A Method for Measurement of the Duration of Picosecond Pulses by Beat Frequency Detection of Laser Output

By W. M. HUBBARD

This paper describes a new method for measuring the duration of mode-locked picosecond laser pulses. It is similar to the Two Photon Fluorescence and Second Harmonic Generation methods in that it measures not the pulse duration directly but rather an autocorrelation function of the signal. It has the advantage in that it can be used for very low-power repetitive light signals.

I. INTRODUCTION

Two methods for determining the duration of picosecond pulses have been reported. The earlier method, which utilizes second harmonic generation (SHG) in nonlinear crystals, was reported independently and almost simultaneously by M. Maier, W. Kaiser and J. A. Giordmaine;[1] J. A. Armstrong;[2] and H. P. Weber.[3] The second method, which utilizes the two-photon fluorescence (TPF) in certain dyes, was first reported by J. A. Giordmaine[4] and his co-workers. Measurements of pulse durations based on these methods have been reported by num-

erous authors. Extensive references to these papers are included in a review article by A. J. DeMaria, W. H. Glenn, Jr., M. J. Brienza and M. E. Mack.[5]

Neither of these methods measures pulse duration directly. Both, in fact, measure the autocorrelation function $G(\tau)$ of the intensity $I(t)$, namely

$$G(\tau) = \int_{-\infty}^{\infty} I(t)I(t + \tau)\, dt.$$

A careful measurement of the contrast ratio between $G(\tau)$ at the peak of the pulse and $G(\tau)$ between pulses is necessary in order to determine whether or not the measured values represent the duration of a pure AM pulse. (This was first pointed out for TPF by H. P. Weber[6] and independently by J. R. Klauder, M. A. Duguay, J. A. Giordmaine and S. L. Shapiro.)[7] If the contrast ratio is not the correct value for pure AM pulses, little concerning the signal waveform can be reliably inferred from a knowledge of $G(\tau)$.

The beat frequency detection (BFD) method proposed in this paper measures a different quantity, namely

$$H_k(\tau) = \tfrac{1}{4}F_k(\tau) + \tfrac{1}{4}F_k(-\tau)$$

where

$$F_k(\tau) = \left| \frac{1}{T} \int_{-T/2}^{T/2} \mathcal{E}^*(t)\, \mathcal{E}(t + \tau) \exp\left(-j2\pi \frac{t}{T} k\right) dt \right|^2$$

and $\mathcal{E}(t)$ is the electric field of the optical signal, $1/T$ is the pulse repetition rate and k/T is the frequency of the harmonic chosen for analysis. It too suffers from the fact that if the correct "contrast ratio"* is not observed, the result is ambiguous.

If the proper "contrast ratio" is observed, then $\mathcal{E}(t)$ is real (i.e., there is no angle modulation of the signal) and in cases of practical interest the pulse duration is short compared with the period T. Thus, for small k, $kt/T \ll 1$ over the region where $\mathcal{E}(t)$ contributes significantly to the integral and

$$F_k(-\tau) = F_k(\tau) = \left| \frac{1}{T} \int_{-T/2}^{T/2} \mathcal{E}^*(t)\, \mathcal{E}(t + \tau)\, dt \right|^2.$$

Thus $H_k(\tau)$ is essentially the autocorrelation function of $\mathcal{E}(t)$ when there is no angle modulation. It is therefore as good a measure of pulse duration as is the autocorrelation function of $I(t)$.

---

* This term will be defined in Section III.

Under certain conditions there is reason to believe that the pulses are "chirped." That is, they have the form

$$\mathcal{E}(t) = \mathcal{E}_0(t) \exp \left[j(\omega_0 t + \alpha t^2) \right]$$

where $\mathcal{E}_0(t)$ is a real, slowly varying envelope function. The measured quantity $H_k(\tau)$ has an interesting and distinctive behavior in this case. This is considered in Section III. It should be possible in some cases to determine by the BFD method how much, if any, "chirp" is present. Neither TPF nor SHG is capable of providing this information.

II. DESCRIPTION AND ANALYSIS OF THE BEAT-FREQUENCY DETECTION METHOD OF PULSE-DURATION MEASUREMENT

The output $\mathcal{E}(t)$ of a laser is a periodic function which can be written in the form

$$\mathcal{E}(t) = \sum_{n=-\infty}^{\infty} A_n \cos \left\{ (n\omega + \omega_0)t + \phi_n \right\} \tag{1}$$

where $2\pi/\omega = T$ is the period, ω_0 is the center frequency of the spectrum and A_n is the amplitude of the nth line in the spectrum. If the phases ϕ_n are related by

$$\phi_n = n\Phi \tag{2}$$

where Φ is any constant, the signal $\mathcal{E}(t)$ is a sequence of pulses with no angle modulation and the pulse envelope has the minimum possible duration for the set $\{A_n\}$. Such a signal is called "mode locked". If equation (2) is not satisfied, angle modulation (and a wider pulse) occurs. In general, if ϕ_n is a set of random numbers a "noiselike" (but periodic) signal is obtained. The pulse duration of the signal $\mathcal{E}(t)$ is therefore dependent on $\{\phi_n\}$.

Consider the circuit shown in Fig. 1. The laser output $\mathcal{E}(t)$ is divided into equal parts by the 3 dB hybrid mirror (beam splitter) and the two parts are recombined in the same mirror after one path has suffered a delay τ relative to the other. Beam transforming lenses may be required in order to assure alignment of the phase fronts of the two interfering beams. Mirror M_2 is mounted on a track so that the delay in one arm of the device can be varied over a range corresponding to one period of the laser signal. Mirror M_1 is mounted on a piezoelectric acoustic modulator so that its position can be varied periodically over a distance of a few microns for reasons which will be explained below. The output of the hybrid is therefore
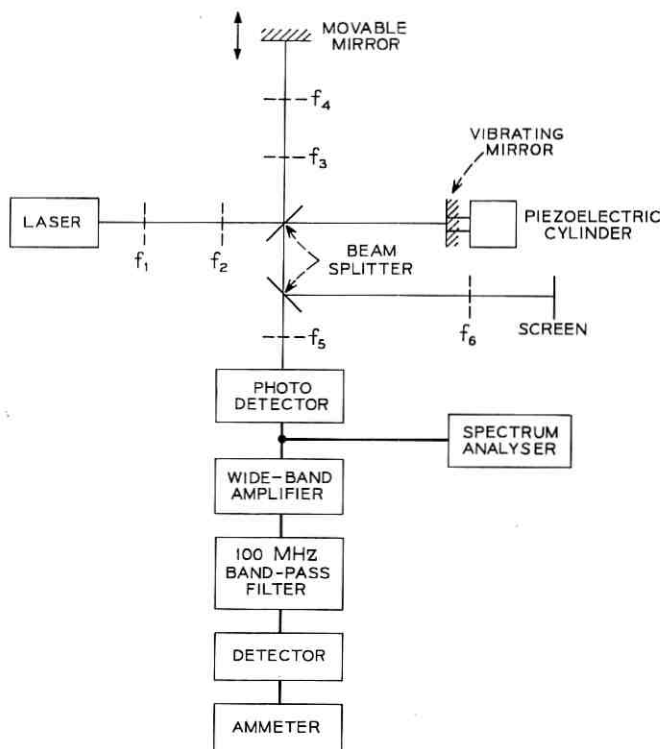
Fig. 1—Diagram of the BFD apparatus.

$$\{\,\mathcal{E}(t) + \mathcal{E}(t - \tau)\}$$

$$= \sum_{n=-\infty}^{\infty} A_n[\cos\{(n\omega + \omega_0)t + \phi_n\} + \cos\{(n\omega + \omega_0)(t - \tau) + \phi_n\}]. \quad (3)$$

If this signal is now applied to a product demodulator (photodetector), one obtains as output:

$$V(t) = [\mathcal{E}(t) + \mathcal{E}(t - \tau)]^2$$

$$= \sum_{n=-\infty}^{\infty} \sum_{\kappa=-\infty}^{\infty} B_n^{(\kappa)} \cos\{\kappa\omega t + \theta_n^{(\kappa)}\}$$

$$\cdot \left[\cos\frac{\kappa\omega\tau}{2} + \cos\left\{\left(n + \frac{\kappa}{2}\right)\omega + \omega_0\right\}\tau\right] \quad (4)$$

where

$$B_n^{(\kappa)} = A_{n+\kappa}A_n\,, \qquad \theta_n^{(\kappa)} = \phi_{n+\kappa} - \phi_n$$

and terms in $2\omega_0 t$ have been ignored as has a time shift of amount $\tau/2$. If this signal is now passed through a bandpass filter which is centered at $k\omega$ and is sufficiently narrow that only the $\kappa = k$ term in the sum is passed, one obtains

$$V_k(t) = \sum_{n=-\infty}^{\infty} B_n^{(k)} \cos\{k\omega t + \theta_n^{(k)}\}$$
$$\cdot \left[\cos\frac{k\omega\tau}{2} + \cos\left\{\left(n + \frac{k}{2}\right)\omega + \omega_0\right\}\tau\right]. \qquad (5)$$

It should be noted that since $\theta_n^{(0)} = 0$ for all n, $V_0(t)$ is independent of $\{\phi_n\}$ and independent of t.

$$V_0(t) = \sum_{n=-\infty}^{\infty} B_n^{(0)}[1 + \cos\{(n\omega + \omega_0)\tau\}]. \qquad (6)$$

One can readily understand this result by realizing that the $V_k(t)$ term in general results from the beats between all possible pairs of lines separated by an amount $k\omega$. In particular $V_0(t)$ results from "beats" between each spectral line and itself. Since no beats between different lines contribute to $V_0(t)$, no relative phases between lines influence the result.

Returning to the general case, square law detection of $V_k(t)$ gives

$$U_k(\tau) = 2\,|\,V_k(\tau)\,|^2$$
$$= \sum_n \sum_m B_n^{(k)} B_m^{(k)} \cos(\theta_n^{(k)} - \theta_m^{(k)})$$
$$\cdot \left[\cos^2\frac{k\omega\tau}{2} + \tfrac{1}{2}\cos(n-m)\omega\tau + \tfrac{1}{2}\cos\{(n+m+k)\omega + 2\omega_0\}\tau\right.$$
$$+ \cos\frac{k\omega\tau}{2}\left[\cos\left\{\left(n + \frac{k}{2}\right)\omega + \omega_0\right\}\tau\right.$$
$$+ \left.\left.\cos\left\{\left(m + \frac{k}{2}\right)\omega + \omega_0\right\}\tau\right]\right]. \qquad (7)$$

We observe that the entire ω_0 dependence is contained in the last three terms. If one of the mirrors in the interferometer is scanned through an optical wavelength, the contribution of these terms is cancelled out (since over an optical wavelength their average value is zero). This is accomplished by mounting one of the mirrors on a piezoelectric transducer and vibrating it. The resultant signal (after passing through a suitable low-pass filter) is given by

$$\bar{U}_k(\tau) = \sum_n \sum_m B_n^{(k)} B_m^{(k)} \cos\{\theta_n^{(k)} - \theta_m^{(k)}\}$$
$$\cdot \left[\cos^2\frac{k\omega\tau}{2} + \tfrac{1}{2}\cos(n-m)\omega\tau\right]. \qquad (8)$$

We pause to evaluate equation (8) for two extreme cases. Consider the mode-locked case $\phi_n = n\Phi$ which implies $\theta_n'^{(k)} - \theta_m^{(k)} = 0$. This gives

$$\bar{U}_k(\tau)_{\substack{\text{mode}\\\text{locked}}} = \cos^2 \frac{k\omega\tau}{2} [\sum_n B_n]^2 + \tfrac{1}{2} | \sum_n B_n \exp (jn\omega\tau) |^2. \qquad (9)$$

In this equation and in the remainder of this paper the superscript (k) on B_n and θ_n is to be understood. The mode-locked case represents one extreme in the sense that it provides the shortest possible pulse consistent with a particular set of amplitudes $\{A_n\}$. The "random case", i.e., ϕ_n a random variable uniformly distributed on $[-\pi, \pi]$, represents the opposite extreme—the expectation value of the signal power is constant in time. The expectation value of $\bar{U}_k(\tau)$ for this case can be obtained by observing that the expectation value of $\cos \{\theta_n - \theta_m\}$ when the $\{\phi_n\}$ are uniformly distributed on $[0, 2\pi)$ is given by

$$\langle\cos (\theta_n - \theta_m)\rangle = 1 \quad \text{if} \quad k = 0 \quad \text{(for all } n, m),$$
$$= 1 \quad \text{if} \quad n = m \quad \text{(for all } k),$$
$$= 0 \quad \text{otherwise.}$$

This gives:

$$\langle\bar{U}_k(\tau)\rangle = \left[\cos^2 \left(\frac{k\omega\tau}{2}\right) + \frac{1}{2}\right] \sum_n B_n^2 , \qquad k \neq 0. \qquad (10)$$

Equations (9) and (10) are illustrated in Fig. 2.

We now return to a consideration of the general case. Equation (8) can be written in the form

$$\bar{U}_k(\tau) = \cos^2 \frac{k\omega\tau}{2} \left| \sum_{n=-\infty}^{\infty} \mathcal{B}_n \right|^2$$

$$+ \frac{1}{4} \left| \sum_{n=-\infty}^{\infty} \mathcal{B}_n \exp (jn\omega\tau) \right|^2 + \frac{1}{4} \left| \sum_{n=-\infty}^{\infty} \mathcal{B}_n \exp (-jn\omega\tau) \right|^2 \qquad (11)$$

where

$$\mathcal{B}_n = B_n \exp (j\theta_n).$$

Similarly set

$$\mathcal{A}_n = A_n \exp (j\phi_n).$$
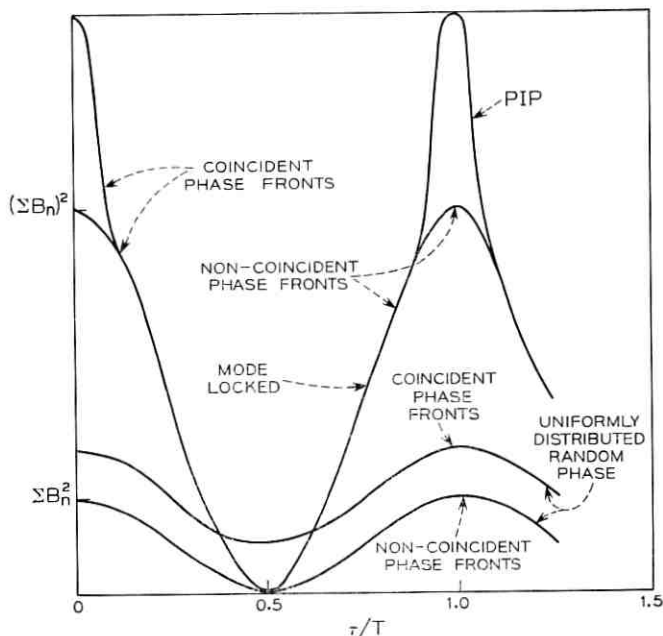
Then

$$\mathcal{B}_n = \mathcal{A}_{n+k}\mathcal{A}_n^* .$$

Fig. 2—Expected experimental results for the four cases discussed in the text.

Consider the expression:

$$\left| \sum_{n=-\infty}^{\infty} \mathcal{B}_n \exp\left(j2\pi \frac{x}{T} n\right) \right|^2$$

$$= \left| \sum_{n=-\infty}^{\infty} \mathcal{C}_{n+k} \mathcal{C}_n^* \exp\left(j2 \frac{\pi x n}{T}\right) \right|^2$$

$$= \left| \frac{1}{T} \int_{-T/2}^{T/2} E^*(-t)E(x-t) \exp\left(j2\pi k \frac{t}{T}\right) dt \right|^2 \quad (12)$$

where

$$E(t) = \sum_{n=-\infty}^{\infty} \mathcal{C}_n \exp\left(j \frac{2\pi t}{T} n\right). \quad (13)$$

Define

$$F_k(x) = \left| \frac{1}{T} \int_{-T/2}^{T/2} E^*(t)E(t+x) \exp\left(-j2\pi k \frac{t}{T}\right) dt \right|^2. \quad (14)$$

Then

$$\bar{U}_k(\tau) = \cos^2\left(\frac{k\omega\tau}{2}\right)F_k(0) + \tfrac{1}{4}F_k(\tau) + \tfrac{1}{4}F_k(-\tau). \quad (15)$$

From equation (15) we can draw an interesting comparison between the BFD method and the TPF and SHG methods. It is clear from equation (14) that $F_k(0)$ is the power spectrum of the intensity and therefore the autocorrelation function of the intensity is simply

$$G(\tau) = \sum_{k=-\infty}^{\infty} F_k(0) \exp\left(-j2\pi k \frac{\tau}{T}\right).$$

Therefore the quantities measured by all three methods depend only on the quantities $F_k(\tau)$. The results of the BFD method depend on the values for a single k and all τ; the results for TPF and SHG on the values for $\tau = 0$ and all k. From equations (12) and (14), we have

$$F_k(\tau) = \left| \sum_{n=-\infty}^{\infty} \mathcal{B}_n \exp\left(j2\pi \frac{\tau}{T} n\right) \right|^2 \tag{16}$$

while from equation (1)

$$I(t) = \left| \sum_{n=-\infty}^{\infty} \mathcal{A}_n \exp\left(j2\pi n \frac{t}{T}\right) \right|^2.$$

Thus, $F_k(\tau)$ and $I(t)$ are *formally* equivalent with the transformations

$$B_n \leftrightarrow A_n ,$$

$$\theta_n \leftrightarrow \phi_n ,$$

$$\tau \leftrightarrow t.$$

Comparing equation (13) with (1) shows that

$$\mathcal{E}(t) = \mathrm{Re}\ \{E(t) \exp(j\omega_0 t)\}.$$

That is, $E(t)$ is the (complex) signal from the laser with the carrier frequency omitted.

For the mode-locked case, the pulses are very short compared with the repetition rate $1/T$. Thus for at least the first few $k = 1, 2, \cdots$ we can write $(2\pi kt)/T \ll 1$ for those values of t for which $E(t)$ is significantly different from zero. This gives (since from equation (16) $F(\tau)$ is even when the B_n are all real):

$$F_k(-\tau) = F_k(\tau) \doteq \left| \frac{1}{T} \int_{-T/2}^{T/2} E^*(t)E(t + \tau)\, dt \right|^2.$$

Thus the first term in equation (15) is just a raised cosine with period T/k, but the sum of the second and third terms is the square of the autocorrelation function of $E(t)$. This sum provides a good measure of the pulse duration when the signal is mode locked.

Since this technique, like TPF and SHG, possesses some ambiguity

if the signal is not mode locked, it is important to be able to distinguish mode-locked behavior from non-mode-locked behavior. In order to do this, consider the quantity

$$\bar{U}_k(0) = \frac{3}{2} \left| \sum_{n=-\infty}^{\infty} \mathfrak{B}_n \right|^2,$$

$$= \frac{3}{2} \left| \sum_{n=-\infty}^{\infty} B_n \right|^2 \quad \text{mode-locked case,}$$

$$= \frac{3}{2} \sum_{n=-\infty}^{\infty} B_n^2 \quad \text{random-phase case.}_{k \neq 0}$$

For $k = 0$, $\theta_n \equiv 0$ and $\mathfrak{B}_n = B_n$ is real for any set $\{\phi_n\}$. Thus

$$\bar{U}_0(0) = \frac{3}{2} \left| \sum_{n=-\infty}^{\infty} B_n \right|^2$$

for all cases. By considering the $k = 0$ (dc) term the quantity

$$\left| \sum_{n=-\infty}^{\infty} B_n \right|^2$$

can be measured experimentally. Then if

$$\bar{U}_k(0) \neq \bar{U}_0(0) \qquad k = 1, 2, 3 \cdots$$

the signal is not mode locked.

III. SPECIAL CASES

It is instructive to consider how much $\bar{U}_k(0)$ differs from $\bar{U}_0(0)$ for various cases. To this end we consider three special situations, namely, a pure FM signal, a model of partial mode locking, and a "chirped" pulse. Consider first the sinusoidally modulated pure FM signal:

$$\mathcal{E}(t) = \cos \{\omega_0 t + \phi \sin \omega t\}$$

$$= \sum_{n=-\infty}^{\infty} J_n(\phi) \cos \{(\omega_0 + n\omega)t\}.$$

This is just a special case of equation (1) with

$$A_n = J_n(\phi), \qquad \phi_n = 0 \quad \text{all } n.$$

The resulting expression for $\bar{U}_k(\tau)$ is therefore

$$\bar{U}_k(\tau) = \cos^2 \frac{k\omega\tau}{2} \left[\sum_{n=-\infty}^{\infty} J_n(\phi) J_{n+k}(\phi) \right]^2$$

$$+ \frac{1}{2} \left| \sum_{n=-\infty}^{\infty} J_n(\phi) J_{n+k}(\phi) \exp(jn\omega\tau) \right|^2.$$

This can be reduced, by means of a well-known Bessel function identity to

$$\bar{U}_k(\tau) = \tfrac{1}{2} J_k^2\!\left(2\phi \sin \frac{\omega\tau}{2}\right).$$

Note that $\bar{U}_0(0) = \tfrac{1}{2}$ whereas $\bar{U}_k(0) = 0$, $k = 1, 2, 3, \cdots$. Also, the largest value that $\bar{U}_k(\tau)$ $k = 1, 2, \cdots$ can assume is more than 9 dB below $\bar{U}_0(0)$ for the FM signal.

The simplest model for partial mode locking is to let the phase ϕ_n of the nth spectral line of the laser be uniformly distributed on the interval $[-\alpha, \alpha] = [-\epsilon\pi, \epsilon\pi]$. Both a Monte Carlo calculation and an analytical calculation were made. The analytical calculation is described below. The results of both calculations are presented in Fig. 3.

Consider equation (8). In order to determine the expectation value (ensemble average) $\langle \bar{U}_k(\tau)\rangle$ we must compute the values of $\langle\cos(\theta_m - \theta_n)\rangle$. Three cases arise (for $k \neq 0$):

Case I: $n = m$

$$\langle\cos(\theta_n - \theta_m)\rangle = 1.$$

Case II: $n = m \pm k$

$$\langle\cos(\theta_n - \theta_m)\rangle = \langle\cos(\phi_{n+2k} - 2\phi_{n+k} + \phi_n)\rangle \equiv a.$$
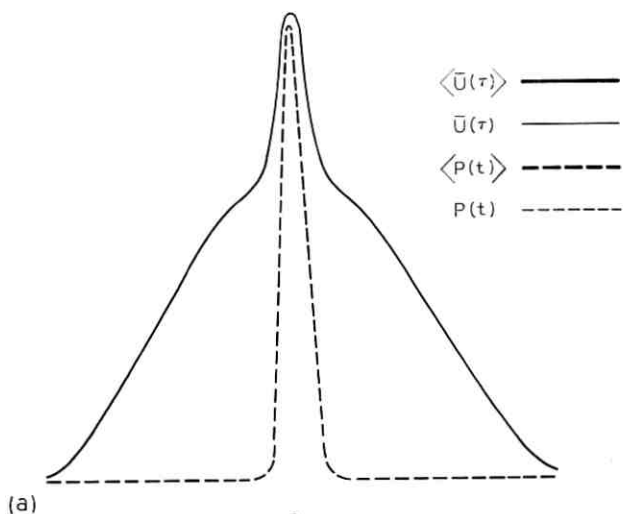
Case III: All subscripts distinct

$$\langle\cos(\theta_n - \theta_m)\rangle = \langle\cos(\phi_{n+k} - \phi_n + \phi_{m+k} - \phi_m)\rangle \equiv b.$$
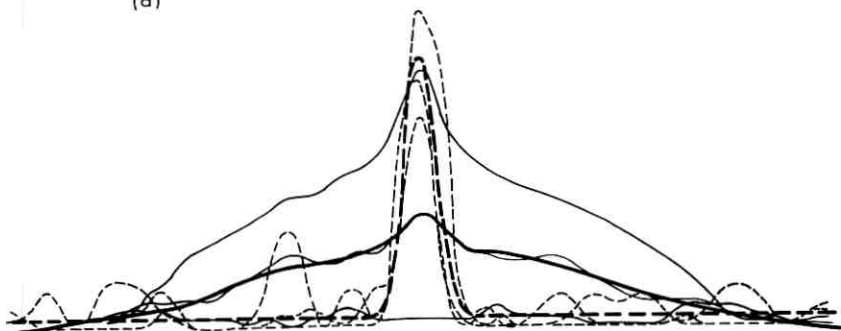
Then

$$\bar{U}_k(\tau) = b \sum_{\substack{n \\ n \neq m \\ n \neq m \pm k}} \sum_m B_n B_m\left[\cos^2\left(\frac{k\omega\tau}{2}\right) + \tfrac{1}{2}\cos(n - m)\omega\tau\right]$$

$$+ \left[\cos^2\left(\frac{k\omega\tau}{2}\right) + \frac{1}{2}\right]\sum_n B_n^2$$

$$+ 2a \sum_n B_n B_{n+k}\left[\cos^2\left(\frac{k\omega\tau}{2}\right) + \tfrac{1}{2}\cos(k\omega\tau)\right].$$
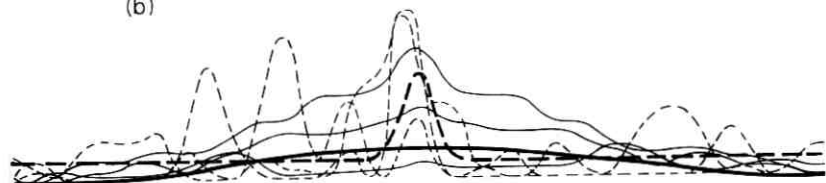
This can be rewritten in the form

$$\bar{U}_k(\tau) = G_1 \cos^2\left(\frac{k\omega\tau}{2}\right)$$

$$+ G_2 + \tfrac{1}{2}b[(\sum_n B_n \cos n\omega\tau)^2 + (\sum_n B_n \sin n\omega\tau)^2]$$
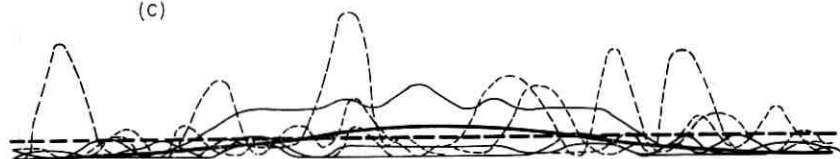
Fig. 3(a)—$\langle \bar{U}(\tau) \rangle$ and $\langle P(t) \rangle$ for ideal mode locking. (b)—$\langle \bar{U}(\tau) \rangle$, $\langle P(t) \rangle$ and three representative curves from the Monte Carlo calculation of $\bar{U}(\tau)$ and $P(t)$ for $\epsilon = 0.5$. (c)—$\langle \bar{U}(\tau) \rangle$, $\langle P(t) \rangle$ and three representative curves from the Monte Carlo calculation of $\bar{U}(\tau)$ and $P(t)$ for $\epsilon = 0.7$. (d)—$\langle \bar{U}(\tau) \rangle$, $\langle P(t) \rangle$ and three representative curves from the Monte Carlo calculation of $\bar{U}(\tau)$ and $P(t)$ for $\epsilon = 1$.

where

$$G_1 = b(\sum_n B_n)^2 + (1 - b) \sum_n B_n^2 + 4(a - b) \sum_n B_n B_{n+k} ,$$

$$G_2 = \tfrac{1}{2}(1 - b) \sum_n B_n^2 - (a - b) \sum_n B_n B_{n+k} .$$

We now turn to the problem of evaluating a and b. It can be shown*
that

$$\langle \cos (x) \rangle = \prod_{i=1}^{n} \frac{\sin \epsilon_i}{\epsilon_i}$$

where $x = \sum_{i=1}^{n} t_i$,
  t_i is uniformly distributed on $[-\epsilon_i , \epsilon_i]$,
  $\{t_i\}$ is a set of statistically independent random variables.

From this, one immediately obtains:

$$a = \left(\frac{\sin \alpha}{\alpha}\right)^2 \left(\frac{\sin 2\alpha}{2\alpha}\right) = \left(\frac{\sin \alpha}{\alpha}\right)^3 \cos (\alpha),$$

$$b = \left(\frac{\sin (\alpha)}{\alpha}\right)^4 .$$

In particular

$$\langle \bar{U}_k(0) \rangle = G_1 + G_2 + \tfrac{1}{2}b \mid \sum_n B_n \mid^2$$

$$= \tfrac{3}{2}b \mid \sum_n B_n \mid^2 + \tfrac{3}{2}(1 - b) \sum_n B_n^2 + 3(a - b) \sum_n B_n B_{n+k} . \quad (17)$$

Let

$$R_k = \langle \bar{U}_k(0) \rangle / (\tfrac{3}{2} \mid \sum_n B_n \mid^2) = \frac{\langle \bar{U}_k(0) \rangle}{\langle \bar{U}_0(0) \rangle}$$

$$= b + \frac{(1 - b) \sum_n B_n^2}{(\sum_n B_n)^2} + 2(a - b) \frac{\sum_n B_n B_{n+k}}{(\sum_n B_n)^2} . \quad (18)$$

Note that for $\alpha = 0$ (mode-locked case) $R_k = 1$ while for $\alpha = \pi$
(random-phase case)

$$R_k = \frac{\sum_n B_n^2}{(\sum_n B_n)^2} \equiv \rho_k .$$

---

* This result, even though it is not new, is derived in the Appendix.

R_k can be thought of as the "contrast ratio" in analogy with the TPF case. Consideration of the above results shows that $R_k \approx 1$ for $\alpha < 0.5$ rad. If we assume $\rho_k = 0.1$ (a reasonable value for about 10 spectral lines contributing to the signal) and neglect the third term in equation (18) [Note that this term is always negative—therefore its omission makes the result somewhat pessimistic] we obtain the following results

α(rad)	R_k
0	1
0.5	0.86
1	0.55
1.5	0.28
2	0.14
2.5	0.10

Thus the "degradation" in "contrast ratio" becomes quite pronounced by the time α exceeds one radian. In this model the expectation value of the actual pulse power is given by

$$ P(t) = \frac{1}{2} \left\{ \left(1 - \frac{\sin^2 \alpha}{\alpha^2} \right) \sum_{n=-\infty}^{\infty} A_n^2 + \frac{\sin^2 \alpha}{\alpha^2} \left| \sum_{n=-\infty}^{\infty} A_n \exp\left(jn\omega t \right) \right|^2 \right\}. $$

Thus as α departs from zero the expectation value of the power separates into two terms, one a constant term proportional to $\sum A_n^2$, the other a replica of the mode locked pulse diminished in amplitude by the factor $\sin^2 \alpha / \alpha^2$. At $\alpha = 1$ radian (where R_k is beginning to fall off as $\sin^4 \alpha / \alpha^4$) the pulse term in $\langle P(t) \rangle$ is down only 1.5 dB and falling half as rapidly.

The expectation values of $U_1(\tau)$ and $P(t)$ are plotted in Fig. 3 for the gaussian case

$$ A_n = \exp\left[-\left(\frac{nf}{f_0} \right)^2 \right] $$

where f is the pulse repetition frequency and f_0 is related to the bandwidth of the pulse. For this figure a value of 0.1 was used for the quantity f/f_0. The results of a Monte Carlo type calculation are also illustrated in this figure.

Finally we consider the case of "chirped" gaussian pulses, i.e., a repetitive train of pulses of the form

$$ E(t) = \exp\left[-\frac{1}{2} \left(\frac{t}{\tau} \right)^2 \right] \exp\left(j\beta t^2 \right). $$

This can be represented by the series

$$E(t) = \sum A_n \exp [j(n\omega t + \gamma n^2)]$$

where $A_n = \exp[-\frac{1}{2}(n\omega/\Omega)^2]$. The following result is derived for arbitrary A_n, however. Here $\phi_n = \gamma n^2$ and $\theta_n = \gamma(n^2 + 2nk + k^2 - n^2) = \gamma k(2n + k)$.

Consider

$$\left| \sum_{n=-\infty}^{\infty} \mathcal{B}_n \exp \left(j\frac{2\pi\zeta n}{T} \right) \right|^2$$

$$= \left| \sum_{n=-\infty}^{\infty} A_n A_{n+k} \exp [j\gamma k(2n + k)] \exp \left(j\frac{2\pi\zeta n}{T} \right) \right|^2$$

$$= \left| \sum_{n=-\infty}^{\infty} A_n A_{n+k} \exp \left(j\frac{2\pi nx}{T} \right) \right|^2$$

where $x = [\zeta + (T/\pi) k\gamma]$.

Comparing this with equation (12), we see that this "chirping" results in a decrease in the amplitude of the raised cosine term and a splitting in the τ term in equation (15). In fact

$$U_k(\tau) = \cos^2 \left(\frac{k\omega\tau}{2} \right) F_k\left(\frac{Tk\gamma}{\pi} \right) + \frac{1}{4}F_k\left(\frac{Tk\gamma}{\pi} + \tau \right) + \frac{1}{4}F_k\left(\tau - \frac{Tk\gamma}{\pi} \right)$$

where in this result $F_k(x)$ is evaluated using the unchirped pulse since the phase term was handled explicitly.

This result is readily understood when one recalls that $F(\tau)$ is formally equivalent to $I(t)$ with this substitution

$$A_n \leftrightarrow B_n ,$$

$$\phi_n \leftrightarrow \theta_n .$$

When A_n is gaussian, B_n is also, but when ϕ_n is quadratic in n, θ_n is linear. A phase shift which is linear in frequency corresponds to a shift of the origin in the "time" domain. This result is illustrated in Fig. 4.

The curve for $\gamma = 0.05$ corresponds to the case in which the pulse duration is ten times the reciprocal bandwidth of the signal. This is comparable to the observed values of pulse duration and bandwidth for Nd-Glass lasers. Thus this technique might provide a useful tool for investigating the amount of chirp on pulses from these lasers.

It is of interest to consider the second-order statistics of $\bar{U}(\tau)$ in order to estimate the amount of fluctuation to be expected in the random non-mode-locked case. In order to derive the second-order statistics,
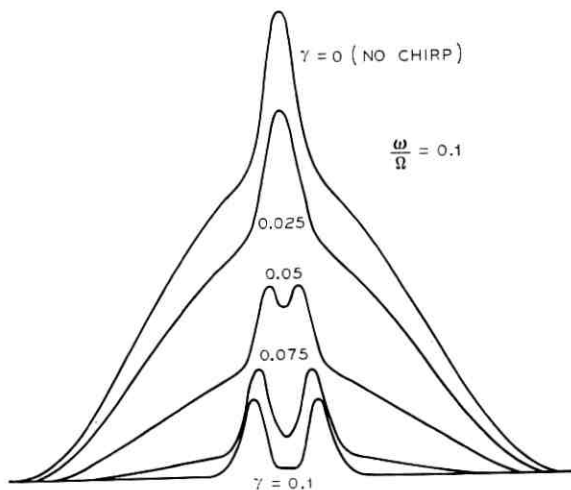
Fig. 4—$U(\tau)$ for a gaussian pulse with various amounts of chirp.

it is necessary to compute $[\bar{U}_k(\tau)]^2$ and to observe that

$$\langle \cos \{(\theta_n - \theta_m) + (\theta_{n'} - \theta_{m'})\}\rangle = 1 \quad n = m, \quad n' = m',$$
$$= 1 \quad n = m', \quad m = n',$$
$$= 0 \quad \text{otherwise.}$$

$$\langle \cos \{(\theta_n - \theta_m) - (\theta_{n'} - \theta_{m'})\}\rangle = 1 \quad n = m, \quad n' = m',$$
$$= 1 \quad n = n', \quad m = m',$$
$$= 0 \quad \text{otherwise.}$$

After some tedious manipulation one obtains the variance

$$\sigma^2 = \sum_{n \neq m} \sum \left[\cos^2\left(\frac{k\omega\tau}{2}\right) + \tfrac{1}{2}\cos(n - m)\omega\tau\right]B_n^2 B_m^2 .$$

This result is plotted in Fig. 5 for the signal illustrated in Fig. 3a.

It is also instructive to consider a somewhat different experiment. If we repeat the above experiment for the case where either the phase fronts of the signals do not coincide or the beams do not overlap at the photodetector we obtain, instead of equation (4),

$$V(t) = |\, \mathcal{E}(t)\,|^2 + |\, \mathcal{E}(t - \tau)\,|^2$$
$$= \sum_n \sum_k B_n \cos \{k\omega t + \theta_n\} \cos \frac{k\omega\tau}{2}$$
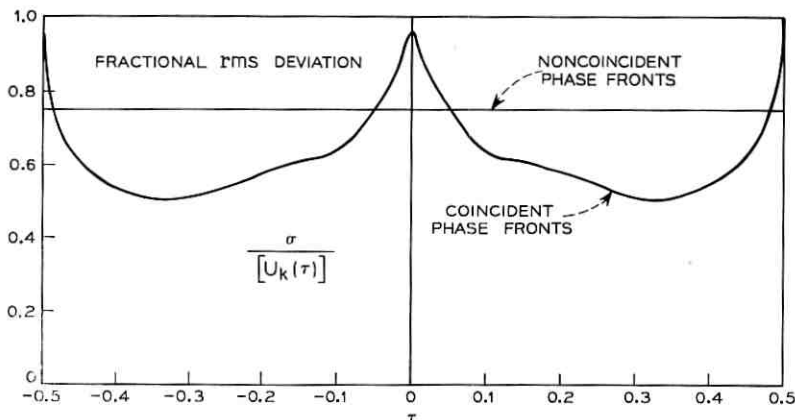
Fig. 5—Fractional rms deviation for the two statistical cases discussed in the text.

and

$$V_k(t) = \sum_n B_n \cos \{k\omega t + \theta_n\} \cos \frac{k\omega\tau}{2}$$

which upon square law detection becomes

$$U_k(\tau) = 2 \mid V_k^{(\tau)} \mid^2 = \cos^2 \frac{k\omega\tau}{2} \sum_n \sum_m B_n B_m \cos (\theta_n - \theta_m).$$

This gives the following result.

Uniformly distributed random phase:

$$\langle U_k(\tau) \rangle = \cos^2 \left(\frac{k\omega\tau}{2}\right) \sum_n B_n^2$$

$$\sigma^2 = \cos^4 \left(\frac{k\omega\tau}{2}\right)[\sum_n' \sum_{n \neq m}' B_n^2 B_m^2].$$

Mode-locked case $\phi_n = n\Phi$)

$$U_k(\tau) = \cos^2 \left(\frac{k\omega\tau}{2}\right)(\sum_n B_n)^2.$$

This case is identical in form with the Hanbury Brown and Twiss[8] experiment.

The results for the mode-locked case and the random-phase case are summarized in Table I and illustrated in Fig. 2.

TABLE I—SUMMARY OF RESULTS FOR MODE-LOCKED CASE AND UNIFORMLY DISTRIBUTED RANDOM PHASE

	Mode-Locked Case ($\phi_n = k\Phi$)	Uniformly Distributed Random Phase
Light beams add coherently on the detector, i.e., phase fronts coincide	$\bar{U}_k(\tau) = \cos^2\left(\frac{k\omega\tau}{2}\right)\left(\sum_n B_n\right)^2 + \frac{1}{2}\left|\sum_n B_n \exp(jn\omega\tau)\right|^2$	$\langle \bar{U}_n(\tau)\rangle = \left[\cos^2\left(\frac{k\omega\tau}{2}\right) + \frac{1}{2}\right]\sum_n B_n^2$ $\sigma^2 = \sum_{\substack{n\ m \\ n\neq m}}' \left[\cos^2\left(\frac{k\omega\tau}{2}\right)\cos^2\left(\frac{k\omega\tau}{2}\right) + \frac{1}{2}\cdot\cos(m-n)\omega\tau\right]^2 B_n^2 B_m^2$
Light beams do not add coherently on the detector, i.e., phase fronts do not coincide	$\bar{U}_k(\tau) = \cos^2\left(\frac{k\omega\tau}{2}\right)\left(\sum_n B_n\right)^2$	$\langle \bar{U}_k(\tau)\rangle = \cos^2\left(\frac{k\omega\tau}{2}\right)\sum_n B_n^2$ $\sigma^2 = \cos^4\frac{k\omega\tau}{2}\left[\sum_{\substack{n\ m \\ n\neq m}}' B_n^2 B_m^2\right]$

IV. EXPERIMENTAL VERIFICATION OF THE METHOD

A scaled experiment was performed using a He–Ne laser with a $c/(2\ell)$ frequency of 100 MHz. The pulse duration was measured by the BFD technique and compared with the actual pulse shape as observed by means of a high-speed photodiode and a sampling oscilloscope. Figure 6 shows the results of the experiment and Fig. 7 shows the oscilloscope display of the pulse. The pulse duration of 0.7 ns as determined from the BFD experiment is in excellent agreement with the value measured with the oscilloscope.

V. CONCLUSIONS

The experiment proposed here should be capable of distinguishing between mode-locked and non-mode-locked behavior of the laser. If mode-locked behavior is observed, a measure of the pulse width is available from the shape of the dc voltage $\bar{U}(\tau)$ versus τ curve.

The experiment is relatively simple to instrument and should be readily applicable to any CW laser whose pulse repetition rate is sufficiently low to allow construction of narrowband "IF" circuits at that frequency.
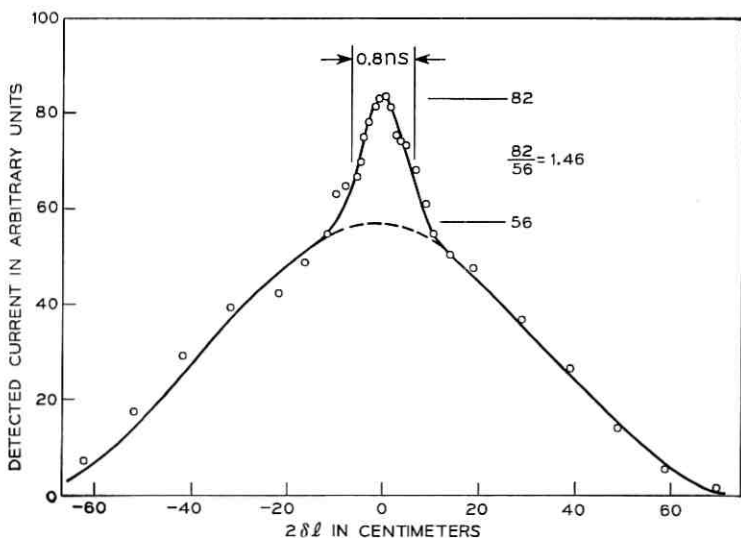

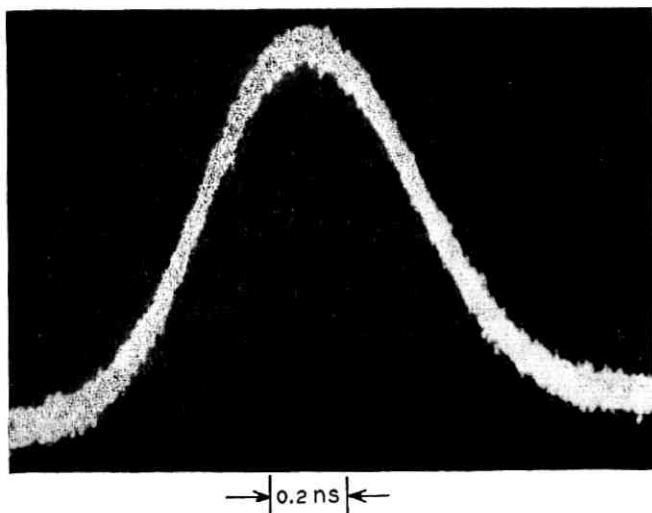Fig. 6—Experimental results for a mode-locked He-Ne laser.
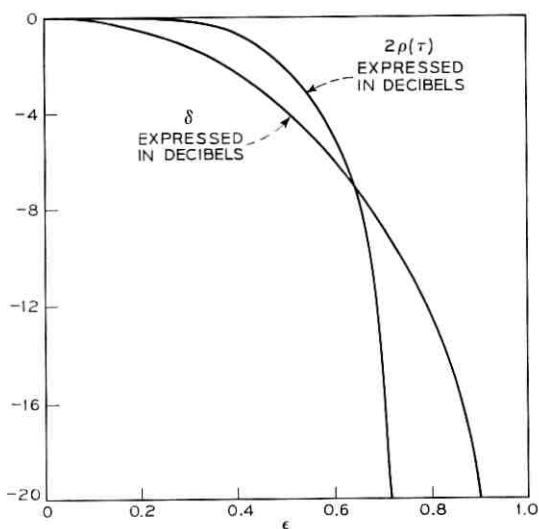
Fig. 7—Mode-locked pulse from He-Ne laser.

Fig. 8—Contrast ratio and pulse degradation as functions of epsilon.

APPENDIX

A general form for the expectation values of $\cos(x)$ such as the ones required for the calculation in the text is derived below. In particular, it is shown by induction that

$$\langle \cos(x) \mid n \rangle = \prod_{i=1}^{n} \frac{\sin(\epsilon_i)}{\epsilon_i} \tag{19}$$

where: $x = \sum_{i=1}^{n} t_i$,

each t_i is uniformly distributed on an interval $[-\epsilon_i, \epsilon_i]$, the density function

$$f_i(t) = 1/(2\epsilon_i), \qquad t \, \varepsilon \, [-\epsilon_i, \epsilon_i]$$

$$= 0 \quad \text{otherwise}$$

all of the t_i's are statistically independent.

$\langle \cos(x) \mid n \rangle$ is the expectation value of $\cos(x)$.

We calculate the $n = 1$ case:

$$\langle \cos(x) \mid 1 \rangle = \frac{1}{2\epsilon_i} \int_{-\epsilon_i}^{\epsilon_i} \cos(x) \, dx = \frac{\sin(\epsilon_1)}{\epsilon_1}. \tag{20}$$

We now assume the form (19) and show that it can be extended to the $n + 1$ case:

$$\langle \cos(x) \mid n + 1 \rangle = \int_{-\infty}^{\infty} \cos(x) F_{n+1}(x) \, dx$$

where $F_{n+1}(x)$ is the distribution function of x for $n + 1$ terms. But

$$F_{n+1}(x) = F_n(x) \, f_{n+1}(x),$$

$$= \frac{1}{2\epsilon_{n+1}} \int_{-\epsilon_{n+1}}^{\epsilon_{n+1}} F_n(x - \tau) \, d\tau.$$

Thus

$$\langle \cos(x) \mid n+1 \rangle$$

$$= \frac{1}{2\epsilon_{n+1}} \int_{-\infty}^{\infty} \cos(x) \int_{-\epsilon_{n+1}}^{\epsilon_{n+1}} F_n(x - \tau) \, d\tau \, dx$$

$$= \frac{1}{2\epsilon_{n+1}} \int_{-\epsilon_{n+1}}^{\epsilon_{n+1}} \int_{-\infty}^{\infty} \{ \cos(x) \cos(\tau) + \sin(x) \sin(\tau) \} F_n(x) \, dx \, d\tau$$

but the second term vanishes because $F_n(x)$, being generated by the successive convolution of even functions, is even. The integral reduces to

$$\langle \cos(x) \mid n+1 \rangle = \frac{1}{\epsilon_{n+1}} \int_{-\epsilon_{n+1}}^{\epsilon_{n+1}} \cos(\tau) \, d\tau \int_{-\infty}^{\infty} \cos(x) F_n(x) \, dx$$

$$= \frac{\sin(\epsilon_{n+1})}{\epsilon_{n+1}} \prod_{i=1}^{n} \frac{\sin(\epsilon_i)}{\epsilon_i}$$

which completes the proof.

REFERENCES

1. Maier, M., Kaiser, W., Giordmaine, J. A., "Intense Light Bursts in the Stimulated Raman Effect," Phys. Rev. Letters, *17*, No. 26 (December 1966), pp. 1275–1277.
2. Armstrong, J. A., "Measurement of Picosecond Laser Pulse Widths," Appl. Phys. Letters, *10*, No. 1 (January 1967), pp. 16–18.
3. Weber, H. P., "Method for Pulsewidth Measurement of Ultrashort Light Pulses Generated by Phase-Locked Laser Using Nonlinear Optics," J. Appl. Phys., *38*, No. 5 (April 1967), pp. 2231–2234.
4. Giordmaine, J. A., Rentzepis, P. M., Shapiro, S. L., and Wecht, K. W., "Two-Photon Excitation of Fluorescence by Picosecond Light Pulses," Appl. Phys. Letters, *11*, No. 7 (October 1967), pp. 216–218.
5. DeMaria, A. J., Glenn, Jr., W. H., Brienza, M. J., and Mack, M. E., "Picosecond Laser Pulses," Proc. IEEE, *57*, No. 1 (January 1969), pp. 2–25.
6. Weber, H. P., "Comments on the Pulse Width Measurement with Two-Photon Excitation of Fluorescence," Phys. Letters, *27A*, No. 5 (July 1968), pp. 321–322.
7. Klauder, J. R., Duguay, M. A., Giordmaine, J. A., and Shapiro, S. L., "Correlation Effects in the Display of Picosecond Pulses by Two-Photon Techniques," Appl. Phys. Letters, *13*, No. 5 (September 1968), pp. 174–176.
8. Hanbury Brown, R., and Twiss, R. Q., "Interferometry of the Intensity Fluctuation in Light I. Basic Theory: The Correlation Between Photons in Coherent Beams of Radiation," Proc. Roy. Soc., *242*, Series A, No. 1230 (November 1957), pp. 300–324.

Measured Attenuation and Depolarization of Light Transmitted Along Glass Fibers

By L. G. COHEN

Loss measurements have been made on glass fibers immersed in index matching oil. The measurements include state of polarization analyses of the light leaving fiber ends. The character of the beam was determined by using a λ/4 plate in conjunction with a polaroid analyser to measure the components of unpolarized and polarized light.

The data indicates that unclad fibers attenuate light much more than similar glass clad fiber waveguides. Mode purity within a multimode glass fiber was estimated from measurements of the admixture of modes leaving its end. Light guided along unclad fibers was considerably less distorted than light propagating along similar clad waveguides.

I. INTRODUCTION

The possibility of using glass fibers as dielectric waveguides in optical communication systems is well known. This paper describes the light ($\lambda = 6328$ Å) transmission properties of unclad glass fibers surrounded by index matching oil. Their characteristics are compared with those of similar clad fibers in which the cladding was glass. Two types of optical measurements were made: the attenuation and the state of polarization of the light from a fiber end.

Measurements were made on several sizes of fibers (28 μm and 13 μm diameter unclad fibers; 19 μm and 3 μm core diameter clad fibers).* The index of refraction of the oil medium surrounding unclad fibers was varied to mismatch the bulk glass by 8.94, 3.05, and 1.17 percent. The clad fibers were surrounded by glass whose refractive index was 0.824 percent lower than that of the bulk material.[1]

---

* All fibers were manufactured by DeBell and Richardson, Inc. of Hazardville, Connecticut. The bulk material of the core was SSK1 glass. Its index of refraction was measured to be 1.614 at 6328 Å.

II. EXPERIMENTAL APPARATUS AND PROCEDURE

Figure 1 illustrates the laboratory set-up used to make measurements on fibers immersed in oil. The input end of the fiber was rigidly clamped relative to the laser beam, while the output end was immersed in oil to reduce light reflections. Power output was measured with a solar cell which was loaded with a 1000 Ω impedance to ensure detector linearity. The output light power was maximized by adjusting the position of the fiber's input end with a three-dimensional micromanipulator.

Several problems arose when unclad fibers were immersed in oil. Since they are not coated, unclad fibers are sensitive to surface contamination. They were cleaned with methyl ethyl ketone and then threaded through twin iris supports. A black cardboard holding the supported fiber was then placed on top of the oil trough. The input end of the fiber was clamped rigidly and fastened to the window of the oil trough. By sliding the cardboard away from the window the supported fiber could be lowered into the trough without touching its walls. Light was focused into the fiber while it was still suspended in air so that the input tip could be easily located once it was immersed
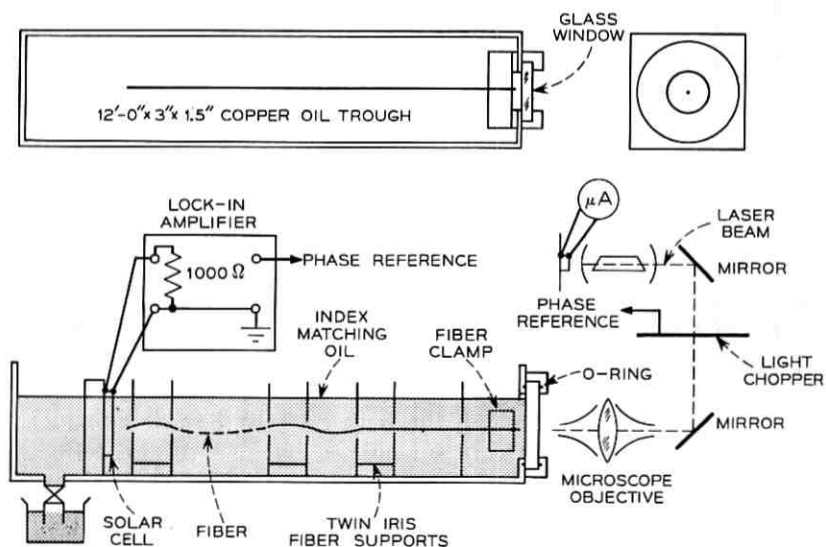

Fig. 1—Experimental arrangement for transmission measurements on fibers immersed in oil.

in index matching oil. Irised supports were employed to ensure that the light scattered from the fiber was a negligible portion of the detected light intensity. Light from the end of a fiber did not vary substantially in power as a function of the number of supports.

The refractive index of the bulk material and oil samples was measured with an Abbe refractometer. However, the refractive index of fibrous glass is in some doubt. The following procedure, illustrated in Fig. 2, was used to make a direct measurement. Laser light was focused into a 28 μm unclad fiber whose input and output tips were suspended in air. A small section of fiber was immersed in oil and its output observed. Energy transmitted through the fiber appeared to be cutoff when the refractive index of the oil cladding was adjusted to be $\Delta = 0.614$ percent less than the index of the bulk glass (i.e., $n_D = 1.6041$ at 6328 Å). This apparent deviation from the bulk glass property may have been caused by the quenching of the glass imposed by the fiber drawing process.[2]

The character of the fiber's output was analyzed by using a $\lambda/4$ plate in conjunction with a polaroid analyzer (refer to Fig. 3a and b) to measure the Stokes parameters[3] (S_T, S_1, S_2, S_3) of the light. Define I_1, I_2, I_3, I_4 to be the light powers transmitted through a polaroid sheet when it is oriented parallel to the vectors e_1, e_2, e_3, and e_4 as shown in Fig. 4. I_5 and I_6 are transmitted powers when the polaroid follows a $\lambda/4$ plate which produces a phase difference of $\pi/2$ between e_1 and e_2.

Four measurements are necessary to determine S_T, S_1, S_2, S_3:

(*i*) The total power of the beam, S_T, without the polaroid.

(*ii*) The degree of plane polarization with respect to two arbitrary orthogonal axes, $S_1 = I_1 - I_2$.

(*iii*) The degree of plane polarization with respect to a set of orthogonal axes oriented at 45° relative to the previous set, $S_2 = I_3 - I_4$.

(*iv*) The degree of circular polarization, $S_3 = I_5 - I_6$.

Partially polarized light may be split into two parts: one of power

$$S_p = (\sqrt{S_1^2 + S_2^2 + S_3^2})^{\frac{1}{2}} \tag{1}$$

which is polarized and the other of power $S_T - S_p$ which is unpolarized. The polarized fraction of light, p, can be computed from:

$$p = \frac{S_p}{S_T}. \tag{2}$$

The character of the polarized fraction of light may be expressed in
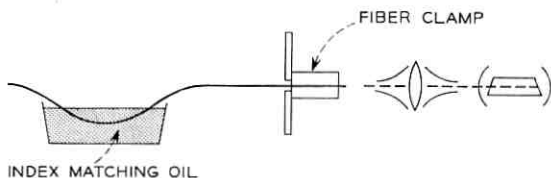
Fig. 2—Arrangement used to determine the refractive index of the unclad fiber core.

terms of the monochromatic electric field strength, **E**, relative to the orthogonal axes $\mathbf{e}_1$ and $\mathbf{e}_2$:

$$\mathbf{E} = b_1 \exp\left[i(\omega t + \delta_1)\right]\left\{\mathbf{e}_1 + \frac{b_2}{b_1} \exp\left[-i(\delta_1 - \delta_2)\right]\mathbf{e}_2\right\}. \qquad (3)$$

Equation (3) may be simplified by defining the complex polarization factor, $\bar{Q} = b_2/b_1 \exp\left[-i(\delta_1 - \delta_2)\right]$, which uniquely describes the
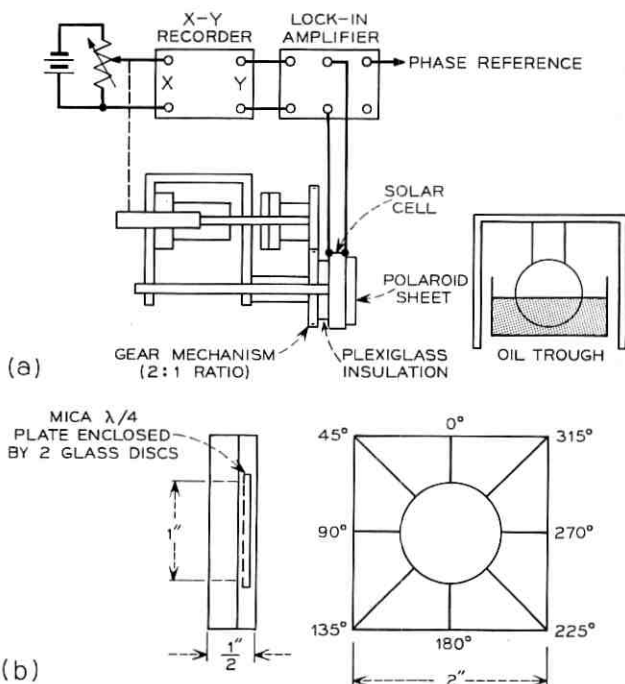
Fig. 3—Schematics of: (a) polaroid analyser and (b) $\lambda/4$ plate.
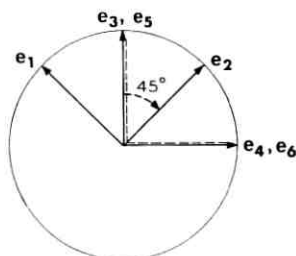
Fig. 4—Polaroid sheet orientations for Stokes parameter measurements. The direction defined by e_3 corresponded to the optic axis of the polaroid when the sheet was adjusted for maximum transmission of linearly polarized laser light. The set of orthogonal axes defined by e_1 and e_2 are oriented at 45° counterclockwise from e_3 and e_4. Axes e_5 and e_6 correspond to measurements made when the polaroid follows a $\lambda/4$ plate which produces a phase difference of $\pi/2$ between directions e_1 and e_2.

polarization of an electromagnetic wave.

$$\mathbf{E} = b_1 \exp\left[i(\omega t + \delta_1)\right]\{\mathbf{e_1} + \bar{Q}\mathbf{e_2}\}. \tag{4}$$

The magnitude and phase of $\bar{Q}$ may be expressed in terms of the Stokes parameters:[4]

$$\bar{Q} = \left|\frac{1 - \dfrac{S_1}{S_p}}{1 + \dfrac{S_1}{S_p}}\right|^{\frac{1}{2}} \exp\left(-i \tan^{-1}\right)\left(-\frac{S_3}{S_2}\right). \tag{5}$$

To make a polarization measurement, a polaroid sheet was rotated in the light to be analysed. The polaroid was mounted on the geared mechanism in Fig. 3a which was used to drive a precision potentiometer. The angular position of the polaroid sheet was recorded on the x-axis of an $x - y$ recorder. The light power transmitted through the polaroid was collected by a solar cell and drove the y-axis of the recorder.

By convention, the direction defined by e_3 in Fig. 4 corresponded to the optic axis of the polaroid when the sheet was adjusted for maximum light transmission. Curve (i) on Fig. 5, made in linearly polarized laser light, is a profile of the transmitted light power, I, as a function of the angle, θ, measured clockwise between e_3 and the polaroid's optic axis.

Curve (ii) in Fig. 5 was obtained when the polaroid followed a $\lambda/4$ plate which produced a phase difference of $\pi/2$ between e_1 and e_2 of Fig.

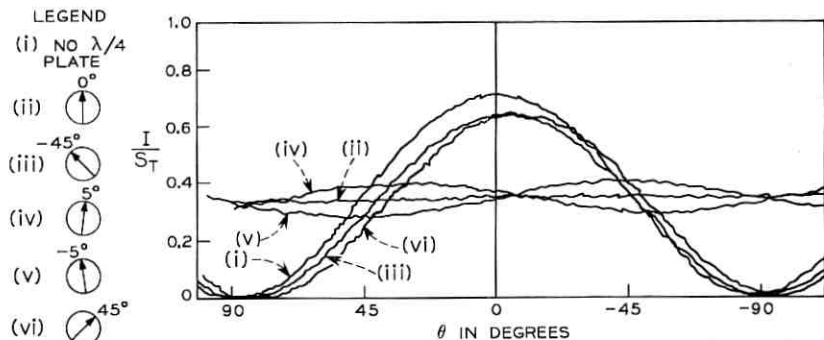Fig. 5—Polarization calibration in linearly polarized laser light. The normalized light power, I/S_T, was measured as a function of θ, the clockwise angle between the polaroid's optic axis and the vertical reference direction, e_3. A reference axis on the $\lambda/4$ plate was aligned with the direction defined by e_3 when the plate was adjusted to produce a phase difference of $\pi/2$ between e_1 and e_2. Curves ii–vi correspond to 5 different orientations of the plate's reference axis relative to e_3.

4. The linearly polarized light which was incident on the plate became circularly polarized and the amount of light transmitted by the polaroid sheet was the same for all orientations. The remaining curves of Fig. 5 correspond to four additional $\lambda/4$ plate orientations relative to e_3 . They are each consistent with a different alignment of Fig. 4. For example, the figure would have to be rotated 45° counterclockwise in order to serve as a reference for the $\lambda/4$ plate orientation corresponding to curve (iii). The polarized power of the light, S_p , should be invariant to rotation of Fig. 4. The quality of the $\lambda/4$ plate was assessed by comparing S_p computed for the five orientations: (ii) through (vi). The maximum deviation arose between (ii) and (vi) and only amounted to 1.82 percent.

Several correcting factors had to be determined before the Stokes parameters could be measured absolutely. Two different solar cells were employed. One detector was used to measure light transmitted through the polaroid (i.e., S_1 , S_2 , S_3). The other was used to measure the total power, S_T . The sensitivity and linearity of both detectors were compared in the same beam. The transmissivity, $T_{r_{pol}} = I_3/S_T = 0.748$, of the polaroid sheet in oil was measured by comparing the light power transmitted when its optic axis was parallel to e_3 with the total power. The transmissivity of the $\lambda/4$ plate in oil, $T_{r_{\lambda/4}} = I/S_T = 0.947$, was measured by comparing the light power transmitted through it with the total power.

Figure 6 illustrates typical measurements of the state of polarization

of the partially polarized light leaving sample fiber lengths. Curve (a) in Fig. 6 offers consecutive measurements for three different end cuts of the same 19 μm clad fiber which had been totally immersed in oil. It illustrates that end effects are negligible for immersed fibers.

Two profiles, of I versus θ, appear for each measurement in Fig. 6. One profile was obtained when the polaroid followed a $\lambda/4$ plate which produced a phase difference of $\pi/2$ between $\mathbf{e_1}$ and $\mathbf{e_2}$. It was used to determine S_3. The other profile was obtained without the $\lambda/4$ plate and was used to determine S_1 and S_2. The three parameters S_1, S_2, and S_3 were normalized relative to the total power, S_T, and were used to compute the polarized fraction of light, p, from equations (1) and (2) along with the complex polarization factor, $\bar{Q}$, from equation (5). Comparison of Figs. 6a and 6c with Fig. 6b illustrates that the degree of circular polarization, determined from the measurement with
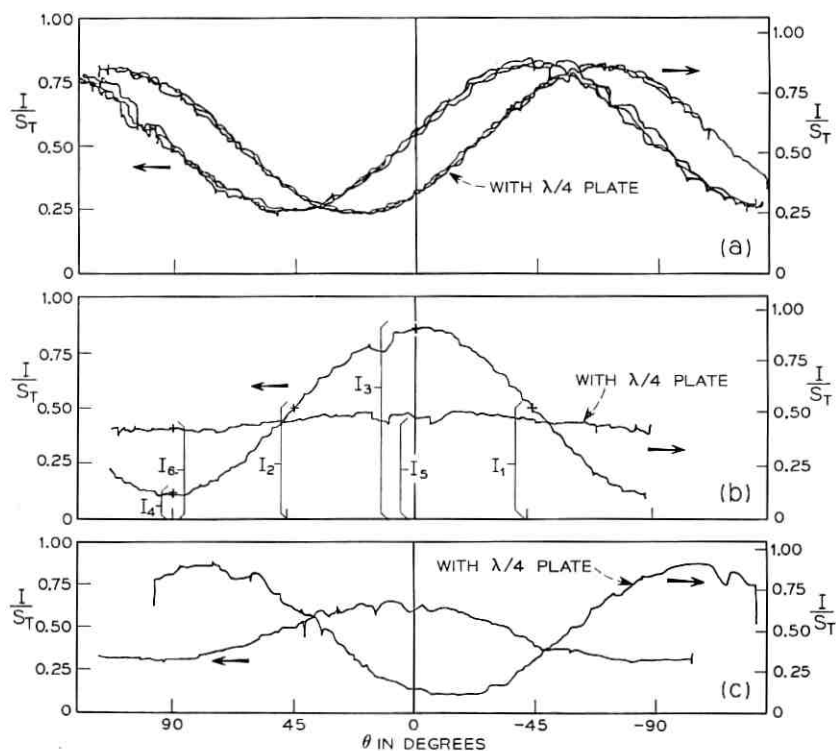
Fig. 6—Sample profiles of I/S_T vs θ for partially polarized light leaving: (a) 19μm clad fiber ($L = 301$ cm, $\Delta = 0.824\%$), (b) 28μm unclad fiber ($L = 238$ cm, $\Delta = 2.44\%$), and (c) 28μm unclad fiber ($L = 197$ cm, $\Delta = 0.56\%$).

the $\lambda/4$ plate, could vary over a wide range. The analysis based on the profiles in Fig. 6b are included in that figure. The analysed light was 75 percent polarized. Its polarized fraction was characterized by $\bar{Q} = 1.04 \exp (i2.9°)$. Such results should be independent of light power and they were approximately so. Figure 7a is a plot of p versus normalized input power for a 31-cm length of 28 μm fiber with $\Delta = 2.44$ percent. Figure 7b shows $\bar{Q}$ plotted in the complex plane with light intensity as the parameter. The points are numbered sequentially to correspond to points of Fig. 7a reading from right to left. If depolarization had not occurred the light would have been 100 percent polarized (i.e., $p \equiv 1$) and all data points in the complex $\bar{Q}$ plane would have been located at 1 | $0°$. The results of Fig. 7a and b indicates that p, arg. $\bar{Q}$, | $\bar{Q}$ | deviated by only ± 3.7 percent, $\pm 2.5°$, ± 13 percent from their mean values when the input power was doubled.
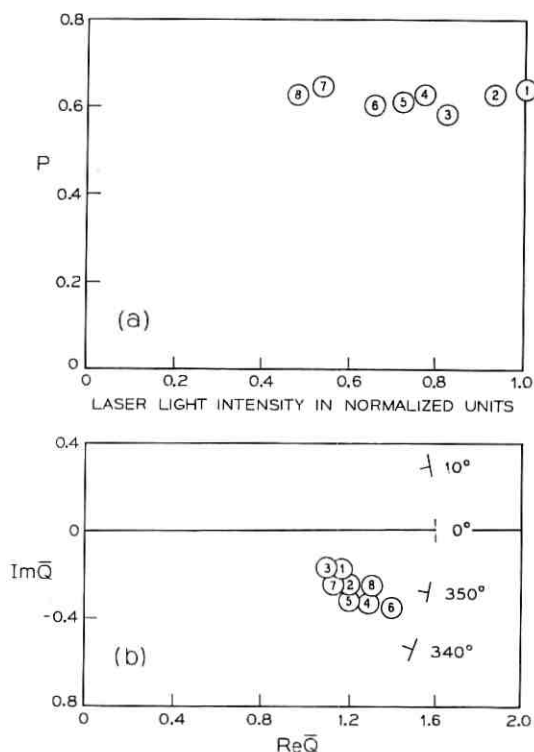
Fig. 7—The state of polarization of a 28μm unclad fiber ($L = 30$ cm, $\Delta = 2.44\%$) analysed as a function of input power: (a) p vs normalized laser light power, (b) $\bar{Q}$ is plotted in the complex plane with light power as the parameter. The points are numbered sequentially to correspond to the points of (a) reading from right to left.

III. MEASUREMENTS ON FIBERS

Optical properties of fibers were measured as a function of fiber length, L, measured from the input tip. The length was changed by cutting off known decrements of fiber. Some figures contain several sets of data points. Each set of points corresponded to a different fiber. Although laser power was not kept constant between runs, it remained stable during the course of a single set of measurements.

The transmission loss coefficient, α, of a fiber is determined from the slope of a semi-logarithmic plot of detector output, I, versus L.

Since:

$$I = I_0 \exp(-\alpha L) \tag{6}$$

then

$$\alpha = -\frac{\ln \dfrac{I}{I_0}}{L}. \tag{7}$$

Figure 8a and b illustrate transmission measurements made on two different 13 μm unclad fibers which had been immersed in oil whose refractive index was 2.44 percent lower than the index of the fibrous core at 6328 Å. The data points, labeled sequentially reading from right to left, are coded to correspond to polarization measurements made on the same fibers and recorded in Fig. 9. Light power, I, was measured in terms of the scale units of a lock-in amplifier. The attenuation coefficient of unclad fibers increased markedly and the power input, I, deviated from an exponential dependence on L for $L < 1m$. This behavior was probably related to launching phenomena. It was inconsequential since measurements were always made on lengths long enough to make transmission properties solely dependent upon fiber parameters. Occasionally, lack of linearity was noted near the ends of unclad fiber. These apparent increases in attenuation occurred along lengths of fiber which had become contaminated upon contact with a dusty surface. Figure 8c illustrates transmission measurements made on a 3 μm clad fiber. The data points are coded to correspond to polarization measurements recorded in Fig. 10. The discontinuity at data point 7 of Fig. 8c was due to a gross imperfection within the core or at the core cladding interface of the 3 μm clad fiber.

Transmission measurement results, for clad and unclad fibers, are summarized in Table I. The loss measurements made on clad fibers served to calibrate the experimental technique. They are consistent
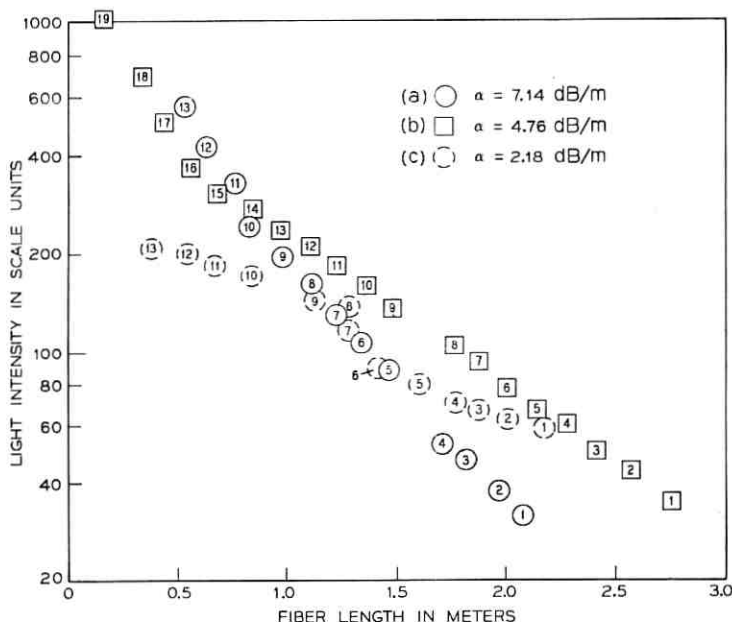
Fig. 8—(a, b) Transmission loss measurements along two different 13 μm unclad fibers immersed in oil (Δ = 2.44%). (c) Transmission loss measurements along a 3μm clad fiber.

with results originally obtained by A. R. Tynes, et al.[1] Fiber diameter (2a), the index mismatch between the core and its surroundings (Δ), the asymptotic value of the polarized fraction of light for long lengths of fiber (p_0), the characteristic number of the fiber waveguide

$$\left(R = \frac{2\pi a}{\lambda} (n_1(n_1^2 - n_2^2))^{\frac{1}{2}} \simeq (2.8) \frac{\pi a}{\lambda} n_1^{\frac{3}{2}} \Delta^{\frac{1}{2}}\right),$$

and the approximate number of propagating modes (N) appear as parameters. Loss coefficients of the oil media surrounding the immersed fibers were measured on an optical loss-gain measuring set.[5] Those results, included in Table II, and the analysis outlined in Appendix B indicate that the fraction of α due to energy propagating within the oil was less than 3 percent.

The data listed in Table I indicates that the attention constant, α, increases for decreasing fiber diameter, 2a, and index mismatch, Δ, but remains approximately constant when the preceding parameters

are changed in a manner which keeps the number of modes constant.
In general, unclad fibers were much more lossy than similar glass fiber
waveguides. Unavoidable surface contamination of the exposed sur-
faces may explain this deviation.

The guiding quality of a fiber may be estimated by measuring the
state of polarization of light leaving its end. Linearly polarized light
having a Gaussian distribution was launched into a series of multi-mode
fibers. The minimum number of propagating modes was seven for clad
fiber and 210 for unclad fiber. The input light conformed to the dominant
mode of each of the fibers tested. At best, for fibers free of irregularities
in geometry and refractive index all propagating energy would remain in
the dominant mode and $p \equiv 1 \equiv \bar{Q} \equiv 1$. For poor guiding structures
the energy originally in the dominant mode should eventually become
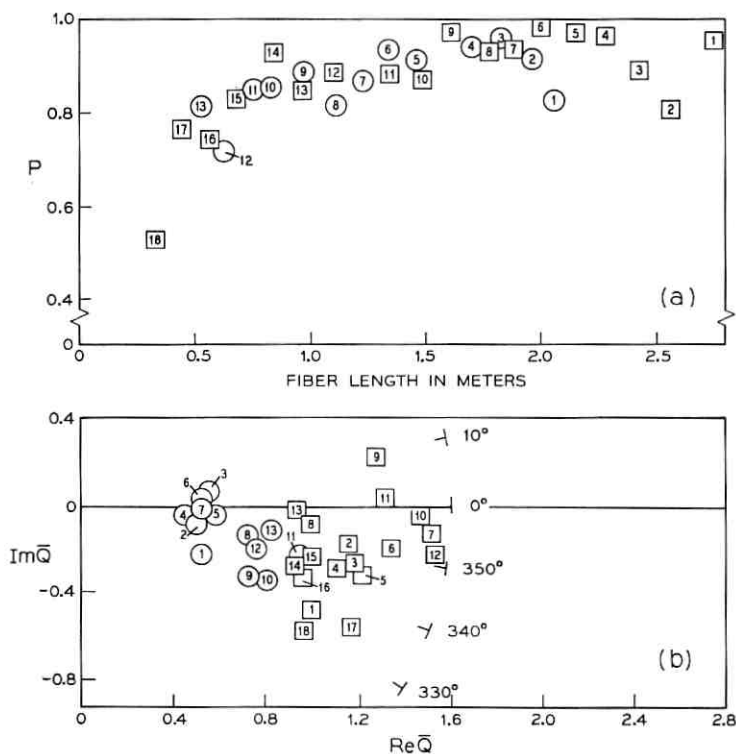
Fig. 9—Analyses of the states of polarization along two different 13μm unclad
fibers ($\Delta = 2.44\%$): (a) p vs fiber length, and (b) $\bar{Q}$ is plotted in the complex plane
with fiber length as the parameter. The points are numbered sequentially to cor-
respond to the points of (a) and Figure 8a, b reading from right to left.
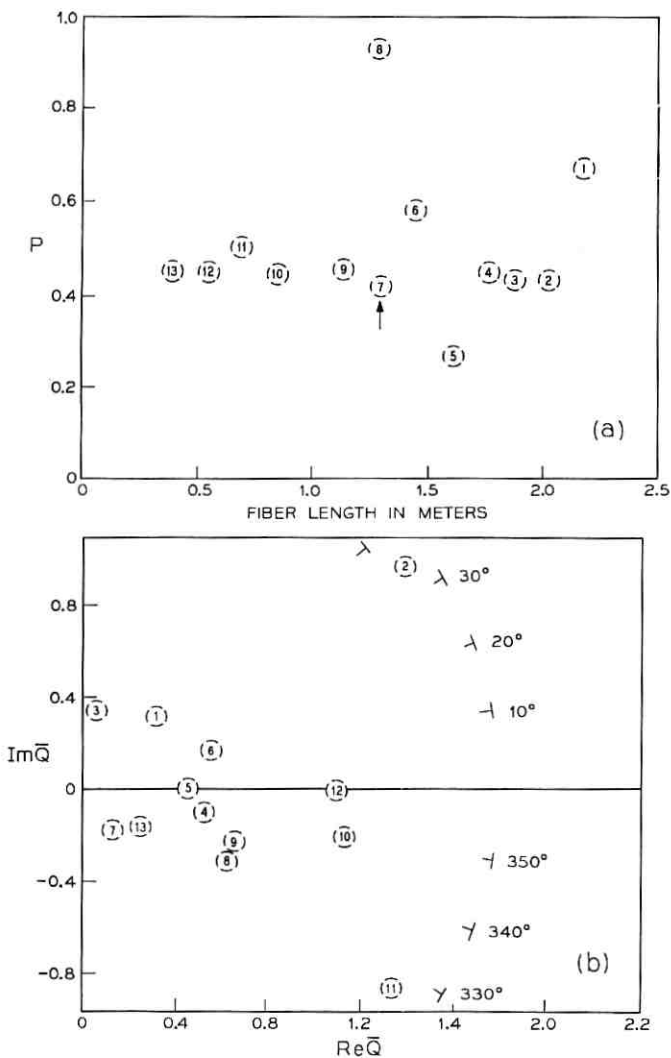
Fig. 10—Analyses of the states of polarization along a 3μm clad fiber (Δ = 0.824%): (a) *p* vs fiber length. The notation ↑ refers to a bright spot indicative of a gross defect within the core or at the core cladding interface. (b) $\bar{Q}$ is plotted in the complex plane with fiber length as the parameter. The points are numbered sequentially to correspond to the points of (a) and Figure 8c reading from right to left.

distributed among all possible propagating modes. The state of polarization for this case is analysed[6] in Appendix B assuming equal power in all modes. The analysis indicates that the fraction of polarized light, $p \to 1/2$ as the number of modes, $N \to \infty$ (i.e., $\lim_{N\to\infty} p = 1/2$).

The laser light launched into all fibers was linearly polarized and consequently the ratio between the polarized and total powers was $p = 1$. Some depolarization was observed in both clad and unclad fibers for measured lengths as small as 5 cm. In unclad fibers, the fraction of polarized light appeared to increase towards an asymptotic value, $p_0 = 1$, for increasing fiber length. The length at which light became completely polarized depended upon the total number of propagating modes. Figure 9a shows sample plots of p versus fiber length for two different 13 μm unclad fibers with $\Delta = 2.44$ percent. Transmission measurements for these fibers are included in Fig. 8a and b. Figure 9b illustrates corresponding plots of $\bar{Q}$. Points in the complex plane are numbered sequentially to correspond to the data points of 9a, reading from right to left. The fact that the imaginary component of $\bar{Q}$ was not negligible for many of the data points in Fig. 9b indicates that the polarized fraction of light, propagating within fibers, can contain significant elliptically polarized components. This type of depolarization may have been due to small guide imperfections which caused the power in the fundamental mode to become distributed among other propagating modes. The phase of a wave can be significantly distorted after propagating short distances when the phase velocity difference between modes is small, as for the slightly elliptical cross-section considered in Appendix C. Ellipticities like that could have been caused by squeezing during the fiber drawing process.

Figure 10 was laid out in the same format as Fig. 9. Corresponding transmission results are included in Fig. 8c. They apply to measurements made on a 3 μm clad fiber which had been immersed in oil to

TABLE I—SUMMARY OF OPTICAL MEASUREMENTS

2a(μm)	Unclad Fibers					Clad Fibers	
	28			13		19	3
Δ (%)	8.32	2.44	0.56	8.32	2.44	0.824	0.824
α (db/m)	2.39	3.92	4.58	3.83	5.71	1.83	2.18
p_0	not measured	1.0	1.0	not measured	1.0	0.50	0.50
R	116	62	29	53	29	25	4
N	>800	800	200	700	200	160	7

TABLE II—TRANSMISSION LOSS IN INDEX MATCHING OILS

Δ (%)	8.32	2.44	0.56
α^* (db/m)	1.85	4.96	4.14

* The transmission loss coefficient, α, of the bulk glass from which all fibers were drawn was measured to be 1.46 dB/m.[1]

eliminate end effects. Light scattering from the cladding was eliminated by matching the oil to the glass cladding. Considerably more unpolarized light propagated along clad fibers than along similar unclad fibers. In addition, the repeatability of results, for clad fibers was not nearly as good and the measurement fluctuations could be much worse than for unclad fibers. Some of these fluctuations resulted from gross imperfections within the core or at the core cladding interface. One such defect is apparent within the 3 μm clad fiber in the vicinity of data point 7.

The 3 μm clad fiber was considerably less multimode than any of the other fibers tested. Its characteristic number was $R = 3.93$ and it could only support seven modes. For some reason, however, the clad fiber was more of a depolarizer since $p \rightarrow p_0$ for lengths as small as $L = 0.4$ m. This was considerably smaller than the corresponding length, $L = 1.5$ m, for unclad fibers with $R = 29$, but was consistent with the decrease in the number of propagating modes.

The asymptotic limit, p_0, of the polarized fraction of light propagating within clad and unclad fibers has been tabulated in Table I. The results imply that unclad fibers are relatively good guiding structures. Clad fibers, by contrast, are extremely poor waveguides since linearly polarized light quickly becomes distributed equally among all possible propagating modes.

IV. CONCLUSIONS

Electromagnetic energy was more highly attenuated within unclad fibers immersed in oil than within glass clad waveguides capable of supporting the same number of modes. The attenuation constant, α, increased for decreasing fiber diameter, 2a, and index mismatch, Δ. However, α remained approximately constant when the preceding parameters were changed in a manner which kept the number of modes constant (i.e., the characteristic number $R \simeq (2.8)(\pi a/\lambda)n_1^{\frac{1}{2}} \Delta^{\frac{1}{2}} =$ constant). For the unclad fibers α varied from a high of 4.58 dB/m for a guide characterized by $R = 29$ to a low of 2.39 dB/m when $R = 116$.

For clad fibers α varied from a high of 2.18 dB/m when $R = 3.93$ to a low of 1.83 dB/m when $R = 25$.

Light leaving the end of a fiber may be divided into three components: unpolarized, linearly polarized, and circularly polarized light. The guiding quality of a multimode fiber waveguide, excited by completely linearly polarized laser light, was estimated by measuring the ratio, p, between the polarized and unpolarized light powers leaving its end. At best, $p = 1$ if the guided energy is not coupled to any unpolarized modes. For poor guiding structures, $p = 1/2$ if the guided energy becomes distributed equally among all propagating modes. Based on this criteria unclad fibers appear to be much better $(p \rightarrow 1)$ guiding structures than clad fibers $(p \rightarrow 1/2)$. This behavior in guiding quality is consistent with the attenuation measurements which indicate that glass clad fibers, with their lower loss coefficients, can contain more light modes than oil clad fibers can.

The polarized fraction of light leaving some fibers had significant circularly polarized components which could have been due to a slight cross-sectional ellipticity caused by squeezing during the fiber drawing process.

V. ACKNOWLEDGMENT

APPENDIX A

Loss Coefficient Error Due to Energy Propagating
Outside the Fiber Core

In a dielectric waveguide all but the TE_{0m} and TM_{0m} modes are hybrid. Both **E** and **H** have nonzero components along the axis of the guide. If E_z makes the dominant contribution to a transverse component of E, then the mode is designated as EH_{nm}. If H_z makes the dominant contribution, then the mode is designated as HE_{nm}.

The number and kind of modes which may propagate in a particular waveguide can be identified if the parameter,

$$R = (2\pi a/\lambda) \sqrt{[n_1(n_1^2 - n_2^2)]^{\frac{1}{2}}},$$

is known. This parameter is an upper bound on the zero of the Bessel function[7] corresponding to the highest-order propagating mode. Table I lists transmission results as a function of waveguide parameters. The

minimum value of R, for an unclad fiber, was $R = 29$. By computing the fractional amount of power propagating outside the fiber core for $R = 29$, an upper bound may be determined for the loss coefficient error.

A total of 18 circularly symmetric modes can propagate (i.e., $TE_{01} \rightarrow TE_{09}$ plus $TM_{01} \rightarrow TM_{09}$). An additional 191 hybrid modes may also propagate (the highest order hybrid modes are the $EH_{23,1}$ and $HE_{25,1}$ modes. They are practically degenerate). The fractional amount of power propagating outside the core, I_{ext}/I_{total}, may be estimated by assuming that power is distributed equally among all the modes and that all external power is carried by the doubly degenerate highest order mode:

$$\frac{I_{ext}}{I_{total}} < 1 - \frac{207}{209} = 0.01. \tag{8}$$

The loss coefficients of the core and the external medium both contribute to the transmission loss coefficient, α:

$$\alpha = \frac{I_{core}}{I_{total}} \alpha_{core} + \frac{I_{ext}}{I_{total}} \alpha_{ext}$$

$$\approx \alpha_{core}\left(0.99 + 0.01 \frac{\alpha_{ext}}{\alpha_{core}}\right). \tag{9}$$

The values of α_{ext} for the oil media in which unclad fibers were immersed are listed in Table II.

$$\left(\frac{\alpha_{ext}}{\alpha_{core}}\right)_{max} = \frac{(\alpha_{ext})_{max}}{\alpha_{bulk}} = 3.4. \tag{10}$$

The loss coefficient error, e, may be estimated after substituting equation (10) into equation (9).

$$e = \frac{\alpha - \alpha_{core}}{\alpha_{core}} \times 100 = 2.4\%. \tag{11}$$

APPENDIX B

Derivation of the State of Polarization of
Light in Terms of Its Mode Content

Consider a square fiber of side, a, which can support N^2 modes. In Section II and Fig. 4 the Stokes parameters (S_T, S_1, S_2, S_3) were defined in terms of power transmitted through a polaroid sheet. The Stokes parameters of guided light leaving the end of a square fiber

will now be expressed in terms of the electric field components of each mode in the transverse plane of the guide.

$$S_T = \int_{-a/2}^{a/2} \int_{-a/2}^{a/2} \{|E_x|^2 + |E_y|^2\}\, dx\, dy,$$

(12)

$$S_1 = I_1 - I_2 = \int_{-a/2}^{a/2} \left\{ \left| \frac{1}{\sqrt{2}}(E_y - E_x) \right|^2 - \left| \frac{1}{\sqrt{2}}(E_y + E_x) \right|^2 \right\}\, dx\, dy,$$

$$S_1 = \int_{-a/2}^{a/2} \int_{-a/2}^{a/2} -2\,|E_y E_x|\, dx\, dy,$$

(13)

$$S_2 = I_3 - I_4 = \int_{-a/2}^{a/2} \int_{-a/2}^{a/2} \{|E_y|^2 - |E_x|^2\}\, dx\, dy.$$

(14)

The polarized fraction of each mode is linearly polarized since there is no phase shift between its x and y field components. Therefore:

$$S_3 = I_5 - I_6 = 0.$$

(15)

The polarized fraction of light, p, can be computed from:

$$p = \frac{[S_1^2 + S_2^2 + S_3^2]^{\frac{1}{2}}}{S_T}$$

$$= \frac{\left[\left(\int_{-a/2}^{a/2} \int_{-a/2}^{a/2} -2\,|E_y E_x|\, dx\, dy \right)^2 + \left(\int_{-a/2}^{a/2} \int_{-a/2}^{a/2} \{|E_y|^2 - |E_x|^2\}\, dx\, dy \right)^2 \right]^{\frac{1}{2}}}{\int_{-a/2}^{a/2} \int_{-a/2}^{a/2} \{|E_x|^2 + |E_y|^2\}\, dx\, dy}.$$

(16)

For small differences in indices of refraction between the core and its cladding and for modes far from cut-off, the transverse field in a dielectric guide closely resembles the field in a similar metallic waveguide. The transverse electric field components corresponding to modes in a square metallic waveguide are:[9]

TM$_{mn}$ modes:

$$\left.
\begin{aligned}
E_{x_{mn}} &= \frac{-j}{\left[1 - \left(\dfrac{k_{c_{mn}}}{k}\right)^2\right]^{\frac{1}{2}}} \frac{k_n k}{k_{c_{mn}}^2} A \sin k_m x \cos k_n y \\[2em]
E_{y_{mn}} &= \frac{j}{\left[1 - \left(\dfrac{k_{c_{mn}}}{k}\right)^2\right]^{\frac{1}{2}}} \frac{k_m k}{k_{c_{mn}}^2} A \cos k_m x \sin k_n y
\end{aligned}
\right\} \exp\left(-ik_{z_{mn}} z + i\omega t\right),$$

(17)

TE$_{mn}$ modes:

$$\left.\begin{array}{l} E_{x_{mn}} = j \sqrt{\dfrac{\mu}{\epsilon}} \dfrac{k_n k}{k_{c_{mn}}^2} B \cos k_m x \sin k_n y \\[3mm] E_{y_{mn}} = -j \sqrt{\dfrac{\mu}{\epsilon}} \dfrac{k_m k}{k_{c_{mn}}^2} B \sin k_m x \cos k_n y \end{array}\right\} \exp\left(-ik_{x_{mn}}z + i\omega t\right) \quad (18)$$

where:

$$k_m^2 + k_n^2 + k_{z_{mn}}^2 = n_1^2 k^2, \tag{19}$$

$$k_m = \frac{m\pi}{a}, \qquad k_n = \frac{n\pi}{a}, \qquad k_{c_{mn}} = (k_m^2 + k_n^2)^{\frac{1}{2}}.$$

The following results apply to both TE and TM modes.

By using the orthogonality property of sines and consines one can show that:

$$\int_{-a/2}^{a/2} \int_{-a/2}^{a/2} | E_{x_{mn}} E_{y_{mn}} | \, dy = 0, \tag{20}$$

$$\int_{-a/2}^{a/2} \int_{-a/2}^{a/2} | E_{x_{mn}} |^2 \, dx \, dy \propto \left(\frac{a}{2}\right)^2, \tag{21}$$

$$\int_{-a/2}^{a/2} \int_{-a/2}^{a/2} | E_{y_{mn}} |^2 \, dx \, dy \propto \left(\frac{a}{2}\right)^2. \tag{22}$$

Substitute equations (20), (21) and (22) into equation (16):

$$p_{mn} = \left| \frac{k_m^2 - k_n^2}{k_m^2 + k_n^2} \right| = \left| \frac{m^2 - n^2}{m^2 + n^2} \right| \tag{23}$$

where: $0 \leqq m \leqq N$, $0 \leqq n \leqq N$ but $m = n \neq 0$.

Equation (23) is an expression for the polarized fraction of any mode within a square metallic waveguide. It is approximately correct for square dielectric guides if $\Delta \ll 1$ and if the mode is far from cut-off.

As an example assume that the total power within the guide is distributed equally among all propagating modes. Equation (23) may be used to compute p by averaging p_{mn} over all propagating modes.

$$p = \frac{1}{N^2} \sum_{\substack{m=0, n=0 \\ m=n \neq 0}}^{m=N, n=N} p_{mn} \approx \frac{1}{N^2} \int_{m=0}^{N} \int_{n=0}^{N} \left| \frac{m^2 - n^2}{m^2 + n^2} \right| dm \, dn. \tag{24}$$

The integration may be simplified after referral to Fig. 11 which illustrates values of p_{mn} projected onto the $m - n$ plane. p_{mn} is constant along radial lines passing through the origin at $m = n = 0$. In the limit $N \to \infty$ the integral over the $N \times N$ square may be replaced
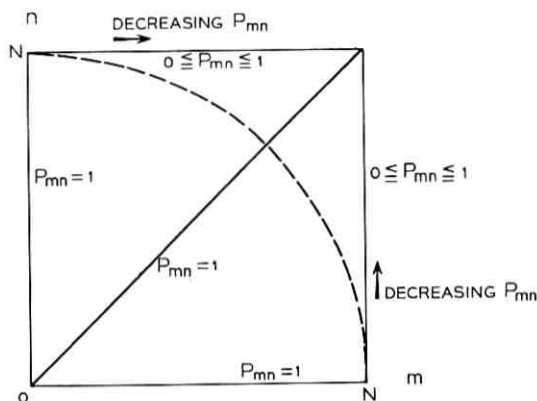
Fig. 11—Values of p_{mn} are shown projected onto the $m - n$ plane. If $N \to \infty$ then the $N \times N$ square domain of definition for p_{mn} may be replaced by the dashed circle of radius N.

by an integral over a circle of radius N and centered at the origin. Define:

$$m = R \cos \Phi, \qquad n = R \sin \Phi, \tag{25}$$

$$p \approx \frac{2}{N^2} \int_{\Phi=0}^{\pi/4} \int_{R=0}^{N} p_{mn} R \, dR \, d\Phi. \tag{26}$$

Now substitute equation (25) into equation (26)

$$\lim_{N \to \infty} p = \frac{2}{N^2} \int_{\Phi=0}^{\pi/4} \int_{R=0}^{N} (\cos^2 \Phi - \sin^2 \Phi) R \, dR \, d\Phi = \tfrac{1}{2}. \tag{27}$$

APPENDIX C

Circular Polarization Induced in a Slightly Elliptical Fiber

Consider a fiber of elliptical cross section having major and minor axes a, b. Two fundamental modes polarized along a and b have a phase shift $(k_z - k_z')L$ after propagating a distance, L.

To calculate $(k_z - k_z')$ we roughly approximate the fiber's elliptical cross-section by a rectangle of sides a and b. From Ref. 8 we get the axial propagation constant of the dominant mode polarized along a:

$$k_z = \frac{2\pi}{\lambda} n_1 \left[1 - \left(\frac{\lambda}{2an_1} \right)^2 \left(1 + \frac{2A}{\pi a} \right)^{-2} \right]^{\frac{1}{2}} \tag{28}$$

where:

$$A = \frac{1}{2^{\frac{1}{2}} n_1} \frac{1}{\sqrt{\Delta}}.$$

In multi-mode guides:

$$\frac{2A}{\pi a} \ll 1 \tag{29}$$

and

$$k_z \approx \frac{2\pi}{\lambda} n_1 \left\{ 1 - \left(\frac{\lambda}{2an_1}\right)^2 \frac{1}{2} \left(1 - \frac{4A}{\pi a}\right) \right\}. \tag{30}$$

Thus:

$$(k_z - k_z')L \approx \frac{\pi\lambda}{4n_1} L \left\{ \left(\frac{1}{b^2} - \frac{1}{a^2}\right) + \frac{4A}{\pi a^2} \left(\frac{1}{a} - \frac{a^2}{b^2}\frac{1}{b}\right) \right\}. \tag{31}$$

Assume:

$$\frac{b-a}{a} \ll 1, \quad \frac{1}{b} = \frac{1}{a+(b-a)} \approx \frac{1}{a}\left\{1 - \frac{b-a}{a}\right\}. \tag{32}$$

Then:

$$(k_z - k_z')L \approx \frac{\pi}{2n_1}\frac{\lambda}{a}\frac{L}{a}\left(\frac{a-b}{a}\right)\left\{1 - \frac{6A}{\pi b}\right\}. \tag{33}$$

Choose as an example:

$$a = 28\mu m, \quad L = 1m, \quad n_1 = 1.604, \quad \left(\frac{a-b}{a}\right) = 0.01, \quad \Delta = 0.0056.$$

The phase shift after one meter is:

$$(k_z - k_z')L = 2.52\pi(1 - 0.127). \tag{34}$$

This phase shift is significant even for the small ellipticity assumed.

REFERENCES

1. Tynes, A. R., Pearson, A. D., and Bisbee, D. L., to be published in J. Opt. Soc., February 1971.
2. Otto, W. H., "Compaction Effects in Glass Fibers," J. Am. Cer. Soc., *44*, No. 2 (February 1961), pp. 68–72.
3. Ditchburn, R. W., *Light*, New York: Interscience Publishers, Inc., 1963, pp. 480–483.
4. Beckmann, P., *The Depolarization of Electromagnetic Waves*, Boulder, Colorado: The Golem Press, 1968, pp. 24–35.
5. Tynes A. R., and Bisbee, D. L., unpublished work.
6. Marcatili, E. A. J., unpublished work.
7. Beattie, C. L., "Table of First 700 Zeroes of Bessel Functions," B.S.T.J., *37*, No. 3 (May 1958), pp. 689–697.
8. Marcatili, E. A. J., "Dielectric Rectangular Waveguide and Directional Coupler for Integrated Optics," B.S.T.J., *48*, No. 7 (September 1969), pp. 2071–2102.
9. Ramo S., and Whinnery, J. R., *Fields and Waves in Modern Radio*, New York: Wiley, Inc., 1953, p. 366.

Excitation of Waveguides for Integrated Optics with Laser Beams

By D. MARCUSE and E. A. J. MARCATILI

We discuss in this paper the excitation of dielectric waveguides used for integrated optical circuitry. Thin optical films are usually excited by means of coupling prisms using frustrated total internal reflection. We propose instead excitation of thin-film waveguides directly by gaussian laser beams. Harmful effects of imperfect substrate edges can be avoided by letting the dielectric waveguide end inside of the substrate. The exciting laser beam is directed towards the end of the guide from the inside of the substrate material.

The analysis is based on neglecting reflection at the end of the thin-film waveguide and on assuming that the thin film is infinitely extended in one dimension. The maximum excitation efficiency predicted with this model is 97 percent. It is expected that the excitation efficiency of more realistic guides for integrated optics can be as high as 90 percent.

I. INTRODUCTION

Communications systems using light waves as the carrier of information need some means of signal processing at the end terminals and perhaps also at intermediate repeater points along the transmission lines. It has been suggested[1,2] to employ integrated optical circuits for the purpose of filtering, amplifying, pulse regeneration, etc., of the optical signal beam.

These introductory remarks make it plausible that the need exists for exciting guided modes in the waveguides used for integrated optical circuits.[3] An efficient method of mode excitation utilizes the evanescent field outside of a high-index prism to couple light energy from a laser beam to one of the guided modes of a thin-film waveguide.[4] This method is particularly suitable for wide thin-film guides. However, for the very narrow light guides that are likely to be used for integrated optical circuits the prism coupler method may become less

efficient since the dielectric optical waveguide may easily be narrower than the laser beam. A very narrow beam must be highly focused and thus has the disadvantage of a high-beam divergence contrary to the requirements of the prism coupler.

An alternate method of exciting the guided modes of dielectric waveguides is by shining the laser beam directly at the end of the guide. It has been shown theoretically that the conversion efficiency obtainable by this method is as high or higher than that of the prism coupler.[5,6] Light injection by direct excitation with a laser beam is usually employed with cladded round optical fibers. The application of the same method to the dielectric waveguides of integrated optical circuits suffers from the disadvantage that the ends of these guides are likely to be of poor optical quality. Figure 1 shows the geometry of one possible form of an integrated optical waveguide. This guide can be produced by diffusion, by sputtering or evaporation techniques.

Unless the end of the substrate of the guide is optically polished after the deposition of the higher refractive index region, the end of the resulting guide can be expected to appear as shown in the figure. The optical quality of the end of the guide can be improved if the guide is not allowed to extend all the way to the end of the substrate but instead terminates inside of it as shown in Fig. 2. Even though this method may avoid the problem of poor optical quality of the edge of the substrate, it introduces the new problem of how to excite the dielectric waveguide whose end is not easily accessible from the outside.

A solution to this problem is shown in Fig. 3. The laser beam is incident not from the outside of the dielectric substrate but from inside of it. The problem of injecting the beam into the substrate is not as severe as the problem of shining a laser beam directly on the end of the waveguide shown in Fig. 1. The beam enters the substrate
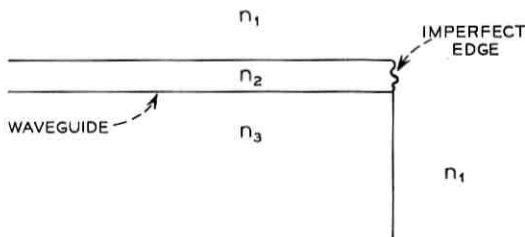
Fig. 1—Thin-film waveguide located directly under the surface of the substrate extending all the way to its edge.
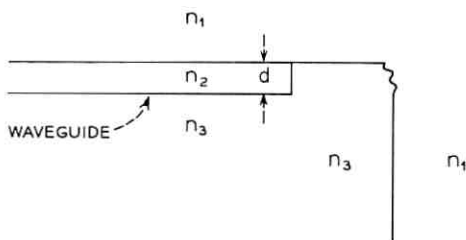
Fig. 2—Thin-film waveguide not reaching to the edge of the substrate.

not at its edge but in a region that is more easily kept free of damage. An external focusing system is, of course, required to achieve a beam of the desired convergence angle.

To study the required beam pattern for the efficient excitation of the optical waveguide, we reverse the problem and study instead the radiation pattern produced by a guided mode reaching the end of the waveguide of Fig. 3. The shape of the laser beam used for mode excitation must be made as similar as possible to the radiation pattern of the waveguide with the only exception that the direction of travel of the light field be reversed.

In the bulk of the paper we discuss the far field radiation pattern of a guided mode leaving the end of the waveguide inside of the substrate material. For simplicity we assume that the waveguide has the form of a slab that is infinitely extended in a direction perpendicular to the plane of Fig. 3. The excitation problem of the actual waveguide is very similar to that of the slab. In particular it is well known that a mode can be excited with high efficiency by an external laser beam. The only complication introduced into the problem under discussion is the presence of the air-dielectric interface of the sub-
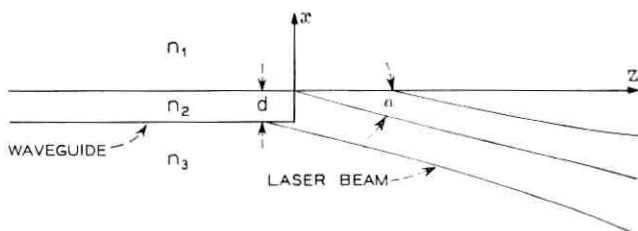
Fig. 3—Thin-film waveguide geometry used for the calculation.

strate material. This complication exists only in the direction parallel to the plane of the drawing. The mode matching problem in the direction perpendicular to the plane of the figure is no more difficult than that treated in earlier publications.[6]

A brief discussion of slightly different waveguide geometries can be found in the section on conclusions at the end of the paper.

II. MODES OF THE DIELECTRIC MEDIUM

The radiation pattern of the guided mode leaving the end of the dielectric waveguide could be computed with the conventional methods of the Kirchhoff-Huygens diffraction integral if it were not for the presence of the air-dielectric interface of the substrate material. A substantial part of the radiation field is reflected at this interface complicating the beam pattern. It appears more natural to employ the normal mode method for the solution of the radiation problem. This method utilizes the fact that any field can be expressed as a superposition of a complete set of normal modes. The air-dielectric interface is thus included automatically. We assume that the difference between the refractive index of the waveguide and the substrate is so slight that reflection at the end of the waveguide can be neglected. The radiation field is then obtained by the requirement that the transverse electric field component must be continuous in the plane $z = 0$ of Fig. 3. The radiation field is expressed as a superposition of the normal modes of the structure to the right of the end of the waveguide. No guided modes exist in this region so that the superposition of modes must be expressed as an integral over the continuum of radiation modes.

Since we have simplified the problem to that of a slab infinitely extended in y direction, we impose the condition that there is no field variation in that direction

$$\frac{\partial}{\partial y} = 0. \tag{1}$$

The guided mode field is assumed to be of the form

$$
\left.
\begin{aligned}
E_y &= A_g \exp(-\gamma x) & 0 \leq x \leq \infty \\
E_y &= A_g\left(\cos \kappa x - \frac{\gamma}{\kappa}\sin \kappa x\right) & -d \leq x \leq 0 \\
E_y &= A_g\left(\cos \kappa d + \frac{\gamma}{\kappa}\sin \kappa d\right)\exp[\theta(x+d)] & -\infty \leq x \leq -d
\end{aligned}
\right\} z \leq 0
\tag{2}
$$

with

$$A_g = 2\kappa \left\{ \frac{\omega\mu_0 P}{\beta_0 \left(d + \frac{1}{\gamma} + \frac{1}{\theta}\right)(\kappa^2 + \gamma^2)} \right\}^{\frac{1}{2}}. \tag{3}$$

A factor $\exp\left[i(\omega t - \beta_0 z)\right]$ has been suppressed. The constants and parameters appearing in these equations have the following meaning:

$\omega = 2\pi f$ radian frequency,
$P =$ power carried by the mode,
$\mu_0 =$ magnetic susceptibility of free space,
$\beta_0 =$ propogation constant,
$d =$ slab thickness,

$$\kappa = (n_2^2 k^2 - \beta_0^2)^{\frac{1}{2}}, \tag{4}$$

$$\gamma = (\beta_0^2 - n_1^2 k^2)^{\frac{1}{2}}, \tag{5}$$

$$\theta = (\beta_0^2 - n_3^2 k^2)^{\frac{1}{2}}, \tag{6}$$

$k = \omega(\epsilon_0\mu_0)^{1/2}$ free space propogation constant,
$n_1 =$ refrective index in the region $x > 0$,
$n_2 =$ refractive index in the region $-d < x < 0$,
$n_3 =$ refractive index in the region $-\infty < x < -d$.

The magnetic field components follow from the relations

$$H_x = \frac{-i}{\omega\mu_0} \frac{\partial E_y}{\partial z}, \tag{7}$$

$$H_z = \frac{i}{\omega\mu_0} \frac{\partial E_y}{\partial x}. \tag{8}$$

The values of the propagation constant β_0 are obtained from the eigenvalue equation

$$\tan \kappa d = \frac{\gamma + \theta}{\kappa\left(1 - \frac{\gamma\theta}{\kappa^2}\right)}. \tag{9}$$

In addition to the guided modes there is also a continuum of radiation modes in the region $z < 0$. However, these modes are of no interest to us.

To the right of the region where the dielectric waveguide has ended ($z > 0$) no guided modes are possible. The continuum of radiation modes can be subdivided into three distinct groups. We use the propa-

gation constant β of the continuum modes for their classification. In the region

$$0 \leq \beta < n_1 k \tag{10}$$

the following radiation modes exist:

$$\mathcal{E}_y^{(1)} = \begin{cases} A_r \cos \rho x & 0 \leq x < \infty, \\ A_r \cos \sigma x & -\infty < x \leq 0, \end{cases} \tag{11}$$

and

$$\mathcal{E}_y^{(2)} = \begin{cases} i A_r \sqrt{\dfrac{\sigma}{\rho}} \sin \rho x & 0 \leq x < \infty, \\ i A_r \sqrt{\dfrac{\rho}{\sigma}} \sin \sigma x & -\infty < x \leq 0. \end{cases} \tag{12}$$

A factor $\exp [i(\omega t - \beta z)]$ has again been suppressed. The amplitude A_r is related to power by the relation

$$A_r = 2\left(\frac{\omega \mu \rho P}{\pi \beta (\sigma + \rho)}\right)^{\frac{1}{2}}. \tag{13}$$

However, it must be pointed out that the power carried by one individual radiation mode is infinite. The power P appearing in equation (13) is defined by the expression

$$P\delta(\rho - \rho')\delta_{\nu\nu'} = \frac{\beta}{2\omega\mu} \int_{-\infty}^{\infty} \mathcal{E}_y^{(\nu)*}(\rho) \, \mathcal{E}_y^{(\nu')}(\rho') \, dx. \tag{14}$$

$[\delta(\rho - \rho')$ is Dirac's delta function, $\delta_{\nu\nu'}$, is the Kronecker delta symbol.] Equation (14) expresses the orthogonality of the radiation modes. It implies that the modes (11) and (12) are not only orthogonal among each other but also that the modes given by (11) are orthogonal to the modes given by equation (12). The parameters ρ and σ are related to the propagation constant β by the following expressions

$$\rho = (n_1^2 k^2 - \beta^2)^{\frac{1}{2}} \tag{15}$$

and

$$\sigma = (n_3^2 k^2 - \beta^2)^{\frac{1}{2}}. \tag{16}$$

Whereas there are two sets of radiation modes in the interval given by equation (10) there is only one set of radiation modes in the interval

$$n_1 k \leq \beta \leq n_3 k. \tag{17}$$

This set is given by the equation

$$\mathcal{E}_y^{(3)} = \begin{cases} B_r \exp(-\delta x) & 0 \leq x < \infty, \\ B_r\left(\cos \sigma x - \dfrac{\delta}{\sigma} \sin \sigma x\right) & -\infty < x \leq 0, \end{cases} \quad (18)$$

with

$$B_r = \left(\frac{4\omega\mu\sigma\delta P}{\pi\beta(\sigma^2 + \delta^2)}\right)^{\frac{1}{2}}. \quad (19)$$

The parameter σ is given by equation (16) while δ is defined as

$$\delta = (\beta^2 - n_1^2 k^2)^{\frac{1}{2}}. \quad (20)$$

The modes (18) are orthogonal among each other as well as to the modes (11) and (12).

Actually the radiation modes (11) and (12) cover also the range along the imaginary β axis from 0 to $i \infty$. However, the modes belonging to this imaginary branch of β values do not carry power and are thus of no interest to our investigation.

A more detailed study reveals that the radiation modes of equations (11) and (12) do not take part in forming the main lobe of the radiation field in the space $z > 0$. The main lobe is obtained by a superposition of the radiation modes (18).

III. CALCULATION OF THE RADIATION PATTERN

We simplify the calculation of the radiation pattern (resulting from the guided mode carried in the waveguide at $z < 0$ spilling over into the half space $z > 0$) by matching the transverse component E_y of the electric field at the plane $z = 0$. The continuity condition for the transverse **H** component is ignored. This approximation is acceptable if reflection at the plane $z = 0$ can safely be neglected.

The continuity condition for the transverse **E** field can be expressed by the equation

$$E_y = \int_0^{(n_2{}^2 - n_1{}^2)^{\frac{1}{2}}k} q(\delta)\, \mathcal{E}_y^{(3)}(x, \delta)\, d\delta. \quad (21)$$

The field on the left side of this equation is the guided mode at $z = 0$ while the field on the right side is the radiation field at $z = 0$ expressed as a superposition of radiation modes. Only the main radiation lobe is included by restricting the mode expansion to the modes of

equation (18). The much smaller side lobes are, however, of no interest to us.

Using the orthogonality of the radiation modes we obtain the expansion coefficient $q(\delta)$ from equation (21)

$$
q(\delta) = \frac{\beta}{2\omega\mu P} \int_{-\infty}^{\infty} E_\nu(x)\, \mathcal{E}_\nu^{(3)}(x,\,\delta)\, dx = 2\kappa(\beta\sigma\delta)^{\frac{1}{2}}(n_2^2 - n_3^2)k^2
$$

$$
\cdot \frac{(\gamma - \delta)(\kappa^2 + \delta^2)^{\frac{1}{2}} + (\kappa^2 + \gamma^2)^{\frac{1}{2}}\left[(\theta + \delta)\cos\sigma d + \left(\dfrac{\theta\delta}{\sigma} - \sigma\right)\sin\sigma d\right]}{\left[\pi\beta_0\left(d + \dfrac{1}{\gamma} + \dfrac{1}{\theta}\right)(\sigma^2 + \delta^2)(\kappa^2 + \gamma^2)(\kappa^2 + \theta^2)\right]^{\frac{1}{2}}(\kappa^2 - \sigma^2)(\gamma^2 - \delta^2)}.
$$

$$
\tag{22}
$$

In principle the radiation problem is thus solved. Substitution of equations (22) and (18) into (21) (restoring the omitted exponential z dependence) allows us to calculate the main lobe of the radiation field at any point $z > 0$. An exact evaluation of the integral is not possible. However, a good far field approximation can be obtained with the method of stationary phase. The result of this far field approximation is $(x < 0, z > 0)$

$$
E_\nu = \frac{xz}{r^{5/2}} \frac{2^{3/2}\kappa(\omega\mu P)^{1/2}(n_2^2 - n_3^2)n_3^{3/2}k^{7/2}\exp\left\{i\left[-n_3kr + \arctan\dfrac{\delta}{\sigma}\right]\right\}}{\left[\pi\beta_0\left(d + \dfrac{1}{\gamma} + \dfrac{1}{\theta}\right)(\kappa^2 + \gamma^2)(\kappa^2 + \theta^2)(\sigma^2 + \delta^2)\right]^{1/2}(\kappa^2 - \sigma^2)(\gamma^2 - \delta^2)}
$$

$$
\cdot\left\{(\gamma - \delta)(\kappa^2 + \theta^2)^{1/2} + \right.
$$

$$
\left. (\kappa^2 + \gamma^2)^{1/2}\cdot\left[(\theta + \delta)\cos\sigma d + \left(\dfrac{\theta\delta}{\sigma} - \sigma\right)\sin\sigma d\right]\right\}. \tag{23}
$$

The parameters entering this equation are now related to the x, z coordinates

$$
\left.\begin{aligned}
\beta &= n_3 k \frac{z}{r} \\[6pt]
\sigma &= -n_3 k \frac{x}{r} \\[6pt]
\delta &= (n_3^2 z^2 - n_1^2 r^2)^{\frac{1}{2}} \frac{k}{r} \\[6pt]
r &= (x^2 + z^2)^{\frac{1}{2}}
\end{aligned}\right\}. \tag{24}
$$

Equation (23) is valid only in the region where δ is real. Equation (23) does not have any poles since the numerator vanishes simultaneously with the denominator.

Figure 4 shows plots of the absolute value of E_y as a function of the angle

$$\alpha = \arctan \frac{|x|}{z} \tag{25}$$

for the case $n_1 = 1$, $n_2 = 1.51$ and $n_3 = 1.5$ and for several values of kd. The most interesting aspect of these curves is their width and the position of their maxima. It is easiest to visualize the meaning of these curves by assuming that the wavelength is fixed but that the guide width d is changing. For small values of d the guided mode field extends far into the substrate medium with refractive index n_3. A wide aperture field leads to a narrow radiation lobe. This explains the narrow radiation pattern for the curve with $kd = 9$. As the width d increases the guided mode contracts at first giving rise to a wider radiation pattern. At $kd = 20$ the radiation pattern assumes its greatest width. As d increases even more the guided mode field again becomes wider and thus creates a narrower radiation lobe.

The position of the maximum is shown as the solid line in Fig. 5. A qualitative explanation for the position of the radiation maximum can be given as follows. The guided mode field inside of the medium
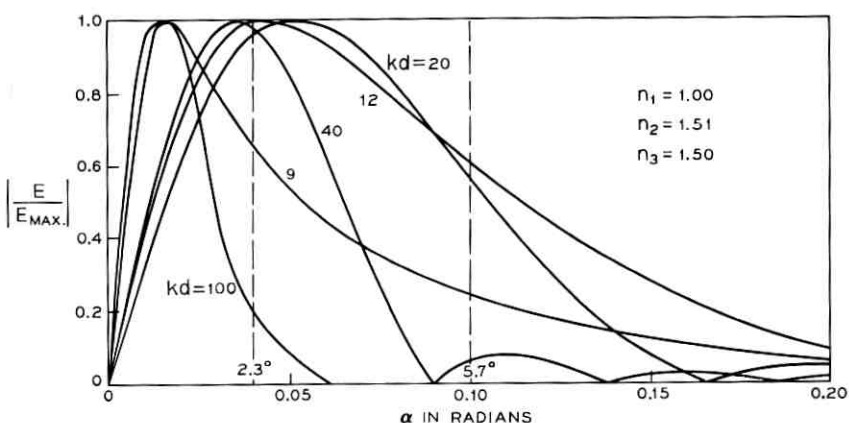
Fig. 4—Ratio of E_y over the peak value of the electric vector of the radiation field as a function of the angle in radians. $n_1 = 1.00$, $n_2 = 1.51$, $n_3 = 1.50$.
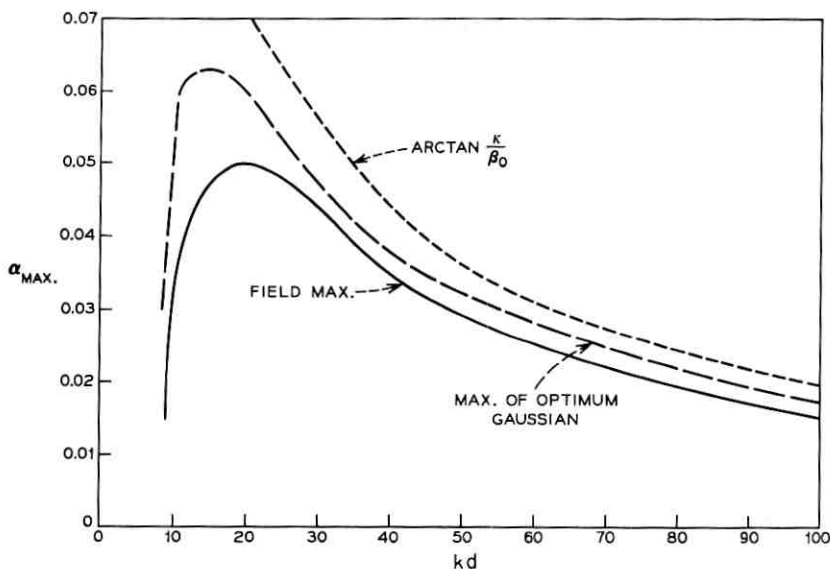
Fig. 5—Location of the peak of the radiation pattern. The solid curve indicates the maxima of the curves of Fig. 4; the dotted curve represents the direction of travel of the plane waves whose superposition comprise the guided mode; the dash-dotted curve indicates the location of the field maxima of the optimum gaussian distribution used to excite the waveguide.

with index n_2 can be decomposed into two plane waves traveling at certain slant angles given by the equation

$$\tan \alpha' = \frac{\kappa}{\beta_0}. \tag{26}$$

It might be expected that the radiation field can be explained by assuming that the two plane waves composing the guided mode simply detach themselves from the guidance structure and travel out into the space $z > 0$. The wave whose slope points up towards the air-dielectric interface is totally reflected and is forced to travel in the same direction as the other plane wave component. The dotted curve of Fig. 5 shows the angle α' of equation (26) as a function of kd. The dotted and solid curves do not coincide indicating that the picture of the plane waves simply detaching themselves to form the radiation pattern is not quite accurate. However, it is also apparent that this explanation provides a first rough idea of the direction into which the radiation is emitted.

Figure 6 is a plot of the phase angle

$$\phi = \arctan \frac{\delta}{\sigma} \tag{27}$$

appearing in equation (23) as a function α. This curve is independent of kd and holds for all radiation patterns shown in Fig. 4. It is apparent that the surface $r =$ const. is not exactly a surface of constant phase. Since the departure from constant phase is linear, Fig. 6 indicates that the radiation pattern does not originate at the coordinate origin but at a slightly shifted point. The shift of the far field source point is so slight however (only a few percent of the width of the guiding region) that its exact position is not important. For all practical purposes it can be assumed that the source point of the far field radiation pattern is located at the air-dielectric interface at the point where the guiding structure ends.

IV. MATCHING THE RADIATION PATTERN WITH A GAUSSIAN BEAM

We are now in a position to return to the primary purpose of our investigation. Knowing the radiation pattern produced by a guided mode we can also answer the question: what is the efficiency with which this mode can be excited? To do this we need to find what fraction of the radiation pattern appears in form of a gaussian beam. The efficiency with which the radiation field excites a given gaussian beam mode of free space is, by the reciprocity theorem, identical to the efficiency with which the radiation field and consequently the
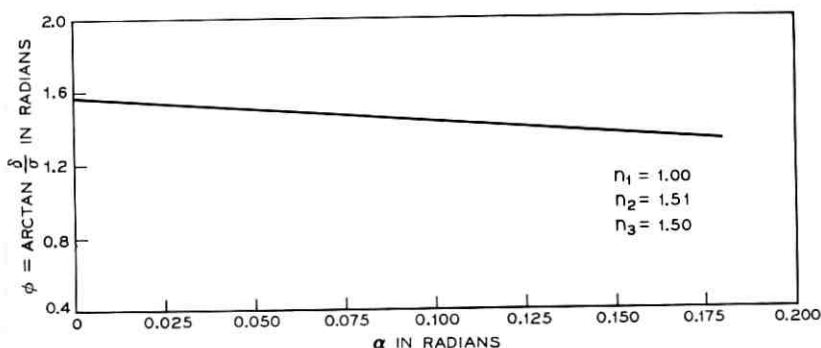
Fig. 6—The phase angle occurring in equation (23) as a function of α.

guided mode producing this radiation field can be excited by a gaussian beam mode.

The radiation field E_y can be expressed as a sum of hermite-gaussian modes $\mathcal{E}_n$

$$E_y = \sum_{n=0}^{\infty} c_n \mathcal{E}_n .$$ (28)

The only coefficient of interest to us is c_0, the excitation coefficient of the lowest order hermite-gaussian mode. Since the zero order hermite polynominal is the constant unity, this mode is purely gaussian. Using the gaussian mode

$$\mathcal{E}_0 = \left(\frac{2}{\pi}\right)^{\frac{1}{4}} \left(\frac{2\omega\mu P}{n_3 k r \theta_0}\right)^{\frac{1}{2}} \exp\left[-\left(\frac{\alpha - \alpha_0}{\theta_0}\right)\right]$$ (29)

with the angular beam half width θ_0 we can express the expansion coefficient c_0 as follows.

$$c_0 = \frac{n_3 k}{2\omega\mu P} \int_0^{\alpha_c} E_y \mathcal{E}_0^* r \, d\alpha$$

$$= \frac{2^{5/4} n_3^2 (n_2^2 - n_3^2) \kappa k^4}{\pi^{3/4} \left[\theta_0 \beta_0 \left(d + \frac{1}{\gamma} + \frac{1}{\theta}\right)(\kappa^2 + \gamma^2)(\kappa^2 + \theta^2)\right]^{1/2}}$$

$$\cdot \int_0^{\alpha_c} \exp\left[-\left(\frac{\alpha - \alpha_0}{\theta_0}\right)^2\right] \frac{\sin\alpha \cos\alpha}{(\sigma^2 + \delta^2)^{1/2}(\kappa^2 - \sigma^2)(\gamma^2 - \delta^2)}$$

$$\cdot \left\{(\gamma - \delta)(\kappa^2 + \theta^2)^{1/2} + (\kappa^2 + \gamma^2)^{1/2}\right.$$

$$\left.\cdot \left[(\theta + \delta)\cos\sigma d + \left(\frac{\theta\delta}{\sigma} - \sigma\right)\sin\sigma d\right]\right\} d\alpha.$$ (30)

The upper limit α_c of the integral corresponds to the angle at which δ of equation (24) vanishes. The transmission coefficient from the free space gaussian beam mode to the guided mode of the waveguide is now given by

$$T = |c_0|^2.$$ (31)

The integral in (30) was evaluated numerically. The results are shown in Fig. 7. The optimum angles for the maximum of the gaussian beam α_0 and the optimum half width θ_0 had to be determined by trial and error. The optimum values α_0 of the peak of the gaussian beam distribution are shown as the dash-dotted curve of Fig. 5. This curve
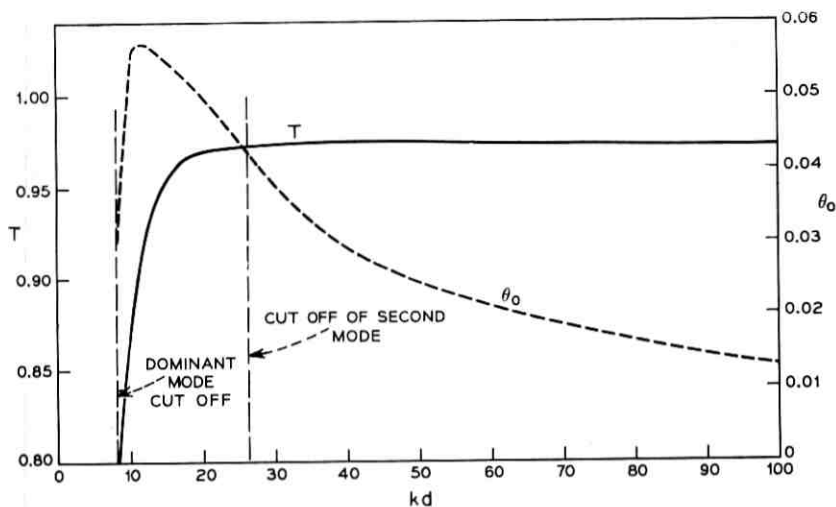
Fig. 7—Optimum transmission (excitation) coefficient T and optimum beam half width θ_0 of the exciting gaussian beam.

falls between the position of the maxima of the radiation pattern and the dotted line indicating the direction of the plane waves composing the guided mode. The optimum position of the gaussian beam mode approaches the dotted curve more closely. The idea of determining the required direction of the incident laser beam from the direction of the plane waves comprising the guided mode is thus more accurately reflected in the maximum position of the optimum gaussian beam.

The optimum half width θ_0 of the gaussian distribution is also shown in Fig. 7. The optimum position of the gaussian beam as well as its optimum width can not be obtained simply from the curves of Fig. 4 because of the asymmetry present in these curves.

The maximum achievable coupling efficiency is quite high. According to Fig. 7, 97 percent of the power can be converted from an optimally placed gaussian beam to the guided mode of the waveguide. We must remember, however, that our present case was oversimplified by assuming that the dielectric waveguide is infinitely extended in y direction. If the waveguide has only a finite width in y direction the gaussian beam must also be focused in that direction in order to match it to the guided mode. However, the matching problem in the y direction is not complicated by a dielectric interface. The radiation pattern that results from the finite beam width in y direction is perfectly

symmetric. Matching to symmetric radiation fields has been discussed in earlier work[5],[6] and was found to be highly efficient. (99 percent of the laser power can theoretically be converted to the dominant mode of a round optical fiber.) The loss in excitation efficiency caused by a finite width of the waveguide in y direction is expected to be slight. It is safe to assume that 90 percent conversion efficiency can be obtained for a more realistic waveguide with finite width in x as well as y direction.

V. CONCLUSION

We have seen that a dielectric optical waveguide constructed as shown in Fig. 2 can be excited by a gaussian laser beam with high efficiency. Based on the assumption that the waveguide is infinitely extended in y direction the maximum obtainable excitation efficiency is 97 percent. For a more realistic guide whose y dimension is also finite the conversion efficiency must be somewhat less. However, based on experience with mode excitation by gaussian beams in symmetrical waveguides we expect the excitation efficiency to be better than 90 percent.

The gaussian beam must be inserted into the waveguide from the side of the substrate. This requires shining the beam into the substrate material at a small angle with respect to its air-dielectric interface. A typical angle is of the order of 2°. It may be necessary to shape the substrate in order to facilitate the injection of the laser beam. The main advantage of this scheme over injection of the beam into a waveguide that reaches to the edge of the substrate consists in avoiding the imperfections in geometry that must be expected to exist at the edge of the substrate material. The waveguide extends only a few microns into the substrate material so that the edge of the substrate would have to be absolutely perfect to within a fraction of a micron.

Other waveguide geometries can be excited by a similar technique. It is possible to create a region of higher refractive index inside of the substrate material by ion implantation using high energy ion beams. A waveguide of this kind is shown schematically in Fig. 8. The waveguide may reach the outer face of the substrate material or it may end inside of the substrate. Excitation of this guide with a gaussian laser beam is possible in either case. This geometry is particularly advantageous since the radiation pattern of this type of waveguide is

n_1
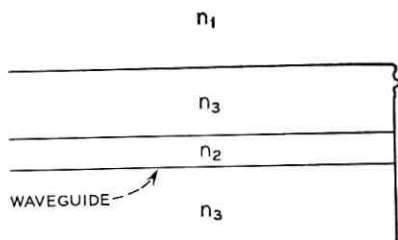
Fig. 8—Alternate dielectric waveguide configuration. The guiding region is located beneath the surface of the substrate.

not influenced by the presence of the air-dielectric interface. High excitation efficiency can be expected.

REFERENCES

1. Miller, S. E., "Integrated Optics: An Introduction," B.S.T.J., *48*, No. 7 (September 1969), pp. 2059–2069.
2. Shubert, R., and Harris, J. H., "Optical Surface Waves on Thin Films and Their Application to Integrated Data Processor," I.E.E.E. Trans. MTT, *16*, (December 1968), pp. 1048–1054.
3. Goell, J. E., and Standley, R. D., "Sputtered Glass Waveguide for Integrated Optical Circuits," B.S.T.J., *48*, No. 10 (December 1969), pp. 3445–3448.
4. Tien, P. K., Ulrich, R., and Martin, R. J., "Modes of Propagating Light Waves in Thin Deposited Semiconductor Films," Appl. Phys. Letters, *14*, (May 1, 1969), pp. 291–294.
5. Stern, J. R., Peace, M., and Dyott, R. B., "Launching into Optical-Fibre Waveguide," Elec. Letters, *6*, No. 6 (March 19, 1970), pp. 160–162.
6. Marcuse, D., "Excitation of the Dominant Mode of a Round Fiber by a Gaussian Beam," B.S.T.J., *49*, No. 8 (October 1970), pp. 1695–1703.

Rain Outage Performance of Tandem and Path Diversity 18-GHz Short Hop Radio Systems

By T. L. OSBORNE

The performance with respect to rain outage time of dual path diversity and non-path diversity (tandem) arrangements for 18-GHz short hop radio systems is computed and compared. The analysis is based on two extrapolations of R. A. Semplak's[1] three-year average of the measured probability distributions for rain attenuation at 18.5 GHz on a 6.4-km hop in New Jersey. The effects of merge hops and joint fading between hops in the diversity system, dependence of the rain attenuation distribution on hop length, and uncertainty in the tail of the distribution are included.

The results show that (i) the performance of tandem systems relative to diversity systems increases as the system length increases, (ii) the difference in the number of repeaters per unit length required for short and long tandem systems is small, (iii) the performance of the diversity system is strongly dependent on the amount of joint fading between parallel paths, and (iv) the performance of the tandem system is strongly dependent on the tail of the attenuation distribution. Neither of the latter two factors is known from rain attenuation measurements, but if the joint attenuation probabilities are sufficiently high, then diversity shows no advantage over tandem for either of the assumed extrapolations. The uncertainty in the tail of the attenuation distribution and the sensitivity of the tandem system performance to it emphasize the need for reliable attenuation measurements out to a probability of about 10^{-7}.

I. INTRODUCTION

The problems of utilizing the frequencies above 10 GHz for radio relay communication have been discussed by L. C. Tillotson[2]; the major technical problem is attenuation by rain. Based on results derived from a rain gauge network in Bedfordshire, England, which

showed that the joint distribution of rain rates for two laterally separated paths is significantly less than the distributions for the individual paths, D. C. Hogg[3] suggested the use of dual path diversity as a means for reducing the magnitude of the problem. However, in deciding the value of path diversity, many other factors must be considered such as variation of fade margin with hop length, merging of the diversity paths at switching points, the nature of the attenuation distributions, and the degree of independence of two laterally separated paths versus separation.

L. T. Gusler[4,5] has previously calculated the reliability and maximum lengths of diversity and nondiversity systems at 11 and 17 GHz based on attenuation distributions derived from six rain gauges on a five-mile path in New Jersey, and assumptions on joint distributions, merge paths, and repeater parameters. He concluded that diversity systems with short hops and moderate fade margins are suitable for long-haul systems, but that nondiversity systems are suitable only for short-haul or low-reliability systems. These results have since been used in studies of long-haul microwave pole line systems.[6]

This paper describes a new comparison of the rain outage performance of 18-GHz dual path diversity (diversity) systems and non-path diversity (tandem) systems including the effects of merge paths in the diversity system, joint fading between diversity paths, dependence of the rain attenuation distribution on hop length, and uncertainty in the tails of the rain attenuation distribution. Characteristics other than rain outage performance, such as path switching, delay equalization of the diversity paths, equipment redundancy, and interference may weigh for or against a particular system, but these factors are not considered here.

II. SUMMARY

The tandem and diversity systems are compared on the basis of equal numbers of identical repeaters in a length which is equal to the length of a diversity switching section; the section geometry is shown in Fig. 1. Lengths of tandem and diversity systems with 0.01 percent total rain outage time are calculated from the rain attenuation distribution as a function of tandem hop length with section length and the repeater fade margin as parameters. Diversity path separation is also a parameter in the diversity case and spacings of 10 and 20 km are used in the computations. System lengths for outage times other than 0.01 percent are inversely proportional to the outage time. To
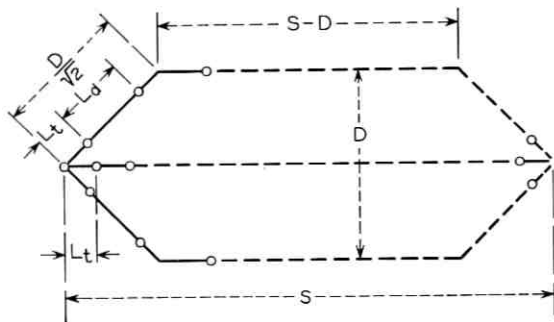
Fig. 1—Geometry of tandem and diversity systems.

approximate the reduced repeater spacing necessary at the merge points in the diversity system, the lengths of the first and last hops in each diversity path are set equal to the hop length in the comparison tandem system.

The rain outage performance of a tandem system depends on the probability distribution of rain attenuation which is a function of the hop length; diversity system performance depends on the joint probability distributions for laterally separated parallel hops. The measured 18-GHz attenuation distribution for a 6.4-km hop in Holmdel, New Jersey, reported by R. A. Semplak[1], is used as the basic distribution. To extrapolate and to include the effects of variations in the tail of the distribution, the measured distribution is approximated by two functions, A and B. The measured distribution for the 6.4-km hop and its approximations are shown in Fig. 2 as curves R1, A1, and B1. A hop length dependence is assumed which causes the approximation to fit the measured 30.9-GHz (1.9-km hop length) distribution reported by Semplak[1] when it is converted to an equivalent 18.5-GHz distribution. This derived 18.5-GHz distribution for a 1.9-km hop and its approximations are shown as curves R2, A2, and B2 in Fig. 2. Figure 3(a) shows the assumed dependence of attenuation on hop length for constant probability and Fig. 3(b) shows the family of distributions obtained from the approximating function B with hop length as a parameter.

Joint fading of the directly opposite and first diagonally opposite hops in the diversity system is included in the following way. When two hops are statistically independent, the distribution of the attenuation which is exceeded jointly on both hops is the product of the indi-
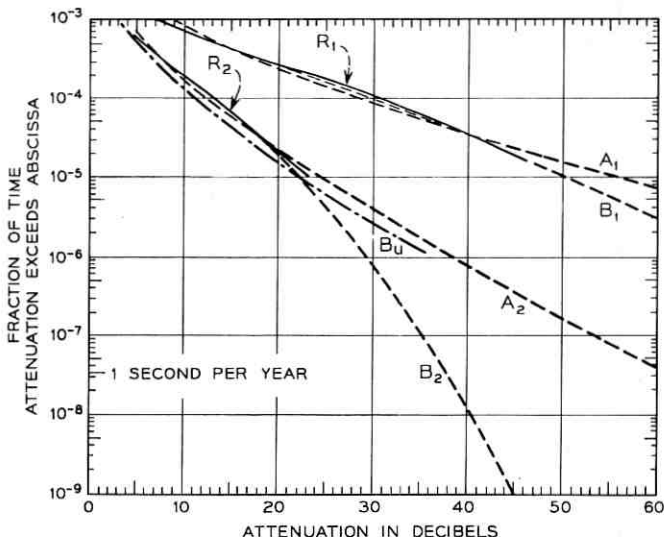
Fig. 2—Probability distributions for rain attenuation at 18 GHz. Curves are: R1, measured at 18.5 GHz on 6.4-km hop (Ref. 1); R2, 18 GHz distribution for 1.9-km hop derived from 30.9-GHz distribution (Ref. 1); A1 and A2, approximating function A for 6.4-km and 1.9-km hop lengths; B1 and B2, approximating function B for 6.4-km and 1.9-km hop lengths; Bu, 1.9-km distribution derived from one minute rain rate distribution (Ref. 13) using equation (1) with $k = 0.1$.

vidual distributions; when the two hops are completely dependent (conditional probabilities are unity) the joint distribution is equal to the individual distribution. Joint distributions lying between these two special cases are assumed to be given by the individual distribution raised to an exponent, c, which has values $c = 2$ in the independent case and $c = 1$ in the completely dependent case. Diversity system performance is calculated for distributions corresponding to c equal to 2.0, 1.8, 1.6, 1.4, and 1.0. These distributions are shown in Fig. 4 for the approximation B1 in Fig. 2. The results of the analysis can be applied, once the measured joint attenuation distribution is known, by finding the corresponding exponent, c, for the measured distribution.

All hops other than the directly opposite hops and the first diagonally opposite hops in the diversity systems are assumed to be statistically independent. However, joint fading of other pairs of hops further decreases the performance of the diversity system relative to the tandem system. Therefore the preceding assumptions establish minimum conditions under which the diversity system must have superior
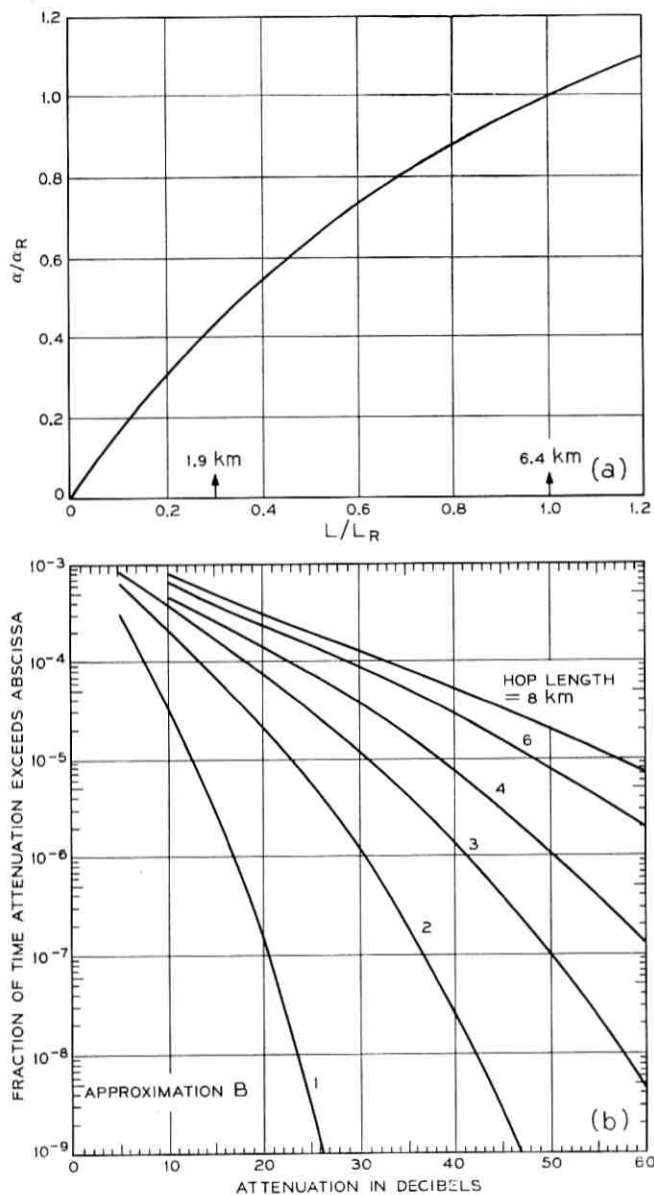
Fig. 3—(a) Attenuation as a function of hop length at constant probability using equation (2) with $\rho = 0.543$. (b) Probability distributions of rain attenuation with hop length as a parameter using approximating function B and equation (2) with $\rho = 0.543$.
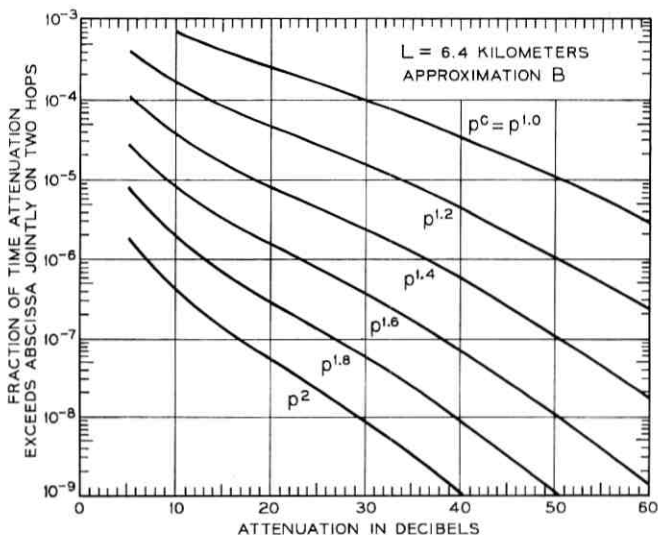
Fig. 4—Joint probability distributions for two paths with c as a parameter using approximating function B and a hop length of 6.4 km.

performance if it is to remain advantageous under more realistic conditions.

The results of this analysis are shown in Figs. 5 through 8. In Figs. 5, 6 and 8, the lengths of tandem and diversity systems with 0.01 percent total rain outage times are plotted as a function of tandem hop length with the joint fading exponent, c, as a parameter of the diversity system. Figure 7 shows the probability that the rain attenuation will exceed the hop fade margin for different hop lengths and fade margins based on approximation B. The results show in general that (*i*) the performance of tandem systems relative to diversity systems increases as the system length increases, making tandem systems more suitable for long systems; (*ii*) the difference in the number of repeaters per unit length required for short and long tandem systems is small; (*iii*) the performance of the diversity system depends strongly on the degree of independence of the two parallel paths in the diversity system; and (*iv*) the performance of the tandem system depends strongly on the tail of the attenuation distribution. Neither of the latter two factors has been measured for rain attenua-

Fig. 5—Lengths of 0.01 percent outage time tandem and diversity systems based on approximating function B with a diversity path separation of 10 km, a section length of 48 km, and a repeater fade margin of (a) 40 dB, and (b) 50 dB.
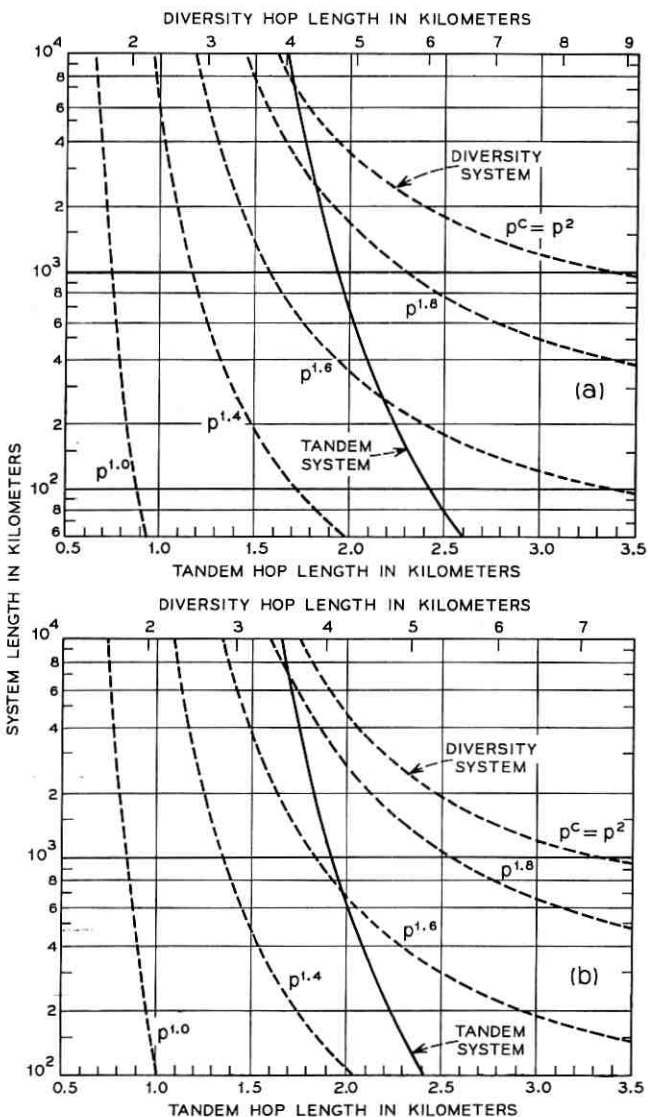
Fig. 6—Lengths of 0.01 percent outage time tandem and diversity systems based on approximating function B with a repeater fade margin of 40 dB, and (a) a 20-km diversity path separation and a 48-km section length, and (b) a 10-km diversity path separation and a 96-km section length.
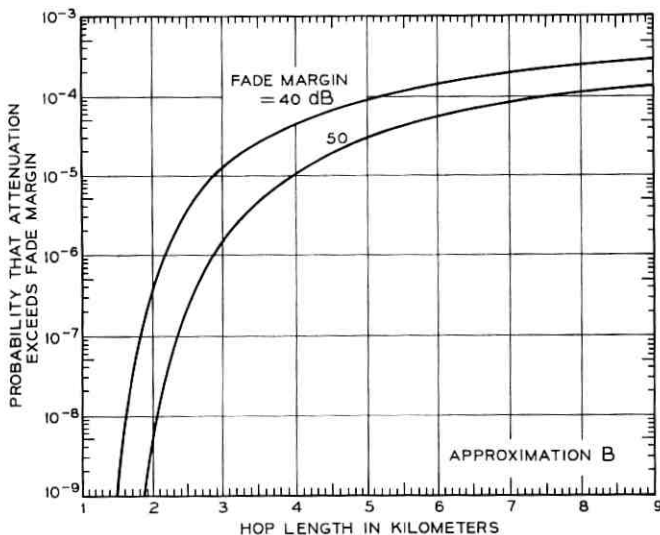
Fig. 7—Probability that the rain attenuation exceeds the hop fade margin as a function of hop length with standard repeater fade margin as a parameter using approximating function B.

tion.* The uncertainty in the tail of the distribution and the sensitivity of the tandem system performance to it results in about a 50 percent uncertainty in the number of repeaters required and emphasizes the need for reliable measurements of the rain attenuation distribution out to a probability of about 10^{-7}. At present, Semplak's measured distribution extends only to about 2×10^{-5}.

III. INTERPRETING THE RESULTS—EXAMPLES

Figure 5(a) shows the system lengths obtained with repeaters having a 40-dB fade margin on a 1-km hop, and with a section length of 48 km and path separation of 10 km. For a tandem hop length of 2 km, the tandem system curve shows that the length of a tandem system with a total rain outage time of 0.01 percent is 650 km. The $c = 2$ diversity curve shows that the length of a diversity system

---

* The joint *rain rate* distribution reported by Mrs. A. E. Freeny and J. D. Gabbe[7] for a hop length of 5.2 km and path separation of 10.4 km corresponds to $c \approx 1.4$. If it is assumed that the joint attenuation distributions have the same degree of independence, then the performance of the diversity system is indicated by the $c = 1.4$ curves in Figs. 5 through 8 and the diversity system is inferior or comparable to the tandem system depending on the tail of the attenuation distribution.
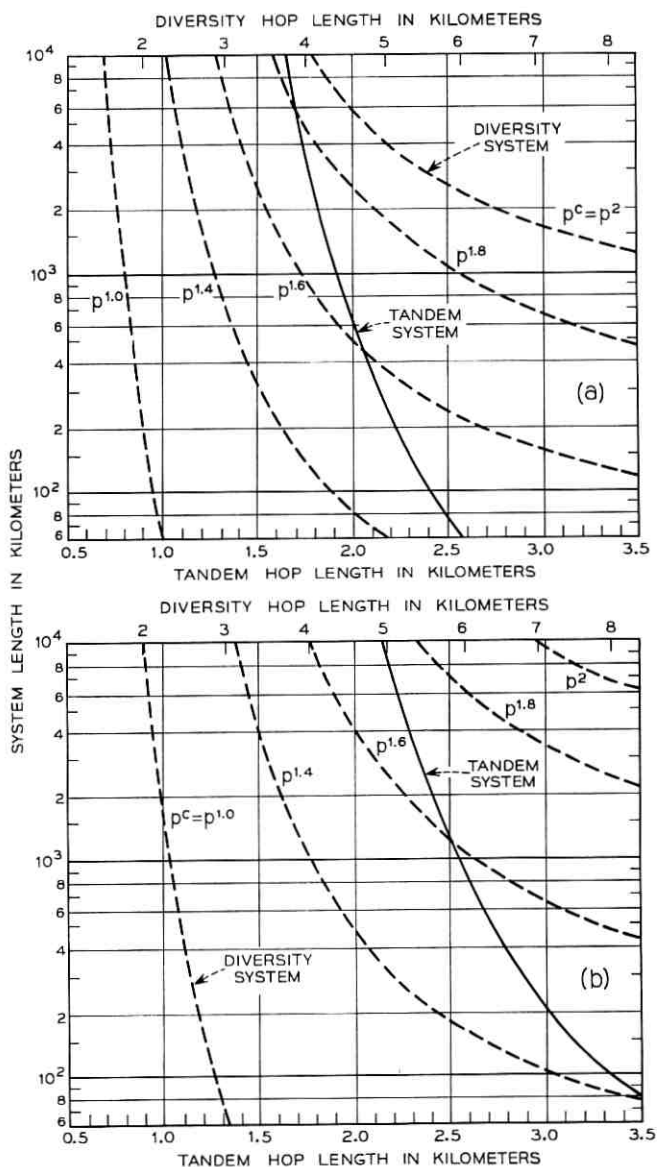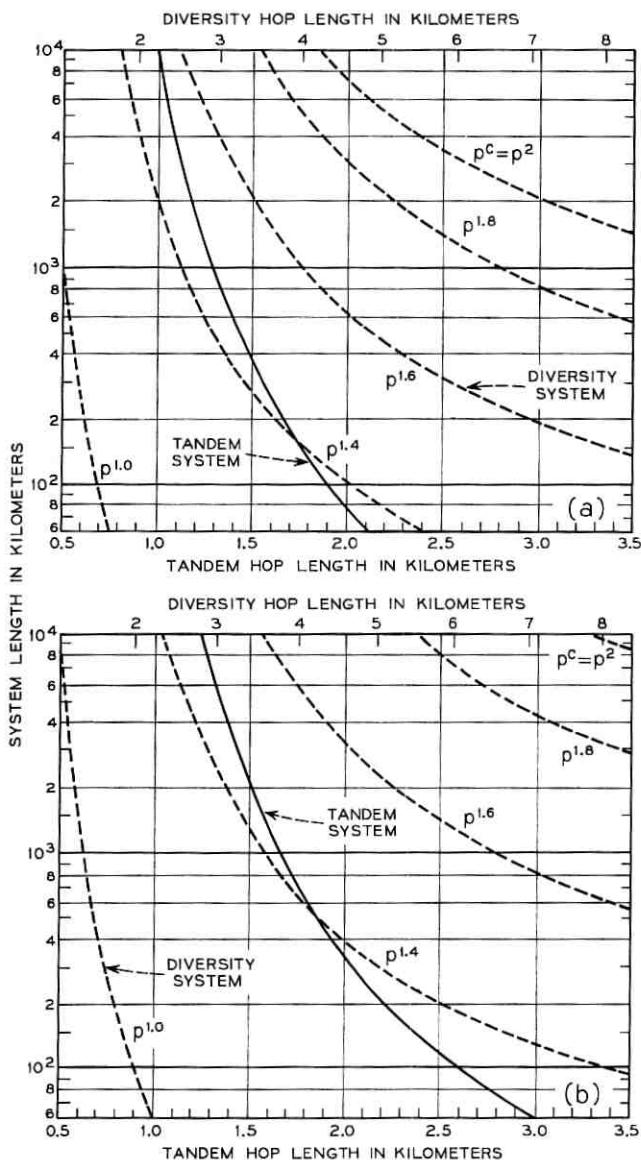
Fig. 8—Lengths of 0.01 percent outage time tandem and diversity systems based on approximating function A with a diversity path separation of 10 km, a section length of 48 km, and a repeater fade margin of (a) 40 dB, and (b) 50 dB.

with the same number of repeaters per units length and the same total rain outage time is 5600 km if all hops have statistically independent fading. However, if there is joint fading of opposite hops corresponding to $c = 1.8$, 1.6, and 1.4, the length of the diversity system decreases to 2450, 510, and 82 km respectively. For joint fading corresponding to $c = 1.6$ the diversity system is better than the tandem system only for system lengths less than 460 km and for $c = 1.4$ the tandem system is better for all lengths.

Figure 5(a) can be interpreted in another way. Suppose a 5000-km system with 0.01 percent rain outage time is required. The tandem system requires a hop length of 1.73 km. Assuming joint fading corresponding to $c = 1.6$, the diversity system requires an equivalent tandem hop length of 1.37 km which is equivalent to a 26 percent increase in the number of repeaters per unit length.

The tandem curve shows another important result. A 500-km short-haul tandem system with 0.01 percent outage time requires a 2.04-km hop length whereas a 5000-km long-haul tandem system with the same outage time requires a 1.73-km hop length. This means that a 10-fold increase in system length requires only an 18 percent increase in the number of repeaters per unit length.

Figures 5(b), 6(a) and 6(b) show the same sets of curves but with different values of the three parameters, repeater fade margin, diversity path separation, and section length. Figure 5 shows that the relative diversity system performance increases with larger fade margins, but the improvement is not appreciable. Figure 6(a) in comparison with Fig. 5(a) can be used to estimate the improvement in joint fading required to offset the larger number of repeaters caused by a wider path separation. Figure 6(b) shows that the relative diversity performance is slightly better for longer section lengths.

The sensitivity of these results to the tail of the attenuation distribution can be seen by comparing Figs. 5 and 8. Figure 8 shows system lengths calculated from approximation A in Fig. 2; otherwise all parameters are the same as in Fig. 5. Because the tandem hops require much smaller fading probabilities than do diversity hops, tandem system performance is very sensitive to the tail of the distribution whereas diversity system performance is fairly insensitive. For example, from Fig. 5(a), using approximation B, a 1000-km tandem system requires a hop length of 1.93 km whereas, from Fig. 8(a) using approximation A, a 1000-km tandem system requires a hop length of 1.29 km, equivalent to a 50 percent increase in the number

of repeaters per unit length. For $c = 1.6$ diversity systems there is no significant difference in hop lengths, reflecting the equality of the two approximations in the region where measured statistics are available.

IV. PROBABILITY DISTRIBUTIONS FOR RAIN ATTENUATION

The rain outage performance of tandem and diversity radio systems depends entirely on the probability distribution of rain induced attenuation on a path of given length and on the joint distributions for multiple paths of different lengths and separations. Such distributions can be obtained in two ways: by direct measurement of attenuation data or by calculation from measured rain rate data. The latter method is much more difficult and unreliable, but one is irresistibly tempted to use it because of the large amount of potentially useful rainfall data available. Consequently, much effort has been spent trying to establish a usable relation between rain rate and attenuation, and, anticipating a usable relation, measurements have been made of the spatial and temporal characteristics of rainfall; but few statistics are available for directly measured rain attenuation. In order to establish the basis for the attenuation distributions used here, these measurements will be discussed briefly.

In 1965, R. G. Medhurst[8] summarized previous measurements and compared measured attenuation with the attenuation computed from measured rain rates using the Mie theory; he found the relationship between them quantitatively inconsistent. Semplak and R. H. Turrin[9] have reported attenuation measurements at 18 GHz on a 6.4-km path in New Jersey and, for data accumulated in 1967, have compared these measurements with attenuations calculated from measured rain rates along the path on a single event basis and also by comparing the annual distribution of measured attenuation with the annual attenuation distribution computed from the annual rain rate distribution. On the single event basis the results were similar to Medhurst's findings. The comparison of the probability distributions showed reasonable agreement if the attenuation was computed from the rain rate by the formula

$$\alpha = k\bar{R}d \qquad (1)$$

where k was 0.059, $\bar{R}$ the average rain rate, and d the hop length. However, Semplak[1] has also reported a similar comparison of the distributions based on 1968 data in which case agreement was obtained

with $k = 0.09$. A value of $k = 0.1$ has also been used.[10] These variations translate directly into at least a 50 percent uncertainty in the number of repeaters required if rain rate statistics are used for system calculations.

Freeny and Gabbe[7] have published a statistical summary of rain rate measurements made in a six-month period in 1967 on the Crawford Hill, New Jersey, rain gauge network. This work was aimed at obtaining a statistical description of the behavior of point rain rates, the relationship of two rain rates separated in space, and the relationship of average rain rates on pairs of paths in various configurations. However, although believing the relative values to be fairly good, the authors do not believe the numerical values should be considered better than an order of magnitude.

Ruthroff[11] has suggested that the discrepancy between measured attenuations and those calculated from rain rate data may lie in the measurement of the rain rates. He has also shown that because of the integrating effect of the radio path, the time over which the rain rate is integrated plays an important part in obtaining path attenuation from rain rates along the path.[12]

Thus, for the present, it appears that neither the statistics of high rain rates nor the quantitative relation between rain rate and attenuation are known accurately enough for rain rate data to be used as a basis for system design. Our recourse is to use the distributions obtained from directly measured attenuations. The problem in this case is that the amount of data is inadequate, in both the long time statistical sense and in the range of values of the parameters involved. Semplak[1] has reported a three-year average annual attenuation distribution for the 18-GHz attenuation on a 6.4-km path in New Jersey, and a two-year average annual distribution for the 30.9-GHz attenuation on a 1.9-km path at the same location, but no distributions for different path lengths or joint distributions for spatially separated paths are available.

With the above factors in mind, the following approach is used here to obtain the required distributions of attenuation for individual and spatially separated paths. The three-year average annual distribution for measured rain attenuation at 18.5 GHz on the 6.4-km path as reported by Semplak[1] is taken as the basic distribution. This distribution is approximated by two functions, called A and B, which represent reasonably high and low values of the tail of the distribution. Semplak's measured 6.4-km distribution and the approximating functions

A and B are shown in Fig. 2 as curves R1, A1, and B1 respectively. Two functions are used in order to show the effects of variations in the tail of the distribution since the measured distribution does not extend to sufficiently low probabilities.

In order to introduce a hop length dependence for the distributions, the functional dependence on hop length

$$\frac{\alpha}{\alpha_R} = \frac{L/L_R}{\rho + (1 - \rho)L/L_R} \tag{2}$$

is assumed for the attenuation at any given probability level, where α is the attenuation for a hop length L, α_R is the reference attenuation at the reference length L_R which are taken from the measured 6.4-km distribution, and ρ is a suitably chosen constant. To evaluate ρ, an attenuation distribution for a shorter hop length is needed; for this purpose the two-year average annual distribution for measured rain attenuation at 30.9 GHz on the 1.9-km hop is converted to an 18.5-GHz distribution by dividing the attenuation at any given probability level by 2.0.* For $\alpha_R = 40$ dB and $L_R = 6.4$ km from the 6.4-km distribution, and $\alpha = 17.5$ dB and $L = 1.9$ km from the derived 1.9-km distribution, both at a probability level of 3.45×10^{-5}, the value $\rho = 0.543$ is obtained. Equation (2) with $\rho = 0.543$ is plotted in Fig. 3. This variation of α with L is approximately linear for small values of L but decreases for larger L as is intuitively expected. The curve is similar in shape but somewhat more nonlinear than that obtained by Hogg from rain rate data.[10] The derived 1.9-km distribution and the approximating functions A and B for a 1.9-km hop length are shown in Fig. 2 as curves R2, A2, and B2 respectively. For comparison, the distribution calculated from H. E. Bussey's[13] one-minute Washington, D. C., rain rate distribution, using equation (1) with $k = 0.1$, is also shown in Fig. 2 for a hop length of 1.9 km.

For calculating the outage time of a diversity system, joint probability distributions for two parallel laterally separated hops of equal length are needed for different hop lengths and path separations. Since no such directly measured attenuation distributions are available these distributions must be assumed. Fortunately, because of the assumed symmetry of the diversity system, the distributions for two special cases can be derived from the individual distributions. Let event a_i be the fading of the ith diversity hop and event b_i be the

---

* This method of obtaining a 1.9-km distribution was first suggested to me by Gusler. The factor of 2.0 has been used previously, but should be viewed with caution since it is derived basically from equation (1).

fading of a hop in the opposite path of the diversity system; the joint probability of both events is $P(a_i, b_i)$. Since the hops are identically distributed, i.e., $P(a_i) = P(b_i) \equiv p$, and only the cases when the variables take the same values are considered, in the special cases of complete dependence $[P(a_i \mid b_i) = P(b_i \mid a_i) = 1]$ and statistical independence the joint probabilities are $P(a_i, b_i) = p$ and $P(a_i, b_i) = p^2$ respectively. Joint probabilities between and including these two cases are represented by

$$P(a_i, b_i) \equiv p^c \qquad 1 \leqq c \leqq 2 \qquad (3)$$

where $c = 1$ gives the case of complete dependence and $c = 2$ gives the case of statistical independence. The joint probability distributions generated in this way for 6.4-km hop lengths using approximating function B are shown in Fig. 4 for different values of c. It would be possible to assume some functional dependence of c on diversity path separation and then compute diversity performance for different path separations; the approach used here, however, is to compute performance for a range of values for c at two practical values of path separation. The diversity system performance can then be deduced once the relative independence of two hops is measured.

V. SYSTEM PERFORMANCE

5.1 *Assumptions*

(i) The same standard repeater is used in both tandem and diversity systems, i.e., transmitter power, antenna gain, and receiver noise figure are the same. The repeater is characterized by its fade margin on a one-km hop.

(ii) The number of repeaters in one section is the same in both tandem and diversity systems, or equivalently, the number of repeaters per unit length is the same in both systems.

(iii) The end hops in each diversity path are the same length as the hops in the comparison tandem path, but in calculating joint fading probabilities involving these hops they are considered the same as the other diversity hops. In other words, we assume that the end hops are shorter but that their closer spacing just compensates causing their joint fading probability to be the same as the other diversity pairs.

(iv) The rain attenuation probability distribution is given by one of the following approximations to Semplak's measured distribution at 18.5 GHz on a 6.4-km hop.

Approximation A:

$$p \equiv p(\alpha \geq \alpha_R) = \frac{.0134}{\alpha_R} \epsilon^{-0.0567\alpha_R} . \tag{4}$$

Approximation B:

$$p \equiv p(\alpha \geq \alpha_R) = \frac{.00719}{\alpha_R} \epsilon^{-0.00103\alpha_R^2} , \tag{5}$$

where α_R is the total rain attenuation. These approximations are shown as curves A1 and B1 respectively in Fig. 2.

(v) The attenuation at a constant probability level is related to the hop length according to equation (2) with $\rho = 0.543$, $L_R = 6.4$ km, and α_R equal to the attenuation on the 6.4-km hop. If the attenuation α on a hop with length L is known, the probability that that attenuation is exceeded is found by finding the reference hop attenuation α_R for the same probability from equation (2) then substituting α_R into equations (4) or (5).

(vi) All hops have statistically independent fading except the directly opposite hops and the first diagonally opposite hops in the two diversity paths; the joint distribution for these pairs of hops is given by equation (3). As shown in the Appendix, the probability of outage for the diversity system is a series of sums and differences of the joint fading probabilities of combinations of two or more hops. In probabilities involving three hops, directly opposite and first diagonally opposite hops are paired and the pair is assumed statistically independent of the third hop. Joint probabilities involving combinations of four or more hops are assumed to be zero with a maximum estimated error of less than 1.4 percent when $c = 1$.

5.2 Tandem System

Let L_t be the tandem repeater spacing and R be the number of tandem repeaters in a section with length S as shown in Fig. 1. The number of hops in the tandem system, N_t, is equal to the number of repeaters, i.e., $N_t = R = S/L_t$. The fade margin of each of the tandem hops is

$$\alpha_t = \alpha_s - 20 \log L_t \tag{6}$$

where α_s is the fade margin of the standard repeater on a 1-km hop.

The fractional rain outage time of the tandem section is

$$P_t = 1 - (1 - p_t)^{N_t} \approx N_t p_t \tag{7}$$

where p_t is found by substituting α_t and L_t into equation (2) and the resulting α_R into equations (4) or (5). The outage time of the section determines the length, L_t, of a system with a specified total outage time, T,

$$L_t = \frac{ST}{P_t}. \tag{8}$$

The system lengths of tandem systems with 0.01 percent total rain outage time are plotted in Figs. 5 through 8.

5.3 *Diversity System*

Referring to Fig. 1, in each diversity path the number of diversity hops, N_d, with the diversity hop length, L_d, is

$$N_d = \frac{R - 3}{2} = \frac{S/L_t - 3}{2} \tag{9}$$

and

$$L_d = \frac{S + (\sqrt{2} - 1)D - 2L_t}{N_d} \tag{10}$$

where R is the total number of diversity repeaters, and D is the diversity path separation.

We have assumed that the probability of joint fading of the merge hops is the same as the probability of joint fading on the other diversity hops by virtue of the combination of a shorter length and reduced spacing. Therefore in calculating the probability of the diversity system not working, the effective number of diversity hops in a diversity path is

$$N_{d_e} = N_d + 2 = \frac{S/L_t + 1}{2}. \tag{11}$$

The fade margin for the diversity hops is

$$\alpha_d = \alpha_s - 20 \log L_d . \tag{12}$$

The fractional rain outage time of the diversity section is (see Appendix)

$$
\begin{aligned}
P_d = {}& [(3N_{d_e} - 2)p_d^{(c-2)} + (N_{d_e}^2 - 3N_{d_e} + 2)]p_d^2 \\
& - [(6N_{d_e}^2 - 16N_{d_e} + 12)p_d^{(c-2)} + (N_{d_e}^3 - 7N_{d_e}^2 + 16N_{d_e} - 12)]p_d^3
\end{aligned}
\tag{13}
$$

where p_d is found by substituting α_d and L_d into equation (2) and the resulting α_R into equations (4) or (5), and where c is the joint fading

exponent defined in Section III. The system length is calculated from equation (8).

Figures 5 through 8 show the lengths of diversity systems with a 0.01 percent total rain outage time plotted as a function of the hop length of the tandem system which has the same number of repeaters per unit length. Curves are shown for five degrees of joint fading between directly opposite and first diagonally opposite hops. The most severe case of joint fading is $c = 1$, the value corresponding to complete dependence of the two paths.

5.4 *Effects of Other Combinations of Joint Fading*

There are two other types of joint fading which affect the relative merit of tandem and diversity systems. As a result of assuming independence in these cases, the diversity-tandem comparison previously given is more favorable to diversity than is actually the case, perhaps substantially so; and secondly the actual outage time calculations for the tandem system are conservative, predicting more outage time than is actually the case.

5.4.1 *Joint Fading of Series Hops*

Joint fading of the hops connected in series occurs on both tandem and diversity paths. In either case, joint fading of series hops in a path decreases the total outage time for the path because more of the individual hop fades occur at the same time. Therefore the assumption of independence of the series hops gives larger outage times than would be calculated with joint fading. However, since the diversity hop lengths are always at least twice as long as the tandem hop lengths, the average distance between series diversity hops is always larger than between series tandem hops causing less joint fading of the series diversity hops. Therefore the assumption of independence of series hops increases the calculated outage times more for tandem than for diversity.

5.4.2 *Joint Fading of Diagonally Opposite Hops in the Diversity Paths*

In a diversity system with 4-km hops spaced 10 km apart, the distance from the center of one hop to the center of the first diagonally opposite hop is 10.4 km; for the next diagonally opposite hops the distances are 10.8 km, 11.2 km, 11.7 km, etc. There is little change in distance for the first few diagonally opposite hops. Therefore the joint fading for these diagonally opposite hops is expected to be comparable in magnitude to the joint fading of the directly opposite hops. Joint

fading of diagonally opposite hops increases the diversity section outage time in the same way as joint fading of the directly opposite hops. Since joint fading of diagonally opposite hops increases the diversity system outage time, and since the joint fading for at least the near diagonally opposite hops is expected to be comparable to the joint fading of directly opposite hops, the results calculated here on the basis of independence are optimistic toward the diversity system.

VI. ACKNOWLEDGMENT

APPENDIX

Assume two parallel laterally separated diversity paths with N hops in each path. Denote the fading of the ith hop in one path by event a_i and the fading of the ith hop in the other path by event b_i. Denote the outage of the path containing hop a_i by event a, and the outage of the other path by event b. The probability of a diversity section outage is then

$$P_{nw} = 1 - P(\bar{a} \cup \bar{b}) = 1 - P(\bar{a}) - P(\bar{b}) + P(\bar{a} \cap \bar{b}) \qquad (14)$$

where the bar denotes the complementary event. Using the DeMorgan laws, the $P(\bar{a})$ and $P(\bar{b})$ terms can be evaluated in terms of the individual hop probabilities, e.g.;

$$P(\bar{a}) = P(\bar{a}_1 \cap \bar{a}_2 \cap \bar{a}_3 \cdots) = 1 - P(a_1 \cup a_2 \cup a_3 \cdots)$$

$$= 1 - [P(a_1) + P(a_2) + P(a_3) + \cdots - P(a_1 a_2) - P(a_1 a_3)$$

$$- P(a_2 a_3) - \cdots + P(a_1 a_2 a_3) + \cdots]. \qquad (15)$$

Similarly, for $P(\bar{a} \cap \bar{b})$ we get

$$P(\bar{a} \cap \bar{b}) = 1 - P(a_1 \cup a_2 \cup a_3 \cdots \cup b_1 \cup b_2 \cup b_3 \cdots). \qquad (16)$$

Expanding equation (16) in a series of joint probabilities of individual events and substituting in equation (14) gives P_{nw} in terms of a series of joint fading probabilities of combinations of two to $2N$ hops excluding joint probabilities involving only hops in the same path. For example, if the $P(\cdot)$ notation is omitted and parentheses are used to indicate pairing of directly opposite and first diagonally opposite hops, P_{nw} has the form

$$P_{nw} = (a_1 b_1) + (a_1 b_2) + a_1 b_3 + \cdots + (b_1 a_2) + b_1 a_3 + \cdots$$
$$- (a_1 b_1) a_2 - (a_1 b_1) b_2 - \cdots - a_1 (a_2 b_2)$$
$$- a_1 a_2 b_3 - \cdots + (a_1 b_1)(a_2 b_2) + (a_1 b_1) a_2 a_3 + \cdots$$
$$+ a_1 (a_2 b_2) a_3 + \cdots + a_1 a_2 a_3 b_5 + \cdots . \tag{17}$$

From equation (3), the paired terms represent probabilities of the form p^c whereas unpaired individual events have probability p; since they are assumed to be statistically independent P_{nw} has the form

$$P_{nw} = B_{21} p^2 p^{(c-2)} + B_{20} p^2 - B_{31} p^3 p^{(c-2)} - B_{30} p^3 + B_{42} p^4 p^{2(c-2)}$$
$$+ B_{41} p^4 p^{(c-2)} + B_{40} p^4 - B_{52} p^5 p^{2(c-2)} - B_{51} p^5 p^{(c-2)} - B_{50} p^5 + \cdots ,$$
$$\tag{18}$$

where the $B_{\ell m}$ coefficients are the number of terms in P_{nw} with the form $p^\ell p^{m(c-2)}$. The $B_{\ell m}$ coefficients are a function of the number of hops N in each diversity path and an exact evaluation of P_{nw} from equation (18) would require evaluating each coefficient for every value of N. Although the total number of terms having the same value of ℓ is easily found, the distribution between coefficients with different m is not. Fortunately the higher order terms get small very rapidly for small p and it can be shown that a conservative estimate of the error involved in assuming that terms of fourth order and higher are zero is less than 1.4 percent if $p \leq 3 \times 10^{-4}$ and $c = 1.0$, the worst cases used here. For the type of pairing assumed here, B_{20}, B_{21}, B_{30}, and B_{31} have been evaluated giving, from equation (18)

$$P_{nw} = [(3N - 2)p^{(c-2)} + (N^2 - 3N + 2)]p^2$$
$$- [(6N^2 - 16N + 12)p^{(c-2)} + (N^3 - 7N^2 + 16N - 12)]p^3 \tag{19}$$

which is the same as equation (13).

REFERENCES

1. Semplak, R. A., "The Influence of Heavy Rainfall on Attenuation," IEEE Trans. Antennas and Propagation, *AP-18*, No. 4 (July 1970), pp. 507–511.
2. Tillotson, L. C., "Use of Frequencies Above 10 GHz for Common Carrier Applications," B.S.T.J., *48*, No. 6 (July–August 1969), pp. 1563–1576.
3. Hogg, D. C., "Path Diversity in Propagation of Millimeter Waves Through Rain," IEEE Trans. Antennas and Propagation, *AP-15*, No. 3 (May 1967), pp. 410–415.
4. Gusler, L. T., "Use of Route Diversity to Avoid the Effects of Rain Attenuation at 11 and 17 GHz," unpublished work.

5. Gusler, L. T., "A Solution for the Rain Attenuation Problem at 11 and 17 GHz," IEEE International Conference on Communications, Philadelphia, Pa., (June, 1968), Abstracts of Informal Papers.
6. Hathaway, S. D., "Possible Terrestrial Common Carrier Applications Above 10 GHz," IEEE International Convention Digest, (March, 1968), p. 66.
7. Freeny, Mrs. A. E., and Gabbe, J. D., "A Statistical Description of Intense Rainfall," B.S.T.J., *48*, No. 6 (July–August 1969), pp. 1789–1852.
8. Medhurst, R. G., "Rainfall Attenuation of Centimeter Waves: Comparison of Theory and Measurement," IEEE Trans. Antennas and Propagation, *AP-13*, No. 4 (July 1965), pp. 550–564.
9. Semplak, R. A., and Turrin, R. H., "Some Measurements of Attenuation by Rainfall at 18.5 GHz," B.S.T.J., *48*, No. 6 (July–August 1969), pp. 1767–1787.
10. Hogg, D. C., "Statistics on Attenuation of Microwaves by Intense Rain," B.S.T.J., *48*, No. 9 (November 1969), pp. 2949–2962.
11. Ruthroff, C. L., "Microwave Attenuation and Rain Gauge Measurements," Proc. IEEE, *57*, No. 6 (June 1969), p. 1235.
12. Ruthroff, C. L., "Rain Attenuation and Radio Path Design," B.S.T.J., *49*, No. 1 (January 1970), pp. 121–135.
13. Bussey, H. E., "Microwave Attenuation Statistics Estimated from Rainfall and Water Vapor Statistics," Proc. I.R.E., *38*, No. 7 (July 1950), pp. 781–785.

Model for Computation of Interference to Radio-Relay Systems From Geostationary Satellites

By A. S. MAY and M. J. PAGONES

A statistical model is suggested for the computation of interference into terrestrial radio-relay systems from geostationary satellites.

The model is general. It incorporates radio-relay characteristics, satellite arrangements, and allowable satellite power flux densities. A program simulator can be used to calculate the satellite power flux density corresponding to a particular radio-relay interference objective.

Interference distributions are computed for AT&T and CCIR radio-relay models using the power flux density that was suggested for study at the 1969 CCIR Interim Meeting at Geneva, Switzerland.

I. INTRODUCTION

Terrestrial radio-relay systems operating in the 4-GHz and 6-GHz bands share these frequencies with communication satellite systems, and additional bands may be designated for shared usage. In the future, as the geostationary orbit occupancy increases, coexistence of these systems will be possible only if controls are imposed on both systems.

Presently, the International Radio Consultative Committee (CCIR) recommends[1] that radio-relay antennas, radiating frequencies normally received by satellites, maintain a specified angular separation with respect to the geostationary orbit or, where this is not practical, the application of power limitation to the terrestrial transmitters. They further recommend[2] that satellite power flux density be limited as a function of the angle of illumination at the earth's surface. The pointing restriction is intended to protect satellites from terrestrial systems. Methods have been devised[3] for calculating beam-orbit separation as a function of azimuth displacement of the antenna from intercept, and

81

the recommendation is being complied with by the Bell System. The satellite power limitation has been challenged as being too restrictive and, in the interest of satellite system economy, an increase at the higher angles of arrival has been proposed.

The derivation of the present satellite power flux-density limitation assumes one radio-relay station in a 50-hop system will have a direct exposure to a geostationary satellite, and the remaining stations will have sufficient antenna discrimination that the additional interference is not significant. On this premise the flux-density limit for a tangential ray was established. A linear escalation for higher angles of arrival accounts only for a differential in transmission loss and satellite antenna gain between the horizon and the subsatellite point. A proposal to greatly increase the flux-density limit at all angles other than zero degrees requires a new evaluation.

At present the number of communication satellites is small and the effective radiated power is below existing limits. Consequently, an experiment with existing satellites interfering into the terrestrial network will not generate fruitful results in determining adequate protection for both systems. A laboratory experiment is more tractable but it is very difficult to simulate actual cases because the interference is a function of the spatial arrangement of the two systems, and the results should be applicable internationally. To place an experimental satellite in the geostationary orbit for the purpose of measuring interference into terrestrial systems is ludicrous for it would yield no more information than a laboratory experiment and would be vulnerable to the same limitations.

A new attempt to evaluate the satellite power flux density limitation originated in the U.S.A. in preparation for the 1969 CCIR meeting in Geneva, Switzerland. It was concluded that an analytical approach is the most promising.

In subsequent sections we present a description of the system model developed by the authors for studying satellite interference into terrestrial systems, along with the analysis, computer simulation, and some results.

II. SYSTEM MODEL

The satellite system is assumed deterministic with all satellites in the geostationary orbit. The spacing between satellites is assumed fixed with each satellite transmitting the same effective radiated power. Moreover, it is assumed the entire orbit is filled.

The assumption that all satellites are in the geostationary orbit is realistic because the number of medium-orbit systems is small, and the additional interference from medium orbit satellites can be accounted for by decreasing the interference allocation of the stationary satellites. Of course, the short-term interference contribution from medium orbit satellites should not be ignored, but this is not the subject of this discussion. Furthermore, it is assumed that if a satellite is visible to a radio-relay station, it will illuminate the area of the radio-relay station with the flux density permitted by the angle of arrival. Actually, this is not always the case. A satellite may be visible but it may use a highly directive antenna that does not illuminate all radio-relay stations with the maximum permissible flux density. Again, the effect of this assumption is not serious because the spot-beam satellites would reduce the interference and their effect can be adequately taken into account by increasing the effective satellite spacing.

The radio-relay system is assumed to be composed of a number of hops. The number of hops depends on the system—50 hops for a CCIR system and 140 hops for an AT&T system. Figure 1 introduces the system model concept.

The azimuth of each radio-relay system, referred to as a trendline, is assumed as a random variable. The distribution that was used for the results given in Section V is uniform between 0 and 2π. However,
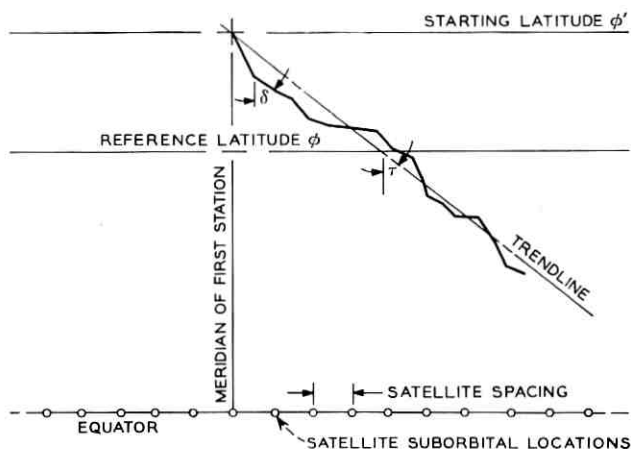

Fig. 1—System model showing one trendline.

the simulation is capable of using any distribution. For the AT&T terrestrial network, a uniform distribution is reasonable; however, for other countries the distributions will be biased (Canada's network is nearly east-west). Since the direction of each station is not the same as the trendline (intra-system interference and geography dictate a zig-zag route), the direction of each station is assumed as a random variable within some range about the direction of the trendline. Again the results in Section V were calculated for a uniform distribution within ±25° of the trendline direction.

The radio-relay antenna pattern can be any appropriate pattern depending on the actual system. However, the angular separation between the beam of the radio-relay station and the satellite does not correspond to the off-beam angle of most radio-relay antenna patterns because the patterns are usually measured in either the horizontal or the vertical direction. The true angular separation is neither, but either pattern can be used because most of the interference results when the separation angle is small, and the pattern is nearly symmetrical in all planes for small off-beam angles.

To summarize, the model incorporates a satellite spacing, radio-relay antenna pattern, number of hops, trendline, distribution, and station pointing distribution. These parameters are the inputs for a computer program which calculates the distribution of the total radio-relay interference.

III. DESCRIPTION OF THE SIMULATION PROGRAM

The location of the radio-relay system need only be identified in latitude. Since the stationary orbit is assumed filled with a constant satellite spacing, the longitude need not be a parameter of the interference. The starting assumption is made that the first radio-relay station is on the same longitude as one of the satellites. Moreover, the latitude of the center of the radio-relay system is used as a parameter to describe the geographical location of the radio-relay system and this parameter is definitely one of the variables of the system interference. The trendline direction is a random variable that was described earlier. This variable will not appear in the final answer because a distribution will be assumed and it will be averaged out. If a symmetrical distribution is assumed about zero degrees, there are some obvious points of symmetry that should be considered in order to save computation time. Assuming that the trendline direction is measured from the south, there is no need to consider negative trendline azi-

muths because duplicate results will be generated. Furthermore, irrespective of the assumed distribution, only the direction of transmission having southerly pointing receivers need be considered because that direction will receive the significant interference. Consequently, the azimuth directions that should be considered are between 0 and $\pi/2$ radians from south for the north latitudes.

The computer program begins by selecting a latitude for the center of the radio-relay system. Results are provided in this memorandum for latitudes between 20° and 70° in 10° increments for 50-hop systems and at 40° for a 140-hop system. Some care is required in selecting the latitudes used in the computer program. Those that are too small may cause the radio-relay system to cross the equator requiring some of the logic of the program to be changed to account for the crossover of the equatorial plane. All the derived geometry, however, is valid. For latitudes beyond about 81.3°, the entire geostationary orbit is below the horizon. Consequently, stations beyond this latitude will receive no interference.*

After a latitude has been selected, a trendline direction, τ, is chosen from a random number generator of numbers between 0 and 1.

$$\tau = 90° \cdot RND(\zeta). \tag{1}$$

The latitude of the first station with respect to the reference latitude, ϕ_r, is

$$\phi' = \frac{T}{2} \cos (\tau) + \phi_r \tag{2}$$

where T is the great-circle angular span of the radio-relay system. The average angular span for a CCIR system (50 hops) is near 22° and for an AT&T system (140 hops) about 61°.

With the trendline and antenna direction determined, the next calculation is the total interference into the first station from all visible satellites. The limits of the visible stationary orbit must be determined and this will, in turn, determine the number of visible satellites. First, the azimuth displacement, A, to the intersection of the geostationary orbit with the horizon is calculated and then the relative longitudinal displacement, λ, from the station is determined (see Fig. 2). The azimuth displacement† (measured from the south) is

$$| A | = \cos^{-1} \left[\frac{\tan \phi}{(K^2 - 1)^{\frac{1}{2}}} \right], \tag{3}$$

---

* Free-space propagation is considered.
† All geometry derivations are given in Appendix A.

Fig. 2—Geometry of the geostationary orbit and a radio-relay station.

where

ϕ = latitude of station,
$K = R/a$,
R = geostationary orbit radius and
a = earth radius.

The above equation assumes zero degree elevation for the radio-relay antenna and zero refraction. The problem could have been carried out with a non-zero elevation and with an appropriate distribution. However, most existing systems have a mean of near zero degrees. Also, taking refraction into account would have no significance in a statistical model. Consequently, any other assumptions would not have been much more realistic.

The relative longitudinal separation of the intercept is

$$\lambda = \sin^{-1} \{\sin A \ [1 - (K^{-1})^2]^{\frac{1}{2}}\}. \tag{4}$$

Since the span of the stationary orbit is symmetrical about the zero degree azimuth line, the total longitudinal span of the orbit is 2λ. Consequently, the number of visible satellites is the total of the one on the same meridian as the first radio-relay station and all other satellites included by incrementing the given satellite separation until the critical longitude, λ, is not exceeded. Also for negative azimuths the satellite spacing is stepped in negative increments until the negative critical longitude is reached.

The problem then reduces to the calculation of the angular separation from the beam of the radio-relay station to each visible satellite, and conversion to the appropriate interference suppression.

The azimuth, A_z, to each visible satellite is

$$A_z = \cot^{-1} \{\cot \lambda_r \sin \phi\}, \tag{5}$$

where λ_r is the relative longitude to the next satellite. The elevation angle to the satellite, assuming zero-degree elevation of the radio-relay antenna, is

$$\theta = \frac{\pi}{2} - [\beta + \Omega], \tag{6}$$

where

$$\beta = \cos^{-1} [\cos \phi \cos \lambda_r] \tag{7}$$

and

$$\Omega = \tan^{-1} \left[\frac{\sin \beta}{K - \cos \beta}\right]. \tag{8}$$

Finally, the separation angle, γ, as shown in Fig. 3, is given as

$$\gamma = \cos^{-1} [\cos \theta \cos (A_z - \delta)], \tag{9}$$

where δ is the random direction of the radio-relay antenna measured
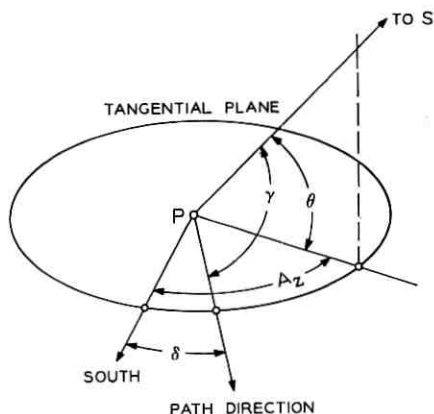
Fig. 3—Geometry determining the off-beam angle to a satellite.

from the south and is

$$\delta = (\tau + 25°) - 50° \cdot \text{RND} (\zeta). \tag{10}$$

ζ is a random number set as described earlier and thus allows δ to vary $\pm 25°$ about the chosen trendline.

The angle, γ, is one of the functions of the interference. However, before computing the interference, it is necessary to determine the location for the next station because it is at a different latitude, and it is improbable that it is on the same meridian as one of the satellites. The procedure is to calculate the new latitude and longitude of each succeeding station and start calculating the interference from the satellite that was used at the first radio-relay station. Naturally new azimuthal limits of the orbit, $|A|$, and longitudinal limits, λ, will be calculated. Furthermore, since the starting satellite is the same as that of the first station, the number of visible satellites on each side of the starting satellite will not be symmetrical.

The latitude shift of the next station is

$$\Delta\phi = \phi - \sin^{-1} \{\sin \phi \cos \rho - \cos \phi \sin \rho \cos \delta\}, \tag{11}$$

where ρ is the great-circle angular span to the next station. Since the total angular span was previously given as 22° for a 50-hop system, ρ is approximately 0.45°.

The longitude shift of the next station is

$$\Delta\lambda = \sin^{-1}\left\{\frac{\sin \rho \sin \delta}{[1 - (\sin \phi \cos \rho - \cos \phi \sin \rho \cos \delta)^2]^{\frac{1}{2}}}\right\}. \quad (12)$$

The geometry of the latitude and longitude shifts is shown in Fig. 4. From these values the azimuthal limit of the orbit and the relative longitude to each satellite can be calculated. Consequently, the beam separation angle from each radio-relay station to each visible satellite can be calculated.

To summarize briefly, the program starts at a given latitude and chooses a trendline direction. Then it shifts the latitude of the first station in order to make the center of the trendline lie near the middle of the chosen latitude, and it picks the beam direction of the first radio-relay station. It calculates the longitudinal limits of the visible stationary orbit and calculates the interference from each visible satellite taking into account the beam separation angle. The coordinates of the next station are calculated and the new set of visible satellites determined. The interference is once again calculated starting at the same satellite as the one used for the first station and proceeding in positive and negative satellite spacing increments until all visible satellites are exhausted. The interference to all radio-relay stations of a system is similarly calculated. The program proceeds with other sample trendline directions and chooses a sufficient number of direc-
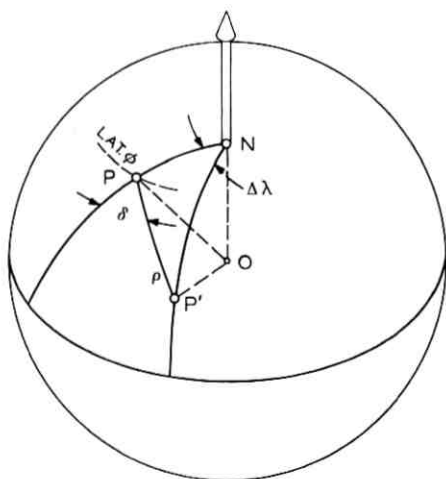

Fig. 4—Geometry determining the latitude and longitude shift.

tions to accurately arrive at an interference distribution. The results in Section V were derived for a sample of 50 trendlines at each latitude. Furthermore, the procedure is repeated at each chosen latitude.

The size of the sample appeared adequate for calculation of the mean; however, it was too small for calculation of the tails. Consequently, some smoothing was applied to the results, and some cases were rerun with larger samples to assure that the suggested trend of the 50-sample results was accurate. Increasing the sample size to greater than 50 was not too meaningful because the uncertainty of the radio-relay and satellite characteristics was large enough so that more accurate computation with higher samples would not have been more enlightening.

IV. COMPUTATION OF INTERFERENCE

The exact calculation of the interference between the satellite and radio-relay system involves the convolution of the two received spectral densities, with the suppression of the interfering spectral density being a function of the beam separation angle, γ. However, an approximate calculation is possible because the satellite signal is one of relative high index. Consequently, the simplifying assumption can be made that the interfering spectrum is substantially flat over the receiver bandwidth. It is possible that this assumption may not be valid when the index is not sufficiently high, or when the top baseband frequency of the satellite system is much smaller than that of the radio-relay system. However, even for these cases an upper bound calculation can be made by assuming a flat spectrum with the same magnitude as the peak of the actual spectrum. Also the assumption is made that carrier dispersal is used during periods of light loading.

The actual convolution could have been easily carried out, but the reason the computations were simplified is that the peak of the baseband interference is a function of the spectral densities, frequency separation of the two signals, and the de-emphasis of the radio-relay system. Hence, it would have introduced additional dimensions in the number of variables. Moreover, the effect in the results would have been insignificant. It is also true that a satellite signal without carrier dispersal and low-index of modulation may generate more interference. In any event, the flat spectrum approach will be used in order to make the results more tractable.

The baseband interference can be related to the thermal noise of the system by

$$\frac{i_4}{n_4} = \frac{i_c}{n_c}, \tag{13}$$

where i_4 and n_4 are the interference and thermal noise power respectively in a 4-kHz bandwidth at the input of the receiver, and i_c and n_c are the interference and noise power respectively in a voice channel. The baseband interference can be calculated as

$$i_c = \frac{i_4}{n_4} n_c . \tag{14}$$

The thermal noise input to the receiver is a function of the radio-relay system noise temperature

$$n_4 = kT_s b, \tag{15}$$

where

  k is Boltzmann's constant,
  T_s is the system noise temperature and
  b is the voice channel bandwidth.

The interference power input to the receiver is a function of the satellite power flux density, f, the elevation angle, θ, from the interfered radio-relay station to the interfering satellite, and the gain of the radio-relay antenna in the direction of the satellite. Hence

$$i_4 = f(\theta)g(\gamma) \cdot \frac{\Lambda^2}{4\pi\ell}. \tag{16}$$

Equation (16) contains the factor $\Lambda^2/4\pi$ because $f(\theta)$ is in units of $W/m^2 \cdot 4$ kHz (Λ is the wavelength of the carrier). Hence, what is required is the radio-relay station effective antenna aperture in the direction of the satellite. The factor ℓ is feeder losses. For the results of Section V, 3dB was assumed.

The baseband interference can now be written as

$$i_c = \frac{f(\theta)g(\gamma)\Lambda^2/4\pi}{kT_s b\ell} \cdot n_c . \tag{17}$$

It can be seen that interference can be expressed as the product of three factors:

  (i) The gain of the radio-relay antenna in the direction of the interfering satellite.
  (ii) The power flux density of the satellite, $f(\theta)(W/m^2 \cdot 4$ kHz):

where θ is the elevation angle from the interfered station to the interfering satellite.

(iii) A constant, μ, that depends on the parameters of the radio-relay system.

$$i_e = \mu f(\theta) g(\gamma). \tag{18}$$

μ can be easily derived from equations (13), (14), (15), (16) and (17) as

$$\mu = \frac{n_e \cdot (\Lambda^2/4\pi)}{kT_s b\ell}. \tag{19}$$

V. INTERFERENCE RESULTS

At the 1969 CCIR Interim Meeting in Geneva, Switzerland, a suggestion was generated to study the effect of a new power flux-density limitation. The proposed limitation is

$$F^*(\theta) = \begin{cases} -152 & 0° \leq \theta \leq 5° \\ -152 + 0.5(\theta - 5) & 5° \leq \theta \leq 25° \text{ dB Rel. } 1W/m^2 \cdot 4 \text{ kHz.} \\ -142 & \theta \geq 25° \end{cases} \tag{20}$$

This limit was used in the simulation program for both the CCIR and AT&T systems.

The parameters used for the CCIR reference system are

Antenna gain =

$$\left. \begin{array}{r} 0° \leq \gamma \leq 0.48° \\ 0.48° \leq \gamma \leq 57.5° \\ \gamma \geq 57.5° \end{array} \right\} G(\gamma) = \begin{cases} 42 \text{ dB} \\ 34 - 25 \log_{10} \gamma \quad \text{dB,} \\ -10 \text{ dB} \end{cases} \tag{21}$$

Hops = 50,

$T_s = 1750°\text{K}$,

$\mu = 5.8 \times 10^{13}$ at 4 GHz,

$M = 10 \log_{10} \mu = 137.6$, \hfill (22)

$I = F(\theta) + G(\gamma) + 137.6$.

The parameters used for the AT&T system approximate a TD-3

---

* Capitals are used for logarithmic quantities.

system and are

Antenna gain =

$$0° \leqq \gamma \leqq 0.575° \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} G(\gamma)^* = \begin{cases} 40 \text{ dB} \\ 34 - 25 \log_{10} \gamma \quad \text{dB}, \\ -10 \text{ dB} \end{cases} \quad (23)$$

Hops = 140,

$T_. = 1160°K$,

$\mu = 7 \times 10^{13}$ at 4 GHz,

$M = 10 \log_{10} \mu = 138.5$,

$$I = F(\theta) + G(\gamma) + 138.5. \tag{24}$$

Calculations were performed for the above systems using satellite spacings of 3° and 6°. Calculations were also carried out with the additional constraint that the radio-relay antenna not point within two degrees of the stationary orbit. This restriction was accounted for in the program by setting the angular separation between the radio-relay station and satellite to two degrees whenever it calculated to be less than two degrees. This extra set of calculations were performed because CCIR already has a recommendation at 6 GHz for new stations not to point within 2° of the stationary orbit whenever possible. Since Bell System 4-GHz routes are, in many cases, also 6-GHz routes, it is informative to evaluate the results of this restriction.

The interference distributions were calculated and are shown in Figs. 5–11. Figs. 5–10 include distributions for CCIR systems at latitudes of 20° through 70° for satellite spacings of 3° and 6°, and with and without the stationary orbit pointing restriction. The AT&T system is calculated at a latitude of 40° only; this result is shown in Fig. 11. Since it is a 4000-mile system, inclusion of other latitudes that span the U.S.A. would not have been more enlightening.

The attached results are not included as a check of the validity of the assumed power flux density, but as example calculations for the assumed statistical models. The program is general and could incorporate other radio-relay azimuthal distributions, satellite spacings, antenna patterns, and radio-relay system characteristics. It is also possible to use the program to calculate an appropriate power flux

---

* This pattern is pessimistic and not representative of the horn reflector antenna which has better off-beam discrimination.

Fig. 5—Distribution of interference into CCIR reference systems centered at 20° latitude with (a) 2° pointing restriction, and (b) no restriction.

density limitation to meet an interference objective. The answer is derived by iteration and it is not unique. However, it converges very quickly to a desired answer. Some parameters of the desired density may also be specified prior to computation and the program can generate the remaining parameters. For example, the power flux density at the horizon may be specified ($\theta = 0$), and the simulator can determine the slope and maximum density. Furthermore, joint distributions of parameters other than latitude can be calculated. Joint distributions with elevation angle as a parameter may be just as useful.

It may well be that simpler models can be formulated, but it may be difficult to convince the radio-relay community of their validity. The

program that has been derived can be made quite realistic, and perhaps the insight that can be gained from this model could lead to simpler models.

VI. CONCLUSIONS

The suggested analytical approach for the evaluation of satellite interference has practical value due to its generality and realism. The calculation techniques have been reported widely in the literature and have been substantiated by experiment. The contribution of this paper is a method for calculation of the probability of exposure to satellite

Fig. 6—Distribution of interference into CCIR reference system centered at 30° latitude with (a) 2° pointing restriction, and (b) no restriction.
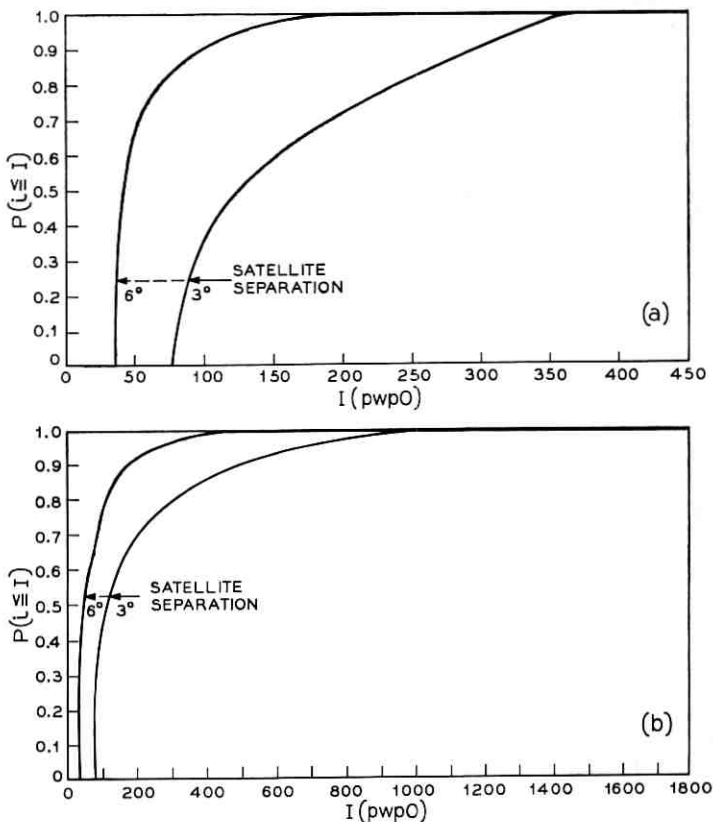
Fig. 7—Distribution of interference into CCIR reference systems centered at 40° latitude with (a) 2° pointing restriction, and (b) no restriction.

interference. The interference is a random variable and is a function of several variables.

The results were presented as a function of latitude because this appeared as a useful presentation for radio-relay operators. However, in attemping to reduce the number of variables, some very important parameters were averaged out.

One of these parameters is the elevation angle to the satellite. The simulation was also used to compute distributions as a function of the elevation angle, and the results showed the dominating effect of a direct exposure. Even at latitudes as low as 20°, more than 50 percent

of the interference was due to elevation angles less than 5°, and for latitudes near 70° the contribution of exposures of less than 5° elevation accounted for nearly 90 percent of the interference (See Fig. 12). The results for the high latitudes are not surprising. However, the lower-latitude results are surprising because the probability of intercepting the geostationary orbit within 5° is not high. Even though the number of off-beam contributions is very high compared to the contributions within 5° elevation angles, the radio-relay antenna directivity suppresses these contributions to account for less than 50 percent of the total interference. Nevertheless, their contribution is not negligible as assumed in previous interference models.
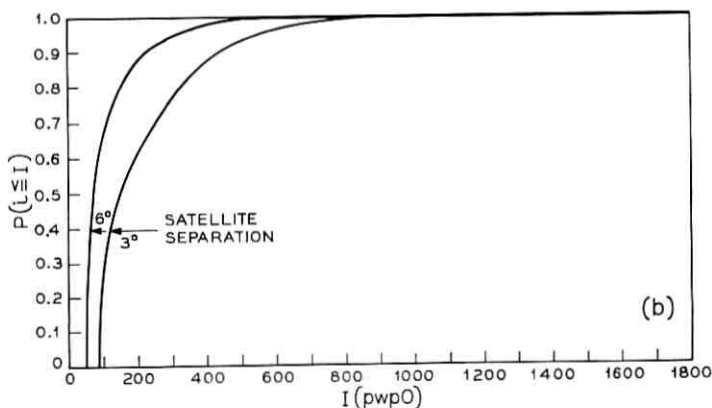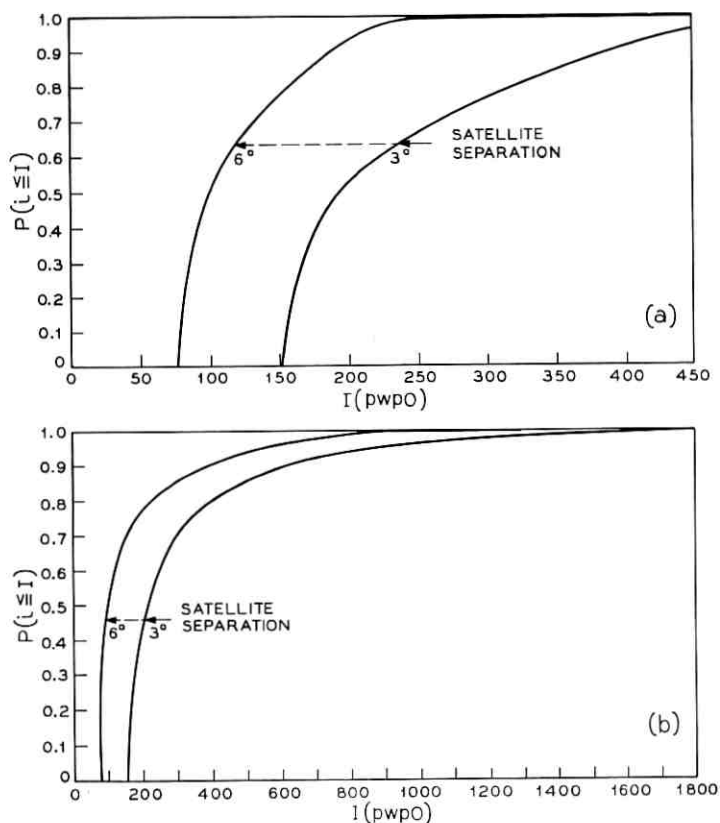

Fig. 8—Distribution of interference into CCIR reference systems centered at 50° latitude with (a) 2° pointing restriction, and (b) no restriction.
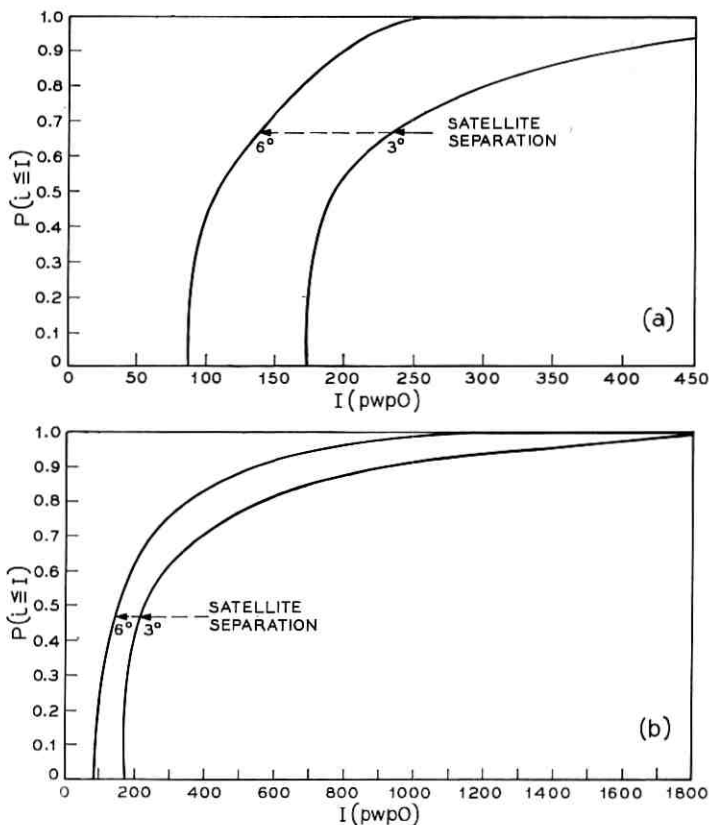
Fig. 9—Distribution of interference into CCIR reference systems centered at 60° latitude with (a) 2° pointing restriction, and (b) no restriction.

VII. ACKNOWLEDGEMENT

APPENDIX A

Derivation of Equations

Figure 2 illustrates the geometry of the problem. The azimuth displacement from the meridian through station P to a satellite suborbital

location S' is identical to angle A of spherical triangle PES'. The elevation angle θ(angle OPS $- \pi/2$) to a satellite at S is dependent upon the relative longitude λ between E and S' and the latitude ϕ.

A.1 Azimuth to Intercept

From the laws for right spherical triangles

$$\cos A = \tan \phi/\tan \beta, \tag{25}$$

where ϕ is the station latitude and β is the arc equivalent of angle O of plane triangle OPS. For the case of orbit intercept, and when θ is
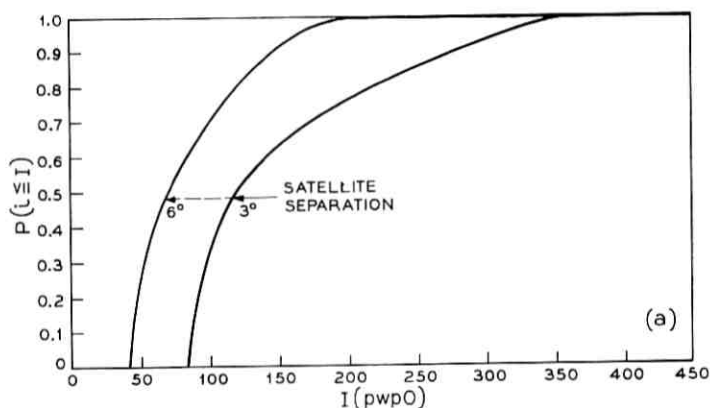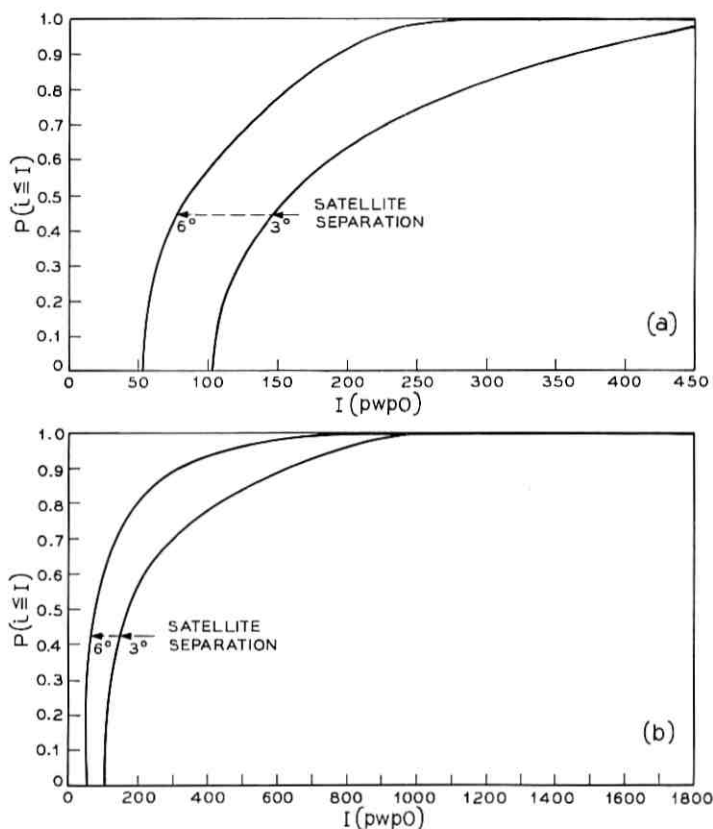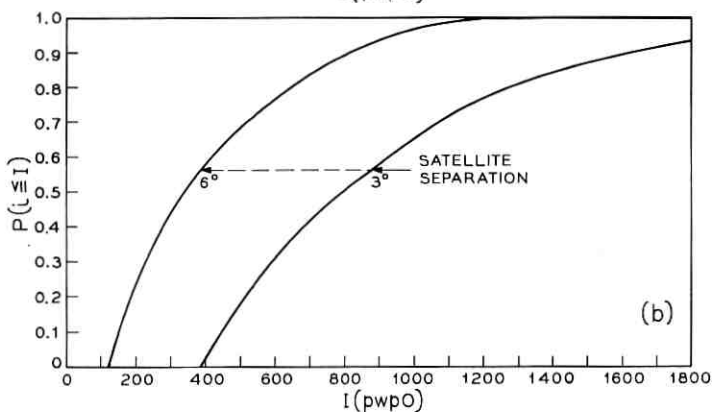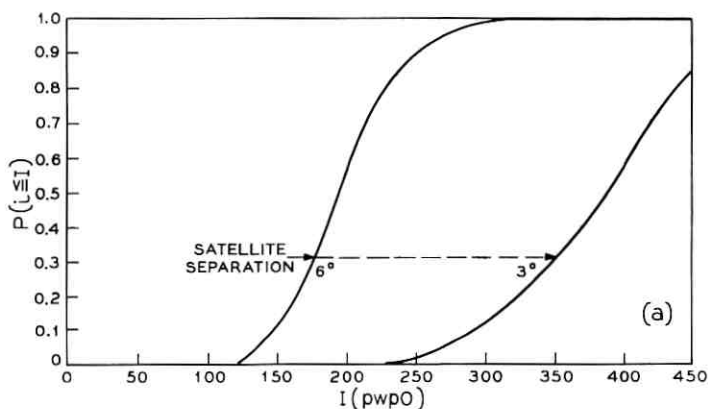

Fig. 10—Distribution of interference into CCIR reference systems centered at 70° latitude with (a) 2° pointing restriction, and (b) no restriction.
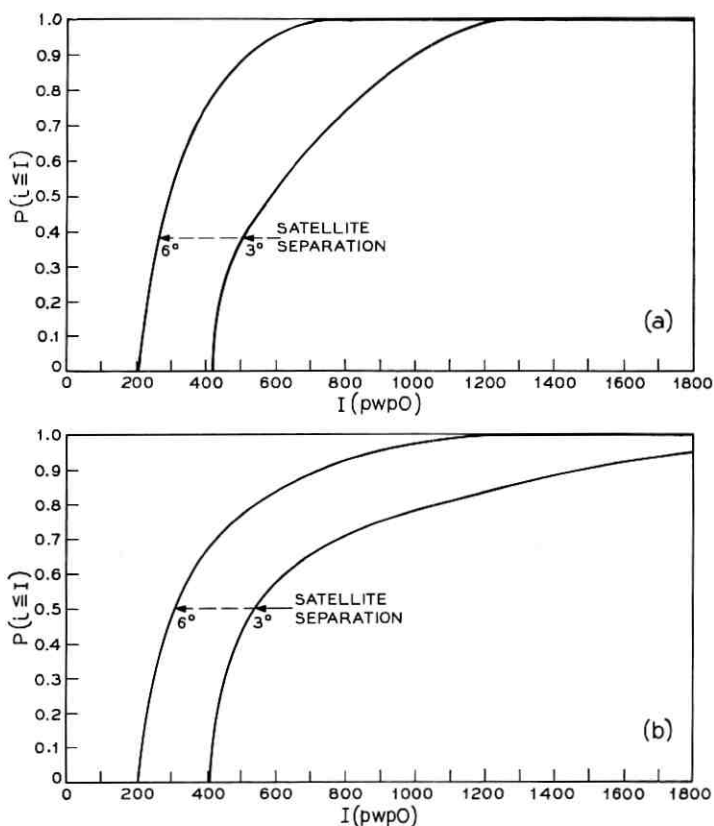
Fig. 11—Distribution of interference into 140-hop systems centered at 40° latitude with (a) 2° pointing restriction, and (b) no restriction.

zero, triangle OPS is a right triangle and β may be expressed as

$$\beta = \cos^{-1}(K^{-1}) \tag{26}$$

where $K = R/a$ (orbit radius/earth radius). From equations (25) and (26) the azimuth to intercept is given by

$$A = \cos^{-1}[\tan \phi/(K^2 - 1)^{\frac{1}{2}}]. \tag{27}$$

A.2 *Relative Longitude to Intercept*

The relative longitude to a suborbital intercept corresponding to azimuth A may also be obtained from spherical triangle PES'. From geometry

$$\sin \lambda = \sin A \sin \beta. \tag{28}$$

Combining equations (26) and (28) gives

$$\lambda = \sin^{-1}\left[\sin A\left(1 - \left(\frac{1}{K}\right)^2\right)^{\frac{1}{2}}\right]. \tag{29}$$

The total longitudinal span is 2λ.

A.3 *Azimuth to a Chosen Orbit Position*

For a chosen orbit position the relative longitude λ_r is known. Then, from triangle PES',

$$A_z = \cot^{-1}[\cot \lambda_r \sin \phi]. \tag{30}$$

A.4 *Elevation Angle to a Chosen Orbit Position*

From Fig. 2 the following relationships are established.

$$\theta = \frac{\pi}{2} - (\beta + \Omega), \tag{31}$$

$$\beta = \cos^{-1}(\cos \phi \cos \lambda_r), \tag{32}$$

$$\Omega = \tan^{-1}[\sin \beta/(K - \cos \beta)]. \tag{33}$$

Manipulating equations (31), (32) and (33) yields:

$$\theta = \frac{\pi}{2} - \cos^{-1}(\cos \phi \cos \lambda_r) - \tan^{-1}\left[\frac{(1 - \cos^2 \lambda_r \cos^2 \phi)^{\frac{1}{2}}}{K - \cos \lambda_r \cos \phi}\right]. \tag{34}$$
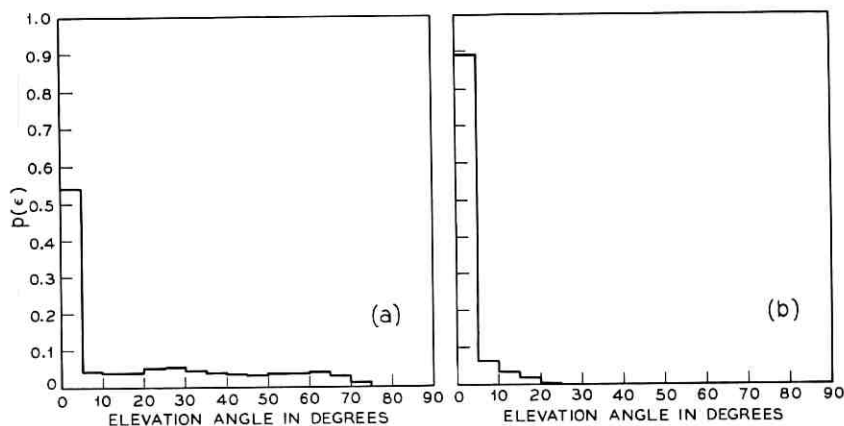
Fig. 12—Interference density vs. elevation angle into CCIR reference system centered at (a) 20° latitude and (b) 70° latitude. Satellite spacing = 6°.

A.5 *Angle Between Beams*

The off-beam angle between the radio-relay direction and a satellite is shown as angle γ in Fig. 3. Since the elevation angle θ has been determined in a plane perpendicular to a tangential plane at P, γ is given by

$$\gamma = \cos^{-1} [\cos \theta \cos (A_z - \delta)], \tag{35}$$

where δ is the direction of the radio-relay beam.

A.6 *Determination of the Latitude and Longitude Shift Between Adjacent Radio-Relay Stations*

In Fig. 4, P and P' represent locations of adjacent stations of a radio-relay system and ρ is the great circle angular span between them. If the average angular span of the entire system is T, then ρ is T/n where n is the number of hops. The value of ρ for a line of sight radio-relay system is about 0.45°.

It is seen that the latitude shift between P and P' is the angular difference between $\widehat{NP}$ and $\widehat{NP'}$ and that the longitude shift is equal to angle N of the spherical triangle NPP'. Since $\widehat{NP} = \pi/2 - \phi$, $\widehat{NP'} = \pi/2 - (\phi - \Delta\phi)$, and angle $NPP' = \pi - \delta$, the law of cosines gives

$$\Delta\phi = \phi - \sin^{-1} (\sin \phi \cos \rho - \cos \phi \sin \rho \cos \delta) \tag{36}$$

and the law of sines

$$\Delta\lambda = \sin^{-1} [\sin \rho \sin \delta / \cos (\phi - \Delta\phi)]. \tag{37}$$

Hence, from equations (36) and (37),

$$\Delta\lambda = \sin^{-1} \left\{ \frac{\sin \rho \sin \delta}{[1 - (\sin \phi \cos \rho - \cos \phi \sin \rho \cos \delta)^2]^{\frac{1}{2}}} \right\}. \tag{38}$$

REFERENCES

1. CCIR Recommendation 406-1, Documents of the XI Plenary Assembly, Oslo, 1966, 4, Part 1.
2. CCIR Recommendation 358-1, Documents of the XI Plenary Assembly, Oslo, 1966, 4, Part 2.
3. Lundgren, C. W., and May, A. S., "Radio-Relay Antenna Pointing for Controlled Interference with Geostationary Satellites," B.S.T.J., *48*, No. 10 (December, 1969), pp. 3387–3422.

Measurements of Atmospheric Attenuation on an Earth-Space Path at 90 GHz Using a Sun Tracker

By G. T. WRIXON

Measurements of 90-GHz attenuation through the total atmosphere due to absorption by clouds, fog and rain are presented and compared with simultaneous measurements at 16 GHz. It is seen that rain is the most important contributor to atmospheric extinction at 90 GHz, light rain causing greater then 25-dB attenuation on a number of occasions during the measurement period.

I. INTRODUCTION

Absorption in the atmospheric window between the 60-GHz oxygen absorption band and an oxygen absorption line at 118.75 GHz is due to absorption by the tails of atmospheric oxygen and water vapor features. Liquid water, when present in the form of rain, cloud and fog, also causes much attenuation in this atmospheric window. Measurements of attentuation through the total atmosphere are of importance in helping to decide the feasibility of earth-satellite communications in this frequency region.

Measurements of attenuation through the clear atmosphere as a function of atmospheric water vapor content have been made by F. I. Shimabukuro[1] on an earth-space path at 90 GHz. The results are in fair agreement with theory. A. W. Straiton and C. W. Tolbert[2] and D. C. Hogg and R. H. Turrin[3] have measured propagation loss through the clear atmosphere along ground-based paths at frequencies near 94 GHz. Their results agree satisfactorily with theoretical predictions provided suitable values are taken for the pressure broadening constants of the oxygen and water-vapor lines.[4] Hogg and Turrin[3] have also measured attenuation due to rain at 70 GHz along ground-based paths, and the results again show reasonably good agreement with theory.[4]

There is however a lack of experimental data on the effects of rain, cloud and fog on attenuation along earth-space paths at frequencies near 90 GHz. Theoretical treatment of the problem is hampered by the lack of detailed knowledge of rain and cloud characteristics. Recent papers by A. E. Freeny and J. D. Gabbe,[5] and R. A. Semplak[6] have discussed rainfall distributions measured in the neighborhood of Crawford Hill, New Jersey, and simultaneous short path attenuation measurements at 18.5 and 30.9 GHz. For the most part, however, little is known about the horizontal and vertical distributions of rain, and even less is understood about those of clouds.

The need for direct measurements of total atmospheric attenuation in this frequency region is thus apparent and it provided the incentive to carry out such measurements recently using the Crawford Hill sun tracker.[7]

II. APPARATUS

A receiver operating at 90 GHz was assembled and calibrated in the laboratory. The sun tracker normally operates at 16 and 30 GHz. For the purpose of this experiment the 90-GHz receiver replaced the 30-GHz receiver. A block diagram showing the receiver connected to the sun tracker can be seen in Fig. 1. Linear and logarithmic chart recorder outputs were available.

The receiver was measured to have a double sideband noise figure of 11.0 dB (i.e., a noise temperature of 3100°K) and a dynamic range of 25 dB when observing the sun.

A noise lamp, coupled into the line via a directional coupler, was calibrated using matched loads at room temperature and liquid nitrogen temperature.

The sun tracker antenna beam at 90 GHz was observed by allowing the sun to drift through it. The result, a convolution of the sun's brightness distribution with the antenna beam pattern, looked quite symmetric and showed no large sidelobes. From this measurement a beamwidth of 25 minutes of arc was computed for the sun tracker at 90 GHz. A similar measurement determined the beam width at 16 GHz to be 65 minutes of arc. Thus, almost the entire solar disc, about 32 minutes of arc, was covered by both beams.

III. MEASUREMENTS

The 90-GHz receiver was attached to the sun tracker during the period from December 2 to 16, 1969. During that time a variety of
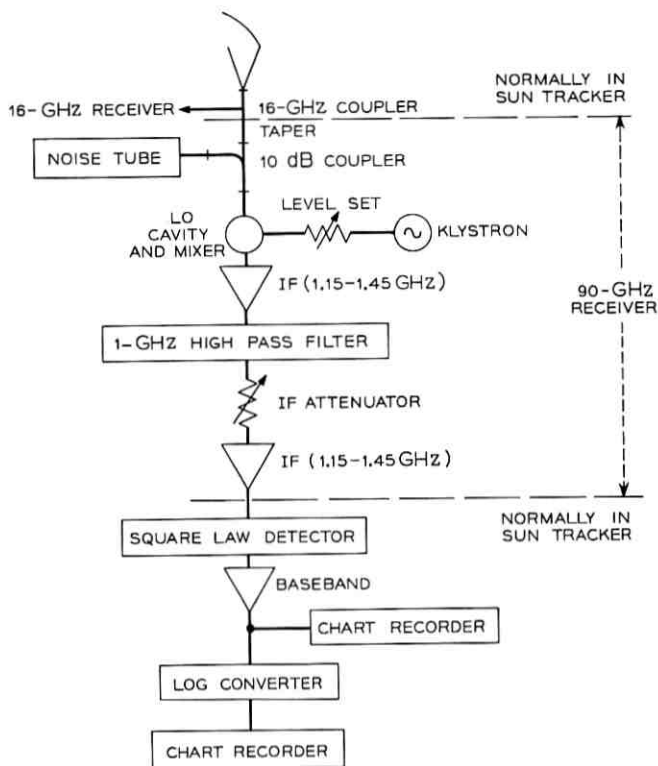
Fig. 1—Block diagram of the 90-GHz receiver in the sun tracker.

weather conditions, viz., cloud, thick fog, rain and some wet snow, occurred at Crawford Hill. The excess attenuation at 90 GHz was measured for all of these. No attempt was made to obtain attenuation statistics. In fact, once the attenuation through the clear atmosphere had been measured and the clear weather levels noted on the chart recorder, the receiver was switched on only during periods of bad weather or when cloud was present; in other words, only when excess attentuation was expected.

A discussion is given of the attenuation at 90 GHz caused by various types of weather conditions. To offset the necessarily qualitative descriptions for most weather conditions (e.g., light cloud, heavy fog, etc.) the excess attenuation at 90 GHz will be compared to that simultaneously measured with the 16-GHz receiver. In this way the 90-GHz data will be seen less as isolated attenuation measurements

due to random weather conditions, and more as results to be viewed in perspective with the more familiar 16-GHz data.

It should be borne in mind that at the time these measurements were made, the sun was very low in the southern sky, its elevation angle varying from about 27° at noon to 15° at 3 P.M. These angles correspond to path lengths through the atmosphere of 2.2 and 4.0 times the path length at the zenith respectively.

Attenuation through the clear atmosphere was measured by observing the intensity of solar radiation versus zenith angle. On a clear cold day (temperature, 32° F; relative humidity, 50%) a value of 0.38 dB was obtained for the attenuation through one atmosphere.

3.1 Excess Attenuation due to Clouds

Four basically different types of cloud cover were encountered, the attenuation characteristics of which are summarized in Table I. The values of attenuation were actually measured from the output of the linear recorder, because small dB changes were more easily measured here than on the output of the logarithmic recorder. Only those values of attenuation measured between 10 A.M. and 2 P.M. are quoted in the table. During these times the atmospheric path length took on values between 2.2 and 2.8 times the zenith path length. The maximum attenuation values quoted in the table did not occur at the time of the longest path length for any of the four types of clouds cover.

The measured ratios* of the 90- and 16-GHz cloud attenuations for the first two cloud cases are about what one would expect based on theoretical calculations. From expressions given by D. E. Kerr[8] for attenuation due to small water droplets, it is easily calculated that this ratio should be 27 for clouds at 20°C and 18 for clouds at 0°C. The ratio actually measured for the first two cloud cases (when the ground temperature was about 9°C), of about 23 is perhaps a little larger than would be expected since the cloud temperature is probably close to 0°C. This is a small discrepancy however.

The large value of attenuation measured in Case 3, i.e. during a period of heavy overcast between rainfalls, (see Fig. 2) was probably due to the fact that since precipitation occurred shortly afterwards, these clouds contained many large water droplets.[9] In this event, attenuation would be greater[10] than in the case of clouds of comparable water content which are not associated with precipitation and whose

---

* These ratios are estimated to be accurate to within about ±10 percent, this uncertainty being due mainly to inaccuracies in measuring the rather low values of 16-GHz attenuation (see Tables I and II).

TABLE I—SUMMARY OF CLOUD ATTENUATION MEASUREMENTS

Case Number and Type of Cloud Cover	Ground Temperature	Relative Humidity	90-GHz Atten.	Corresponding 16-GHz Atten.	Ratio $\frac{\text{90-GHz Atten.}}{\text{16-GHz}}$	Comments
1. Individual Cumulous Clouds in an otherwise clear sky	9°C	45%	Maximum: 3.7 dB Typical: 2.1 dB	0.16 dB 0.1 dB	~22	See text for note on maximum attenuation.
2. Overcast sky (no rain)	8°C	55%	Maximum: 1.6 dB	0.07 dB	~23	Signal continuously fluctuates as cloud cover moves across sky. Typical fluctuation ~1 dB at 90 GHz. No fluctuation at 16 GHz.
3. Heavily overcast sky between periods of rain (see Fig. 2)	10°C	100%	Maximum: 7.2 dB Typical: 5.2 dB	0.3 dB 0.21 dB	~24	
4. Overcast sky on cold day (no rain)	1°C	65%	Maximum: 1.64 dB	0.12 dB	~14	Most of the time there was little attenuation (<0.5 dB at 90 GHz) due to this cloud cover and no continuous fluctuations were present as in No. 2 above.
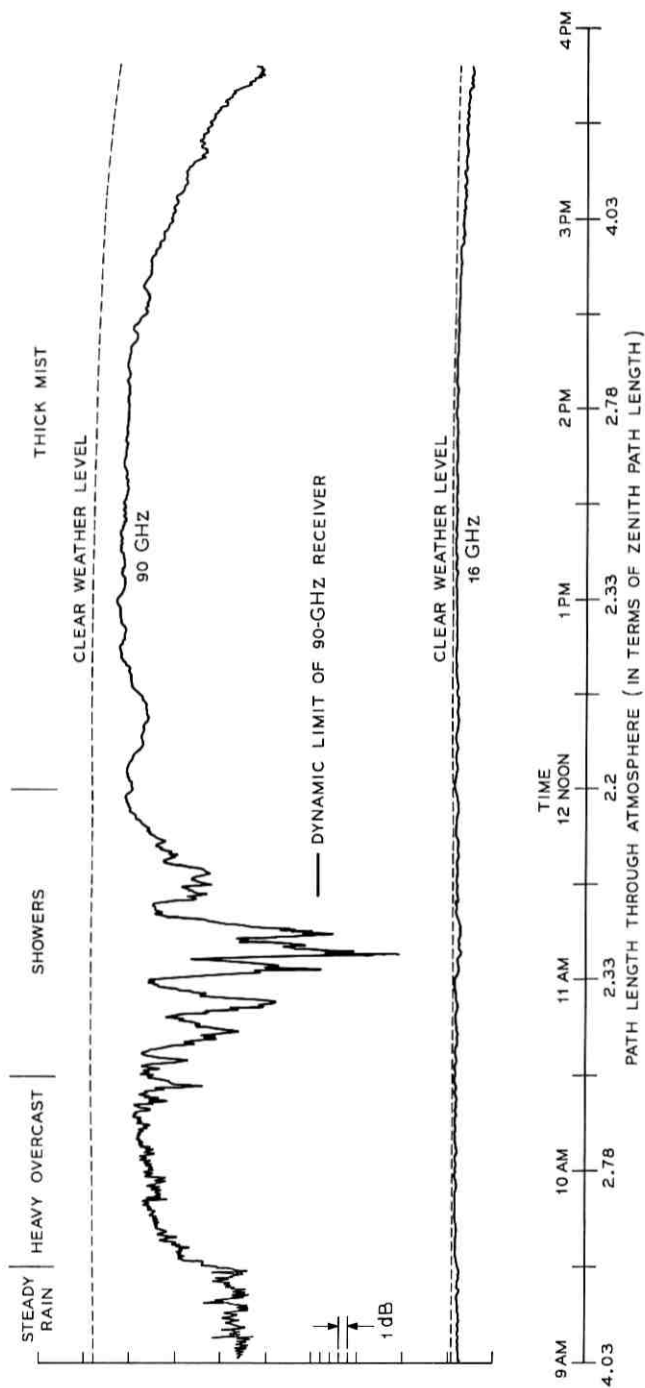
Fig. 2—Logarithmic record of 90-GHz and 16-GHz attenuation for December 8, 1969.

droplet size is very much less than a wavelength.[9] Another possibility is that part of the beam intercepted some rain; this would, of course, also have increased the attenuation.

Case 4 is different from the previous three cases because of the lower measured ratio of 90- to 16-GHz attenuation. Since the ground temperature was close to freezing during this measurement the cloud temperature was probably below $-10°C$. If the clouds consisted of pure ice, one would expect very little attenuation at either frequency[8] and, indeed, during much of this time, no measureable attenuation was present on either the 90- or 16-GHz channels in spite of the completely cloud-covered sky. This is in marked contrast to Case 2 when an overcast sky gave rise to continuous fluctuations of the 90-GHz signal.

On a number of occasions during the Case 4 experiment, however, some attenuation did occur, data for which are given in Table I. From what follows it seems clear that this attenuation was caused by clouds containing supercooled water droplets. Such droplets have been observed in clouds at temperatures as low as $-30°C$.[9]

Saxton[11] has deduced, from laboratory measurements, the dielectric properties of supercooled water at $-10°C$. If one substitutes these dielectric properties into Kerr's[8] expressions for attenuation due to small water droplets, an expected ratio of 90- to 16-GHz attenuation of 15 is obtained. This agrees very favorably with the measured result of 14 and suggests that clouds of supercooled droplets were indeed responsible for the attenuation experienced in Case 4.

3.2 Excess Attenuation due to Fog

Excess attenuation due to heavy fog was measured on one day, December 8. On this particular day attenuation due to cloud, rain and fog was measured. Figure 2 shows a copy of the logarithmic recorder outputs for both the 90- and 16-GHz channels for this data. Also shown is the clear weather signal level. It is seen from the figure that during the period of heavy fog, the signal fluctuations were much less rapid than during the period of heavy overcast. This suggests that the atmosphere was much more stable during the period of fog. Between noon and 2 P.M., the average attenuation was 3.2 dB at 90 GHz and 0.2 dB at 16 GHz, giving a ratio of 16.

If the prevailing ground temperature of $11°C$ is taken to be the absorbing temperature of the water droplets in the fog (since, typically, fogs do not extend very far above the ground) then from Kerr,[8] the expected ratio of 90- to 16-GHz fog absorption is calculated

to be about 24. The low measured ratio of 16 is probably due to the fact that during a fog the atmosphere is 100 percent saturated with water vapor. Since the ratio of attenuation at 90 and 16 GHz due to water vapor in the atmosphere is about 8.5 at a ground temperature of 15°C,[12] its presence in fog will tend to lower predicted ratios of attenuation based on small droplet absorption alone.

Indeed, when one calculates the attenuation due to the extra water vapor in the atmosphere on December 8th (i.e., over the amount present when the clear weather level was established), and subtracts it out, values of 0.1 dB and 2.35 dB are obtained for the average attenuation, at 16 and 90 GHz respectively, due to the fog water droplets alone. This gives a value of 23.5 for the ratio of 16- to 90-GHz attenuation which agrees with the calculated value of 24.

3.3 *Excess Attenuation due to Rain*

Attenuation due to rain was measured on three days, December 8, 10, and 14. During these times, rain rates not exceeding 1 mm/hour were measured using a rain gauge located on Crawford Hill, about 400 ft. from the sun tracker.

Table II summarizes the results of the measurements of rain attenuation. In Fig. 2, the logarithmic record for December 8, is seen the effect of attenuation due to light steady rain, between 9 and 10 A.M., and due to intermittent showers, between 10:30 A.M. and noon.

On December 10, a light rain was observed from 1:30 P.M. until 2:30 P.M. when the dynamic range (25 dB) of the 90-GHz receiver was exceeded. During this time the attenuation, both at 16 and 90 GHz, increased steadily between the limits shown in Table II and the dB ratio of the attenuation, fluctuating randomly between 25 and 29, had an average value of about 27.

T. S. Chu[12] has calculated expected rain attenuations for different rain rates using a Laws-Parsons drop size distribution. From Chu's work it is observed that as the rainfall rate increases, the ratio of attenuation at 90 and 16 GHz decreases. This can be seen in Fig. 3 where the ratio is plotted against rain rate. The decrease can be understood by first noting that in a Laws-Parsons distribution (see e.g., Chu and Hogg[13]) the average raindrop size increases with increasing rain rate. Now for each frequency there is a certain drop size which is most effective in absorbing electromagnetic radiation[13]; the higher the frequency, the smaller the most effective rain-drop size. At 16 GHz the most abundant raindrops have a smaller size than those which are most effective; consequently, the increasing size in a heavier rain

TABLE II—SUMMARY OF RAIN ATTENUATION MEASUREMENTS

Date Type of Rain	90-GHz Attenuation	Corresponding 16-GHz Attenuation	Ratio $\frac{\text{90-GHz Atten.}}{\text{16-GHz}}$	Comments
12/10: Light rain observed from 1:30 P.M. to 2:30 P.M. Between 1:30 P.M. and 3 P.M., a rainfall of 0.5 mm was measured by the Crawford Hill Rain Gauge.	Min. 5.22 dB Max. >25 dB	0.2 dB 0.9 dB	dB Ratio varied randomly between 25 ~ 29	Attenuation increased monotonically both at 16 and 90 GHz during this period.
12/14: Rain mixed with snow, turning to rain observed from 12:00 P.M. to 2 P.M. No rain gauge reading obtained as gauge was full of snow. Ground temperature, 35° to 37°F.	Min. 7.00 dB (end of rain) Max. 13.6 dB	0.3 dB 0.54 dB	Ratio decreased from 28 to about 17 and increased again to 23	The maximum 90-GHz attenuation was measured before the minimum ratio of 90/16 GHz attenuation was attained.
12/8: (a) Steady rain observed from 9 to 9:30 A.M. Rain gauge measurement ~0.8 mm/hr.	16.5 dB	0.6 dB	~27	Attenuation remained reasonably constant at both frequencies during this rain (see Fig. 2).
(b) Between 10:15 A.M. and noon, showers were observed. Total amount of rain that fell during this period (from rain gauge) was ~0.5 mm.	15 dB 20 dB 24 dB >25 dB	0.65 dB 0.75 dB 0.9 dB 1 dB	23 27 27 >25	These are the values of attenuation measured at the peak of various showers (see Fig. 2).
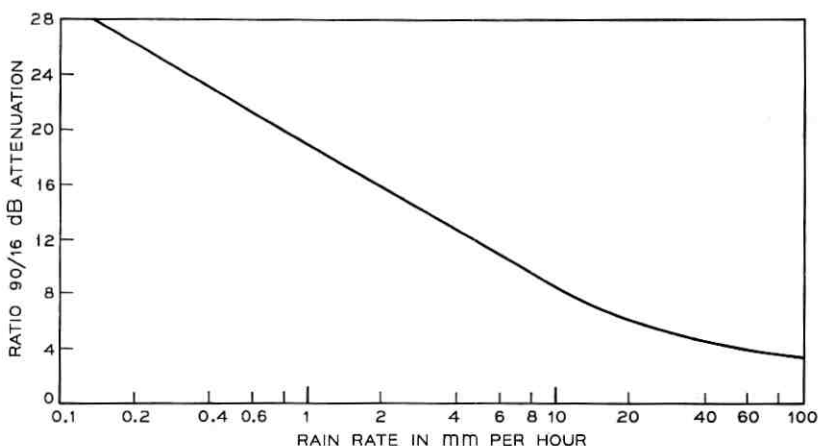
Fig. 3—Expected ratios of 90- to 16-GHz attenuation for different rain rates.

will increase the absorption effectiveness. At 90 GHz however, the most abundant drops are already larger than those which are most effective; consequently, the increasing size in a heavier rain will decrease the effectiveness.[10] The net result is to decrease the ratio of 90- to 16-GHz rain attenuation as the rain rate increases.

The observed ratio of 27 on December 10 indicates (see Fig 3) that attenuation due to a very light rain was being measured. It is interesting to note that even though the 90-GHz attenuation increased drastically during the measuring period, the ratio remained reasonably constant. This suggests that the increased attenuation was not caused by increasingly heavy rainfall at one point, but by the steadily increasing occurrence of light rain along the beam path through the atmosphere.

On December 14 the attenuation through rain mixed with snow, gradually all turning to rain, was observed. The ratio of 90- to 16-GHz attenuation showed a definite trend, beginning at a value of 28, decreasing to 17 and increasing again as the rain stopped to about 23. The maximum value of 90-GHz attenuation (13.6 dB) did not occur at the same time as the maximum value of 16-GHz attenuation (0.7 dB). The latter occurred while the ratio was at its minimum of 17 and about 10 minutes after the maximum attenuation at 90 GHz was recorded. If the effect of the snow is neglected, the above results can be interpreted, using Fig. 3, as indicating an increasing rain rate

along a decreasing path length, followed by a gradual tapering off of the rain. This is because even though the dB ratio decreased, indicating an increasing rain rate, the 90-GHz attenuation actually decreased, indicating a shorter path length through the rain. While attenuation due to snow is probably very slight at both 90 and 16 GHz,[8] its effect relative to rain is still unknown. This should be borne in mind when evaluating the above interpretation.

The data for December 8 are shown in Fig. 2 and summarized in Table II. The dB ratios are again consistant with the low measured rainfall, although it is perhaps surprising that during the showers the ratio did not decrease somewhat as one might have expected a higher rate of rainfall then.

IV. SUMMARY

Measurements of 90-GHz attenuation through the total atmosphere due to absorption by clouds, fog and rain have been presented.

Measured ratios of 90- to 16-GHz attenuation have been compared to theoretical predictions of the ratio in an effort to discover whether these predictions are reliable enough so that 90-GHz attenuation could be inferred from them. For the case of fog and cloud, this has in fact proved to be the case, provided suitable precautions are taken. In the case of rain, because of a lack of moderate or heavy rains at the sun tracker during the measuring period, it was not definitely established that the dB ratio decreased with increasing rain rate as predicted by theory, although variations in the observed values of the ratio are satisfactorily explained on this basis. However, the good agreement between theoretical and measured values of the dB ratio for the light rains encountered during the experiment make it reasonable to expect similar agreement at higher rain rates.

V. CONCLUSIONS

At 90 GHz it is seen that attenuation due to rain is the most important contributor to atmospheric extinction. This may be contrasted with the 8-14μ range where even cloud causes serious attenuation.[14] However, from the measurements given in this paper, it is seen that rain attenuation at 90 GHz is a very serious problem and that if an earth satellite communications channel is to operate in this band, sophisticated diversity schemes will have to be devised if outages are to be kept from exceeding permissable levels.

REFERENCES

1. Shimabukuro, F. I., "Propagation Through the Atmosphere at a Wavelength of 3.3 mm," IEEE Trans. Ant. and Prop., AP-14, No. 2 (March 1966), pp. 228–235.
2. Straiton, A. W., and Tolbert, C. W., "Factors Affecting Earth-Satellite Millimeter Wavelength Communications," IEEE Transactions, MTT-11, No. 5 (September 1963), pp. 296–301.
3. Hogg, D. C., and Turrin, R. H., unpublished work.
4. Hogg, D. C., "Millimeter-Wave Communication through the Atmosphere," Science, 159, No. 3810 (January 1968), pp. 39–46.
5. Freeny, A. E., and Gabbe, J. D., "A Statistical Description of Intense Rainfall," B.S.T.J., 48, No. 6 (July–August 1969), pp. 1789–1851.
6. Semplak, R. A., "The Influence of Heavy Rainfall on Attentuation at 18.5 and 30.9 GHz," IEEE Transactions, PGAP, AP-18, No. 4 (July 1970), pp. 507–511.
7. Wilson, R. W., "Sun Tracker Measurements of Attenuation by Rain at 16 and 30 GHz," B.S.T.J., 48, No. 5 (May–June 1969), pp. 1383–1404.
8. Kerr, D. E., (editor), Propagation of Short Radio Waves, New York: McGraw-Hill, 1951, Chapt. 8.
9. Fleagle, R. G., and Businger, J. A., An Introduction to Atmospheric Physics, New York: Academic Press, 1963, Chapt. III.
10. Van De Hulst, H. C., Light Scattering by Small Particles, New York: Wiley, 1962, Chapt. 20.
11. Saxton, J. A., "Dielectric Dispersion in Pure Polar Liquids at Very High Radio Frequencies, Part II," Proc. Royal Society, 213A (1962), pp. 473–492.
12. Chu, T. S., unpublished work.
13. Chu, T. S., and Hogg, D. C., "Effects of Precipitation on Propagation at 0.63, 3.5, and 10.6 Micron," B.S.T.J., 47, No. 5 (May 1968), pp. 723–759.
14. Wilson, R. W., "Measurement of Attenuation on an Earth-Space Path in the 8–14μ Atmospheric Window," Science, 168, No. 3938 (June 1970), pp. 1456–1459.

Effects of Transmission Delay on Conversational Behavior on Echo-Free Telephone Circuits

By PAUL T. BRADY

The present study sought to determine the effects of echo-free transmission delays of 600 ms and 1200 ms on conversational behavior when subjects were unaware that delay was present. Sixteen pairs of male subjects conversed 10 minutes each over an echo-free telephone circuit with zero delay and 10 minutes each with 600 ms round-trip delay; a second group of 16 pairs conversed 10 minutes each on zero delay and 1200 ms round-trip delay. Subjects noticed nothing unusual about the circuit, but the delay caused a statistically significant increase in frequency of confusions and in amount of both double talking (simultaneous speech from both speakers) and mutual silence. Analysis by means of an on-off pattern generating model revealed that subjects seem to make some adjustments in their speaking behavior; they tend to wait longer for responses and keep talking longer when interrupted. The two delay values of 600 ms and 1200 ms produced virtually identical effects in the measures made here. All of the subjects in the study were inexperienced with delay circuits; this experiment does not address itself to the possibility that experienced subjects might react differently to the delay. Further work will investigate the effects of training subjects to notice delay.

I. INTRODUCTION

1.1 *Purpose of Study*

This paper describes experiments that examined effects on the conversational behavior of subjects who talked on telephone circuits containing round-trip transmission delays of 600 ms and 1200 ms, typical of those in one- and two-hop synchronous satellites. However, satellite circuits also contain echoes and echo suppressors, which create added difficulties in conversation. The present study examines behavior on

echo-free circuits, which presumably represent the circuits that would result if echoes could be eliminated from existing circuits. In particular, answers are sought to the following questions:

(*i*) Does the introduction of delay on a circuit produce any measurable effects on speaking behavior, even though the subjects notice nothing unusual about the circuit?

(*ii*) Do the subjects experience difficulty in talking over delay circuits, even though they do not notice the difficulty?

This paper reports quantitative measures of the effects of delay indicating that the answer to both of the above questions is "yes."

1.2 *Background*

For several years, studies have been directed toward measuring customer acceptability of circuits containing both transmission delay and echo suppressors. For example, these circuits were used in one study,[1] in which hundreds of customer interviews were taken just after the customers placed transatlantic calls over such a circuit. A delay of 600 ms (as opposed to the cable transmission round-trip delay of 90 ms) produced a significant increase in the number of customers who reported having difficulty, and in addition produced an increase in other complaints and comments such as "cutting on and off," and "fading." However, such customer interview studies have limited application to echo-free delay circuits because (*i*) the echo suppressors are widely believed to contribute most of the degradation and probably overwhelm the delay effects, and (*ii*) as will be shown here, many of the delay effects are subtle, unnoticed by the conversants, and hence are not reported in interviews.

Regarding delay on 4-wire* echo-free circuits, only a few studies have been conducted. These have been limited to laboratory experiments because of the difficulty of providing 4-wire service to customers in the field.[2] P. D. Bricker[3] reported a laboratory study in which subjects who were talking over a standard circuit were suddenly presented with delay, introduced during conversation but not so as to disrupt a speech burst and provide an artificial cue. Subjects were told that delay would be introduced and were asked to try to detect its presence. It commonly took more than one minute just to detect a 1.2 s round-trip delay. R. M. Krauss and Bricker[4] had subjects con-

---

* The terms "pure delay," "echo-free delay," and "4-wire" are equivalent terms in this study, and refer to delay circuits in which echo is prohibited by using two separate voice paths, one from speaker *A* to *B* and one from *B* to *A*. Further, each path is connected through at all times so that full duplex communication is allowed.

verse and solve puzzles over a circuit with round-trip delays of zero, 600 ms, and 1800 ms. Measurements of word count and subjective opinions showed no significant differences between the zero and 600 ms delays, but the 1800 ms delay caused an increase in reports of "difficulty in communicating due to the circuit."

E. T. Klemmer[5] summarized some experiments conducted at Bell Laboratories in which certain people used 4-wire delay circuits in their regular telephone calls to each other, and could reject the circuit if they considered it unsatisfactory. They were not told the nature of the experimental circuit. Klemmer found that users were very seldom disturbed by delays of 600 ms and 1200 ms.

The previous studies seem to indicate that subjects are generally unaware of and undisturbed by round-trip delays up to 1200 ms. However, by its nature the delay does introduce changes in the temporal patterns of two-way conversation. The present study sought to determine the effects of these changes by comparing many measurements, especially those related to temporal patterns, on conversations on both standard and delay circuits.

II. THE BASIC MECHANISM OF DELAY INTERFERENCE

One might guess that the most obvious effect of introducing delay into a transmission link would be to make the conversation more sluggish; for example, responses to questions are delayed by the round-trip delay value. In fact, this effect is rarely noticed at 1200 ms and is essentially never reported at 600 ms.

There is another delay effect that is far more important at 600 ms and 1200 ms delay. If speaker A momentarily pauses, especially in such a way as to suggest a response from speaker B, A will not receive this response for the normal speaker response delay plus an additional 600 ms or 1200 ms. In the meantime, A, not hearing B's response at the expected time, may proceed on the assumption that B will not respond. Then, B's response will arrive and interrupt A's new talkspurt, thus causing an interference which at times the author has observed to momentarily disrupt normal conversation. D. L. Richards[6] has termed this an *involuntary interruption*, since from B's point of view, he never interrupted A.

Sometimes, during a lull in a conversation, both speakers begin to talk at approximately the same time, and each person will have generated several syllables or even words before he is conscious of the other person's delayed speech. This can cause confusion.

The principal deleterious effect of delay is, therefore, to cause speech

from one conversant to arrive at the other at unexpected times. This effect causes serious problems when echo suppressors are included,[7] but even on echo-free circuits the added difficulties are readily apparent, as is shown in Section IV.

III. PROCEDURE

3.1 *Experimental Circuit and Test Conditions*

Figure 1 is a diagram of the experimental circuit, which is designed to represent a standard long distance telephone circuit. The telephones were in two separate sound treated rooms; the experimenter could insert a 600 ms or 1200 ms delay from speaker *B* to speaker *A*.

It is well established, but not intuitively obvious, that the total round-trip delay can be distributed in any desired manner on the two transmission paths (*A* to *B* and *B* to *A*), as long as the total delay remains the same. Thus, the Fig. 1 configuration is identical to the subjects as a circuit in which half the delay is in each path. This point is explained in a previous paper.[8]

Sixteen pairs of male Bell Laboratories employees conversed ten minutes on the standard circuit and ten minutes on the 600 ms delay circuit, and 16 different pairs conversed on the standard and 1200 ms delay circuits for ten minutes each. In all, there were 32 different pairs of men. (Women were not used to avoid lengthening the experiment by adding a new variable.) Each pair consisted of good friends with mutual interests; the author has found that such pairs converse readily. The subjects were told to talk for 20 minutes on any topics they
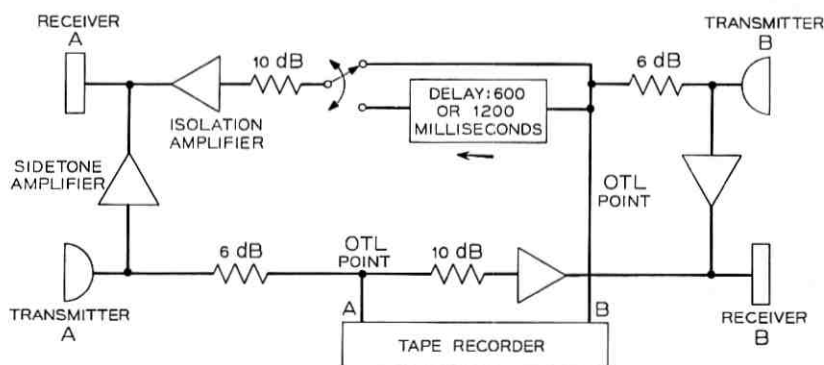
Fig. 1—Circuit over which subjects talked.

wished, but that they both should converse; monologues would be unacceptable. They were specifically instructed not to discuss the circuit, and that our interest was centered in recording a conversation for computer analysis of conversational behavior. They were told nothing about delay. Half the pairs began on the standard circuit, and half began on delay. After 10 minutes the circuit was changed, but in such a way to be undetectable to the subjects. (The delay from B to A was added or removed while A was talking.)

To obtain subjects inexperienced with delay, the subjects were chosen from areas of Bell Laboratories that had not been concerned with transmission delay. It was not possible to ask the subjects about delay prior to the experiment, as this would disclose the nature of the experiment. On interviews afterwards, one subject stated that he was familiar with delay circuits. His conversation was replaced with one that involved a new pair of subjects.

3.2 *Measurements Made on the Calls*

The following five measures were made on the calls.

3.2.1 *Subjective Measurements*

3.2.1.1 *Opinion of Subjects.* Immediately on completion of the call, the subjects were asked if they noticed anything unusual about the circuit. They were interviewed in separate rooms and did not hear each other's answers.

3.2.1.2 *Confusion Measures by Observers.* The author, assisted by a clerk, listened to the first five minutes of each condition. This covered 32 conditions with zero delay, 16 with 600 ms delay, and 16 with 1200 ms delay. A "confused situation" was scored whenever either speaker *reacted* to a double talk occurrence (simultaneous speech from A and B). The "reaction" had to contain a clear break in normal speech flow, such as a request for repeat ("What?"), or a self-generated repeat, or a sudden halt in the middle of an utterance. If both speakers simply continued to double talk, a confused situation was not scored, even though subsequent conversation indicated a lack of contextual understanding. Both observers had to agree on any confused situation for it to be scored.

Since the scoring instructions were imprecise, subjective judgment was used by the observers. To check validity, two other observers were given the above scoring rules and, without further instructions or practice sessions, were asked to observe four five-minute conversations.

3.2.2 *Objective Measurements*

3.2.2.1 *Speech Levels.* The equivalent peak level (epl) of a speech sample is an estimate of the peak of the instantaneous level distribution, based on a power measurement.[9,10] Epl readings were made of the first five minutes of every condition.

3.2.2.2 *Pattern Analysis.* An on-off speech pattern analysis was performed over the entire 10-minute length of each condition using a computer program written in 1963 by Mrs. N. W. Shrimpton. The program defines ten "events" related to the timing of the speech patterns, such as talkspurts, pauses, and double talk periods. The program produces considerable statistical data on the speech patterns, as illustrated in a previous paper.[11]

The patterns are strongly influenced by the design of the speech detector. In the present study, the detector was the same as that used in Ref. 11, with the threshold set to −40 dBm. (See Ref. 11 for further details on threshold settings.)

To show the effects of delay on speech patterns, the author has chosen to include data from two measurements. The first consists of the average lengths of each of the ten events. The second consists of the percent time each speaker spent in each of three states, namely, talking, double talking, and mutual silence. The talking time includes double-talking time. (These three measures are sufficient to calculate three others. A speaker is either talking or silent, hence, silent time is 100 percent minus talking time. "Solitary talking" time (no double talking) is talking minus double-talking time. "Listening" time, when a speaker is silent while the other is talking, is 100 percent minus talking minus mutual silence time.)

3.2.2.3 *On-Off Pattern Model Analysis.* The author has recently developed a stochastic model for generating on-off patterns that are statistically similar to those occurring in real conversation.[12] Using this model, only a few numbers are required to specify the conversational statistics. The model is described in Section 2.1.2 of Ref. 12. For convenience, the diagram of the model is included here in Fig. 2. Speaker A is considered to be in one of six states, as determined from the on-off speech patterns that appear at his side. The six parameters governing A's behavior are related to his probabilities for changing his talk-silence status. Numerically, these parameters, labeled α and β with suitable sub- and superscripts, specify the rates at which A departs from each state, as determined by Poisson processes. The larger the

parameter, the sooner A will tend to leave the state. Thus, high values for parameters indicate short durations for the corresponding states. The subscript abbreviations stand for *sol*itary, interrup*ted*, interrup*tor*, *pause*, *alt*ernate and *int*errupt.

In the present experiment, the entire 10-minute length of each condition was examined to obtain the six parameters for each speaker. These parameters are determined by measuring the frequencies of changing from state to state. The method of obtaining the parameters is described in more detail in Section 2.2 of Ref. 12.

3.3 Technique of the On-Off Pattern Analysis

The pattern analysis and model behavior, which are objective measures described in Sections 3.2.2.2 and 3.2.2.3 above, both use on-off speech patterns that are obtained with a speech-detector technique
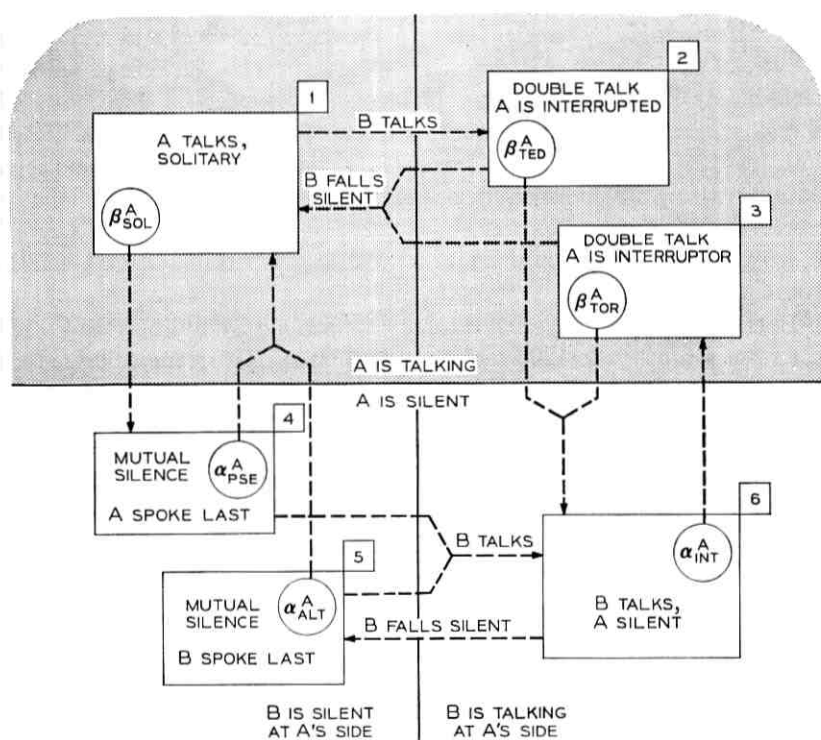
Fig. 2—The six-state model used in this study. Vertical transitions are due to Poisson processes at A's side. Horizontal transitions, due to B, are in A's external environment and are not generated by A's model.

described in previous papers.[11-13] The absolute values of any of the measures, such as average talkspurt length or the α and β values, are dependent on the speech detector parameters. For this reason, the absolute measure values were not used as an indication of the effects of delay. Instead, the *changes* in the values of the measures for each pair of subjects were noted when delay was inserted. By having each pair of subjects talk on both circuits in one session, and by balancing the pairs between delay-circuit-first and delay-circuit-second, changes in the measures from effects other than delay were minimized.

IV. RESULTS

The following is a comparison of the measurements on the standard and delay circuits.

4.1 *Opinion of Subjects*

When asked if they noticed anything unusual about the circuit, none of the 32 subjects on 600 ms delay commented on the delay. (A few stated they heard some background noise at low level.) Two pairs of the 32 subjects on 1200 ms delay made delay-related comments. These four people said they had "confused situations," but they noticed nothing wrong with the circuit. Some comments were, "I occasionally asked the other if he was still there," or "Sometimes, if we both started talking at the same time, I heard the other person but couldn't understand him."

It appears that inexperienced subjects are highly unlikely to notice the insertion of up to 1200 ms round-trip delay. This conclusion agrees with those of Bricker[3] and Klemmer.[5]

4.2 *Confusion Measures by Observers*

On listening to the conversations, both observers agreed there were many places in which confusion occurred when the delay was introduced. Table I summarizes the average number of confusions occurring in the first five-minutes of each condition.

Coincidentally, the 16 pairs of speakers on 600 ms delay had ex-

TABLE I—CONFUSIONS IN 5 MINUTES OF CONVERSATION, AVERAGED
OVER 16 CONVERSATIONS

	Standard Circuit	Delay
600 ms pairs	1.2	3.7
1200 ms pairs	1.2	4.3

actly the same average number of confusions on the standard circuit as did the speakers on 1200 ms delay. Thirteen of the 16 pairs experienced more confused situations on delay than on the standard circuit, for both values of delays. This increase is significant at the 0.05 level (sign test). Note that the average number of confusions on delay for 1200 ms is not much greater than for 600 ms. The difference between 3.7 and 4.3 confusions on the two different delays is not significant (t-test).

The four five-minute segments monitored by other observers consisted of two pairs of zero and 1200 ms delay. For these four segments, the author and clerk recorded 1, 5, 1, and 6 confusions (for zero, 1200 ms, zero, and 1200 ms delay). The other two observers, as a team, recorded 0, 6, 0, and 9 confusions. Though not in exact agreement with the author's measures, these measures also indicate an increase of confusions on the delay circuits.

Summarizing this Section, when delay was added, the subjects experienced a significant increase in confused situations or difficulties, although they noticed no change in circuit quality. The 1200 ms delay effect was not significantly different from the 600 ms delay effect.

4.3 *Speech Level Measurements*

Delay had little or no effect on speaking levels. For the 600 ms delay circuits, 19 out of 32 speakers had lower epl values on the standard circuit than on delay. For the 1200 ms delay circuits, 20 out of 32 speakers had lower values on the standard circuit. Neither of these ratios is significantly different from chance.

Since half the pairs received the delay circuit first, and half received the standard first, the standard versus delay comparison does not show a possible effect of time on speaking levels. We therefore tested the difference in speaking levels between the first and second parts of each conversation. For the 600 ms delay pairs, 11 out of 32 were quieter on the first segment, and for the 1200 ms delay pairs, the ratio was 12 out of 32. Neither is significant at the 0.05 level.

The average epl for all 64 speakers was −11.6 dBm, at the OTLP point of a simulated toll circuit. (The zero transmission level point is an arbitrary reference level used to establish relative levels in a telephone circuit.)

4.4 *Pattern Analysis*

Table II lists the average lengths of the ten events, defined in Ref. 11, for all conversations. In every case, the average length is taken over 32 speakers.

TABLE II—AVERAGE LENGTH OF TEN EVENTS ON STANDARD AND DELAY CIRCUITS
(See Ref. 10 for definitions of the events.)

Event	Average Length, Seconds		Significance Level*	Average Length, Seconds		Significance Level*
	Std. Circuit	600 ms Delay		Std. Circuit	1200 ms	
1. Talkspurt	1.426	1.344	—	1.271	1.241	—
2. Pause	2.174	1.994	—	2.007	2.192	—
3. Double Talk	0.264	0.324	0.01	0.274	0.357	0.01
4. Mutual Silence	0.490	0.550	0.01	0.558	0.614	0.01
5. Alternation Silence†	0.433	0.422	—	0.496	0.505	—
6. Pause in Isolation	0.520	0.560	0.1	0.580	0.645	0.05
7. Solitary Talkspurt	1.424	1.325	—	1.230	1.127	—
8. Interruption	0.726	0.795	—	0.805	0.690	—
9. Speech after Interruption	0.852	1.001	—	0.865	0.976	0.1
10. Speech before Interruption	1.245	0.964	—	0.929	0.931	—

* Significance levels are determined from sign tests on the number of subjects that change in the indicated direction.
† In Fig. 6 of Ref. 10, the alternation silences illustrated in the sample patterns are all incorrectly labeled. The A's and B's are transposed. An alternation silence from B to A occurs when B stops talking, a mutual silence ensues, and then A talks, as determined from analysis at A's side. These are counted for all speakers A. Analysis at B's side produces the A to B alternation silences, which are counted for speakers B.
Note: Each average length represents 32 subjects. The two "Standard Circuit" columns represent two different groups of 32 subjects.

When analyzing the patterns, a separate analysis must be made at each speaker's side, because the delay, when present, causes the temporal patterns to differ at the two sides. For example, with no delay, both speakers will have an identical number of double talks of the same average length. With delay, double talking at one side does not imply double talking at the other. Thus, on the delay circuits, a separate analysis is required for each speaker, not just for each conversation.

Significance tests were made from sign tests on the lengthening or shortening of events. For example, 30 out of 32 speakers experienced an increase in average double talking length when going from a standard circuit to the 600 ms delay circuit.

Table II shows that when delay was inserted, there was a statistically significant increase in the average lengths of double talks, mutual silences, and pauses in isolation. Event 9, speech after interruption, is not very much affected since it is significant at only the 0.1 level, and then only on the 1200 ms delay.

Consider now the amount of time spent in the talking, double-talking, and mutual-silence states. These results are shown in Table III. When delay is inserted, the percent time spent in both double talking and mutual silence increases significantly.

A fairly simple mechanism can account for most of the results shown in Tables II and III. The increase in duration of double talks and mutual silences, and also of pauses in isolation, which are a subset of mutual silences, can generally be explained by noting that speech from the distant talker (e.g., B) will sometimes arrive at A's side at inappropriate times. Specifically, responses to A's speech can arrive when A has resumed talking rather than during A's brief silences. Hence, the previous talking sequence A—mutual silence—B—mutual silence—A has been replaced with A—longer mutual silence—A—double talking, resulting in more time in mutual silence and in double talking.

There was little difference in the magnitude of the effect on patterns of 1200 ms versus 600 ms delay. From Table III, the double talking time for 600 ms delay was 5.9 percent versus 4.2 percent for the standard condition, a ratio of 1.40. This ratio was 1.41 for the 1200 ms case. The mutual silence ratios were 1.11 and 1.07 for 600 and 1200 ms, respectively. From Table II, the ratios of increases for average length of events for 600 ms and 1200 ms delays were 1.23 versus 1.30 for double talks, 1.12 versus 1.10 for mutual silences, and

TABLE III—PERCENT TIME IN THREE STATES ON STANDARD AND DELAY CIRCUITS, AVERAGED OVER 32 SPEAKERS

State	Standard	600 ms Delay	Significance* Level	Standard	1200 ms Delay	Significance* Level
Talking	40.5%	40.0%	n.s.	38.5%	38.4%	n.s.
Double Talking	4.2%	5.9%	0.01	4.4%	6.2%	0.01
Mutual Silence	23.0%	25.5%	0.01	27.2%	29.0%	0.01

* Significance based on sign test on number of subjects that changed in the indicated direction.

1.08 versus 1.11 for pauses in isolation. None of the differences in ratios between 600 and 1200 ms is significant.

To summarize the pattern analysis results, insertion of 600 or 1200 ms delay produced a statistically significant increase in the durations of double-talk and mutual-silence periods, and in the percent time spent in these states. There was no difference between the delays on the magnitude of the effect.

4.5 On-Off Pattern Model Analysis

Table IV lists the average values of the model parameters for all subjects. Two parameters change significantly. Let us first consider α_{pse}, the tendency* to resume talking after falling silent. Its value is lowered, indicating a tendency to remain silent longer, possibly in order to give the other speaker a chance to respond. The second changing parameter, β_{ted}, is also lowered, indicating a greater tendency to keep talking when interrupted.

The remaining four parameters remain unchanged. In particular, α_{int} hardly is affected; people's tendency to interrupt remains the same.

Colloquially, the downward changes in α_{pse} and β_{ted} as delay is introduced indicate that people exhibit an increased tendency to wait for replies, and have a greater tendency to ignore interruptions.

Table IV shows little difference between the effects of a 600 ms and 1200 ms delay, as noted in the ratio differences in parameters for the delay circuit versus the standard circuit.

The model concerns itself only with that aspect of human behavior reflected by a person's tendency to remove himself from a state once he enters it. The delay may rearrange speech patterns from speaker B so that, for example, A may be interrupted and thus double talk more often than he would without delay. This would increase the time A spends in state 2 of Fig. 2. In measuring A's behavior, the model is unconcerned that with delay, A may spend more time in state 2. The model does indicate, however, that when A is in state 2, he is less likely to leave when on a delay circuit than when on a standard circuit. Thus, when the model indicates that A's behavior changes with delay, it is referring to A's tendencies to talk and fall silent, and not the change in interaction of A's and B's speech patterns, which are largely controlled by delay.

---

*The α's and β's are not probabilities in a mathematical sense; they are instead arrival rates for "pulses" to change state. However, they are proportional to probabilities for very small time increments, as defined in the Poisson process. Here, they are colloquially referred to as "tendencies."

TABLE IV—ON-OFF SPEECH PATTERN MODEL PARAMETERS ON STANDARD AND DELAY CIRCUITS

	600 ms delay subjects				1200 ms delay subjects			
	Std. Circuit	Delay	Ratio Delay/Std.	Significance Level	Std. Circuit	Delay	Ratio Delay/Std.	Significance Level
β_{sol}	0.68	0.71	1.04	—	0.76	0.78	1.03	—
β_{ted}	1.93	1.41	0.73	0.01	1.88	1.33	0.71	0.01
β_{tor}	2.37	2.01	0.85	—	2.30	2.14	0.93	—
α_{pse}	2.13	1.79	0.84	0.01	1.87	1.59	0.85	(0.1)
α_{alt}	0.95	0.91	0.96	—	0.75	0.77	1.03	—
α_{int}	0.24	0.23	0.96	—	0.24	0.22	0.92	—

Note: Units of the parameters are "events per second," as explained in Ref. 11. Each parameter is averaged over 32 people. Significance level is determined from a sign test on the number of speakers that changed in the indicated direction.

The on-off pattern model can be applied to delay analysis in another way. Instead of looking at the way a speaker behaves, we can examine the way he *seems* to behave as viewed by the other speaker. For example, on the delay circuit, speaker B may talk during A's silence at B's side, but B's talkspurt, on arriving at A, may interrupt A's speech.

To perform the required analysis, all we need do is examine all the B speakers' patterns at the A terminals and vice versa. (Clearly, in the no-delay case the results will be the same as before since the patterns are identical at both sides.) Table V shows the results of the model analysis, viewing each speaker from his conversant's viewpoint.

Table V presents a striking contrast to Table IV. The first two entries β_{sol} and β_{ted}, are similar in both tables; β_{sol} remains unchanged and β_{ted} decreases. However, the parameter β_{tor} drops in Table V and remains unchanged in Table IV. This means, for example, that if, on a delay circuit, speaker A perceives B as an interruptor, B appears to A to tend to keep interrupting when in fact B is just as anxious to terminate double talking as he was on the standard circuit.

The effects on the α's are completely reversed in Tables IV and V. The drop in α_{pse} in Table IV is absent in Table V, implying that A perceives B as not tending to wait for replies when in fact he is. The drop in α_{alt} of Table V implies sluggishness; on delay, B seems much less likely to reply (or alternate) to A. In fact, his behavior is unchanged, as seen from α_{alt} in Table IV, but his replies now arrive after A has resumed talking. Thus, A perceives these as interruptions, as indicated by the rise in α_{int} in Table V. Whereas the tendency to interrupt remains the same (α_{int} is unchanged in Table IV), it appears to increase with delay.

As with all other results in this study, Table V shows no difference between 600 and 1200 ms delay on the apparent changes in the parameters.

We summarize this section by compiling a "profile" of the typical speaker conversing on a delay circuit, as expressed in terms of the model. From Table IV, his behavior changes slightly; it is as if he were more polite in letting his conversant have more time to respond (α_{pse} drops), and more persistent in continuing talking if interrupted (β_{ted} drops). But, from Table V he appears to the other speaker to be sluggish to respond (α_{alt} drops), interrupt more often (α_{int} increases about 50 percent), and once engaged in double talking, he appears reluctant to stop no matter who interrupted. This is an unflattering

TABLE V—MODEL PARAMETERS AS VIEWED FROM DISTANT SPEAKER

	600 ms delay subjects				1200 ms delay subjects			
	Std. Circuit	Delay	Ratio Delay/Std.	Significance Level	Std. Circuit	Delay	Ratio Delay/Std.	Significance Level
β_{aol}	0.68	0.69	1.01	—	0.76	0.78	1.03	—
β_{ted}	1.93	1.50	0.78	0.01	1.88	1.28	0.68	0.01
β_{tor}	2.37	1.83	0.77	0.01	2.30	1.63	0.71	0.01
α_{pse}	2.13	2.12	1.00	—	1.87	1.72	0.92	—
α_{alt}	0.95	0.61	0.64	0.01	0.75	0.47	0.63	0.01
α_{int}	0.24	0.35	1.46	0.01	0.24	0.39	1.63	0.01

portrait; to the extent that the model parameters correspond to aspects of human behavior, these effects may give rise to the confused situations that are observed on the delay circuits.

V. THE EFFECTS OF AMOUNT OF DELAY

On seeing the results of this experiment, one might be puzzled by the lack of difference in the effects of 600 vs 1200 ms delay. This may be reasonable, however, in the light of the following experiment conducted with simulated conversations.

Values of all α and β parameters were measured from 8 no-delay conversations, and parameter averages were taken for the A and B speakers separately. This was done to obtain a profile of two typical speakers. (These two profiles turned out to be very similar, as they should if the A and B recording and measuring circuitry were identical.) We then had "typical speaker" A engage in a series of computer-simulated 20 minute conversations with "typical speaker" B over delays up to 2400 ms. (The simulator is described in Ref. 12.) For each delay value, a separate simulation run was made such that the simulated speakers were substituted for the speakers shown in Fig. 1. Note that it would have been meaningless to record patterns of a simulated no-delay conversation and then replay those particular patterns with one speaker's patterns delayed with respect to the other's. In this case, each model would fail to respond to delay-induced changes in the other's speech.

An analysis of the type in Table IV applied to the simulated conversations would simply yield the α and β values used to generate the patterns. That is, the "true" values of α and β were fixed for all delays. But the apparent values, as seen from the other speaker, did change, and these ratio changes are plotted in Fig. 3. Each curve represents a single parameter, and each point on each curve is the average of two values, one for speaker A and one for speaker B.

On reading Fig. 3, remember that real people make changes in their true values of β_{ted} and α_{pse} when delay is introduced, while the simulated speakers did not. In Fig. 3, all of the six apparent parameters change, while in Table V only four changed. The discrepancy may be due to holding the β_{ted} and α_{pse} parameters constant in the simulated conversations; had these parameters been varied in the same manner as observed in real subjects, Fig. 3 might be more consistent with Table V.

The most important result shown in Fig. 3 is that the results tend to
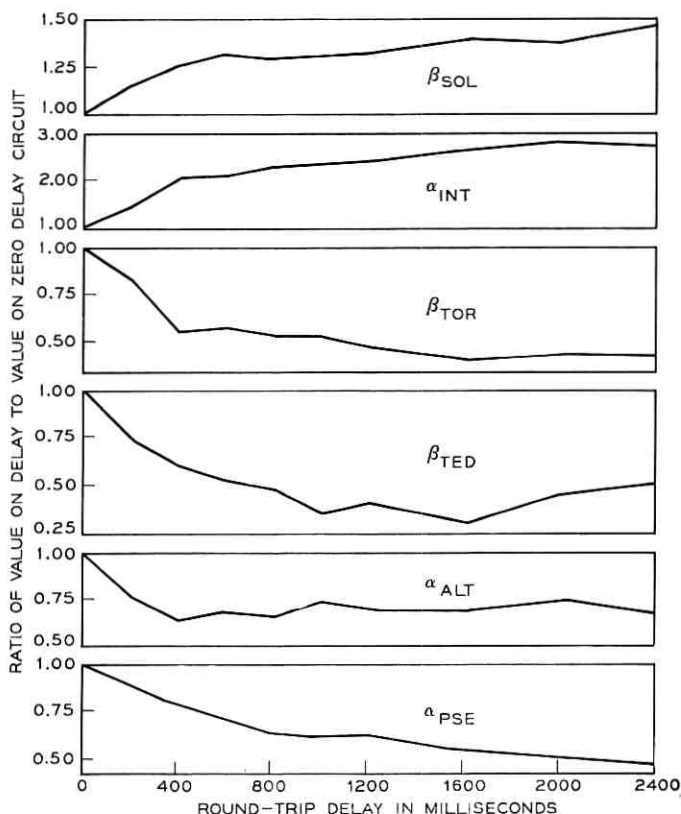
Fig. 3—Changes in the apparent values of the 6 model parameters when two simulated "speakers," with fixed parameters, "converse" over different delays. Each point is the average of the two speakers' apparent values.

asymptote with large delay values, and that they reach this asymptote in roughly 600 ms. *Therefore, a 600 ms round-trip delay is sufficient to produce all of the effects of apparent parameter change due strictly to delay.*

It is obvious that if *very* long delays are introduced, people will begin to notice the delay in responses to questions and will probably object to the delay for reasons other than confusion difficulties. Klemmer found a marked increase in dissatisfaction with 2400 ms delay.[5] The simulated speakers of Fig. 3 respond only to on-off patterns, and not contextual cues, and will never "notice" this sluggishness.

VI. CONCLUSIONS

When delay is introduced into a two-way telephone conversation between people who are inexperienced on delay circuits, measurable changes occur in their speaking behavior, even though the subjects notice no change in the circuit. The subjects become confused more often, engage in more double talking and mutual silence, and exhibit certain changes in their on-off pattern generation behavior. In addition, when a talker's speech patterns are analyzed at his conversant's side, he appears more likely to interrupt, less likely to terminate interruption, and less likely to respond to questions or pauses.

A round-trip delay of 600 ms seems large enough for all of the above effects to reach their asymptotic values. As delay becomes longer than 600 ms, additional degradation of the type measured here does not increase, but the subjects eventually may notice the sluggishness caused by very long delays.

If echoes and echo suppressors are present, other studies[1] indicate that the circuits are perceived to be unsatisfactory by some people, and that the effects get worse as delay is increased from 600 to 1200 ms. This is probably due to the mutilating effects on speech, which are caused by the echo suppressors. If, however, echoes are controlled to the point that the circuit appears to be a 4-wire circuit, inexperienced subjects or customers might behave in a manner similar to those subjects in the present study.

An important question not answered by this study is what will happen as subjects become experienced on the delay circuits. The question of behavior after acquiring experience provides a natural direction for future research that the author intends to pursue on this problem.

REFERENCES

1. Helder, G. K., "Customer Evaluation of Telephone Circuits with Delay," B.S.T.J., 45, No. 7 (September 1966), pp. 1157–1191.
2. Emling, J. W., and Mitchell, D., "The Effects of Time Delay and Echoes on Telephone Conversations," B.S.T.J., 42, No. 6 (November 1963), pp. 2869–2892.
3. Bricker, P. D., Satellite Communications Physics, Chapter 5, Foster, R. M., editor, Murray Hill, N. J.: Bell Telephone Laboratories, 1963, pp. 70–77.
4. Krauss, R. M., and Bricker, P. D., "Effects of Transmission Delay and Access Delay on the Efficiency of Verbal Communication," J. Acoust. Soc. Am., 41, No. 2 (February 1967), pp. 286–292.
5. Klemmer, E. T., "Subjective Evaluation of Transmission Delay in Telephone Conversations," B.S.T.J., 46, No. 6 (July-August 1967), pp. 1141–1147.
6. Richards, D. L., "Conversational Performance of Speech Links Subject to Long Propagation Times," Proc. I.E.E. (London), 116 (1969), pp. 955–963.

7. Brady, P. T., and Helder, G. K., "Echo Suppressor Design in Telephone Communications," B.S.T.J., *42*, No. 6 (November 1963), pp. 2893–2918.
8. Brady, P. T., "A Stochastic Model of Message Interchange on a Channel with Transmission Delay," I.E.E.E. Trans. on Commun. Technology, *COM-15*, No. 3 (June 1967), pp. 405–412.
9. Brady, P. T., "Equivalent Peak Level: A Threshold-Independent Speech-Level Measure," J. Acoust. Soc. Am., *44*, No. 3 (September 1968), Pages 695–699.
10. Brady, P. T., "A Comparison of the Equivalent Peak Level with Other Speech Level Measures," unpublished work.
11. Brady, P. T., "A Statistical Analysis of On-Off Patterns in 16 Conversations," B.S.T.J., *47*, No. 1 (January 1968), pp. 73–91.
12. Brady, P. T., "A Model for Generating On-Off Patterns in Two-Way Conversations," B.S.T.J., *48*, No. 7 (September 1969), pp. 2445–2472.
13. Brady, P. T., "A Technique for Investigating On-Off Patterns of Speech," B.S.T.J., *44*, No. 1 (January 1965), pp. 1–22.

Inherent Load-Balancing in Step-By-Step Switching Systems

By M. M. BUCHNER, JR., and S. R. NEAL

(Manuscript received June 19, 1970)

Questions have been raised over the years regarding the validity of the traffic-engineering tables used by the Bell System and others for the graded-multiple trunk groups within the step-by-step switching system. The tables indicate that the capacity of a graded multiple is increased when it is imbedded in the step-by-step system, the increase becoming larger as the number of switches connected to the graded multiple decreases. The increase in capacity supposedly occurs because of a "finite-source effect."

In this paper, we investigate in detail the flow of traffic through the step-by-step system to determine the validity of the tables. By a combination of simulation and analysis, we show that the increase in capacity arises not from a finite-source effect but from an inherent load-balancing that results from the clever manner in which the switches and trunks are interconnected. We conclude that, when used for engineering in the presence of day-to-day variations, the tables are adequate when the gradings are used with the large selector groups. For the minimum number of selectors, the tables estimate fairly accurately the capacity of the larger gradings, slightly overestimate the capacity for the medium-sized gradings, and overestimate the capacity of some of the smaller gradings by about 20 percent.

I. INTRODUCTION

The graded-multiple trunk groups that connect the successive switching stages of the step-by-step switching system* are commonly engineered according to a set of tables furnished by American Telephone and Telegraph Company.[1] Because the tables, which were prepared some years ago, are based upon certain approximations and a good deal of engineering judgment and because the arguments used to

---

* It is assumed that the reader is familiar with the step-by-step system. For those who are not, a summary of the system is given in Appendix A.

justify the tables are not convincing, their validity has been questioned.[†]

Specifically, the tables indicate that the traffic capacity of a graded multiple is increased when it is imbedded in the step-by-step system and that the increase becomes more significant as the number of switches connected to the grading decreases. The common "explanations" for the increase in capacity usually relate in some imprecise manner to a "finite-source effect."

Several common incorrect arguments[2-4] that relate to the capacity of the graded multiples in the step-by-step system are summarized in Appendix C. It is important to realize the nature of these arguments in order to appreciate the uncertainty that has persisted concerning the validity of the tables and to view the present study in the proper perspective.

To clarify the issues involved, we examine in detail the flow of traffic through the step-by-step system. We demonstrate that the capacity of a graded multiple is increased when it is imbedded in the step-by-step system and that the increase becomes larger as the number of switches connected to the grading decreases. We show that the increase in capacity results from an "inherent load-balancing"[*] caused by the way in which the switches and trunks are interconnected rather than from a finite-source effect. By a combination of simulation and analysis, we conclude that the accuracy of the tables depends upon the size of the grading and the number of selectors.

II. INHERENT LOAD-BALANCING

In this section, we show qualitatively why the capacity of a graded multiple increases when traffic is offered to the graded multiple through the line finders and first selectors.

Two fully equipped line-finder groups and four first-selector half-shelves are pictured in Fig. 1. For clarity, the line-finder to first-selector connections are shown only for one line-finder group. The graded multiple in Fig. 1 is connected to a particular level of the first selectors (we refer to the level as "our level"). The graded multiple is used by calls directed to our level; calls to other levels and, thus, to other destinations use trunk groups that are not shown.

---

† There are indications that part of the uncertainty surrounding the tables is due to an incomplete knowledge of the mathematical models used to generate them. Since this information is not readily available, some of the history of the development of the tables is presented in Appendix B.

* A phrase suggested by E. E. Sellars, American Telephone and Telegraph Company.

Assume that the five calls indicated by dashed lines in Fig. 1 are in progress. Now, suppose that a subscriber served by the top line-finder group originates a call. Of the 16 idle line finders, only one can direct the call into the top selector half-shelf, i.e., the relatively congested part of the system. However, there are 15 line finders that will direct the call into one of the lower selector half-shelves, i.e., the relatively uncongested part of the system. The important point is that the call is much more likely to be routed through the relatively
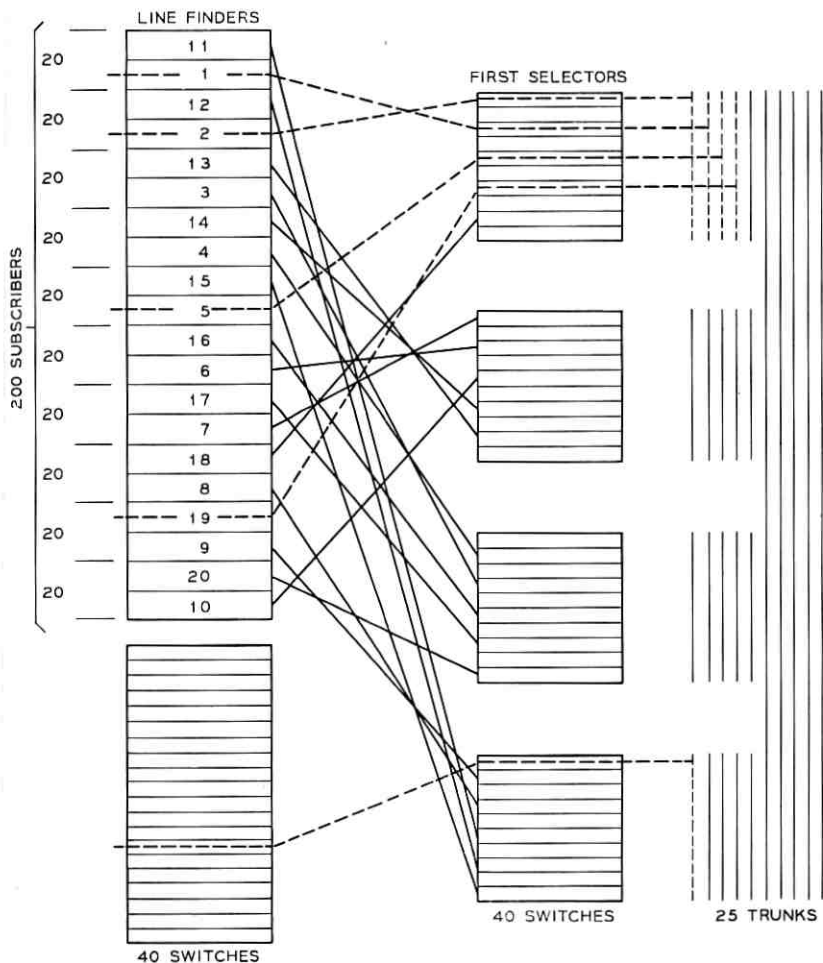

Fig. 1—Inherent load-balancing.

free part of the system than through the relatively busy part. The bias toward the less occupied selector half-shelves increases as the imbalance in occupancy increases. Therefore, we have the desirable situation wherein new arrivals tend to be offered to the less occupied first-choice subgroups. This results in a more uniform filling of the grading and a corresponding reduction in the chance that a call is blocked at one first-choice subgroup of the grading when there are idle trunks in some other subgroup. That is, the capacity of the grading is increased. The term "inherent load-balancing" is used to refer to this effect.

III. MATHEMATICAL MODELS

We wish to measure quantitatively the increase in grading capacity that results from the inherent load-balancing and to determine the validity of the tables.

3.1 *Complete Model*

For the complete model, the step-by-step system is considered in detail. Traffic is offered to the graded multiple through the complex interactions of the line finders and first selectors. Included in the model are all physical characteristics of the switching system deemed pertinent to the determination of the capacity of the grading. (These characteristics are specified in Appendix A.) The following assumptions are made concerning the traffic.

(*i*) Each line-finder group is fully equipped (i.e., contains 20 switches) and serves 200 subscribers.* Requests for service from each group of 200 subscribers are approximated by a Poisson process. The separate processes for the several groups are assumed to be independent with the same mean.

(*ii*) The 200 subscribers served by a line-finder group are divided into 10 subgroups of 20 subscribers each. For the ith subgroup of the jth line-finder group, let $b_{i,j}$ denote the number of busy subscribers. Given that the next arrival occurs in the jth line-finder group, the probability that the next arrival occurs in the ith subgroup is

$$\frac{20 - b_{i,j}}{200 - \sum_{i=1}^{10} b_{i,j}}.$$

---

* In practice, the number of working lines is not greater than 194. However, we assume 200 lines both because of the resulting numerical convenience and because of the negligible effect upon our results.

(*iii*) The line finders operate as described in Section A.2. Calls blocked at the line finders are delayed. An important aspect of the complete model is that the actual line-finder to first-selector wiring patterns are used. Therefore, traffic is distributed over the first selectors exactly as in the physical system.

(*iv*) Arrivals go to our level with probability p_l , i.e., p_l is the proportion of the calls arriving at the first selectors that require a trunk in our graded multiple.

(*v*) Holding times are independent and identically distributed according to a negative-exponential distribution.

(*vi*) Calls blocked at the graded multiple leave the system immediately and do not return. Calls directed to other levels are never blocked and remain in the system for one holding time.

3.2 *Approximate Model*

The complexity of the complete model arises both from the interactions occurring in the subscriber to line-finder network and from the interconnections between line finders and first selectors. To reduce complexity (and simulation computing time), an approximate model has been developed. In the model, the line finders and the line-finder to first-selector interconnections are modeled by assuming that an arrival seizes a selector at random from the group of idle first selectors.

The model is illustrated in Fig. 2. Suppose that the five calls indicated by dashed lines are in progress when a new call arrives. The probability that the call is served by a selector in the top half-shelf is 5/35 whereas the probability that the call is served by a selector in, say, the second half-shelf is 10/35. Because the arrival tends to be directed to the relatively uncongested part of the system, the model provides a good characterization of the inherent load-balancing.

The following assumptions are made concerning the traffic in the approximate model.

(*i*) All subscribers who can originate calls to our grading are considered as one large group. Requests for service from the group of subscribers are approximated by a Poisson process.

(*ii*) An arrival has full access to the first selectors and seizes a selector at random from the group of idle first selectors.

(*iii*) Assumptions (*iv*) and (*v*) from the complete model are also used here.

(*iv*) Calls blocked at the selectors or at the graded multiple leave the system immediately and do not return. Calls directed to
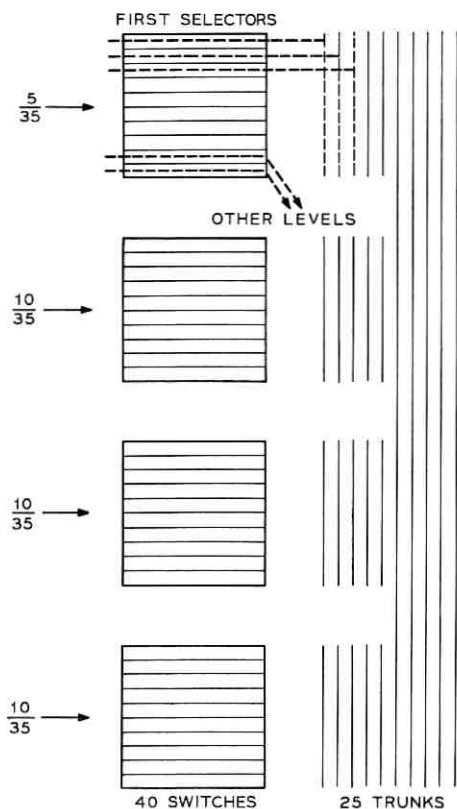
Fig. 2—Approximate model.

other levels are never blocked and remain in the system for one holding time.

3.3 Isolated Model

We wish to measure the increase in capacity that results from imbedding a graded multiple in the step-by-step system. Thus it is necessary to determine the capacity of the graded multiple when all traffic effects that arise because of the line finders and first selectors are ignored. This is called the isolated model. The load-loss relations for the isolated model can be computed by means of the equivalent-random method.[5] The following assumptions are made concerning the traffic.

    (i) Calls arrive at the grading according to a Poisson process.

    (ii) Arrivals are uniformly distributed over the first-choice subgroups.

(*iii*) Calls blocked at the graded multiple leave the system immediately and do not return.

(*iv*) Holding times are independent and identically distributed according to a negative-exponential distribution.

IV. NUMERICAL RESULTS AND EVALUATION OF TABLES

It is difficult to analyze a graded multiple even without the complexity of the step-by-step system. The approximations most often employed (such as the equivalent-random method[5]) are not directly applicable when the inherent load-balancing affects the capacity of the grading. Thus, two random-walk simulations were constructed, one using the complete model and one using the approximate model.

Considerable effort was devoted to determining the number of calls that should be processed at various loads to achieve a reasonably homogeneous coefficient of variation for the blocking probability. By a detailed statistical analysis of the simulation results, it was shown that, for a load a, the desired homogeneity is obtained by processing $1500/[p_t B(a)]$ calls where $B(a)$ is the blocking probability. The resulting coefficient of variation fell in the range of four to five percent. Since the number of calls processed (and, thus, the computing time) increases rapidly as p_t becomes small, our results are for $p_t \geqq 0.5$.

Because the computing time is sensitive to p_t, it would be desirable to show that meaningful results can be obtained when the multilevel aspects of the selectors are ignored, i.e., $p_t = 1$. In particular, it would be very useful if satisfactory estimates of blocking could be obtained from the approximate model with $p_t = 1$. This idea was tried successfully; the details are given in Section 4.1.

One purpose of this study is to determine the effect of the number of selectors connected to a graded multiple. Thus, two extreme situations are considered: the maximum and minimum number of selectors used with each grading. The maximum is always 320 selectors. The minimum is determined by the requirement that there be at least ten selectors for each first-choice subgroup.

For the graded multiples considered below, the tables in Ref. 1 were consulted to obtain the load-loss relations presently in use. The load-loss relations for the isolated model were obtained from the results in Ref. 6.

4.1 *Comparison of the Complete and Approximate Models*

The models are compared for both a small and a large graded multiple. In Figs. 3 and 4, results are presented for a 25-trunk and 45-trunk
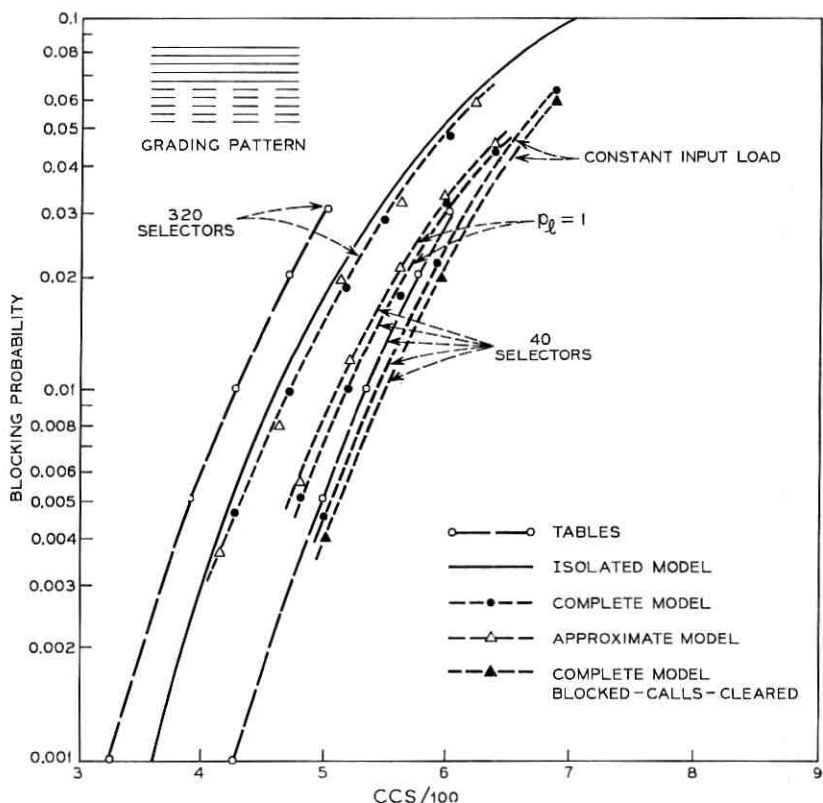
Fig. 3—Load-loss relations for a 25-trunk graded multiple.

graded multiple. Let us begin by describing how the four curves for 40 selectors in Fig. 3 (indicated by dashed lines) were obtained from the complete and approximate models: subsequently we shall comment on the significance of the curves. The top two curves, labeled $p_\ell = 1$, were generated by letting $p_\ell = 1$ in both the complete model and the approximate model. The next dashed curve was obtained by using the complete model when the load offered to the switching system was held constant at 983 CCS (27.3 erlangs) and the load offered to the grading was varied by changing p_ℓ^* (six to seven percent of the calls were delayed at the line finders). A similar curve was obtained using

---

* Changing the load to the grading by varying p_ℓ keeps the load (and the blocking) approximately constant at the line finders. This seems to be the most realistic way to generate load-loss relations for the gradings.

the approximate model but, because the curve fell ever so slightly to the right of the $p_\ell = 1$ curve, it is not shown.

The bottom curve was also obtained by holding the load offered to the system constant at 983 CCS. However, in this case the complete model was modified so that the line finders operated on a blocked-calls-cleared basis; i.e., if all line finders in a line-finder group are busy when a call arrives at the group, the blocked call is immediately cleared (about two percent of the calls were blocked at the line finders). The system actually would operate between the latter two curves because, although blocked calls are delayed, defections can occur.

Consider the cause of the disparity in the curves. Because there are ten selectors per first-choice subgroup, no calls can be blocked at the line finders when $p_\ell = 1$. Therefore, the only difference in the models
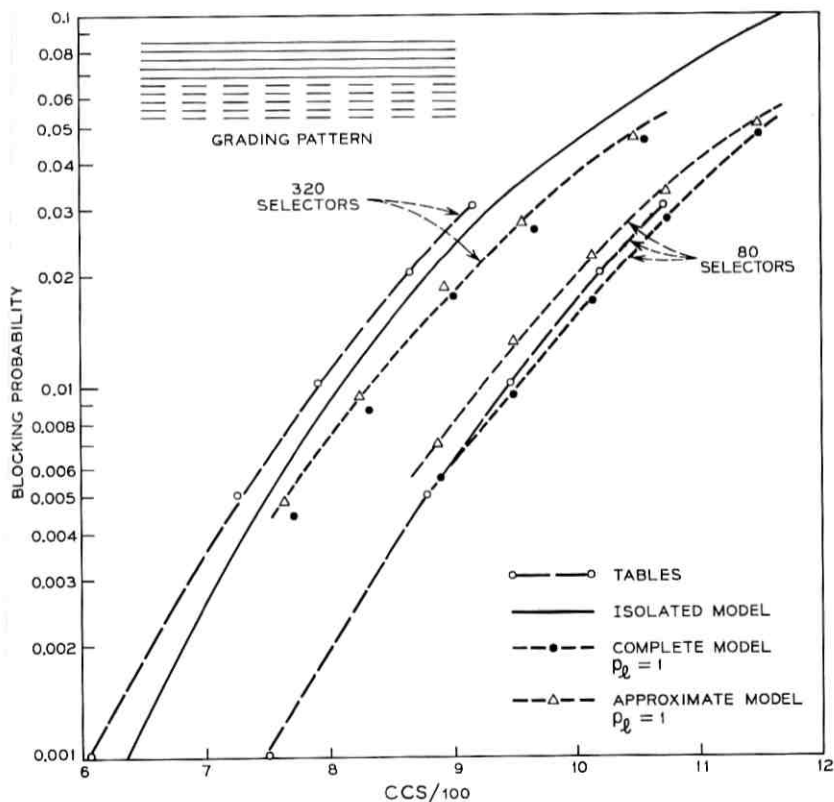

Fig. 4—Load-loss relations for a 45-trunk graded multiple.

is the representation of the inherent load-balancing, i.e., in complete detail versus our approximation. The capacity predicted by the approximate model is about 2.5 percent less than the capacity obtained from the complete model. We conclude that, in this case, the approximate model is a good characterization of the inherent load-balancing and provides a useful upper bound for the load-loss relation.

Recall that, in the approximate model, arrivals have full access to the 40 selectors. In the complete model, requests for service are directed to one of the two line-finder groups, thereby providing only partial access. It follows that, when $p_t < 1$, the traffic offered to the grading is somewhat smoother in the complete model than in the approximate model. The difference in smoothing appears to be the reason the complete model deviates from the $p_t = 1$ curve when $p_t < 1$ whereas the approximate model is relatively insensitive to changes in p_t.

Furthermore, the two curves for constant input-load differ because blocked calls receive different treatment in the blocked-calls-cleared characterization of the line finders than in the blocked-calls-delayed operation. However, in both cases of constant input-load, the blocking at the line finders is somewhat higher than Bell System design and the deviation from the $p_t = 1$ curve is, therefore, somewhat exaggerated.

For 320 selectors, observe that the complete and approximate models give almost identical results for $p_t = 1$. It is expensive to evaluate completely the effects of smoothing because, in order to achieve significant smoothing at the line finders, p_t must be small. One test was run with an input load to the switching system of 5500 CCS and $p_t = 0.12$ (only 0.3 percent of the calls were delayed at the line finders). In this case, the smoothing did not change the blocking probability from the $p_t = 1$ case.

In Fig. 4, the models are compared for 80 and 320 selectors connected to a 45-trunk grading. As above, the approximate model underestimates the capacity by about 2.5 percent for 80 selectors but, for 320 selectors, the two models are in good agreement.

We conclude that the approximate model provides a good characterization of the inherent load-balancing. For the minimum number of selectors, the approximate model underestimates the single-hour capacity by about 2.5 percent but, for larger numbers of selectors, it is in almost exact agreement with the complete model. When the originating traffic is smoothed somewhat by the line finders, the approximate model predicts blocking probabilities that are slightly higher than those obtained with the complete model. Thus, for the range of blocking at the line finders encountered in Bell System designs, the

approximate model with $p_\ell = 1$ furnishes a reasonable upper bound for the load-loss relations. Consequently, we feel secure in using the approximate model for testing the other cases of interest.

4.2 Results and Evaluation of Tables

The validity of the tables is determined in this section. The approach is to begin with the small trunk groups and progress to the larger graded multiples. In Appendix D we present analytical results which provide useful relations between several of the system parameters. The results are obtained for the approximate model with $p_\ell = 1$ under the assumption that the graded-multiple gain is independent of the gain due to the inherent load-balancing.

Using the results of Section II, we see that the inherent load-balancing occurs only when a trunk group has at least two first-choice subgroups: the line-finder to first-selector interconnections do not enhance capacity in step-by-step trunk groups of sizes one through ten trunks (analytical justification for this result is provided in Appendix D). The tables agree with this point for nine or fewer trunks (where standard Poisson* capacities are used) but not for the case of ten trunks. For ten trunks, the tables underestimate single-hour capacity by as much as ten percent for 320 selectors, are fairly accurate for 40 selectors, but, for 20 selectors, overestimate single-hour capacity by varying amounts ranging from seven percent at three percent blocking to 22 percent at 0.1 percent blocking.

Next, consider the smaller gradings with two first-choice subgroups, i.e., 11 through 19 trunks. The results in Appendix D imply that the capacity of a graded multiple of N trunks is bounded above by the capacity of a full-access group of N trunks having Poisson input. However, for 20 selectors, the tables show a capacity that, to varying degrees, exceeds the capacity of the corresponding full-access groups: the discrepancy is greatest for 11 trunks and decreases with increasing trunk-size until, for 19 trunks, the tables show a capacity that is very close to the capacity of the 19-trunk full-access group.

Figures 5 and 6 give results for gradings of 11 and 19 trunks, respectively. In each figure, the load-loss relation for the isolated viewpoint and the Erlang-B load-loss curve represent upper and lower bounds,

---

* R. I. Wilkinson[7] has shown that the blocked-calls-cleared assumption (i.e., the Erlang-B load-loss relation) gives good estimates of single-hour losses for full-access trunk groups. However, when typical day-to-day variations are included, the increased average loss is better approximated by the blocked-calls-held assumption (i.e., the Poisson load-loss relations).
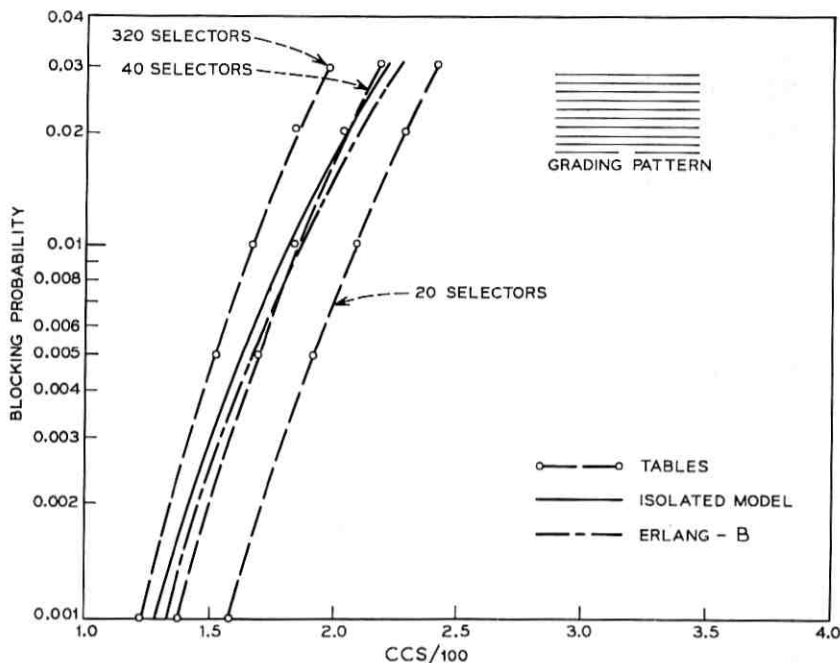
Fig. 5—Load-loss relations for the 11-trunk graded multiple.

respectively, for the actual load-loss relations. Because these bounds are close for 11 trunks, there can be little inherent load-balancing. The tables underestimate by about ten percent the single-hour capacity of the 11-trunk grading for 320 selectors, are nearly correct for 40 selectors, but overestimate by about 20 percent the single-hour capacity for 20 selectors. The results in Fig. 6 for the 19-trunk grading indicate that the tables are accurate for 20 selectors (this 19-trunk grading is only used with 20 selectors).

Now, consider some representative larger gradings. Results for 19-trunk and 25-trunk gradings each with four first-choice subgroups are given in Figs. 7 and 3, respectively. Similarly, results for 37-trunk and 45-trunk gradings each with eight first-choice subgroups are shown in Figs. 8 and 4, respectively. In each case, the tables are fairly close to the simulation results for the minimum number of selectors but, to varying degrees, are conservative for the 320-selector case. These comparisons are made on the basis of single-hour capacity.

Because the inherent load-balancing increases as the number of

first-choice subgroups increases, one might expect the capacity of gradings with 12 subgroups to be considerably higher than the tables indicate. To check this point, a grading of 45 trunks having 12 first-choice subgroups was tested (see Fig. 9). The grading is used with either 120 or 240 selectors. The simulation results indicate that the tables do underestimate capacity by about ten percent.

Because "grade of service" for the gradings in the step-by-step system means average service in the busy-season busy hours, the tables are generally used in situations where the effects of day-to-day variations are included. From Ref. 7 we observe that the effects of day-to-day variations cause a decrease in capacity of approximately seven percent. Consequently, from the above remarks concerning the single-hour capacities of the gradings, it follows that the accuracy of the tables for engineering in the presence of day-to-day variations also depends upon the size of the gradings and the number of selectors. Specifically, we find that the tables are fairly accurate when the
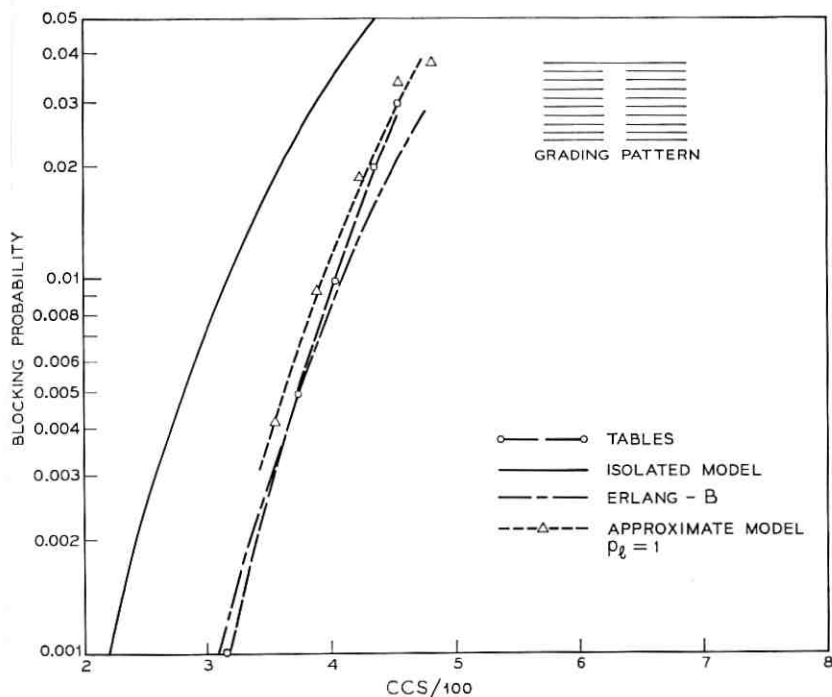
Fig. 6—Load-loss relations for a 19-trunk graded multiple with 20 selectors.
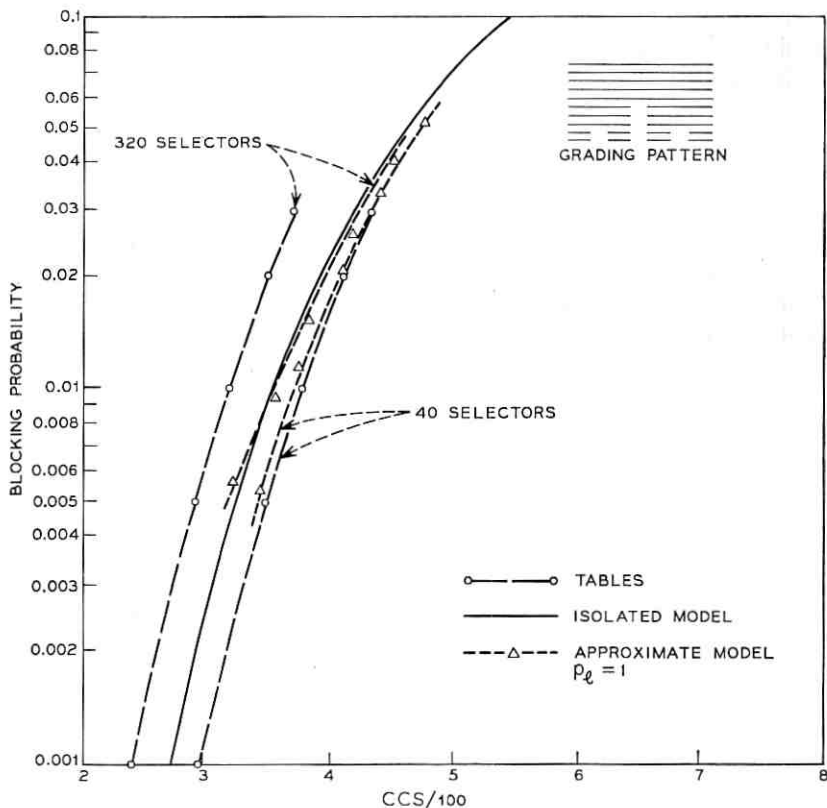
Fig. 7—Load-loss relations for a 19-trunk graded multiple.

gradings are used with the large selector groups. For the minimum number of selectors, the tables estimate fairly accurately the capacity of the larger gradings, slightly overestimate the capacity of the medium-sized gradings, and overestimate the capacity of some of the smaller gradings by about 20 percent. The tables are correct for trunk groups of one through nine trunks but, for ten trunks, they incorrectly indicate that the capacity varies with the number of selectors connected to the trunk group.

We conclude that the capacity of a step-by-step graded multiple does increase as the number of selectors connected to the grading decreases. The reason seems to be that, with fewer selectors, an imbalance in the loading of the grading results in a larger probability that a new arrival is directed to the relatively idle areas of the grading.

The simulations also indicate that the inherent load-balancing is more pronounced in the larger gradings having many first-choice subgroups. Analytical results in Appendix D imply that the inherent load-balancing does increase as the number of first-choice subgroups increases, provided the number of selectors per subgroup is held constant. The tables reflect this phenomenon too.

Notice that the simulations imply that the inherent load-balancing is more pronounced in the 45-trunk grading than the 37-trunk multiple even though both gradings have eight first-choice subgroups. In Appendix D, it is also shown that for a fixed number of first-choice subgroups and selectors, the inherent load-balancing increases as the number of trunks in these subgroups increases. Since the 37-trunk grading effectively has fewer trunks in the first-choice subgroups, the
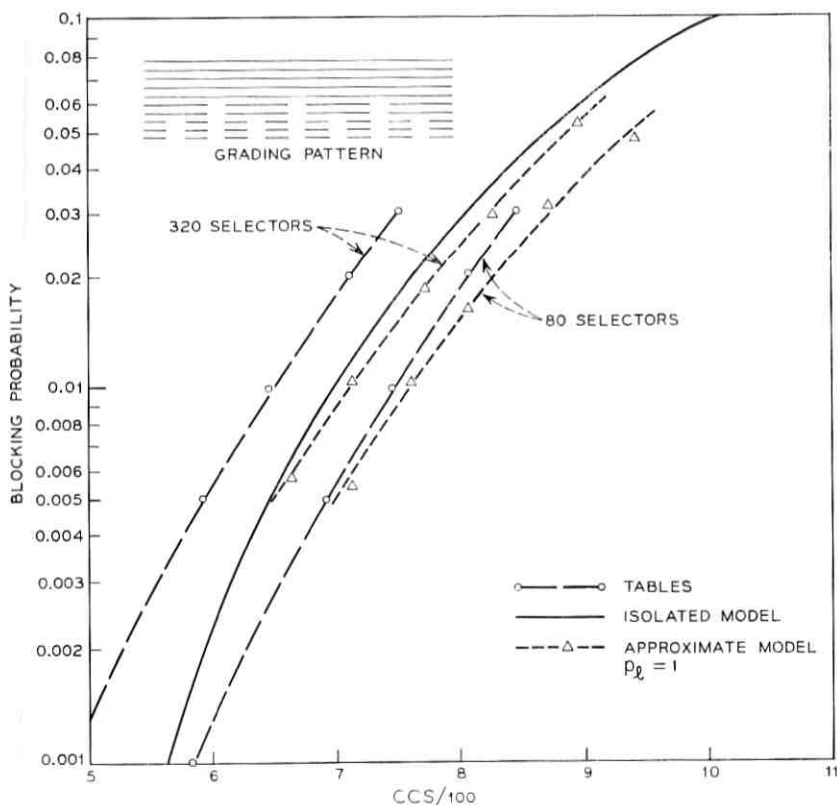
Fig. 8—Load-loss relations for a 37-trunk graded multiple.
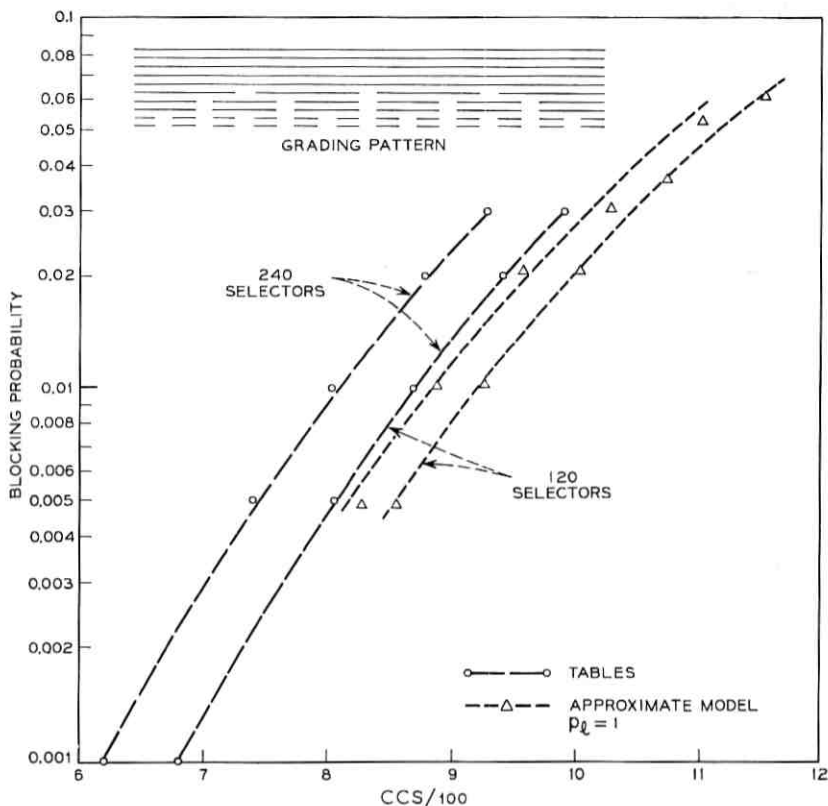
Fig. 9—Load-loss relations for a 45-trunk graded multiple.

simulation results are consistent. The tables also agree with this conclusion.

V. CONCLUSIONS

We have shown that the capacity of a graded multiple is increased when it is imbedded in the step-by-step system. The increase occurs because of the inherent load-balancing that arises from the interconnections between subscribers, line finders, and first selectors. Furthermore, the increase becomes larger as the number of selectors connected to the grading decreases.

The accuracy of the tables depends upon the size of the gradings and the number of selectors. Specifically, we find that, when used for engineering in the presence of day-to-day variations, the tables are

reasonably accurate when the gradings are used with the large selector groups. For the minimum number of selectors, the tables estimate fairly accurately the capacity of larger gradings, slightly overestimate the capacity of the medium-sized gradings, and overestimate the capacity of some of the smaller gradings by about 20 percent. The tables are correct for trunk groups of one through nine trunks but, for ten trunks, they incorrectly indicate that the capacity varies with the number of selectors connected to the trunk group.

Revised tables incorporating the improvements achievable by the methods described herein are being generated. The revised tables consist of two sections, one for single-hour engineering and one which includes the effects of day-to-day variations. The second section is of particular importance because "grade of service" for the graded multiples in the step-by-step system means average service in the busy-season busy hours.

This study considered the graded multiples that are connected to the first stage of selectors. However, the connections between selector stages are arranged so that there is also inherent load-balancing for the gradings which appear at subsequent stages. Consequently, if the smoothing effect of the preceding selector stages is negligible, the results presented herein are also valid for graded multiples at all stages in the distribution network. Thus, our results furnish upper bounds for the blocking on subsequent stages.

APPENDIX A

Traffic Flow in the Step-by-Step Switching System

We give a summary of the operation of the step-by-step switching system. An attempt is made to show how the physical design of the

switching system influences traffic flow. Because the size and layout of a step-by-step system can vary, we describe one particular configuration to illustrate the necessary points. The material in this appendix has been gathered from many sources. A large portion has been obtained from Refs. 3 and 8 and is presented without further specific reference.

A.1 *General Summary*

The step-by-step switch is used in three different modes, as: a line finder, a selector, and a connector. Each subscriber line appears on the terminal bank of several line finders. The exact number of line finders is determined by traffic considerations. When the calling party goes off-hook, one of the line finders serving the subscriber is chosen to hunt for and find the requesting line.* Each line finder is permanently connected to a first selector.† Thus, when the line finder locates the requesting line, the system is ready to receive the first dialed digit and dial tone is returned.

In response to the first digit, the first selector steps vertically to the appropriate level and then hunts horizontally across the terminals on the level for an idle trunk to the second selector stage. Upon locating an idle trunk, the first selector stops. The subscriber is thereby connected to a second selector through the line finder, the first selector, and the trunk from the first selector to the second selector. The second digit, which determines the movement of the second selector, can now be supplied by the subscriber. The remaining selector stages operate similarly with each selector switch controlled by a different digit.

After passing through the required number of selectors,‡ the call reaches the connector switch. The next to the last digit causes the connector to step vertically to the appropriate level and the last digit causes the connector to move horizontally across the level to the appropriate terminal. If the line is free, ringing current is applied. If the line is busy, busy tone is returned. During the establishment of a connection, it is possible that a selector may find all available trunks busy. When this occurs, the selector returns the reorder tone.

---

* In Section A.2, we specify which line finder is chosen.
† This applies to a basic step-by-step system not equipped with *Touch-Tone*® service or common control.
‡ Because seven digits are the standard number dialed and because such a large selection is frequently not required, some of the digits are not needed to complete the call. The step-by-step system can perform "digit absorption." Thus, the number of selector stages is variable and is determined by local considerations.

A.2 *Line Finders*

In the more recent step-by-step offices, 200-point line-finder switches* are used. Hence, the subscriber lines are divided into groups of 200 lines.† Each group of lines is served by a group of at most 20 line finders.‡

For a line finder to have access to 200 subscriber lines, each of the 100 switch positions serves two lines. A relay circuit associated with the line finder determines which subscriber line is calling. The 200 lines are divided into ten groups of 20 subscribers and are arranged on the switch banks in a slipped multiple, as shown in Fig. 10. In Fig. 10, each rectangle represents a group of 20 subscriber-lines. For example, the particular group of 20 lines represented by the shaded rectangle appears on level 5 of switch 1, level 4 of switch 2, . . . , level 1 of switch 5, level 10 of switch 6, . . . , and level 6 of switch 20. Thus, each group of 20 subscriber lines will terminate on the first level of two switches, the second level of two switches, . . . , and the tenth level of two switches. In order to minimize hunting time, the line finder chosen to service a request is the available line finder with the requesting line on the lowest level. Because each subscriber line appears on two line finders for each level, a priority is established to determine which of the two line finders is used first. In some systems, the preferred line finder is that with the lowest number; in other systems, the preference alternates on successive calls.

For our study, we model the selection rules for the line finders as follows. The 200 subscribers served by a line-finder group are subdivided into ten groups of 20 subscribers. For each subgroup, first- and second-choice line finders are specified as shown in Fig. 11 (here, the line finder with the lowest number is used first). For example, the subgroup indicated by an asterisk in Fig. 11 uses line finder 5 as the first choice and line finder 15 as the second choice. If line finders 5 and 15 are both busy, the system then searches for an idle line finder in the order 4, 14, 3, 13, 2, 12, . . . , 6, 16.

When all line finders are busy and new requests for service arrive, the system acts as a delay system with defections; i.e., the requests

---

* We restrict our attention to 200-point line finders, although 100-point line finders are used in some offices.

† Although 200-point line finders can serve 200 lines, in practice the number of working lines is not greater than 194. However, we assume that the subscribers are divided into groups of 200 both because of the resulting numerical convenience and because the assumption has a negligible effect upon our results.

‡ For minor classes of lines with high calling rates, it is possible to use a line-finder group of 30 switches.
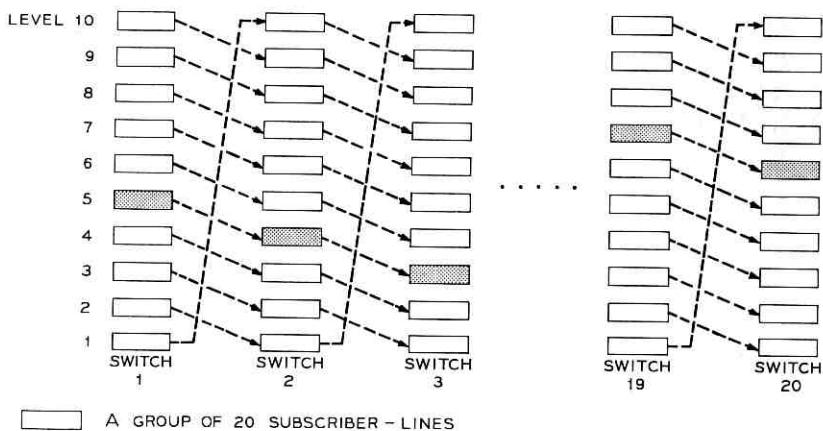
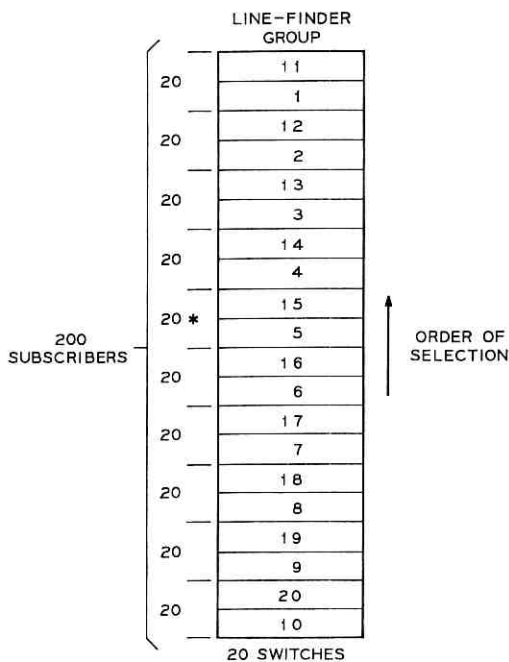Fig. 10—Slipped multiple on terminal banks of line-finder switches.

Fig. 11—Line-finder selection sequence.

join a queue and wait as long as necessary for dial tone unless the subscriber tires of waiting and hangs up. (For our study, we assume that there are no defections and no retrials.)

The order of service for requests in the queue is rather interesting. The queue is a backlog of requests for line finders to hunt for the requesting lines. Thus, when a line finder becomes free, it will immediately begin to hunt. The requesting line that is served is the one that appears lowest on the terminal bank of the hunting line finder.

A.3 *First Selectors*

The selector switches are mounted on frames with a capacity of 320 switches and are arranged on shelves with a capacity of 20 selectors per shelf. Each shelf is divided into two half-shelves of ten switches each. The corresponding terminals of all selectors on a half-shelf are permanently multipled together. It would be possible to provide a small independent trunk-group from each level of each half-shelf of first selectors to the appropriate second-selector switches. However, it is more efficient to arrange trunks in a graded multiple than in small independent trunk-groups.[9] Accordingly, the first selectors for a particular class of service are divided into groups of 20, 40, 60, 80, 120, 160, 240 or 320 switches depending upon the traffic and the number of switches in the class. On each level, the trunks from a group of first selectors to the second selector stage are arranged in a graded multiple. The graded multiple used on a particular level is determined by the amount of traffic directed to the level.

If a call is offered to a graded multiple and an idle trunk is not available, the reorder tone is returned. Thus, calls blocked at the graded multiples are cleared. Although there is actually a six- to ten-second delay, we assume that blocked calls are cleared immediately.

A.4 *Line-Finder to First-Selector Interconnections*

Each line finder is permanently connected to a unique first selector. In order to distribute the traffic evenly over the first selector stage, the links from the line finders are distributed over the first selectors in fixed patterns. The patterns attempt to connect a set of line-finder groups to a set of first-selector half-shelves such that the traffic from a line-finder group is uniformly distributed over a number of selector shelves.

Two fully equipped line-finder groups and four first-selector half-shelves are shown in Fig. 1. For clarity, the line-finder to first-selector

connection pattern is shown only for one line-finder group. In Sections III and IV, we show that the patterns and the selection rules for the line finders are extremely important in establishing the traffic capacity of the graded multiples that connect the first selectors to the second selectors.

APPENDIX B

History of the Development of the Tables

The early step-by-step systems did not use graded multiples. Consequently, corresponding analysis of trunk capacity concerned itself only with a consideration of the effect of the switching apparatus located at the input of a trunk group. Since the selectors only had ten terminals, large trunk groups were split into smaller subgroups not exceeding ten trunks each. The original mathematical efforts on this problem were carried out by E. C. Molina[10,11] in the early twenties. His results are known as the splits-in-the-multiple theory.

The original step-by-step systems used a primary-secondary switching arrangement instead of line-finder groups. Molina was aware of the importance of the various switch interconnections. However, it appears that he may not have been quite certain of their actual effect since he presented three different mathematical models for the engineers to consider. After the construction of load-loss relations for each of the three models, it was decided empirically that the (calls distributed) collectively-at-random model was most appropriate.

The principal feature of the collectively-at-random model is that the busy selectors are distributed over the group of all selectors according to a hypergeometric distribution. More precisely, assume that the outgoing trunk group is split into g groups of c trunks which are connected to s selectors; that is, there are cg trunks and sg selectors. Now, let M denote the (random) number of busy selectors in the first subgroup, and N the number in the remaining $(g - 1)c$ subgroups. Then, the collectively-at-random model assumes that

$$P\{M = m, N = n \mid M + N = m + n\} = \frac{\binom{s}{m}\binom{(g - 1)s}{n}}{\binom{gs}{m + n}}.$$

Some time later (in 1941), Wilkinson observed that the load-loss relations furnished by the splits-in-the-multiple theory were well ap-

proximated by the relations resulting from the assumption that the selectors were independent traffic sources. That is, for engineering purposes, the step-by-step system could be viewed as a system with finite-source input, the number of sources being taken as the number of selectors. Thus, the increase in trunk capacity (due to the primary-secondary to selector interconnections and explained by the splits-in-the-multiple theory) came to be known as the finite-source effect. (This terminology may have caused some of the confusion cited in the Introduction and Appendix C.)

Graded multiples were placed in use in step-by-step systems during the late twenties. The original step-by-step gradings used a maximum of 19 trunks. This number was increased to 37 in 1941, and to the present level of 45 in 1949.

A mathematical basis for engineering graded multiples was furnished by Molina some time around 1925. His graded-multiple work is still referred to as the no-holes-in-the-multiple theory. The name arises from the assumption that calls on common trunks are transferred to individuals as soon as an individual trunk becomes idle. Since such a grading would have a higher capacity than is actually achievable, Molina suggested using only one-half the predicted increase in capacity (over a full-access 10-trunk group). In 1927, a data analysis of gradings in a panel system indicated that one-half gain was a little low,[9] but the difference was not great enough to cause much concern. In 1931, Wilkinson[9] published a set of engineering curves (which he called hump-back curves) giving the increase in capacity for several classes of simple graded-multiples. These curves became the basis for engineering graded multiples in the Bell System.

It appears that the original step-by-step selector-multiple tables were constructed in 1925 by P. P. Coggins and J. R. Ferguson of the AT&T Company. These tables (for gradings up to 19 trunks) used the splits-in-the-multiple theory to estimate the increase in capacity due to primary-secondary interconnections. A graded-multiple gain was obtained from "Graded Random 'B' " curves developed by Ferguson in 1924. The results were then combined to give a total gain in capacity over a full-access ten-trunk group having Poisson input.

When the size of the gradings was increased to 37 (in 1941), Wilkinson's half-gain hump-back curves were used to obtain the graded-multiple gain. At this time, a "finite-source gain" was estimated by curves which were apparently constructed from a combination of splits-in-the-multiple theory, binomial capacity and some smoothing "based on engineering judgment."

Although one can raise several objections to the preceding approach, the results were reasonably good. Moreover, the material in Section IV indicates that the original work is still fairly good for trunk engineering in present step-by-step systems, even though line finders are now used in place of a primary-secondary switching system. The results in Appendix D show that the main reason the original work is still adequate can be traced to the fact that assumption two in the approximate model (Section 3.2) is implicitly accounted for in Molina's collectively-at-random model.

APPENDIX C

Common Incorrect Arguments

We present several common arguments that supposedly explain the capacities of the graded multiples in step-by-step systems.

c.1 *Finite-Source Effect*

From the finite-source effect viewpoint, the increase in the capacity of a grading is attributed to the finite number of selectors located before the trunk group. Since originating traffic must pass through the selectors prior to reaching a trunk group, it is reasoned that the traffic to the group will be similar to that from a finite number of independent traffic-sources.

Below, we give two statements of the argument based on the finite-source effect. The first is taken from Ref. 2, page 15; the second, from Ref. 3 (in slightly modified form).

   (i) Limited source is the constrictive effect of forcing traffic to enter a trunk group through a limited number of sources, in our case, selectors. The maximum effect of limited sources is illustrated where 10 selectors have access to 10 trunks. No matter how much traffic is offered, the trunks themselves cannot be the cause of any blocking; and their capacity is not 149 CCS (for P.01 grade-of-service using blocked-calls-held assumption) but 360 CCS, which represents 100-percent usage of each trunk. If 11 selectors have access to these 10 trunks, this maximum trunk capacity would no longer be obtained since, with 10 calls in progress through the system, the eleventh selector may offer a call resulting in a blockage. However, only one of the 11 selectors can produce this condition at any one time, hence, the

capacity, though not 360, is still much greater than 149. And so as we increase the number of selectors to exceed the number of trunks more and more, the capacity of the group of 10 trunks approaches more closely the basic unlimited source capacity of 149 CCS.

(ii) It will be found that 10 trunks with varying numbers of selectors feeding traffic thereto will carry traffic as follows: 320 selectors on 10 trunks, 149 CCS; ···; 40 selectors on 10 trunks, 164 CCS. The apparent increase in capacity as the number of selectors (traffic sources) is reduced is the gain that results from the "limited source" effect. That is, as each trunk becomes busy, the remaining sources are reduced by one, reducing to this extent the probability of all remaining trunks becoming busy.

There is considerable dissent about the accuracy of these statements. The most common counter-argument claims that the traffic offered to the trunk group does not decrease gradually as the selectors become busy but is concentrated through the idle selectors, i.e., the entire load is shifted to the idle selectors. As a result, the traffic capacity of the trunk group does not decrease gradually as the number of selectors increases from ten. Instead, the capacity should change drastically as one goes from ten to eleven selectors and, thereafter, remain approximately constant as the number of selectors increases from eleven. Of course, one is free to speculate how the finite-source effect argument might apply to a graded multiple.

Notice that both the argument and the counter-argument (and all variations on these themes) differ substantially from the correct viewpoint based on inherent load-balancing (as described in Section II). Also, the arguments are very imprecisely stated and can vary considerably depending on how the missing details are chosen. The underlying models bear little resemblance to the step-by-step system and, thus, one cannot point to a few specific defects.

c.2 *The Effect of Clipping*

Now, we present the arguments of those who believe the finite-source effect to be negligible. Consider a graded multiple on a particular selector-level. The arrivals to the grading come from half-shelves of selectors. Assuming that calls for the entire system arrive according to a Poisson process,* one might assume that the arrivals from a half-

---

* Only high-usage gradings (first-choice routes) are considered in this study.

shelf of selectors also constitute a Poisson process until all the selectors in the half-shelf become busy (due to calls on the several levels), at which time no new arrivals can occur. That is, clipping takes place, leading to the term "clipped Poisson process."

R. R. Mina[4] has examined the effect of clipping. Specifically, he considered the following situation:

(i) Exactly one selector half-shelf (ten selectors) is connected to each first-choice subgroup.

(ii) Calls arrive at each selector half-shelf according to a Poisson process. The several processes are independent with the same mean.

(iii) All arrivals go to the level under consideration (i.e., the multi-level aspect of the system is ignored).

(iv) Holding times are independent and identically distributed according to a negative-exponential distribution.

Using this model, Mina has shown (using a simulation) that grading capacity is not significantly enhanced by the clipping. Because Mina believes that his model adequately represents the step-by-step system, he further asserts [Ref. 4, page 12]:

This simulation shows that the effect of the limitation of the number of input switches is insignificant and that the increase in the traffic carrying capacity of graded trunk tables used in the U.S.A. is not realistic.

Furthermore, the structure of the AT&T tables suggests that small gradings, since they can be formed with fewer groups, are more efficient than larger gradings. This is obviously not true as it is apparent that large gradings are more efficient than small gradings.

Also [Ref. 4, page 24],

The principle of limited sources, on which the AT&T tables are based, is not realistic and underestimates the quantity of trunks.

Mina's comments would be valid if the input to a grading were adequately approximated by a clipped Poisson process. However, such a model overlooks a significant aspect of step-by-step systems: the inherent load-balancing that results from the subscriber to line-finder network and the interconnection of line finders and first selectors [see Section II].

APPENDIX D

Inherent Load-Balancing

The purpose of this appendix is to relate Molina's collectively-at-random model[10,11] to the approximate model with $p_l = 1$ presented in Section 3.2. To accomplish this task, the graded-multiple gain is assumed to be independent of the inherent load-balancing. We model the system as follows:

A service system S has $c_1 + \cdots + c_g$ servers separated into g independent groups $G_1, \cdots, G_g$. The group G_i contains c_i servers, and associated with G_i is a switch-group S_i containing s_i switches, $s_i \geqq c_i$: a switch in S_i is allotted to each customer being served in G_i.

Customers arrive at S according to a Poisson process with intensity a. If n_i denotes the number of busy servers in G_i, $i = 1, \cdots, g$, at an arrival epoch, the customer is sent to G_k with probability

$$\frac{s_k - n_k}{(s_1 + \cdots + s_g) - (n_1 + \cdots + n_g)}, \qquad k = 1, \cdots, g.$$

If a customer arrives to find all positions occupied in all switch-groups, he leaves the system and does not return. If a customer is directed to G_i when at least one of the c_i servers is available, a server is selected and service commences immediately. An arrival occurring when all c_i servers are busy leaves the system and does not return.

All service times are assumed to be independent and identically distributed according to a negative-exponential distribution having unit mean.

In order to gain some insight into the problem, the case $g = 2$ is examined first. Assume that the system is in statistical equilibrium, let N_i denote the number of busy servers in G_i at an arbitrary instant, and define

$$p(n_1, n_2) = P\{N_1 = n_1, N_2 = n_2\}.$$

The following relations determine p. For $0 \leqq n_i \leqq c_i - 1$,

$$(a + n_1 + n_2)p(n_1, n_2)$$

$$= a\frac{s_1 - n_1 + 1}{s_1 + s_2 - n_1 - n_2 + 1} p(n_1 - 1, n_2)$$

$$+ a\frac{s_2 - n_2 + 1}{s_1 + s_2 - n_1 - n_2 + 1} p(n_1, n_2 - 1)$$

$$+ (n_1 + 1)p(n_1 + 1, n_2) + (n_2 + 1)p(n_1, n_2 + 1). \qquad (1)$$

If $0 \leq n_1 \leq c_1 - 1$,

$$\left(a \frac{s_1 - n_1}{s_1 + s_2 - n_1 - c_2} + n_1 + c_2 \right) p(n_1, c_2)$$

$$= a \frac{s_1 - n_1 + 1}{s_1 + s_2 - n_1 - c_2 + 1} p(n_1 - 1, c_2)$$

$$+ a \frac{s_2 - c_2 + 1}{s_1 + s_2 - n_1 - c_2 + 1} p(n_1, c_2 - 1)$$

$$+ (n_1 + 1) p(n_1 + 1, c_2). \tag{2}$$

An expression similar to (2) can be written for $0 \leq n_2 \leq c_2 - 1$. Finally, for $n_1 = c_1$, $n_2 = c_2$,

$$(c_1 + c_2) p(c_1, c_2) = a \frac{s_1 - c_1 + 1}{s_1 + s_2 - c_1 - c_2 + 1} p(c_1 - 1, c_2)$$

$$+ a \frac{s_2 - c_2 + 1}{s_1 + s_2 - c_1 - c_2 + 1} p(c_1, c_2 - 1). \tag{3}$$

Moreover,

$$\sum_{n_1=0}^{c_1} \sum_{n_2=0}^{c_2} p(n_1, n_2) = 1. \tag{4}$$

Equations (1) through (3) can be solved by using the definition of conditional probability to write

$$p(n_1, n_2)$$

$$= P\{N_1 = n_1, N_2 = n_2 \mid N_1 + N_2 = n_1 + n_2\}$$

$$\cdot P\{N_1 + N_2 = n_1 + n_2\}.$$

Now, taking note of the way the arrivals are distributed over G_1 and G_2, one might guess that

$$P\{N_1 = n_1, N_2 = n_2 \mid N_1 + N_2 = n_1 + n_2\} = \frac{\binom{s_1}{n_1} \binom{s_2}{n_2}}{\binom{s_1 + s_2}{n_1 + n_2}}, \tag{5}$$

and try

$$P\{N_1 + N_2 = n\} = k \frac{a^n}{n!},$$

where k is an appropriate constant. Direct substitution shows that

$$p(n_1, n_2) = p(0, 0) \frac{\binom{s_1}{n_1}\binom{s_2}{n_2}}{\binom{s_1 + s_2}{n_1 + n_2}} \frac{a^{n_1 + n_2}}{(n_1 + n_2)!} \tag{6}$$

satisfies equations (1) through (3). Equations (4) and (6) can be used to determine $p(0, 0)$. Since a solution to equations (1) through (4) is unique, the system state probabilities are determined. It is interesting that the distribution (6) can be viewed as a modification of a product of two Engset distributions $k_i \binom{s_i}{n_i} a^{n_i}$, $i = 1, 2$.

If necessary, these state probabilities can be generated numerically from the relations

$$p(n_1 + 1, n_2) = \frac{a}{n_1 + 1} \frac{s_1 - n_1}{s_1 + s_2 - n_1 - n_2} p(n_1, n_2)$$

and

$$p(n_1, n_2 + 1) = \frac{a}{n_2 + 1} \frac{s_2 - n_2}{s_1 + s_2 - n_1 - n_2} p(n_1, n_2),$$

which follow from (6).

The call congestion is given by

$$B(a; c_1, c_2) = (s_1 - c_1) \sum_{n_2 = 0}^{c_2} \frac{p(c_1, n_2)}{s_1 + s_2 - c_1 - n_2}$$

$$+ (s_2 - c_2) \sum_{n_1 = 0}^{c_1} \frac{p(n_1, c_2)}{s_1 + s_2 - n_1 - c_2}. \tag{7}$$

Perhaps the most important aspect of this model is the appearance of the hypergeometric distribution shown in equation (5). As noted in Appendix B, Molina's collectively-at-random model was characterized by this distribution. Consequently, it is no longer surprising that the tables (based on Molina's model) adequately reflect the inherent load-balancing in most of those cases tested.

There are two limiting cases of interest; namely, $s_i \to \infty$ and $s_i \to c_i$. Setting $s_1 = s_2 = s$ yields $(s_i - n_i)/(s_1 + s_2 - n_1 - n_2) \to 1/2$ as $s \to \infty$. Hence, when the number s of selectors is very large, the system operates as two independent systems with c_i servers and Poisson input of intensity $a/2$.

Equation (7) shows that there is no blockage at either service area when $s_i = c_i$, $i = 1, 2$ (since all blockage takes place at the switch

groups). However, the actual capacity of the system S (consisting of switch groups and servers) is identical to that obtained from a full-access group of $c_1 + c_2$ servers.

Thus, when the number s_i of switching positions (selectors) exceeds c_i, the capacity of the system is bounded above by that of a full-access group of $c_1 + c_2$ servers, and below by that of two independent groups of c_1 and c_2 servers respectively; that is the system capacity increases as the number of switches decreases.

The preceding results can be extended in a straightforward fashion to any number g of service groups. Letting N_i denote the number of busy servers in G_i, $i = 1, \cdots, g$, it can be shown that

$$p(n_1, \cdots, n_g) = P\{N_i = n_i, i = 1, \cdots, g\}$$

$$= \frac{\binom{s_1}{n_1} \cdots \binom{s_g}{n_g}}{\binom{s_1 + \cdots + s_g}{n_1 + \cdots + n_g}} \frac{a^{n_1 + \cdots + n_g}}{(n_1 + \cdots + n_g)!} p(0, \cdots, 0).$$

The system capacity is bounded below by the capacity of g independent groups $G_1, \cdots, G_g$ of servers having $c_1, \cdots, c_g$ servers, respectively, and is bounded above by the capacity of a full-access group of $c_1 + \cdots + c_g$ servers.

Applying this last result to the step-by-step graded multiples, it follows that the inherent load-balancing is more pronounced in the graded multiples with the larger number of first-choice subgroups. Also, for gradings with a fixed number (at least two) of first-choice subgroups, the inherent load-balancing increases as the numbers of trunks in these subgroups increase. (It is evident that there is no inherent load-balancing in full-access trunk groups.) These phenomena were observed in the simulations described in Section IV.

It is not particularly difficult to extend the model treated above to include the multilevel aspects of the step-by-step selectors. However, there is no indication that such results would furnish additional insight into the problem.

REFERENCES*

1. "Traffic Facilities Practices," Division D, Section 4h, American Telephone and Telegraph Co., New York, N. Y., May 1951.

---

* Since the history of our subject is not reflected in the public record, only four of the references are in the open literature. Copies of the other references, while not generally available, are in our possession, and arrangements can be made for investigators in the teletraffic field to examine them.

2. "Interdepartmental Dial Equipment Management; Step-by-Step," Division V, Section B, New York Telephone Co., March 1968.
3. Dustin, G. E., "Step-by-Step Systems—Interswitch Trunking Arrangements and Associated Traffic Engineering Considerations," unpublished work, Bell Telephone Laboratories, 1961.
4. Mina, R. R., "Gradings," lecture given at Advanced Telephone Traffic Engineering Conference, Michigan State University, July 31, 1968.
5. Wilkinson, R. I., "Theories for Toll Traffic Engineering in the U.S.A.," B.S.T.J., *35*, No. 2 (March 1956), pp. 421–514.
6. Neal, S. R., "An Extension of the Equivalent Random Theory with Application to the Analysis of Graded Multiples Having Nonrandom Offered Traffic," unpublished work, Bell Telephone Laboratories, July 14, 1969.
7. Wilkinson, R. I., "Some Recent Developments in Trunking Theory," unpublished work, Bell Telephone Laboratories, March 9, 1965.
8. "Switching Systems," Manual for Course B-328, Communications Development Training Program, Bell Telephone Laboratories, 1964.
9. Wilkinson, R. I., "The Interconnection of Telephone Systems—Graded Multiples," B.S.T.J., *10*, No. 4 (October 1931), pp. 531–564.
10. Molina, E. C., "The Theory of Probabilities Applied to Telephone Trunking Problems," B.S.T.J., *1*, No. 2 (November 1922), pp. 69–81.
11. Coggins, P. P., "Theory of Probability Applied to Telephone Problems," New York Telephone Co., Out-of-Hours Course Notes, 1927–1928.

New Results on Avalanche Multiplication Statistics with Applications to Optical Detection

By S. D. PERSONICK

(Manuscript received July 21, 1970)

In this paper, we derive statistics of the random gain of two types of avalanche diode optical detectors. A simple optical binary receiver which could employ these devices is analyzed. In particular, we determine the moment generating function of the random gain probability density for a diode with equal hole and electron collision probabilities and for a diode with unilateral gain. For the unilateral gain case, we invert the moment generating function to obtain the probability density which turns out to be a shifted Bose Einstein density. Using the Chernoff bound, we analyze the performance of a simple binary receiver using the above devices. In addition, we exactly analyze a receiver with a deterministic gain device. We upper bound the degradation incurred from the use of a random gain rather than a deterministic gain. For the devices above, the degradation can be as small as a dB or less in certain ranges of parameter values discussed in the text.

I. INTRODUCTION

Practical optical direct detection receivers can employ detectors with internal gain (more than one output electron on the average per optically or thermally generated primary electron) to overcome thermal noise in amplifying stages following the detector. Since the gain is a random variable, its statistics affect the system performance. In certain systems (e.g., linear analog intensity modulation with linear processing and a mean square error risk criterion) it is sufficient to know the mean and variance of the random gain to determine performance. Various authors[1,2] have calculated those statistics for a variety of avalanche multiplier models. In digital systems with a probability of error risk, one needs to know the probability density of the gain to evaluate performance.

Up to now, little work has been published on these statistics.* In this paper, we derive the moment generating function (and in one case the density) of the multiplication statistics for two special cases of avalanche diode detectors. We apply results to performance evaluation for a simple binary optical receiver.

II. A REVIEW OF THE AVALANCHE MULTIPLICATION PROCESS

To understand the results which follow, we briefly review some results which have appeared in the literature.

We can model an avalanche diode as follows.[1] Incident light or thermal agitation causes the generation of a hole-electron pair in a portion of the semiconductor bulk of the device. An electric field across this bulk not necessarily uniform in strength causes the carriers to drift in appropriate directions. In certain regions of the bulk where the field is sufficiently high, carriers of one type or the other pick up enough energy to suffer ionizing collisions which result in the generation of additional hole-electron pairs. Each of these secondaries may of course generate additional secondaries. We wish to find some statistics of the total number of secondaries that result from a given primary (initially generated) pair.

To solve this problem, we must pick the right bookkeeping scheme. That is, after making more detailed assumptions about the above process, we must find a scheme which will lead to an algebraically tractable way of keeping track of all of the secondaries. We shall solve for some important statistics of two particular types of avalanche multipliers in the next section.

First, for both of these types, we shall make the following assumption. We require that the field strength and bulk characteristics be such that with high probability the interval (in time or distance as you wish) between ionizing collisions of a given carrier is sufficiently large that previous collisions do not appreciably affect the statistics of a current collision. That is, we assume that we can model all ionizing collisions of a given carrier on its way through the high field region of the bulk as independent. We shall assume that each carrier type (hole and electron) has an associated probability per unit length of suffering an ionizing collision. These probabilities may be functions of position if the field in the bulk is not uniform. As a result, the

---

* To the best of the author's knowledge, the only other results on the complete statistics are due to R. J. McIntyre[1] and were derived independently of these results. At the writing of this manuscript they are not yet published.

number of secondaries directly generated by a given carrier (not including secondaries generated by these secondaries) is Poisson distributed with mean equal to the integral of the generation probability per unit length over the path travelled by that carrier.

We now study two important special cases of the above. By special case it is meant that assumptions on each carrier will be made about the ionization probability per unit length.

Case 1 will correspond to a diode in which holes and electrons have equal collision ionization probabilities per unit length. In such a diode, we have a feedback situation where carriers travelling in one direction create carriers through collision travelling in the opposite direction which in turn create carriers travelling in the original direction, etc. For such a case, we shall see that the mean number of pairs directly resulting from any given pair, which is a function of the applied voltage, must be less than one in order to have a stable (non-infinite) multiplication. Case 2 will correspond to a diode where only one type of carrier causes collision generation of new pairs. Such a diode is said to have unilateral gain.

III. CASE 1: THE EQUAL IONIZATION PROBABILITY DIODE

Suppose we assume that the ionization probabilities for holes and electrons (at a given velocity) are equal. We also assume that the shape of the high field region of the bulk is such that its length parallel to the field is constant. As a result no matter where in the bulk a hole-electron pair is created the sum of the distances in the high field region travelled by the two carriers is the same (since they travel in opposite directions). More important, the total number of secondaries directly generated by the two carriers is Poisson distributed with mean equal to the integral of the ionization probability per unit length over the total distance in the high field region parallel to the field travelled by both carriers, which is constant (see Fig. 1). The fact that the statistics of the secondaries directly generated by a given hole-electron pair is independent of where that pair itself is generated is the bookkeeping aid we shall next exploit. We discuss a type of random multiplier which is a generalized case of the above multiplication process.

IV. A RANDOM MULTIPLIER

Consider the following random multiplier. Into the device we send an initial "count" (which for the avalanche diode corresponds to a
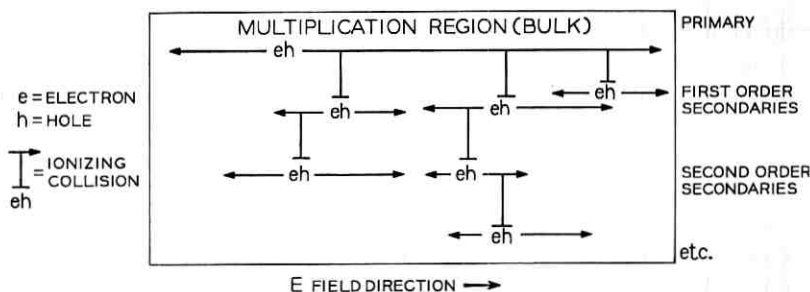
Fig. 1—Model of equal ionization device.

hole-electron pair). This count generates "directly" a random number of first-order secondary counts. What we mean by "directly" shall be clear presently. Each first-order secondary count independently generates a random number of second-order secondary counts. Analogously, each secondary of order j (if any of this order are created) generates a random number of $(j + 1)^{\text{th}}$ order secondaries. For any order j, the probability density governing the number k of $(j + 1)$th order secondaries generated by that count is $f(k)$. We emphasize that an important assumption in the ensuing analysis is the fact that a secondary of any order generates higher-order secondaries independently of what order it is and how many other counts of other orders there are.

We are interested in determining the statistics of the total number of counts of all orders (we shall call the input count an order zero secondary for convenience). Our procedure will be as follows. We shall find the statistics of the sum of the counts up to order, say, n in terms of the statistics of the sum of counts up to order $n - 1$. Since we know the statistics of the sum of the counts up to order zero (namely there is one count) and order 1, we use induction to get an expression for the statistics of the total sum of all counts of all orders.

Call the number of counts of order j the random variable S_j. Remember that these counts are randomly generated by all counts of order $j - 1$. Now call the sum of the S_j from zero to n inclusive the random variable G_n. We seek the statistics of the random variable G_∞ which is the limit of the G_n as n goes to infinity and which is assumed to exist. (We shall soon see that for G_∞ to have finite mean, the average number of direct secondaries per primary must be less than 1.)

As mentioned, we wish to determine some statistics defined upon

the partial sums G_n. For reasons which will soon be clear, we define $f_{n,\,n-1}(x, y)$ as the discrete joint density on G_n and G_{n-1} (i.e., the probability that simultaneously G_n equals x and G_{n-1} equals y). Now we ask the question: How can we have $G_n = x$ given $G_{n-1} = y$ assuming x greater than y? Well, by definition S_n must equal the difference $x - y$. That is, the number of nth-order secondaries must be $x - y$. How can this happen? We know $G_{n-1} = y$. Suppose $G_{n-2} = z$ of course less than y (or equal). Therefore $S_{n-1} = y - z$. It is these secondaries (i.e., S_{n-1}) which must contribute $x - y$ nth-order secondaries. The probability that $S_{n-1} = u$ and that $G_{n-1} = y$ is just $f_{n-1,\,n-2}(y, y - u)$. The probability that the number of nth-order secondaries created by these $u(n - 1)$th-order secondaries is $x - y$ is the probability that u independent random variables whose distributions are all $f(k)$ add up to $x - y$. This is simply $f^{*u}(x - y)$ where $f^{*u}(x - y)$ is the convolution of $f(z)$ with itself u times, the result evaluated at $z = (x - y)$. (A well-known result about sums of independent random variables.) Thus the probability that G_n equals x and $G_{n-1} = y$ is equal to the probability that $G_{n-1} = y$ and $G_{n-2} = y - u$ and that u secondaries of order $n - 1$ result in $x - y$ secondaries of order n—the preceding averaged over all possible values of u. In symbols

$$f_{n,n-1}(x, y) = \sum_z f^{*(y-z)}(x - y)f_{n-1,n-2}(y, z) \tag{1}$$

where $f^{*(x)}$ is the convolution of f with itself x times.

The next steps we shall take are not intuitive. They come from experience with dealing with equations like equation (1).

First to change convolutions to products, Laplace transform equation (1) on both variables x and y using the corresponding transform variables s and t.

$$M_{k,k-1}(s, t) = \sum_{x,y} f_{n,n-1}(x, y) \exp(sx + ty) = \sum_x \sum_y \sum_z f^{*(y-z)}(x - y) \;,$$

$$\cdot f_{n-1,n-2}(y, z) \exp(sx + ty)$$

$$= \sum_y \sum_z \exp[(\psi_0(s))(y - z)] \exp(sy) \exp(ty) f_{n-1,n-2}(y, z),$$

$$= M_{n-1,n-2}(\psi_0(s) + s + t, -\psi_0(s)), \tag{2}$$

where $\psi_0(s)$ is given by*

$$\psi_0(s) = \ln \sum_{k=0}^{\infty} f(k) \exp(sk). \tag{2a}$$

---

* ln $x =$ natural logarithm of x.

Lemma: The iterative relationship in equation (2) implies the iterative relation

$$M_{n+1,n}(s, t) = \exp [s + t + \psi_0(\ln (M_{n,n-1}(s, t)))]. \qquad (3)$$

Proof: Using induction. We know the joint moment generating function of G_1 and G_0 since $G_0 = 1$ with probability 1 and $G_1 = 1 + $ a random variable with density $f(k)$.

$$M_{1,0}(s, t) = \exp [s + t + \psi_0(s)]. \qquad (4)$$

Using equation (2) we obtain

$$
\begin{aligned}
M_{2,1}(s, t) &= \exp [\psi_0(s) + s + t - \psi_0(s) + \psi_0(s + t + \psi_0(s))] \\
&= \exp [s + t + \psi_0(s + t + \psi_0(s))]
\end{aligned}
\qquad (5)
$$

which is also the result predicted by equation (3).

Now assume equation (3) holds for some arbitrary n. Then using equation (2), we obtain

$$
\begin{aligned}
M_{n+2,n+1}(s, t) &= M_{n+1,n}(\psi_0(s) + s + t, - \psi_0(s)) \\
&= \exp [s + t + \psi_0(s) - \psi_0(s) \\
&\quad + \psi_0(\ln (M_{n,n-1}(s + t + \psi_0(s), - \psi_0(s))))]
\end{aligned}
\qquad (6)
$$

which using our result (2) immediately yields

$$M_{n+2,n+1}(s, t) = \exp [s + t + \psi_0(\ln (M_{n+1,n}(s, t)))]. \qquad (7)$$

Thus if equation (3) holds for $n = j$, it holds for $n = j + 1$. Furthermore equation (3) holds for $n = 1$. Thus the lemma is proven true.

If we set $t = 0$, then we have the moment generating function of G_n in terms of the moment generating function of G_{n-1}.

$$
\begin{aligned}
p_n(m) &= \text{density of } G_n , \\
M_n(s) &= \Sigma p_n(m) \, exp \, (sm) = \Sigma f_{n,n-1}(m, r) \exp (sm) \exp (0r), \\
&= M_{n,n-1}(s, 0).
\end{aligned}
\qquad (8)
$$

Thus

$$
\begin{aligned}
M_n(s) &= M_{n,n-1}(s, 0) = \exp [s + \psi_0(\ln (M_{n-1,n-2}(s, 0)))], \\
&= \exp [s + \psi_0(\ln (M_{n-1}(s)))].
\end{aligned}
\qquad (9)
$$

Now assume that as order gets higher and higher, the probability of a secondary of that order goes down fast enough so that $M_k(s)$ con-

verges (i.e., we get a finite number of electron-hole pairs out in response to the initial pair). Then we obtain

$$M_\infty(s) = \exp [s + \psi_0(\ln (M_\infty(s)))].$$ (10)

We obtain the differential equation*

$$M'_\infty(s) = M_\infty(s)\left[1 + \psi'_0(\ln (M_\infty(s))) \frac{M'_\infty(s)}{M_\infty(s)}\right],$$

$$= \frac{M_\infty(s)}{1 - \psi'_0(\ln (M_\infty(s)))} ; \qquad M_\infty(0) = 1.$$ (11)

In principle we can solve for $M_\infty(s)$ using equation (11) and we can take the inverse Laplace transform to find $p_\infty(m)$ the probability of the sum of all the secondaries. In addition we can solve for all the moments of G_∞ without inverse transforming by repeated differentiation of equation (11) as will be shown in the next section.

V. APPLICATION TO AVALANCHE DIODE

Getting back to the avalanche diode with equal carrier collision ionization probabilities, it is clear that each secondary (pair) created in the multiplication process discussed in Section III produces higher-order secondaries as discussed in Section IV with $f(k)$ a Poisson density with mean equal to the integral of the ionization probability per unit length over the path of an electron pair in the high field region of the bulk. That is

$$f(k) = Z^k \exp (-Z)/k!$$ (12)

where Z is the mean of the density $f(k)$.

From equation (2a) we obtain

$$\psi_0(s) = Z[\exp (s) - 1].$$ (13)

Plugging into equation (11), we get

$$M'_\infty(s) = \frac{M_\infty(s)}{1 - ZM_\infty(s)}.$$ (14)

If we call the total number of pairs in the response G (for Gain) we have (G is a random variable equal to the total number of electron-

---

* Throughout the text $x'(\cdot)$ means differentiation with respect to the total argument in parenthesis.

hole pairs at the output resulting from an input pair)

$$\bar{G} = E(G) = M'_\infty(s)\Big|_{s=0} = (1 - Z)^{-1} = \text{mean secondary pairs in response.}$$

$$\overline{G^2} = E(G^2) = M''_\infty(s)\Big|_{s=0} = \frac{M'_\infty(s)}{1 - ZM_\infty(s)}\Big|_{s=0} + \frac{ZM_\infty(s)}{(1 - ZM_\infty(s))^2}\Big|_{s=0}$$

$$= \left(\frac{1}{1-Z}\right)^2 + \frac{Z}{(1-Z)^3} = \left(\frac{1}{1-Z}\right)^3$$

$$= \text{mean square number of output pairs.} \tag{15}$$

This corresponds to results by previous authors[1,2] which were limited to calculating these first two moments. We can calculate all order moments, and in addition we can numerically integrate equation (14) to obtain $M_\infty(s)$.

VI. CASE 2: SINGLE CARRIER AVALANCHE DIODE

Returning to the discussion of Section III, suppose that only one type of carrier causes ionizing collisions. Further, assume that the probability per unit length of an ionizing collision as a carrier of that type travels through the high field region of the bulk is constant-β. Also assume that the initial carrier of the ionizing type starting the multiplication process will travel a total distance W on its way through the multiplication region. We shall now set up bookkeeping.* Divide the multiplication region into increments of length Δ. Model the multiplication process as follows (see Fig. 2).

Let $p_k(n)$ be the probability density of the number of carriers n present at the beginning of the kth increment when the carriers pass that point. Assume each of these carriers generates a Poisson distributed number of new carriers in the kth increment with mean $\beta\Delta$. The total number of carriers at the beginning of the $(k + 1)$th increment given N are present at the beginning of the kth increment is N plus a random variable whose density is the above Poisson density convolved with itself N times. Therefore we have (by argument similar to those in Section IV)

$$p_{k+1}(n) = \sum_j p_k(j)P_\Delta^{*i}(n - j) \tag{16}$$

---

* The reader is cautioned that our bookkeeping scheme for case 1 is quite different from that for case 2.
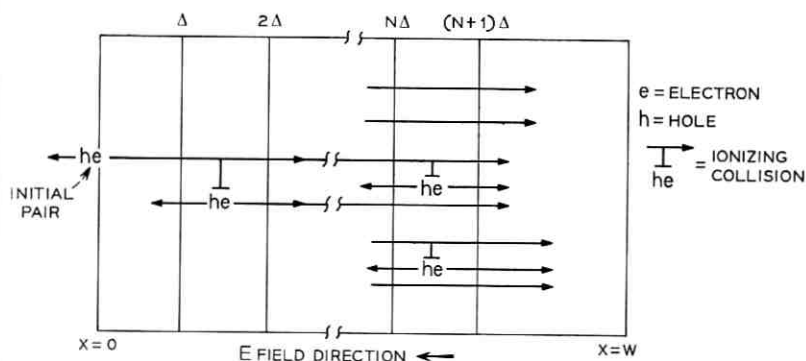
Fig. 2—Model of unilateral gain device.

where $P_\Delta^{*i}(n)$ is the j-fold convolution of

$$P_\Delta(n) = \frac{(\beta\Delta)^n \exp(-\beta\Delta)}{n!}$$

the result evaluated at the argument n. Taking the Laplace transform of equation (16) we obtain

$$M_{k+1}(s) = \sum_n p_{k+1}(n) \exp(sn) = \sum_i p_k(j) M_\Delta^i(s) \exp(sj)$$

$$= M_k(\psi_\Delta(s) + s)$$

(17)

where

$$M_\Delta(s) = \sum_n \frac{(\beta\Delta)^n}{n!} \exp(-\beta\Delta) \exp(sn) = \exp[(\beta\Delta)(\exp(s) - 1)]$$

and

$$\psi_\Delta(s) = \beta\Delta(\exp(s) - 1).$$

Lemma: The solution of equation (17) satisfies the iterative relation

$$M_k(s) = M_{k-1}(s) \exp(\beta\Delta[M_{k-1}(s) - 1]).$$

(18)

Proof: At the input, we have one count, therefore

$$M_1(s) = \exp(s).$$

By equation (17) we have

$$M_2(s) = \exp(\beta\Delta(\exp(s) - 1) + s).$$

(19)

But equation (19) is the result predicted by equation (18). Thus (18)

is true for $k = 1$. Next, assume equation (18) is true for $k = \ell$, by (18)

$$M_\ell(s) = M_{\ell-1}(s) \exp (\beta\Delta[M_{\ell-1}(s) - 1]). \tag{20}$$

Using equation (17), we obtain

$$M_{\ell+1}(s) = M_{\ell-1}(\psi_\Delta(s) + s) \exp (\beta\Delta[M_{\ell-1}(\psi_\Delta(s) + s) - 1])$$

$$= M_\ell(s) \exp (\beta\Delta[M_\ell(s) - 1]).$$

Thus by induction, the lemma is true.

We now wish to let Δ approach zero. Defining

$$\psi_k(s) = \ln M_k(s) \tag{21}$$

we obtain

$$\psi_k(s) = \psi_{k-1}(s) + \beta\Delta[\exp (\psi_{k-1}(s)) - 1].$$

Taking the limit as $\Delta \to 0$ (Δ is the width of the increment in position x along the bulk)

$$\frac{\partial}{\partial x} \psi(x, s) = \beta[\exp (\psi(x, s)) - 1]. \tag{22}$$

(Where x replaces $k\Delta$ as the variable of position in the multiplication region.) And where

$$\psi(0, s) = s$$

(i.e., $M(0, s) = \exp(s)$ since there is one pair at $x = 0$). Solving equation (22) we obtain

$$M(x, s) = \frac{1}{1 - \exp (\beta x)(1 - \exp (-s))} \tag{23}$$

which can be checked by differentiation.

If we evaluate equation (23) at $x = W$, we obtain

$$M(W, s) = \frac{1}{1 - \bar{G}(1 - \exp (-s))} \tag{24}$$

where $\bar{G} = \exp (\beta W) = (\partial/\partial s)M(W, s) \mid_{s=0}$ = mean multiplication = average number of pairs at the output in response to an initial pair.

The corresponding probability density of the total carriers leaving the multiplication region is the inverse Laplace transform of $M(W, s)$ as given by equation (24).

$$P_W(n) = \frac{1}{\bar{G}} \left(\frac{\bar{G} - 1}{\bar{G}}\right)^{n-1} \qquad n = 1, 2, 3, \cdots . \tag{25}$$

VII. APPLICATION TO OPTICAL COMMUNICATION

In the following sections, we shall obtain bounds to the received energy per pulse required to achieve a desired error rate in a simple optical binary receiver. That is, there is some optimal adjustment of system parameters which requires a minimum of signal energy to achieve exactly a desired error rate. We cannot exactly calculate the minimum energy but we can find a range which we prove that energy falls into. We call upon the Chernoff bound[3] to upper bound the energy requirements. We obtain a lower bound by studying a receiver which we prove performs better than the receiver we wish to obtain the bounds for, but which is easy to analyze. Hopefully, the upper and lower bounds will be close enough to each other to be useful for determining the actual energy requirements. The assumed receiver is shown in Fig. 3.

VIII. THE CHERNOFF BOUND

The Chernoff bound is useful for studying the tails of distributions. It is simply derived as follows.

Let $p(x)$ be a probability density. Suppose we wish to know $Q(\gamma)$ defined as

$$Q(\gamma) = \int_{\gamma}^{\infty} p(x) \, dx. \tag{26}$$

Let s be a real number. Clearly

$$\frac{\exp (sx)}{\exp (s\gamma)} > 1 \qquad \text{for} \qquad s > 0, x > \gamma. \tag{27}$$

Therefore we have

$$Q(\gamma) \leqq \int_{\gamma}^{\infty} \frac{\exp (sx)}{\exp (s\gamma)} p(x) \, dx \leqq \int_{-\infty}^{\infty} \frac{\exp (sx)}{\exp (s\gamma)} p(x) \, dx \qquad \text{for} \quad s > 0. \tag{28}$$
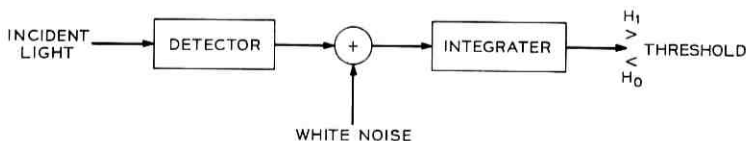

Fig. 3—Receiver.

Define the moment generating function of x, $M_x(s)$ as

$$M_x(s) = \int_{-\infty}^{\infty} \exp{(sx)}p(x) \, dx = \exp{(\psi_x(s))}, \tag{29}$$

where $\psi_x(s)$ is a natural log of $M_x(s)$ and is called the semi-invariant moment generating function (S.I.M.G.F.) of the random variable x.

Examining the bound of equation (28), we can minimize the right side (obtain the tightest bound) by choosing an optimal s within the constraint region. Differentiating we obtain

$$\psi_x'(s) = \gamma \text{ as the optimal } s \text{ provided } s > 0 \tag{30}$$

(prime denotes derivative). It can be shown that a unique s satisfying equation (30) exists and is greater than zero if γ is greater than the mean of the distribution of the random variable x. This situation will always be satisfied for cases of interest here. We obtain the bound

$$Q(\gamma) \leqq \exp{(\psi_x(s) - s\psi_x'(s))} \big|_{\psi'_x(s) = \gamma} . \tag{31}$$

Similarly, if we define

$$P(\gamma) = \int_{-\infty}^{\gamma} p(x) \, dx \tag{32}$$

where γ is less than the mean of x, then we obtain the bound

$$P(\gamma) \leqq \exp{(\psi_x(s) - s\psi_x'(s))} \big|_{\psi'_x(s) = \gamma} \tag{33}$$

(where s will turn out to be negative).

Equations (31) and (33) constitute the Chernoff bounds.

IX. APPLICATION OF THE CHERNOFF BOUND TO A SIMPLE ADDITIVE GAUSSIAN NOISE PROBLEM

Suppose an observer must distinguish between two hypotheses $H(1)$ and $H(0)$ based upon his received value of a random variable X. He desires to minimize his probability of error. For the case of interest here the best (minimal error test) technique simply compares the received value of the random variable to a threshold—and a decision $H(1)$ or $H(0)$ is made.

Define p_m, the miss probability, as the probability that the decision is $H(0)$ when $H(1)$ was actually the true hypothesis. Define p_f, the false alarm probability, as the probability that the decision is $H(1)$ when $H(0)$ is the true hypothesis. For a communication system where $H(1)$ and $H(0)$ are *a priori* equally probable, the error probability p_e.

equals $\frac{1}{2}$ the miss probability p_m plus $\frac{1}{2}$ the false alarm probability p_f . (An error occurs if either a miss or a false alarm occurs.) In this paper we shall assume that we desire $p_m \cong p_f \cong p_e$.

Suppose we have chosen a threshold, and we decide $H(1)$ if the received random variable is above threshold and $H(0)$ if the received random variable is below threshold. Let $p_1(x)$ be the probability density of the received random variable X given $H(1)$ is true. Let $p_0(x)$ be the density given $H(0)$ is true. Then clearly if the threshold is γ, we have

$$p_m = \int_{-\infty}^{\gamma} p_1(x)\, dx,$$

$$p_f = \int_{\gamma}^{\infty} p_0(x)\, dx.$$

(34)

We can bound p_m and p_f using the Chernoff bound if we know the semi-invariant moment generating function of the received statistic X on both hypotheses.

For cases of interest here, X will consist of the sum of two independent random variables. There will be a gaussian random variable of zero mean and known variance σ^2 representing thermal noise contributions. This gaussian component will be independent of which hypothesis is true. In addition there will be added a random variable Y, whose statistics depend upon the hypothesis. From the definition in equation (29) it is easy to see that the semi-invariant moment generating function of the sum of two independent random variables is simply the sum of the two separate semi-invariant moment generating functions. The S.I.M.G.F. of a gaussian zero mean random variable is $s^2\sigma^2/2$. Thus to apply the Chernoff bound, we must know the S.I.M.G.F. of Y under both hypothesis. As we shall see, we can use equation (34) to pick the threshold and the required received energy to achieve at least a desired performance.

X. DETAILS OF THE RECEIVER

Figure 3 is a block diagram of the system. Figure 4 schematically depicts the system in somewhat more detail. The current $i_s(t)$ is generated when incident light is present. The presence of incident light is from now on called hypothesis one. If the incident light is from a coherent or highly incoherent source (the required bandwidth depends upon the intensity), then $i_s(t)$ consists of electrons arriving as in a Poisson process with rate $\eta P/hf$ per second. η is the quantum efficiency
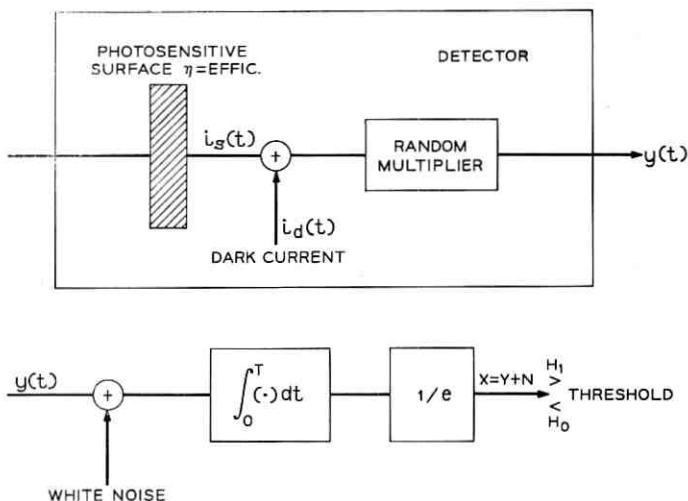
Fig. 4—Schematic.

of the light sensitive surface; P is the incident light intensity and hf is Planck's constant multiplied by the incident light center frequency. If the light is partially coherent, then the statistics of the current $i_s(t)$ are considerably more complicated. We shall not study this case here, but shall assume the Poisson statistics for both incoherent and coherent light sources.

The current $i_d(t)$ arises when electrons are spontaneously emitted in the detection device due to thermal effects. In the literature it is usually called dark current. It shall be modelled as consisting of electrons arriving as a Poisson process with rate λ_d.

The devices we are interested in employ internal current gain. That is, every electron in the processes $i_s(t)$ and $i_d(t)$ enters a random multiplication mechanism whereby it generates a random number of secondary electrons. We shall ignore the time dispersion of the secondaries around the arrival time of a primary. We are only interested here in the statistics of the number of secondaries per primary. (The electrons in $i_s(t)$ and $i_d(t)$ make up the primary process.)

We call the current at the output of the random multiplier $y(t)$. To $y(t)$ we add white gaussian noise representing the thermal noises of the receiver following the detector. The sum is integrated over the pulse duration and the resulting random variable normalized by the electron charge, e, is compared to a decision threshold. The random variable X consists of a gaussian random variable whose mean is

zero and whose variance is the thermal noise spectral height N_0 times the pulse duration T^* and an added random variable Y dependent upon the hypothesis, equal to the number of secondary electrons leaving the random multiplier during the pulse duration interval. As discussed in Section III, we need the S.I.M.G.F. of Y under both hypothesis. (Note the electron charge e has been absorbed by normalization into N_0.)

XI. OBTAINING THE S.I.M.G.F.

As discussed above, the random variable Y is generated when two Poisson processes—one due to signal if present and the other due to dark current—drive a random multiplier. In Sections II–VI, statistics of two avalanche multiplier processes are discussed. These results will be called upon soon.

Suppose we wish to know the S.I.M.G.F. of the total number of secondary electrons leaving the random multiplier in an interval. We know that the number of electrons incident upon the multiplier has Poisson statistics with mean proportional to the interval length T. Let Λ equal this mean. Let $P_\Lambda(n)$ equal the Poisson density with mean Λ. Let $P_m(x)$ equal the density of the number of secondaries resulting from a given primary. The density of the total secondaries is given by

$$P_{\text{total sec}}(x) = \sum_{n=0}^{\infty} P_\Lambda(n) P_m(x)^{*n} \tag{35}$$

where the notation $P_m(x)^{*n}$ denotes convolution of $P_m(x)$ with itself n times.

If $M_Y(s)$ is the moment generating function of the total secondaries, then transforming equation (35) we obtain

$$M_Y(s) = \sum_0^\infty P_\Lambda(n) M_m^n(s)$$

$$= \sum_0^\infty P_\Lambda(n) \exp\left(n\psi_m(s)\right)$$

$$= M_\Lambda(\psi_m(s)) \tag{36}$$

where

$M_m(s)$ is the moment generating function of the multiplication statistics, $\psi_m(s)$ is the corresponding S.I.M.G.F. [natural log of

---

*If $n(t) =$ thermal noise and $N = \int_0^T n(t)\ dt$ where $E[n(t)n(u)] = N_0\delta(t - u)$, then we have $E[N^2] = E[\int_0^T \int_0^T n(t)n(u)\ dt\ du] = N_0 T$.

$M_m(s)$] and $M_\Lambda(s)$ is the moment generating function associated with the Poisson density.

But

$$M_\Lambda(s) = \sum_0^\infty [\Lambda^n \exp(-\Lambda)/n!] \quad \exp(sn) = \exp(\Lambda(\exp(s) - 1)). \quad (37)$$

Thus

$$M_Y(s) = \exp(\Lambda(M_m(s) - 1)), \qquad (38)$$
$$\psi_Y(s) = \Lambda(M_m(s) - 1).$$

Thus, if we know $\lambda_d T$ the component of Λ due to dark current, and if we know $M_m(s)$, then we have the S.I.M.G.F. of Y under both hypotheses.

XII. APPLICATION OF CHERNOFF BOUND

Using equation (38), we know that the S.I.M.G.F. of the statistic X under hypothesis zero is

$$\psi_{X0}(s) = \lambda_d T(M_m(s) - 1) + N_0 T s^2/2. \qquad (39)$$

Under hypothesis one, the S.I.M.G.F. is

$$\psi_{X1}(s) = (\lambda_d + P\eta/hf)T(M_m(s) - 1) + N_0 T s^2/2. \qquad (40)$$

Since equation (30) is independent of the signal power P, we can use the following Chernoff bound

$$p_f \leq \exp[\psi_{X0}(s) - s\psi'_{X0}(s)]\,|_{\psi'_{x0}(s) = \gamma} \qquad (41)$$

to determine an upper bound to the required γ to achieve the desired false alarm probability. That is, we evaluate (41) at equality to determine γ given the desired p_f.

Having obtained γ in this way, we use the bound

$$p_m \leq \exp[\psi_{X1}(s) - s\psi'_{X1}(s)]\,|_{\psi'_{x1}(s) = \gamma} \qquad (42)$$

to determine the upper bound on the power P. That is, using the value of γ found above and using (42) we determine an upper bound to the required power P. It is an upper bound since γ is larger than is actually required, and (42) itself is a bound given γ.

XIII. OBTAINING A LOWER BOUND

Examine the system shown in Fig. 5. Essentially we have there the system of Fig. 4, except that the random multiplier has been replaced
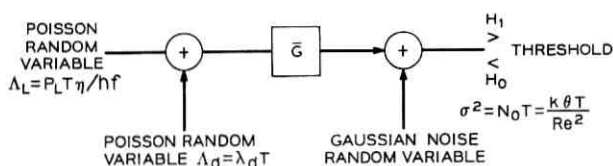
Fig. 5—Lower bounding receiver.

by a deterministic gain. It is intuitive and straightforward to prove that the system of Fig. 5 performs better than the system of Fig. 4, for a given energy per pulse and a given dark current. Therefore, if we determine the energy per pulse required for the system of Fig. 5 to achieve a given error rate, that value is a lower bound to the corresponding quantity for the system of Fig. 4.

The technique we use to analyze the system in Fig. 5 depends upon the parameter values. If under hypothesis one or under both one and zero, the mean number of electrons per pulse is large enough (greater than 250), we can simply replace the Poisson random variable by its mean. This results in a slightly looser lower bound, since we are eliminating another random quantity. Calculations indicate that the difference between the lower bounds for such large mean count numbers is small. On the other hand, for small mean count numbers on either hypothesis zero or zero and one, we should retain the Poisson statistics and use a computer to calculate the required threshold and light energy. It should be noted that the system is indifferent to a change of noise standard deviation σ and gain G provided the ratio of these two quantities is fixed. A calculation of the required mean number of counts per pulse $\Lambda_L = P_L T \eta / h f_0$ in the presence of light, assuming no dark current and a 10^{-9} error rate is displayed in Fig. 6.

XIV. APPLICATION TO AVALANCHE DIODE DETECTORS

In Sections II through VI, the moment generating function we call $M_m(s)$ (of the multiplication probability density) was derived for two types of avalanche diodes. The first is a two-carrier device with equal hole and electron ionization probabilities (e.g., a Germanium diode). For such a device $M_m(s)$ must satisfy the differential equation

$$d/ds[M_m(s)] = M_m(s) \left/ \left(1 - \left(\frac{\bar{G} - 1}{\bar{G}} \right) M_m(s) \right) \right. ; \qquad M_m(0) = 1 \qquad (43)$$

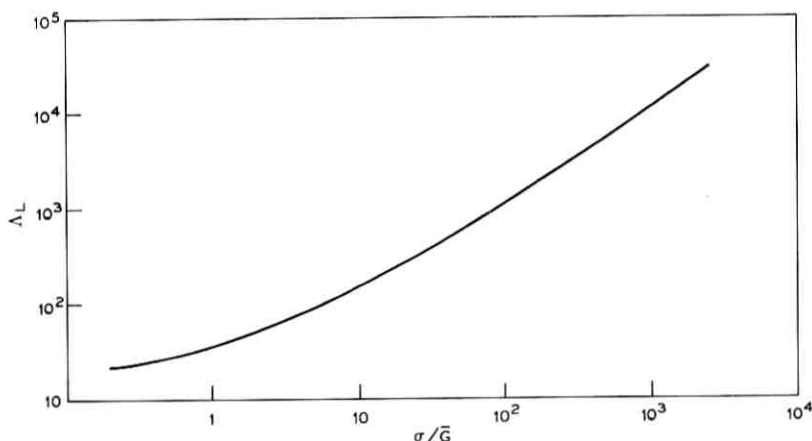where $\bar{G}$ is the mean multiplication.

Fig. 6—Lower bound to required mean detected photons per pulse vs $\sigma/\bar{G}$. $\sigma = \sqrt{N_0 T}$.

The second type of diode is a unilateral gain device (single carrier ionization). The corresponding moment generating function is

$$M_m(s) = 1/[1 + \bar{G} (\exp (-s) - 1)]. \tag{44}$$

For $\bar{G}$ larger than about 10 the probability density (25) of the multiplication for this device is approximately exponential (it is exactly a shifted Bose–Einstein) and would have moment generating function

$$M_m(s) \underset{\text{expon. approx.}}{=} 1/(1 - s\bar{G}). \tag{45}$$

Suppose we wish to apply the bounds to these devices. Let us take the following special case for our calculations. We shall assume zero dark current. We will set the desired error rate to 10^{-9}. Under these assumptions the probability of a false alarm is the probability that the thermal noise of variance $N_0 T$ drives the statistic X above threshold. Since the noise is gaussian, it is straightforward to see that the threshold is very nearly (to two decimal places) six standard deviations, i.e., $6(N_0 T)^{\frac{1}{2}}$, for a 10^{-9} error rate. Using this threshold, we calculate the value of P, the light power needed to achieve a 10^{-9} miss probability.

We have

$$10^{-9} = p_m \leq \exp ((PT\eta/hf)(M_m(s) - 1 - M'_m(s)) - N_0 T s^2/2) \tag{46}$$

where we must have

$$(PT\eta/hf)M_m'(s) + N_0Ts = 6(N_0T)^{\frac{1}{2}}.$$

Solving equation (45) with equality for P yields an upper bound.

The lower bound, calculated as described in Section VII as a function of the noise standard deviation over the mean gain is plotted in Fig. 6.

Without loss of generality, we solve for the upper and lower bounds to $\Lambda_s = TP_s\eta/hf$, thus eliminating the quantum efficiency and carrier frequency as parameters.

Now we must specify N_0 and T. The thermal noise has been normalized by the electron charge squared so that Y could represent the total number of received electrons rather than the total charge per pulse. Therefore

$$N_0 = k\theta/(Re^2) \qquad (47)$$

where

  k is the Boltzman constant,
  R is the equivalent receiver input resistance,
  e is the electron charge,
  θ is the absolute receiver temperature (assumed 300°K).

Assuming T roughly in the range of 10^{-8} to 10^{-11} seconds, and R roughly 10 to 100 ohms we find that the square root of N_0T the noise standard deviation is in the range of 100 to 10,000.

If we apply these parameter values to equation (46) using the multiplication statistics of (43), (44) and (45), we obtain the following results.

For the single carrier diode, and mean gains of 20 and 100 in the multiplier, the actual moment generating function (44) and exponential approximation (45) yield equivalent results within desired computational accuracy. This leads to an interesting consequence. The upper bound is independent of changes in the noise standard deviation and the mean gain, provided the ratio of these two quantities is the same. (This is a consequence of the exponential multiplication and is not a general result.) Thus, a single curve can be plotted where the abscissa represents this ratio, and the ordinate is the percent by which the upper bound Λ_U to the required counts Λ_s exceeds the lower bound Λ_L. In fact, we can also plot the same ordinate against the lower bound Λ_L as abscissa. Such plots are shown in Figs. 7 and 8 respectively.
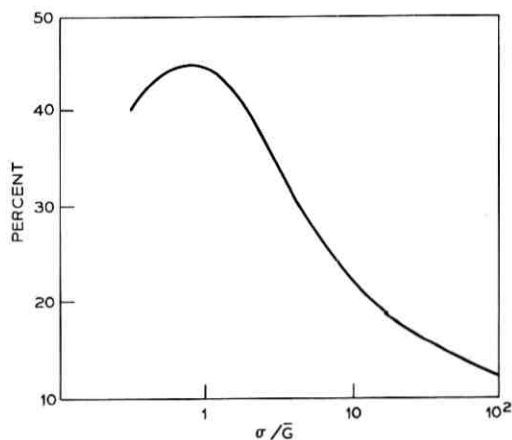
Fig. 7—Percent by which upper bound exceeds lower bound vs $\sigma/\bar{G}$ for unilateral gain diode. Percent $= ((\Lambda_U - \Lambda_L)/\Lambda_L) \times 100$.
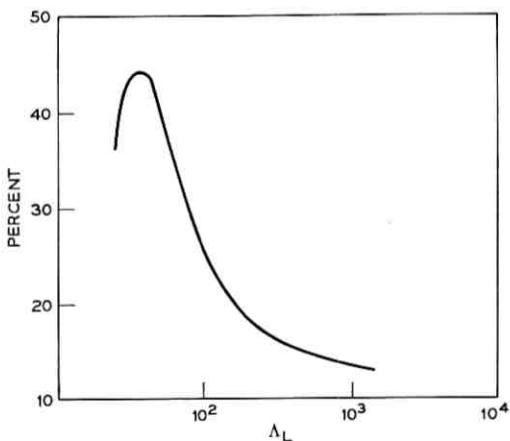
Fig. 8—Percent by which upper bound exceeds lower bound vs lower bound for unilateral gain diode.

For the two carrier diode, the above simplification does not occur. A family of curves similar to the ones above must be plotted for different mean multiplications. Two such curves (for $\bar{G} = 100$ and $\bar{G} = 20$) are plotted in Fig. 9.

14.1 *Example:*

To use the curves, assume for instance the following parameter values

mean avalanche gain $\bar{G} = 100$,
light pulse duration $T = 10^{-10}$,
dark current—negligible,
equivalent noise resistance of circuitry following avalanche detector $R = 100\Omega$,

then we have

$$N_0 = \frac{k\theta}{Re^2} = \frac{4.14 \times 10^{-21}}{2.56 \times 10^{-36}} = 1.6 \times 10^{15} ,$$

$$\sigma = (N_0 T)^{\frac{1}{2}} = 4 \times 10^2 ,$$

$$\sigma/\bar{G} = 4.$$

Fig. 9—Percent by which upper bound exceeds lower bound vs $\sigma/\bar{G}$ for equal ionization diode.

From Fig. 6, the lower bound $\Lambda_L = \eta P_L T / h f_0$ to the required mean number of detected photons for a 10^{-9} error rate is roughly 100. From Fig. 7, the upper bound for a unilateral gain device is 30 percent more or 130 photons.

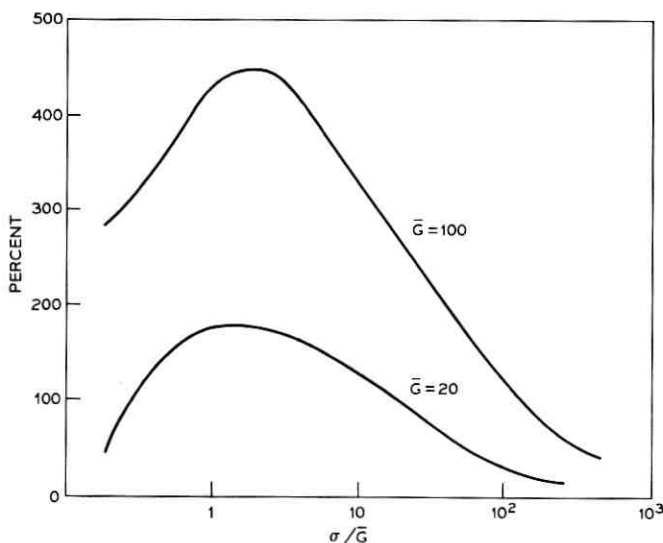From Fig. 9, the upper bound for an equal ionization avalanche diode is 500 percent more than the lower bound or 600 photons. The upper bound gives an indication of the degradation associated with random gain rather than deterministic gain.

14.2 *Comments*

The curves of Figs. 6–9 are intended as an example of the use of equations (39) through (42) with devices having statistics given by (43) or (44) and (45) assuming no dark current and an error rate of 10^{-9}.

For other error rates and nonnegligible dark current, we must return to equations (39) through (45) and generate new curves.

For other types of random gain we can still use equations (39) through (42). The cases of equal electron-hole ionization and unilateral gain are probably bounds to the practical case of unequal nonzero ionizations. That is, if results on mean square gain[1,2] are any indication, the degradation for an unequal ionization device should lie between the degradations (compared to deterministic gain) of the devices studied above. To test this conjecture, we need the statistics (moment generating functions) of such unequal ionization devices.

XV. CONCLUSIONS

We have derived the moment generating functions for two special case avalanche diodes. The important intermediate cases of unequal ionization coefficients is an important area for further study.

Application to the direct detection receiver indicates that for important ranges of parameter values we can tightly bound the energy required for a desired error rate. Further, the nature of the lower bound indicates that there may not be much degradation, for certain parameter values, due to the use of random rather than deterministic gain.

APPENDIX

Glossary of Terms

$\bar{G}$ = mean avalanche gain
T = input pulse duration

e = electron charge
R = post detector circuit equivalent noise resistance
θ = 300°K
k = Boltzmann constant
N_0 = $(k\theta/Re^2)$ = post detection circuit thermal noise spectral height
σ = $\sqrt{N_0 T}$
Λ_s = required mean number of detected photons to achieve a desired error rate
Λ_U = upper bound to Λ_s
Λ_L = lower bound to Λ_s

REFERENCES

1. McIntyre, R. J., "Multiplication Noise in Uniform Avalanche Diodes," IEEE Trans. on Elec. Devices, *ED- 13*, No. 1 (January 1966), pp. 164–168.
2. Tager, A. S., "Current Fluctuations in a Semiconductor Under Conditions of Impact Ionization and Avalanche Breakdown," Sov. Phys–Solid State, *8* (1965), pp. 1919–1925.
3. Van Trees, H. L., *Detection Estimation and Modulation*, Vol. I, New York: Wiley and Sons, 1967.

Exchange of Spatial and Temporal Resolution in Television Coding

By R. F. W. PEASE and J. O. LIMB

We describe in this paper a television system in which only a fraction of the picture elements is sampled. A movement detecting circuit examines the frame-to-frame difference signals and divides the picture into "moving" and "stationary" areas. In stationary areas the unsampled elements retain their value from the previous frame; in moving areas the unsampled elements take interpolated values from the current frame. When one-half of the elements is sampled, the resulting pictures are difficult to distinguish from the original, fully sampled, picture. When the fraction is only one quarter, the degradation is visible and disturbing. We mention several possible improvements.

I. INTRODUCTION

Television coders are usually designed to meet simultaneously the worst contingencies with respect to contrast, sharpness and movement. That is, a fast moving subject can be reproduced with the full spatial resolution afforded a stationary subject and with the full contrast resolution afforded a low detail subject. By contrast resolution we mean the accuracy with which the coded signal represents the amplitude of the input signal for a particular picture element.

We want to reduce the channel capacity required for transmitting television signals by coding the signal so that full spatial resolution is only available in stationary areas of the picture and full temporal resolution is only available in moving areas of the picture.

Exchange of spatial and contrast resolution has been appreciated and demonstrated for some time.[1] One of the first examples of a coder that exchanged spatial and contrast resolution was demonstrated by E. R. Kretzmer[2]; more quantizing levels were assigned to low frequency signal components and fewer levels were assigned to high frequency signal components.

A. J. Seyler[1] has pointed out that this principle can be extended to exchanging temporal and spatial resolution because we can tolerate blurring of moving objects. This is evident because the television camera integrates the video signal on the target for 1/30th of a second. Thus, if the object being imaged on the target moves during this period, the image will be blurred and resolution will be destroyed in the direction in which the object moves. The amount of blurring can be quite large. For example, if the object moves at a speed which would require two seconds to cross the television screen, then in one frame-time a picture element would average light from eight different picture elements on the object (at broadcast television rates). Consequently, if the object is moving in the horizontal direction, the horizontal resolution is reduced approximately eightfold; there is even more reduction if there is camera lag.

It is not clear whether people tolerate such blurring because of the psycho-physical integration of moving objects[3] or because of long exposure to television, but the fact that we do tolerate this loss in resolution means that we should be able to reduce the spatial sampling rate in moving areas of a television picture without degrading picture quality. We refer to sampling at a reduced rate along a scan line (or along a vertical line) as spatial subsampling. The unsampled picture elements are replaced by interpolating between the sampled elements.

If an object moves slowly enough, temporal resolution can be greatly reduced without impairing picture quality. A simple method of achieving this is by frame repeating, in which one new frame of information out of n is transmitted and for the remaining $(n - 1)$ frames this one frame is just repeated.[4] As would be expected if the subject moves fast enough, the perceived motion is very jerky. In alternative, and more pleasing, methods for reducing temporal resolution $1/n^{th}$ of the points are replenished in each frame in a dot inter-laced fashion.[4-6] The type of picture degradation for these schemes differs from frame repeating in that when a subject moves, the edges become blurred and exhibit a checkerboard texture. We refer to sampling a given picture element at the reduced frame rate as temporal subsampling.

By temporally subsampling in stationary areas and spatially sub-sampling in moving areas we should be able to reduce the channel capacity required to transmit a satisfactory television signal. We have made a digital television system in which the picture is divided into stationary and moving areas which are appropriately subsampled.

This apparatus, the resulting pictures and the saving of required channel capacity are described below.

II. APPARATUS

An outline schematic is shown in Figure 1. The scan format is 30 frames per second with a 2:1 interlace. Each frame has 271 lines, each comprising 248 picture elements. The television signal source is a vidicon camera; the video signal is digitized to 8-bit accuracy in the Pulse Code Modulation (PCM) encoder and stored in the digital ultrasonic wire delay lines of the frame memory. The output of the frame memory is subtracted from the PCM encoder output to yield a frame difference signal for each picture element. These signals are examined in the movement detector whose output determines whether the current element (entering the area) is to be regarded as being in a moving or stationary area. The output changes from a "stationary state" $(s = 0)$ to a "moving state" $(s = 1)$ when more than n frame difference signals out of a sequence of m exceed a threshold. The output will return to the stationary state when all m samples exhibit insignificant frame difference signals.

Thus the movement detector exhibits hysteresis reducing the frequency of mode changes so that the picture is segmented into relatively few contiguous moving and stationary areas.

In areas judged to be stationary, alternate elements in each line are sampled according to the pattern of Fig. 2a where the elements tabulated A are sampled in one frame and the elements tabulated B are sampled in the next. The unsampled elements retain the value of
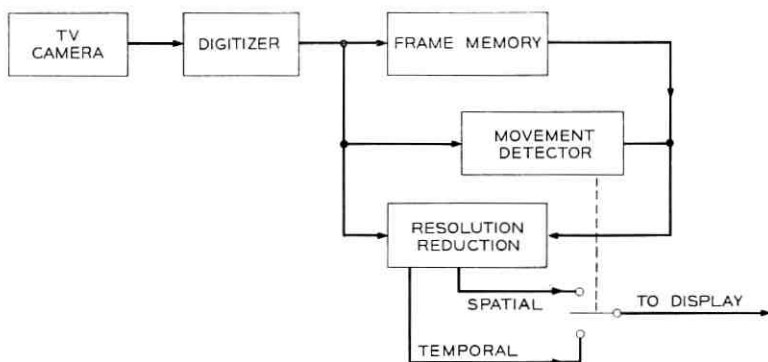

Fig. 1—System for resolution exchange.

the previous frame. In the moving areas the same sampling pattern could be used, but we found that the pattern shown in Fig. 2b gave a better picture; the unsampled elements (marked 0) are given the mean value of the two neighboring elements. When the sampling rate along the line is reduced, we rely on the blurring caused by the motion to bandlimit the video signal. A bandlimiting filter should also be used on the output signal. However, with the sampling pattern of Fig. 2b linear interpolation between the sampled elements gave satisfactory pictures although aliasing patterns were just visible when the subject moved at a certain speed. Further filtering removed these patterns but caused blurring which we judged to be more objectionable.

Strickly speaking, frame-to-frame comb filters should be used in "stationary" areas because of the reduced temporal sampling rate but such filters would involve frame delays and were felt to be impractical. The aforementioned checkerboard texture which can be seen on moving edges in temporally subsampled areas is probably due to the resulting aliasing. The experiments carried out with this apparatus and the resulting pictures are described in the next two sections.

In a second series of experiments only every fourth element was sampled and the sampling patterns for the stationary and moving areas are shown in Figs. 2c and 2d. Only fast movement can now sufficiently bandlimit the video signal and so there are switched filters before and after the subsampling.

Some experiments were also conducted with a 4-bit frame difference quantized signal. The characteristics of the quantizer are:

INPUT LEVELS

±	0, 1	2–5	6–11	12–17	18–27	28–37	38–52	54–255

OUTPUT LEVELS

±	0	2	8	14	22	32	44	60

III. PROCEDURE

Each scene viewed by the television camera was a head and shoulders view of a person talking; the movement varied from 'gentle' (only lip movement) to 'very active' (arm movement and the person getting up and walking away).

We viewed the resulting pictures on a television monitor. The unblanked raster size was $5\frac{1}{2}$ inches horizontal by 5 inches vertical. The viewing distance was approximately 40 inches and the ambient

```
                                                A O O O A O O O A
                                                A O O O A O O O A
A B A B A B A B A     A O A O A O A O A     A C B D A C B D A     O O A O O O A O O
A B A B A B A B A     A O A O A O A O A     A C B D A C B D A     O O A O O O A O O
B A B A B A B A B     O A O A O A O A O     B D A C B D A C B     A O O O A O O O A
B A B A B A B A B     O A O A O A O A O     B D A C B D A C B     A O O O A O O O A
(a) TEMPORAL 2:1      (b) SPATIAL 2:1       (C) TEMPORAL 4:1      (d) SPATIAL  4:1
```

Fig. 2—Sampling patterns. ABCD refer to consecutive frames. 0 denotes average of neighboring, sampled elements.

illumination was about the average office illumination (70 footcandles).

For each scene we first viewed the picture resulting from the 8-bit PCM signal and then the processed pictures in the following order: continuously temporally subsampled; continuously spatially subsampled; temporally subsampled in stationary areas and spatially subsampled in moving areas. As a check we sometimes 'flagged' the area judged moving by the movement detector. For most work the detector was adjusted to switch to spatial subsampling when four out of eight picture elements (pels) examined exhibited frame-to-frame difference signals exceeding a threshold of four (out of 255 levels); for a return to the temporal subsampling mode, none of the eight examined pels exhibited a frame difference exceeding this threshold.

For recording still photographs a model head was swung as a pendulum bob so that the object speed would be known. One television frame (selected at the bottom of the swing) was stored, displayed continuously and photographed with an exposure time of $\frac{1}{4}$ second.

IV. RESULTS

The pictures which were temporally subsampled (by a factor of 2) all over, were not only excellent for stationary scenes but also satisfactory for scenes with an object (e.g., a head) speed up to one pel per frame interval for an object of normal contrast. At higher speeds the checkerboard pattern already referred to and described by others[4] became visible; at object speeds of two pels per frame interval and above, this pattern was annoying for most scenes.

When the pictures were spatially subsampled by a factor of two all over, the loss of resolution was just visible for most scenes; the loss was more obvious when viewing graphical material.

When the sampling frequency was halved temporally in stationary areas and spatially in moving areas, the resulting pictures were

usually indistinguishable from the fully sampled picture. The scenes where the difference was visible contained a contrasty edge moving slowly (one pel per frame interval or less); in this case the continuity of the edge was disturbed. This effect is caused by erratic movement detection so that parts of the edge are sometimes spatially subsampled and sometimes temporally subsampled. More sophisticated movement detection should reduce this edge breakup by changing modes only when the blurring due to temporal subsampling is comparable with the blurring due to spatial subsampling.

Some still photographs are shown in Fig. 3. In Fig. 3a the whole scene is stationary and the picture is fully sampled. In Fig. 3b the head is moving at about three pels per frame interval (estimated accuracy $+0$, $-\frac{1}{2}$ pels per frame interval); the picture is still fully sampled and the blurring introduced by the movement can be seen by comparison with Fig. 3a. Fig. 3c shows an equivalent scene with a 2:1 exchange of resolution. There is very little difference between this picture and the fully sampled one (Fig. 3b). Fig. 3d shows the bright flags indicating that part of the scene judged to be moving during one frame.

The results using 4:1 subsampling were less encouraging. With overall 4:1 temporal subsampling (according to the pattern of Fig. 2c) the blurring and checkerboard patterns were objectionable even in such slowly moving areas as someone's mouth when he was talking. In the pictures resulting from 4:1 spatial subsampling (according to the pattern in Fig. 2d), the blurring was again objectionable, especially in stationary areas.

In the pictures resulting from exchanging resolution according to the movement detector, the breakup of edges was much more apparent and the movement detection was erratic; even some parts of the background or other stationary areas were judged moving. These errors arise because the digital filter in the loop gives rise to frame difference signals which may be interpreted as movement by the movement detector. This effect keeps the digital filter switched in even in stationary areas. Again, more sophisticated movement detection should reduce this effect.

The pictures coded as a frame difference quantized signal (instead of 8-bit PCM) were usually similar to those coded as 8-bit PCM. The only difference was with scenes containing fast movement (speeds of four pels per frame interval or more) of contrasty objects; these objects appeared somewhat noisier than in those pictures coded as eight-bit PCM.
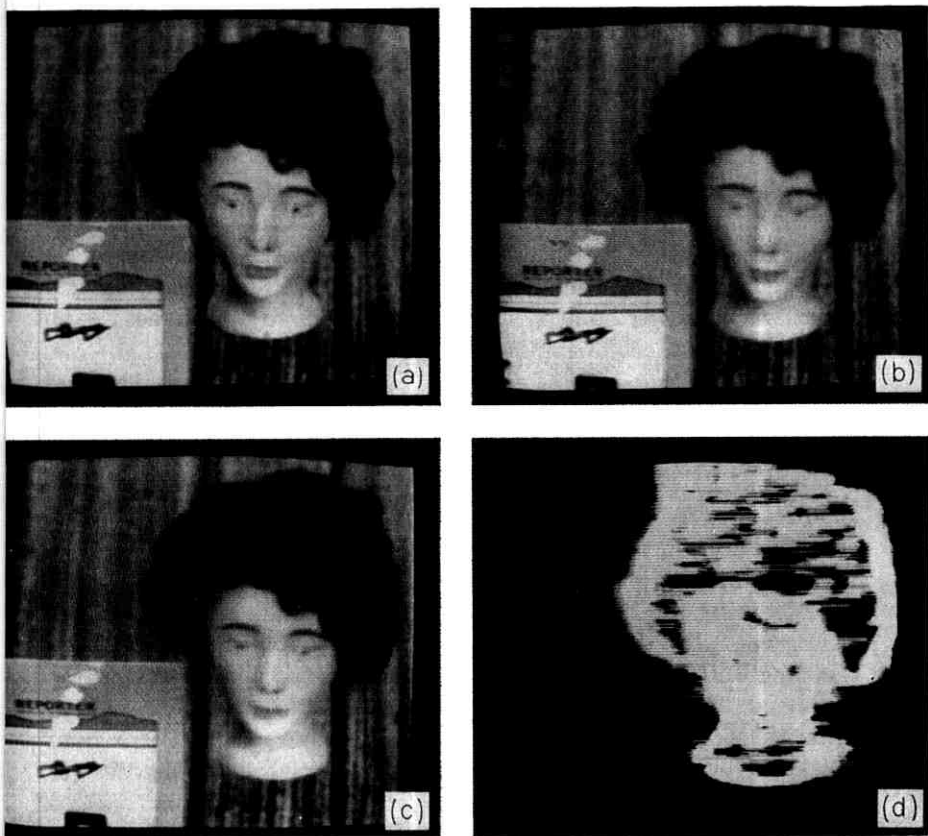
Fig. 3—(a) Fully sampled picture showing stationary head. (b) Fully sampled picture showing moving head. (c) Subsampled (2:1) picture using exchanged resolution. (d) Flags showing area deemed moving. Glossy prints of this figure may be obtained by writing to the authors.

V. DISCUSSION

5.1 Comparison of 2:1 and 4:1 Subsampling

Although the simple techniques described above yielded a satisfactory picture when the overall sampling rate was halved, the equivalent techniques did not give a satisfactory picture when the sampling rate was quartered. The unpleasant effect of the latter technique appears mostly at the boundaries between areas treated differently. However, there are many possibilities using a 4:1 ratio reduc-

tion that remain to be tried: (i) Use an intermediate mode in which the second of the three unsampled elements is assigned the value from the previous frame, and the first and third are assigned the average value of their neighbors. This is a simultaneous reduction by a factor of two in both spatial and temporal resolution and should result in a less visible boundary between areas encoded in the different modes. (ii) Modify the spatial subsampling patterns so that vertical and horizontal resolution is reduced equally. At present the horizontal loss is more than the vertical loss. (iii) Use an improved movement detector which will indicate the local speed of movement so that modes are not switched until the blurring is similar in each mode.

5.2 *Combination of Subsampling with Other Coding Techniques*

Much digital television employs coders which give either element differential PCM in which the interelement difference along a line is quantized[7] or frame differential PCM in which the difference between consecutive frames is quantized.[8,9]

If the information describing when to switch modes can be transmitted during the blanking interval, then we could use a 4-bit frame difference PCM signal combined with 2:1 resolution exchange for encoding television signals. In our scan format, which resembles that envisaged for *Picturephone*® service, this would require a channel capacity of four megabits per second. We have measured the number of mode changes per line for a number of different scenes and found that it was rarely more than six. If we used a 4-bit word to describe each mode change, then the time required for transmitting the information is rarely more than six microseconds. This is considerably less than the line blanking interval of 13 microseconds for our present scan format. Such a system, although requiring a frame memory, is essentially bufferless as a 6-word shift register suffices to buffer the above mode switching information. Applying spatial subsampling techniques to element differential PCM is not straightforward because when we halve the horizontal sampling frequency we no longer have such near neighbors for predicting the value of an upcoming element and therefore quantizing noise increases. Two possible alternatives are:

(i) Use vertically adjacent elements for predicting the upcoming element. Because of interlace such elements are temporally displaced by one field time from the current line and may be unsatisfactory for predicting values in the current line in moving

areas. This problem would not arise with a sequential, non-interlaced, scan format.

(ii) Reduce only vertical resolution by sampling alternate lines in each field. Such a process may give rise to an unpleasant rolling pattern in the picture.

We describe experiments in which resolution exchange is applied to an element differential PCM signal in a forthcoming publication.

A third form of coding which is being considered is conditional picture element replenishment[10] in which only the signal corresponding to the moving part of the picture is transmitted. Such a system requires a large buffer store which can become overloaded when the pictures contain large areas of rapid movement. The use of spatial subsampling, either 2:1 or 4:1, in such moving areas is one way of reducing or preventing buffer overflow.[9,11] If we are aiming for a bit rate of one-bit per picture element per frame we can simulate the worst possible case by spatially subsampling by 4:1 the whole picture encoded as 4-bit frame differential PCM. The resulting picture looks much better than the picture tearup which occurs when a buffer overloads[10] and, hence, this mode can be used whenever the buffer approaches overflow.

VI. ACKNOWLEDGMENTS

REFERENCES

1. Seyler, A. J., "The Coding of Visual Signals to Reduce Channel-Capacity Requirements," Proc. IEE, 109, Part C, No. 16 (September 1962), pp. 676–684.
2. Kretzmer, E. R., "Reduced-Alphabet Representation of Television Signals," 1956 IRE Conv. Rev., Part 4, pp. 140–147.
3. Van den Brink, G., "The Visibility of Details of a Moving Object," Acta Electronica, 2 (1957–58), pp. 44–49.
4. Brainard, R. C., Mounts, F. W., and Prasada, B., "Low Resolution TV: Subjective Effects of Frame Repetition and Picture Replenishment," B.S.T.J., 46, No. 1 (January 1967), pp. 261–271.
5. Deutsch, S., "Pseudo Random Dot Scan Television Systems," Trans. On Broadcasting IEEE, BC-11, No. 1 (July 1965), pp. 11–21.
6. L. C. Jesty, "Television as a Communication Problem," Proc. IEE, 99, Part 3A, No. 20 (1952), pp. 761–770.
7. Limb, J. O., and Mounts F. W., "Digital Differential Quantizer for Television," B.S.T.J., 48, No. 7 (September 1969), pp. 2583–2599.
8. Schaphorst, R., "Frame to Frame Coding of N.T.S.C. Color TV," Symposium on Picture Bandwidth Compression, April 1969, M.I.T.
9. Candy, J. C., Franke, M. A., Haskell, B. G., and Mounts, F. W., to be published.

10. Mounts, F. W., "A Video Encoding System Using Conditional Picture-Element Replenishment," B.S.T.J., *48*, No. 7 (September 1969), pp. 2545–2554.
11. Medlin, J. E., "The Prevention of Transmission Buffer Overflow in Telemetry Data Compressors," IEEE Trans. Comm. Tech., *COM-16*, No. 1 (February 1968), pp. 94–107.

On Beneš Rearrangeable Networks

By F. K. HWANG

V. E. Beneš considered a class of multi-stage switching networks[1] and proved that if the linkage pattern between two stages is chosen in a specific way, then the resulting networks are rearrangeable. We offer a simpler proof by pointing out the relation between Beneš class networks and the Slepian–Duguid Theorem on three-stage Clos networks.

I. INTRODUCTION

V. E. Beneš[1] considered the class, denoted here by $B(n_1, n_2, \cdots n_{t+1})$, of all connecting networks ν with the following properties:

(*i*) ν is two sided, with N terminals on each side where $N = \prod_{i=1}^{t+1} n_i$.

(*ii*) ν is built of an odd number $s = 2t + 1$ of stages ζ_k , $k = 1, \cdots 2t + 1$ connected as specified by permutations $\varphi_1, \cdots \varphi_{2t}$. In the notation of Beneš,

$$\nu = \zeta_1 \varphi_1 \zeta_2 \cdots \varphi_{2t} \zeta_{2t+1} .$$

(*iii*) ζ_k consists of N/n_k identical square switches of size n_k .

(*iv*) $\zeta_k = \zeta_{2t+2-k}$ for $k = 1, \cdots t$.

Beneš proceeded to prescribe a specific way of choosing φ_k (See page 113 of Ref 1):

"Order the switches of each stage; to define φ_k for a given $1 \leq k \leq t$, take the first switch of ζ_k , say with n_k outlet and n_k a divisor of N, and connect these outlets one to each of the first n_k switch of ζ_{k+1} ; go on to the second switch of ζ_k and connect its n_k outlets one to each of the next n_k switches of ζ_{k+1} ; when all the switches of ζ_{k+1} have one link on the inlet side, start again with the first switch; proceed cyclically in this way till all the outlets of ζ_k are assigned."

Beneš also specified $\varphi_k = \varphi_{2t+1-k}$ for $k = t + 1, \cdots 2t$. We call a network $\nu \, \varepsilon \, B(n_1 , n_2 , \cdots n_{t+1})$ constructed in this manner a cyclic Beneš network.

Beneš proved that a cyclic Beneš network is rearrangeable. The proof given by Beneš in Ref. 1 is, however, quite lengthy and involved. In the present work, we offer an alternate proof by pointing out the links between cyclic Beneš networks and the Slepian–Duguid Theorem on three-stage Clos networks. It is hoped that the simpler proof will lead to new insights into the problem of constructing rearrangeable networks.

II. MULTISTAGE REARRANGEABLE NETWORKS

We show a way to construct $\nu \ \varepsilon \ B(n_1, n_2, \cdots n_{t+1})$ from $\nu' \ \varepsilon$ $B(n_2, n_3, \cdots n_{t+1})$ such that if ν' is rearrangeable then ν is rearrangeable.

We construct a three-stage network by having the first stage and the third stage each consist of N/n_1 copies of $n_1 \times n_1$ square switches, $A_1 \cdots A_{N/n_1}$, $C_1 \cdots C_{N/n_1}$ say, where the second stage consists of n_1 copies of ν', say, $B_1, \cdots B_{n_1}$. Each switch in the first stage and the third stage is then linked to every B_j in the second stage. (It does not matter which inlet or outlet of B_j is linked to which A_i or which C_i.) This gives a network $\nu \ \varepsilon \ B(n_1, n_2, \cdots n_{t+1})$. Now if ν' is also constructed in this manner and so on down to the three-stage network, we then call ν a Beneš network. That a Beneš network is rearrangeable will follow from a multistage version of the Slepian–Duguid Theorem on three-stage Clos networks. To be complete, we state this multistage version and give a proof.

Theorem 1: Let $\nu \ \varepsilon \ B(n_1, n_2, \cdots n_{t+1})$ be a Beneš network. Then ν is rearrangeable.

Proof: For $t = 0$, we have a special case of a $n_1 \times n_1$ square switch which is clearly rearrangeable and no construction is needed. Supposing that Theorem 1 is true for $t' = t - 1 \geqq 0$, we prove Theorem 1 for $t' = t$.

A maximal assignment between each inlet terminal and each outlet terminal is a permutation φ of the set of numbers $\{i : i = 1, 2, \cdots N\}$ where $N = \prod_{j=1}^{t+1} n_j$. (Following Beneš, we need only consider maximal assignments.) A given maximal assignment can also be viewed as a set of defining relationships between each switch in the first stage and each switch in the third stage (treating the set of ν' as the second stage). Set $N_2 = n_2 \cdot n_3 \cdot \cdots n_{t+1}$, note that $N = n_1 N_2$. Let A_i be the ith switch in the first stage, $A = \{A_i : i = 1, 2, \cdots N_2\}$ and C_j be the jth switch in the third stage, $C = \{C_j : j = 1, 2, \cdots N_2\}$. Consider a particular first stage switch A_i. Suppose the n_1 inlet terminals of

A_i are assigned by φ to third-stage outlet terminals, y_{i1}, y_{i2}, $\cdots y_{in}$. Denote by $Y_{ik}(Y_{ik} \; \varepsilon \; C)$ the third-stage switch that contains y_{ik}, $k = 1, 2, \cdots n_1$, and let $S_i = \{Y_{ik} : k = 1, 2, \cdots n_1\}$. Since $\bigcup_{i=1}^{x} S_i$ has a total of $x \cdot n_1$ elements and since each distinct element has only n_1 repetitions, there are at least x distinct elements in $\bigcup_{i=1}^{x} S_i$ for each $x = 1, 2, \cdots N_2$. Hence the condition of P. Hall's Theorem[2] on distinct representation of subsets is satisfied and there exists a set $Z = \{Z_i : i = 1, 2, \cdots N_2\}$ such that $Z_i \; \varepsilon \; S_i$ and $Z = C$. Then x_i, $i = 1, \cdots N_2$ can be chosen such that $x_i \; \varepsilon \; A_i$, $\varphi x_i \; \varepsilon \; Z_i$. Now we may choose to route the n_1 calls $x_i \to \varphi x_i$ through the first second-stage switch B_1, which by induction hypothesis is rearrangeable. The problem is then reduced to that of a maximal assignment in a network of type $B(n_1 - 1, n_2, \cdots n_{t+1})$. By repeatedly applying the same argument, we obtain subassignments on each B_k.

III. CYCLIC BENEŠ NETWORKS

Theorem 2: Every cyclic Beneš network is a Beneš network, hence is rearrangeable.

Proof: A single stage cyclic Beneš network is a $N \times N$ square matrix, which is a Beneš network. Suppose Theorem 2 is true for $(2t - 1)$ stage network where $t \geq 1$. We prove a $(2t + 1)$ stage cyclic Beneš network is a Beneš network.

Let $\nu = \zeta_1 \varphi_1 \zeta_2 \cdots \varphi_{2t} \zeta_{2t+1} \; \varepsilon \; B(n_1, n_2, \cdots n_{t+1})$ be a cyclic Beneš network. We shall show that the complex $p = \zeta_2 \varphi_2 \zeta_3 \cdots \varphi_{2t-1} \zeta_{2t}$ can be decomposed into n_1 copies of ν' where $\nu' \; \varepsilon \; B(n_2, n_3 \cdots n_{t+1})$ is also a cyclic Beneš network. Furthermore φ_1 and φ_{2t} are such that each switch in ζ_1 and ζ_{2t+1} is linked to each ν' as is required by the definition of a Beneš network. By our inductive hypothesis, ν' is a Beneš network. This is enough to prove that ν is a Beneš network.

Let the notation $\{A/\zeta_i(\sigma)\}$ denote a set of switches $A \; \varepsilon \; \zeta_i$ in the network σ. Similarly, let $\{B/\zeta_{i+1}(\sigma)\}$ denote a set of switches $B \; \varepsilon \; \zeta_{i+1}$ in σ. We define a relation R on the two sets A and B and write

$$\{A/\zeta_i(\sigma)\} R \{B/\zeta_{i+1}(\sigma)\}$$

if every switch of A is linked to every switch of B in the network σ.

For the cyclic Beneš network $\nu = \zeta_1 \varphi_1 \zeta_2 \cdots \varphi_{2t} \zeta_{2t+1}$, let the switches in each stage be ordered. Note that the φ_i, for $i = 1, 2, \cdots t$, in ν can be described by,

$$\{x(\bmod f_i)/\zeta_i(\nu)\} R \{n_i(x - 1) + l_i : l_i = 1, 2, \cdots n_i/\zeta_{i+1}(\nu)\}$$

for

$$x = 1, 2, \cdots f_i \quad \text{where} \quad f_i = N/n_i n_{i+1},$$

and

$$\varphi_i = \varphi_{2t+1-i} \quad \text{for} \quad i = t + 1, \cdots 2t.$$

Decompose the compex $P = \zeta_2 \varphi_2 \zeta_3 \cdots \varphi_{2t-1} \zeta_{2t}$ into $\{\nu_1, \nu_2, \cdots \nu_{n_1}\}$ where ν_h, $h = 1, 2, \cdots n_1$, consists of those sets of switches:

$$\{h (\text{mod } n_1)/\zeta_2(\nu)\}$$

$$\left\{(h - 1) \prod_{j=2}^{k-1} n_j + m_k \cdot m_k = 1, 2, \cdots \prod_{j=2}^{k-1} n_j \Big/ \zeta_k(\nu)\right\}$$

$$\text{for} \quad k = 3, 4, \cdots 2t - 1$$

and

$$\{h (\text{mod } n_1)/\zeta_{2t}(\nu)\}.$$

Then clearly $\nu_h \, \varepsilon \, B(n_2, n_3 \cdots n_{t+1})$, $h = 1, 2, \cdots n_1$.

In each ν_h, let the switches in each stage be ordered as is consistent with their orderings in ν. Suppose s_i is a switch $\varepsilon \, \zeta_i(\nu_h)$ and g'_i its coordinate in ν_h (i.e., s_i is the g'^{th}_i switch in $\zeta_i(\nu_h)$). Then $s_i \, \varepsilon \, \zeta_i(\nu)$ and has the coordinate g_i in ν.

If we write g'_i uniquely as

$$g'_i = C \prod_{j=2}^{i-1} n_j + d'_i \quad \text{for} \quad C \geqq 0, \quad 1 \geqq d'_i \geqq \prod_{j=2}^{i-1} n_j \quad (1)$$

$$\left(\text{using the convention } \prod_{j=2}^{i-1} n_j = 1 \quad \text{if} \quad i = 2\right)$$

then g_i can be uniquely written as

$$g_i = C \prod_{j=1}^{i-1} n_j + d'_i + (h - 1) \prod_{j=2}^{i-1} n_j. \quad (2)$$

Vice versa, if $s_i \, \varepsilon \, \zeta_i(\nu)$, $i \neq 1, 2t + 1$, has coordinate g_i as expressed in equation (2), then s_i is also a switch in $\zeta_i(\nu_h)$ with coordinate g'_i as expressed in equation (1).

Next we show that ν_h is a cyclic Beneš network, i.e., let $\nu_h = \zeta'_2 \varphi'_2 \zeta'_3 \cdots \varphi'_{2t-1} \zeta'_{2t}$, then φ_i, $i = 2, 3 \cdots t$ can be described by

$$\{x'(\text{mod } f'_i)/\zeta'_i(\nu_h)\} R\{n_i(x' - 1) + l_i : l_i = 1, 2, \cdots n_i/\zeta'_{i+1}(\nu_h)\} \quad (3)$$

for

$$x' = 1, 2, \cdots f'_i$$

where

$$f' = N'/n_i n_{i+1} \quad \text{and} \quad N' = N/n_1 ,$$

and

$$\varphi_k = \varphi_{2t+1-k} \quad \text{for} \quad k = t + 1, \cdots 2t - 1.$$

(Note that in equation (3), the two sets of switches are written in coordinates g' of ν_h .) From equations (1) and (2),

$$g_i = n_1 g_i' - (n_1 - 1)d_i' + (h - 1) \prod_{j=2}^{i-1} n_j . \tag{4}$$

Let $s_i \ \varepsilon \ \zeta_i'(\nu_h)$ be a switch having coordinate $g_i' \equiv x' \pmod{f_i'}$. Then we have the unique expression

$$g_i' = q \cdot f_i' + x' \quad \text{for some} \quad q \geq 0. \tag{5}$$

Since $\prod_{j=1}^{i-1} n_j$ divides f_i', from equations (1) and (5), there exists a unique u, $0 \leq u \leq \prod_{j=i+1}^{t+1} n_j$, such that

$$x' = u \prod_{j=2}^{i-1} n_j + d_i' . \tag{6}$$

From equation (4), the corresponding g_i is

$$g_i = n_1(q f_i' + x') - (n_1 - 1)d_i' + (h - 1) \prod_{j=2}^{i-1} n_j$$

$$= q n_1 f_i' + n_1 x' - (n_1 - 1)d_i' + (h - 1) \prod_{j=2}^{i-1} n_j$$

$$= q f_i + x \tag{7}$$

where $f_i = n_1 f_i'$ and

$$x = n_1 x' - (n_1 - 1)d_i' + (h - 1) \prod_{j=2}^{i-1} n_j$$

$$= n_1 \left(u \cdot \prod_{j=2}^{i-1} n_j + d_i' \right) - (n_1 - 1)d_i' + (h - 1) \prod_{j=2}^{i-1} n_j , \quad \text{by} \quad (5)$$

$$= u \cdot \prod_{j=1}^{i-1} n_j + d_i' + (h - 1) \prod_{j=2}^{i-1} n_j . \tag{8}$$

It can be easily verified that $0 \leq x < f_i$.

Equation (7) says that a switch $s_i \ \varepsilon \ \zeta_i(\nu_h)$ which has coordinate $g_i' \equiv x_i' \pmod{f_i'}$ has coordinate $g \equiv x \pmod{f_i}$ in ν.

Next we show that if the two sets $\{G_{i+1}'/\zeta_{i+1}'(\nu_h)\}$ and $\{G_{i+1}/\zeta_{i+1}(\nu)\}$

are such that

$$\{x'(\bmod f_i')/\zeta_i'(\nu_h)\}R\{G_{i+1}'/\zeta_{i+1}'(\nu_h)\} \tag{9}$$

and

$$\{x(\bmod f_i)/\zeta_i(\nu)\}R\{G_{i+1}/\zeta_{i+1}(\nu)\} \tag{10}$$

hold, then G_{i+1}' and G_{i+1} are coordinates of the same set of switches. Equation (9) implies

$$\{G_{i+1}'/\zeta_{i+1}'(\nu_h)\}$$

$$= \{n_i(x' - 1) + l_i : l_i = 1, 2, \cdots n_i/\zeta_{i+1}'(\nu_h)\},$$

$$= \left\{n_i\left(u \cdot \prod_{j=2}^{i-1} n_j + d_i' - 1\right) + l_i : l_i = 1, 2, \cdots n_i/\zeta_{i+1}'(\nu_h)\right\},$$

$$= \left\{u \prod_{j=2}^{i} n_j + (d_i' - 1)n_i + l_i : l_i = 1, 2, \cdots n_i/\zeta_{i+1}'(\nu_h)\right\}.$$

Equation (10) implies

$$\{G_{i+1}/\zeta_{i+1}(\nu)\} = \{n_i(x - 1) + l_i : l_i = 1, 2, \cdots n_i/\zeta_{i+1}(\nu)\},$$

$$= \left\{n_i\left(u \cdot \prod_{j=1}^{i-1} n_j + d_i' + (h - 1)\prod_{j=2}^{i-1} n_j - 1\right)\right.$$

$$\left. + l_i : l_i = 1, 2, \cdots n_i/\zeta_{i+1}(\nu)\right\},$$

$$= \left\{u \prod_{j=1}^{i} n_j + (d_i' - 1)n_i + (h - 1)\prod_{j=2}^{i} n_j\right.$$

$$\left. + l_i : l_i = 1, 2, \cdots n_i/\zeta_{i+1}(\nu)\right\}.$$

But from equation (2), if $g_i' = u \prod_{j=2}^{i} n_j + (d_i' - 1)n_i + l_i$, $1 \leq l_i \leq n_i$, then its corresponding g_i is

$$g_i = u \cdot \prod_{j=1}^{i} n_j + (d' - 1)n_i + l_i + (h - 1)\prod_{j=2}^{i} n_j .$$

Therefore $\{G_{i+1}'/\zeta_{i+1}'(\nu_h)\}$ and $\{G_{i+1}/\zeta_{i+1}(\nu)\}$ are clearly coordinates of the same set of switches.

That each switch in $\zeta_1(\nu)$ is linked to each ν_h' by φ_1 is a direct result of

$$\{x(\bmod f_1)/\zeta_1(\nu)\}R\{n_1(x - 1) + l_1 : l_1 = 1, 2, \cdots n_1/\zeta_2(\nu)\}$$

since

$$| \{h(\bmod f_1)\} \cap \{n_1(x - 1) + l_1 : l_1 = 1, 2, \cdots n_1\} | = 1$$

for each $h = 1, 2, \cdots n_1$.

Since $\nu = \zeta_1 \varphi_1 \zeta_2 \cdots \varphi_{2t} \zeta_{2t+1}$ is symmetric with respect to its middle stage, and $\nu_h \, \varepsilon \, B(n_2, n_3, \cdots n_{t+1})$, clearly $\varphi'_k = \varphi'_{2t+1-k}$ for $k = t + 1$, $\cdots 2t - 1$. And finally, again by an argument of symmetry, each switch in ζ_{2t+1} is linked to each ν'_h by φ_{2t}.

REFERENCES

1. Beneš, V. E., *Mathematical Theory of Connecting Networks and Telephone Traffic*, Chapter 3, New York: Academic Press, 1965.
2. Hall, P., "On Representative of Subsets," J. London Math. Soc., *10* (1935), pp. 26–30.

Contributors to This Issue

PAUL T. BRADY, B.E.E., 1958, Rensselaer Polytechnic Institute; M.S.E.E., 1960, Massachusetts Institute of Technology; Ph.D., 1966, New York University; Bell Telephone Laboratories, 1961—. Mr. Brady has worked in modeling on-off speech patterns and speech-level distributions, especially as they occur in two-way conversation over circuits containing voice-operated devices and transmission delay. Member, Acoustical Society of America, Sigma Xi.

MORGAN M. BUCHNER, JR., B.E.S., 1961, and Ph.D., 1965, The Johns Hopkins University; U. S. Army Active Duty, 1966–1968; Bell Telephone Laboratories, 1965–66 and 1968—. At Bell Laboratories, Mr. Buchner initially worked on the design and performance of data transmission systems. He is presently involved in studying the flow of traffic through telephone networks. He is a Supervisor in the Traffic Research Department. Member, IEEE, Tau Beta Pi, Sigma Xi, Eta Kappa Nu.

LEONARD G. COHEN, B.E.E., 1962, City College of New York; Sc.M., 1964, and Ph.D. (Engineering), 1968, Brown University; Bell Telephone Laboratories, 1968—. At Brown University, Mr. Cohen was engaged in research on plasma dynamics. At Bell Telephone Laboratories, he has concentrated on the study of optical transmission techniques. Member, IEEE, Sigma Xi, Tau Beta Pi, Eta Kappa Nu.

W. M. HUBBARD, B.S., 1957, Georgia Institute of Technology; M.S., 1958, University of Illinois; Ph.D., 1963, Georgia Institute of Technology; Bell Telephone Laboratories, 1963—. Mr. Hubbard's work has included analyses related to the design of millimeter-wave solid-state repeaters for use in a waveguide transmission system and the construction of prototype high-speed repeaters for this type of system. He is currently engaged in optical transmission research with emphasis on repeater techniques. Member, Sigma Xi, Tau Beta Pi, Phi Kappa Phi, American Physical Society.

FRANK K. HWANG, B.A., 1960, National Taiwan University; M.B.A., 1964, City College of New York; M.E.S., 1966, and Ph.D., 1968, North Carolina State University; Bell Telephone Laboratories, 1967—. Mr.

Hwang has been working in the area of statistics and combinatorial mathematics, including such problems as sorting, merging and group testing. Currently he is teaching at the National Ching-Hua University, Taiwan, on a six-month leave of absence. Member, A.S.A.

JOHN O. LIMB, B.E.E., 1963, and Ph.D., 1967, University of Western Australia; Bell Telephone Laboratories, 1967—. Mr. Limb has worked on the coding of television signals to reduce channel capacity requirements. He is currently working on methods of reducing frame-to-frame redundancy in moving pictures for *Picturephone®* visual telephone applications.

E. A. J. MARCATILI, Aeronautical Engineer, 1947, and E. E., 1948, University of Cordoba (Argentina); research staff, University of Cordoba, 1947–54; Bell Telephone Laboratories, 1954—. He has been engaged in theory and design of filters in multimode waveguides and in waveguide systems research. More recently he has concentrated on optical transmission media. Fellow, IEEE.

DIETRICH MARCUSE, Diplom Vorpruefung, 1952, Dipl. Phys., 1954, Berlin Free University; D.E.E., 1962, Technische Hochschule, Karlsruhe, Germany; Siemens and Halske (Germany), 1954–57; Bell Telephone Laboratories, 1957—. At Siemens and Halske, Mr. Marcuse was engaged in transmission research, studying coaxial cable and circular waveguide transmission. At Bell Telephone Laboratories, he has been engaged in studies of circular electric waveguides and work on gaseous masers. He spent one year (1966–1967) on leave of absence from Bell Telephone Laboratories at the University of Utah where he wrote a book on quantum electronics. He is presently working on the transmission aspect of a light communications system. Member, IEEE, Optical Society of America.

A. S. MAY, B.S.E.E., 1939, West Virginia University; Bell Telephone Laboratories, 1939–1962; American Telephone and Telegraph Company, 1962—. At Bell Labs. Mr. May was engaged in the design of radar equipment and served as Supervisor in the development of microwave radio-relay systems. At AT&T, he has been engaged in microwave and guided wave planning and in studies of frequency sharing by terrestrial and satellite communications systems. He currently is a Manager in research and development studies.

SCOTTY R. NEAL, B.A. (Mathematics), 1961, M.A. (Mathematics), 1963, and Ph.D. (Mathematics), 1965, University of California, Riverside; Research Mathematician, Naval Weapons Center, China Lake, California, 1964–1967; Bell Telephone Laboratories, 1967—. Since coming to Bell Laboratories, Mr. Neal has been primarily concerned with the analysis of various aspects of telephone traffic systems. He has also worked on applications of optimal linear estimation theory and certain aspects of communication theory. Member, American Math Society, SIAM, Sigma Xi.

THOMAS L. OSBORNE, B.S.E.E., 1961, M.S.E.E., 1963, Auburn University; Bell Telephone Laboratories, 1963—. Mr. Osborne has been involved in research on solid-state microwave radio systems and associated circuits. Member, IEEE, Sigma Xi, Phi Kappa Phi, Tau Beta Pi, Eta Kappa Nu, Pi Mu Epsilon.

MICHAEL J. PAGONES, B.S.E.E., 1961, and M.S.E.E., 1962, University of Illinois; Bell Telephone Laboratories, 1964—. Mr. Pagones has been concerned with analytical studies of satellite and radio-relay systems. At present he is engaged in interference studies of digital modulation. Member, Tau Beta Pi, Phi Kappa Phi, Eta Kappa Nu, IEEE.

R. F. W. PEASE, B.A., 1960, M.A. and Ph.D., 1964, University of Cambridge; Bell Telephone Laboratories, 1967—. Mr. Pease held a faculty appointment at the University of California at Berkeley prior to joining Bell Labs and worked on electron microscopy. He is now trying to efficiently encode moving and still pictures.

S. D. PERSONICK, B.E.E., 1967, City College of New York; S.M. in E.E., 1968, E.E., 1969, and Sc.D., 1969, Massachusetts Institute of Technology; Bell Telephone Laboratories, 1967—. Mr. Personick is engaged in work on wire systems engineering.

GERARD T. WRIXON, B.E., 1961, National University of Ireland, Cork; Royal Netherlands Aircraft Factory, Amsterdam, 1961–1963; M.S., 1964, California Institute of Technology; Ph.D., 1969, University of California, Berkeley; Bell Telephone Laboratories, 1969—. Since coming to Bell Laboratories, Mr. Wrixon has been engaged in antenna studies, atmospheric propagation studies and radio astronomy. Member, IEEE, American Astronomical Society.

B. S. T. J. BRIEF

An Image Band Interpretation of Optical Heterodyne Noise

By S. D. PERSONICK

I. INTRODUCTION

Optical heterodyning and homodyning (heterodyning to baseband) have been studied by various authors.[1-3] The intermediate frequency or baseband signal that is obtained by ideal optical heterodyning or homodyning consists of the signal that would be expected from classical analysis plus an added gaussian quantum noise. The variance of this quantum noise is twice as large in the heterodyning case as it is in the homodyning case. Since classically heterodyning need not have a noise disadvantage over homodyning, this result at first seems puzzling. We shall show below that the disadvantage of heterodyning can be interpreted as an unavoidable "image band" quantum noise. We shall argue this both heuristically and formally using quantum field theory.

II. HEURISTIC ARGUMENTS

Consider the classical heterodyne system. Two radio frequency (RF) bands can contribute to the intermediate frequency (IF) output. One contains the desired signal plus associated noises occupying the signal band. The other is the image band containing image band noises. Classically, to avoid the noises at the IF due to the image band, we filter out the RF image band before it can enter the receiver front end. In the quantum case, we have quantum noise at the IF due to the signal band and the image band. Unlike the classical case, we cannot eliminate the image band quantum noise. Thus heterodyning has twice the quantum noise of homodyning (which has no image band).

III. RIGOROUS RESULTS

We shall now show that the "noisy" quantum heterodyne measurement of the signal on a mode of a quantum field often described in

terms of an overcomplete set of measurement states,[4] can be described in terms of a compatible pair of measurements on the product space of that mode and an extraneous "image mode."

Suppose that the density operator for a single "signal" mode of the optical field in a bounded region of space is (using the notation of R. J. Glauber[5])

$$\rho_s^{(\beta)} = \int \frac{1}{\pi N} \exp \left\{ |(\alpha_s - \beta)|^2 / N \right\} |\alpha_s > < \alpha_s| \, d^2\alpha_s$$

where N is proportional to the thermal noise variance, α_s and β are complex numbers (β represents signal) and $|\alpha_s >$ is an eigenvector of the annihilation operator a_s which along with its adjoint a_s^+ satisfies

$$[a_s, a_s^+] = 1.$$

We could estimate the real or imaginary part of β by measuring

$$M_1 = \frac{a_s + a_s^+}{2}$$

or

$$M_2 = \frac{a_s - a_s^+}{2j} \, ;$$

but they do not commute and cannot be measured simultaneously. Consider next another field mode as "image mode" which does not depend on β and which has arbitrarily small thermal noise, i.e., its density operator is

$$\rho_I = \int \frac{1}{\pi M} \exp \left\{ |\alpha_I|^2 / M \right\} |\alpha_I > < \alpha_I| \, d^2\alpha_I$$

where $|\alpha_I >$ is an eigenvector of the image mode annihilation operator a_I

$$[a_I, a_I^+] = 1.$$

Since the annihilation and creation operators of the signal mode commute with the annihilation and creation operators of the image mode, the following two operators defined on the product space of the two modes commute

$$L_1 = \tfrac{1}{2}(a_s + a_s^+) + \tfrac{1}{2}(a_I + a_I^+),$$

$$L_2 = \frac{1}{2j}(a_s - a_s^+) - \frac{1}{2j}(a_I - a_I^+).$$

The density operator of the product space is simply $\rho_s{}^{(\beta)}\rho_I$ the product of the density operators.

The joint moment generating function of the random variable which results from simultaneous measurement of L_1 and L_2 is

$$M_{12}(u, v) = E\{\exp [uL_1(\text{outcome})] \exp [vL_2(\text{outcome})]\}$$

$$= TR[\rho_s^{(\beta)}\rho_I \exp (uL_1) \exp (vL_2)]$$

using the following properties[6]

$$\exp (A) \exp (B) = \exp (A + B) \exp [\tfrac{1}{2}[A, B]]$$

if

$$[A, [A, B]] = 0$$

and

$$[B, [A, B]] = 0$$

we obtain

$$TR[\rho_s(\beta)\rho_I(\beta) \exp (uL_1) \exp (vL_2)] = M_{12}(u, v)$$

$$= \exp [uRl(\beta)] \exp [v \text{ Im } (\beta)] \exp [\{(u^2 + v^2)/2\}(N/2 + M/2 + \tfrac{1}{2})]$$

$$= \exp [uRl(\beta)] \exp [v \text{ Im } (\beta)] \exp [\{(u^2 + v^2)/2\}(N/2 + \tfrac{1}{2})]$$

$$\text{for} \quad M \ll N$$

which is the result obtained by physical heterodyning[1-3] (note that $M \ll N$ corresponds to filtering image band noise).

If we had measured $M_1 = (a^+ + a)/2$ alone, the outcome random variable would have moment generating function

$$M_3(u) = E \exp [uM_1(\text{outcome})]$$

$$= TR\rho_s(\beta) \exp (uM_1) = \exp (uRl(\beta)) \exp \left[\frac{u^2}{2} (N/2 + \tfrac{1}{4})\right]$$

which is the result obtained by physical homodyning.[1-3]

We see that the homodyne measurement has half the gaussian quantum noise added to the real part of β (which is the outcome mean) as the heterodyne measurement has. However, if we must know both the real and imaginary parts of β, we must heterodyne, since homodyning can only yield one part or the other (unless we split the received signal in half and homodyne each part).

IV. CONCLUSIONS

We see from the above that the fact that heterodyning results in twice the quantum noise of homodyning is not strange at all. The added noise is heuristically and rigorously associated with an image band which can contribute extra noise classically as well.

REFERENCES

1. Oliver, B. M. "Signal to noise ratios in photoelectric mixing," Proc. IRE, *49*, No. 12 (December 1961), pp. 1960–1961, and *50*, No. 6 (June 1962), pp. 1545–1546.
2. Haus, H., and Townes C. H., "Comments on noise in photoelectric mixing," Proc. IRE, *50*, No. 6 (June 1962), pp. 1544–1545.
3. Personick, S., "Efficient Analog Communication Over Quantum Channels," Massachusetts Institute of Technology, R.L.E. Technical Report 477.
4. Gordon, J. P., and Louisell, W. H., "Simultaneous Measurements of Non-commuting Observables," *Physics of Quantum Electronics*, New York: McGraw-Hill, 1966, pp. 833–840.
5. Glauber, R. J., "Coherent and Incoherent States of the Radiation Field," Phys. Rev., *131*, No. 6 (September 1963), pp. 2766–2788.
6. Louisell, W. H., *Radiation and Noise in Quantum Electronics*, New York: McGraw-Hill, 1964.

Erratum

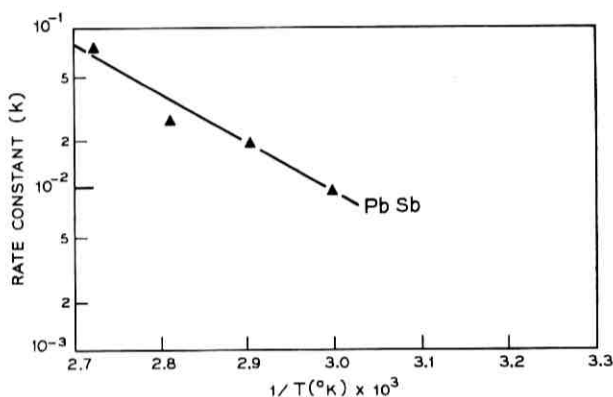The log scale shown on Fig. 14 (p. 1291) for the rate constant (k) is in error by a factor of ten. It should be:

Fig. 14—Rate constant (k) on log scale vs $1/T$ (°K) for 5-5/8" × 5-7/8" × 1/4" thick PbSb grids.

This is a plotting error only and does not alter the data nor the conclusions shown.