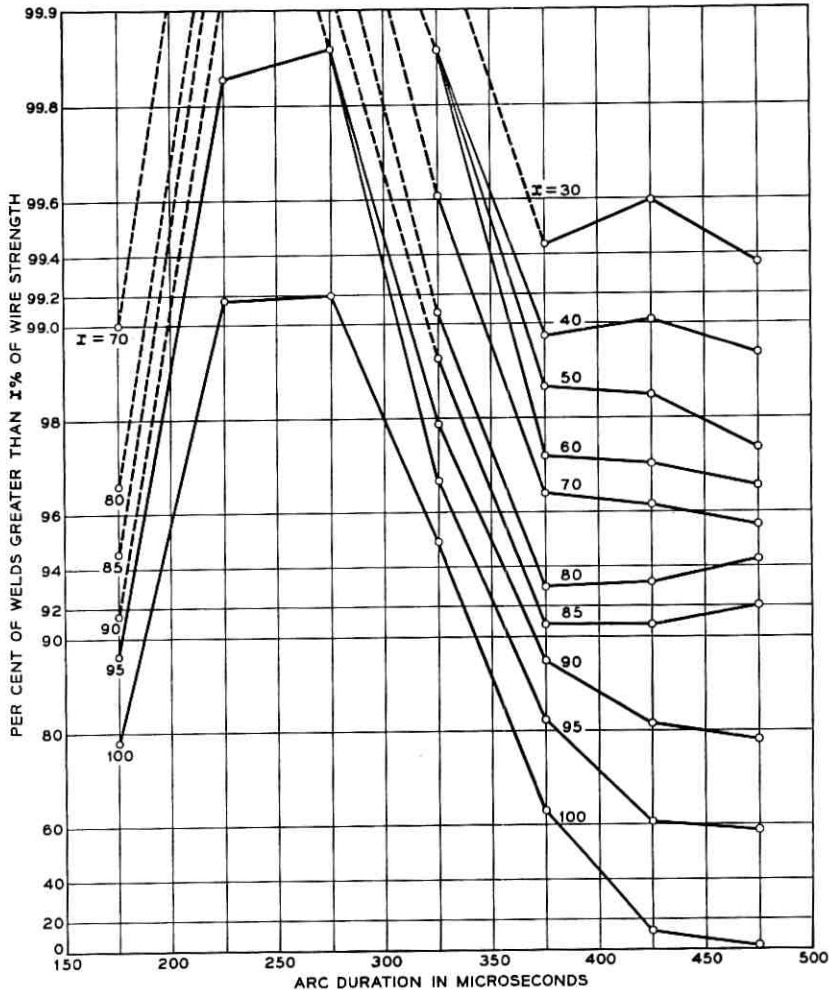
Fig. 6 — Results of test No. 1, weld strength vs arc duration: 24-gauge copper wire, nickel-silver terminal, capacitance 2300 μf, resistance 0.04 ohm, diagonal-cut wire tip, variable wire velocity.

The estimated limits of the optimum range, 300 and 200 microseconds, agree closely with both the results of previous tests and the analysis. At arc durations less than 200 microseconds there is a relatively fast drop-off in weld quality. At arc durations longer than 300 microseconds there is also a dropoff in quality, although not as rapid as in the first case. As the arc duration deviates further from optimum toward longer times, the

percentage of weaker welds begins to level off. Although this was unexpected, it may be explained plausibly on the basis of varying solidification time.

The time interval that can safely elapse after $t_{a(\max)}$ depends, of course, on the heat contained in an unknown thickness of molten film over the electrodes. This thickness is probably quite variable. At normal welding velocities, it is quite thin; however, there is evidence that it generally increases on the wire for longer arc durations (slower wire speed).[5] Therefore, the time to solidify the molten film will also be variable and will, in general, become longer at longer arc durations.

The results of the second phase of testing, shown in Fig. 7, indicate a rapid loss of weld quality as the burnback is decreased. In agreement with the preliminary calculations, the weld quality was relatively good for arc durations greater than 320 microseconds or burnback in excess of 0.020 inch.

5.2 *Mapping Over-all Test on Burnback vs Arc Duration Axes*

The results of the over-all test program can best be discussed with the aid of Fig. 8. On this figure, the over-all test range is mapped into five regions and the monitoring limits visually displayed with respect to these regions on a burnback versus arc duration axis. Some further explanation will be helpful in interpreting this figure.

For any weld, the average wire velocity during arcing and the duration of the arc may be measured. These measurements determine a point on Fig. 8, where the product of the measurements (burnback) is the ordinate and arc duration is the abscissa. As wire velocity is varied and a number of points are plotted, a curve takes shape. This curve relates the melted wire distance to the time after arc initiation for a particular set of circuit conditions and wire tip. Accordingly, the upper curve of Fig. 8 is the sketched-in melting curve for the welding conditions in the first phase of the test program. The lower set of curves is for Phase II, where each curve in the set would be determined by a corresponding value of circuit resistance. The intersection of a particular melting curve with a line of constant velocity locates the arc duration and burnback of a weld for the welding conditions corresponding to that melting curve. This point will be referred to as an "operating point" of the welding process. Naturally, there will be a statistical scattering of the data about the operating point due to normal test variations.

The first test phase extends over regions A, B, and C while the second test extends over regions D and E. Note that in Phase I the operating points coincide with a melting curve, while in Phase II the operating
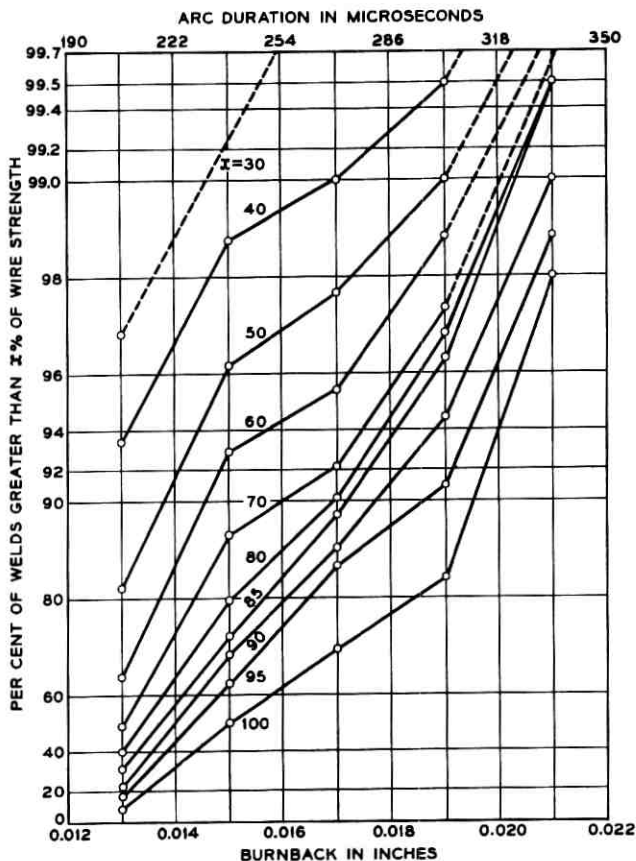
Fig. 7 — Results of test No. 2, weld strength vs burnback. Conditions same as test No. 1, except that wire velocity is constant 63 in/sec and resistance is varied.

points lie along the 63 inches per second velocity line. Table III summarizes the test results for the regions of Fig. 8 and indicates the monitoring inequalities which were violated in each case. The "order of merit" of weld quality is based on the test results in Figs. 6 and 7. Table III, in conjunction with Fig. 8, demonstrates the necessity of each of the monitoring inequalities.

In Fig. 8 it is seen that the constant burnback line of 0.020 inch forms the boundary between regions D and E. The fact that this boundary, only 0.001 inch more than the average tip length, effectively separates the good from the poor welds implies that burnback is the controlling factor in weld quality in Test Phase II. An independent
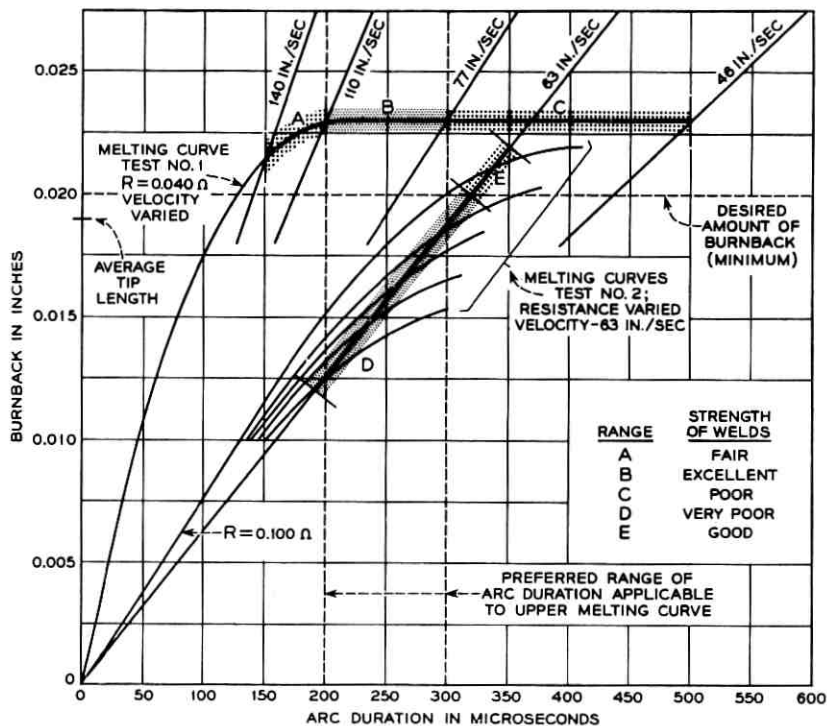
Fig. 8 — Graphical representation of over-all test program.

$t_{a(min)}$ limit therefore is obviated in this case and the entries in the table agree with the burnback column. On the other hand, in region A, weld quality is poorer, despite an average burnback of 0.022 inch. Here $t_{a(min)}$ is the controlling limit.

It is significant that there is this distinction in the controlling lower limit of arc duration, and the question is asked, Why does weld quality

TABLE III

Test Phase	Region	Weld Quality Order of Merit	$t_a < t_{a(max)}$	$t_a > t_{a(min)}$	$b > 0.020$, or $t_a > t_v$
I	A	3	Satisfied	Violated	Satisfied
I	B	1	Satisfied	Satisfied	Satisfied
I	C	4	Violated	Satisfied	Satisfied
II	D	5	Satisfied	Violated*	Violated
II	E	2	Satisfied	Satisfied*	Satisfied

* See Section 5.2.

drop off in region A? A comparison of welding conditions leads to some speculations as to the reasons. In region A, the electrodes close at the highest current density and with the fastest wire speed. It is reasonable that extremes in these variables could be undesirable. For instance, during the high current portion of the arcing period, the melting process and temperatures are undergoing rapid transitions. Probably some vaporization is taking place. If the wire speed is fast, closure occurs during this high current portion of the arcing period. Under such volatile conditions, the weld process is apt to be erratic. A relatively slower speed would lengthen the arc duration and allow conditions to become more stable.

In general, poor weld quality will result if the wire speed is allowed to become too slow. This is seen clearly in Fig. 8 by letting the operating point move to the right along the upper melting curve. By the time the wire speed has decreased to 63 inches per second, the "operating point" is well into the poor area in region C. However, if we now hold the wire speed constant and add 0.060 ohm to the welding circuit, the operating point shifts into region E, and to a different melting curve, resulting in an improvement in weld quality. Hence, a change of circuit resistance in this case will restore the good weld quality which had been lost by slowing the wire speed. This suggests that the slow wire speed is responsible for the poor weld quality, not through any mechanical effects, but rather by lengthening the duration of the arc past a maximum safe limit for a particular melting curve.

5.3 *Statistical Interpretation of Data*

The data of Fig. 6 are replotted in Fig. 9. There are seven distributions shown, one for each 50-microsecond increment of arc duration. In general, all the distributions consist of two portions. The first portion is the distribution of welds which failed near the weld at less than wire strength, while the other is the distribution of welds which failed in the wire. In each case the portions representing weld failures are fairly linear, with a steep slope; however, in some cases the lines curve as they join the wire break portions. Because of the choice of ordinate, the portions representing wire failures are horizontal straight lines at 100 per cent of wire strength.

To aid in interpreting these curves, the wire strength is thought of as an upper cutoff boundary on the weld strength. The following example is an analogous situation: suppose pennies were being pitched to a line. The density of pennies would be normally distributed on either side of the line. Now suppose that an adhesive-coated wall were placed at the line, all other things remaining the same. All the pennies which would
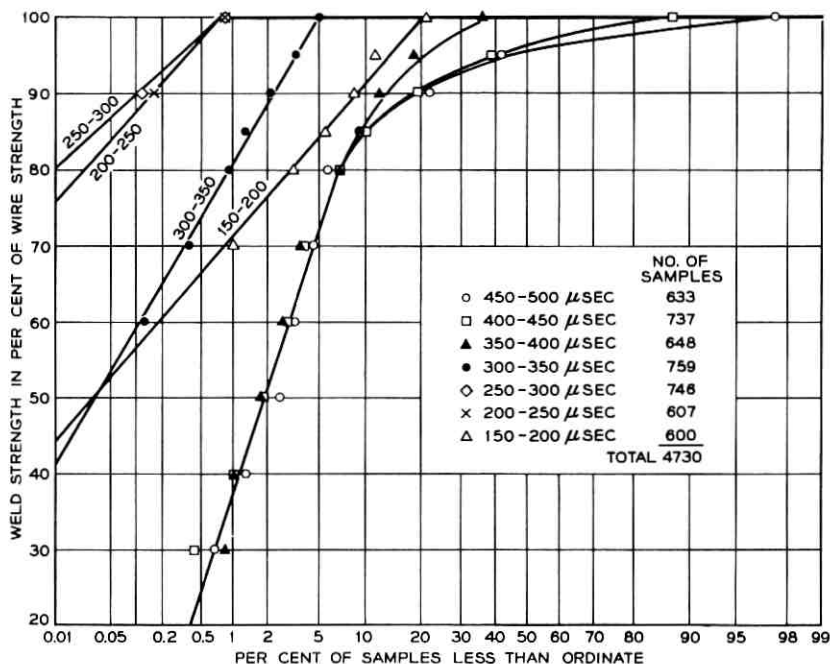
Fig. 9 — Cumulative distribution of data from test No. 1.

ordinarily pass over the line now hit the wall and stick to it. A portion of a normal distribution is left to one side of the wall and all the remaining pennies stuck to the wall. Plotting the distribution of pennies around the wall would result in a distribution similar to that found in the case of percussive-welded connections. The pennies which stick to the wall represent wire breaks, and the truncated normal distribution of pennies to one side of the wall represents weld breaks.

It appears, therefore, that the distributions of weld strengths in Fig. 9 are normal below some value of weld strength since the lower strength portions of all the curves are essentially straight lines. Therefore, the probabilities associated with the lower tails of the distribution can be read directly. The two curves which represent the welds satisfying the monitoring inequalities approach the 0.01 percentage point of the distribution at approximately 76 per cent and 80 per cent of wire strength. All the other curves are 45 per cent of wire strength or less at this equivalent point on the distributions. Thus these data indicate that, with monitoring, there is a much smaller probability of finding weak welds than in an equivalent size sample of unmonitored welds.

Based on these data, it appears that percussive welding is capable of meeting present quality objectives of a maximum defect rate of one in ten thousand, a defect being defined as a connection which is weaker than 60 per cent of wire strength.

5.4 *Generalizations Regarding Weld Strength Data*

The following are some typical characteristics of the strength distribution of low-voltage percussive welds. These generalizations are not drawn solely from the present data, but have been observed repeatedly in various sets of data.

(a) When the process is optimized, the distribution of weld strength is bunched tightly around the wire strength.

(b) When the process is off optimum, a long tail develops in the distribution of weld strength. This tail is made up of relatively weak welds which usually comprise a small percentage of the total number of welds in the distribution. The tail is an incomplete normal distribution, bounded on the upper side by the wire strength.

(c) The over-all quality of a population of percussive welds is defined completely in terms of the tail of the distribution.

The following factors all tend to limit an effective assurance of quality based on small sample size destructive testing:

(a) Only a very small percentage of the data yields useful information regarding the tail of the distribution.

(b) The large standard deviation of the tail makes large sample sizes necessary if confidence in the sampling is not to be sacrificed.

(c) The allowable percentage of defects is small (one in ten thousand). Therefore, large sample sizes are required in order to gain a reasonable degree of confidence in the sampling.

The conclusion, based on these factors, is that small sample size destructive testing alone cannot assure quality levels of less than one defect in ten thousand. The sample size required to assure this degree of connection reliability, with a reasonable degree of confidence, would be prohibitive from a cost standpoint.

VI. SUMMARY AND CONCLUSIONS

1. It is characteristic of the low-voltage percussive welding process that the relatively weak welds in a population will be contained in a long tail of the distribution. The welds in this tail are distributed normally, with large variance, and are bounded on the high side by the wire strength. The vast majority of welds are not contained in the tail and have essentially the distribution of the wire strength. This double

nature of the over-all distribution, unfortunately, severely limits the effectiveness of small sample size destructive testing.

2. In view of the difficulties encountered in statistical sampling by destructive testing and because of the lack of any effective and practical method of testing nondestructively, it was sought to attain consistently reliable percussive welds by continuous control of the process, i.e., by monitoring certain important welding parameters as each weld is made.

3. Preliminary tests indicated that the time duration of the arc was the most sensitive over-all indicator of variations in the process. The effectiveness of arc duration is further increased if the measurement is related to the approach velocity of the wire. Since both these parameters are relatively easy to measure, arc duration and wire velocity were chosen as the most effective and practical measurements for purposes of monitoring.

4. It was found that there is a range of arc duration for which the process is optimized. Weak welds which occur at longer arc durations are attributed to excessive heat conduction losses which result in premature solidification of the molten metal film over the arcing electrodes. An analytical derivation is made for a maximum allowable arc duration, which is defined as that time when heat conduction losses begin to exceed power input to the weld.

5. At excessively short arc durations, weak welds can be attributed to insufficient arc energy, which results in insufficient burnback on the wire to produce a full-area weld. Accordingly, the minimum allowable arc duration is shown to be the time required for the moving wire to travel a distance equal to its tip length.

6. In addition to satisfying the proper limits of arc duration, other conditions should be satisfied in order to attain good welding results. These include a properly and consistently made wire tip, no excessive spreading of the arc, and sufficient force between electrodes at contact to affect good mating and fusion.

7. A test program of nearly 6000 welds confirmed that the low-voltage percussive welding process is optimized when the arc duration is between an upper and a lower value. Furthermore, these limits, as indicated by the data, agree with the predicted values from the analysis and earlier testing.

8. Based on the data of weld strength for welds satisfying the monitoring inequalities, the low-voltage percussive welding process is capable of meeting present quality assurance objectives of less than one defective connection in ten thousand, where a defective connection is defined as one weaker than 60 per cent of the wire strength.

VII. ACKNOWLEDGMENTS

The author gratefully acknowledges the guidance and encouragement given by C. B. Brown and H. M. Knapp. Thanks are also due to G. W. Mills for many valuable discussions and to J. J. Dunbar for assistance in the laboratory.

REFERENCES

1. Sumner, E. E., Some Fundamental Problems in Percussive Welding, B.S.T.J. **33,** July, 1954, p. 885.
2. Gellatly, J. S., Johnson, K. F., and Quinlan, A. L., Low Voltage Percussive Welding, The Western Electric Engineer, **3,** July, 1959, p. 22.
3. Boyle, W. S., and Smith, J. L., unpublished work.
4. Boyle, W. S., and Smith, J. L., Low Voltage Arc Welding Circuit for Use with Percussion Hand Welder, U.S. Patent No. 2,836,703, 5/27/58.
5. Coyne, J. C., unpublished work.
6. Dobrjanskyj, L., unpublished work.

A Theorem on the Distribution of Weights in a Systematic Code†

By JESSIE MACWILLIAMS

A systematic code of word length n is a subspace of the vector space of all possible rows of n symbols chosen from a finite field. The weight of a vector is the number of its nonzero coordinates; clearly any given code contains a certain finite number of vectors of each weight from zero to n. This set of integers is called the spectrum of the code, and very little is known about it, although it appears to be important both mathematically and as a practical means of evaluating the error-detecting properties of the code.

In this paper it is shown that the spectrum of a systematic code determines uniquely the spectrum of its dual code (the orthogonal vector space). In fact the two sets of integers are related by a system of linear equations. Consequently there is a set of conditions which must be satisfied by the weights which actually occur in a systematic code. If there is enough other information about the code, it is possible to use this result to calculate its spectrum.

In most systems of error correction by binary or multiple level codes the minimum distance between two code words is an important parameter. (The distance between two code words is the number of coordinate places in which they differ.) Much attention has been given to devising codes which have an assigned minimum distance.

The weight of a code word is its distance from the origin. The distance between two code words is the weight of the vector obtained by subtracting one from the other, coordinate by coordinate. If the code words form a vector space, this vector is itself a member of the code. Such codes are called systematic codes. The set of integers specifying the weight of each code word is then exactly the same collection of numbers as the set of integers specifying the distance between each pair of code words.

---

† This paper formed part of a thesis presented to the Department of Mathematics, Harvard University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Thus it is customary to talk about weight properties rather than distance properties of systematic codes.

In many cases, practically all that is known explicitly about the distribution of weights in a code is that the weight has a certain minimum value. Recent studies have shown that it would be useful (e.g., in the study of real life channels) to have more information. We would like to be able to answer questions of the following sort:

i. Given a method (implemented or theoretical) for constructing a systematic code, how many elements of each weight will be obtained? (It is a safe assumption that nobody will want to write out the code vectors and count them.)

ii. Given a set, u_1, u_2, $\cdots$, u_s, of positive integers, is it possible to construct a systematic code with elements of these weights only?

In theory there exists a method of answering these questions.[1,2,3] Unfortunately this method is quite difficult to apply. The purpose of this paper is to give a different method which is in some ways more useful.

We show that the spectrum of a systematic code determines uniquely the spectrum of the dual code (the orthogonal vector space). In fact, the two sets of integers are related by a system of linear equations. Our main theorem shows how to obtain this system of equations.

In Section I we give definitions and statements of the main theorem and of some corollaries. Section II contains proofs of these theorems. Section III describes how the results of Section I may be applied.

I. DEFINITIONS, NOTATION AND A STATEMENT OF THE MAIN THEOREM

Let F be a finite field of q elements; q is a prime power. Let F^n denote the direct sum of n copies of F. F^n is the set of all possible row vectors of length n, in which each coordinate is an element of F. Addition of two vectors is defined coordinate by coordinate, under the rules prevailing in F.

F^n is a vector space of dimension n over F. Choose a basis consisting of the n vectors

$$\epsilon_1 = (1, 0, 0, \cdots, 0)$$
$$\epsilon_2 = (0, 1, 0, \cdots, 0)$$
$$\cdots\cdots\cdots\cdots\cdots$$
$$\epsilon_n = (0, 0, 0, \cdots, 1).$$

An element u of F^n is then expressed uniquely as

$$u = \sum_{i=1}^{n} u_i \epsilon_i, \qquad u_i \epsilon F.$$

We write $u = (u_1, u_2, \cdots, u_n)$.

The weight of u is defined to be the number of u_i which actually appear in this sum — i.e., the number of nonzero coordinates in the vector u.

An alphabet is any subspace of F^n; a vector belonging to the alphabet α is called a letter of α. It may happen that every letter of α has zero as the jth coordinate — this case is not excluded.

The scalar product of two vectors,

$$u = \sum_{i=1}^{n} u_i \epsilon_i, \quad v = \sum_{i=1}^{n} v_i \epsilon_i, \qquad u_i, v_i \epsilon F,$$

is $u*v = \sum_{i=1}^{n} u_i v_i$, where the multiplication and addition are carried out in F. If F is the field of two elements $0, 1$, for example, the scalar product of $(1, 1, 0)$ with itself is $1 \cdot 1 + 1 \cdot 1 = 0$.

Two vectors u, v are orthogonal if their scalar product is zero. In the example above, $(1, 1, 0)$ is orthogonal to itself.

The orthogonal complement of an alphabet α is the set of all vectors of F^n which are orthogonal to every vector of α. It is clear that these vectors also form an alphabet, say $\mathcal{B}$, which is called the dual alphabet of α. If k is the dimension of α, the dimension of $\mathcal{B}$ is $m = n - k$.

The main result of this paper is as follows (the proof is given in Section II):

Let α be an alphabet of dimension k, and $\mathcal{B}$ the dual alphabet of dimension m. Let A_i, B_i denote the number of letters of weight i in α, $\mathcal{B}$ respectively. Of course, $A_0 = B_0 = 1$. Set $\gamma = q - 1$. Let z be an indeterminate.

Theorem 1: The quantities defined above are related by the equation

$$\sum_{i=0}^{n} A_i (1 + \gamma z)^{n-i} (1 - z)^i = q^k \sum_{i=0}^{n} B_i z^i.$$

Remarks:

i. The formula above is symmetric in the sense that, setting $(1 - z)/(1 + \gamma z) = \hat{z}$, we obtain by straightforward algebra

$$\sum_{i=0}^{n} B_i (1 + \gamma \hat{z})^{n-i} (1 - \hat{z})^i = q^m \sum_{i=0}^{n} A_i \hat{z}^i.$$

ii. Theorem 1 is a statement about equivalent classes;[3,4] it is still true if α, $\mathcal{B}$ are replaced by equivalent alphabets.

An alphabet α is said to be decomposable,[3,4] with respect to the basis

ϵ_1, ϵ_2, $\cdots$, ϵ_n of F^n, if it is the direct sum of two alphabets α_1, α_2, where α_1 contains n columns of zeros and α_2 occupies these columns only. For example, the alphabet

$$
\begin{matrix}
0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 \\
1 & 1 & 1 & 1
\end{matrix}
$$

is decomposable, with

$$
\alpha_1 = \begin{matrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{matrix} \qquad \alpha_2 = \begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1. \end{matrix}
$$

In general, α_1 is a k_1-dimensional alphabet in F^{n_1}, and α_2 a k_2-dimensional alphabet in F^{n_2}, with $n_1 + n_2 = n$, $k_1 + k_2 = k$, $k_1 \leq n_1$, $k_2 \leq n_2$.

The dual alphabet of a decomposable alphabet is also decomposable; in fact $\mathcal{B} = \mathcal{B}_1 + \mathcal{B}_2$, where $\mathcal{B}_i$ is the dual alphabet of α_i in F^{n_i}, $i = 1, 2$. (The example above is self-dual.)

Corollary 1.1: If α is decomposable, say $\alpha = \alpha_1 + \alpha_2$, the equation

$$
\sum_{i=0}^{n} A_i (1 + \gamma z)^{n-i} (1 - z)^i = q^k \sum B_i z^i
$$

is reducible in the obvious sense; the factors are the equations pertaining to α_i, $\mathcal{B}_i$ in F^{n_i}, $i = 1, 2$.

For the example above we have

$$
[(1 + z)^4 + (1 - z)^2]^2 = 2^2 (1 + z^2)^2.
$$

Corollary 1.2: A necessary condition for the existence of an alphabet containing letters of weights w_i, $i = 1, 2, \cdots, s$, and no other, is that there exists a set of integers α_i, $i = 1, 2, \cdots, s$, such that the expression

$$
(1 + \gamma z)^n + \gamma \sum_{j=1}^{s} \alpha_i (1 + \gamma z)^{n-w_i} (1 - z)^{w_i},
$$

when expanded in powers of z, takes the form

$$
q^k + \gamma q^k \sum_{i=1}^{n} \beta_i z^i,
$$

where the β_i are positive integers.

Unfortunately, this condition is not sufficient. For example,

$$
(1 + z)^8 + 7(1 + z)^6 (1 - z)^2 + 7(1 + z)^2 (1 - z)^6 + (1 - z)^8
$$
$$
= 2^4 (1 + 7z^2 + 7z^6 + z^8),
$$

but it is not possible to construct a binary alphabet containing 7 letters of weight 2 and no letters of weight 4.

If $A_1 = A_2 = \cdots = A_{2j} = 0$, every vector of weight $\leq j$ in F^n appears as a coset leader for α, and conversely. Another way of saying this is that, for all pairs of distinct letters a, a', of α and any $i \leq j$, the set of vectors at distance i from a is disjoint from the set of vectors at distance i from a'. In this case we can enumerate these vectors by weights as follows:

Let $f_{s,i}$ denote the number of vectors of weight s in F^n which are at distance i from some letter of α. Write

$$(1 + \gamma z)^{n-i}(1 - z)^i = \sum_{i=0}^{n} \Psi(i, j)z^j$$

Corollary 1.3: If α contains no letter of weight $< 2j + 1$, then

$$\sum_{s=1}^{n} f_{s,j}x^s = \sum_{i=0}^{n} B_i \Psi(i, j)(1 + \gamma x)^{n-i}(1 - x)^i.$$

The proofs of corollaries 1.1 to 1.3 are given in Section II.

II. PROOFS

If α is an alphabet of F^n and $\mathfrak{B}$ the orthogonal complement of α, the weights of the letters of $\mathfrak{B}$ are, of course, uniquely determined by the letters of A. However, a much stronger statement can be made: the set of integers specifying the number of letters of each weight in $\mathfrak{B}$ is related by a system of linear equations to the set of integers similarly defined for α. This section will consist of proofs of this statement and of some of its consequences.

Two proofs are given. The first is short and easy; the second is longer and more sophisticated. However, it incidentally produces a more general result and gives some insight into what is going on.

We make the following conventions for notation: α shall be a k-dimensional alphabet in F^n; $\mathfrak{B}$ shall be the orthogonal complement of α of dimension $m = n - k$; γ shall denote the quantity $q - 1$. A_i, B_i denote the number of letters of weight i in α, $\mathfrak{B}$ respectively. The binomial coefficient $\binom{r}{s}$ is understood to be zero if $s > r$.

Let $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ be the usual basis of F^n. Let $s = (s_1, s_2, \cdots, s_\nu)$ be a set of ν different indices, $1 \leq s_i \leq n$, and let $t = (t_1, t_2, \cdots, t_{n-\nu})$ be the complementary set of indices. Denote by F_s^ν, $F_t^{n-\nu}$ the spaces generated by $\epsilon_{s_1}, \cdots, \epsilon_{s_\nu}$ and $\epsilon_{t_1}, \cdots, \epsilon_{t_{n-\nu}}$. Clearly, F_s^ν, $F_t^{n-\nu}$ are orthogonal complements in F^n. Let $|H|$ stand for the number of vectors in a space H.

Lemma 2.0:

$$| \, \alpha \cap F_t^{n-\nu} | = q^{k-\nu} | \, \mathfrak{B} \cap F_s^{\nu} |.$$

Proof: The orthogonal complement of $\alpha \cap F_t^{n-\nu}$ is the smallest space containing $\mathfrak{B}$ and F_s^{ν}. This is the lattice theoretic union of $\mathfrak{B}$ and F_s^{ν}, which we write $\mathfrak{B} \cup F_s^{\nu}$. Then

$$| \, \mathfrak{B} \cup F_s^{\nu} | \cdot | \, \alpha \cap F_t^{n-\nu} | = q^n = q^{m+k}.$$

The number of vectors in $\mathfrak{B} \cup F_s^{\nu}$ is $q^m q^{\nu}/| \, \mathfrak{B} \cap F_s^{\nu} |$.
Hence

$$q^{m+\nu} | \, \alpha \cap F_t^{n-\nu} | = q^{m+k} | \, \mathfrak{B} \cap F_s^{\nu} |$$

or

$$| \, \alpha \cap F_t^{n-\nu} | = q^{k-\nu} | \, \mathfrak{B} \cap F_s^{\nu} |.$$

Denote by $\{(\epsilon_{s_1}, \cdots, \epsilon_{s_\nu}), a\}$ a pair consisting of ν basis vectors of F^n and a vector a of α which is orthogonal to each of $\epsilon_{s_1}, \cdots, \epsilon_{s_\nu}$.

Lemma 2.1:

i. For a fixed set of indices $s_1, \cdots, s_\nu$ the number of pairs

$$\{(\epsilon_{s_1}, \cdots, \epsilon_{s_\nu}), a\} \quad is \quad | \, \alpha \cap F_t^{n-\nu} |.$$

ii. The total number of such pairs for all choices of ν distinct basis vectors is $\sum_{i=0}^n A_i \binom{n-i}{\nu}$.

Proof:

i. $F_t^{n-\nu}$ consists of exactly those vectors of F^n which are orthogonal to $\epsilon_{s_1}, \cdots, \epsilon_{s_\nu}$; hence $\alpha \cap F_t^{n-\nu}$ consists of exactly those vectors of α which are orthogonal to $\epsilon_{s_1}, \cdots, \epsilon_{s_\nu}$.

ii. If $a \, \epsilon \, \alpha$ is of weight i, then a is orthogonal to $n - i$ of the vectors $\epsilon_1, \cdots, \epsilon_n$. A set of ν vectors may be chosen from these $n - i$ in $\binom{n-i}{\nu}$ ways. Hence the total number of pairs

$$\{(\epsilon_{s_1}, \cdots, \epsilon_{s_\nu}), \quad a\} \text{ is } \sum_{i=0}^n A_i \binom{n-i}{\nu}.$$

Let $\sum_s$ indicate summation over all possible choices of ν indices $s_1, \cdots, s_\nu$; similarly, $\sum_t$ denotes summation over all the complementary sets $t_1, \cdots, t_{n-\nu}$. Lemma 2.1 is equivalent to

$$\sum_t | \, \alpha \cap F_t^{n-\nu} | = \sum_{i=0}^n A_i \binom{n-i}{\nu}.$$

Replace α by $\mathfrak{B}$, ν by $n - \nu$ and s by t. The same argument then gives

$$\sum_s |\mathfrak{B} \cap F_s^{\nu}| = \sum_{i=0}^{n} B_i \binom{n - i}{n - \nu}.$$

Lemma 2.2

$$\sum_{i=0}^{n} A_i \binom{n - i}{\nu} = q^{k-\nu} \sum_{i=0}^{n} B_i \binom{n - i}{n - \nu}.$$

Proof: For a fixed set s (which determines, of course, a fixed set t) we have by 2.0

$$|\alpha \cap F_t^{n-\nu}| = q^{k-\nu} |\mathfrak{B} \cap F_s^{\nu}|.$$

Thus

$$\sum_t |\alpha \cap F_t^{n-\nu}| = q^{k-\nu} \sum_s |\mathfrak{B} \cap F_s^{\nu}|,$$

which, by 2.1 is the same thing as

$$\sum_{i=0}^{n} A_i \binom{n - i}{\nu} = q^{k-\nu} \sum_{i=0}^{n} B_i \binom{n - i}{n - \nu}.$$

The equation of 2.2 holds for $\nu = 0, 1, \cdots, n - 1$. This is, in fact, one form of the promised set of linear equations between the quantities A_i, B_i.

We now give the second proof.

Let $\mathcal{G}$ be a finite Abelian group. A character χ of $\mathcal{G}$ is a homomorphism of $\mathcal{G}$ into the multiplicative group of complex numbers of absolute value 1. The characters of $\mathcal{G}$ form a group $\mathcal{G}^*$ which is isomorphic to $\mathcal{G}$, there being in general no canonical isomorphism.†

If α is a subgroup of $\mathcal{G}$, the characters such that $\chi(a) = 1$ for all a of α form a subgroup $\mathfrak{B}^*$ of $\mathcal{G}^*$. $\mathfrak{B}^*$ is precisely the character group of $\mathcal{G}$ *mod* α.

Suppose now that $\mathcal{G}$ is the additive group of a finite field. $\mathcal{G}^*$ is just a multiplicative copy of $\mathcal{G}$, and the characters can be labeled by the elements of $\mathcal{G}$ in a symmetric way; that is, if $r, s, \cdots$ are elements of $\mathcal{G}$ we have

$$\chi_r(s) = \chi_s(r) = \chi(r, s).$$

If $\mathcal{G}$ is the additive group of a prime field of order q, we take r, s etc.

---

† For prime fields, the proof can be given without mentioning the word character. The presentation here is an uneasy compromise with conscience — we wish to indicate possible extensions to nonprime fields without doing too much work.

to be the integers $mod\ q$, and set $\chi(r, s) = \zeta^{rs}$ where ζ is a primitive qth root of unity.

We have from the general theory of characters

$$\chi(r, 0) = \chi(0, s) = 1,$$

$$\sum_{r=1}^{\gamma} \chi(r, s) = -1 \quad \text{if } s \neq 0.$$

Let $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ be the fixed basis of F^n, and $u = (u_1, u_2, \cdots, u_n)$ the coordinates of a vector of F^n with respect to this basis. The character group of F^n is, of course, a multiplicative copy of F^n. We label the characters by elements of F^n as follows:

$$\psi_u(v) = \prod_{i=1}^{n} \chi(u_i, v_i) = \psi_v(u) = \psi(u, v).$$

Let $\mathcal{C}$ be a subspace of F^n. The characters such that $\psi_b(a) = 1$ for all a of $\mathcal{C}$ form a subgroup $\mathcal{B}^*$ of the character group. $\mathcal{B}^*$ is exactly the character group of F^n mod $\mathcal{C}$. The elements b which label these characters form a subspace $\mathcal{B}$ of F^n, isomorphic to F^n mod $\mathcal{C}$. In our notation, the equation $\psi(a, b) = 1$ holds for all a of $\mathcal{C}$ and all b of $\mathcal{B}$, and given either $\mathcal{C}$ or $\mathcal{B}$, the other is uniquely† determined by this condition.

Lemma 2.3: Let $\mathcal{C}$, $\mathcal{B}$ *be related as above. Then*

i. $$\sum_{a \, \epsilon \, \mathcal{C}} \psi(v, a) = q^k \text{ if } v \, \epsilon \, \mathcal{B}.$$

ii. $$\sum_{a \, \epsilon \, \mathcal{C}} \psi(v, a) = 0 \text{ if } v \notin \mathcal{B}.$$

Proof: Part i is obvious, since by definition $\psi(v, a) = 1$ if $v \, \epsilon \, B$. For ii we observe that for $a \, \epsilon \, \mathcal{C}$, $\psi(v, a) = \psi_a(v)$ is a character of F^n mod $\mathcal{B}$. If $\bar{v}$ denotes a coset of F^n mod $\mathcal{B}$, $\sum_{a \epsilon \mathcal{C}} \psi_a(\bar{v}) = 0$ for $\bar{v} \neq \mathcal{B}$. Now $\psi_a(v) = \psi_a(\bar{v})$ for any v in $\bar{v}$; hence

$$\sum_{a \, \epsilon \, \mathcal{C}} \psi(v, a) = \sum_{a \, \epsilon \, \mathcal{C}} \psi_a(\bar{v}) = 0 \quad \text{if } v \notin \mathcal{B}.$$

Lemma 2.4: If F is a prime field, then $\mathcal{C}$, $\mathcal{B}$ are related as in 2.3 if and only if they are orthogonal complements.

Proof: $\psi(a, b) = \zeta^{\sum a_i b_i}$ where ζ is a primitive qth root of unity. Hence $\psi(a, b) = 1$ implies that a is orthogonal to b. Since $\mathcal{C}$ is isomorphic to F^n mod $\mathcal{B}$ the dimensions of $\mathcal{C}$, $\mathcal{B}$ add up to n. Thus $\mathcal{B}$ is the orthogonal complement of $\mathcal{C}$.

---

† That is, if F is a prime field. Otherwise we must fix the basis of F over its prime field before we claim uniqueness.

Let $f(i, s)$ denote a function of the integers i, s with values in a ring R. The values of $f(i, s)$ may be added and multiplied, and these operations obey the two distributive laws. $f(i, v_i)$ denotes the same function of i and the ith coordinate of v.

Lemma 2.5:

$$\sum_{v \, \epsilon \, F^n} \prod_{i=1}^{n} f(i, v_i) = \prod_{i=1}^{n} \sum_{r=0}^{\gamma} f(i, r).$$

Proof: If $n = 1$ the statement is

$$\sum_{s=0}^{\gamma} f(1, s) = \sum_{r=0}^{\gamma} f(1, r),$$

which is obvious. Assume the truth of the lemma for F^{n-1}.

Let F_r^n, $0 \leq r \leq \gamma$, denote the set of vectors of F^n which have last coordinate r. Clearly the F_r^n are a partition of F^n. Then

$$\sum_{v \, \epsilon \, F^n} \prod_{i=1}^{n} f(i, v_i) = \sum_{r=0}^{\gamma} \sum_{v \, \epsilon \, F_r^n} \left[\prod_{i=1}^{n-1} f(i, v_i) f(n, r) \right]$$

$$= \sum_{r=0}^{\gamma} f(n, r) \sum_{v \, \epsilon \, F^{n-1}} \prod_{i=1}^{n-1} f(i, v_i)$$

$$= \sum_{r=0}^{\gamma} f(n, r) \prod_{i=1}^{n-1} \sum_{r=0}^{\gamma} f(i, r) \quad \text{(by induction)}$$

$$= \prod_{i=1}^{n} \sum_{r=0}^{\gamma} f(i, r).$$

Let $z^{(r)}$ be a set of (commuting) indeterminates, $r = 0, 1, \cdots, \gamma$. To each vector $v = (v_1, v_2, \cdots, v_n)$ of F^n associate a monomial $\prod_{i=1}^{n} z^{(v_i)}$. The monomial associated with v describes how many times each field element appears as a component of v. Let R be the ring of polynomials in $z^{(0)}, z^{(1)}, \cdots, z^{(\gamma)}$ over the complex numbers. Let $u = (u_1, u_2, \cdots, u_n)$ be a fixed vector of F^n.

Lemma 2.6:

$$\sum_{v \, \epsilon \, F^n} \psi(u, v) z^{(v_1)} z^{(v_2)} \cdots z^{(v_n)} = \prod_{j=1}^{n} \sum_{r=0}^{\gamma} \chi(u_j, r) z^{(r)}.$$

Proof: Set $f(j, v_j) = \chi(u_j, v_m) z^{(v_j)}$, which is in R. Then

$$\psi(u, v) z^{(v_1)} z^{(v_2)} \cdots z^{(v_n)} = \prod_{j=1}^{n} f(j, v_j).$$

By 2.5,

$$\sum_{v \,\epsilon\, F^n} \psi(u, v) z^{(v_1)} z^{(v_2)} \cdots z^{(v_n)} = \prod_{j=1}^{n} \sum_{r=0}^{\gamma} f(j, r)$$

$$= \prod_{j=1}^{n} \sum_{r=0}^{\gamma} \chi(u_j, r) z^{(r)}.$$

Lemma 2.7: Let $\mathfrak{a}$, $\mathfrak{B}$ be orthogonal complements in F^n, as usual. Then

$$\sum_{u \,\epsilon\, \mathfrak{a}} \prod_{j=1}^{n} \sum_{r=0}^{\gamma} \chi(u_j, r) z^{(r)} = q^k \sum_{v \,\epsilon\, \mathfrak{B}} z^{(v_1)} z^{(v_2)} \cdots z^{(v_n)}$$

Proof: We evaluate the quantity

$$F(u, v) = \sum_{u \,\epsilon\, \mathfrak{a}} \sum_{v \,\epsilon\, F^n} \chi(u, v) z^{(v_1)} z^{(v_2)} \cdots z^{(v_n)}$$

in two ways, which give the two sides of the equation.
By 2.6

$$F(u, v) = \sum_{u \,\epsilon\, \mathfrak{a}} \prod_{j=1}^{n} \sum_{r=0}^{\gamma} \chi(u_j, r) z^{(r)}.$$

Also

$$F(u, v) = \sum_{v \,\epsilon\, F^n} z^{(v_1)} z^{(v_2)} \cdots z^{(v_n)} \sum_{u \,\epsilon\, \mathfrak{a}} \psi(u, v).$$

By 2.4 and 2.3

$$\sum_{u \,\epsilon\, \mathfrak{a}} \psi(u, v) = \begin{cases} q^k & \text{if } v \,\epsilon\, \mathfrak{B} \\ 0 & \text{if } v \,\not\epsilon\, \mathfrak{B}. \end{cases}$$

Hence

$$F(u, v) = q^k \sum_{v \,\epsilon\, \mathfrak{B}} z^{(v_1)} z^{(v_2)} \cdots z^{(v_n)}.$$

Theorem 2.8: Let $\mathfrak{a}$ be a k-dimensional alphabet of F^n, and $\mathfrak{B}$ the orthogonal complement of dimension $m = n - k$. Let A_i, B_i denote the number of letters of weight i in $\mathfrak{a}$, $\mathfrak{B}$. Then

$$\sum_{i=0}^{n} A_i (1 + \gamma z)^{n-i} (1 - z)^i = q^k \sum_{i=0}^{n} B_i z^i.$$

Proof: In 2.7 set $z^{(r)} = \begin{cases} z & \text{if } r \neq 0 \\ 1 & \text{if } r = 0. \end{cases}$

If $u_j = 0$, $\chi(u_j, r)$ is 1 and $\sum_{r=0}^{\gamma} \chi(u_j, r) z^{(r)}$ becomes $(1 + \gamma z)$.

If $u_j \neq 0$ $\sum_{r=1}^{\gamma} \chi(u_j, r)$ is -1, and $\sum_{r=0}^{\gamma} \chi(u_j, r) z^{(r)}$ becomes $(1 - z)$.

Let $|u|$ denote the number of nonzero u_j.

Then $\prod_{j=1}^{n} \sum_{r=0}^{\gamma} \chi(u_j, r)z^{(r)}$ goes into $(1 + \gamma z)^{n-|u|}(1 - z)^{|u|}$;

$|u|$ is of course the weight of $u = (u_1, u_2, \cdots, u_n)$, so that the left-hand side of 2.7 becomes

$$\sum_{i=0}^{n} A_i(1 + \gamma z)^{n-i}(1 - z)^i.$$

The right-hand side of 2.7 is clearly

$$q^k \sum_{i=0}^{n} B_i z^i,$$

which proves the theorem.

Innumerable sets of linear equations between the quantities A_i, B_i may be obtained from theorem 2.8. The following two are sometimes useful.

Lemma 2.9: For $\nu = 0, 1, \cdots, n$,

i. $$\sum_{i=0}^{n-\nu} A_i \binom{n-i}{\nu} = q^{k-\nu} \sum_{i=0}^{n} B_i \binom{n-i}{n-\nu}.$$

(*These are the equations of 2.2.*)

ii. $$\sum_{i=\nu}^{n} A_i \binom{i}{\nu} = q^{k-\nu} \sum_{i=0}^{\nu} (-1)^i B_i \gamma^{\nu-i} \binom{n-i}{n-\nu}.$$

i. is obtained by setting $(1 + \gamma z)/(1 - z) = 1 + y$,

ii. by setting $(1 - z)/(1 + \gamma z) = 1 + y$. The algebraic details are easy to verify.

This process is reversible, i.e., (i) or (ii) imply 2.8. Before exploring the consequences of theorem 2.8, we give a different specialization of 2.7.

Theorem 2.10: Let $B_s^{(1)}$ be the number of letters of $\mathcal{B}$ which contain s coordinates equal to 1. Let A_{0s} be the number of letters u in $\mathcal{C}$ of weight s for which $\sum_{i=1}^{n} u_i = 0$. Let A_{1s} be the number of letters u in $\mathcal{C}$ of weight s for which $\sum_{i=1}^{n} u_i \neq 0$. (Clearly $A_s = A_{0s} + A_{1s}$). Then

$$\sum_{s=0}^{n} B_s^{(1)} z^s = (A_{0s} - A_{1s}/\gamma)(z - 1)^s(z + \gamma)^{n-s}.$$

Proof: In 2.7 set

$$z^{(r)} = \begin{cases} z & \text{if } r = 1 \\ 1 & \text{if } r \neq 1. \end{cases}$$

Then

$$q^k \sum_{v \,\epsilon\, \mathfrak{B}} z^{(v_1)} z^{(v_2)} \cdots z^{(v_n)} \text{ becomes } q^k \sum_{s=0}^{n} B_s^{(1)} z^s,$$

$$\sum_{r=0}^{\gamma} \chi(u_i, r) z^{(r)} \text{ becomes } \chi(u_i, 1)z + \sum_{r=0}^{\gamma} \chi(u_i, r) - \chi(u_i, 1)$$

$$= \chi(u_i, 1)(z - 1) + \sum_{r=0}^{\gamma} \chi(u_i, r)$$

$$= \begin{cases} \chi(u_i, 1)(z - 1) & \text{if } u_i \neq 0 \\ z + \gamma & \text{if } u_i = 0 \end{cases}$$

$$\prod_{i=1}^{n} \sum_{r=0}^{\gamma} \chi(u_i, r) z^{(r)} \text{ becomes } (z + \gamma)^{n-|u|}(z - 1)^{|u|} \prod_{u_i \neq 0} \chi(u_i, 1).$$

Now if u is a letter of A, so are also the letters $2u, \cdots, \gamma u$, and these have the same weight as u. We sum first over these letters

$$\sum_{s=1}^{\gamma} \prod \chi(su_i, 1) = \sum_{s=1}^{\gamma} \chi(s\Sigma u_i, 1)$$

$$= \begin{cases} -1 & \text{if } \Sigma u_i \neq 0 \\ \gamma & \text{if } \Sigma u_i = 0 \end{cases}$$

The sum of

$$(z + \gamma)^{n-|u|}(z - 1)^{|u|} \prod_{u_i \neq 0} \chi(u_i, 1)$$

over all letters in $\mathfrak{A}$ of the same weight as u is thus

$$(A_{0|u|} - (A_{1|u|}/\gamma))(z + \gamma)^{n-|u|}(z - 1)^{|u|}.$$

Hence the left-hand side of 2.7 becomes

$$\sum_{s=0}^{n} (A_{0s} - (A_{1s}/\gamma))(z + \gamma)^{n-s}(z - 1)^{s}.$$

We return now to the consequences of theorem 2.8. As remarked in the proof of 2.10, if $\mathfrak{A}$ contains a letter u, it contains also the letters $2u, \cdots, \gamma u$; that is, the number of letters of weight i in $\mathfrak{A}$ is divisible by γ for $i > 0$. We have then

Lemma 2.11: A necessary condition for the existence of an alphabet containing letters of weights w_i, $i = 1, 2, \cdots, s$, and no other, is the existence of a set of integers α_i, $i = 1, 2, \cdots, s$, such that the expression

$$(1 + \gamma z)^n + \gamma \sum_{j=1}^{s} \alpha_i (1 + \gamma z)^{n-w_i} (1 - z)^{w_i},$$

when expanded in powers of z, takes the form

$$q^k + \gamma q^k \sum_{i=1}^{n} \beta_i z^i,$$

where the β_i are positive integers.

It has been pointed out before that this condition is not sufficient.

Suppose now that $\alpha = \alpha_1 + \alpha_2$ is a decomposable alphabet. α_j is a k_j-dimensional alphabet in F^{n_i}, $j = 1, 2$, with orthogonal alphabet B_j. $k_1 + k_2 = k$, and $n_1 + n_2 = n$. Let $A_i^{(1)}$, $A_i^{(2)}$ and $B_i^{(1)}$, $B_i^{(2)}$ be the number of letters of weight i in α_1, α_2, $\mathcal{B}_1$, $\mathcal{B}_2$.

Lemma 2.12:

$$\sum_{i=0}^{n} A_i (1 + \gamma z)^{n-i} (1 - z)^i$$

$$= \left[\sum_{i=0}^{n_1} A_i^{(1)} (1 + \gamma z)^{n_1-i} (1 - z)^i \right] \left[\sum_{i=0}^{n_2} A_i^{(2)} (1 + \gamma z)^{n_2-i} (1 - z)^i \right]$$

$$= \sum_{i=0}^{n} \left[q^{k_1} \sum_{i=0}^{n_1} B_i^{(1)} z^i \right] \left[q^{k_2} \sum_{i=0}^{n_2} B_i^{(2)} z^i \right].$$

Proof: The number of letters of weight s in $\mathcal{B}_1 + \mathcal{B}_2$ is

$$\sum_{\sigma+\rho=s} B_\sigma^{(1)} B_\rho^{(2)},$$

which is the coefficient of z^s in $\sum_{i=0}^{n_1} B_i^{(1)} z^i \sum_{i=0}^{n_2} B_i^{(2)} z^i$. Similarly,

$$\sum_{\sigma+\rho=s} A_\sigma^{(1)} A_\rho^{(2)}$$

is the coefficient of $(1 + \gamma z)^{n-s} (1 - z)^s$ in

$$\left[\sum_{i=0}^{n_1} A_i^{(1)} (1 + \gamma z)^{n_1-i} (1 - z)^i \right] \left[\sum_{i=0}^{n_2} A_i^{(2)} (1 + \gamma z)^{n_2-i} (1 - z)^i \right].$$

We define the coset leader of a coset of α in F^n to be an element of least weight in the coset. The weight of a coset is defined to be the weight of its coset leader.

If $A_i = 0$ for $i = 1, 2, \cdots, 2e$ every vector of weight $\leq e$ in F^n appears as a coset leader for α and conversely. Another way of saying this is: for all pairs of distinct letters a, a' of α, the set of vectors at distance $i \leq e$ from a is disjoint from the set of vectors at distance i from a'.

Let $c_1^{(i)}$, $c_2^{(i)}$, $\cdots$, $c_\nu^{(i)}$ be the cosets of A of weight i; we assume that

$\nu = \gamma\binom{n}{i}$, i.e., that all vectors of weight i appear as coset leaders. Let $f_{s,i}$ be the number of vectors of weight s contained in the set-theoretic union $\bigcup_{j=1}^{\nu} c_j^{(i)}$. The polynomial $\sum_{s=0}^{n} f_{s,i} x^s$ is called the enumerator (by weight) of this set of vectors. We propose to show that theorem 2.8 gives a convenient expression for this enumerator. We need the following preliminary lemma.

Lemma 2.13: Let u be a fixed vector of weight i. Let d_{st} be the number of vectors of weight s which are at distance t from u. Then

$$\sum_{s=0}^{n} \sum_{t=0}^{n} d_{st} x^s y^t = (1 + \gamma xy)^{n-i} [x + y + (\gamma - 1)xy]^i.$$

Proof: Suppose first that $u = (u_1, u_2, \cdots, u_i)$ is a vector of weight i in F^i. We show that under these circumstances

$$\sum_{s=0}^{i} \sum_{t=0}^{i} d_{st} = [x + y + (\gamma - 1)xy]^i.$$

This is obvious for $i = 1$; we suppose it true for $i - 1$. Let $v = (v_1, v_2, \cdots, v_{i-1})$ be a vector of weight s distant t from $(u_1, u_2, \cdots, u_{i-1})$ in F^{i-1}. From v we obtain:

i. One vector $(v_1, v_2, \cdots, v_{i-1}, 0)$, weight s, distant $t + 1$ from u.

ii. One vector $(v_1, v_2, \cdots, v_{i-1}, u_i)$ weight $s + 1$ distant t from u.

iii. $\gamma - 1$ vectors $(v_1, v_2, \cdots, v_{i-1}, v_i)$ $v_i \neq 0$, $v_i \neq u_i$ which have weight $s + 1$, and are distant $t + 1$ from u.

Hence the enumerator for i is obtained by multiplying that for $i - 1$ by $[x + y + (\gamma - 1)xy]$, and the lemma is proved for $n = i$.

We now apply induction to $n - i$. Let u be a vector of weight i in F^n, and u' a vector of weight i in F^{n-1} obtained from u by omitting one zero coordinate. Let v' be a vector of F^{n-1} which has weight s and is distant t from u'. From v' we obtain in F^n

i. One vector of weight s, distant t from u, by adding a zero coordinate to v'.

ii. γ vectors of weight $s + 1$, distant $t + 1$ from u, by adding a non-zero coordinate to v'.

This corresponds to multiplication by $(1 + \gamma xy)$. Hence the lemma is proved.

Lemma 2.14: Suppose that $A_i = 0$, $i = 1, 2, \cdots, 2e$, and take $t \leqq e$. Then the enumerator, $\sum_{s=0}^{n} f_{s,t} x^s$, of vectors in cosets of weight t of F^n mod $\mathfrak{a}$ is the coefficient of y^t in

$$\sum_{i=0}^{n} A_i (1 + \gamma xy)^{n-i} (x + y + (\gamma - 1)xy)^i.$$

Proof: The cosets of weight t in F^n *mod* α are disjoint, and contain all vectors of F^n which are at distance t from some letter of α.

Set $(1 + \gamma z)^{n-i}(1 - z)^i = \sum_{r=0}^{n} \Psi(i, n, r)z^r$. Let $\mathcal{B}$, B_i have their usual meaning. Assume the conditions of 2.14.

Lemma 2.15:†

$$q^m \sum_{s=0}^{n} f_{s,i}x^s = \sum_{i=0}^{n} B_i \Psi(i, n, t)(1 + \gamma x)^{n-i}(1 - x)^i$$

Proof: Set

$$z = \frac{x + y + (\gamma - 1)xy}{1 + \gamma xy},$$

then

$$1 + \gamma z = \frac{(1 + \gamma x)(1 + \gamma y)}{(1 + \gamma xy)}, \qquad 1 - z = \frac{(1 - x)(1 - y)}{1 + \gamma xy}.$$

Make this substitution in the equation

$$\sum_{i=0}^{n} B_i(1 + \gamma z)^{n-i}(1 - z) = q^m \sum_{i=0}^{n} A_i z^i$$

we obtain

$$\sum_{i=0}^{n} B_i(1 + \gamma x)^{n-i}(1 - x)^i(1 + \gamma y)^{n-i}(1 - y)^i$$

$$= q^m \sum_{i=0}^{n} A_i(1 + \gamma xy)^{n-i}(x + y + (\gamma - 1)xy)^i.$$

Equating coefficients of y^t gives us

$$\sum_{i=0}^{n} B_i\Psi(i, n, t)(1 + \gamma x)^{n-i}(1 - x)^i = q^m \sum_{s=0}^{n} f_{s,i}x^s$$

III. APPLICATIONS

The easiest application of theorem 1 is to a generalized Hamming alphabet, that is, a close-packed 1-error correcting alphabet over a field of q elements. Such an alphabet exists for $n = (q^m - 1)/\gamma$, all $m > 1$.[3]

The dual alphabet is of dimension m, and contains $(q^m - 1)$ letters of weight q^{m-1}. The spectrum of a generalized Hamming alphabet is thus given by the expansion of

$$(1 + \gamma z)^n + (q^m - 1)(1 + \gamma z)^{n-u}(1 - z)^u$$

where $n = (q^m - 1)/\gamma$, $u = q^{m-1}$

† A similar formula $(q = 2)$ is found by Lloyd[5] for close-packed codes which are not assumed to be group codes.

TABLE I — DISTRIBUTION OF WEIGHTS IN THE TWO GOLAY CODES

The first table is for the 3-error-correcting $(23, 12)$ alphabet over Z_2, the second for the 2-error-correcting $(11, 6)$ alphabet over Z_3. In both cases, i stands for weight, B_i for the number of letters of weight i in the dual alphabet, A_i for the number of letters of this weight in the Golay alphabet.

i	B_i	A_i
0	1	1
7	0	23×11
8	23×22	23×22
11	0	23×56
12	23×56	23×56
15	0	23×22
16	23×11	23×11
23	0	1

i	B_i	A_i
0	1	1
5	0	2×66
6	2×66	2×66
8	0	2×165
9	2×55	2×55
11	0	2×12

Theorem 1 may, in fact, be used to calculate the number of letters of each weight in any close-packed code. The results for the two Golay[6] codes are given in Table I.

Anything which is known about the structure of an alphabet or its dual may be used with theorem 1 to limit the number of possible weight distributions. Such items of information are a very diversified character, and no general method has been developed. However, the results obtained by hand calculation indicated that it is probably worthwhile to make a systematic computer study of the known classes of alphabets.

ACKNOWLEDGMENTS

REFERENCES

1. Slepian, D., A Class of Binary Signaling Alphabets, B.S.T.J., **35**, January, 1956, pp. 203–234.
2. Bose, R. C., and Kuebler, R. R., Jr., A Geometry of Binary Sequences Associated with Group Alphabets in Information Theory, Ann. Math. Stat. **31**, 1960, p. 113.
3. MacWilliams, F. J., Error-correcting Codes for Multiple-Level Transmission, B.S.T.J., **40**, January, 1961, pp. 281–308.
4. Slepian, D., Some Further Theory of Group Codes, B.S.T.J., **39**, September, 1960, pp. 1219–1252.
5. Lloyd, S. P., Binary Block Coding, B.S.T.J., **36**, March, 1957, pp. 517–535.

Surface Effects of Radiation on Transistors*

By D. S. PECK, R. R. BLAIR, W. L. BROWN and F. M. SMITS

Observation of surface effects of ionizing radiation on several types of transistors indicates that in reverse-biased devices these effects occur at much lower radiation dosage than in unbiased devices or bulk semiconductor material. Further, the total radiation dose rather than dose rate seems often to be the more important factor in the effect. The type of particle used in irradiation is unimportant; the significant factor is the ionization it produces. The effects seem to arise from ionization of gases within a transistor encapsulation and interaction between the ionized gas and residual semiconductor surface contaminants. This results in inversion layers at the device surface and thus in alteration of junction characteristics. The changes in device properties are not permanent, but the recovery after removal from radiation is complex and proceeds with characteristic times between seconds and days.

Different types of devices may respond quite differently to exposure, and the response is different even between different batches and individuals, indicating a dependence upon device processing.

I. INTRODUCTION

A wide variety of effects of high-energy radiation on semiconductor materials and devices has been recognized and studied for a number of years. The major emphasis in this field has been on effects that involve the bulk properties of semiconductors. A great deal of progress has been made in understanding the processes that control bulk radiation phenomena[1] and in understanding the implications of these phenomena for semiconductor devices.[2] Radiation effects on semiconductor surfaces also have been observed,[3] and this paper is concerned with some special aspects of surface phenomena that have recently come to light. In contrast to bulk effects, the surface radiation effects are very poorly under-

stood and in general even poorly characterized. This paper will attempt to shed a little light on these complexities and indicate a type of measurement program that has been found appropriate for dealing with devices intended for use in a radiation environment such as that of the Van Allen belts in space. It will fall short of providing a satisfactory understanding of the processes involved.

The paper is arranged as follows: in Section II, for perspective, a brief discussion of the two broad classes of bulk radiation effects is included. Section III describes the early observations that provoked the present work. Section IV is a proposed model of the basic process. Section V describes results of a number of experiments carried out to test the mechanisms of the process. Section VI characterizes effects with significant numbers of devices. Section VII describes the process of testing and selection undertaken for Telstar devices. Section VIII is a summary of the paper.

II. BULK RADIATION EFFECTS

Bulk radiation effects can be placed in two broad classes that arise from (a) hole-electron pairs produced in the crystal by ionizing radiation and (b) defects in the semiconductor lattice produced by high-energy particles. These phenomena often occur together but they result from quite different interactions of radiation with the solid and they have very different consequences in semiconductors.

2.1 *Hole-Electron Pair Formation*

Fig. 1(a) illustrates the first case. Any charged particle passing through a solid produces ionization through collisions with the bound electrons. These collisions excite electrons to the conduction band and leave holes in the valence band, producing electron-hole pairs in exact analogy with the production of pairs by light. Neutrons and gamma rays also cause ionization effects through intermediate reactions that produce charged particles. As far as effects which depend on ionization are concerned, the particular particle involved is incidental; all that matters is how much energy is lost in the solid. The number of hole-electron pairs produced is proportional to this energy loss.

The generated pairs tend to recombine with a time constant that is the conventional lifetime. Hence, all effects in this class are transient and persist only for the order of a lifetime after the excitation is removed. The pairs produced alter the conductivity of semiconducting materials. They also contribute currents in p-n junction diodes and transistors.

Under pulsed ionizing radiation these effects can alter conductivities and currents by many orders of magnitude. On the other hand these effects can be very small in response to a single energetic particle and special p-n junction diodes may be required even to detect them.[4] A case of intermediate magnitude has been considered by Rosenzweig[5] who has used silicon solar cells to measure the intensity of moderate radiation fields.

2.2 Lattice Damage

Fig. 1(b) illustrates the other type of bulk radiation effect in semiconductors that arises from collisions of energetic particles with the nuclei of the lattice. If such a collision transfers sufficient energy to the struck atom, it is capable of moving the atom from its normal lattice site to some interstitial position in the crystal. These events are rare by comparison with the ionization events of Fig. 1(a), but they create permanent or at least semipermanent defects in the structure of the lattice. The most important consequence of these defects is reduction in the carrier lifetime of the material. Increases in diode reverse current and decreases in transistor current gain are produced as functions of the time integral of the flux of particles. In contrast to the pair production of the paragraph above, this radiation damage is extremely dependent on the particular particle involved. Energetic protons, for example, are much more effective in producing damage than energetic electrons. This type of radiation effect is of major importance to the long-term power conversion efficiency of solar cells in space. Detailed consideration of this problem can be found elsewhere.[6]

III. THE SURFACE PROBLEM

Surface effects on semiconductor devices have an illustrious history of subtlety and perversity and it is no surprise to learn that radiation is an environmental factor that must be considered. Several years ago, before the present sophistication in surface processing, experiments were carried out in an attempt to characterize radiation surface effects[3] (changes in junction current, breakdown voltage, current gain, etc.). No systematic picture evolved, although surface cleanliness seemed certain to play some role.

Transistors that have evolved from refinements in junction formation and surface treatment techniques have, in the last few years, been examined in a variety of high-energy neutron and gamma-ray environments. Particular attention has been given to bulk radiation damage
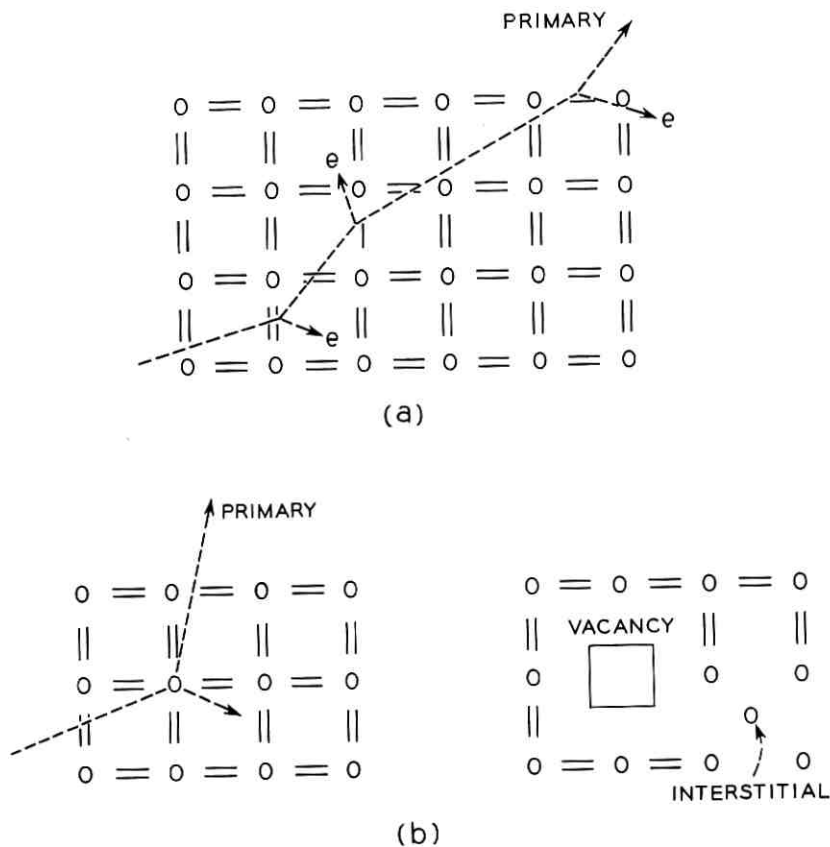
Fig. 1. — (a) Production of hole-electron pairs by collision of a charged particle with electrons of a semiconductor. (b) Production of lattice defects by collision of particles with the nuclei of a semiconductor.

effects on current gain since these effects turned out to be serious (particularly for silicon) at the flux levels of current interest in the vicinity of nuclear reactors. Very substantial improvements in radiation tolerance have been found; however, no recent comprehensive work on the surface effects has been reported.

The relative stability of the characteristics of a particular type of diffused silicon transistor under radiation is illustrated in Fig. 2. The collector reverse current is displayed because it is a particularly sensitive indication of surface stability at the very low currents that are conventional in present silicon devices. The gamma-ray radiation* used in this

* Radiation dose is frequently measured in "rads," which are a measure of the
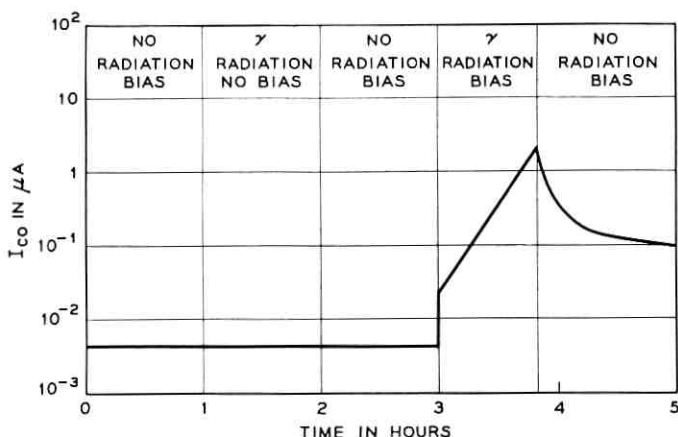
Fig. 2. — The response of I_{CBO} of a diffused silicon transistor to either radiation or bias alone or to both together. Radiation dose rate 8.5×10^5 rads/hr.

case is representative of ionizing radiation in general. The figure shows that after one-hour tests with neither bias nor radiation or with either alone, the current is essentially unchanged. (Actually much longer tests under bias are conventionally made in checking device reliability with the same result as shown here.) However, when bias and gamma radiation are simultaneously applied two effects appear: there is a sudden rise in current and then an upward drift of current over many minutes. The sudden rise of about 0.02 μa is largely due to ionization of the gas in the device can. The current drifts upward by another decade in the 45-minute exposure following its rapid rise. When the gamma radiation is removed the current does not immediately drop back to its pre-irradiation value, but gradually declines and even after an hour is more than a decade too high. A drop of about 0.02 μa must still occur at gamma-ray turnoff, but this is so small compared with the level to which the current has drifted, that the drop is not visible.†

The drift up and the slow decay of the collector currents are surface effects produced by the radiation only when the collector junction is reverse-biased. This aspect of radiation sensitivity had not previously

---

energy absorbed per unit mass of material. One rad is equivalent to absorption of 100 erg/gm. Such a dose would produce approximately 2×10^9 ion-electron pairs/cm³ in atmospheric air or about 4×10^{13} electron-hole pairs/cm³ in silicon.

† Ion currents to the electrical leads either in the device encapsulation or in the gamma radiation chamber pose a serious problem to measurements on devices at very low leakage currents in high radiation fields (10^5-10^6 rad/hr). It has often proved necessary to remove the device from the radiation environment momentarily for measurement. Because of recovery effects care must be taken to obtain measurements quickly and at uniform times after removal.

been reported. It represents a factor that must be considered for reliability of devices that must operate in any radiation environment including the high-energy electron and proton belts found in outer space. The conclusion that these effects arise at the semiconductor surface can be reached in a number of ways. If they were bulk effects the marked influence of applied bias is quite unreasonable. Radiation defects created in a solid have been influenced by extremely high electric fields, but only to the extent of causing them to migrate very short distances in long times.[7] Furthermore, the decay of the effect occurs in a time that is very short for defect annealing in silicon. But more convincing is the sensitivity of the effect to the surface environment of the device inside its encapsulating can.

Fig. 3 shows four typical npn diffused silicon transistors of two types, each with two kinds of ambient atmosphere. Increases in collector current are observed in all four, but at quite different integrated gamma-ray doses. In both device types early current increases are associated with gas filling. In type B the evacuated device shows no measurable change until the integrated dose is in a range expected to cause substantial decreases in bulk lifetime. In this case, the influence of bias, although not shown in the figure, is practically nonexistent.

Not only is there a variation in the response depending on the gas filling of transistor cans, but there is also great variability among devices with a single type of filling. All evacuated units are slow to respond, but in the type B transistor only about half of the gas-filled units respond quickly. The other half are almost as stable as the evacuated
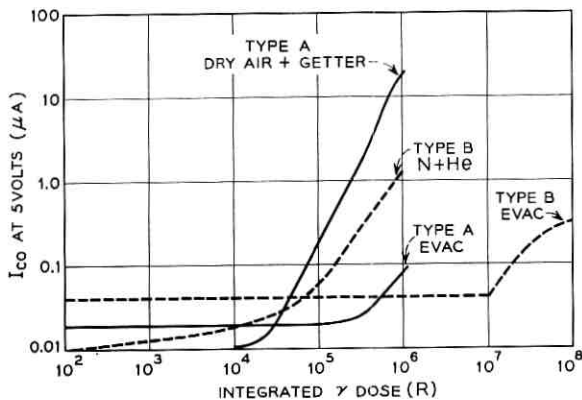

Fig. 3. — The radiation degradation of I_{CBO} of two types of diffused silicon transistors, evacuated or with gas filling. Radiation dose rate 8.5×10^5 rads/hr.

units. This points again to a surface effect and to a broad spread in response arising from rather subtle differences in the surface chemistry. This lack of reproducibility complicates the study of the process involved and necessitates the use of statistical experiments, some of which will be reported in later sections.

Before continuing to discuss the experimental observations we will introduce a simple model which describes some, but not all, of the effects, and provides a framework for the later discussion.

IV. A MODEL OF THE PROCESS

Radiation, gas encapsulation, and device bias that seem to be essential to the effects shown in Section III can be combined in a simple model of the process. Fig. 4 illustrates the ingredients: the fringing field of a reverse-biased collector junction on an npn transistor and ions and electrons produced by gamma radiation in the gas of the encapsulation. The fringing field separates the electron-ion pairs, depositing electrons on the collector side of the junction and positive ions on the surface of the base. On both regions these charges tend to produce inversion layers at the surface, the effects being analogous for pnp and npn devices. For simplicity only the inversion layer on the p-type base of a npn transistor will be considered. A magnified view of the edge of the device might be as shown in Fig. 5. The positive ions induce an electron-rich inversion layer or "channel" on the base and in effect extend the collector region
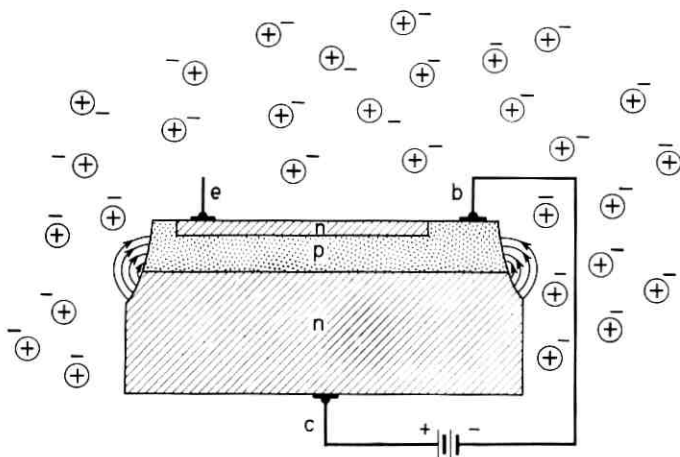

Fig. 4. — A model of a reverse-biased transistor in a gas atmosphere ionized by radiation.
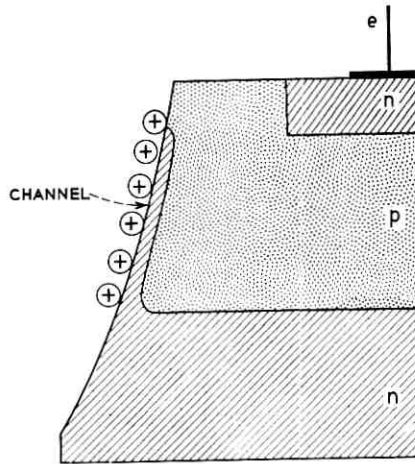
Fig. 5. — Formation of a channel on the base of a transistor by positive ion collection on the surface.

out over the base. The channel represents a grossly different surface than existed before. Since the junction between the electron-rich channel and the base material constitutes an extension of the collector-base junction, it contributes to the collector saturation current. In part this is simply because of the extra junction area. More importantly, because the channel junction is very close to the surface, the surface generation process can yield much more current per unit area than for the junction in the bulk. Furthermore, if the channel extends to the emitter it can add additional current to the collector by serving as a conducting path between the two. Channel effects on transistors and diodes have been studied previously in considerable detail[8] entirely unrelated to the presence of radiation. The basic channel characteristics of emitter-to-collector conductance, high emitter floating potential, and channel pinchoff have all been observed in connection with the present surface radiation effects on diffused silicon transistors.

We have so far considered only formation of a channel by charge collection in the collector fringing field, but a second possibility is shown in Fig. 6. In all the npn diffused silicon transistors used in these experiments the collector is electrically tied to the encapsulating can of the device. Under collector reverse bias a field then exists throughout the whole can of such a sign as to drift positive ions toward the surface of the device base. This feature can be expected to increase the ion-collection efficiency.
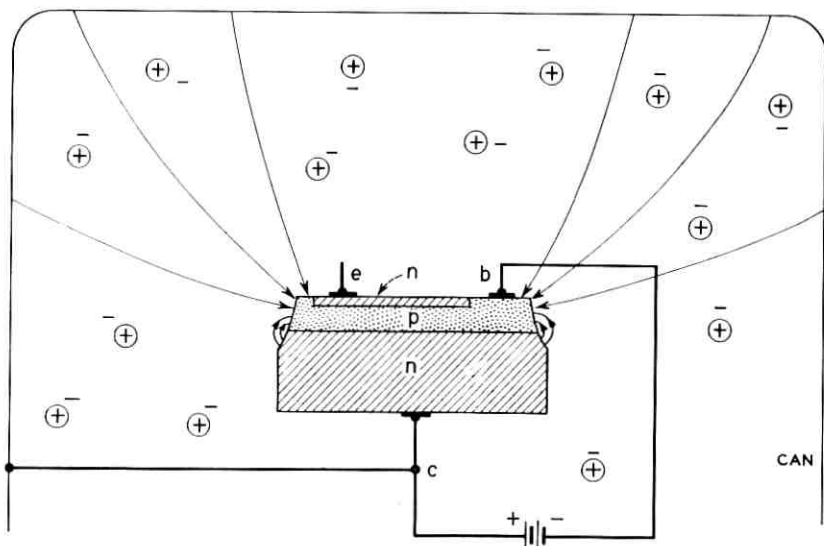
Fig. 6. — The enhancement of ion collection at a transistor surface due to electron fields between the device and its encapsulating case.

One may ask if there are enough ions produced in the gas of the can, by the radiation doses that have been used, to provide large channel effects. At an integrated dose of 10^4 rads (at which these effects may be substantial) and at atmospheric gas pressure (the normal device filling) a total of about 2×10^{13} ions per cc will have been produced. With a device enclosure of about 3×10^{-3} cc and a base layer area of about 10^{-3} cm², if all the ions were collected on the base, their concentration would be about 10^{14}/cm². On typical base material only about 10^{11} surface charges/cm² will be required to produce a channel. These several orders of magnitude margin are probably quite important because the efficiency of the surface charging process seems likely to be quite low. Furthermore, the process is far from a perfect charge integrator as we shall have occasion to observe in connection with reciprocity experiments in Section 5.2.

The lack of reproducibility noted in Section III in connection with devices of the same type and same gas filling seems to necessitate an elaboration on the model. If ions of the gas are sufficient to produce a channel, why are there some devices with gas that are as stable as devices without? The gas ions themselves must not be the tenacious charge on the surface that forms the channel. The gas ions probably

exchange their charge with residual contaminants on the device surface. Ionization of the surface contaminants directly is apparently too rare to be observed, since vacuum-encapsulated devices show uniformly high surface stability. This is roughly reasonable since the probability of ionization of any single atom is estimated to be only about 10^{-5} at 10^4 rads. Even with a monolayer of residual surface contamination, the surface ion concentration would then be only about $10^9/cm^2$. Of course, if the gas ions are to do the job, they must be reasonably effective in finding and exchanging charge with the residual impurities. There seems to be margin for inefficiencies in these very rough numbers.

This model predicts a number of effects that can be tested:

1. The effect depends only on ionization, not on incident particle type.
2. The simplest form of the model suggests that the effect is cumulative and depends on total dose, not on dose rate.
3. The effect should be more pronounced at higher collector bias.
4. The electric field between the can and the semiconductor may influence ion collection.
5. The decay of the effect should be faster if the device is not under continuous bias and should be accelerated in the presence of radiation without bias.

The results of the experimental tests of these predictions are contained in the following section.

V. TESTS OF THE MODEL

5.1 *Ionization*

If ionization in the gas of the device encapsulation is essential, then just as in the case of the ionization effect in a semiconductor discussed in Section 2.1, the type of energetic particle should not matter. The experiments in Section III were carried out with Co^{60} gamma rays which ionize through photo or Compton electrons that they produce. We have tested the model by comparing the results of 18 Mev proton irradiation with the Co^{60} gamma rays. The individual device response scatters so widely that the behavior of a number of similar devices is examined in each case. Fig. 7 shows the collector reverse current versus radiation dose, for gamma rays in dashed lines and protons in solid lines. The dose is calculated simply from the amount of energy deposited in a gas (the gas of the encapsulation) by gamma rays and protons. The radiation intensity (dose per unit time) was approximately 10^6 rads/hr for the protons as well as for the gamma rays.

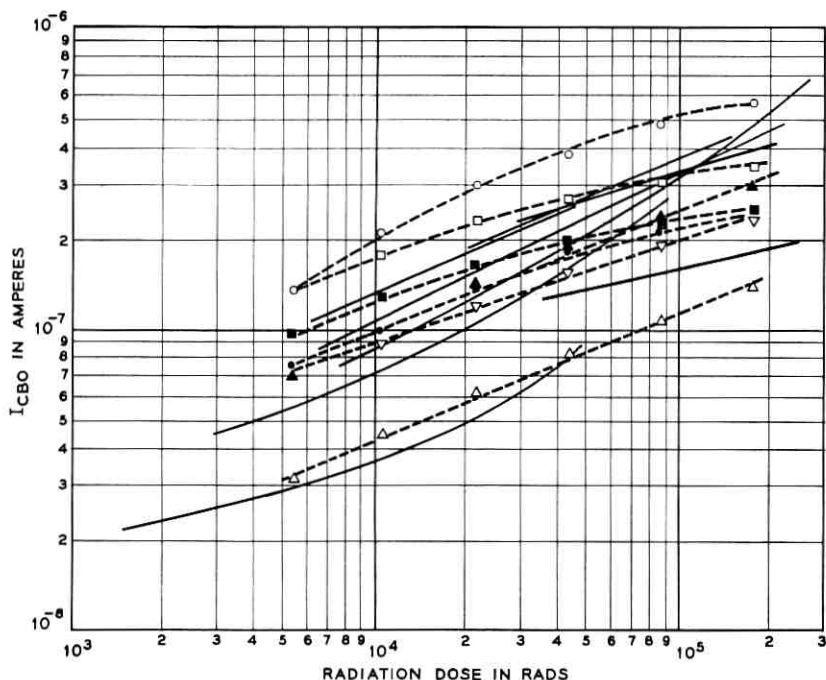Within the spread of response observed, there is no significant differ-

Fig. 7. — Comparison of I_{CBO} changes under radiation by gamma rays (dashed lines) and protons (solid lines).

ence between protons and gamma rays. If bulk damage in silicon were involved, one could expect a factor of 10 to 100 greater effect for the protons than for the gamma rays. (The bulk damage effect per particle comparing 18 Mev protons and 1.25 Mev gamma rays would be a factor of 10^4 to 10^5, but the scale in Fig. 7 is not particles but ionization, or energy loss, and the protons lose energy at a much higher rate than gamma rays.) We conclude that ionization is essential to the mechanism of this surface radiation effect.

5.2 *Reciprocity*

The surface ionization effects may depend only on the total radiation dose or they may also depend on the dose rate. In the first case there is reciprocity between dose rate and time. Information on this point is relevant to understanding the process and is of immediate practical importance as well. The experiments are most easily performed at high dose rates, but the device reliability is also of concern in modest radiation fields.

In Fig. 2 recovery of the collector current following radiation is illustrated. The fact that recovery occurs at all when a device is removed from radiation but kept on bias proves that the observed result does not depend solely on the total dose. The ionized state of the surface at any time is not simply related to the total number of ions that have reached that surface.

Fig. 8 shows the change of collector current vs dose at two different high dose rate levels. The experiment was started at 8×10^5 rads/hr. After approximately half a minute, when a dose of 6×10^3 rads had been delivered, the transistor was placed inside a lead shield that attenuated the gamma radiation by a factor of six. When a dose totaling 1.8×10^4 rads had been given the device, the attenuator was removed and the higher level radiation continued. After each attenuator change, the collector current continues to rise in a smooth extension of the earlier portions of the curve. If the dose rate were important, the middle segment of the curve would be expected to have a different slope than the two ends. During the attenuator changes the device was out of the radiation for approximately a minute and the current had started to recover, as indicated by the first point taken at the lower dose rate. A comparable drop would have appeared at the second attenuator change, but the
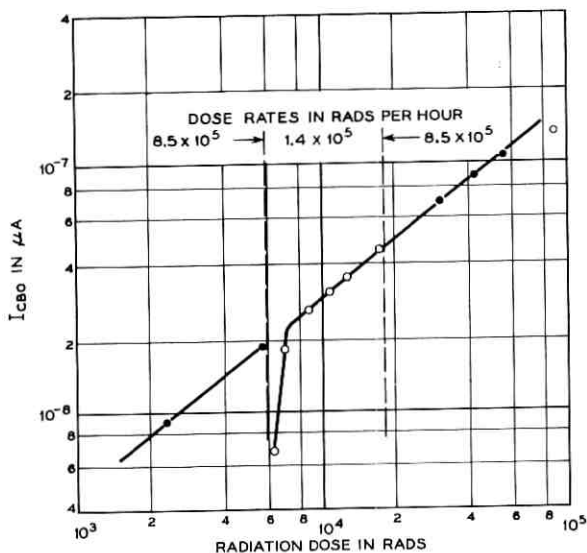

Fig. 8. — Reciprocity of dose rate and time at two high radiation dose rate levels.

earliest measurement after reinsertion in the radiation field was not obtained soon enough for it to show.

The rapid reestablishment of a previous high response to radiation, after some recovery from it, represents a memory in the process. This kind of memory has been seen repeatedly. Devices that were irradiated and have apparently completely recovered their original collector characteristics by standing out of radiation and even off bias, will still tend to reestablish their former response on a second radiation exposure. Even after several weeks, a device seems to retain a sensitivity to subsequent radiation as a result of an earlier exposure. There is some evidence that the memory can be removed by baking at approximately 100° C for a few hours.

The ionization produced by the radiation apparently has two functions: first, to produce some chemical species that are capable of ionization and second, to keep these species in an ionized state. The number of these centers would reflect the total radiation dose, but some minimum level of radiation would be required to keep them active.

We can draw the conclusion from Fig. 8 that, in the high-intensity region, dose is the important variable. Fig. 9 illustrates quite a different
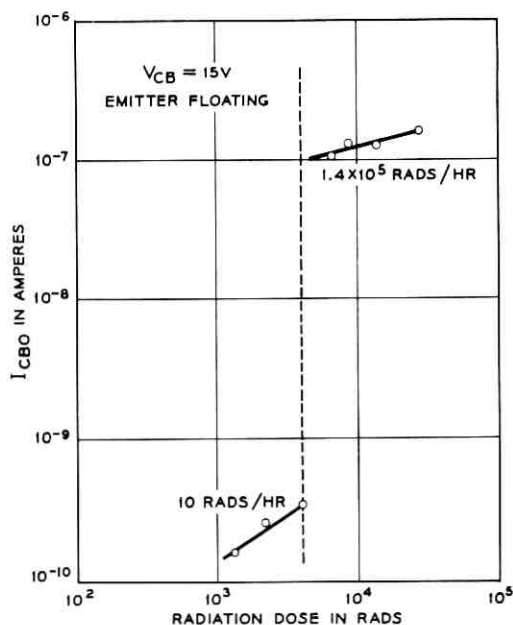

Fig. 9. — Lack of reciprocity between high and low dose rates.

range. Here a device started its radiation history at only 10 rads/hr and established a pattern of current increase that is already determined at a dose less than 10^4 rads. Putting the device into the attenuated high-intensity source does not produce a simple continuation of the earlier curve, but rather produces a curve two and one-half orders of magnitude higher in current. This illustration is extreme and not many devices show this large a discontinuity, but essentially none give results that could be interpreted as simple reciprocity. Apparently over a dose-rate range this great, using the concept of the preceding paragraph, the ionization-sensitive chemical entities are not fully ionized in the low radiation field.

5.3 Surface Effect vs Collector Bias

Since the surface effects we have been considering are absent except under the simultaneous application of radiation and collector reverse bias, we expect to find that changing the bias will alter the effect. This can occur because of increases in efficiency of charge collection, but also because the increased junction field tends to bind the ions more tightly to the surface or distribute them to form a more extensive inversion region.

Fig. 10 shows the effect of a sudden change in bias from 5 to 15 volts. One gains the impression that the current suddenly adjusts to a level and to a rate of change that are what they would have been if the entire dose had been given the device at the higher bias. If this impression is valid it would indicate that the effect at higher bias had not been retarded by the initial dose at low bias, and hence that ion collection was no different. In a plot like Fig. 10 it must be realized that the significance of the first dose rapidly diminishes in comparison with the total as one moves out along the second branch of the curve. It does seem possible to infer that a given number of ions on the surface gives a larger current contribution at higher bias. Rearrangement in the new field and effective extension of the length of a surface channel can occur.

Fig. 11 shows the response of transistors under radiation at different biases. Up to a dose of about 10^4 rads the current increase depends strongly on bias. Between 10^4 and 10^5 rads the slopes of the curves on this log-log plot are quite similar. Although only 4 units are illustrated in Fig. 11, the total group in the experiment was 80 and this pattern of behavior was quite consistent. Even beyond 10^5 rads where the slopes increase there is considerable uniformity between individuals and across the voltage range. This point is illustrated in a different way in Fig. 12 where the median behavior of a group of 10 to 12 devices at each bias is
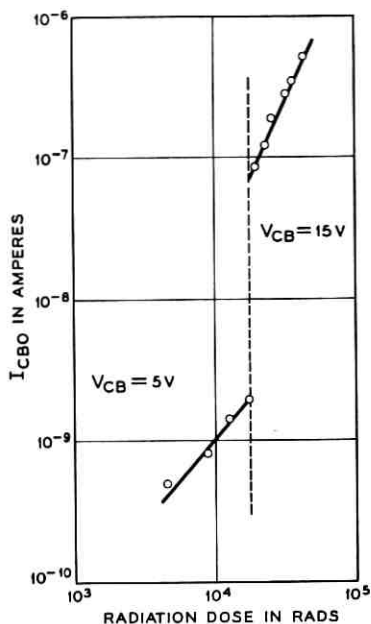
Fig. 10. — Influence of change in collector bias on degradation of I_{CBO}. Dose rate 8.5×10^5 rads/hr.

plotted at 6×10^3 rads and 6×10^4 rads. The pattern of bias dependence is already established at the lower dose and is smoothly maintained after the further order-of-magnitude change in dose. The lump at 10 volts appears to be significant. If it is real, its explanation will require a clever refinement in the model.

We do indeed observe the anticipated increase in device degradation at higher bias. The consistency of the shape of the voltage dependence at doses greater than 10^4 rads strongly suggests that the dependence does not arise only from a difference in the numbers of available surface ions, but also from variation in the arrangement of ions and the influence of this arrangement on the current characteristics of the channel. The collection of charge does not seem to be the dominant variable over the voltage range examined.

Another indication that junction bias affects the arrangement of charge on a device surface is shown in Fig. 13. The characteristics observed after a total dose of about 10^7 rads and at a high dose rate may show considerable structure, depending on the rate at which the characteristic is swept out. Furthermore, the differences between the charac-
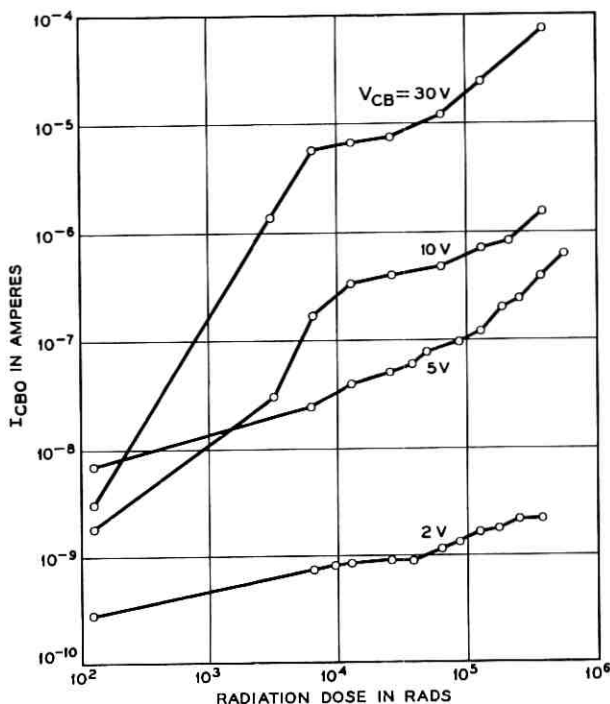
Fig. 11. — Dependence of degradation in I_{CBO} on collector bias.

teristics with slow and fast sweep show that rearrangements are not instantaneous. Note that here also a peak in I_{CBO} occurs in the vicinity of 10 volts.

5.4 *Influence of Can Potential*

It has been suggested that the electric field existing between the can and the transistor base by virtue of the bias on the collector junction may alter the collection of ions at the semiconductor surface. This possibility has been examined using diffused silicon diodes, encapsulated in the same gaseous atmosphere as the diffused silicon transistors. Devices with the p-region common to the can and others with the n-region common were studied. There was no substantial difference in sensitivity to ionization in these cases. On the other hand it was found that by alteration of the device processing, devices of either polarity were insensitive. It is clear that in this case the chemical surface condition is of more vital importance than the can-to-diode bias.
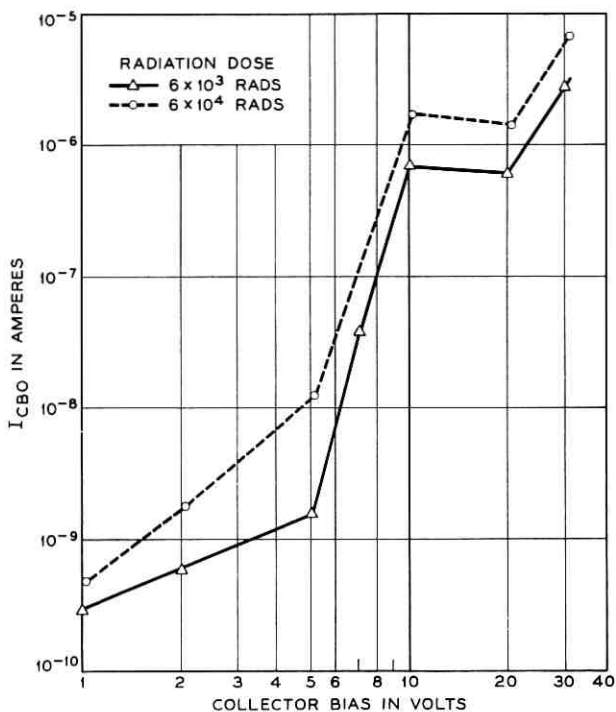
Fig. 12. — Median behavior of I_{CBO} with collector bias at two integrated doses.

No silicon transistor with reversible polarity to the can has been available, but a germanium transistor with all leads insulated from the can has been irradiated with reversal of the sign of can bias. The results are shown in Fig. 14. The device was irradiated at 8×10^5 rads/hr for one minute with the can negative with respect to the transistor base (the collector of which was continuously reverse biased) and then for one minute with the can positive and so on for longer times as the irradiation proceeded. The leakage current is very clearly much more sensitive to positive potential on the can than to negative, a result consistent with the original picture of the role of positive ions in a channel on the transistor base. This experiment has been repeated several times with some lack of reproducibility. In some cases can bias makes much less difference than as shown in Fig. 14 and in two cases where the whole radiation effect was smaller, the dependence on can bias polarity was reversed.

There seems to be no question that the can-to-device potential can be important, but clear-cut evidence that positive ions rather than electrons
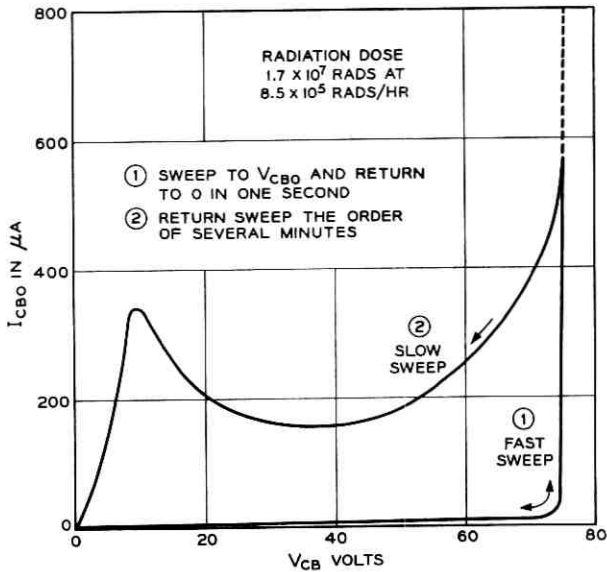
Fig. 13. — Structure and drift in the reverse characteristic of a heavily irradiated transistor under radiation.

(or negative ions) are always the important particle in transistor surface ionization effects has not been demonstrated.

5.5 Recovery

Recovery of the surface effects after an exposure to radiation has already been mentioned in connection with Figs. 2 and 8. The recovery is not exponential. In many cases, recovery under bias is well represented by a straight line on a log I_{CBO} vs log t plot, although the points at less than 0.5 min tend to fall below such a curve. The recovery represents a loss of charge at the surface, perhaps by neutralization from the interior in the absence of a radiation flux, but the memory effects suggest the active species do not actually leave the surface.

The comparison between recovery with and without bias is shown in Fig. 15. A transistor given a dose of 4.6×10^3 rads at 15 volts is shown, first recovering at this same bias and then recovering with the collector circuit open except momentarily for measurement. The decay is substantially accelerated in the zero-bias condition. This is consistent with the idea that the ions are bound in place by the field but in its absence are free to diffuse on the surface and will no longer be concentrated to produce an inversion layer It has also been observed that with reestab-
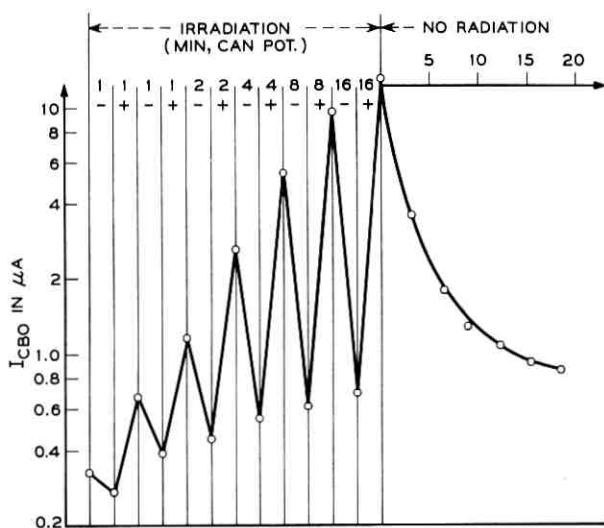
Fig. 14. — The influence of can-to-semiconductor potential on degradation in I_{CBO} of a germanium transistor.

lishment of the bias the current rises before continuing its decay, suggesting that the ions are subject to recapture by the field. There is a loss in this process however, that may arise from the enhanced neutralization of the ions by electrons from the inversion layer when they are more numerous in the absence of reverse bias.

Fig. 16 illustrates the effect of radiation on recovery. The upper curve was taken first and is a normal recovery with bias. The lower curve shows recovery from the same starting value in the 8×10^5 rads/hr gamma field without bias. The device is momentarily removed from
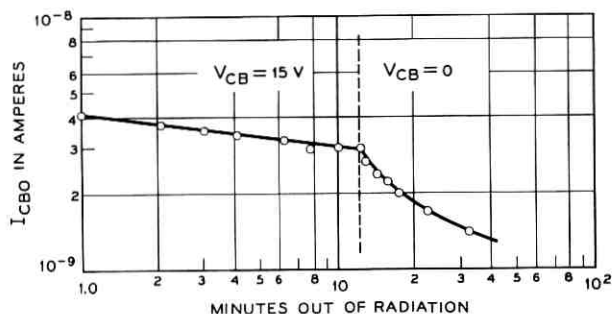

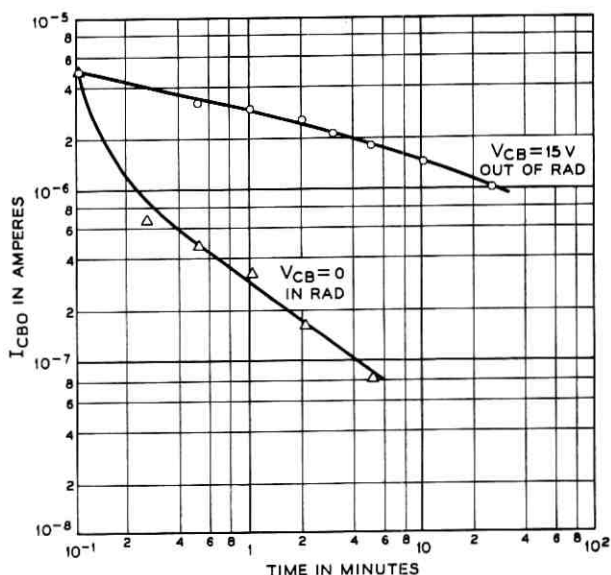Fig. 15. — Decay of the radiation effect with and without bias.

Fig. 16. — Enhancement of the rate of recovery of I_{CBO} with radiation but without bias.

radiation for measurement at 15 volts reverse bias. The recovery is substantially enhanced by the ionization in the absence of bias. This seems to be explained through neutralization of the charges on the semiconductor surface by the newly formed ions. Under bias the gas ions are directed by the junction field to strike the surface in places where they add to an inversion layer. In the absence of a field from applied bias, these ions tend to go wherever they can reduce fields produced by surface charges.

VI. CHARACTERIZATION OF EFFECTS WITH SIGNIFICANT NUMBERS OF DEVICES

With the above background concerning the surface effects of ionizing radiation, and with recognition of the limited knowledge of the range of exact surface conditions existing in semiconductor devices, further tests were planned to establish the extent of the ionization effect in larger samples of devices. For the purpose of such a test program, a transistor type was selected which showed particular sensitivity to the combined effects of electrical bias and ionizing radiation. These transistors normally have quite low collector reverse currents, I_{CBO} (in the order of

10^{-10} amperes). They are therefore good subjects for study at relatively low gamma dose levels since small increases in I_{CBO} can readily be recognized. A large number of devices were available with sufficient power aging to indicate stability of characteristics, so that changes could clearly be attributed to the radiation exposure.

Although the measurements of gain of these transistors under gamma radiation indicate some change with dose, the changes are relatively smaller than those in the junction reverse current and are hence less subject to recognition of a specific pattern of change. The data presented here, therefore, relate specifically to the changes in I_{CBO} .

6.1 *Typical Pattern of Degradation with Dose*

In order to determine the proper conditions for large scale evaluation, and to have some estimate of changes to be expected, it is desirable to establish the pattern of change in characteristics. From the data presented in Section V on the various factors affecting degradation, there develops a typical pattern of degradation, at least for the type of diffused silicon transistor used, as shown in Fig. 17. In this log-log plot of I_{CBO} vs integrated dose, four recognizable regions are indicated, an initial region of stability followed by three regions of change. The boundaries of the regions as indicated in Fig. 17 depict only representative variations and do not indicate actual limits of the device type.
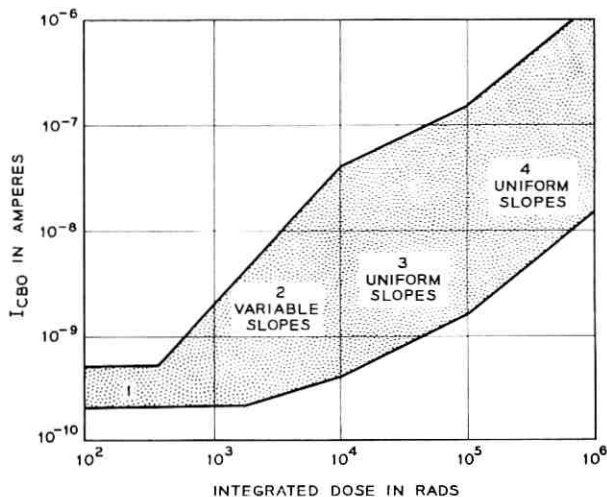

Fig. 17. — The typical pattern of I_{CBO} degradation under radiation for a particular type of diffused silicon transistor.

As discussed in Section 5.3, the slopes of the degrading I_{CBO}'s become quite uniform among all units of this type beyond a total dose of about 10^4 rad, or beyond the variable Region 2. In Region 3 all units show slopes of approximately $\frac{1}{2}$, and at about 10^5 rads there is a fairly distinct increase in slope which distinguishes Region 4. Although these two regions are recognizable, they are of lesser importance in judging comparative usefulness of the devices in radiation than are the region of stability and the slope in Region 2.

The region of stability is indicated even in the early tests (see Fig. 3) and continues to show up, even at the lowest initial current levels, when the dose rate is low enough to allow measurements at sufficiently low total dose. The extent of Region 1, or the dose value at which an individual unit starts to degrade, is quite variable among units of a type, as well as between types, and is a significant point for further consideration of devices.

The transition into Region 2, where the units show a typical linear increase in log I_{CBO} with log of dose, is quite distinct in all observed cases. For this reason, data which show only the degrading slope in Region 2 (perhaps because a high dose rate causes the first measurement to be beyond Region 1) can be extrapolated with some confidence back to the preradiation value to achieve an estimate of the transition dose. Major features of Region 2 are extreme variations between individual devices of a type and considerable dependence on collector voltage. This dependence is shown in Fig. 11 by the variations in I_{CBO} values achieved at about 10^4 rad, apparently about the end of Region 2 for this type.

In general, there appears to be a consistency between Regions 1 and 2 in that those units which show the greatest change in Region 2 also tend to indicate the shortest period of stability in Region 1; and conversely, those showing smaller changes in Region 2 appear to have the longer period of initial stability. Within the range of actual initial values (from 8×10^{-11} to 2×10^{-9} amperes) however, there appears to be no correlation between the initial value and the subsequent severity of degradation.

6.2 Distributions of Dose at Initiation of Degradation

With the recognition of the typical response of this transistor type to radiation, a question can first be asked regarding the variability in the extent of the region of stability, Region 1.

Fig. 18 shows several plots of the distribution, on a normal probability scale, of the integrated dose at which degradation is initiated (the end of Region 1). The three lowest curves show the results obtained from ex-
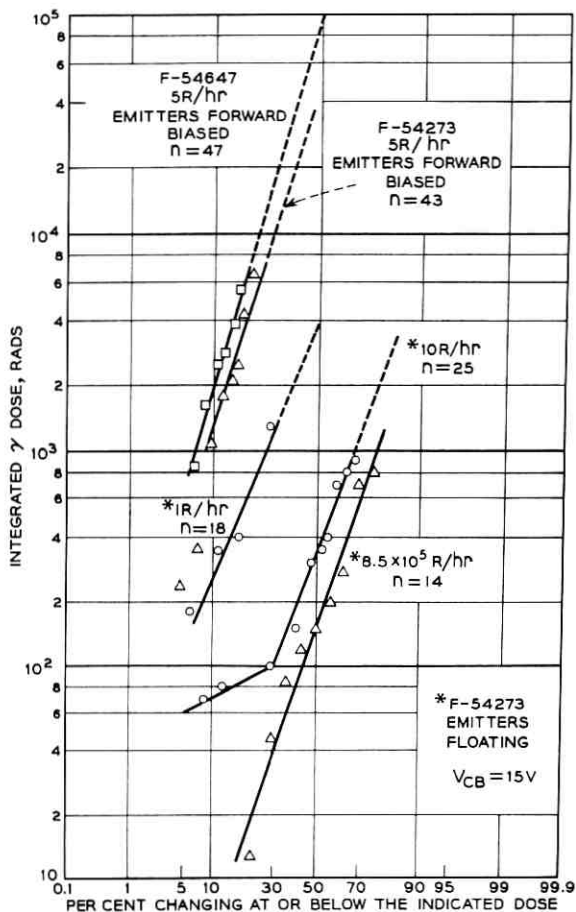
Fig. 18. — Distribution of the radiation dose at initiation of degradation.

posure to gamma radiation at 8.5×10^5 rads/hr., 10 rads/hr., and 1 rad/hr., all with a collector bias of 15 volts and with no connection to the emitters. There appears to be only about a factor of 2 difference between the 8.5×10^5 rad/hr. and the 10 rad/hr. results. This evidence of reciprocity between the 8.5×10^5 rad/hr. dose and the 10 rad/hr. dose at the end of Region 1 is in contrast to the lack of reciprocity indicated in Fig. 9 which compares reverse currents in Region 2. On the other hand, the difference between the 1 rad/hr. and 10 rad/hr. distributions seems to indicate a breakdown of reciprocity, at this level, in the dose required to initiate degradation.

The two upper curves of Fig. 18 are obtained at 5 rad/hr. and with a forward bias on the emitters during radiation in addition to the reverse collector bias. One of these curves is on the same transistor type (F-54273) as that of the lower distributions. The other distribution at 5 rad/hr. is on another transistor type (F-54647) which is basically the same but with a somewhat higher distribution of initial current gain. It would appear that the application of forward emitter bias causes an increase of roughly two orders of magnitude in the dose required for the onset of I_{CBO} degradation. This could result from the neutralization of positive charges on the surface of the base region by electrons injected into the base by the forward biased emitter.

6.3 *Distributions of I_{CBO} Increase*

Fig. 19 shows the distribution of the increase in collector reverse current resulting after 1.4×10^4 rads, which dose is past the completion of the variable Region 2. Here the I_{CBO}, on a log scale, is plotted on the normal probability scale. These data are on the same types of units represented in Fig. 18 and are obtained from radiation with the emitters forward biased. As in Fig. 18, the higher-gain type evidences a lesser degradation, having a longer region of stability as well as a smaller increase in I_{CBO} during Region 2. The excellent match to a log-normal distribution (except at very low currents where the data reflects the inaccuracy of measuring the difference between two nearly equal numbers) lends confidence to the interpretation that the break at the upper end of the distribution is caused by the transition into Region 3 of those units which have changed the most in the variable Region 2.

The broad distributions of Figs. 18 and 19 point up the necessity for relatively large-scale experiments in order to make valid comparisons between different test conditions or device types.

6.4 *Tests of Other Semiconductor Devices*

Similar kinds of distributions and general responses to the various conditions affecting degradation under radiation have been found in other codes of diffused silicon transistors and silicon diodes which have been encapsulated with a gas filling.

Fig. 20 shows, for example, the distribution of reverse current increase of a $\frac{1}{4}$-watt diffused silicon diode after 3.7×10^4 rads at 7.4×10^5 rads/hr. The broken shape of the distribution may simply indicate that the devices did not come from a single product run. The lack of the distinct break in the distribution at high currents suggests that the
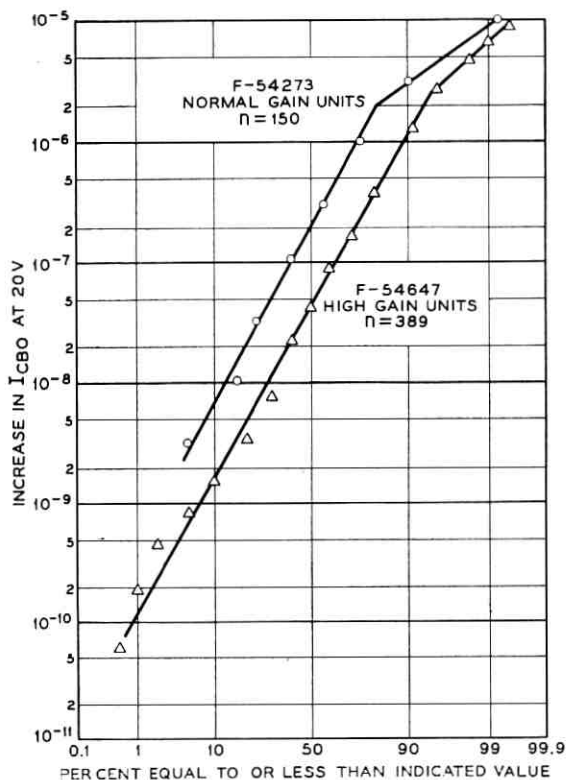
Fig. 19. — Distribution of current degradation after a dose of 1.4×10^4 rads at 8.5×10^5 rads/hr.

transition to Region 3 may not exist, or exists at a different dose. It would appear that each type must be examined in detail for a good understanding, even empirical, of its performance under radiation.

One interesting variation to the kind of response indicated above is that of a type of silicon alloy transistor which contains a silicone grease. In this case a relatively minor degradation is observed during short periods at a high gamma dose rate, but more severe degradation occurs during the period subsequent to completion of the exposure. This is indicated in Fig. 21, a plot of collector reverse current vs time after repeated exposures on an individual device. The device was exposed for several consecutive periods at 1.4×10^5 rad/hr., with current measurements subsequent to each exposure. The procedure was then continued at 8.5×10^5 rad/hr. Here each curve represents the I_{cBO} measurements subsequent to radiation for the total time indicated on the curve. I_{cBO}
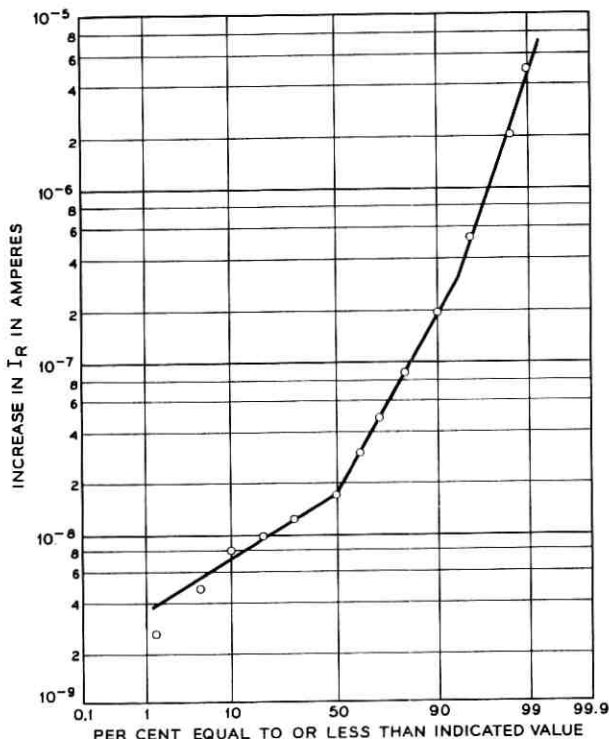
Fig. 20. — Distribution of degradation of current in a type of diffused silicon diode at 3.7×10^4 rads.

increases subsequent to the exposure, but may actually be decreased during the next exposure. After sufficient exposure, the devices evidence a saturation of the subsequent I_{CBO} , and also a saturation of the effect with continued exposure.

This saturation effect is confirmed at the 5 rad/hr. dose rate in Fig. 22, showing the I_{CBO} curves vs integrated dose for two typical transistors of this type. Both show an eventual saturation of the curve and one shows an ultimate reduction which appears also to be typical of this type of transistor.

Several types of diffused germanium transistors have also been examined, with the general result that increases in I_{CBO} are relatively small (of the order of a factor of 3 or 4), and relatively consistent, until the dose reaches about 10^6 rads, where more drastic increases may be expected. The dependence on V_{CB} was relatively insignificant. The effect of processing is evident in these types, however, with one group showing
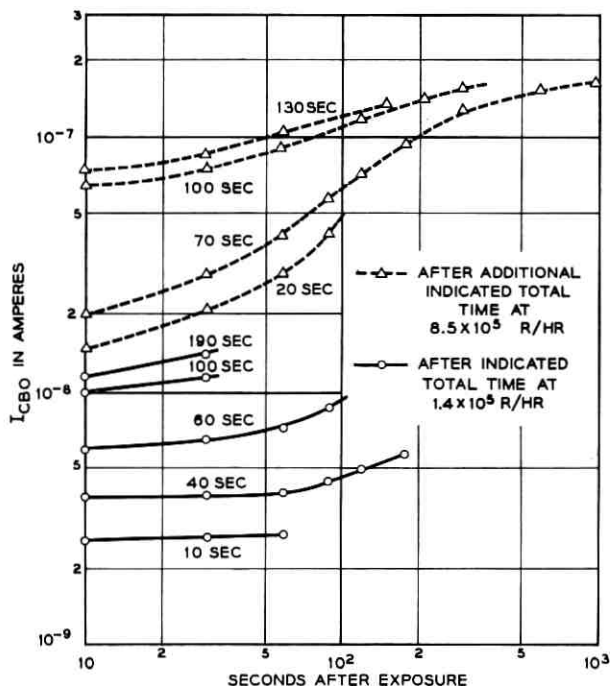
Fig. 21. — The pattern of collector current response in a grease-covered silicon transistor.

marked changes at about 10^4 rads, and another group having essentially no changes at 10^7 rads.

VII. TESTING AND SELECTION FOR TELSTAR DEVICES

The process of selecting the semiconductor components for use in the Telstar experimental communication satellite consisted of:

1. Qualification of design for reliability and performance as required in each application,
2. Fabrication,
3. Screening and preaging to assure satisfactory operation in the system environment,
4. Life testing,
5. Selection of the most stable devices.

At the time of recognition of surface effects from ionizing radiation (about October 1961), much of the life testing was already in process. Since the Telstar satellite was to orbit through the Van Allen radiation
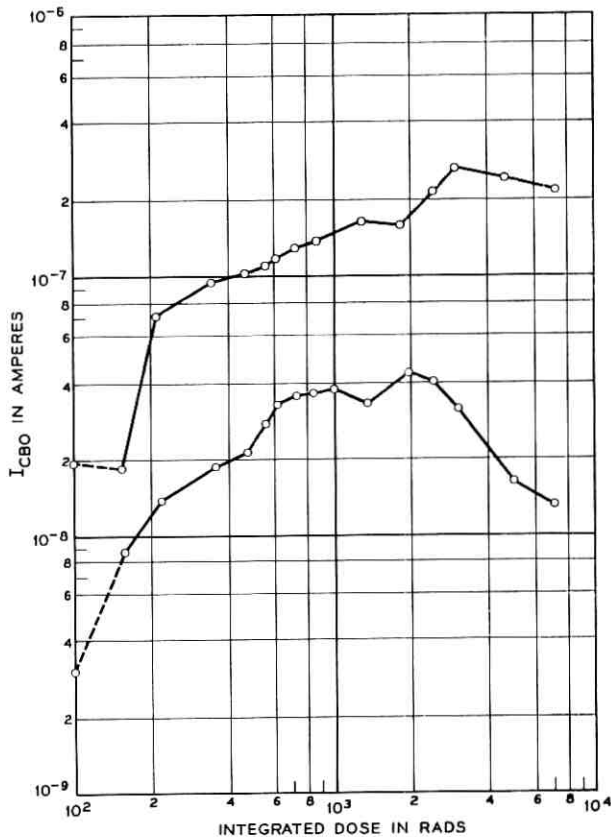
Fig. 22. — The response of a grease-covered silicon transistor under prolonged radiation at 5 rads/hr.

belt, the addition of radiation as an environmental factor was essential. Steps were therefore taken to determine the qualification of all types in the program and to study the screening and selection techniques that would be useful.

Fig. 23 shows a plot of the estimated ionizing radiation intensity for components inside a satellite, as a function of the shielding of the components from the external environment. These curves refer to the ionization produced by protons only and were developed from existing estimates of the Van Allen belt particle fluxes. The upper and lower limit curves reflect the uncertainties in this estimate. For shielding thicknesses of less than 0.1 inch the electron contribution to ionization should be considered as well, but for Telstar, typical components are shielded by
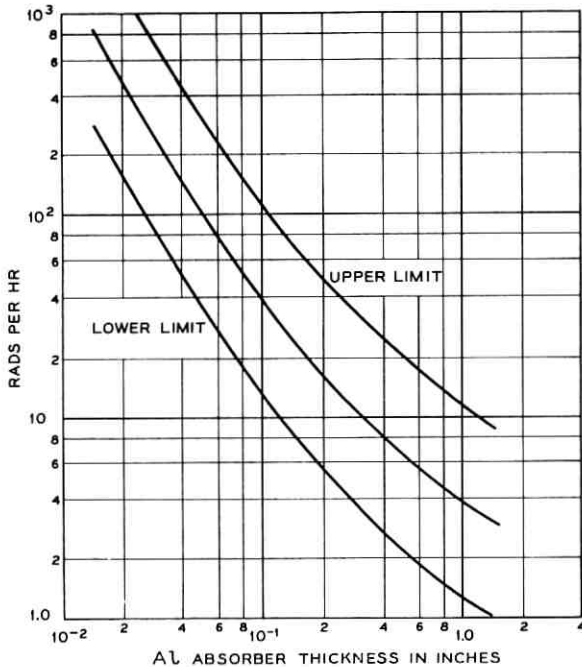
Fig. 23. — The maximum radiation intensity anticipated in space due to high-energy protons, as a function of the thickness of aluminum shielding.

the equivalent of 0.5 to 1.0 inch of aluminum and only high-energy protons are significant. The curves of Fig. 23 indicate that at this typical shielding, the radiation intensity would be a maximum of 10–20 rads/hr in the heart of the Van Allen belt. This results in a maximum average of 3–5 rads/hr over an entire orbit, considering Telstar's approximate 25 per cent effective exposure to Van Allen radiation.

With this estimate of the maximum radiation environment, all device types were given a gamma exposure at 8.5×10^5 rads/hr for one minute (the equivalent of at least three months in orbit), followed by at least one week at 3 rads/hr. Any device type showing evidence of change in either condition was replaced or subjected to individual selection or screening. Device types showing no change were considered satisfactory. All of these devices were subject to a 15-v reverse bias on the collector and an emitter current corresponding to the application.

The diffused silicon transistors F-54273 and F-54647 were used for experimental studies of screening and selection procedures. It is noted that the one-minute dose of 1.4×10^4 rads (at 8.5×10^5 rads/hr) is not

much beyond the variable Region 2 of Fig. 17 and should be very effective in providing a comparison between individuals. The study of the effectiveness of such a dose as a screening procedure was carried out through the following program:

1. Preradiation of a number of these transistors to screening exposures between 10 seconds and 6 minutes at 8.5×10^5 rad/hr and
2. Subsequent exposure at 5 rad/hr.

This permits an evaluation of the effect of the low dose rate after a screening dose.

Fig. 24 shows the collector reverse current measurements of two units from this program, these being generally typical of the results of all of the devices. In this case, the measuring equipment was limited in sensitivity to about 10^{-9} amperes and the initial values prior to radiation are shown at this value although they may have been somewhat lower. Unit No. 762 is representative of those units which suffered a relatively minor increase in reverse current during the preradiation dose. During the subsequent radiation at 5 rad/hr, this unit returned quite rapidly to its original value, remaining there until the dose became somewhat
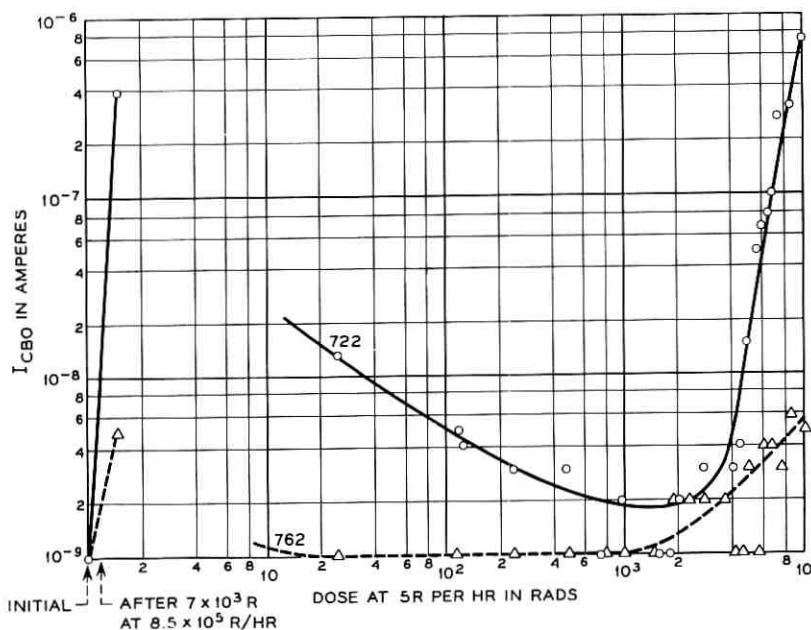

Fig. 24. — The typical pattern of change in I_{CBO} in low-level radiation after a screening dose.

greater than 10^3 rads, at which time it began to degrade in a fashion very similar to that expected from earlier tests. Unit No. 722 suffered a much greater change during the preradiation and took an appreciably longer time recovering toward its initial value. Before it fully recovered it reached the point of onset of final degradation and began to change quite rapidly. (A third type of response was seen in an occasional unit which degraded so severely during the preradiation that the subsequent low dose resulted in further increase in current and no recovery was evident at all.)

The influence of the screening dose on the distribution of onset of the final I_{CBO} degradation (at 5 rad/hr) is shown in Fig. 25 for four similar groups of 12 units each. One of these groups was irradiated at 5 rad/hr with no initial radiation at the high dose level. The other three groups were given an initial radiation at 8.5×10^5 rad/hr for different lengths of time to achieve the initial radiation dose indicated in the figure. It is seen that all of these distributions are essentially the same, indicating that the 5 rad/hr dose will cause a modification of the surface condition established by the high initial dose, causing the units to look eventually as if they had not received the initial dose. This is evidence of a contradiction to the principle of reciprocity of dose rate and time, in that the
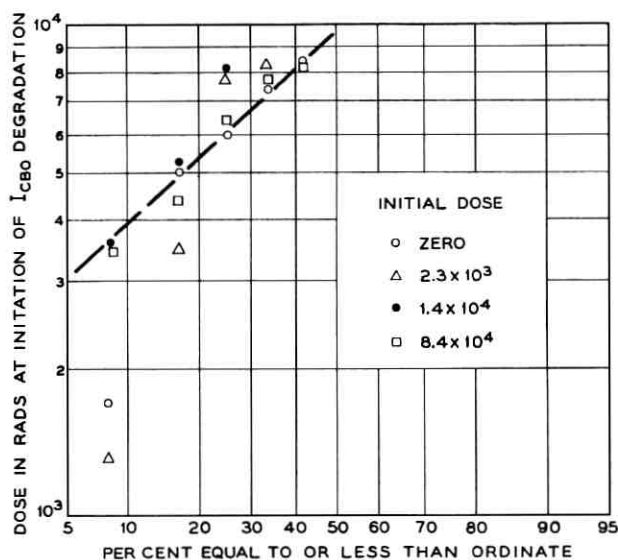

Fig. 25. — Distribution in dose for initiation of degradation with and without pre-irradiation.

additional dose added at the 5 rad/hr rate does not normally cause a continuation in the degradation produced in the initial dose.

Another point of interest in Fig. 24 is the comparison of the I_{CBO} value resulting from the preradiation dose with that subsequently occurring after an equal dose at 5 rad/hr. Inspection of the two curves on Fig. 24 reveals that the currents at 7×10^3 rad are approaching those resulting from the preradiation dose. Fig. 26 shows a plot of the results of all the units so treated. The dashed line indicates the one-to-one correlation between the two current values, those points to the right of the line representing the units which during the 5 rad/hr exposure did not develop reverse currents as high as those obtained in the initial dose. This line was not extended below about 4×10^{-9} amperes because this was approaching the limit of sensitivity of equipment which was then in use. It is noted that all but two of the significant readings fall on the side of the one-to-one correlation line corresponding to larger changes after the preradiation dose than after the subsequent low dose rate exposure.

In some cases the lack of correlation is quite significant, the currents after the low dose rate exposure remaining below 10^{-8} amperes although quite large changes were observed after the screening dose. In these
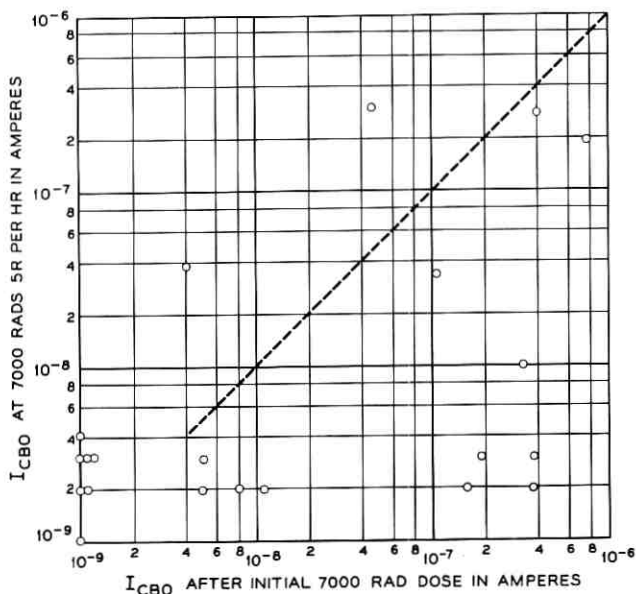

Fig. 26. — Correlation of change in I_{CBO} after a screening dose at 8×10^5 rads/hr with subsequent changes at 5 rads/hr.

units, however, the changes in current gain after the low dose rate exposure were found to be appreciably greater than in those units which were more stable in the initial screening dose. Consequently, it is found that selecting of those units with I_{CBO} less than 10^{-8} amperes after the screening dose would have eliminated 9 of the 10 units which subsequently degraded to either I_{CBO} values greater than 10^{-8} amperes or gains less than 50 per cent of the initial gain. Of the 12 units good in these respects after the low dose rate exposure, only two would have been eliminated by selection following the screening dose. It is thus shown that selection on the basis of I_{CBO} after a screening dose is effective for this device type.

An indication of the effect of this screening on the distribution in degradation of I_{CBO} under low-level radiation exposure is given in Fig. 27. Both distributions are of I_{CBO} after exposure, first to 30 seconds at 8.5×10^5 rad/hr and then to 10^4 rad at 5 rad/hr. One distribution is of
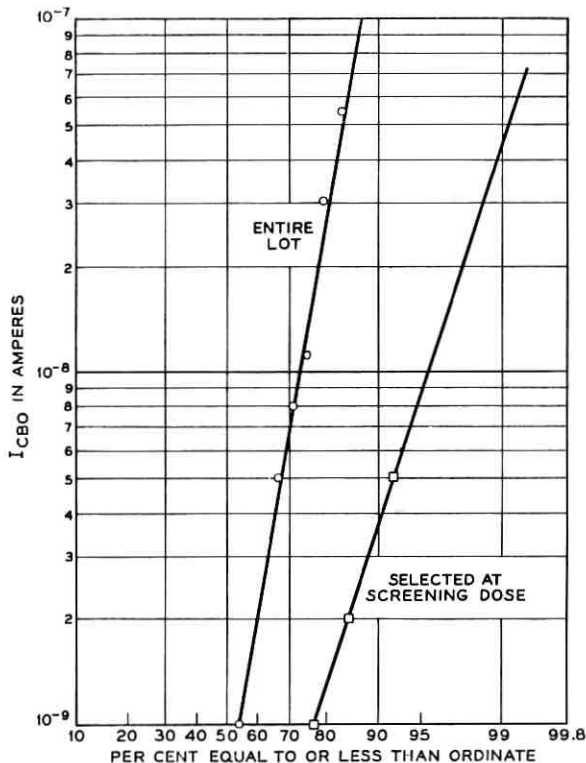

Fig. 27. — Distribution of radiation response in screened and unscreened devices.

the entire sample of devices and the other is of only those which were selected for I_{CBO} less than 10^{-8} amperes after the initial dose. It is seen that an appreciable improvement is achieved by the screening procedure.

VIII. SUMMARY

It has been found that changes can occur in semiconductor characteristics, because of surface effects of ionizing radiation in a device under electrical bias, at much lower doses than those required to produce surface effects in the absence of bias or to produce changes in the bulk of the semiconductor. These effects apparently arise from ionization in the gas of a device encapsulation and collection of ions on the device surface. The gas ions probably serve to produce and activate chemical species on the semiconductor which induce surface inversion layers that alter the junction characteristics. The chemical condition of the surface previous to irradiation is apparently involved in the process in a sensitive way. The effects observed depend on junction bias, envelope potential, and in many respects on total radiation dose rather than dose rate. The major features of these observations and of observations of the characteristic recovery of the surface effects after radiation under bias can be described by the ion-induced inversion layer model of the process.

In the time since the first direct observation of this effect, a comprehensive study of all types of semiconductor devices has not been attempted. Observations of several types, however, indicate that different types may respond quite differently to radiation, that the response may be quite dependent upon processing (and therefore upon production periods or batches), and that within a type the response may be quite variable between individual units.

One type of transistor was used for much of the experimental work of defining the radiation effect, and the I_{CBO} response was studied because the large changes facilitated comparative measurements. It is recognized that current gain and other surface-dependent characteristics can also be affected by radiation, and they should be observed in any evaluation of types, if critical in specific applications. It is hoped that the studies presented here will be a guide in formulating such evaluations.

It has also been found that at least some devices can be screened for sensitivity to radiation by means of a short-time, high-level dose, with correlation to subsequent low-level exposure results. Here, too, it is recognized that variations exist among types, and the usefulness of such screening operations should be evaluated for each type of interest.

It is hoped that further work will serve to provide more definitive results relating the radiation effects to a physical model and also to the

surface conditions or processes contributing to the effect. These preliminary results can at least serve to expose the problem and to suggest the lines of further study and action.

IX. ACKNOWLEDGMENTS

The material presented here includes the contributions of many of the authors' colleagues. Of particular significance have been the efforts of E. P. Moyer, E. R. Schmid, and J. Lange, of Bell Laboratories and G. L. Miller of Brookhaven National Laboratory.

REFERENCES

1. See Conference on Radiation Effects in Semiconductors, J. Appl. Physics, **30,** August, 1959, in particular, Watkins, G. D., Corbitt, J. J., and Walker, R. M., Spin Resonance in Electron-Irradiated Silicon, p. 1198.
2. Easley, J. W., and Dooley, J. A., Fast Neutron Bombardment Reduction of Transistor Current Gain, J. Appl. Phys., **31,** June, 1960, p. 1024.
3. Proceedings of Second Conference on Nuclear Radiation Effects on Semiconductor Devices, Materials, and Circuits, September 17 and 18, 1959.
4. I.R.E. Transactions on Nuclear Science, Vol. **NS-8,** January 1, 1961.
5. Rosenzweig, W., Diffusion Length Measurement by Means of Ionizing Radiation, B.S.T.J., **41,** September, 1962, p. 1573.
6. Smits, F. M., Smith, K. D., and Brown, W. L., Solar Cells for Communication Satellites in the Van Allen Belt, Jour. Brit. I.R.E., **22,** 1961, p. 161.
7. Baruch, P., Mobility of Radiation Induced Defects in Germanium, J. Appl. Phys., **32,** April, 1961, p. 653.
8. Kingston, R. H., Review of Germanium Surface Phenomena, J. Appl. Phys., **27,** February, 1956, p. 101.

Development of Solderless Wire Connector for Splicing Multipair Cable

By H. J. GRAFF, J. M. PEACOCK, and J. J. ZALMANS

This paper describes the development and laboratory testing of a new high-reliability solderless connector for splicing cable conductors. It includes discussion of the physical parameters which influence the performance of electrical contacts in general. The experimental and analytical techniques which evolved as part of this project permit important reductions in the amount of experimental data required to make reliability and aging predictions, and should be useful in other problems dealing with the appraisal of electrical contacts.

I. INTRODUCTION

The development of new connectors and new techniques for making wire connections has been actively pursued at Bell Telephone Laboratories for some time. Current estimates indicate that well over two billion wire connections are made each year in the Bell System. These connections obviously represent a large investment, and the problems of cost and reliability are becoming increasingly difficult as modern circuits grow in complexity. Performance testing of a connector to establish its reliability is a challenging task.

1.1 *B Wire Connector*

The purpose of this paper is to describe the development and performance appraisal of the B Wire Connector recently released for splicing outside plant multipair cable conductors. As illustrated in Fig. 1, this connector consists of a thin springy liner with sharp tangs on the internal surface, and a brass outer shell encased in a plastic jacket. In use, the conductors to be joined are inserted without removing insulation, and the connector is pressed with a pneumatic tool. The sharp tangs penetrate the insulation and establish contact with the conductors as illustrated in Fig. 2. The individual tangs act as springs, storing en-
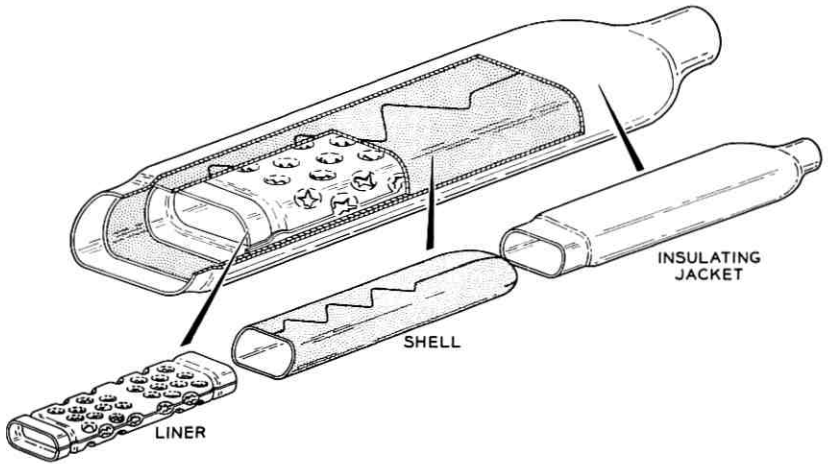
Fig. 1 — B Wire Connector.

ergy to maintain intimate contact with the conductors over the long service life of the device. Electrically, the contact is equivalent to a soldered joint, and therefore is suitable for use in any type circuit where multipair cable may be used. Demand for the B Wire Connector is expected to reach about 250 million per year with resulting annual savings of several million dollars.

1.2 Reliability and Aging Requirements

The application for which the B Wire Connector is intended demands the utmost in reliability. The connector, of course, is a series element; a typical cable circuit has a splice every 500 feet or so, often as many as 20 splices in a line. The connections are made under field conditions,

Fig. 2 — B Wire Connector — cross section after pressing on wires.

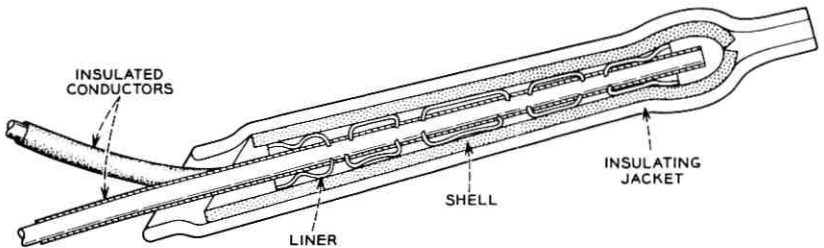and are expected to match the service life of the cable, frequently well over 40 years. The splices are enclosed in splice cases or lead sleeves, many of which are hermetically sealed.

When reduced to quantitative values, the reliability requirements are very high indeed. Consider a one-mile run of 2424-pair cable, with ten splices. 48,480 connections are needed, and if 0.1 per cent or 48 of these are bad initially, or become bad in time, 48 or 2 per cent of the circuits would be lost, neglecting the possibility that two bad connections could occur in the same circuit. Similarly, at a failure rate of 0.01 per cent or 1 in 10,000, five circuits would not be available. To assure satisfactory field acceptance, and to allow margin for other sources of failure, target values for the inherent reliability of the connector were established at 1 in 100,000 after 40 years of aging in the outside plant environment.

1.3 *Performance Appraisal Methods*

The appraisal of a device for reliability and performance requirements such as these is always a very difficult and challenging task. As part of this project, a number of advances have been made in the techniques usually used for contact appraisals so as to reduce the amount of experimental data required and to improve the precision of the aging predictions. Some of these techniques are applicable to development work on other types of connectors, and will be discussed in this paper. First it is necessary to describe the mechanisms which influence an electrical contact.

II. CONTACT MECHANISMS

Wire connectors in use in the Bell System today take many forms, ranging from the time-proven binding post, through the solderless wrapped devices, to the newer insulation-piercing, quick-connect terminals. Although radically different in many of their features, these devices have much in common — the basic mechanisms that influence the behavior of the contact are the same. In all of them the circuit is established by clamping together two metallic elements, the wire and the terminal. Circuit continuity is established through the areas of metal-to-metal contact. This contact can be considered to have two components of "contact resistance": (1) constriction resistance due to the convergence of current paths within the parent metal to the relatively small points of contact, and (2) interfacial resistance due to the possible presence of resistive surface films.

A number of variables influence the actual value of the contact resistance in a given connection. Some of these can be controlled quite well through proper selection of materials and platings, geometry, etc. Others tend to be more random in nature: for example, the fine-grain details of surface roughness or oxide coverage which would influence the actual areas in contact at a given loading. Due to the random nature of some of these mechanisms, a distribution of contact resistance may be expected. The spread or range of resistance is a measure of the control or reproducibility of a connector, and thus tends to indicate its reliability. To hold the degree of variability to a minimum, it is necessary to understand the various environmental and aging mechanisms applying to a particular design.

Having established an initial contact of a given quality, one must be alert to possible degradation of the contact with time. This is likely to happen as a result of corrosion, stress relaxation, "breathing" of oxidizing atmosphere due to contraction and expansion of connecting parts consisting of different materials, or mechanical disturbance. A good degree of mechanical rigidity, with adequate contact force and energy storage to maintain these forces over the years, are necessary design features to achieve long service objectives. With inadequate provision for these requirements, high resistance and noisy connections are inevitable, and under dry circuit conditions (low-voltage, low-current signals) barrier films can easily cause complete circuit failure. The requirements for sufficient areas of intimate "gas tight" contact and for energy storage have been well known for some time and were studied in some detail during the development of the solderless wrapped connection.[1]

III. MEASURING TECHNIQUES

Measurement of the resistance distribution and aging trends suggested above poses quite a challenge. For a "good" connector design, i.e., one with low contact resistance, the variation of contact resistance from sample to sample may be in the order of a few microhms. Resistance measuring techniques commonly used in the past for the appraisal of contact devices have been capable of resolving joint resistances only to the low milliohm range. With these techniques it is possible to single out "high-resistance" joints but not to study the statistical distribution of "good" joints. The sample sizes needed to establish reliability rates and to draw conclusions about designs and design variables with these relatively insensitive measuring techniques are prohibitively large. On the other hand, with experimental data of sufficient precision, the be-

havior of the many variables can be studied in much greater detail, and analytical techniques can be brought to bear so as to reduce drastically the amount of experimental data required. For effective analysis by statistical techniques with reasonable sample sizes, the data must be collected with instrumentation sensitive enough to permit the quantitative measurement of differences in resistance levels from joint to joint. Also, for analysis of the aging mechanisms in types of connectors where this is of concern, changes in contact resistance with time must be measurable at short time intervals (several days or a week). If the above requirements are met, statistical study of contact resistance distributions is possible — both initially and after any desired interval of normal or accelerated aging — and reliability rates can be estimated with small sample sizes.

Techniques for measuring fractions of a microhm of resistance have been known for years. The problem in measuring contact resistance on actual devices to these levels of accuracy lies principally in separating the variables in such a way that the experimental errors do not swamp out the resistance to be measured. A typical device, for example a connector joining two wires, may be thought of as a two-terminal black box, on which only external measurements may be made. The various resistive elements within the box include the internal resistances of the lead wire and connector in addition to the contact resistances between wires and connector. The normal tolerances on the internal resistances, due to uncertainties in wire diameter, length, temperature and resistivity, are generally larger than the true contact resistance of connectors of high enough quality to be of interest for long-term service on dry circuits. These uncertainties must be either eliminated or accurately compensated if the desired measurements are to be obtained.

A very simple circuit has been developed to permit resistance measurements of the kind contemplated here. This circuit, as shown in Fig. 3, is a modified Kelvin bridge in which the unknown resistance of the contact is compared to a "standard," which is an accurately controlled length of one of the lead wires.

The circuit has a number of features which reduce experimental errors:

1. Variations in lead wire diameter and resistivity are compensated exactly by proportionate changes in resistance of the standard.

2. Resistance measurements are insensitive to temperature changes, at least to the extent that the unknown has the same temperature coefficient as the standard.

3. The effects of thermal potentials can be balanced out in dc bridge measurements by using normal and reverse polarity of test current.
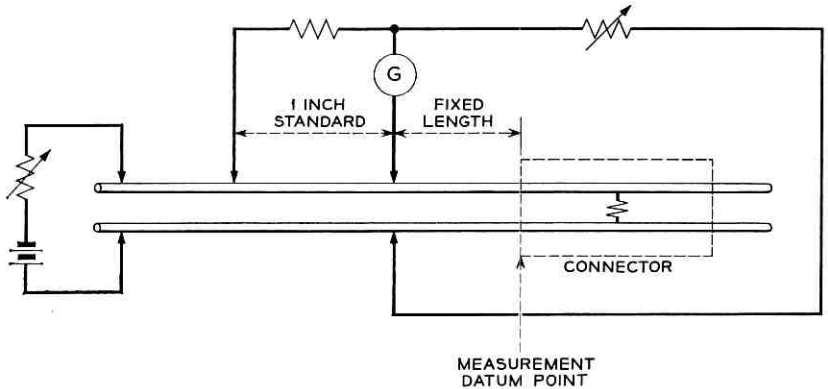
Fig. 3 — Modified Kelvin bridge — setup for loop resistance measurement.

4. The effect of variable resistances due to the connection of the sample to the test set (frequently a source of appreciable error) may be made arbitrarily small by putting them in series with the high impedance arms of the bridge.

The loop resistance measurement, which includes the internal resistance of the lead wires and connector in addition to the contact resistance, Fig. 3, can be used effectively to determine variations in contact resistance between connections, and to measure changes in contact resistance with time. To keep the experimental errors to a minimum, these measurements require very close control of the reference points from which resistance measurements are made. With the measuring equipment used, mechanical tolerances are such that the average variation in loop resistance due to these factors is about ±4 microhms (for 19-gauge conductors and 100 milliamperes test current).

In some cases it is possible to make a more cumbersome measurement than the simple "loop" measurement shown in Fig. 3 so as to separate the various "black box" resistance elements from the contact resistance itself. Modifications in the test set connections as shown in Fig. 4 have been made for this purpose. The unknown bridge arm in Fig. 4(a) includes the resistance of one of the lead wires in to the point of contact and the contact resistance between the two wires. In Fig. 4(b) the unknown contains merely the resistance of the lead wire to the "centroid" of the contacts. The difference between these two readings gives us the contact resistance by itself. These multiple measurements are quite accurate for simple point-contact configurations. For more complicated structures having distributed contacts, the observed resistance must be interpreted in terms of the distribution, somewhat analogous to the
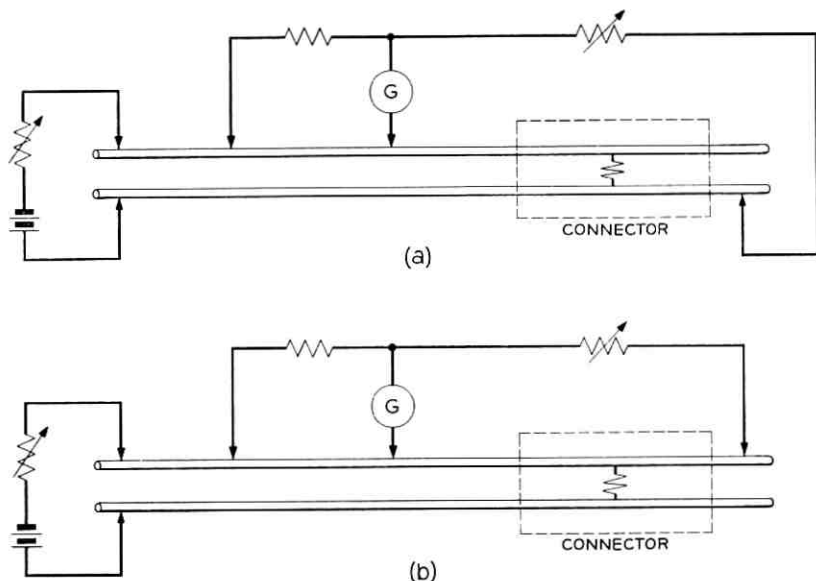
(a)

(b)

Fig. 4 — Modified Kelvin bridge. Circuit (a) is used to measure the resistance of the top lead wire in to the point of contact, plus the resistance between the two wires. Circuit (b) measures merely the lead wire resistance. The contact resistance is calculated from these two measurements.

input observations on the distributed parameters of a transmission line. In many cases the primary interest is in changes that occur with time rather than in the absolute value of contact resistance, and the more straightforward loop resistance measurement is sufficient for these purposes.

IV. AUTOMATED MEASURING APPARATUS

The need for automated measurements and streamlined data handling was recognized in the early development stages of the B Wire Connector. The manual bridge described in Section III was automated to read and record joint resistances of groups of up to twenty joints automatically. This equipment is shown in Fig. 5. A block diagram of the automated bridge giving major components is shown in Fig. 6. The basic circuit is the same as for the manual bridge. The self-balancing feature is obtained by using the bridge unbalance voltage as the input to the servo-system of a self-balancing potentiometer. The shaft position of the potentiometer, directly related to the resistance measured, is encoded and converted into digital form for direct recording on punched cards or for
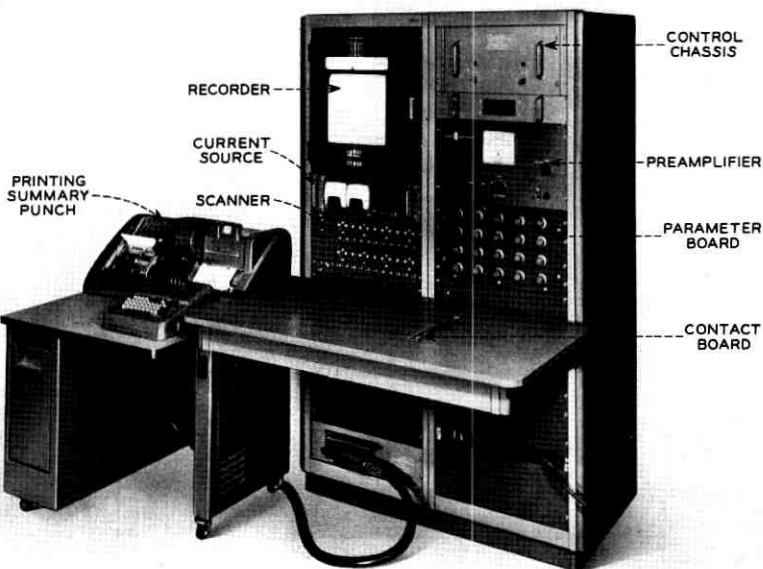
Fig. 5 — Automated, self-balancing bridge for repetitive joint resistance measurements.

visual readout. A timing unit controls the readout, current reversal and the scanner switch by which connections to individual joints mounted on a board are made.

Fig. 7 shows a mounting board with 16 joints. The board also includes four looped conductors, which are used as controls to detect possible errors due to misalignment or shifting of the mounting board relative to knife-edge contacts on the contact board. The use of controls is arbitrary, of course. In this instance, it will be noted that the length of the loop was random. This mounting arrangement provides good control of reference points and saves considerable time on repetitive measurements.

As noted in Fig. 3, the loop resistance of the joint is compared to the resistance of a "standard" length of wire, in this case, one inch. It has proven convenient to read out the resistance in thousandths of an inch of conductor, measured from the entrance end of the connector, rather than in ohms. For the purpose of this paper, the resistance values plotted have been converted to milliohms. It should be noted that the wire resistance varies with the wire gauge used, and this must be taken into account in comparing one plot with another. The precision or repeatability of the system is well within ten microhms, the actual value
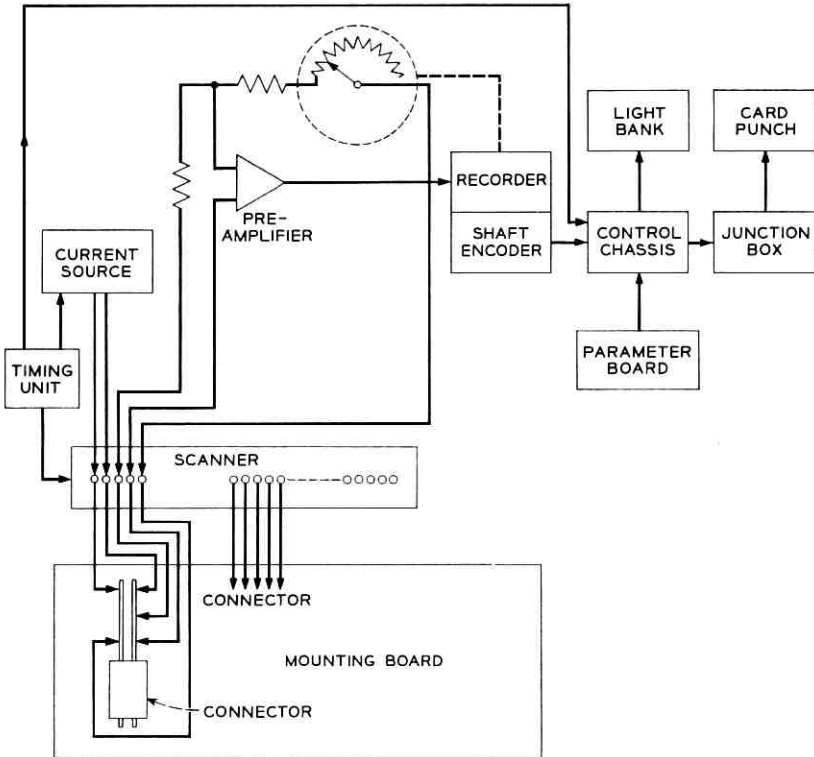
Fig. 6 — Block diagram of automated bridge.

depending on the gauge of the wire being tested and the test current used.

V. EVALUATION OF B WIRE CONNECTOR

The objectives in the design of the new connector were to improve the electrical performance and reliability of the previous standard "pigtail" cable splice, and to achieve economies through time savings by simplifying the splicing methods. The program which led to the final design of the B Wire Connector shown in Fig. 1 will be described in this section.

5.1 Evolution of Basic Design

After considering a number of wire joining methods, it appeared that a device which could be pressed onto the ends of unskinned conductors

Fig. 7 — B Wire Connectors on mounting board used in automated bridge.

offered the most attractive solution. For this particular application, initial studies showed that such a device should consist of two metallic members: an inner member of relatively hard material with some form of penetrators to pierce the insulation, and a soft, ductile outer member to hold the penetrators in contact with the metallic conductors after the pressing force is removed, and thus minimize the effects of springback. With this pri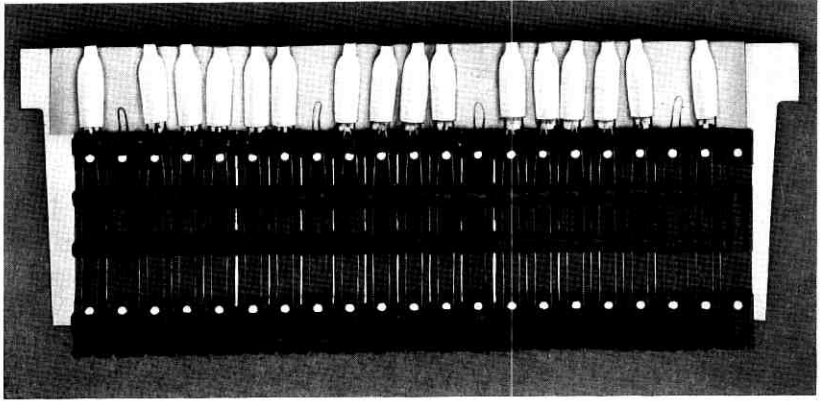nciple in mind, two alternative designs were proposed. These two designs used the same soft outer brass sleeve, but differed in the design of the inner member. The liner in connector X consisted of a helix formed from hard drawn rectangular brass wire that had been twisted on its own axis. Connector Y had a liner of thin spring bronze pierced with numerous holes so as to form a multiplicity of tang-like penetrators on the inner surface. At this early stage of the design effort, a number of questions arose, the principal one being, "Which design approach offers the most promise?"

To appraise the performance of connectors X and Y, loop resistance measurements were obtained on small groups of comparably made joints of both types. Measurements were repeated periodically, during which time the connections aged naturally at laboratory ambient. Minute differences in joint resistance resulting from statistical variations in contact characteristics and the aging phenomena were readily discernible with these measurements. Mean connector resistances and standard deviations were plotted vs log time and are shown in Fig. 8.

A study of the aging characteristics, as illustrated by the mean resistance vs time, shows a significant difference between the two connectors: connector Y has a lower aging rate. The plot also gives some
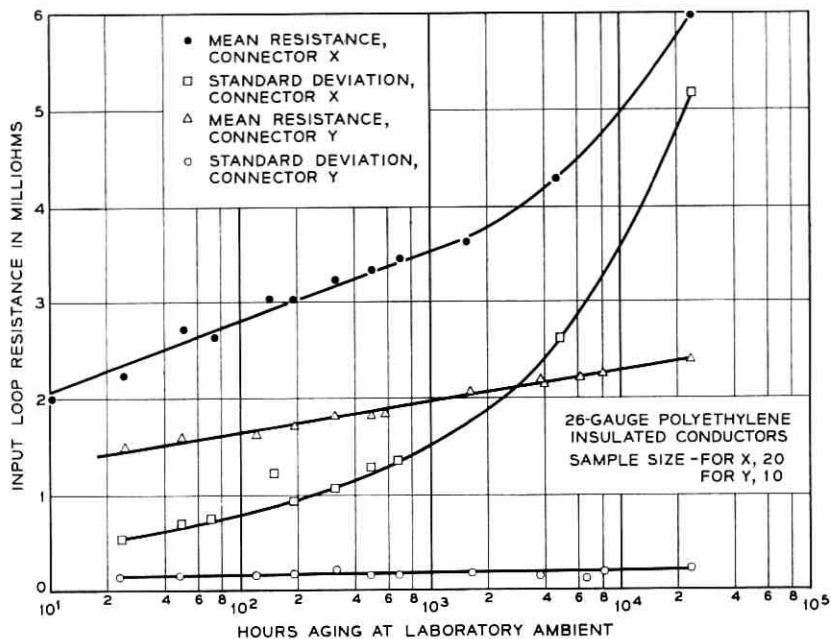
Fig. 8 — Comparative resistance plots for X and Y connectors showing performance under natural aging conditions.

insight as to the fundamental mechanisms affecting the performance of connector Y. The linear relationship of resistance and log time indicates the aging phenomenon is following some exponential law. Further analyses of these and other data suggest that the principal component of aging results from stress relaxation of the soft copper conductor at the various points of contact. Stress relaxation has been studied extensively, is predictable, and can be compensated by the provision of increased energy storage in the connector.

The standard deviation plots are also of interest since they provide an indication of the reliability that may be expected from larger samples. Any significant increase in the deviation with aging is an indication of instability and poor reliability. In view of these considerations, it can be concluded from Fig. 8 that connector Y has both a potentially longer service life and a higher reliability rate than connector X. Fortunately, connector Y is also more manufacturable.

It is of interest to note that the initial performance of the two connector designs is not radically different. Based on the initial results, either design might be expected to give good results. In less than one

month of natural aging, however, significant differences in the aging rates of the two designs became apparent, in spite of the fact that only a handful of samples were under study. As verified by subsequent data, a valid choice could have been made between the connectors at this time. It is apparent that the ability of the test method to resolve and detect small changes in resistance is of considerable assistance in predicting the long-time performance of connector designs with a minimum number of test samples.

If the resolving power of the test method were degraded to ±1 milliohms, as an example, a significant difference between the two designs could be detected eventually, after the lapse of a much longer aging time. Also, it is possible to improve the accuracy of the predictions of the standard deviation, in spite of the poor accuracy of the test method, by testing many more samples, and by applying the appropriate statistical principles. Neither alternative offers an attractive substitute for good resolving power in the test method.

Still another alternative, in case the resolving power of the test method is poor, is to resort to accelerated aging. A good example of an accelerated test was used in developing the solderless wrapped connector.[2] Here, the connections were exposed to 105°C for 150 days, with intermittent mechanical disturbance. One criterion for judging the performance of the connector is the magnitude of resistance instability resulting from the mechanical disturbance. The measurements are repeated periodically as the stress relaxation proceeds. Due to the mechanical stability and rigidity of connectors X and Y, mechanical disturbance does not produce significant resistance fluctuations, certainly not any that could be detected by measurements of 1.0- or even 0.5-milliohm sensitivity.

5.2 Development of Final Design

To complete the design of connector Y, now known as the B Wire Connector, a number of experiments were conducted in which the pertinent design parameters were varied to determine their effects on the performance. An interesting example is a Latin Square experiment performed to study the geometric variables involved in the design of the perforated liner. In this experiment, three levels of pierced hole size and spacing were studied for each of three different material thicknesses. The results of this experiment are shown in Table I, which is a tabulation of the resistance change after 1400 hours of aging at laboratory ambient for each of the design combinations. An analysis of variance indicates that the 0.005 inch thickness is inferior from the standpoint of contact resistance to either the 0.006- or 0.007-inch thickness at the

TABLE I — LATIN SQUARE EXPERIMENT: RESISTANCE CHANGES DUE TO AGING
FOR VARIOUS LINER THICKNESSES AND PIERCED HOLE SIZES
(26-Gauge Pulp-Insulated Conductors)

Hole Size in Inches:	Liner Thickness in Inches			Row Sum:
	0.005	0.006	0.007	
0.025	0.093*	0.001	0.012	
	0.047	0.004	0.001	
	0.026	−0.005	0.001	
	0.020	0.019	0.000	
	0.186	0.019	0.014	0.219
0.030	0.045	0.024	0.009	
	0.024	0.019	0.018	
	0.076	0.019	0.010	
	0.080	0.002	0.003	
	0.225	0.064	0.040	0.329
0.035	0.042	0.003	0.007	
	0.025	0.012	0.005	
	0.019	0.009	0.007	
	0.034	−0.001	−0.002	
	0.120	0.023	0.017	0.160
Column Sum:	0.531	0.106	0.071	0.708

* Resistance increase in equivalent inches of 26-gauge conductor after 1400
hours of aging. This conductor has a nominal resistance of 3.48 microhms/mil.

1 per cent level of significance. The analysis further indicates that the
variations occurring between pierced hole sizes are probably due to
chance causes. Although differences in performance could be noted with
the measuring precision available, this experiment, and others of a
similar nature designed to study other variables such as material hard-
ness, tang shape, etc., showed quite clearly that the connector perform-
ance was not critically dependent on the minor details of the liner
fabrication, and that reasonable manufacturing tolerances were not
likely to upset the reliability estimates.

In other experiments, accelerated aging tests were used to good ad-
vantage to magnify the effects being studied. One such experiment was
used to demonstrate the advantage of tin plating the liner. In this ex-
periment, a group of connections was subjected to cyclically varying
humidity. The pulp insulation used on cable conductors is quite sensi-
tive to changes in humidity. Pulp is trapped under compression in the
connector, and humidity changes cause an alternate shrinking and
swelling of the pulp and consequently a mechanical working of the
joint. This test has proven very effective in evaluating the stability of

the various contacts within the connector. Fig. 9 shows graphically the behavior of two small groups of joints under this accelerated test. One group has liners with tin plating; the other group has liners without plating.

Analysis of these results led to a good understanding of the manner in which the individual contacts are made. Fig. 2 is a cross-sectional view which shows the liner tangs contacting and digging into the conductor. Due to spring-back, immediately after pressing, small voids occur here and there between the tang and conductor. The mechanical working of the joint by humidity changes causes sufficient movement to dislodge the contacts. Tin plating, however, provides a soft, smearable surface which is evidently quite effective in filling these voids and preventing motion. In addition to better stability, the increased contact area due to tin plating reduces the contact resistance quite significantly, and at the same time reduces the statistical deviation among similar joints. Also, previous work[3] has shown that tin-plated surfaces under sufficient contact pressures are subject to solid-state diffusion, which, of course, would improve the electrical contact as a function of time.

5.3 Accelerated Testing

The design and interpretation of accelerated tests, although necessary in many cases, are always both difficult and controversial. Under
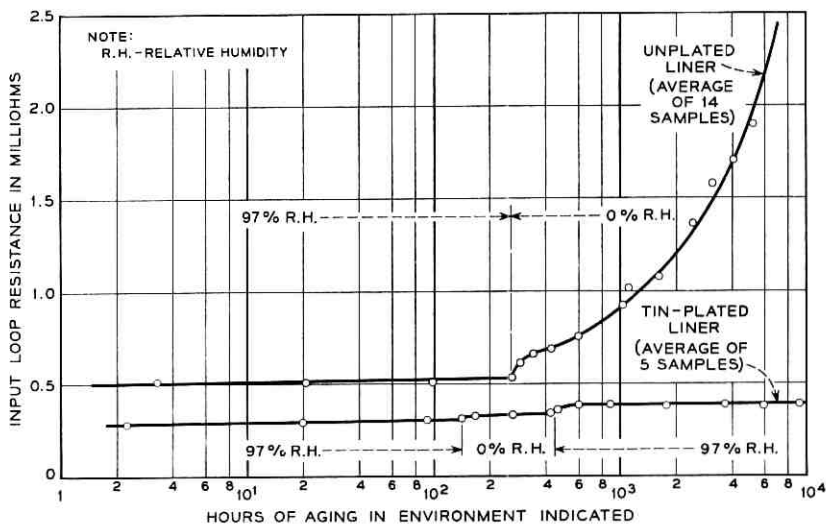

Fig. 9 — Humidity cycling test results which show the superior performance of tin-plated contact surfaces.

all normal environments, a number of aging mechanisms are involved, with various degrees of interaction between them. An accelerated test should ideally accelerate each of the mechanisms equally and preserve the interaction factors; all the pertinent mechanisms must be included or accounted for, and any new mechanisms that may emerge in the future must be anticipated. The development of a test that will satisfy all these requirements is a Utopian goal, and as a practical matter, some degree of engineering risk is inevitable in the use of accelerated tests.

During the development of the B Wire Connector, considerable effort was devoted to the identification of the pertinent aging mechanisms and to the design of accelerated tests to study their influence. One interesting approach has been to design a test that is obviously too severe; if the connector survives this test, then a good factor of safety is assured, and the quantitative calibration of the test becomes unnecessary. The humidity cycling test already mentioned is an example. In its proposed usage on pulp-insulated conductors, the B Wire Connector will always age in a hermetically sealed, low-humidity cable system; and by comparison, the accelerated humidity cycling test is quite severe. The good performance of the tin-plated inserts gives us confidence that the connector has an adequate margin of safety against the effects of changing humidity without further calibration.

In other cases, it has been necessary to attempt a more precise calibration. For instance, we are concerned with the aging mechanism resulting from daily and seasonal temperature fluctuations. These changes cause a mechanical working of the contacts brought about by the differences in coefficient of thermal expansion of the conductor-connector-insulation components. This type of aging could result in contact instability. In addition, the mechanical working of the contact parts may expose the contact areas to the entry of gas and to the formation of oxides or insulating films.

In evaluating the performance of the connector on polyethylene-insulated conductors, the temperature change becomes an especially important variable. The insulation trapped under compression in the connector causes considerable mechanical working of the contacts due to its coefficient of thermal expansion, which is much larger than the coefficients of the metallic parts of the connector. To determine the effect this mechanism has on connector performance, two temperature cycling tests have been devised. These tests have the following temperature ranges: (1) $-320°$ to $+160°F$, and (2) $-40°$ to $+140°F$.

The first test was chosen in order to subject joints to a thermal shock. It consists of heating the joints to $+160°F$ and chilling them rapidly by

immersion in liquid nitrogen. This test proved effective in causing an increased rate of aging. Fig. 10 shows the accelerated aging curve of polyethylene-insulated 19-gauge copper conductor joints as compared to pulp-insulated conductor joints. Correlation of this and natural aging data helps to estimate the number of cycles representing the expected service life. Relationship between number of temperature cycles of $-320°$ to $+160°F$ and hours of natural aging for pulp- and polyethylene-insulated cable conductors is shown in Fig. 11. These curves indicate that 1 year of natural aging is approximated by something in the order of 50 cycles, and that the equivalent natural aging time increases as the cube of the number of temperature cycles. Assuming this correlation exists for future time, extrapolation of these curves shows that 40 years natural aging is approximated by some 200 cycles.

Since the temperature changes in natural environment are more gradual, the connectors will not be exposed to thermal shocks of the magnitude present in the first temperature cycling test. To represent actual temperature extremes in the field, the $-40°$ to $+140°F$ tempera-
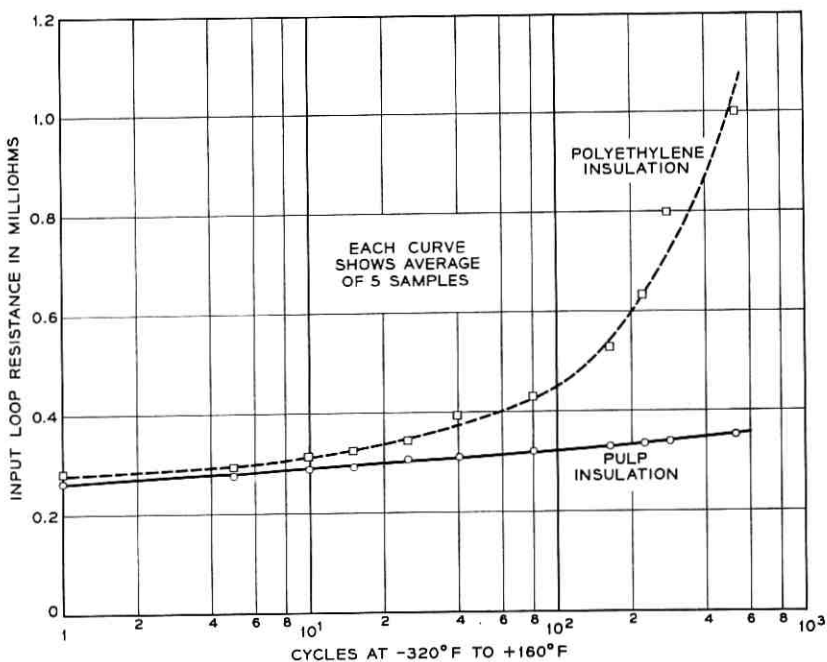
Fig. 10 — Liquid nitrogen temperature cycling test results showing the acceleration of aging for polyethylene-insulated conductors as compared to pulp.
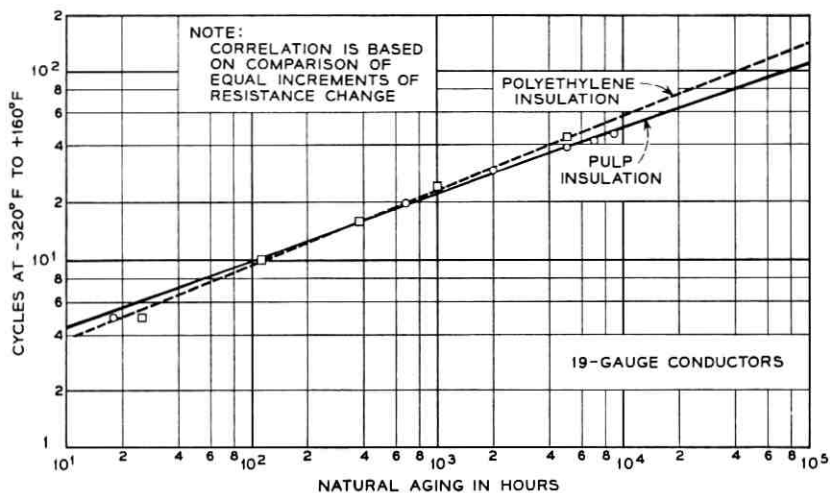
Fig. 11 — Correlation of liquid nitrogen cycles to natural aging.

ture cycle was chosen. The smaller temperature range has the advantage of simulating the natural environment more closely than the thermal shock test and, therefore, is being used exclusively for the final evaluation of plastic-insulated conductor joints. The experimental correlation obtained between natural aging and cycles of temperature cycling between $-40°F$ to $+140°F$ is shown in Fig. 12. Early in the test fewer cycles are required, but the straight-line relationship observed between
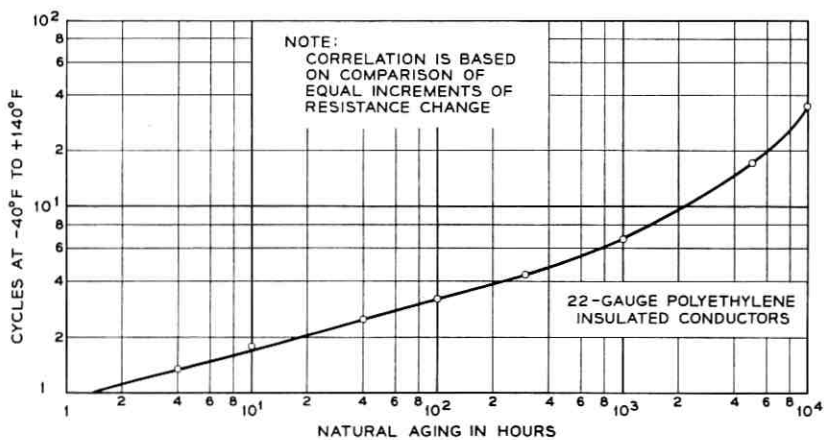
Fig. 12 — Correlation of $-40°F$ to $+140°F$ temperature cycles to natural aging.

liquid nitrogen cycles and natural aging does not hold for the longer −40°F to +140°F temperature cycles. The explanation for this deviation is not known, but it is suggested that presence of nitrogen within the connector may exclude oxygen or other corrosive gases from the critical contact regions of the joint; also, the considerably longer cycle time for the −40°F to +140°F test may have a bearing on this deviation — about 5 cycles per day against 50 cycles for the liquid nitrogen to +160°F test.

An important factor contributing to aging and not included directly in the above tests, is corrosion. Tests where joints are subjected to corrosive atmospheres at constant temperature do not simulate the "breathing" that occurs between contact surfaces due to uneven expansion and contraction of contact members with changing temperature. Calibration of room temperature corrosion tests based on corrosion film thickness may therefore be misleading, because the penetration of the corrosion products between contact surfaces during the short exposure time is likely to be insignificant.

To simulate the "breathing" of corrosive gases between contact surfaces, temperature cycling in the presence of corrosive atmosphere has been tried with reasonable success. This test approaches more nearly a true accelerated aging test because, in addition to forcing a mechanical working of the contacts, it accelerates the corrosion process. Ideally, the same degree of acceleration should be used for the two mechanisms. Attempts to correlate the aging of corrosion-cycled B Wire Connectors and natural aging have not been successful to date. The heavy corrosive films that form in a short period of time interfere with the test-set probes and also significantly change the resistance of the conductor used as a "standard" in the bridge measurement. The test has been useful, however, for comparison testing of different types of connectors.

As indicated above, stress relaxation of the soft copper conductor at the various contact points in the B Wire Connector is a possible aging mechanism. Fundamental experiments have established that relaxation does occur in stressed copper structures and that it is possible to accelerate this process with elevated temperature exposure. For example, previous research has shown that 3 hours at 175°C is equivalent to the stress relaxation during 40 years of natural time.[1] This test when used on the B Wire Connector indicated that the connector contacts were not affected. Although encouraging, we do not consider that this test necessarily approximates 40 years relaxation in the connector, since the elevated temperature may well have changed the true force pattern significantly. For instance, in polyethylene-insulated conductor joints

the plastic insulation softens and flows within the connector, changing the internal force distribution.

The philosophy used in applying accelerated tests to the B Wire Connector, and attempts to correlate these with natural aging, are based on correlating equal increments of change in resistance rather than on correlation of a known mechanism, such as stress relaxation, to expected service life. The aging mechanisms are complex, but observations of the resistance behavior itself with time is believed to be more meaningful in this case than a fundamental analysis of the contact members themselves. About a year is required to establish with reasonable accuracy the correlation between accelerated and natural aging relationships using equal increments of resistance change as the criterion. This correlation has been made possible by the development of the precise measuring techniques already described. In any of these accelerated tests, however, a certain degree of engineering risk is taken when extrapolation to long periods of time is made.

5.4 *Reliability Appraisal*

As discussed earlier, the reliability and life requirements for a connector to be used in cable splicing are very demanding indeed. Considerable effort has been expended to determine if the final design of the B Wire Connector meets these requirements. Statistical techniques have been used to predict the performance of large numbers of connections based on the performance of a relatively few connectors. The validity of these predictions depends on the accuracy with which one can define the nature of the distribution which describes the population of the connections and on how this distribution changes with time. In studying the resistance data for the B Wire Connector, it appears that the log normal distribution provides a fairly good fit. Analysis of contact resistance data from simple configurations, such as the classical crossed cylinder, supports this conclusion. In this respect, our conclusions are not directly comparable to those reached in the evaluation of the solderless wrapped connections, where a Poisson distribution is assumed.[2] It should be noted that in the present work we are describing the absolute value of the resistances with a log normal distribution, whereas in the work on the solderless wrapped connection, the degree of resistance instability with mechanical disturbance was described with a Poisson distribution.

The reliability of the connector after 40 years of aging in the outside plant environment is of prime importance. As a starting point we have established from resistance measurements on a few hundred connections

the initial reliability of the connectors. From the log-probability plot shown in Fig. 13, it would appear that less than one connection in 100,000 would have an input "loop" resistance as great as two milliohms when first made. As discussed previously, the loop measurement, in addition to measuring contact resistance, also includes the resistance of the conductors that are within the connector. Due to the complexity of the contact configuration, it is difficult to determine precisely what portion of the loop resistance is wire resistance. However, from a study of the geometry, the maximum and minimum possible wire resistances within the connector can be calculated, and are 1.73 and 0 milliohms, respectively, for 24-gauge conductor joints. In practice the wire resistance will always be greater than zero, or conversely, the pure contact resistance will always be less than that plotted in Fig. 13. To this extent, the reliability prediction in terms of pure contact resistance would be better than that quoted above.

Although it may be somewhat premature, we believe predictions of the future reliability of these connections can be made with reasonable confidence based on the early test results of the final design. From resistance measurements during a 7000-hour period on a small sample of connections, aging characteristics for the mean loop resistance and standard deviations have been established, as shown in Fig. 14. From the extrapolation of these plots, it is predicted that after 40 years the mean loop resistance and standard deviation will be 0.780 and 0.110
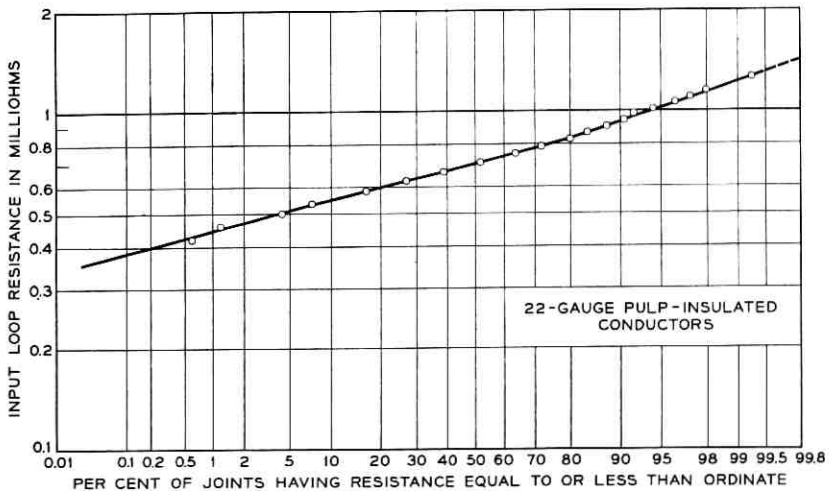
Fig. 13 — Initial resistance distribution of 100 B Wire Connector joints.
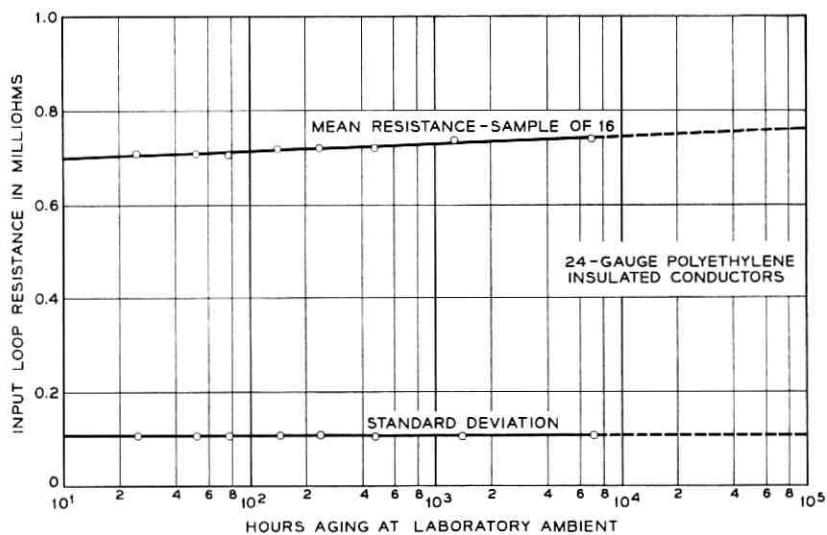
Fig. 14 — Performance of B Wire Connector joints under natural aging.

milliohm, respectively. The distribution of resistance of these connections initially and after 7000 hours of aging at laboratory ambient are shown in Fig. 15. From the predicted values of the mean resistance and standard deviation obtained from Fig. 14, a probability plot showing the distribution after 40 years has been constructed, as shown in Fig. 15.
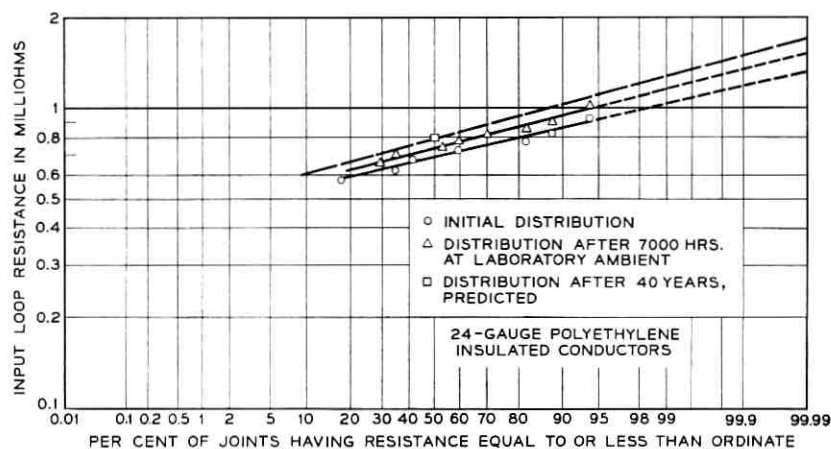
Fig. 15 — Predicted changes in the distribution of B Wire Connector joints with aging.

From the extrapolation of the log-probability plots, it is predicted that about 1 connection in 10,000 can be expected to have a loop resistance as great as 1.3 milliohms initially, and 1.5 and 1.7 milliohms after approximately 1 year and 40 years, respectively.

The tests referred to in this paper are only a few of the many that have been made to optimize the design and study the performance of the B Wire Connector. The other tests include studies made with connectors having only a few penetrators to downgrade the connector performance and make the resistance more sensitive to contact degradation, connectors made with reduced holding forces as a result of thin outer shells, and connectors without tinned liners to accelerate oxidative effects at the critical contact areas. This whole body of test results supports the reliability estimates given here. It is believed that all of the important aging mechanisms have been identified and sufficiently taken into account in the design. The device has been released for field use and so far the inherent reliability is at least as good as expected.

VI. CONCLUSION

The engineering risks involved in appraising a device design for very high reliability and long service life can probably never be reduced to zero. The techniques described here are believed to have minimized these risks in the B Wire Connector, however, and at the same time they have permitted substantial reductions in the laboratory effort required to develop and prove-in this device. Early field experience is considered good confirmation of the basic statistical predictions of initial reliability that were made. Final confirmation of the aging predictions must, of course, await the passage of time.

The experimental techniques described are based largely on the use of laboratory measurements of improved sensitivity which permit the collection of data suitable for more rigorous statistical and experimental analysis than otherwise possible. The data analysis itself has two distinct goals: (1) the study of the fundamental mechanisms which influence the physical behavior of the device, and (2) the formulation of reliability and aging estimates based on a minimum of experimental data through the use of various established statistical and other analytical techniques. In this broad sense, then, the general principles employed in the development of the connector are not new. The degree to which these basic principles have been applied in the past to the design of connectors has been limited, at least in part, by the experimental difficulties in making resistance measurements on actual devices to a sensitivity of much better than ± 0.5 milliohm. The measuring circuit and apparatus discussed

here are simple and permit an improvement in resolving power or refinement in sensitivity by at least two orders of magnitude. The work on the B Wire Connector clearly shows that important information is lost without this refinement. Many of the specific methods presented here are directly applicable to a broad class of engineering problems having to do with electrical contacts.

VII. ACKNOWLEDGMENTS

REFERENCES

1. McRae, J. W., Mallina, R. F., Mason, W. P., Osmer, T. F., and Van Horn, R. H., Solderless Wrapped Connections, B.S.T.J., **32,** May, 1953, p. 523.
2. Elliott, S. J., Evaluation of Solderless Wrapped Connections for Central Office Use, B.S.T.J., **38,** July, 1959, p. 1033.
3. Mason, W. P., and Anderson, O. L., Stress Systems in the Solderless Wrapped Connection and their Permanence, B.S.T.J., **33,** Sept., 1954, p. 1093.

The Fabry-Perot Electrooptic Modulator

by E. I. GORDON and J. D. RIGDEN

The Fabry-Perot modulator, consisting of Fabry-Perot etalon plates separated by an electrooptic material such as KDP, is analyzed in detail. Time-dependent perturbation theory is used to describe the coupling of the axial modes by spatial and time varying perturbations in the dielectric constant. The perturbations are produced by the applied microwave modulating field. It is shown that the correct choice of the spatial variation of the microwave modulating field is essential to achieve efficient modulation and the choice is equivalent to matching the phase velocities of the microwaves and the light.

Power requirements, heating, and bandwidth are discussed and a comparison is made to the traveling-wave modulator described by Kaminow.

Calculations indicate that bandwidths of several hundred megacycles, centered at any microwave frequency, can be obtained with the expenditure of several watts of modulating power.

I. INTRODUCTION

Light modulators, operating at microwave frequencies, have been receiving considerable attention. Recently, Kaminow,[1] using the Pockel's effect in KDP, produced usable amounts of modulation at an X-band frequency. The modulator was operated on a pulse basis because of the large power dissipated per unit volume in the KDP. One solution for the problem of heat dissipation as discussed by Kaminow[1] and others,[2] requires careful matching of the phase velocity of the microwaves to the light velocity in the electrooptic material, and calls for the construction of a long modulator. The resulting large coherence volume decreases the power dissipation to a value low enough to allow continuous operation.

The mechanical difficulties, as well as the transmission loss associated with long rods, among other problems, has made the realization of such low power modulators impractical. On the other hand, the advent of sources of monochromatic light energy, such as the optical maser, provides alternatives not available with conventional light sources.
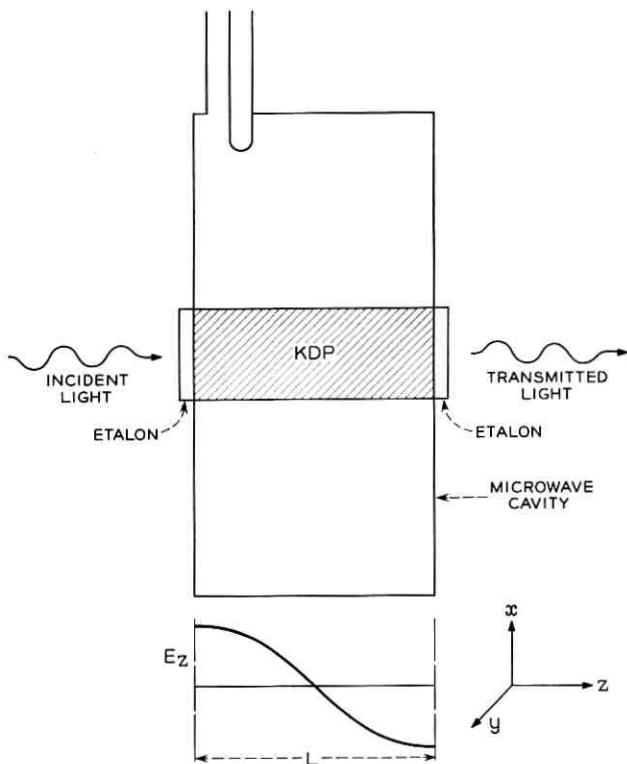
Fig. 1 — Schematic of Fabry-Perot modulator.

We refer, in particular, to resonance or cavity techniques analogous to those employed at microwave frequencies. Here, the coherence volume of the material is made effectively larger by virtue of the multiple reflections within the resonant cavity. The particular technique which we wish to describe is the use of a Fabry-Perot etalon (FPE) in which is included an electrooptic material, as shown in Fig. 1. Although the bandwidth over which modulation can be accomplished is limited by the use of an optical cavity, a bandwidth of several hundred megacycles centered at any microwave frequency should be readily attainable.

Before proceeding with an analysis of the modulator it is worthwhile to consider some simple points of view which indicate the correctness of this approach. The relative transmission I_t through a lossless FPE can be written[3]

$$I_t = [1 + 4\Gamma^2(1 - \Gamma^2)^{-2} \sin^2 \varphi]^{-1} \qquad (1)$$

in which Γ^2 is the power reflectivity of the reflectors and φ is the phase shift for a single pass. The phase shift, for normal incidence, can be written $\varphi = 2\pi f L \mu / c + \theta$, in which f is the optical frequency, L is the etalon spacing, c is the vacuum light velocity, μ is the index of refraction and θ is the phase shift upon reflection.[3] It should be possible to construct a modulator in which L, θ or μ is varied, thereby varying φ and modulating the transmitted intensity. We shall be concerned with variations in μ through the electrooptic effect. In particular we will be interested only in phase shifts associated with volume effects rather than the surface effects associated with variations in reflectivity. The latter are many orders of magnitude less significant.

In the absence of modulation, the transmission characteristic as a function of frequency is given by (1) and is shown in Fig. 2. The transmission peaks occur at frequencies $f_a = ac/2L\mu$ for which $\varphi = a\pi$. The integer a is the number of half wavelengths in the cavity. The frequency difference between adjacent peaks is $c/2L\mu$. For the purpose of discussion assume that the frequency of the incident light corresponds to one of the transmission peaks. Assume also that the index of refraction has a varying component that takes the form of a cosinusoidal standing wave as shown in Fig. 3(a), with a frequency $\omega_m/2\pi = c/2L\mu$ and wavelength $2L$. The variation is equivalent to oppositely traveling waves with phase velocity equal to the light velocity, c/μ. With a peak variation $\delta\mu \cos \omega_m t$, the amplitude of each wave is $\frac{1}{2}\delta\mu$. Under these circumstances, the perturbed phase shift travels in synchronism with the light producing a cumulative phase shift,

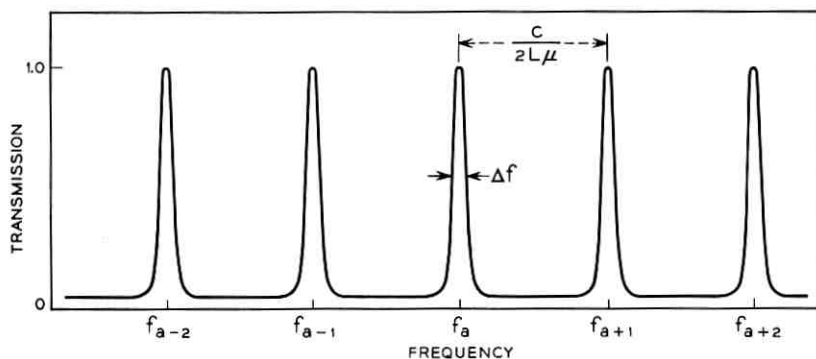$$\delta\varphi = \pi f_a L c^{-1} \delta\mu \cos \omega_m t = \frac{1}{2} a\pi (\delta\mu/\mu) \cos \omega_m t.$$

Fig. 2 — Transmission vs frequency for a Fabry-Perot etalon.

Since $\delta\mu/\mu$ is of order 10^{-6} for practical cases, and $a\pi$ is of order 10^{5}, it follows that $\delta\varphi < 10^{-1}$ and (1) can be written

$$I_t \approx [1 + \Gamma^2(1 - \Gamma^2)^{-2}(a\pi\delta\mu/\mu)^2 \cos^2 \omega_m t]^{-1}. \tag{2}$$

Choosing the reflectivity $\Gamma^2 = 0.9$ indicates that the coefficient of $\cos^2 \omega_m t$ can easily achieve the value unity, leading to large intensity modulation. Note that the modulation rate is twice the modulating frequency, which is consistent with the fact that the transmitted intensity is an even function of the induced phase shift.

From (1) the separation of half-power frequencies or pass band of the FPE, shown as Δf in Fig. 1, is given by

$$\Delta f \approx f_a(1 - \Gamma^2)/a\pi\Gamma. \tag{3}$$

For a one centimeter long etalon, with $\Gamma^2 = 0.9$, and $c/2L\mu = f_a/a = 10^{10}$, $\Delta f \approx 300$ mc/s. One wonders whether this represents a limiting rate at which the light can be modulated, since the pass band Δf represents the maximum rate at which the stored energy in the cavity can be varied. When the perturbation in dielectric constant is uniform across the etalon, it will be shown that the maximum possible modulation rate is of order $\Delta f/2$. However, when the phase velocities are matched as described above, there is no limiting modulation rate.

We can obtain an intuitive appreciation of the lack of a limiting modulation rate by considering Fig. 3. The perturbed part of the dielectric constant ϵ, shown as $\delta\epsilon$ in Fig. 3(a), is depicted as a standing wave. This standing wave corresponds to the microwave modulating electric field which is assumed to have a field component normal to the reflectors. This component produces the variation $\delta\epsilon$. The optical field, excited by the incident light, is shown as the standing wave, E_a in Fig. 3(b). In addition to the displacement current, $\epsilon\partial E_a/\partial t$, there will be a perturbed displacement current, $\partial\delta\epsilon E_a/\partial t$, shown in Fig. 3(c). It can be seen that the perturbed displacement current has the appropriate spatial distribution to excite the modes $E_{a\pm1}$ shown in Fig. 3(d) and 3(e).

The microwave modulating frequency $f_m = \omega_m/2\pi$ is chosen to have the value $c/2L\mu$, so that the sum and difference frequencies appearing in the time dependence of the perturbed part of the displacement current correspond to the frequencies, $f_{a\pm1}$. As a result, components of the perturbed part of the displacement current have the proper spatial variation and frequency to excite the sideband modes. The transmitted light will then have upper and lower sideband frequencies which result from the oscillating sideband modes.
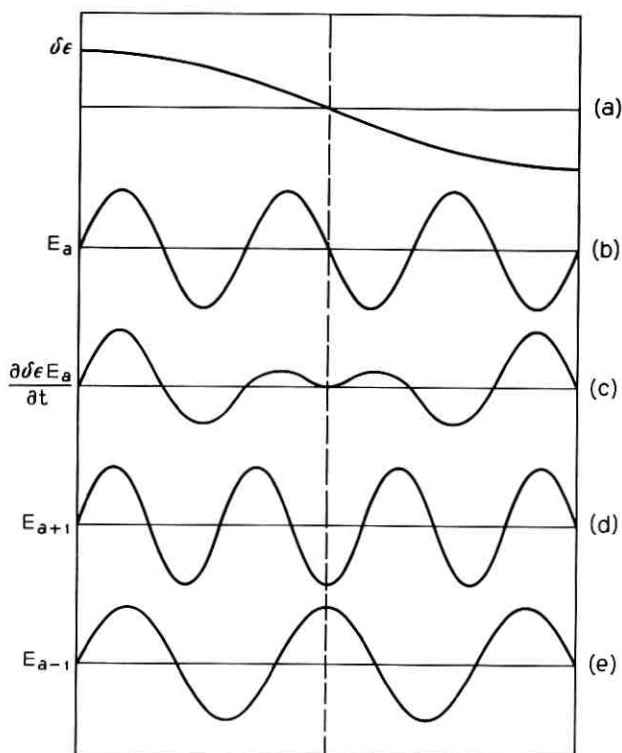
Fig. 3 — Mode coupling resulting from perturbations in the dielectric constant.

From the point of view of coupling of modes, there is no inherent limitation to the rate at which modulation can be accomplished, except that afforded by the electrooptic effect itself. However, the range over which the microwave modulating frequency can be varied must correspond to the bandwidth of the FPE modes, otherwise the sideband modes cannot be excited strongly.

The conditions under which modulation is possible can be summarized by the equations

$$f_{a\pm m} = f_a \pm f_m$$
$$\beta_{a\pm m} = \beta_a \pm \beta_m \tag{4}$$

in which β_n is the propagation constant of the nth mode, in this case f_m and β_m are also the microwave frequency and propagation constant. Equation (4) is recognizable as the Tien $\omega - \beta^4$ relation for reactive

frequency mixing. Equation (4) also implies that the microwave phase velocity equals the light phase velocity, assuming negligible dispersion.

It is not clear yet that the variation in transmitted intensity given by (2), which follows from relatively simple considerations, is consistent with the coupled mode picture. The next sections will be devoted to an analysis of the modulator using time-dependent perturbation theory to determine the transmitted field amplitudes and phases. The variation in transmitted intensity will be computed and shown to be consistent with (2). Other modes of operation will be described which enhance the modulation efficiency and which can yield linear intensity modulation at the microwave frequency. Power requirements and heating will be considered for an X-band modulator for light at 6328 Å.† Some of the optical requirements of the Fabry-Perot cavity will be discussed with respect to available maser beams. Finally a comparison will be made to the traveling-wave modulator described by Kaminow.

II. THEORY

The analytical description will be similar to that used in time-dependent perturbation theory. The optical fields within the FPE will be expanded in a set of eigenfunctions appropriate to perfectly reflecting boundaries. It will be seen that the expansion coefficients are time dependent and satisfy "harmonic-oscillator-like" linear differential equations which are coupled by the perturbed part of the dielectric constant. The mixing or coupling terms involve integrals over the coupled modes and the perturbed part of the dielectric constant.

The microwave fields will not explicitly enter into the calculation; rather they will serve only to produce the perturbed part of the dielectric constant. Although the dielectric constant in electrooptic materials often exhibits tensor properties we will suppress this feature since the notation would be burdened without any real gain in generality.

The basis of the calculation will be the normal mode formulation given by Slater.[6] In the absence of the perturbation in the dielectric constant the electric and magnetic fields within the FPE are divergence-free. Thus, we will use only solenoidal eigenfunctions which are defined in the volume V appropriate to the FPE by the equations

$$k_a \mathbf{E}_a(\mathbf{r}) = \text{curl } \mathbf{H}_a(\mathbf{r}), \qquad k_a \mathbf{H}_a(\mathbf{r}) = \text{curl } \mathbf{E}_a(\mathbf{r}) \qquad (5)$$

in which k_a is the eigenvalue of the ath mode. A consistent set of boundary conditions on $\mathbf{E}_a$ and $\mathbf{H}_a$ are taken to be

---

† As the modulation efficiency increases with decreasing wavelength, we have used the wavelength of the presently shortest wavelength maser.[5]

$$\mathbf{n} \times \mathbf{E}_a = 0, \quad \mathbf{n} \cdot \mathbf{H}_a = 0 \quad \text{on} \quad S_s$$
$$\mathbf{n} \times \mathbf{H}_a = 0, \quad \mathbf{n} \cdot \mathbf{E}_a = 0 \quad \text{on} \quad S_o \tag{6}$$

in which $\mathbf{n}$ is a unit vector normal to the outer surface of the volume V. The surface S_s corresponds to the short circuit boundary while S_o is the open circuit boundary; together they represent the total surface of the volume V. Under these conditions, Slater shows that the functions $\mathbf{E}_a$ and $\mathbf{H}_a$ represent an orthogonal normalized set satisfying the equations

$$\nabla^2\mathbf{E}_a + k_a^2\mathbf{E}_a = 0 \quad \text{and} \quad \nabla^2\mathbf{H}_a + k_a^2\mathbf{H}_a = 0.$$

The orthogonality relations take the form

$$\int_V \mathbf{E}_a \cdot \mathbf{E}_b dV = \int_V \mathbf{H}_a \cdot \mathbf{H}_b dV = \mathcal{E}_{ab}$$

The starting point for the analysis is the Maxwell equations

$$\operatorname{curl} \mathbf{E} = -\mu_0 \partial \mathbf{H}/\partial t$$
$$\operatorname{curl} \mathbf{H} = \partial \epsilon' \mathbf{E}/\partial t + \sigma \mathbf{E} \tag{7}$$

in which σ is the volume conductivity associated with the transmission loss of the electrooptic material, μ_0 is the vacuum permeability and ϵ' is the permittivity. As shown by Slater the various terms in (7) have expansions of the form

$$\mathbf{E}(\mathbf{r}, t) = \sum_a e_a(t)\mathbf{E}_a(\mathbf{r})$$

$$\mathbf{H}(\mathbf{r}, t) = \sum_a h_a(t)\mathbf{H}_a(\mathbf{r})$$

$$\operatorname{curl} \mathbf{E} = \sum_a \mathbf{H}_a \left[k_a e_a + \int_{S_s} (\mathbf{n} \times \mathbf{E}) \cdot \mathbf{H}_a dS \right] \tag{8}$$

$$\operatorname{curl} \mathbf{H} = \sum_a \mathbf{E}_a \left[k_a h_a + \int_{S_o} (\mathbf{n} \times \mathbf{H}) \cdot \mathbf{E}_a dS \right]$$

in which dS is an element of the surface. The mode amplitudes e_a and h_a are defined by

$$e_a(t) = \int_V \mathbf{E}(\mathbf{r}, t) \cdot \mathbf{E}_a dV$$

$$h_a(t) = \int_V \mathbf{H}(\mathbf{r}, t) \cdot \mathbf{H}_a dV \tag{9}$$

in which dV is a volume element. Substituting the relations in (8) into

(7) and using the orthogonality and normality of the functions $\mathbf{E}_a$ and $\mathbf{H}_a$ yields equations for the coefficients of the form

$$\mu_0 \partial h_a / \partial t + k_a e_a = - \int_{S_s} (\mathbf{n} \times \mathbf{E}) \cdot \mathbf{H}_a dS \tag{10}$$

and

$$\epsilon \partial e_a / \partial t + \sigma e_a - k_a h_a - \int_{S_g} (\mathbf{n} \times \mathbf{H}) \cdot \mathbf{E}_a dS \tag{11}$$

$$+ \partial \sum_b \langle a \mid \delta\epsilon \mid b \rangle e_b / \partial t = 0$$

in which we have written $\epsilon' = \epsilon + \delta\epsilon$, ϵ being the unperturbed part of the dielectric constant, and we have used the notation

$$\langle a \mid f(\mathbf{r}) \mid b \rangle = \int_V \mathbf{E}_a \cdot \mathbf{E}_b f(\mathbf{r}) dV. \tag{12}$$

Iterating (10) and (11) yields

$$c'^{-2} \partial^2 e_a / \partial t^2 + k_a^2 e_a + \mu_0 \sigma \partial e_a / \partial t + c'^{-2} \partial^2 \sum_b \langle a \mid \delta\epsilon/\epsilon \mid b \rangle e_b / \partial t^2$$

$$= \mu_0 \partial \left(\int_{S_g} (\mathbf{n} \times \mathbf{H}) \cdot \mathbf{E}_a dS \right) \Big/ \partial t - k_a \int_{S_s} (\mathbf{n} \times \mathbf{E}) \cdot \mathbf{H}_a dS \tag{13}$$

in which $c' = (\mu_0 \epsilon)^{-\frac{1}{2}}$ is the unperturbed velocity of light in the dielectric. A similar redundant equation obtains for h_a. The surface integrals contain two types of field terms: (i) external optical fields incident on the cavity as they appear on the inside boundary wall of the reflector, which serve to excite the FPE modes, and (ii) internal fields which account for energy lost by radiation, in particular the transmitted beam.

In the following treatment of the modulator we will assume that the reflecting surfaces of the etalon can be characterized by a reflection coefficient for field amplitude Γ and that the surface is a "short," i.e., a surface S_s. The fields will be assumed to be parallel to the etalon which is of sufficient lateral extent that the axial modes of the FPE are essentially TEM waves with no diffraction loss. The dielectric modulation is assumed to be uniform across the cross section. Later we will consider nonuniform excitation of the dielectric in the finite cross section, and it will be seen that the lowest-order axial modes can be coupled to the off-axis higher-order modes.

Under these assumptions, the electric and magnetic field on the inside face of the reflectors constituting the surface S_s can be shown to

be related by

$$\mathbf{n} \times \mathbf{E} = c'\mu_0 \mathbf{H}(1 - \Gamma)/(1 + \Gamma) \tag{14}$$

neglecting the extremely small variation in Γ produced by the variation of the dielectric constant within the etalon. First we consider the loss terms in the cavity. For the right-hand side of (13) we may write, using (14) and (8)

$$-k_a \int_{S_s} (\mathbf{n} \times \mathbf{E}) \cdot \mathbf{H}_a dS$$

$$= -k_a c' \mu_0 (1 - \Gamma)(1 + \Gamma)^{-1} \sum_b h_b \int_{S_s} \mathbf{H}_b \cdot \mathbf{H}_a dS \tag{15}$$

there being no surface S_o over which $\mathbf{n} \times \mathbf{H} \cdot \mathbf{E}_a$ is nonzero. Substituting (11) into (15) to eliminate the term in h_b yields

$$-k_a \int_{S_s} (\mathbf{n} \times \mathbf{E}) \cdot \mathbf{H}_a dS = -k_a c'^{-1}(1 - \Gamma)(1 + \Gamma)^{-1}$$

$$\times \sum_b k_b^{-1} \left[\partial e_b/\partial t + \sigma e_b/\epsilon + \partial \sum_d \langle b \mid \delta\epsilon/\epsilon \mid d\rangle e_d/\partial t \right] \tag{16}$$

$$\int_{S_s} \mathbf{H}_b \cdot \mathbf{H}_a dS.$$

For the axial modes of the FPE, the appropriate normalized eigenfunctions are given by

$$H_a = (2/AL)^{\frac{1}{2}} \cos k_a z$$

$$E_a = (2/AL)^{\frac{1}{2}} \sin k_a z \tag{17}$$

in which A is the cross-sectional area, L the length of the modulator, and $k_a = \pi a/L$. The integer a is the number of half wavelengths within the etalon. Remembering that there are two faces of the etalon composing the surface S_s, we note that the integral

$$\int_{S_s} \mathbf{H}_b \cdot \mathbf{H}_a dS = 4/L \tag{18}$$

when b differs from a by an even number, while the integral has the value zero when b differs from a by an odd number. The implication is that the surface loss terms couple only modes of the same longitudinal symmetry. The integral is zero when H_b is one of the off-axis FPE modes because of the orthogonality. Defining

$$Q_a = \tfrac{1}{4}\pi a(1 + \Gamma)^2/(1 - \Gamma^2) \tag{19}$$

(16) becomes

$$
-k_a \int_{S_s} (\mathbf{n} \times \mathbf{E}) \cdot \mathbf{H}_a dS
$$
$$
= -k_a c'^{-1} \sum_b{}' Q_b^{-1} \left[\partial e_b/\partial t + \sigma e_b/\epsilon + \partial \sum_d \langle b \mid \delta\epsilon/\epsilon \mid d\rangle e_d/\partial t \right] \tag{20}
$$

in which the prime implies that only those values of b which differ from a by an even number are to be included in the sum. The third term in the square bracket of (20) is of order $\delta\epsilon/\epsilon$ compared to first. The second term is of order $\sigma/\epsilon\omega_a \sim 10^{-6}$ compared to the first. Since the first term is as small as any other term in (13) because of the factor $Q_b^{-1} \sim 10^{-6}$, we can safely neglect the other terms. In the absence of internal losses the term Q_a is known as the quality factor of the cavity. It defines the resolution of the FPE as an interferometer and defines the bandwidth by the relation $\Delta f \approx f_a/Q_a$ which differs from the bandwidth defined in (3) by the factor $4\Gamma/(1 + \Gamma)^2 \approx 1$.†

Considering the incident wave amplitude as transformed to the inside surface of the reflector it can be shown in similar fashion‡ that

$$
-k_a \int_{S_s} (\mathbf{n} \times \mathbf{E}) \cdot \mathbf{H}_a dS = -\tfrac{1}{2}k_a c'^{-1} Q_a^{-1} \partial e/\partial t \tag{21}
$$

in which e is proportional to the magnetic field of the incident wave on the surface of the reflector. The factor $\tfrac{1}{2}$ appears because the incident fields appear only on one face. Thus (13) becomes

$$
\partial^2 e_a/\partial t^2 + \omega_a^2 e_a + \omega_a Q_{ad}^{-1}\partial e_a/\partial t
$$
$$
+ \omega_a \sum_b{}' Q_b^{-1}\partial e_b/\partial t + \partial^2 \sum_b \langle a \mid \delta\epsilon/\epsilon \mid b\rangle e_b/\partial t^2 \tag{22}
$$
$$
= -\tfrac{1}{2}\omega_a Q_a^{-1}\partial e(\omega t)/\partial t
$$

in which $\omega_a = k_a c'$ is the resonant frequency of mode a and $Q_{ad} = \omega_a\epsilon/\sigma$ is the dielectric quality factor or inverse loss tangent. An alternate expression can be written $Q_{ad} \approx \omega_a \ell/c'$ in which ℓ is the distance for which the energy decays to e^{-1}.

Next, we evaluate the quantities $\langle a \mid (\delta\epsilon/\epsilon) \mid b\rangle$. Since the spatial variation of $\delta\epsilon/\epsilon$ can take the form of $\sin m\pi z/L$ or $\cos m\pi z/L$ in which m is some integer, which corresponds to the appropriate microwave

---

† The factor $4\Gamma/(1 + \Gamma)^2$ differs from unity by less than one per cent for $\Gamma^2 > 0.7$.
‡ The mirror acts as a transformer with turns ratio $(1 + \Gamma)^{\frac{1}{2}}/(1 - \Gamma)^{\frac{1}{2}}$.

field variation, we evaluate two types of coupling terms:

$$\left\langle a \left| \begin{matrix} \sin \pi mz/L \\ \cos \pi mz/L \end{matrix} \right| b \right\rangle$$

$$= (2/L) \int_0^L \left(\sin \pi az/L \, \frac{\sin \pi mz/L}{\cos \pi mz/L} \, \sin \pi bz/L \right) dz \tag{23}$$

in which we have used (17). Performing the integration yields

$$\langle a \mid (\sin \pi mz/L) \mid b \rangle$$

$$= \frac{1}{2} \left[\frac{[1 - \cos (a - b + m)\pi]}{(a - b + m)\pi} + \frac{[1 - \cos (b - a + m)\pi]}{(b - a + m)\pi} \right. \tag{24}$$

$$\left. + \frac{[1 - \cos (a + b - m)\pi]}{(a + b - m)\pi} - \frac{[1 - \cos (a + b + m)\pi]}{(a + b + m)\pi} \right].$$

Terms like $(a + b \pm m)^{-1}$ will be of order 10^{-5} and the last two terms can be neglected. When $a - b \pm m = 0$ or an even integer the remaining terms vanish. When $a - b \pm m$ is an odd integer the terms are finite, yielding

$$\langle a \mid \sin \frac{m\pi z}{L} \mid b \rangle = -\frac{2m/\pi}{(a - b)^2 - m^2}, \tag{25}$$

$$a - b \pm m = \text{odd integer}$$

the largest values corresponding to $a = b$, $m = 1$ with a value $2\pi^{-1}$. Likewise,

$$\langle a \mid \cos \frac{m\pi z}{L} \mid b \rangle = \frac{1}{2} \left[-\frac{\sin (a + b - m)\pi}{(a + b - m)\pi} + \frac{\sin (a - b - m)\pi}{(a - b - m)\pi} \right.$$

$$\left. + \frac{\sin (a - b + m)\pi}{(a - b + m)\pi} - \frac{\sin (a + b + m)\pi}{(a + b + m)\pi} \right] \tag{26}$$

$$= \tfrac{1}{2} \quad \text{for} \quad |a - b| = m.$$

The sine and cosine cases are mutually exclusive. Many modes are coupled by the sine variation since for a given value of a and m there are multiple values of b satisfying (25). There are only two values of b satisfying (26) for the cosine variation. Note that for $m = 0$, corresponding to uniform perturbation of the dielectric, the only coupling is for $a = b$. That is, the modulation can occur only within the mode. This justifies our earlier statement that the modulation rate is restricted to the bandwidth of the FPE mode for uniform excitation.

The cosine distribution yields the largest coupling terms except for

the case when $a = b$ and $m = 1$. The case $m = 1$ implies that the sine variation has a space average part and it is this part which produces the coupling for $a = b$. This is borne out by the fact that for $a = b$, m must be odd and the strength of the coupling decreases with increasing m. Notice that even though the modes of interest are standing waves these results correspond precisely to that predicted by the β condition of (4). The factor of $\frac{1}{2}$ for the cosine distribution arises because only half the amplitude of the standing wave is effective for a given traveling-wave component. The factor $2\pi^{-1}$ for the sine variation can be shown to arise from the fact that although the β condition is not precisely satisfied, the interaction with each traveling-wave component is down by π^{-1}.

In the next sections, we apply these results to calculate the amplitude of the modulation sidebands, their bandwidth, and the intensity modulation of the light.

III. AMPLITUDE OF THE MODULATION SIDEBANDS

Assuming that we have used the appropriate $\cos m\pi z/L$ distribution for that component of the microwave field which varies ϵ', since we want large coupling for $a \neq b$, we can write (22) as

$$\frac{\partial^2}{\partial t^2} e_a + \omega_a{}^2 e_a + \omega_a Q_{ad}{}^{-1} \frac{\partial e_a}{\partial t} + \frac{1}{2} \frac{\partial^2}{\partial t^2} \frac{\delta\epsilon(\omega_m t)}{\epsilon} [e_{a+m} + e_{a-m}]$$
$$+ \omega_a \sum_b{}' Q_b{}^{-1} \frac{\partial e_b}{\partial t} = -\tfrac{1}{2}\omega_a Q_a{}^{-1} \frac{\partial e(\omega t)}{\partial t} \tag{27}$$

which holds for all values of a.† We note that modes e_a and $e_{a\pm m}$ are coupled together. Since the Q of the modulator is so extremely high, it is approximately correct to say that any driving terms at frequencies outside the pass band of a given mode have negligible effect on that mode. Later we will see that for very strong modulation (a case of little practical interest) the high-Q approximation leaves a little to be desired. Hence, to good approximation, we may write for any of the modes a,

$$\sum_b{}' Q_b{}^{-1} \partial e_b/\partial t \simeq Q_a{}^{-1} \partial e_a/\partial t.$$

All the neglected terms are at frequencies well outside the passband of mode a. With these definitions, the quantity $1/Q_{aL} = 1/Q_a + 1/Q_{ad}$

---

† It is also possible to have $|a - b| = 1$, $m = 2$ yielding a coupling term $4/3\pi$ which is only slightly less than $\frac{1}{2}$. We will not consider this case which is considerably more complicated.

is properly called the loaded Q of the resonator. The bandpass in the absence of mode coupling is defined by $\Delta\omega = \omega_a/Q_{aL}$. The quantity $2Q_a$ is also properly called the external Q.

We will assume that the carrier frequency corresponds to the resonant frequency of the ath mode, $\omega = k_a c' = \omega_a$. There is no loss of generality in this assumption since the carrier frequency would normally be fixed. The microwave modulating frequency, on the other hand, must vary over some frequency range. Ideally the microwave frequency has the value $mc'/2L$, since under this circumstance the mixing terms, represented by the fourth term of (27), vary at rates corresponding to the resonant frequencies of the coupled modes. However the microwave frequency must be able to vary over some frequency range. We define the microwave angular frequency as

$$\omega_m = (m\pi c'/L) + \omega_a\delta/Q_{aL} \tag{28}$$

and the parameter δ measures the excursion from center band frequency. In view of the fact that $\omega_a/Q_{aL} \simeq \Delta\omega_a$, the half power frequencies are given approximately by $\delta = \pm\frac{1}{2}$. The $\pm n$th sideband frequencies are given by

$$\omega_a \pm n\omega_m = \omega_a \pm nm\pi c'/L \pm n\delta\omega_a/Q_{aL} = \omega_{a\pm nm} \pm n\delta\omega_a/Q_{aL}. \tag{29}$$

The pertinent mode amplitudes are defined by

$$e_{a\pm nm} = E_{a\pm nm} \exp i(\omega_a \pm n\omega_m)t + \text{complex conjugate}$$

$$\delta\epsilon/\epsilon = [\delta\epsilon/\epsilon] \exp i\omega_m t + \text{complex conjugate} \tag{30}$$

$$e = E \exp i\omega_a t + \text{complex conjugate}$$

which when substituted into (27) yield

$$-[\chi^* E_{a+m} + \chi E_{a-m}] + iE_a = -\tfrac{1}{2}iEQ_{aL}/Q_a \tag{31}$$

$$2n\delta E_{a-nm} - [\chi^* E_{a-(n-1)m} + \chi E_{a-(n+1)m}] + iE_{a-nm} = 0 \tag{32}$$

$$-2n\delta E_{a+mn} - [\chi^* E_{a+(n+1)m} + \chi E_{a+(n-1)m}] + iE_{a+nm} = 0. \tag{33}$$

The coefficients are complex quantities and the asterisk denotes complex conjugate. The modulus of the coefficients equals half the peak value and the argument represents the phase of the related real variable. The quantity

$$\chi = \tfrac{1}{2}Q_{aL}[\delta\epsilon/\epsilon] \tag{34}$$

is a "modulation index" whose modulus equals the product of Q_{aL} and one quarter of the peak variation of $\delta\epsilon/\epsilon$. The argument of χ equals the

phase of the microwave field. Note that we have set all the terms $Q_{a\pm nm}$ equal to Q_a. Reference to (19) indicates that the variation in Q_a from one mode to the next is of order $a^{-1} \cong 10^{-5}$ and thus the subscript is superfluous. Likewise, all other terms of order $\pi c'/2L\omega_a = a^{-1}$ have been neglected.

It is helpful to observe that (33) is redundant. If we define

$$E_{a+nm} = (-1)^n E_{a-nm}{}^* \tag{35}$$

and substitute into (33), we obtain (32).

Defining the normalized field amplitude

$$g_n = (i/\chi^*)^n E_{a-nm}/(\tfrac{1}{2}EQ_{aL}/Q_a) \tag{36}$$

and substituting into (31) and (32) yields

$$2 \mid \chi \mid^2 R_e g_1 + g_0 = -1 \tag{37}$$

$$(1 - i2n\delta)g_n - g_{n-1} + \mid \chi \mid^2 g_{n+1} = 0 \tag{38}$$

which is a three term recursion relation with nonconstant coefficients. Placing the question of the microwave bandwidth aside momentarily we may assume that the microwave frequency is at its appropriate center band value corresponding to $\delta = 0$ in (38). Under this condition solutions for the recursion formula, (38), of the form $g_n = b^n$, yield as the defining equation for b, $\mid \chi \mid^2 b^2 + b - 1 = 0$. The general solution may be written $g_n = Ab_+{}^n + Bb_-{}^n$ with

$$b_\pm = \left[-\frac{1}{2 \mid \chi \mid^2} \pm \sqrt{\left(\frac{1}{2 \mid \chi \mid^2}\right)^2 + \frac{1}{\mid \chi \mid^2}} \right]. \tag{39}$$

We note three things about g_n: (i) it is real, since g_0 is real and all the coefficients in the recursion formula are real; (ii) it is a function of $\mid \chi \mid^2$ and (iii), since for $n > 0$, $(E_{a-nm}/E) \to 0$ as the modulation index $\mid \chi \mid \to 0$, so must $g_n \mid \chi \mid^n \to 0$ as follows from (36).

The solution b_- has the dependence $b_-{}^n \sim \mid \chi \mid^{-2n}$ and therefore does not have the correct behavior as $\mid \chi \mid \to 0$. Thus, we choose the constant $B = 0$. The constant A is evaluated by substituting $g_n = Ab_+{}^n$ into (37), yielding $A = -(1 + 4 \mid \chi \mid^2)^{-\frac{1}{2}}$, so that we may write

$$g_n = -(\tfrac{1}{2})^n \frac{1}{\mid \chi \mid^{2n}} [\sqrt{1 + 4 \mid \chi \mid^2} - 1]^n / \sqrt{1 + 4 \mid \chi \mid^2} \tag{40}$$

and from (36)

$$E_{a-nm}/\tfrac{1}{2}E = -(i2\chi)^{-n}(Q_{aL}/Q_a)[\sqrt{1 + 4 \mid \chi \mid^2} - 1]^n$$
$$\div \sqrt{1 + 4 \mid \chi \mid^2}. \tag{41}$$

The same expression with $\chi \rightarrow \chi^*$ applies for $E_{a+nm}/\frac{1}{2}E$. We note that the amplitude of the carrier wave, $n = 0$, is given by

$$E_a/\tfrac{1}{2}E = -(Q_{aL}/Q_a)[1 + 4\,|\,\chi\,|^2]^{-\frac{1}{2}} \qquad (42)$$

and is reduced in amplitude from its value of $-Q_{aL}/Q_a$ in the no modulation limit $|\,\chi\,| = 0$. The value $E_a = -\frac{1}{2}E$ in the no modulation case, with $Q_{aL} = Q_a$ corresponding to $Q_{ad} = \infty$, is eminently reasonable. Remembering that the surface is a "short" the implication is that the magnetic field of mode a of the resonator exactly cancels out the half of the total carrier magnetic field E on the inner reflecting surface which is associated with the reflected wave of amplitude $E/2$. The incident wave also has amplitude $E/2$, yielding a total field E. Likewise the magnetic field on the opposite reflecting surface is $\pm E/2$. Thus, the incident carrier field $E/2$, which measures the incident energy, correctly accounts for the energy leaving through the opposite face, and we have 100 per cent transmission through the etalon and no reflection. In microwave parlance, the resonator is matched to the incident wave. This always occurs when the external Q equals twice the loaded Q. In the presence of internal loss, $Q_{ad} < \infty$, the resonator cannot be matched if the etalon plates have equal reflectivity. In the presence of the modulation, the carrier mode is loaded by the sidebands even without internal loss and the FPE is no longer matched to the carrier; consequently carrier energy is reflected as well as transmitted. This shows up as a reduction in transmitted power as well as imperfect cancellation of the reflected wave since $|\,E_a\,| < |\,E/2\,|$.

The relative amplitude of sideband to carrier leaving the opposite face of the etalon is given by

$$(E_{a-nm}/E_a) = (i2\chi)^{-n}[\sqrt{1 + 4\,|\,\chi\,|^2} - 1]^n. \qquad (43)$$

In the limit of small $|\,\chi\,|$

$$(E_{a-nm}/E_a)_{|\chi|\rightarrow 0} = (\chi^*/i)^n \qquad (44)$$

and the amplitude is down by $|\,\chi\,|^n$. The power in the first sideband is proportional to $|\,\chi\,|^2$ which is proportional to the microwave modulating power. For large $|\,\chi\,|$

$$(E_{a-nm}/E_a)_{|\chi|\rightarrow\infty} = (|\,\chi\,|/i\chi)^n = \exp - in(\pi/2 + \text{Arg }\chi) \qquad (45)$$

so that the modulation approaches 100 per cent. As we may note from (41), the absolute amplitude of all of the modes decreases with increasing $|\,\chi\,|$ as $(1 + 4\,|\,\chi\,|^2)^{-\frac{1}{2}}$. Thus we might expect maximum sideband energy for some finite value of χ. Maximizing E_{a-nm}/E shows

that for the nth sideband the value of $| \chi |$ for which the energy is a maximum is given by the positive real solution of

$$16 | \chi |^4 - 4n^2 | \chi |^2 - n^2 = 0.$$

For the first sideband the maximum occurs at

$$| \chi | = \tfrac{1}{2}[\tfrac{1}{2}(1 + \sqrt{5})]^{\frac{1}{2}} = 0.636.$$

The transmitted light beam contains 23 per cent of its energy in the first sideband, and in the absence of internal loss the transmitted carrier energy is reduced to 38 per cent of the original carrier energy. As will be seen later, this modulation can be accomplished with only a few watts of microwave power over a bandwidth of several hundred megacycles at X-band.

The microwave bandwidth over which modulation can be accomplished is determined from a solution of (38) with $\delta \neq 0$. The value of δ for which the energy in the first sideband falls to half its value at center band is determined by a solution of $| g_1(\delta) |^2 / | g_1(0) |^2 = \tfrac{1}{2}$. The details of this calculation are too lengthy to be included here. The calculation was performed by first noting that when the first sideband mode is driven near its band edge the second sideband mode is driven well off resonance. As a result, the values of $g_n(\delta)$ decrease rapidly as $\delta \to \tfrac{1}{2}$. It follows from (38) that setting $g_n(\delta) = 0$ is equivalent to neglecting terms of order $(2n)^{-2} \ll 1$. We assumed that $g_3 = 0$ and solved the three relations that follow from (37) and (38). The result can best be expressed in the form

$$\delta = \pm\tfrac{1}{2}(1 + 4 |\chi |^2 + \cdots) \tag{46}$$

which indicates that δ is increased over its zero modulation value of $\pm\tfrac{1}{2}$. This result is understandable if we consider that the modulation process tends to load the carrier mode and decrease its amplitude. This process is more effective when the sidebands are excited strongly. The sideband amplitude is proportional to the carrier amplitude; thus the center of the sideband modes decreases more rapidly than the sides for a given modulation index yielding, in effect, a broader modulation bandwidth. From another point of view we may say that the coupling of a given mode to adjacent modes increases the rate at which energy can be dissipated in the given mode. This decreases the effective loaded Q and increases the bandwidth.

For radiation at 6328 Å, $f_a \simeq 5 \times 10^{14}$ cps, $f_m = 10^{10}$ cps, it follows that $a = 5 \times 10^4$ and for $\Gamma^2 = 0.9$ we have $Q_a \simeq 1.5 \times 10^6$. Thus the modu-

lation bandwidth is at least 3×10^8 cps. For large values of $| \chi |$ the actual bandwidth may be considerably larger but the modulation will not be linear. In any case the high-Q approximation, which is fundamental in this analysis, is probably no longer valid and all conclusions for large $| \chi |$ must be considered as only qualitative.

IV. EFFECT OF NONUNIFORM DIELECTRIC EXCITATION

So far we have assumed that the dielectric constant variation induced by the microwave field depends only on z and t. However, it is necessary to match the microwave phase velocity to the light phase velocity. Since the dielectric constant at the microwave frequency is, in general, considerably larger than it is at the light frequency, the microwave field cannot propagate as a plane wave with no transverse variation. The transverse variation could probably be avoided by making the diameter of the Fabry-Perot cavity very small so that the microwave fields are only partially slowed by the dielectric, but this is not usually possible. As a consequence the microwave field must have a variation of amplitude across the Fabry-Perot cavity. The resultant transverse variation of the perturbed part of the dielectric constant will cause coupling of the FPE modes.

In general we must describe the ath FPE mode in terms of a series of transversely varying eigenfunctions. Thus

$$\mathbf{E}_a = \sum_{lm} A_{lm} \mathbf{R}_l(\mathbf{k}_{lm} \cdot \mathbf{r}) \sin \pi a z/L \qquad (47)$$

in which A_{lm} is a constant determined by the geometry. The frequency of the lmth mode is

$$f_{alm} = (c'/2\pi)[(\pi a/L)^2 + k_{lm}^2]^{\frac{1}{2}} = f_a[1 + k_{lm}^2/(\pi a/L)^2]^{\frac{1}{2}} \qquad (48)$$

so that each of the modes characterized by the number of half wavelengths equal to a has a different resonant frequency. However, k_{lm} is of order $m\pi/D$ in which D is some characteristic transverse dimension of the FPE. The ratio $k_{lm}/(\pi a/L)$ is of order $(m/a)\,(L/D)$. Since L/D will be of order ten or less and m/a will be order 10^{-5} for the lower-order transverse variations, we see that the frequencies f_{alm} differ from the fundamental frequency by about five parts in 10^9. For 6328 Å radiation this amounts to about three megacycles, well within the passband of the FPE.

The microwave modulating signal designed to mix adjacent FPE modes will not produce mixing of different transverse modes with the

same a number but it can produce transverse mixing between adjacent modes. The degree of mixing will depend on integrals of the sort

$$\int \mathbf{R}_l(\mathbf{k}_{lm} \cdot \mathbf{r}) \cdot (\delta\epsilon(\mathbf{r})/\epsilon) \mathbf{R}_{l'}(\mathbf{k}_{l'm'} \cdot \mathbf{r}) dS \qquad (49)$$

evaluated over the cross section of the FPE. The functions $\mathbf{R}_l(\mathbf{k}_{lm} \cdot \mathbf{r})$ are assumed to be suitably normalized and orthogonal. It is apparent that when $\delta\epsilon/\epsilon$ has only a slow variation over the cross section then both l' and m' cannot differ much from l and m. Thus the distribution of fields over the output face of the FPE will not differ much from that associated with the incident carrier wave and the angular spread of the output carrier and the modulation sidebands will not be seriously deteriorated. In addition the amplitude of the modulation index will not be drastically reduced. Thus it appears that even though the microwave field may vary from a maximum value at the center of the FPE to a value approaching zero at its transverse boundaries, no serious effects will result. Very rapid transverse variations of the microwave field, on the other hand, could cause difficulties.

V. AMPLITUDE MODULATION

In the traveling-wave modulator proposed by Kaminow[1] and the FPE modulator described here, the fundamental effect is a phase modulation. By suitable use of polarizers and birefringent plates, the phase modulation of the traveling-wave modulator can convert linearly polarized light into elliptic polarization and hence into amplitude or intensity modulation. It is therefore of interest to look into the intensity modulation properties of the FPE modulator. We require an expression for the total electric field associated with the transmitted light E_t. Using (28), this is given by

$$E_t = E_a \exp i\omega_a t \, [1 + \sum_{n=1}^{\infty} ([E_{a-nm}/E_a] \exp - in\omega_m t \qquad (50)$$
$$+ [E_{a+nm}/E_a] \exp in\omega_m t)]$$

in which E_a is assumed to be polarized parallel to one of the privileged axes which we shall denote as the x-axis of Fig. 1. Combining (35) and (41), (50) becomes for the x-component of the FPE field,

$$E_{tx} = E_{ax} \exp i\omega_a t \, [1 + \sum_{n=1}^{\infty} (y^n + [-y^*]^n)]$$
$$= E_{ax} \exp i\omega_a t \, [(1 - y)^{-1} + (1 + y^*)^{-1} - 1] \qquad (51)$$
$$= E_{ax} \exp i\omega_a t \, F(y)$$

in which

$$y = [E_{a-m}/E_a] \exp - i\omega_m t \tag{52}$$

$$= (i2\chi)^{-1} [\sqrt{1 + 4 \mid \chi \mid^2} - 1] \exp - i\omega_m t,$$

$$F(y) = [(1 - y)^{-1} + (1 + y^*)^{-1} - 1]$$

and we have used the identity $\sum_{n=1}^{\infty} y^n = (1 - y)^{-1} - 1$. Since the privileged axes are driven push-pull, the expression for the transmitted electric field when the incident light is polarized parallel to the other privileged axis is identical, except that χ must be replaced by $-\chi$, and thus y must be replaced by $-y$. Therefore, we may write

$$E_{ty} = E_{ay} \exp i\omega_a t \, F(y)^*. \tag{53}$$

Since the incident field has the value $E/2$ when the light is polarized along one of the privileged axes, the relative intensity of the transmitted light is given by $I_t = \mid 2E_{tx}/E \mid^2$, and we can write, using Eqs. (42), (51) and (52)

$$I_t = \mid F(y) \mid^2 (Q_{aL}/Q_a)^2/[1 + 4 \mid \chi \mid^2]. \tag{54}$$

It can be shown after some manipulation that

$$F(y) = \sqrt{1 + 4 \mid \chi \mid^2} \frac{[1 - i2 \mid \chi \mid \cos (\omega_m t + \varphi)]}{[1 + 4 \mid \chi \mid^2 \cos^2 (\omega_m t + \varphi)]} \tag{55}$$

in which φ is the argument of χ which corresponds to the phase of the microwave field. Substituting (55) in (54) yields

$$I_t = (Q_{aL}/Q_a)^2/[1 + 4 \mid \chi \mid^2 \cos^2 (\omega_m t + \varphi)] \tag{56}$$

for the intensity modulation. For comparison with (2), which is the intensity modulation derived from the Fabry-Perot equation, we assume that the internal loss is negligible so that $Q_{aL} = Q_a$. Using (34) and (19), and writing $\delta\epsilon/\epsilon = 2\delta\mu/\mu$, it follows that

$$2 \mid \chi \mid = \frac{\pi a}{4} \frac{(1 + \Gamma)^2}{1 - \Gamma^2} \frac{\delta\mu}{\mu}. \tag{57}$$

We note that the equivalent term in (2) is $\Gamma(\pi a \delta\mu/\mu)/(1 - \Gamma^2)$, which differs by a factor $4\Gamma/(1 + \Gamma)^2 \approx 1$ from $2 \mid \chi \mid$. The negligible difference from unity can be attributed to the high-Q approximation. Thus the relatively naive considerations which lead to (2) have been shown to be precisely correct.

Other, more efficient, types of intensity modulation can be achieved through the use of polarizers. Suppose that the incident light is polar-

ized at 45° to the privileged axis, and that the modulator is followed by an analyzer which is set parallel to the polarizer. Under these circumstances $E_{az} = E_{ay}$ and

$$E_t = (E_{tx} + E_{ty})/\sqrt{2} = \tfrac{1}{2}E_a \exp i\omega_a t \, [F(y) + F(y)^*].$$

Thus we may write

$$
\begin{aligned}
I_t &= [R_eF(y)]^2(Q_{aL}/Q_a)^2/[1 + 4\,|\,\chi\,|^2] \\
&= (Q_{aL}/Q_a)^2/[1 + |\,\chi\,|^2 \cos^2\,(\omega_m t + \varphi)]^2
\end{aligned}
\tag{58}
$$

which we note differs from (55) in that the denominator is squared. As a result the depth of modulation is roughly twice that attained without the use of polarizer and analyzer. The extra efficiency arises from the fact that the cavity also enhances the induced birefringence and this effect is brought out by the use of the polarizers.

When the analyzer is set at 90° to the polarizer

$$E_t = (E_{tx} - E_{ty})/\sqrt{2} = \tfrac{1}{2}E_a \exp i\omega_a t \, [F(y) - F(y)^*]$$

and

$$
\begin{aligned}
I_t &= [I_mF(y)]^2(Q_{aL}/Q_a)/[1 + |\,\chi\,|^2] \\
&= (Q_{aL}/Q_a)^2 4\,|\,\chi\,|^2 \cos^2\,(\omega_m t + \varphi)/[1 + 4\,|\,\chi\,|^2 \cos^2\,(\omega_m t + \varphi)]^2
\end{aligned}
\tag{59}
$$

and we note that the light is 100 per cent intensity modulated. The peak transmitted intensity occurs for $|\,\chi\,| \cos\,(\omega_m t + \varphi) = 0.5$.

So far, the intensity modulation has been at twice the microwave frequency. It is possible to modulate at the microwave frequency by introducing a quarter-wave plate between the polarizer and analyzer, with the polarizers either parallel or crossed. Under these circumstances,

$$E_t = (E_{tx} \pm iE_{ty})/\sqrt{2} = \tfrac{1}{2}E_a \exp i\omega_a t \, [F(y) \pm iF(y)^*],$$

and

$$
\begin{aligned}
I_t &= \tfrac{1}{2}[R_eF(y) \pm I_mF(y)]^2(Q_{aL}/Q_a)^2/[1 + 4\,|\,\chi\,|^2] \\
&= \tfrac{1}{2}(Q_{aL}/Q_a)^2\,[1 \mp 2\,|\,\chi\,| \cos\,(\omega_m t + \varphi)]^2/ \\
&\qquad\qquad\qquad [1 + 4\,|\,\chi\,|^2 \cos^2\,(\omega_m t + \varphi)]^2
\end{aligned}
\tag{60}
$$

which has a component at the microwave frequency. The introduction of a quarter-wave plate provides a bias, or steady retardation, which allows linear modulation.

VI. MODULATION POWER

The microwave power required for modulation is a function of the percentage modulation. For the sake of discussion we will assume the modulation index is 0.636, corresponding to 23 per cent modulation or maximum energy in the first sideband as described earlier. In terms of the peak variation of the index of refraction μ, the modulation index is given by $\chi = \frac{1}{2}Q_{aL}\delta\mu/\mu$, as follows from (34). For light polarized perpendicular to the z-axis of KDP[7] and parallel to one of the privileged axes

$$\delta\mu = \pm\tfrac{1}{2}\mu^3 r_{63} E_p \tag{61}$$

in which E_p is the peak microwave electric field parallel to the z-axis and r_{63} is the electrooptic constant. The peak electric field can be related to the microwave power absorbed in the KDP by

$$P \approx \tfrac{1}{4}\sigma E_p^2 V \tag{62}$$

in which P is the absorbed power, V is the volume of KDP and σ is the microwave conductivity given by

$$\sigma = \omega\epsilon/Q_d \tag{63}$$

in which ω is the angular microwave frequency, ϵ is the microwave dielectric constant and Q_d is the dielectric Q or inverse loss tangent. It has been assumed that the energy loss associated with microwave electric fields transverse to the z-axis is negligible, a condition which is easily satisfied. The factor $\frac{1}{4}$ arises from a time and space average of the microwave standing wave pattern. Combining (61) to (63) yields

$$P \approx (\omega\epsilon/Q_d)(2\chi/Q_{aL}\mu^2 r_{63})^2 V. \tag{64}$$

At a frequency of 10^{10} cps (X-band) the dielectric constant $\epsilon \approx 20\epsilon_0$, in which ϵ_0 is the dielectric constant of free space, and $Q_d \approx 150$.† For $f_a = 5 \times 10^{14}$ cps, $r_{63} \approx 10^{-11}$ meters/volt, $\Gamma^2 = 0.9$, $Q_{aL} \approx 10^6$ and $\mu \approx 1.5$, (64) yields a power requirement of 240 watts/cc of KDP for $|\chi| = 0.636$. For modulation in the first sideband we must have $c'/2L = 10^{10}$ or $L = 1$ cm. For modulation powers of order ten watts the cross-sectional area of the KDP must be under 0.04 cm^2, corresponding to a diameter of about 0.2 cm.

The temperature rise ΔT from surface to center of a uniformly heated rod which is long compared to its diameter can be written

---

† Measured at room temperature.

$$\Delta T \approx \tfrac{1}{4} P R^2 / K V \tag{65}$$

in which R is the radius, K is the thermal heat conductivity and P/V is defined by (64). For KDP, $K \approx 10^{-2}$ watts/cm°K yielding a temperature rise of about 60°K for a diameter of 0.2 cm. Clearly the diameter of the rod should be made as small as possible.

Another alternative is that the length of the FPE can be doubled. By modulating in the second sideband, the same X-band modulation can be accomplished with one-half the power and one-quarter the power dissipation per unit volume. The modulation bandwidth would be halved.

The choice of $|\chi| = 0.636$ was motivated by the fact that this yields maximum energy in the first sideband. For superheterodyne systems this is a desirable objective. For envelope detection, amplitude modulation at the modulating frequency is desirable. In this case the combination of polarizer analyzer and quarter-wave plate is required yielding the modulation described by (60). For this case $|\chi| \approx 0.1$ would be quite adequate, and the heat dissipation, temperature rise and required power would be reduced by a factor of four.

VII. OPTICAL CONSIDERATIONS

It has been supposed that the light beam to be modulated is perfectly collimated, thereby allowing the energy to be completely coupled into the fundamental FPE mode. Since no light beam is perfectly collimated, we would like to examine this assumption and to determine what limitations are placed on the optical cavity.

When a diverging beam of monochromatic light is incident on the FPE, the transmitted beam will form the typical Fabry-Perot fringes[3] at infinity. The central fringe corresponds to one of the axial FPE modes. Thus we can be assured that the incident energy will excite only the axial mode if the angular spread of the beam is less than the angular spread of the central fringe. The phase difference between successive averaging rays moving at an angle θ within the FPE is given by[3]

$$2\varphi = 4\pi f L \mu c^{-1} \cos \theta.$$

Substituting into (1) indicates that the transmitted intensity will fall to half its $\theta = 0$ value for a bright central fringe at an angle $\theta_{\frac{1}{2}}$ given by $\theta_{\frac{1}{2}} \approx [c(1 - \Gamma^2)/2\pi f_a L \mu \Gamma]^{\frac{1}{2}} \approx Q_a^{-\frac{1}{2}}$. Thus the angular spread of the incident beam must be less than $2\mu Q_a^{-\frac{1}{2}}$. Taking the internal losses of the FPE into account, we would find the somewhat larger angle $2\mu Q_{aL}^{-\frac{1}{2}}$. For the FPE cavity discussed earlier with $Q_{aL} \approx 10^6$, $2\mu Q_{aL}^{-\frac{1}{2}} \approx 3 \times 10^{-3}$

corresponding to 10 minutes of arc. For a beam whose angular divergence is diffraction-limited, the required aperture of the FPE is

$$D \approx 1.2 c Q_{aL}^{\frac{1}{2}} / 2 \mu f_a$$

or about 0.25 mm for the example given above. When "walkoff" of light rays incident at the angle $\mu\theta_{\frac{1}{2}}$ is considered, the required aperture is increased by approximately $2\theta_{\frac{1}{2}} L/(1 - \Gamma^2) = 0.15$ mm. Since KDP rods with a diameter much less than about 1 mm do not seem practical for the present, no constraints will be imposed by any diffraction limited beam such as that from an optical maser.

One important constraint imposed on the optical maser is that it oscillate in only one longitudinal mode or, if not that, that the modes be separated by at least the frequency Δf but less than $c/2L\mu$.

In a manner similar to that by which the half angle $\theta_{\frac{1}{2}}$ was found it can be seen that the quantity μL must be controlled such that

$$\Delta(\mu L)/\mu L < (2Q_a)^{-1} \approx 3 \times 10^{-7}.$$

Changes in temperature of the cavity will tend to change the product μL. Our best guess, in view of the lack of data, is that $\Delta\mu L$ per degree centigrade is of order 2×10^{-6} for one centimeter length. Thus, the temperature must be controlled to better than 0.1°C. Likewise, the etalon should be flat and parallel to within 3×10^{-7} cm.

VIII. COMPARISON TO TRAVELING-WAVE MODULATORS

In traveling-wave reactive modulators the light makes one traversal of the electrooptic material. The transmitted electric field amplitude, for light polarized along one of the privileged axes, is given by

$$E_t = E_0 \exp{(i\gamma \sin{\omega_m t})} \tag{66}$$

in which $\gamma = (\omega L/c')(\delta\mu/\mu)$ and E_0 is the amplitude of the incident light wave at frequency ω. The comparison to the FPE modulator is made by relating the amplitudes of the various sidebands. Thus we may write for E_t

$$E_t = E_0 \sum_{n=-\infty}^{+\infty} J_n(\gamma) \exp{in\omega_m t} \tag{67}$$

in which J_n is the Bessel function of order n.† The amplitude of the nth sideband compared to the carrier is

---

† We have used the identity $\exp{iz \sin{\theta}} = \sum_{n=-\infty}^{+\infty} J_n(z) \exp{in\theta}$.

$$E_n/E_0 = J_n(\gamma)/J_0(\gamma) \approx \gamma^n/2^n n! \quad \text{for} \quad \gamma \ll 1 \qquad (68)$$

which should be compared to (44). We note that for first sideband modulation $\gamma/2$ should be compared to $|\chi|$. Thus we compare

$$\tfrac{1}{2}(\omega L/c')(\delta\mu/\mu)$$

and $\tfrac{1}{8}[(1 + \Gamma)^2/(1 - \Gamma^2)](\omega L/c')(\delta\mu/\mu)$. This indicates that the FPE modulation is more efficient by the factor $\tfrac{1}{4}(1 + \Gamma)^2/(1 - \Gamma^2) \approx \Gamma(1 - \Gamma^2)^{-1}$. For $\Gamma^2 = 0.9$ and equal modulation lengths this is a factor of ten in $\delta\mu/\mu$, requiring one hundred times the power for equal modulation. The advantage of multiple reflection can be compensated somewhat by increasing the interaction length of the traveling-wave modulator. For example, increasing L by $\Gamma(1 - \Gamma^2)^{-1}$ equates $|\chi|$ and $\gamma/2$ for equal $\delta\mu/\mu$ but requires $\Gamma(1 - \Gamma)^{-1}$ times as much power, since for equal $\delta\mu/\mu$ the power dissipation per unit length must be the same. Increasing L by $\Gamma^2(1 - \Gamma^2)^{-2} \approx 100$ and decreasing $\delta\mu/\mu$ by $(1 - \Gamma^2)^{-1}$ again equates $\gamma/2$ and $|\chi|$. In this case the power requirement for either modulator is the same. The potentially greater bandwidth of the traveling-wave device in this case is probably obviated by the difficulty of matching microwave and optical phase velocities closely because of the long interaction length. The power dissipation per unit volume is of course, lower by $\Gamma^{-2}(1 - \Gamma^2)^{+2}$ in the traveling-wave device and this is an extremely important advantage. However, the difficulties associated with fabricating a long rod of electrooptic material appear to be formidable at the moment. In addition, the optical loss of the KDP, as has been recently pointed out by Kaminow,[8] is not completely negligible. A rod sufficient to achieve the last mentioned advantages would be about one hundred centimeters long at X-band modulation frequencies. Since the energy decay distance is about 35 centimeters,† the light intensity would fall to $\approx e^{-3}$ of its initial value, i.e., about 13 db down.

In order to determine the transmission loss through the FPE modulator we may consider the ratio

$$(Q_{aL}/Q_a)^2 = [1 + Q_a/Q_{ad}]^{-2}.$$

Since

$$Q_a/Q_{ad} \approx L/\ell(1 - \Gamma^2) \approx \tfrac{2}{3}$$

the transmitted intensity is decreased by a factor slightly less than three, which is comparable to that of the intermediate length modulator.

---

† A rough measurement at 6328 A.

corresponding to 10 minutes of arc. For a beam whose angular divergence is diffraction-limited, the required aperture of the FPE is

$$D \approx 1.2 c Q_{aL}^{\frac{1}{2}} / 2\mu f_a$$

or about 0.25 mm for the example given above. When "walkoff" of light rays incident at the angle $\mu\theta_{\frac{1}{2}}$ is considered, the required aperture is increased by approximately $2\theta_{\frac{1}{2}}L/(1 - \Gamma^2) = 0.15$ mm. Since KDP rods with a diameter much less than about 1 mm do not seem practical for the present, no constraints will be imposed by any diffraction limited beam such as that from an optical maser.

One important constraint imposed on the optical maser is that it oscillate in only one longitudinal mode or, if not that, that the modes be separated by at least the frequency Δf but less than $c/2L\mu$.

In a manner similar to that by which the half angle $\theta_{\frac{1}{2}}$ was found it can be seen that the quantity μL must be controlled such that

$$\Delta(\mu L)/\mu L < (2Q_a)^{-1} \approx 3 \times 10^{-7}.$$

Changes in temperature of the cavity will tend to change the product μL. Our best guess, in view of the lack of data, is that $\Delta\mu L$ per degree centigrade is of order 2×10^{-6} for one centimeter length. Thus, the temperature must be controlled to better than 0.1°C. Likewise, the etalon should be flat and parallel to within 3×10^{-7} cm.

VIII. COMPARISON TO TRAVELING-WAVE MODULATORS

In traveling-wave reactive modulators the light makes one traversal of the electrooptic material. The transmitted electric field amplitude, for light polarized along one of the privileged axes, is given by

$$E_t = E_0 \exp (i\gamma \sin \omega_m t) \tag{66}$$

in which $\gamma = (\omega L/c')(\delta\mu/\mu)$ and E_0 is the amplitude of the incident light wave at frequency ω. The comparison to the FPE modulator is made by relating the amplitudes of the various sidebands. Thus we may write for E_t

$$E_t = E_0 \sum_{n=-\infty}^{+\infty} J_n(\gamma) \exp in\omega_m t \tag{67}$$

in which J_n is the Bessel function of order n.† The amplitude of the nth sideband compared to the carrier is

––––––––––

† We have used the identity $\exp iz \sin \theta = \sum_{n=-\infty}^{+\infty} J_n(z) \exp in\theta$.

Thus a traveling-wave modulator with a length about $\Gamma(1 - \Gamma^2)^{-1}$ times that of the FPE modulator is comparable with respect to heat dissipation and transmission loss but requires $\Gamma(1 - \Gamma^2)^{-1}$ times the modulation power. An increase in length by another factor $\Gamma(1 - \Gamma^2)^{-1}$ is somewhat drastic, since the transmission loss increases too rapidly.

IX. CONCLUSION

The modulation properties of an electrooptic material placed in a Fabry-Perot etalon have been analyzed in detail. The main conclusion of the analysis is that the FPE modulator offers considerable advantage over the traveling-wave modulator with respect to modulation power when the traveling-wave modulator is not extremely long.

The bandwidth of the modulator is limited by the passband of the FPE, which for practical purposes is several hundred megacycles. In general this bandwidth is compatible with the bandwidth of the microwave modulating cavity (without wasting modulation power) since the dielectric Q of the material is 150, yielding cavity Q's of order 75 and a bandwidth of approximately 130 mc/s. When the electrooptic material does not fill the entire cavity or when additional microwave bandwidth is required, additional loss must be coupled into the cavity. Likewise multimesh broadbanding techniques can be useful.

The requirement of a stable monochromatic source of light for the modulator represents no real disadvantage, since we suppose that the modulator would be used in a communication system which needs such a source anyway.

A major requirement for the successful operation of the FPE modulator is the choice of the correct microwave field configuration which, in effect, matches the microwave phase velocity to that of the light in the electrooptic medium.

REFERENCES

1. Kaminow, I. P., Phys. Rev. Letters, 6, May 15, 1961, p. 528.
2. Pershan, P. S., and Bloembergen, N., Advances in Quantum Electronics, Ed. J. R. Singer, New York, Columbia Univ. Press, 1961, p. 187.
3. Jenkins, F. A., and White, H. E., Fundamentals of Optics, New York, McGraw-Hill, 1950, p. 258.
4. Tien, P. K., J. Appl. Phys., 29, 1958, p. 1357.
5. White, A. D., and Ridgen, J. D., Proc. I.R.E., 50, July, 1962, p. 1697.
6. Slater, J. C., Microwave Electronics, Princeton, D. Van Nostrand, p. 57.
7. Seitz, F., and Turnbull, D., Solid State Physics, Vol. 4, New York, Academic Press, p. 90.
8. Kaminow, I. P., private communication.

Contributors to this Issue

CHARLES D. BEACH, B.S. in E.E., 1959, University of Colorado; M.S. in E.E., 1961, New York University; Bell Telephone Laboratories 1959–1962. His work at the Laboratories involved theoretical analyses of interchannel modulation distortion in tropospheric scatter radio systems. He also participated in an experimental study of intermodulation performance of such systems. Since February, 1962, he has been on leave of absence from the Laboratories while teaching and pursuing doctoral studies at the University of Colorado. Member Tau Beta Pi, Eta Kappa Nu, Sigma Tau.

ROYER R. BLAIR, B.S. in E.E., 1930, Rose Polytechnic Inst.; M.A. in Physics, 1939, Columbia University; Bell Telephone Laboratories, 1930—. Mr. Blair's initial work at the Laboratories in transmission research led to the design of circuits for broadband amplifiers, regulators and test equipment for coaxial systems. During World War II he was concerned with circuits for proximity fuses, radar indicators and other security-classified devices. There followed a period in apparatus development on germanium and silicon varistors and on processing and test equipment used in that field. He later developed high-speed transistor switching circuitry for digital computer applications for the military. Since 1958, he has been investigating and studying the effects of nuclear radiation on various semiconductor devices and on their performance in a wide variety of applications. He is a senior member of the Institute of Radio Engineers and a member of Tau Beta Pi.

W. F. BODTMANN joined Bell Telephone Laboratories in 1941 and during 1942–1945 served in the U.S. Air Corps. After returning to the Laboratories in 1945, he worked on various circuits and components for the TD-2, TJ and TL radio relay systems, and on transistor circuits for wideband amplifiers. Later he was concerned with the FM feedback receivers for Project Echo, and more recently has been working on the frequency compression demodulators for Project Telstar.

W. L. BROWN, B.S., 1945, Duke University; A.M., 1947, Ph.D., 1951, Harvard University; Bell Telephone Laboratories 1950—. Mr. Brown

has been engaged in research on the physical properties of semiconductor surfaces and the nature of defects produced in semiconductors by high energy radiation. He has recently been responsible for radiation damage experiments aboard the Telstar satellite. Fellow of American Physical Society, Sigma Xi and Phi Beta Kappa.

JAMES C. COYNE, A.B., 1953, B.S., 1954, and M.S., 1958, Columbia University; Bell Telephone Laboratories, 1958—. Mr. Coyne participated in the planning for the Columbus Branch Laboratory and was engaged in development work on welded connections. At present he is engaged in reliability and application studies of solderless wrapped and percussive welded connections. Tau Beta Pi.

EUGENE I. GORDON, B.S., 1952, City College of New York; Ph.D., 1957, Massachusetts Institute of Technology; Bell Telephone Laboratories, 1957—. A member of the Electron Device Department, he is engaged in research in optical masers. Member American Physical Society, Phi Beta Kappa, Sigma Xi.

H. JAMES GRAFF, B.S.M.E., 1955, University of Notre Dame; Bell Telephone Laboratories 1956—. Mr. Graff has been concerned primarily with the development of a mechanical splicing method for polyethylene insulated conductor cable. At present he is in charge of a group responsible for the development of load coils, coil cases, and other outside plant items.

JESSIE MACWILLIAMS, B.A., 1939, M.A., 1941, Cambridge (England); Ph.D., 1962, Harvard; Bell Telephone Laboratories 1956—. Mrs. MacWilliams has been concerned with writing computer programs for the analysis and synthesis of transmission networks. She is now engaged in data systems studies, particularly the study of algorithms for decoding systematic error-correcting codes. Member Mathematical Association of America.

JOHN M. PEACOCK, B.S.E., 1947, Princeton University; Western Electric Co., 1947–50; Bell Telephone Laboratories 1953—. At Bell Laboratories Mr. Peacock first worked on the design of antennas, radio towers and other structures. Later he was engaged in fundamental development

work on expanded dielectrics for cable conductor insulation. At present he is supervisor in charge of outside plant wire joining and cable splicing activities. Member, American Wire Association.

D. STEWART PECK, B.S.E., 1939, M.S., 1940, University of Michigan; Bell Telephone Laboratories 1947—. Mr. Peck has been concerned in the design for production of gas-filled electron tubes such as rectifiers, thyratrons and cold-cathode tubes. More recently, he has been in charge of work on reliability studies, applications engineering, and specifications for electron devices. On Project Telstar Mr. Peck's department has been generally responsible for the specification, aging and selection programs for semiconductor components used in the satellite. Member A.I.E.E., American Standards Association, Tau Beta Pi, Eta Kappa Nu, Sigma Xi and Phi Kappa Phi.

ARNOLD PFAHNL, Dr. Phil., University of Graz, Austria, 1948, D.Sc., University of Paris, France, 1956; G. Massiot & Cie. and Massiot-Fluor, Courbevoie, France, 1950–57, Bell Telephone Laboratories, 1957—. At the Laboratories, Mr. Pfahnl has worked on the study of the properties of fast-decay cathode-ray tube phosphors, the development of ZnO phosphors with application to cathode-ray tubes used in electronic switching, and the study of injection electroluminescence of gallium phosphide.

J. DANE RIGDEN, B.Sc., 1955, and Ph.D., 1958, Reading University, Reading, England; National Research Council of Canada, 1958–60; Bell Telephone Laboratories 1960—. Mr. Rigden has specialized in the development of gaseous optical masers and is the co-developer of the visible helium-neon maser.

C. L. RUTHROFF, B.S., in E.E., 1950, and M.A., 1952, University of Nebraska; Long Lines, 1946–52; Bell Telephone Laboratories, 1952—. Since transferring from Long Lines in Lincoln, Nebraska, where he was a central office maintenance man, Mr. Ruthroff has concentrated on radio research with emphasis on FM problems and transistor circuits for microwave components and systems. Member, I.R.E.

FRIEDOLF M. SMITS, Dipl. Phys., 1950, Dr. rer. nat., 1950, University of Freiburg, Germany; research associate, Physikalisches Institut, Uni-

versity of Freiburg, 1950–54; Bell Telephone Laboratories, 1954–62. Mr. Smits went to the Sandia Corporation in May 1962. His work at Bell Telephone Laboratories included studies of solid-state diffusion in germanium and silicon, device feasibility, and process studies, as well as the development of UHF semiconductor devices. He supervised a group that conducted radiation damage studies on components, particularly solar cells, used in the Telstar experimental satellite. Member of the American Physical Society and the German Physical Society.

JOSEPH M. TRECKER, B.S. in E.E., 1949, Iowa State College; Bell Telephone Laboratories, 1949—. His early assignments involved studies of information requirements and planning of communication systems for military applications. More recently, he was engaged in studies of transmission characteristics of tropospheric scatter radio systems. He is now in charge of a group studying information flow and ground communication network arrangements for the Apollo space mission. Member Tau Beta Pi, Eta Kappa Nu.

J. JOHN ZALMANS, B.S.M.E., 1953, University of Arizona; Bell Telephone Laboratories, 1953—. Mr. Zalmans is engaged in the development of connecting devices. American Society of Mechanical Engineers, Tau Beta Pi, Pi Mu Epsilon.