THE BELL SYSTEM
TECHNICAL JOURNAL

Bell Laboratories Fiftieth Anniversary

Since its formation in January 1925, Bell Laboratories has taken over the central role of scientific research and technical development from its predecessors in the Bell System. The pages of *The Bell System Technical Journal* have reflected a steady flow of Bell Laboratories technological contributions over the intervening years (see "Fifty Years of B.S.T.J." in the July-August 1972 issue). It is, therefore, particularly fitting that we give some recognition to this half-century anniversary of Bell Laboratories.

With its preeminent contribution to such broad fields as information theory, solid-state technology, and electronic data processing, Bell Laboratories has been the creative force in a vital and closely integrated Bell System team for transforming technical possibilities into economical telecommunications realities. Examples include the development and deployment of today's vast microwave carrier network (now containing over 300 million circuit miles), the burgeoning array of electronic switching systems—providing a new degree of flexibility and reliability in the handling of telephone calls—and the provision and use of an increasing family of versatile, sophisticated, computer-based systems for better network implementation and maintenance as well as for more efficient business operations.

In an era when the Bell System is being attacked for its monolithic structure, while at the same time being generally praised for the quality of its service and its record in keeping costs down, we should carefully note that our spectacular progress is largely a result of the close ties among Bell Laboratories, Western Electric, the Operating Telephone Companies, and the AT&T parent company, and is not an independent phenomenon.

The second half-century of Bell Laboratories scientific and engineering activity is beginning with the same exciting promise as did the first. Innovative research in basic materials and processes flourishes. At the system level, trends in both switching and transmission are pointing to a steadily growing digital network with important advantages in quality and flexibility of service. Giant projects such as toll ESS switching and millimeter waveguide transmission are harbingers of this future. Optical fiber transmission will form the basis for another avenue of important telecommunications advances. Low-cost electronics are extending the fruits of the solid-state revolution to the loop plant. The realization of control and processing functions in ever smaller, cheaper electronic packages opens the door to increased sophistication in system operation while maintaining uncomplicated interfaces with those who use and maintain such systems. An integrated Bell System team is the key to successful transformation of such technological promises into telecommunications service to the nation and, less directly, to the world.

W. E. Danielson
Chairman,
B.S.T.J. Editorial Committee

Mode Coupling in an Optical Fiber With Core Distortions

By D. MARCUSE and H. M. PRESBY

(Manuscript received May 17, 1974)

The variations in the geometry of a step-index optical fiber are deter-mined as functions of position along the axis of the fiber by an analysis of the backscattered light produced when a beam from a CW laser is incident perpendicular to the fiber axis. The power spectrum computed from this distortion function is then utilized with coupled-mode theory to predict the mode coupling, the reduction in pulse dispersion, and the accompany-ing increased radiation loss of the fiber. The theoretical calculations support experimental observations and account for a partial reduction in the multi-mode pulse dispersion.

I. INTRODUCTION

Optical-fiber communication systems utilizing incoherent light sources such as light-emitting diodes require the use of multimode waveguides to insure efficient excitation of the guide. Such guides, however, suffer from multimode pulse dispersion, because modes with higher group velocity arrive at the receiver earlier than modes with lower group velocity, limiting the information-carrying capacity of the fiber.

Thus, in a fiber of length L, uniform core of index n, uniform cladding of index n $(1 - \Delta)$, and constant cross section, a short pulse feeding equal amounts of power to every mode at the input will arrive at the other ends with a width $\tau = n\Delta L/c$, where c is the speed of light in free space.[1] For example, a fiber with $\Delta = 0.01$ and $n_2 = 1.5$ will have a delay spread of $\tau/L = 50$ ns/km, a serious limitation on either high-capacity or long-distance transmission.

Multimode pulse distortion can be reduced by introducing coupling between the guided modes.[2] The reduction in the pulse length comes about because some power traveling in a fast mode is eventually trans-ferred to a slow mode, while power starting out in a slow mode finds

itself at least partially in a fast mode, so that the extremes of the group velocity spread are partly equalized. The root-mean-square delay spread is proportional to the square root of the product of fiber length L and the coupling length L_c associated with steady-state power transfer.[2] The multimode delay spread τ is thus reduced by a factor of $(L_c/L)^{\frac{1}{2}}$ which, in the case of strong coupling, $L_c \ll L$, can be significant.

One coupling mechanism that has been investigated in detail is geometric variations of the fiber along the longitudinal direction z.[3] If the deformation function $f(z)$, which is defined by these variations, is expanded in a Fourier series, two guided modes μ and ν with propagation constants β_μ and β_ν will be coupled by the Fourier component whose spatial frequency is given by[3] $\Omega = \beta_\mu - \beta_\nu$. In other words, the spatial period is the beat wavelength between the μth and νth modes. Using $n_1 = 1.5$ and $\Delta = 0.01$ as before and assuming a signal wavelength of $1.0\ \mu$m and a fiber core $50\ \mu$m wide, the beat wavelength for the adjacent lowest-order modes is 10 mm and for the adjacent modes near cut-off is 0.7 mm. Spatial periods greater than 10 mm will have very little effect on mode mixing and spatial periods less than 0.7 mm will create signal loss by coupling guided modes to the radiation field.

In this paper, the observation is reported of interface irregularities "unintentionally" introduced into a fiber during the pulling process. Distortions that are on the order of several microns are detected and measured by a backscattered light analysis technique. With the distortion function in hand, the power spectrum is computed and utilized to predict mode coupling, reduction in pulse dispersion, and accompanying increased radiation loss of the fiber. This represents the first time that the distortion function of a real fiber has actually been measured and utilized to predict transmission behavior.

II. EXPERIMENT

The fiber studied was pulled from a preform produced by a chemical vapor deposition (CVD) process.[4] The core of the fiber is elliptical, with major and minor axes on the order of $50\ \mu$m and $30\ \mu$m, respectively. The corresponding measurements for the outside fiber dimensions are $118\ \mu$m and $110\ \mu$m. A microinterferogram and a plot of the index profile are shown in Fig. 1. The profile is close to that of a step-index fiber with a very slight modification at the center. The maximum index difference between the core and the cladding is about $n_1 - n_2 = 0.0135$ or $\Delta = 0.0093$. It is important to point out that the refractive index profile did not vary by measurable amounts (less than one part in 10^4) over distances of interest for mode mixing in the fiber (0.2 cm to 2.0
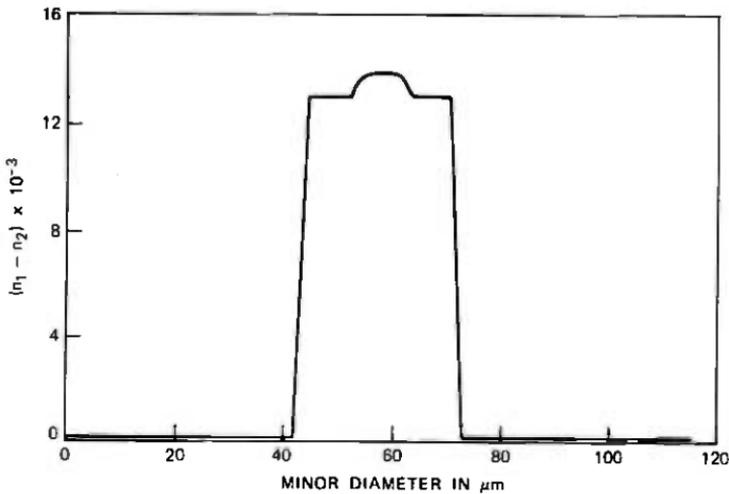
Fig. 1—Microinterferogram and index profile of cvD step-index fiber.

cm), though it did vary from one end of the 1-km fiber to the other.

If the geometry of the fiber remained uniform along its length as in Fig. 1, the fiber would propagate a given set of modes that would not couple among each other, producing a pulse spread of approximately 45 ns/km. The measured pulse width at the 10-dB point, however, was about 10 ns after nearly 1 km of fiber.[5]

To determine the core-cladding interface distortion, which we believe is at least partially responsible for this reduction in pulse dispersion, a backscattered light analysis technique was utilized.[6] This method can be used to detect the parameter b/a, where b is the radius of the fiber core and a is the radius of the cladding, assuming constant indexes of

refraction for the core and the cladding.[7] The technique has been extended to make observations on extended lengths of fiber by the set-up shown in Fig. 2. Light from a cw He-Ne laser is directed to oscillating mirror M_2 by means of fixed mirror M_1. The oscillating mirror serves to transform the ~1-mm circular beam into a line 1-mm wide, with length determined by the amplitude of oscillation. This line impinges upon the fiber, and the backscattered light is detected with photographic film.

A typical backscattered light distribution is shown in Fig. 3. Figure 3a is the overall pattern arising from a 12-cm length of fiber, and Fig. 3b is an expansion of the section on which measurements are made. The magnification in Fig. 3b is approximately 1 to 1. Figure 3a is symmetric about the midpoint with sharp cut-offs at the ends, typical of the backscattered light distribution.[6] As we go from the center of the pattern outward, we observe a region of enhanced fringe intensity. The location of the last fringe in this region is determined by the parameter b/a, assuming constant n_1 and n_2.[7] Here we interpret b and a as an average diameter. If the fiber were of uniform geometry along its length, the fringes would be straight parallel lines; the departure of this fringe from straightness gives a measure of the variation of b/a, and hence the distortion function of the fiber.

Measurements were made on several 14-cm lengths of fiber taken from both ends of the 1-km length. The appearance of the fringe dis-
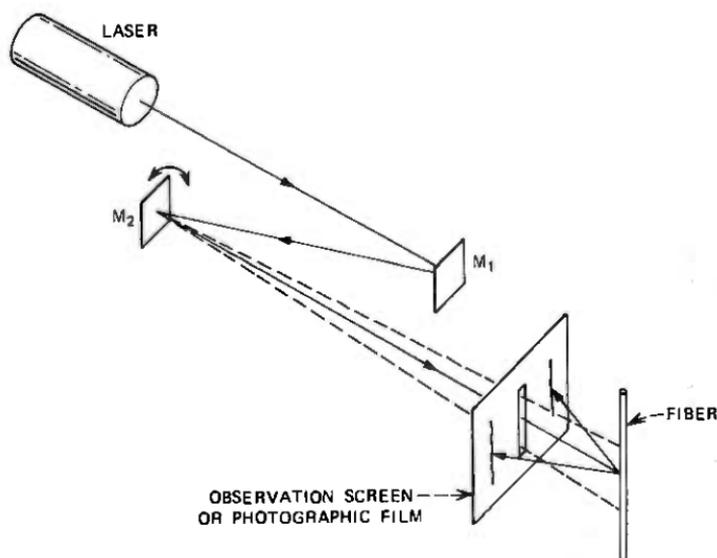

Fig. 2—Optical arrangement to determine core distortion function.

(a)
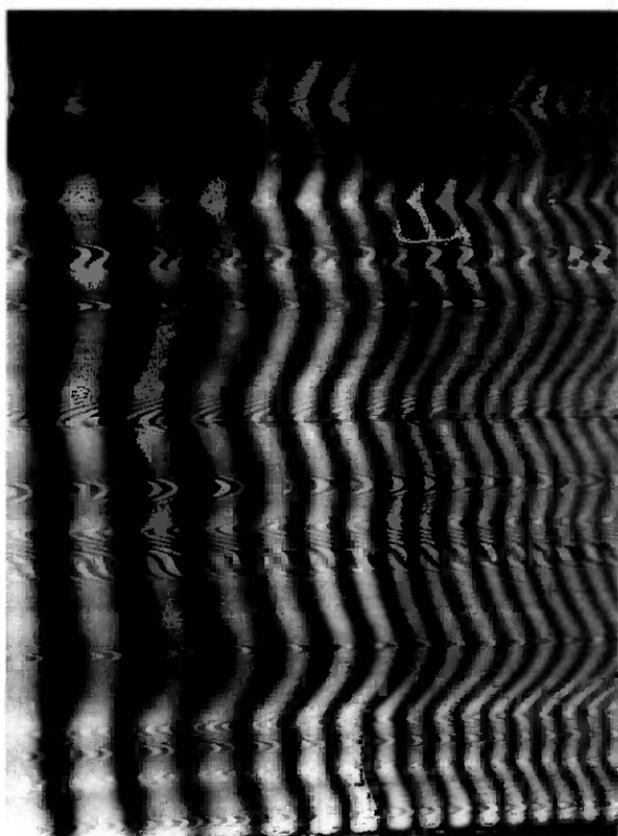
(b)

Fig. 3—Backscattered light patterns. (a) Complete pattern from 12.0-cm section of fiber. Slit in center allows for passage of incident light and bright spots in center are photographic artifacts. (b) Expansion of section of interest.

tortion was similar visually for other sections taken from intermediate portions of the fiber, but detailed measurements were not made on them. Figure 4 is a representative curve of the core distortion. The curve shows $\Delta b = f(z)$ plotted versus position along the fiber axis, with measurements made every 2 mm. The $f(z)$ variations are on the order of several percent.

We believe that the core-cladding interface distortions observed were introduced into the fiber during the pulling process. The preform was reduced to fiber form in an oxyhydrogen flame-pulling apparatus. Either the instability of the flame as a heat source or slight movements of the fiber because of air currents and back-flame effects or combinations of these are probably the contributing factors. Nonuniform pulling speeds and varying preform geometry may be additional factors, but they are expected to introduce variations with a much longer periodicity.

III. ANALYSIS

Given the core distortion function, the Fourier transform is then obtained by computer. A typical transform plot is shown as a bar
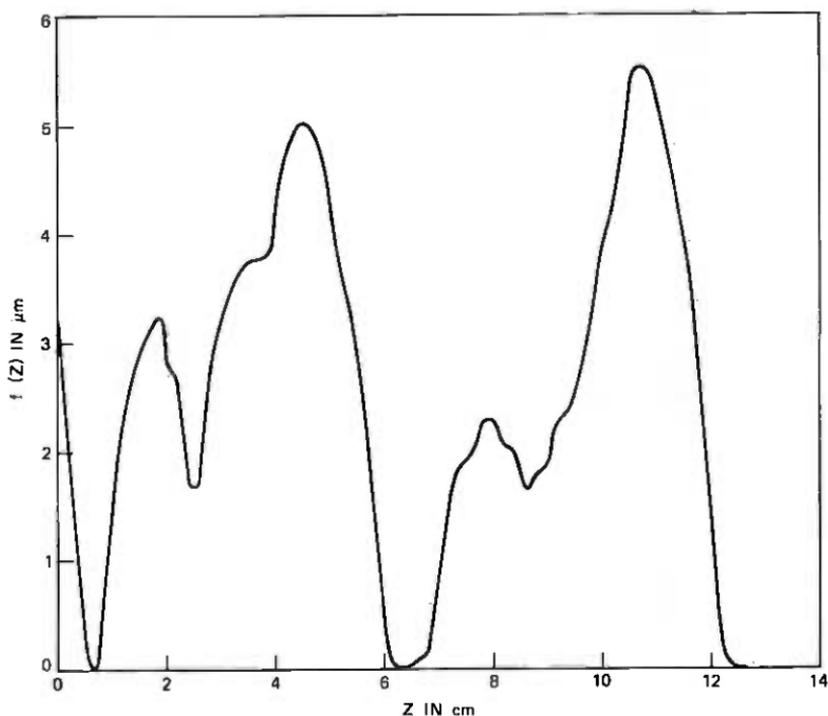

Fig. 4—Representative core distortion function.

graph in Fig. 5. We see that significant components exist in the coupling regions of interest. It is our aim to relate this spatial Fourier spectrum of the distortion function $f(z)$ to the pulse width reduction caused by mode coupling. The theory is described in detail in Ref. 8 for a fiber whose core maintains its circular cross section but is randomly bent. The core of the fiber examined in this paper has an elliptical cross section, and the observed variations consist of changes in the ellipticity. Modes of the elliptical fiber are very complicated and are difficult to apply to a mode-mixing analysis. For this reason, we simplify the problem by assuming that the fiber core nominally has a circular cross section that deforms itself randomly into an ellipse. It is assumed that this model is capable of yielding order-of-magnitude estimates of the performance of the actual fiber.

The core-cladding boundary r of a fiber with elliptical deformations can be described by the function[9]

$$r = b + f(z) \cos 2\phi. \tag{1}$$

The constant radius of the perfect fiber is b, $f(z)$ is the distortion function shown in Fig. 4 whose Fourier power spectrum is shown by the vertical bars in Fig. 5, and ϕ is the angle of the cylindrical coordinate system. A fiber with the core-cladding deformation of (1) couples modes according to the selection rule $\Delta \nu = \pm 2$.[8] The label ν, indicating the azimuthal mode number, enters the field expressions[3] via $\cos \nu\phi$ or $\sin \nu\phi$.
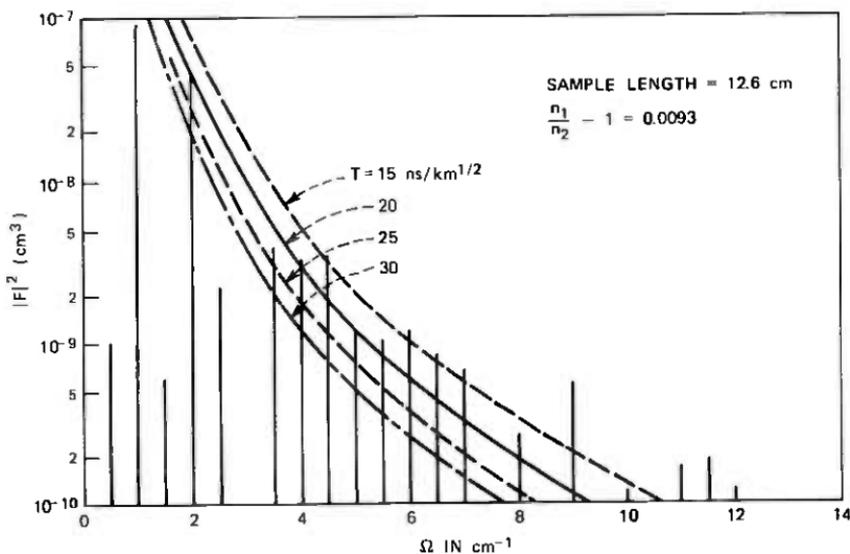

Fig. 5—Fourier transform of core distortion function.

If the power spectrum of $f(z)$ drops off sufficiently rapidly to limit mode coupling to nearest neighbors, the coupled power equations reduce to the form[10]

$$\frac{\partial P_{\nu,m}}{\partial z} = h_{\nu+2,m;\nu,m}(P_{\nu+2,m} - P_{\nu,m}) - h_{\nu,m;\nu-2,m}(P_{\nu,m} - P_{\nu-2,m})$$
$$+ h_{\nu+2,m-1;\nu,m}(P_{\nu+2,m-1} - P_{\nu,m}) - h_{\nu,m;\nu-2,m+1}(P_{\nu m} - P_{\nu-2,m+1}). \quad (2)$$

$P_{\nu,m}$ is the average power carried by the mode labeled ν, m. In a crude approximation (applicable for small ν values and modes far from cutoff), the transverse propagation constant can be expressed as[11]

$$\kappa_{\nu,m} = (\nu + 2m) \frac{\pi}{2b}. \quad (3)$$

In addition to the already mentioned azimuthal mode number ν, we have also introduced the radial mode number m. Together, these two parameters form the compound mode number

$$M = \nu + 2m. \quad (4)$$

The expression in the second line on the right-hand side of (2) describes coupling among modes whose compound mode numbers differ by $\Delta M = 0$. In the crude model indicated by the approximation (3), modes with the same value of M have the same propagation constant and the same group velocity. This is not strictly true, since the degeneracy does not exist in the exact theory. However, the modes with the same M values are coupled very strongly, since coupling among them is caused by the large amplitudes at low spatial frequencies seen in Fig. 5. It is thus reasonable to assume that mode mixing among modes with identical values of M is so rapid that these modes carry equal amounts of power. This assumption causes the differences of the power with equal M values to disappear so that only the first line on the right-hand side of (2) remains. This coupling process is implicitly taken care of by requiring that modes with equal values of M carry identical amounts of power. Using the compound mode number, we write

$$P_{\nu,m} = P_M \quad P_{\nu+2,m} = P_{M+2} \quad P_{\nu-2,m} = P_{M-2} \quad (5)$$
$$h_{\nu+2,m;\nu m} = h_{M+2,M} \quad h_{\nu m;\nu-2,m} = h_{M,M-2}. \quad (6)$$

Equation (2) now assumes the form

$$\frac{\partial P_M}{\partial z} = h_{M+2,M}(P_{M+2} - P_M) - h_{M,M-2}(P_M - P_{M-2}). \quad (7)$$

We proceed by adding all equations of type (7) with the same value of M.[12] However, in addition to this simple summation we use the fact

that there are M modes with mode number M, but $M + 2$ modes with number $M + 2$. By replacing the differences occurring in (7) by differentials, we obtain

$$M \frac{\partial P_M}{\partial z} = 2(M + 2)h_{M+2,M} \frac{\partial P_{M+2}}{\partial M} - 2Mh_{M,M-2} \frac{\partial P_M}{\partial M}. \quad (8)$$

The replacement of the difference terms with derivatives is an approximation that is valid if there are very many guided modes, so that the M numbers are large and can be regarded as a quasi-continuum. With this approximation, (8) assumes the form

$$M \frac{\partial P_M}{\partial z} = 4 \frac{\partial}{\partial M} \left(M h_M \frac{\partial P_M}{\partial M} \right). \quad (9)$$

We can express the compound mode number in terms of the cone angle θ, at which modes with the same values of M appear in the far-field radiation pattern that escapes from the end of the fiber.[3,12] Using[13]

$$M = \frac{2kb}{\pi} \theta \quad (10)$$

and

$$\frac{\partial \theta}{\partial M} = \frac{\pi}{2kb}, \quad (11)$$

with $k = $ free space propagation constant and $b = $ core radius, we obtain finally from (9) with $h_M = h = h(\theta)$

$$\frac{\partial P(z, \theta)}{\partial z} = \left(\frac{\pi}{kb} \right)^2 \frac{1}{\theta} \frac{\partial}{\partial \theta} \left(\theta h \frac{\partial P}{\partial \theta} \right). \quad (12)$$

The propagation of pulses is described by the time-dependent partial-differential equation

$$\frac{\partial P}{\partial z} + \frac{1}{v} \frac{\partial P}{\partial t} = \left(\frac{\pi}{kb} \right)^2 \frac{1}{\theta} \frac{\partial}{\partial \theta} \left(\theta h \frac{\partial P}{\partial \theta} \right), \quad (13)$$

where $v = v(\theta)$ is the group velocity of the mode labeled θ. Solutions to this equation are expressed in the form

$$P(z, t, \theta) = \sum_{j=1}^{N} \int_{-\infty}^{\infty} c_j(\omega)B(\omega, \theta)e^{-\rho^{(j)}(\omega)z}e^{i\omega t}d\omega. \quad (14)$$

The function $B_j(\omega, \theta)$ and the parameter $\rho^{(j)}(\omega)$ are obtained as solutions of an eigenvalue problem.[3] The eigenvalue $\rho^{(j)}$ is expanded into a perturbation series

$$\rho^{(j)}(\omega) = \sigma^{(j)} + \frac{i\omega}{v_a} + i\omega\rho_1^{(j)} + \omega^2\rho_2^{(j)} + \cdots. \quad (15)$$

The second-order perturbation of the first eigenvalue, $\rho_2^{(1)}$, determines the width of the equilibrium pulse via the formula[14]

$$T = 4\sqrt{\rho_2^{(1)}L}, \tag{16}$$

where T is the full width of the gaussian-shaped impulse response of the multimode fiber measured between the $1/e$ points and L is the length of the fiber. It was assumed that the coupling coefficient h is independent of θ. This assumption corresponds to a spatial Fourier power spectrum of the distortion function $f(z)$ of the form

$$\langle |F(\theta)|^2 \rangle = \frac{C}{\Omega^4}. \tag{17}$$

Numerical differences between our present case and the example treated in Ref. 3 arise from the fact that the mode spacing (in β space) of nearest neighbors coupled by elliptical core deformations is twice as large as the spacing between nearest neighbors coupled by random bends. Taking these differences into account, we arrive at the following formula for the width of the equilibrium pulse:

$$T = \frac{1.26n_2[(n_1/n_2) - 1]^{\frac{3}{2}}}{c\sqrt{C}} \sqrt{L}, \tag{18}$$

in which n_1 and n_2 are the refractive indices of core and cladding, c is the velocity of light in vacuum, C is the constant defined in (17), and L is the length of the fiber.

For comparison with the length τ of a pulse carried by uncoupled modes, as mentioned earlier,

$$\tau = \frac{L}{c} n_2\Delta, \tag{19}$$

we use the "improvement factor,"

$$R = \frac{T}{\tau} = \frac{1.26\Delta^{\frac{1}{2}}}{\sqrt{CL}}, \tag{20}$$

with $\Delta = n_1/n_2 - 1$. The numerical values of R are physically meaningful only if $R < 1$. If $R > 1$, the guide length L or the coupling strength are too small for an equilibrium pulse to have established itself. The relative decrease of the pulse width achieved by mode coupling improves with increasing fiber length.

IV. DISCUSSION OF EXPERIMENTAL RESULTS

The step-index fiber used for this study exhibited unintentional, random fluctuations of the core-cladding boundary resulting in mode

coupling. According to (19), an uncoupled pulse width of $\tau = 45$ ns/km should have been observed for $\Delta = 9.3 \times 10^{-3}$; the actually observed pulse width was $\tau = 10$ ns/km. The observed core-cladding irregularities are able to account for a substantial part of the pulse shortening.

The function $f(z)$ of (1) is shown in Fig. 4 for a 12.6-cm section of the fiber. We consider the distortion of this fiber section as typical and representative of this particular fiber. The absolute square values of the Fourier components of the function shown in Fig. 4 are plotted in the bar graph of Fig. 5. It is clear that Fig. 5 does not represent the spatial power spectrum of the distortion function that enters the coupled-mode theory. The required power spectrum would have to be obtained by computing the Fourier spectrum of the function $f(z)$ for the entire length of the fiber. An approximation of this function could be obtained by computing Fourier spectra for a large number of shorter fiber sections and averaging. This procedure is time-consuming if it must be done manually, and an automated process would be required to determine the spatial Fourier spectra of the core-cladding interface distortion of the fiber.

In the absence of more information, we used the Fourier power spectrum shown in Fig. 5 to extract information about the mode coupling process. Since a theory based on a fourth-power law [see (17)] was already available, we approximated the data in Fig. 5 by fourth-power-law curves. The four curves plotted in the figure are possible approximations that may be roughly guessed from the bar graph. The numbers used to label the curves, $T = 15, 20, 25,$ and 30 ns/km$^{\frac{1}{2}}$, are the result of determining the constant C of (17) from the curves and using it to calculate the width of the equilibrium pulse according to (18). Which of the four curves is the most plausible approximation to the actual Fourier power spectrum is open to discussion. It appears to us that the curves between $T = 20$ and $T = 25$ ns/km$^{\frac{1}{2}}$ seem to approximate the power spectrum reasonably well. The Fourier components for high spatial frequencies tend to exceed the values of the curves, while the curves are a little high for small spatial frequencies. Perhaps the fourth-order power law is not the best approximation to the Fourier power spectrum. On the other hand, we do not have enough information to obtain an accurate power spectrum. The high spatial frequencies shown in the bar graph result from rapid fluctuations of the curve $f(z)$ that are partly noise of the measurement process. It can thus be expected that the amplitudes of the high-frequency components appear exaggerated.

The most important point of this discussion is the observation that the measured core boundary distortions are indeed of the right order of magnitude to help explain the observed pulse shortening. Our data

can easily account for an improvement of the pulse width by $R = 0.5$. Since an improvement of $R = 0.2$ has been observed, it appears clear that the core boundary fluctuations are an important contributor to the observed pulse coupling behavior. The remaining amount of coupling may easily have been contributed by the bends of the fiber axis caused by the support mechanism and/or tension on the drum.[15]

Mode coupling with a fourth-power Fourier spectrum of (17) results in a loss penalty of[16]

$$R^2 \sigma^{(1)} L = 0.5 \text{ dB}. \tag{21}$$

$\sigma^{(1)}$ is the steady-state loss coefficient resulting from mode mixing. For an $R = 0.5$, we thus expect an additional fiber loss of $\sigma^{(1)} L = 2$ dB. The observed pulse width improvement of $R = 0.2$ results in a loss penalty of $\sigma^{(1)} L = 10$ dB. The fiber losses actually observed are $\alpha = 30$ dB/km at a wavelength of 0.9μm. The difference between these loss values is attributable to absorption losses in the fiber material.

The improvement factor R of (20) can be expressed in terms of a coupling length L_c,[2]

$$R = \left(\frac{L_c}{L} \right)^{\frac{1}{4}}. \tag{22}$$

For $R = 0.5$ we have $L_c = L/4$ or $L_c = 250$ m, since $L = 1$ km was assumed.

V. CONCLUSIONS

Mode coupling in multimode fibers may be caused by a number of fiber irregularities. Random index fluctuations, random bends, and core-cladding interface deformations are the most likely candidates. In this paper we have considered mode coupling by core-cladding interface deformations observed by a light-scattering technique that extracts the necessary information from the backscattered light of a laser that impinges on the fiber at right angles to its axis. The measured information was used to estimate the amount of pulse shortening that might be caused by this coupling mechanism. We found that the observed magnitude of the core-cladding boundary irregularities can explain some observed pulse shortening. The remainder may be caused by random bends introduced by the surface roughness of the drum on which the fiber is supported, or by refractive index changes along the fiber.

It appears that the backscattering technique used for the determination of the core-cladding interface irregularities may be a useful tool not only for monitoring the precision of fiber drawing processes but also for predicting the amount of mode coupling and consequently the pulse performance of multimode optical fibers. As shown here, to a

good degree this information can be extracted from a fiber only 12.6 cm long.

REFERENCES

1. S. E. Miller, E. A. J. Marcatili, and Tingye Li, "Research Towards Optical Fiber Transmission Systems," Proc. IEEE, *61*, No. 12 (December 1973), pp. 1703–1751.
2. S. D. Personick, "Time Dispersion in Dielectric Waveguides," B.S.T.J., *50*, No. 3 (March 1971), pp. 843–859.
3. D. Marcuse, "Theory of Dielectric Optical Waveguides," New York: Academic Press, 1974.
4. J. B. MacChesney, P. B. O'Connor, J. Simpson, and F. V. DiMarcello, "Multimode Optical Wave Guides Having a Vapor Deposited Core of Germania Doped Borosilicate Glass," Ceramic Bulletin Paper 31-6-73F, *52*, No. 9 (September 1973), p. 704.
5. L. G. Cohen, unpublished data.
6. H. M. Presby, "Refractive Index and Diameter Measurements of Unclad Optical Fibers," J. Opt. Soc. Am. *64*, No. 3 (March 1974), pp. 280–284.
7. H. M. Presby and D. Marcuse, "Refractive Index and Diameter Measurements of Step-Index Optical Fibers and Preforms," to be published in Appl. Opt., Dec. 1974.
8. Ref. 3, Section 5.6.
9. Ref. 3, Eq. (3.6-1).
10. Ref. 3, Eq. (5.2-19).
11. Ref. 3, Eqs. (2.2-10) and (2.2.71).
12. D. Gloge, "Optical Power Flow in Multimode Fibers," B.S.T.J., *51*, No. 8 (October 1972), pp. 1767–1783.
13. Ref. 3, Eq. (5.6-12).
14. Ref. 3, Eq. (5.5-25).
15. D. B. Keck, "Observation of Externally Controlled Mode Coupling in Optical Waveguide," Proc. IEEE, *62*, No. 5 (May 1974), pp. 649–650.
16. Ref. 3, Eq. (5.6-66).

Quasi-Ray Analysis of Crosstalk Between Multimode Optical Fibers

By A. H. CHERIN and E. J. MURPHY

(Manuscript received April 3, 1974)

A model based on frustrated total reflection of waves in a multilayered medium has been developed to analyze the crosstalk between multimode optical fibers. Kappa (κ), the parameter indicating the power distribution among the modes of the fiber, and fiber cladding thickness play very important roles in determining the crosstalk isolation between fibers. Significant but less dominant effects on crosstalk are due to variations in fiber numerical aperture, length, and transmitting wavelength.

I. INTRODUCTION

Crosstalk between communications circuits has long been a problem concerning engineers and designers in the telecommunications industry. In a digital communications system using optical fibers as a transmission medium, at least a 30-dB signal-to-crosstalk ratio will be needed. In those applications where a lossy jacket around the fiber is undesirable, proper design of optical fiber cables and circuits requires an understanding of the parameters that control crosstalk between the fibers.

A number of models can be found in the literature to describe crosstalk between optical fibers.[1-8] Maxwell's equations are the usual starting point for these models, and a field theory approach describing the coupling coefficients between individual modes results. Each of these models provides insight into a possible mechanism for describing crosstalk between optical fibers, and each is operationally useful for calculating crosstalk between single-mode fibers or guides with a small number of propagating modes. A very interesting model describing crosstalk resulting from scattering from a rough core-cladding interface is also described in the literature.[9]

In this paper, a meridional quasi-ray tracing procedure is used to describe crosstalk between highly multimoded optical fibers. This approach is an extension of work developed by H. P. Yuen[10] and N. S. Kapany.[11-14] The mechanism for crosstalk coupling of energy between fibers is that of frustrated total internal reflection of waves in a multi-

layered medium. Integral expressions for crosstalk and transmitted power are developed in terms of the geometry of the system, the absorption loss of the fiber cores, the transmission and reflection coefficients at the core-cladding interface, and the energy distribution at the launching end of the fiber. It is assumed in this work that the propagating modes within a fiber are uncoupled. The error resulting from this assumption should be small when the excitation is chosen to coincide with the equilibrium energy distribution and the coupling between the fibers is weak. A computer evaluation has been made of the integral expressions for crosstalk and transmitted power. Results of a study showing the functional relationship between far-end equal-level crosstalk (FEXT) and cladding thickness, mode energy distribution, length, numerical aperture, and wavelength are included in this paper, along with a discussion of future work in this area.

II. DERIVATION OF GENERAL TRANSMISSION AND CROSSTALK FORMULAS

For the system of fibers shown in Fig. 1, we present a general derivation of the transmission of energy within a fiber and the crosstalk behavior between fibers. We assume that each fiber in an assembly is excited by a source that focuses its power on the center of the entrance end of a fiber, exciting meridional rays as shown in Fig. 2. The ray incident at an angle θ_o to the axis of the fiber is refracted into the fiber at an angle θ which can be simply related to θ_o by the law of refraction. Consider a distribution of input radiant intensity or input power per unit solid angle $F_o(\theta, \phi)$ for some polar angle defined on the entrance surface. A certain fraction of the power, $T_o(\theta, \phi)$, will be transmitted into the fiber because of refraction at the entrance surface. Within the incremental solid angle, $\sin \theta d\theta d\phi$, the power coupled into the fiber, $dP(\theta, \phi)$, can be written as

$$dP(\theta, \phi) = F(\theta, \phi) \sin \theta d\theta d\phi = F_o(\theta, \phi) T_o(\theta, \phi) \sin \theta d\theta d\phi \qquad (1)$$

or

$$\left(\frac{dP(\theta, \phi)}{d\Omega} \right)_{\text{entrance}} = F(\theta, \phi). \qquad (1a)$$

We assume here that an absorption coefficient α per unit length is defined for the fibers. The coefficient α takes into account both bulk absorption loss and scattering loss.

The power flux just prior to hitting the core-cladding interface of the fiber for the first time is given by

$$\left(\frac{dP}{d\Omega} \right)_{P_1} = \epsilon^{-\alpha(d/2)\csc\theta} F(\theta, \phi), \qquad (2)$$

where d is the fiber core diameter.
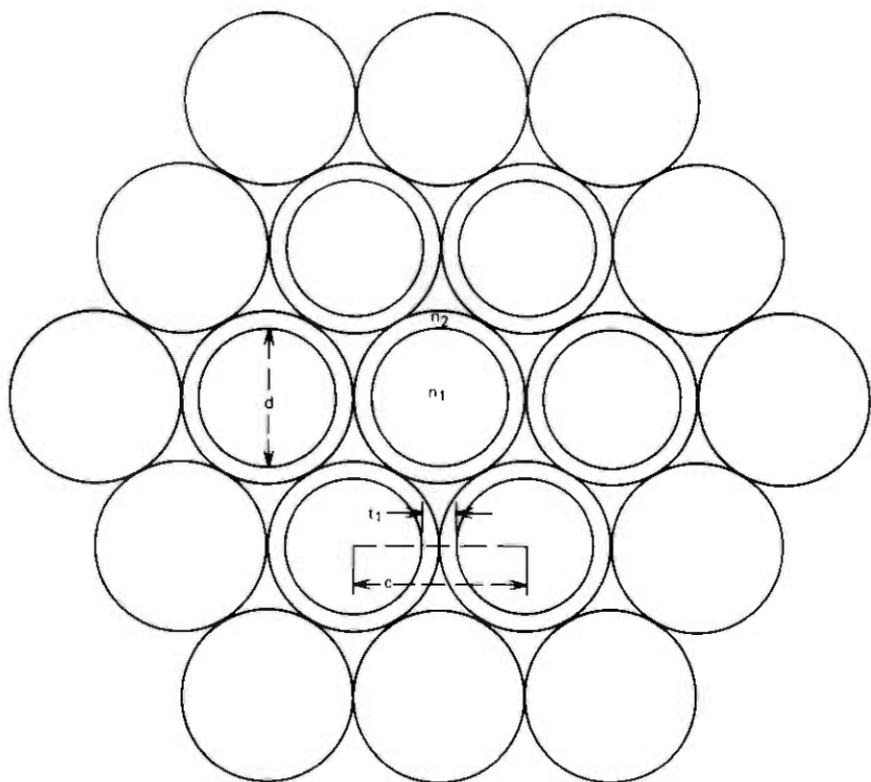
Fig. 1—System of optical fibers.

The power reflected from the first reflection at the core-cladding interface can then be written as

$$\left(\frac{dP}{d\Omega}\right)_{R_1} = \epsilon^{-\alpha(d/2)\csc\theta} R(\theta, \phi) F(\theta, \phi). \qquad (3)$$

The power transmitted into a neighboring fiber from the first reflection at the interface is given by:

$$\left(\frac{dP}{d\Omega}\right)_{T_1} = \epsilon^{-\alpha(d/2)\csc\theta} T(\theta, \phi) F(\theta, \phi). \qquad (4)$$
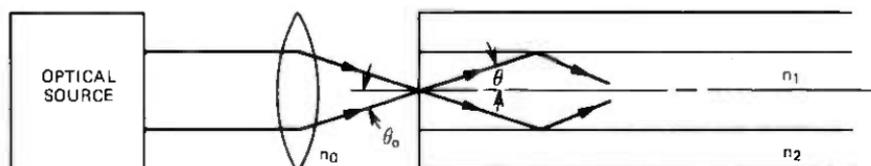

Fig. 2—Meridional ray fiber excitation.
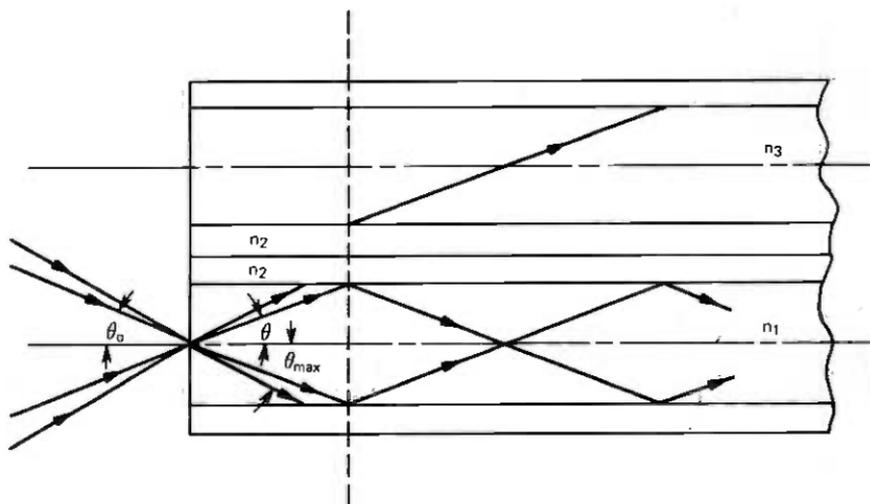
Fig. 3—Ray tracing of transmitted and crosstalk rays in adjacent optical fibers.

The next section shows how $R(\theta, \phi)$, the reflection coefficient, and $T(\theta, \phi)$, the transmission coefficient, are calculated. $T(\theta, \phi)$ in this study is, in effect, the mechanism for obtaining crosstalk between fibers. Crosstalk is shown to be caused by frustrated total reflection of plane waves at the interface of a multilayered medium. The reflected rays continue their propagation down the fiber until they hit the wall again. A fraction of energy is then tunneled as crosstalk to the neighboring fibers, and another fraction is reflected. This process is repeated until the rays reach the exit end of the fiber, as illustrated in Fig. 3. Let us make the following definitions:

$$\left(\frac{dP}{d\Omega} \right)_{P_n} \equiv \text{incident power density at the } n\text{th hit}$$

$$\left(\frac{dP}{d\Omega} \right)_{T_n} \equiv \text{transmitted power density at the } n\text{th hit}$$

$$\left(\frac{dP}{d\Omega} \right)_{R_n} \equiv \text{reflected power density at the } n\text{th hit}.$$

The following power expressions for the nth hit clearly hold.

$$\left(\frac{dP}{d\Omega} \right)_{T_n} = \left(\frac{dP}{d\Omega} \right)_{P_n} T(\theta; \phi) \tag{5}$$

$$\left(\frac{dP}{d\Omega} \right)_{R_n} = \left(\frac{dP}{d\Omega} \right)_{P_n} R(\theta, \phi) \tag{6}$$

$$\left(\frac{dP}{d\Omega} \right)_{P_{n+1}} = \left(\frac{dP}{d\Omega} \right)_{R_n} \epsilon^{-\alpha d \csc \theta}. \tag{7}$$

For a ray of angle θ relative to the fiber axis, the total number of hits M is the largest integer smaller than $(L/d)(\tan \theta) + \frac{1}{2}$.

$$M = \left[\frac{L}{d} \tan \theta + \frac{1}{2} \right] \qquad (8)$$

and the total path length is $L \sec \theta$. The output power distribution at the exit end is therefore

$$\left(\frac{dP}{d\Omega} \right)_{\text{output}} = \epsilon^{-\alpha L \sec \theta} R^M(\theta, \phi) F(\theta, \phi), \qquad (9)$$

so that the total output power of the fiber is

$$P_{\text{output}} = \int_o^{2\pi} \int_o^{\theta_{\text{max}}} \sin \theta F(\theta, \phi) R^M(\theta, \phi) \epsilon^{-\alpha L \sec \theta} T_o(\theta, \phi) d\theta d\phi. \qquad (10)$$

If the exit end of the fiber is not matched to the surrounding medium, there is a transmission factor, $T_o(\theta, \phi) \neq 1$. From eqs. (2), (5), (6), and (7), we obtain

$$\left(\frac{dP}{d\Omega} \right)_{T_n} = \epsilon^{-\alpha d(n-\frac{1}{2}) \csc \theta} [R(\theta, \phi)]^{n-1} T(\theta, \phi) F(\theta, \phi). \qquad (11)$$

Let each crosstalk ray at the nth hit suffer a further attenuation $A'(\theta, \phi, L, n)$ before it reaches the exit end. The total crosstalk output power is then given by

$$P_{xt} = \int_o^{2\pi} \int_o^{\theta_{\text{max}}} \sin \theta F(\theta, \phi) T(\theta, \phi)$$
$$\times \sum_{n=1}^{M} [R(\theta, \phi)]^{n-1} A'(\theta, \phi, L, n) \epsilon^{-\alpha d(n-\frac{1}{2}) \csc \theta} d\theta d\phi. \qquad (12)$$

Note that M is a function of θ. Equations (10) and (12) provide the general formula for signal and crosstalk output power at the end of the fiber as a function of the system parameters. For a specific model of the crosstalk transfer, it is necessary to provide explicit expressions for $F(\theta, \phi)$, $R(\theta, \phi)$, $T(\theta, \phi)$, and $A'(\theta, \phi, L, n)$. A very simple model is discussed later in this paper.

That quantity of particular interest in system design is the signal-to-crosstalk power ratio defined by

$$\left(\frac{S}{N} \right)_{\text{XT}} = \frac{P_{\text{output}}}{P_{\text{XT}}}. \qquad (13)$$

Far-end equal-level crosstalk in decibels is then defined as

$$\text{FEXT} = 10 \log_{10} \frac{P_{\text{output}}}{P_{\text{XT}}}. \qquad (14)$$

III. FRUSTRATED TOTAL INTERNAL REFLECTION IN A MULTILAYERED MEDIUM—A CROSSTALK MECHANISM

In this section, a discussion is given of frustrated total reflection of plane waves in a multilayer medium. This mechanism is the one responsible for crosstalk in our model and serves as the basis for our crosstalk power transmission calculation.

Consider the usual representation of total internal reflection shown in Fig. 4. The ray is completely reflected when ψ exceeds the critical angle, $\psi_c = \sin^{-1} (n_2/n_1)$. The field amplitude decays exponentially in the optically rarer medium. This picture is quite different when we look at the energy flow of the waves more closely.[15–17] The Poynting vector in the rarer medium is by no means zero; only its time average vanishes. This energy flow is depicted in Fig. 5, which also shows that a displacement Δ exists, because of the Goos-Hänchen shift,[15,16] between the incident and reflected beam. For plane waves, this shift is readily calculated.[15] It increases with decreasing angle of incidence ψ, as expected. The Poynting vector also makes a deeper penetration when ψ is smaller.

With this picture we can see that, if a third medium of refractive index, n_1, is brought into close proximity to the first surface from below, as shown in Fig. 6, some energy will be trapped by it as the incident rays make their penetration. This trapped energy will then propagate
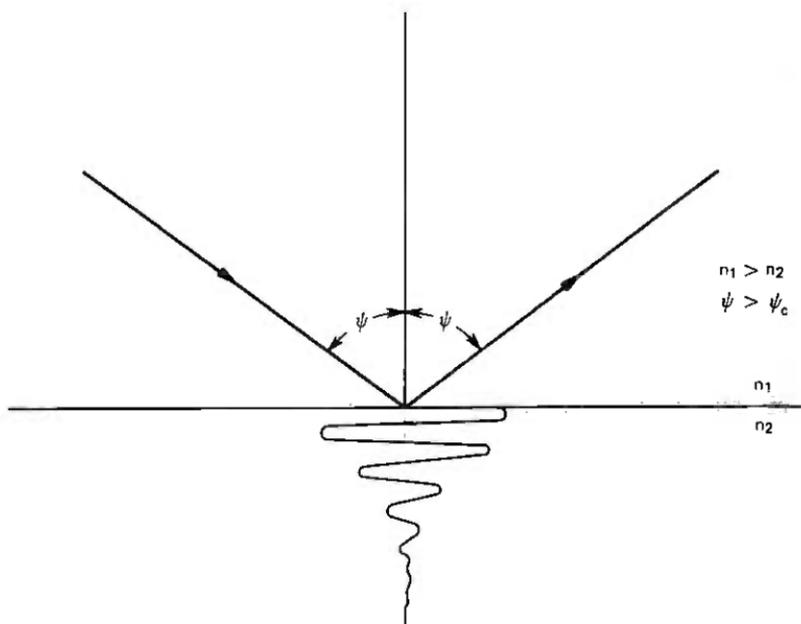

Fig. 4—Total internal reflection from plane dielectric boundary.
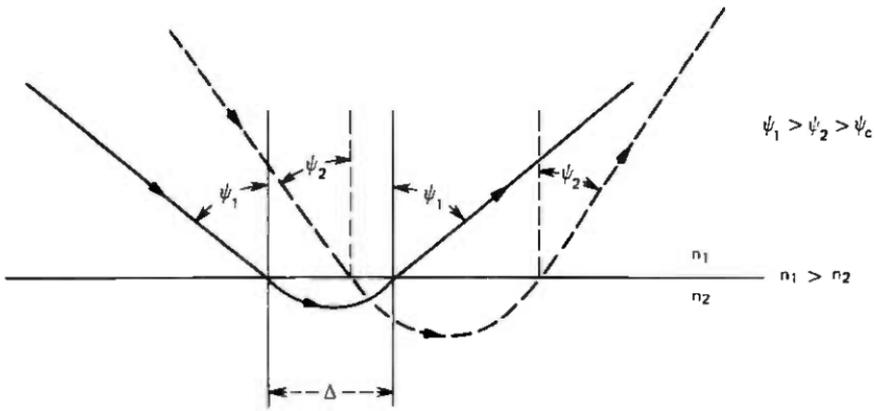
Fig. 5—Displacement of ray—the Goos-Hänchen shift.

into the third medium. This phenomenon of energy transmission through a lower refractive index slab for incidence greater than the critical angle is called frustrated total internal reflection. It is a wave phenomenon outside the domain of pure geometrical optics and is exactly analogous to the tunneling of quantum mechanical particles through a potential barrier, which is classically forbidden.

Consider now three homogeneous media that might represent the core, claddings, and core of adjacent fibers and model them by two half spaces of refractive index, n_1, separated by a third medium, n_2, of thickness, t. We have three homogeneous media for which the transmission coefficient can be obtained by matching boundary conditions directly. The result, for plane wave incidence, is well known in the
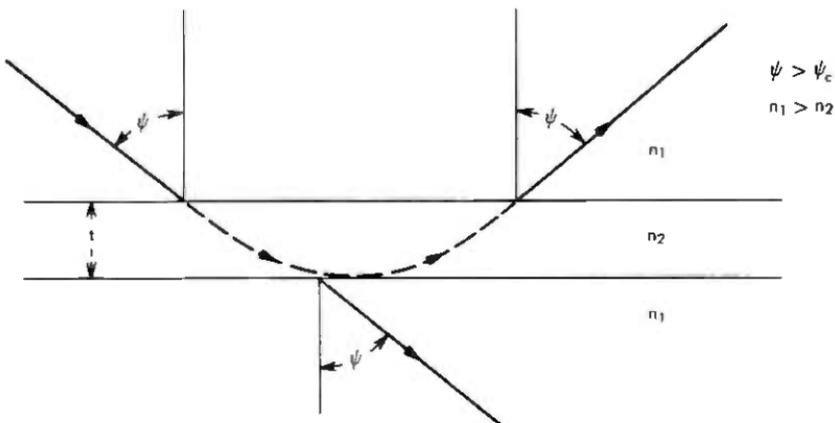

Fig. 6—Frustrated total internal reflection in a multilayered dielectric medium.

literature[15] and has been derived in the appendix for the N media case and specialized to the three-media case discussed here. We assume that, in the vicinity of the core-cladding interface, the signal can be represented as a plane wave and quote the result for the lossless case for the geometry and notation shown in Fig. 7.

Let

$$\gamma = \left[\left(\frac{n_1}{n_2} \right)^2 \sin^2 \theta_1 - 1 \right]^{\frac{1}{2}}. \tag{15}$$

The transmission coefficient, $T(\theta, t)$, which can be defined as the ratio of the transmitted power flux to the incident power flux, is given by

$$T(\theta, \phi) = \left[A_1 \frac{1}{\cosh^2 \beta + \left[(n_2\gamma^2 - n_1 \cos^2 \theta_1)/2n_1n_2 \cos \theta_1\gamma \right]^2 \sinh^2 \beta} \right]$$
$$+ A_2 \left[\frac{1}{\cosh^2 \beta + \left[(n_2^2 \cos^2 \theta_1 - \gamma 2n_1)/2 \cos \theta_1 n_1 n_2\gamma \right]^2 \sin^2 h\beta} \right], \tag{16}$$

where

$A_1 =$ the fraction of incident power polarized perpendicular to the plane of incidence.

$A_2 =$ the fraction of incident power polarized parallel to the plane of incidence.

$$\beta = \frac{2\pi}{\lambda} n_2 t\gamma. \tag{17}$$

Formula (16) is written in terms of θ_o, the incident angle at the input of the fiber, and is used in the calculation of crosstalk power later in this paper.
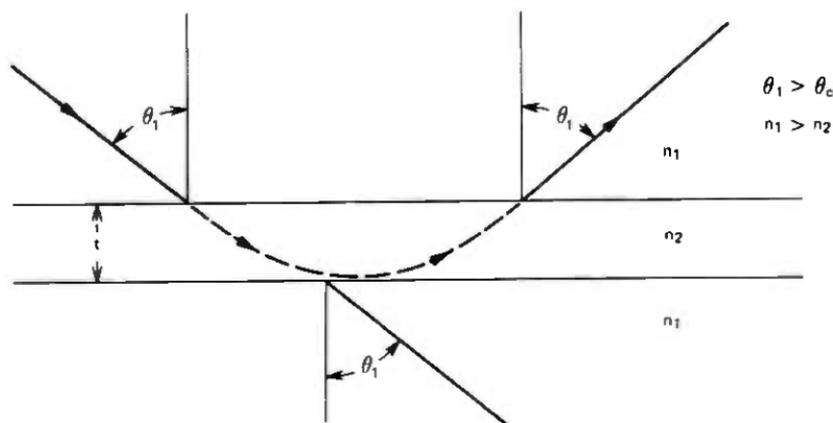

Fig. 7—Geometry for calculation of transmission and reflection coefficients in a single-layered three-medium dielectric.

IV. CROSSTALK FORMULA ADAPTED TO CIRCULAR GEOMETRY

We assume now that the only crosstalk that is modeled here is that which is due to frustrated total reflection. The geometry of the rays near the fiber wall at a hit is illustrated in Fig. 8, which is a cross-sectional view of a fiber and one of its neighbors. Here, ϕ is the polar angle which also serves to define the length t. Using simple geometry we see that

$$t = c \cos \phi - d/2 - [(d/2)^2 - c^2 \sin^2 \phi]^{\frac{1}{2}} \tag{18}$$

and that the angle ϕ_m for which the line ON is tangent to the circle O' is

$$\phi_m = \sin^{-1} d/2c. \tag{19}$$

Let E denote the center point of the entrance end of the fiber. Then a meridional ray ES hitting the wall of the fiber lies in the plane of inci-
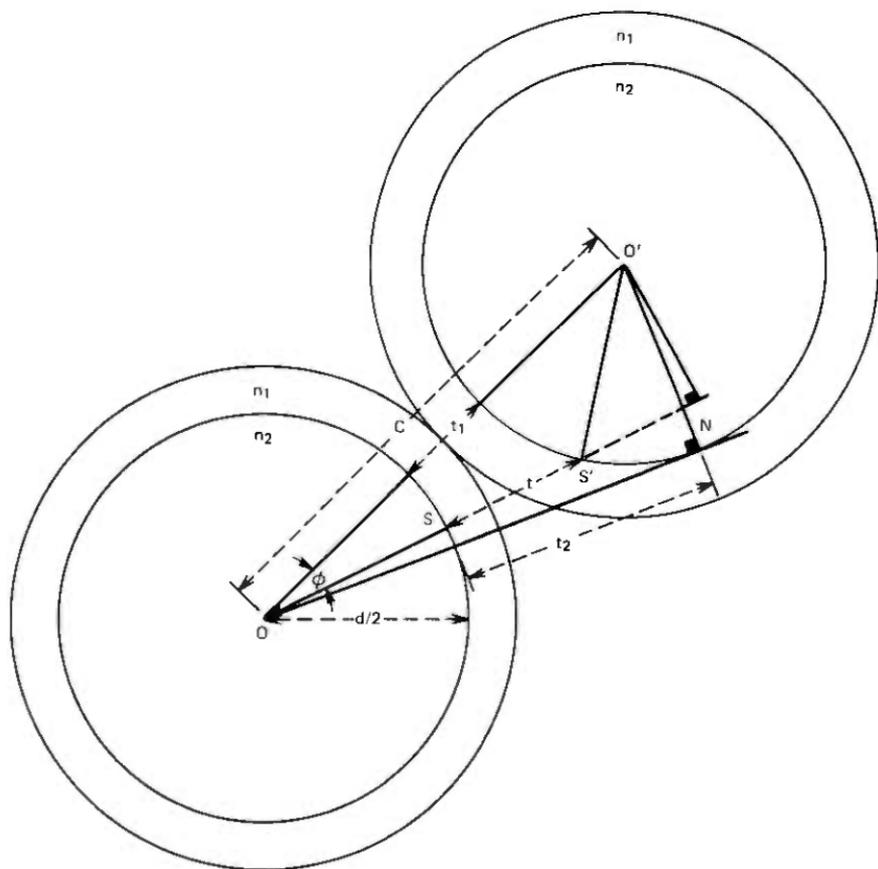

Fig. 8—Geometry of two adjacent round optical fibers.

dence EOS, which cuts the fiber O' at a line passing through S' parallel to the axis of the fiber. The transmission coefficient is calculated in the form of plane wave tunneling through a slab of thickness t, as in Fig. 7. When the medium between is homogeneous and lossless, eq. (16) can be used.

The transmitted ray will propagate as a skew ray and emerge from a point a number of wavelengths from S' depending on θ, as is clear from Figs. 5 and 7. Since the displacement Δ is small compared to the path length between bounces and very small compared to the length of fiber we are interested in, we can essentially regard the ray to emerge from S' for the purpose of calculating the path length of the transmitted ray.

Skew ray analysis shows[12] that the length of the skew ray crosstalk path lengths are also given by $L \sec \theta$. We can now calculate the expressions for P_{XT} and P_{output}. From the above discussion, substitution into the crosstalk integral, eq. (12), yields

$$A'(n) \epsilon^{-\alpha d(n-\frac{1}{2}) \csc \theta} = \epsilon^{-\alpha L \sec \theta} \tag{20}$$

and

$$\sum_{n=1}^{M} \left[R(\theta, \phi) \right]^{n-1} A'(n) \epsilon^{-\alpha d(n-\frac{1}{2}) \csc \theta} = \epsilon^{-\alpha L \sec \theta} \sum_{n=1}^{M} R(\theta, \phi)^{n-1}$$

$$= \epsilon^{-\alpha L \sec \theta} \left[\frac{1 - R(\theta, \phi)^M}{1 - R(\theta, \phi)} \right]. \tag{21}$$

The crosstalk integral simplifies to

$$P_{XT} = \int_{o}^{\theta_{\max}} d\theta \int_{o}^{2\pi} d\phi \, \sin \theta F(\theta, \phi) \epsilon^{-\alpha L \sec \theta} T(\theta, \phi)$$

$$\times \left[\frac{1 - R(\theta, \phi)^M}{1 - R(\theta, \phi)} \right]. \tag{22}$$

Making a change of variable to t with

$$\frac{d\phi}{dt} = \frac{c^2 - (d/2)^2 - (t + d/2)^2}{(t + d/2)\{(c^2 - t^2)[(t + d)^2 - c^2]\}^{\frac{1}{2}}}, \tag{23}$$

we can then write

$$P_{XT} = N \int_{o}^{\theta_{\max}} d\theta \int_{t_1}^{t_2} dt \left(\frac{d\phi}{dt} \right) \sin \theta F(\theta, t) \epsilon^{-\alpha L \sec \theta} T(\theta, t)$$

$$\times \left[\frac{1 - R(\theta, t)^M}{1 - R(\theta, t)} \right], \tag{24}$$

where $N = 2$ times the number of fibers adjacent to the excited fiber.

V. SIMPLIFIED CROSSTALK FORMULAS USED FOR INITIAL COMPUTER STUDY

A specific model requires a knowledge of $F(\theta, \phi)$, $R(\theta, \phi)$, and $T(\theta, \phi)$ in eqs. (24) and (10) to calculate the crosstalk between fibers. As a first-order approximation, the following assumptions were made:

(*i*) The interfiber medium is lossless and homogeneous so that eq. (16) is valid as well as:

(*ii*)
$$T(\theta, \phi) + R(\theta, \phi) = 1. \tag{25}$$

(*iii*) The input angular power distribution of the fiber is a gaussian function of the form

$$F(\theta, \phi) = \epsilon^{-(\theta/\kappa\theta_c)^2}, \tag{26}$$

where κ is a parameter that is a measure of the width of the beam and also an indication of how the power is distributed between the modes of the fiber. θ_c is the critical angle of the fiber.

(*iv*) The air spaces between the fibers are replaced by cladding material for the purposes of calculating the transmission coefficient, $T(\theta, \phi)$. A partial check of this assumption showed that the primary contribution to crosstalk occurred for small values of ϕ. Replacing the air spaces by cladding material caused a maximum error in crosstalk of 4.5 dB.

Under the assumptions mentioned above, the crosstalk integral, (24), becomes

$$P_{\mathbf{XT}} = N \int_o^{\theta_{\max}} d\theta \int_{t_1}^{t_2} dt \left(\frac{d\phi}{dt} \right)$$
$$\times \sin \theta \epsilon^{-(\theta/\kappa\theta_c)^2} \epsilon^{-\alpha L \sec \theta} \{1 - [1 - T(\theta, t)]^M\}. \tag{27}$$

The total output power of the transmitting fiber becomes

$$P_{\text{output}} = \int_o^{2\pi} d\phi \int_o^{\theta_{\max}} \sin \theta \epsilon^{-(\theta/\kappa\theta_c)^2} \epsilon^{-\alpha L \sec \theta} [1 - T(\theta, \phi)]^M d\theta. \tag{28}$$

VI. COMPUTER STUDY—SUMMARY OF RESULTS

A computer program was written and the integrals (27) and (28) were evaluated for typical fiber parameters. A number of studies were made to determine how FEXT [eq. (14)] varies as a function of cladding thickness, numerical aperture, length, κ (kappa), and wavelength. Table I is a guide that defines the parameters and relates the variables

Table I

Far-End Crosstalk as a Function of	Figure Number	Parameters Held Constant	Range of Independent Variable
$\dfrac{d}{c} - \dfrac{\text{Core Diam}}{\text{Cladding Diam}}$	9	NA = 0.10, d = 25.4 μm, L = 1 km, λ = 0.6328 μm α = 20 dB/km, κ = 0.25, 0.35, 0.4, 0.5, 10.0	$\dfrac{d}{c}$ — 0.3 to 0.7
$\dfrac{d}{c}$	10	NA = 0.15, d = 25.4 μm, L = 1 km, λ = 0.6328 μm α = 20 dB/km, κ = 0.25, 0.35, 0.4, 0.5, 10.0	$\dfrac{d}{c}$ — 0.3 to 0.7
$\dfrac{d}{c}$	11	NA = 0.20, d = 25.4 μm, L = 1 km, λ = 0.6328 μm α = 20 dB/km, κ = 0.25, 0.35, 0.4, 0.5, 10.0	$\dfrac{d}{c}$ — 0.3 to 0.7
$\dfrac{d}{c}$	12	NA = 0.10, d = 50.8 μm, L = 1 km, λ = 0.6328 μm α = 20 dB/km, κ = 0.25, 0.35, 0.4, 0.5, 10.0	$\dfrac{d}{c}$ — 0.5 to 0.9
$\dfrac{d}{c}$	13	NA = 0.15, d = 50.8 μm, L = 1 km, λ = 0.6328 μm α = 20 dB/km, κ = 0.25, 0.35, 0.4, 0.5, 10.0	$\dfrac{d}{c}$ — 0.5 to 0.9
$\dfrac{d}{c}$	14	NA = 0.20, d = 50.8 μm, L = 1 km, λ = 0.6328 μm α = 20 dB/km, κ = 0.25, 0.35, 0.4, 0.5, 10.0	$\dfrac{d}{c}$ — 0.5 to 0.9
$\dfrac{d}{c}$	15	NA = 0.10, d = 76.2 μm, L = 1 km, λ = 0.6328 μm α = 20 dB/km, κ = 0.25, 0.35, 0.4, 0.5, 10.0	$\dfrac{d}{c}$ — 0.5 to 0.9
$\dfrac{d}{c}$	16	NA = 0.15, d = 76.2 μm, L = 1 km, λ = 0.6328 μm α = 20 dB/km, κ = 0.25, 0.35, 0.4, 0.5, 10.0	$\dfrac{d}{c}$ — 0.5 to 0.9
$\dfrac{d}{c}$	17	NA = 0.20, d = 76.2 μm, L = 1 km, λ = 0.6328 μm α = 20 dB/km, κ = 0.25, 0.35, 0.4, 0.5, 10.0	$\dfrac{d}{c}$ — 0.5 to 0.9
Numerical aperture	18	d = 50.8 μm, c = 85 μm, α = 20 dB/km, L = 1 km, λ = 0.6328 μm, κ = 0.25, 0.35, 0.40, 0.50, 10.0	NA = 0.05 to 0.30
Fiber length	19	d = 50.8 μm, c = 85 μm, α = 20 dB/km, NA = 0.15 λ = 0.6328 μm, κ = 0.25, 0.35, 0.4, 0.5, 10.0	L = 100 m to 5 km
κ (Kappa)	20	L = 1 km, d = 50.8 μm, c = 85 μm, α = 20 dB/km NA = 0.15, λ = 0.6328 μm	κ = 0.1 to 1.0
Wavelength	21	L = 1 km, d = 50.8 μm, c = 85 μm, α = 20 dB/km NA = 0.15, κ = 0.25, 0.35, 0.4, 0.5, 10.0	λ = 0.6328 to 1.06 μm
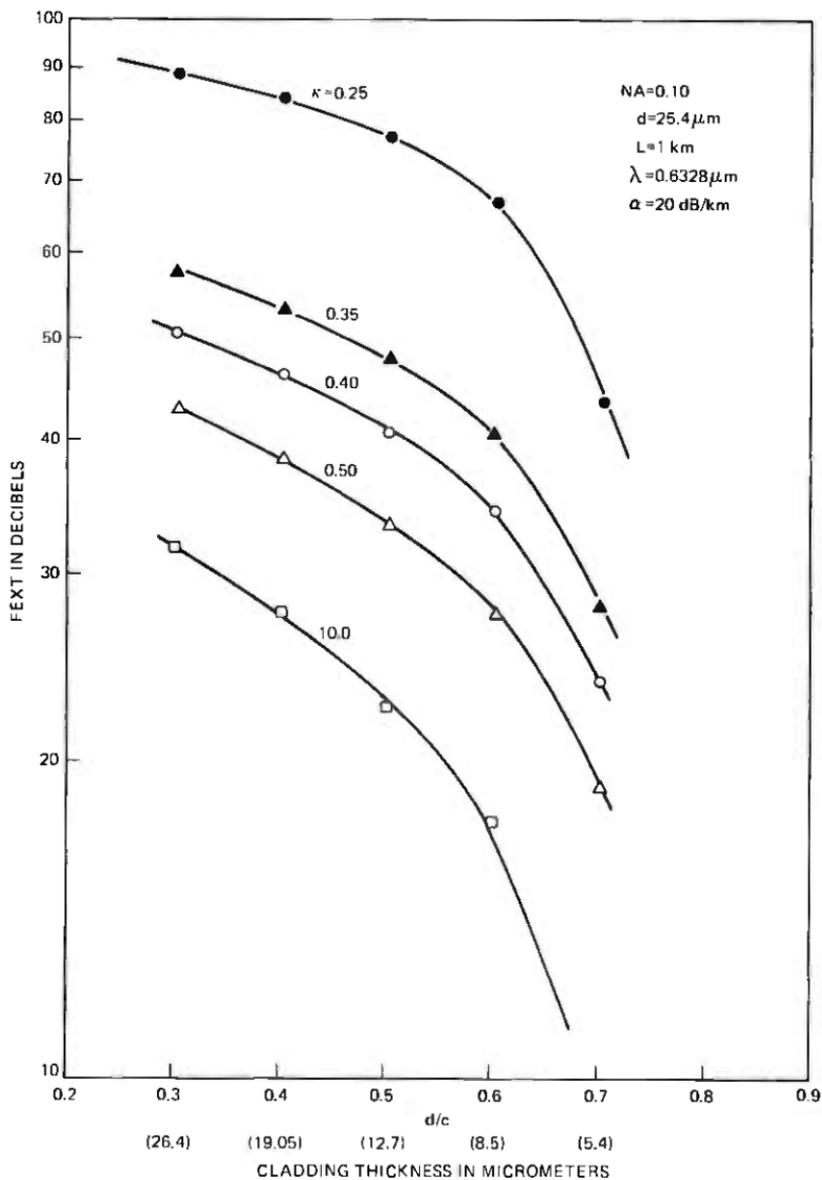
Fig. 9—FEXT vs. d/C (cladding thickness), $d = 25.4$ μm, NA $= 0.10$, $L = 1$ km, $\lambda = 0.6328$ μm, $\alpha = 20$ dB/km.

in the study to the figure numbers appearing in this paper. Figures 9 to 11 show, for numerical apertures of 0.10, 0.15, and 0.20, respectively, the relationships between FEXT and cladding thickness for a 1-mil fiber core diameter. For the same numerical apertures, Figs. 12 to 14 and 15 to 17 show this relationship for 2- and 3-mil core diameters.
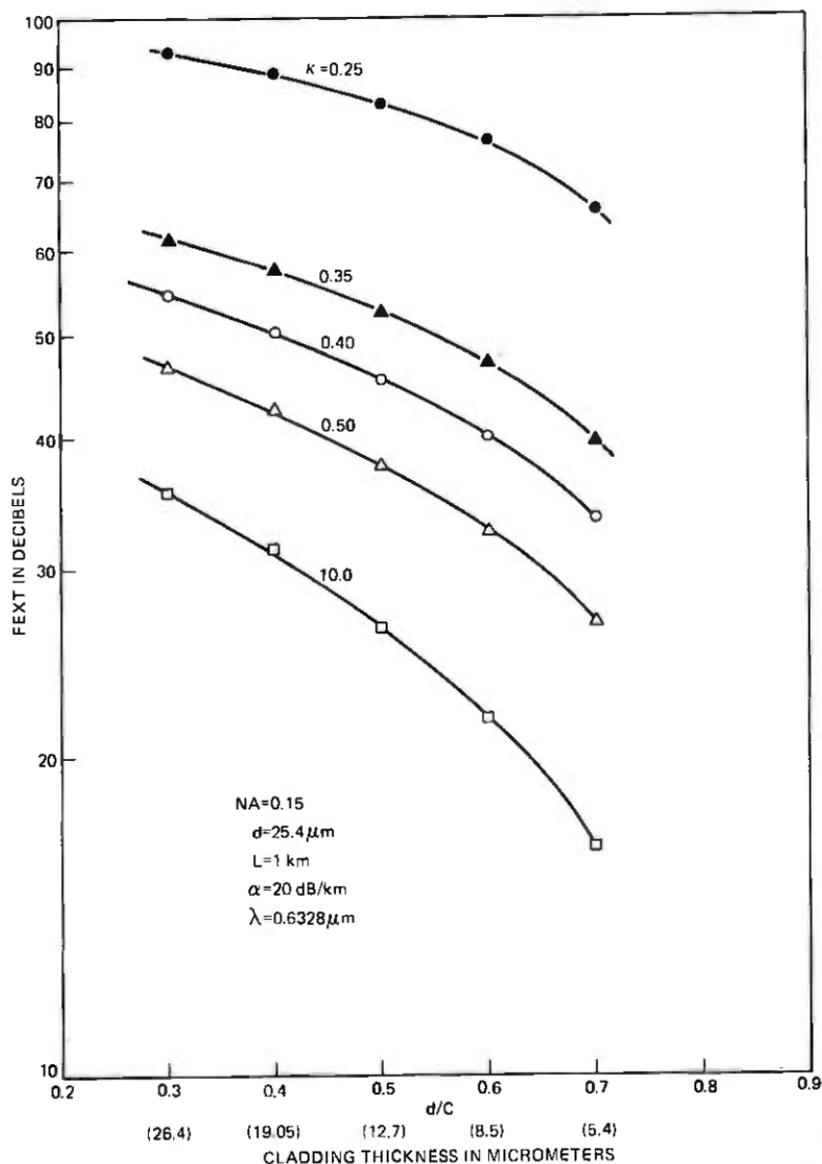
Fig. 10—FEXT vs. d/C (cladding thickness), $d = 25.4$ μm, NA = 0.15, $L = 1$ km, $\lambda = 0.6328$ μm, $\alpha = 20$ dB/km.
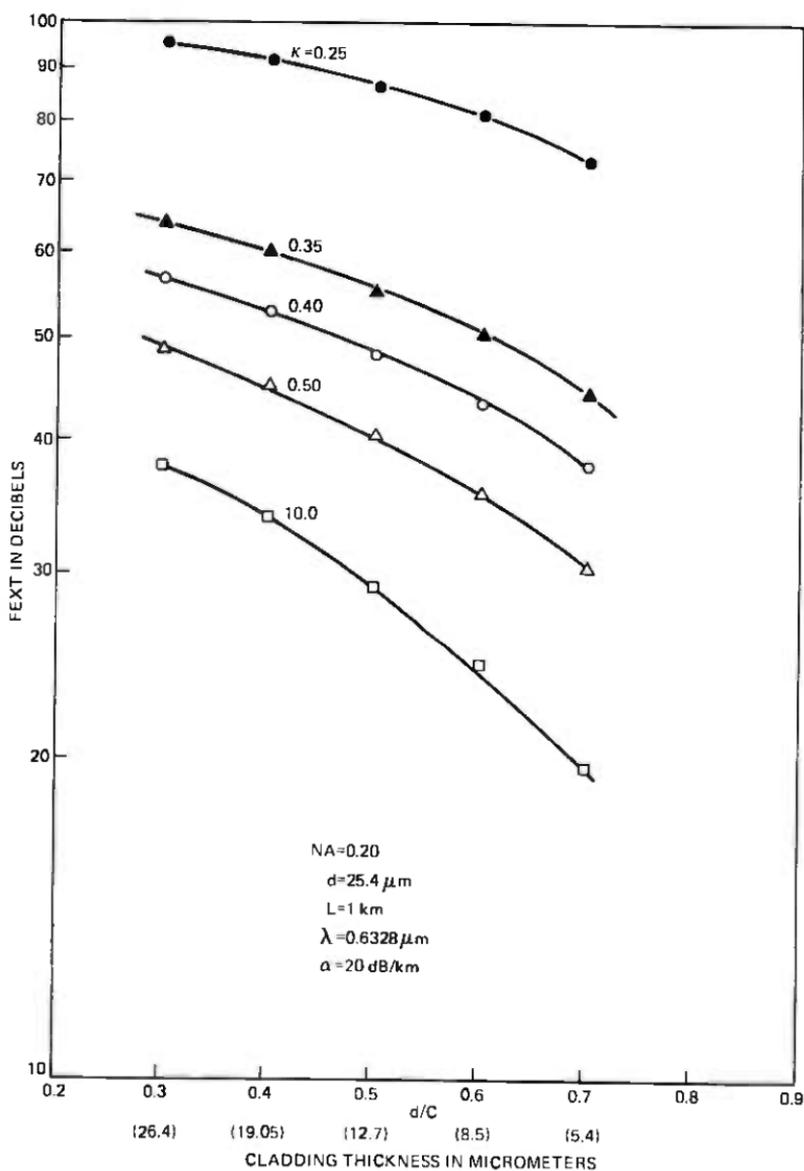
Fig. 11—FEXT vs. d/C (cladding thickness), $d = 25.4$ μm, NA = 0.20, $L = 1$ km, $\lambda = 0.6328$ μm, $\alpha = 20$ dB/km.

Fig. 12—FEXT vs. d/C (cladding thickness), $d = 50.8$ μm, NA = 0.10, $L = 1$ km, $\lambda = 0.6328$ μm, $\alpha = 20$ dB/km.

Fig. 13—FEXT vs. d/C (cladding thickness), $d = 50.8$ μm, NA $= 0.15$, $L = 1$ km, $\lambda = 0.6328$ μm, $\alpha = 20$ dB/km.
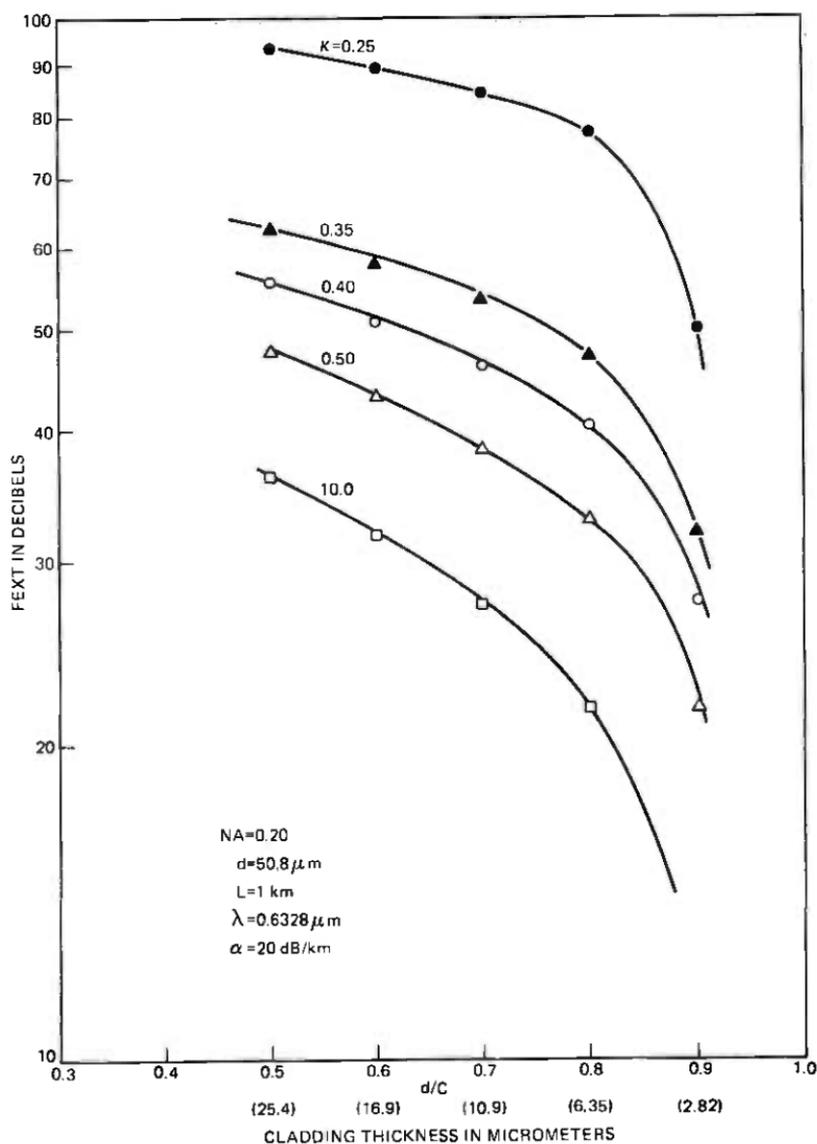
Fig. 14—FEXT vs. d/C (cladding thickness), $d = 50.8\ \mu\text{m}$, NA $= 0.20$, $L = 1$ km, $\lambda = 0.6328\ \mu\text{m}$, $\alpha = 20\ \text{dB/km}$.
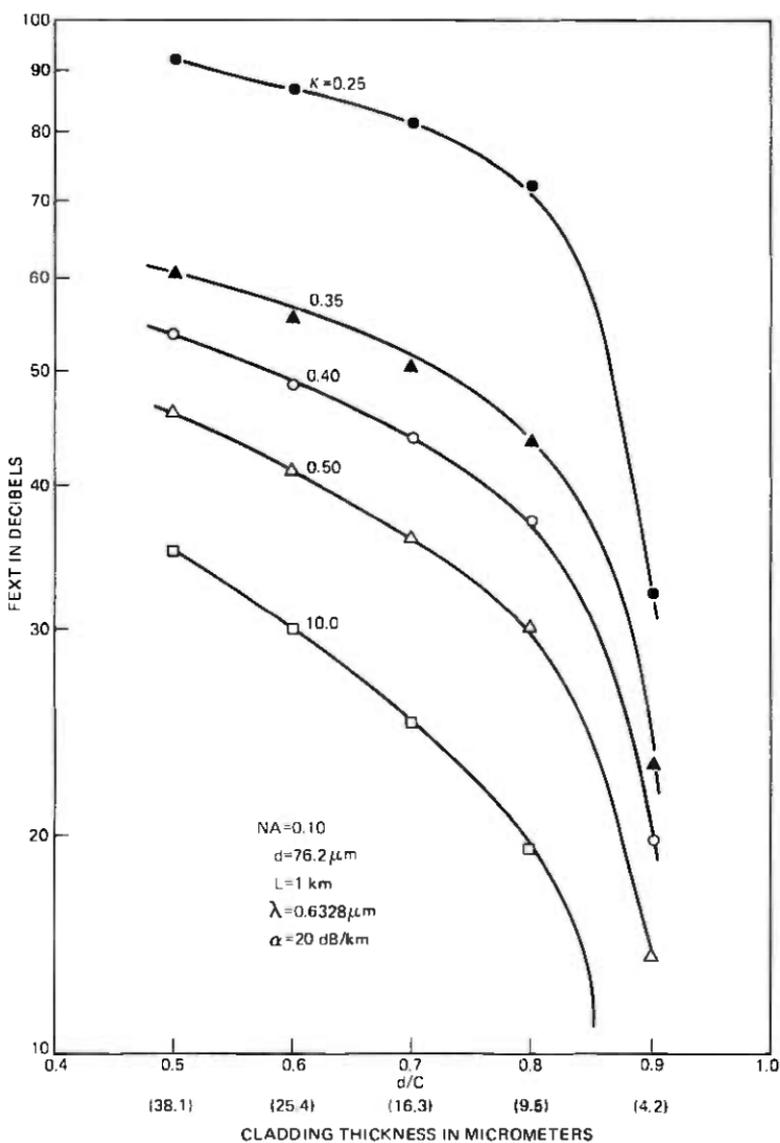
Fig. 15—FEXT vs. d/C (cladding thickness), $d = 76.2$ μm, NA $= 0.10$, $L = 1$ km, $\lambda = 0.6328$ μm, $\alpha = 20$ dB/km.
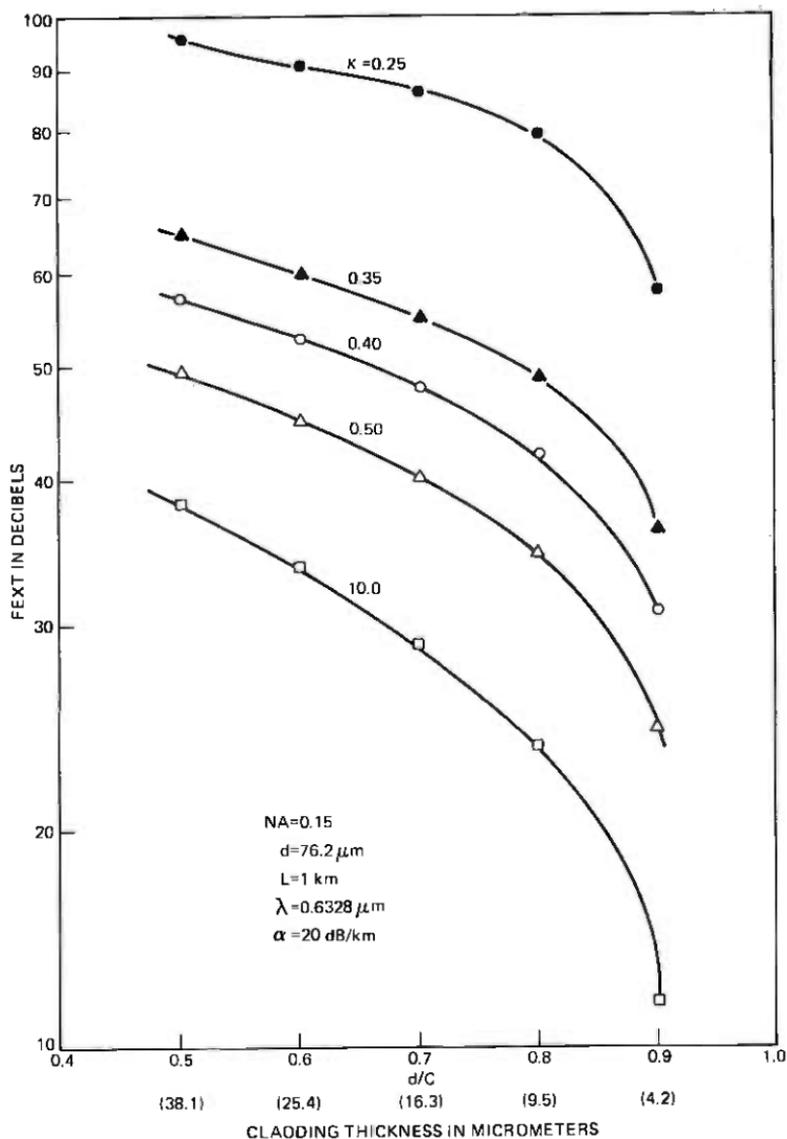
Fig. 16—FEXT vs. d/C (cladding thickness), $d = 76.2$ μm, NA $= 0.15$, $L = 1$ km, $\lambda = 0.6328$ μm, $\alpha = 20$ dB/km.
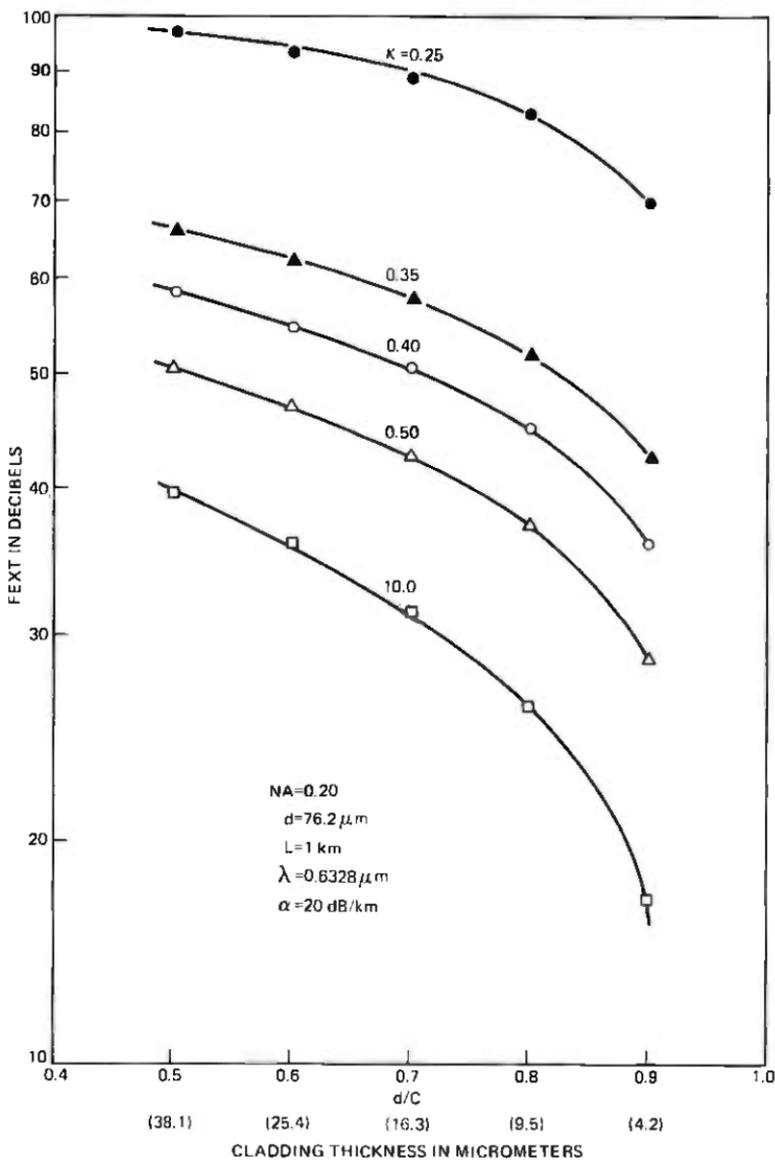
Fig. 17—FEXT vs. d/C (cladding thickness), $d = 76.2$ µm, NA = 0.20, $L = 1$ km, $\lambda = 0.6328$ µm, $\alpha = 20$ dB/km.

For practical claddings greater than 12 microns in thickness, increasing the cladding thickness will improve crosstalk isolation by approximately 0.8 dB/micron. Cladding thickness is an important design parameter for improving crosstalk isolation between optical fibers. If we attempted to eliminate crosstalk by coating the cladding with an opaque substance, cladding thickness would also play an
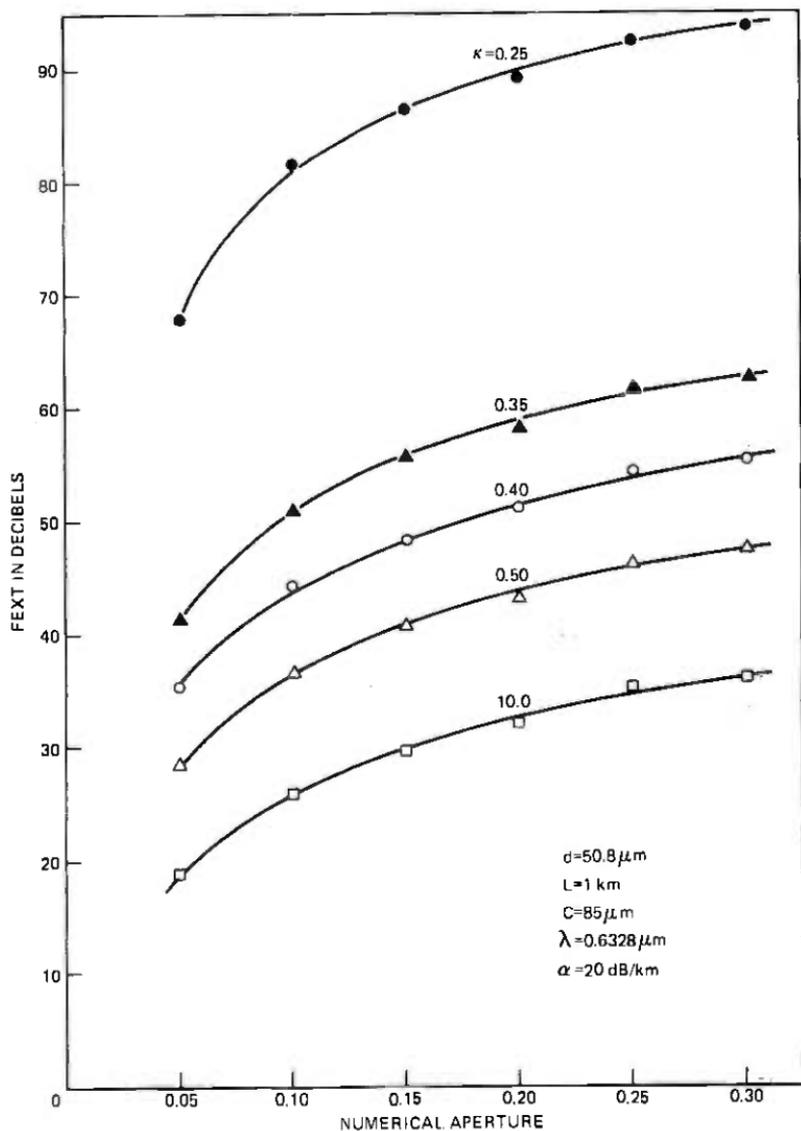
Fig. 18—FEXT vs. numerical aperture, $d = 50.8$ μm, $C = 85$ μm, $\alpha = 20$ dB/km, $L = 1$ km, $\lambda = 0.6328$ μm, $\kappa = 0.25, 0.35, 0.4, 0.5, 10.0$.

important role in determining the amount of loss suffered by the transmitted energy in the core owing to the lossy coating. A cross-check of Figs. 9 to 17 shows a very weak dependence of FEXT on core diameter.

A fiber core diameter of 50.8 μm and cladding diameter of 85 μm was chosen to represent a typical fiber geometry in the remainder of this study. Figure 18 shows the relationship between FEXT and numerical aperture with a fixed cladding refractive index of 1.458. For a given kappa, crosstalk isolation improves by approximately 6 dB by doubling the numerical aperture in the range from 0.10 to 0.30. For an NA
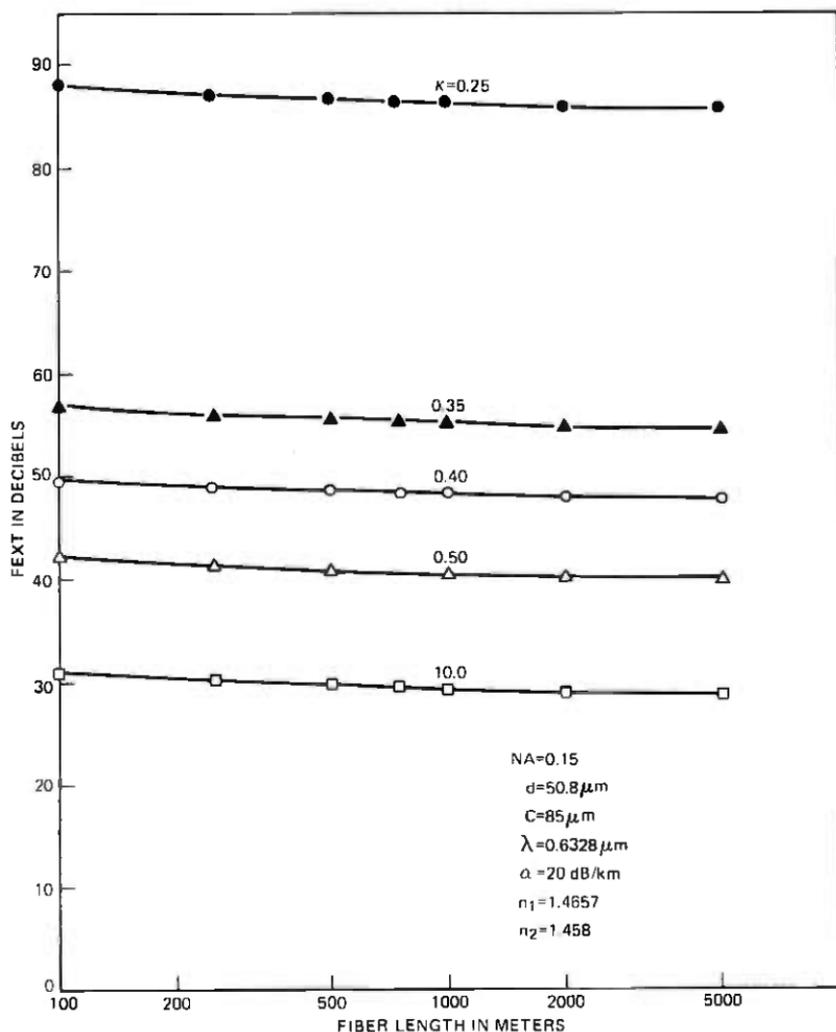
Fig. 19—FEXT vs. fiber length $d = 50.8$ μm, $C = 85$ μm, $\alpha = 20$ dB/km, NA $= 0.15$, $\lambda = 0.6328$ μm, $\kappa = 0.25, 0.35, 0.40, 0.50, 10.0$.

= 0.15, crosstalk was evaluated as a function of length as shown in Fig. 19. For a fixed kappa, crosstalk isolation decreased by approximately 2 dB per decade of length. The effect of kappa on crosstalk is shown in Fig. 20. Kappa is an indication of the distribution of power among the modes of the fiber. As kappa decreases in value, the energy distributed in the lower-order modes of a fiber increases. A knowledge of the steady-state mode distribution should enable us to estimate an effective κ for use in this model. As illustrated in Fig. 20, κ is an extremely important parameter in determining the crosstalk between fibers. For kappas less than 0.5, crosstalk isolation is greater than 40 dB for a kilometer length of fiber. For kappas less than 0.40, crosstalk isolation is greater than 50 dB. The final study, illustrated in Fig. 21, shows the relationship between FEXT and wavelength. For a given kappa, doubling the wavelength of the transmitting signal decreases crosstalk isolation by approximately 7 dB. It is envisioned that typical fiber losses in a future optical transmission system will be less than the 20-dB/km loss reported here. We will extend, in the future, our computer study to include fiber losses ranging from 2 to 20 dB/km, but do not expect this loss parameter to change the general conclusions drawn in this paper.

To determine if the primary mechanism for crosstalk in optical fibers is frustrated total reflection of waves in a multilayered medium, a crosstalk experiment on a long length of parallel fibers must be performed. In this experiment, crosstalk and kappa should be measured
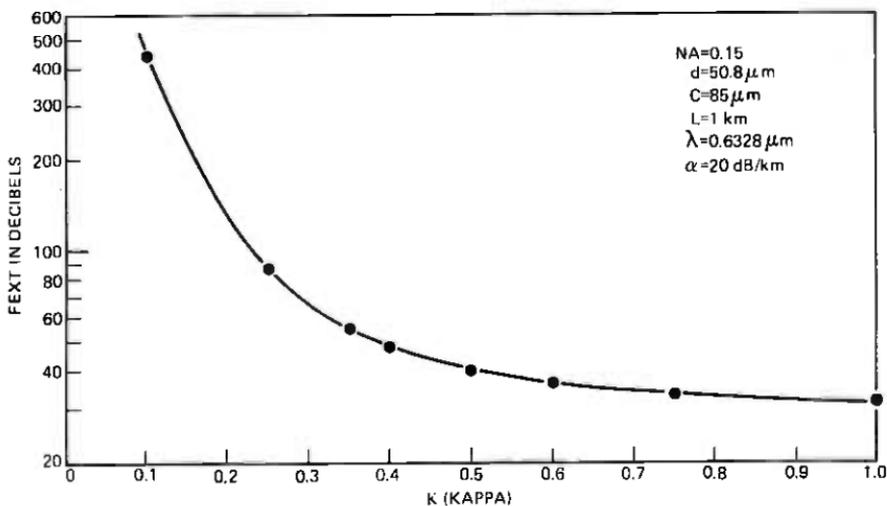
FIG. 20—FEXT vs. κ (kappa), $L = 1$ km, $d = 50.8$ μm, $C = 85$ μm, $\alpha = 20$ dB/km, NA $= 0.15$, $\lambda = 0.6328$ μm.
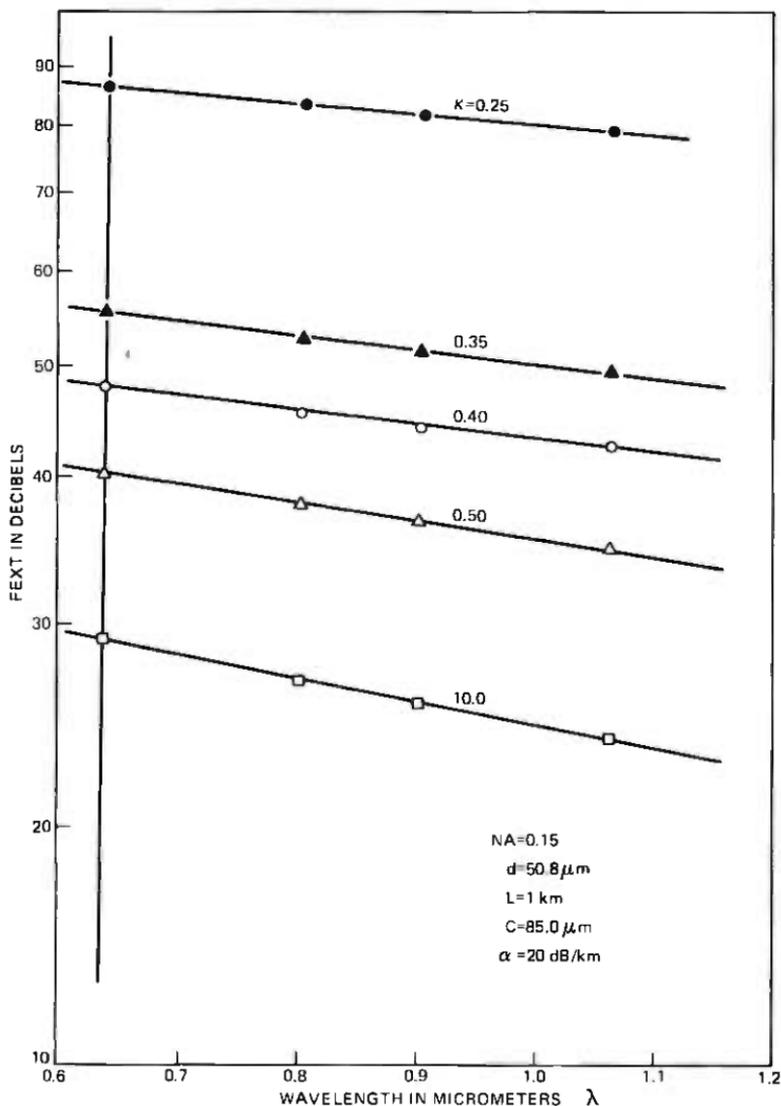
Fig. 21—FEXT vs. λ (wavelength), $L = 1$ km, $d = 50.8$ μm, $C = 85$ μm, $\alpha = 20$ dB/km, NA = 0.15, κ = 0.25, 0.35, 0.40, 0.50, 10.0.

as a function of length at a number of different wavelengths. This type of experiment will enable us to check many aspects of the model.

APPENDIX A

Calculation of Transmission Coefficients in a Multilayered Medium

The general expression for the transmission coefficients in a multi-layered dielectric medium is well known in the literature[15,17,18] and is

presented here. This expression is simplified for the three-medium case and, ultimately, eq. (16) in the text is derived.

Consider the geometry shown in Fig. 22. Let us suppose that between two semi-infinite media, denoted by 1 and $n + 1$, there are $n - 1$ layers of dielectric material denoted by 2, 3, $\cdots$, n. Let a plane wave be incident on the last layer at an angle of incidence θ_{n+1} and let the plane of incidence be the $x - z$ plane. As a result of multiple reflections at the boundaries of the layers, two waves exist in each medium with the exception of medium 1. Our problem will be to determine the amplitude of the transmitted wave in medium $n + 1$ and hence the transmission coefficient. The following notation will be used.

z_j = the coordinate of the boundary between the jth and $(j + 1)$st layers.

$d_j = z_j - z_{j-1}$ = the thickness of the jth layer.

$k_j = (2\pi/\lambda)(n_j)$ = the wave number in the jth medium.

$\alpha_j = k_j \cos \theta_j = z$ component of the wave vector in the jth layer.

$\phi_j = \alpha_j d_j$ = the phase change in the jth medium.

Z_j = the self-impedance of the jth layer.

Z_{in}^j = the input impedance looking into the jth medium from the $j + 1$ medium.

The electric and magnetic fields in the jth medium can then be written as

$$E_{jy} = A_j \exp\left[-i\alpha_j(Z - Z_{j-1})\right] + B_j \exp\left[i\alpha_j(Z - Z_{j-1})\right] \tag{29}$$

$$H_{jz} = \frac{1}{Z_j} \{A_j \exp\left[-i\alpha_j(Z - Z_{j-1})\right] - B_j \exp\left[i\alpha_j(Z - Z_{j-1})\right]\}. \tag{30}$$
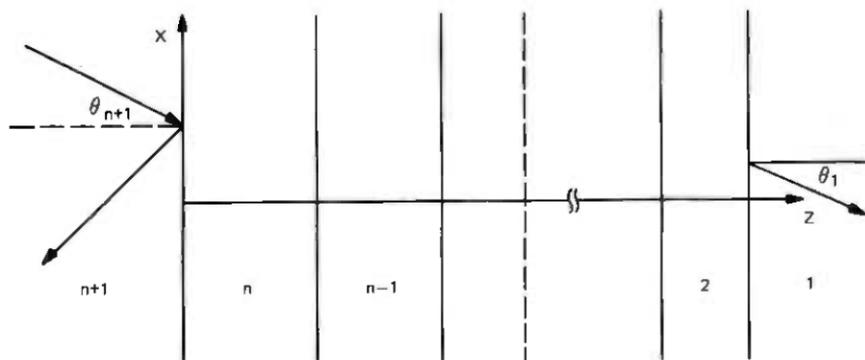
Fig. 22—Geometry used for calculation of transmission coefficients in multilayered dielectric media.

The x and t dependency in this case is omitted for the sake of brevity, but assumes the general form:

$$\exp i(k_{n+1}x \sin \theta_{n+1} - wt).$$

A_j and B_j, $B_1 = 0$ are the amplitudes of the incident and reflected waves in the jth medium. The amplitude A_{n+1} of the incident wave is assumed to be known. To obtain the transmission coefficient of interest,

$$\tau_{n+1} = \frac{A_1}{A_{n+1}} = \frac{A_1}{A_2}\frac{A_2}{A_3} \cdots \frac{A_{n-1}}{A_n}\cdot\frac{A_n}{A_{n+1}}. \tag{31}$$

We can write $2n$ boundary equations for the tangential components of the fields and solve these equations for $A_1, A_2, \cdots, A_{n+1}$; $B_1, B_2, \cdots, B_{n+1}$. When these coefficients are known, the transmission and reflection coefficients for the multilayered medium are obtained.

A second approach, and one that develops an iterative scheme more suitable to a digital computer, describes the transmission coefficient in terms of a generalized input impedance.[19] This is the approach that will be followed here. It is straightforward to derive, and it is shown in the literature that:

$$\frac{A_j}{A_{j+1}} = \frac{Z_j + Z_{in}^j}{Z_{j+1} + Z_{in}^j} \exp(i\phi_j). \tag{32}$$

Substituting into eq. (31) yields

$$\tau_{n+1} = \prod_{j=1}^{j=n} \frac{Z_j + Z_{in}^j}{Z_{j+1} + Z_{in}^j} \exp(i\phi_j), \tag{33}$$

with $d_1 = 0$, where

$$Z_{in}^j = \frac{Z_{in}^{j-1} - iZ_j \tan \phi_j}{Z_j - iZ_{in}^{j-1} \tan \phi_j}\cdot Z_j. \tag{34}$$

Specializing eq. (33) for the three-medium case discussed in the text, we can obtain, with some algebraic manipulation, the following formula:

$$\tau_2 = \frac{4Z_1Z_2}{(Z_1 - Z_2)(Z_2 - Z_3)\epsilon^{i\phi_2} + (Z_1 + Z_2)(Z_2 + Z_3)\epsilon^{-i\phi_2}}. \tag{35}$$

For electric fields parallel and perpendicular to the plane of incidence, eq. (35) will take different forms. For $E_\perp$,

$$Z_1 = \frac{Z_o}{n_1 \cos \theta_1} = Z_3 \tag{36}$$

$$Z_2 = \frac{Z_o}{n_2 \cos \theta_2} = \frac{Z_o}{in_2\gamma}, \tag{37}$$

where

$$\gamma = \left[\left(\frac{n_3}{n_2} \right)^2 \sin^2 \theta_1 - 1 \right]^{\frac{1}{2}} \tag{38}$$

$$Z_o = \sqrt{\frac{\mu}{\epsilon}} \tag{39}$$

$$\beta = k_2 d_2 \gamma. \tag{40}$$

After some algebraic manipulation,

$$\tau_{2\perp} = \frac{K_1^2 \cosh \beta - i K_1 K_2 \sinh \beta}{K_1^2 \cosh^2 \beta + K_2^2 \sinh^2 \beta}, \tag{41}$$

where

$$K_1 = 2n_1 n_2 \cos \theta_1 \gamma \tag{42}$$

$$K_2 = n_2^2 \gamma^2 - n_1^2 \cos^2 \theta_1, \tag{43}$$

for $E_{\shortparallel}$

$$Z_1 = \frac{Z_o \cos \theta_1}{n_1} = Z_3 \tag{44}$$

$$Z_2 = \frac{Z_o \cos \theta_2}{n_2} = \frac{i Z_o \gamma}{n_2}, \tag{45}$$

and $T_{2\shortparallel}$ becomes

$$\tau_{2\shortparallel} = \frac{K_1^2 \cosh \beta + i K_1 K_3 \sinh \beta}{K_1^2 \cosh^2 \beta + K_3^2 \sinh^2 \beta}, \tag{46}$$

where

$$K_3 = n_2^2 \cos^2 \theta_1 - \gamma^2 n_1^2. \tag{47}$$

Utilizing eqs. (41) and (46), we can obtain the power transmission coefficients directly:

$$T_{2\perp} = |\tau_{2\perp}|^2$$
$$= \frac{1}{\cosh^2 \beta + [(n_2^2 \gamma^2 - n_1^2 \cos^2 \theta_1)/2 n_1 n_2 \cos \theta_1 \gamma]^2 \sinh^2 \beta} \tag{48}$$

$$T_{2\shortparallel} = |\tau_{2\shortparallel}|^2$$
$$= \frac{1}{\cosh^2 \beta + [(n_2^2 \cos^2 \theta_1 - \gamma^2 n_1^2)/2 \cos \theta_1 n_1 n_2 \gamma]^2 \sinh^2 \beta}. \tag{49}$$

Equations (48) and (49) were used to obtain eq. (17) in the text of this paper.

REFERENCES

1. A. W. Snyder, "Coupled-Mode Theory for Optical Fibers," J. Opt. Soc. Am., *62*, No. 11 (November 1972), pp. 1267–1277.
2. D. Marcuse, "The Coupling of Degenerate Modes in Two Parallel Dielectric Waveguides," B.S.T.J., *50*, No. 6 (July–August 1971), pp. 1791–1816.

3. A. L. Jones, "Coupling of Optical Fibers and Scattering in Fibers," J. Opt. Soc. Am., *55*, No. 3 (March 1965), pp. 261–271.
4. D. Marcuse, *Light Transmission Optics*, New York: Van Nostrand Reinhold Co., 1972.
5. R. Vanclooster and P. Phariseau, "The Coupling of Two Parallel Dielectric Fibers I—Basic Equations," Physica, *47*, No. 4 (June 1970), pp. 485–500.
6. R. Vanclooster and P. Phariseau, "The Coupling of Two Dielectric Fibers II—Characteristic of Coupling in Two Fibers," Physica, *47*, No. 4 (June 1970), pp. 501–514.
7. R. Vanclooster and P. Phariseau, "Light Propagation in Fiber Bundles," Physica, *49*, No. 4 (November 1970), pp. 493–501.
8. M. Matsuhara and N. Kumagai, "Theory of Coupled Open Transmission Lines and Its Applications," IEEE Trans. Microwave Theory Tech., *MTT-22*, No. 4, April, 1974.
9. D. Marcuse, "Crosstalk Caused by Scattering in Slab Waveguides," B.S.T.J., *50*, No. 6 (July–August 1971), pp. 1817–1831.
10. H. P. Yuen, unpublished work.
11. N. S. Kapany and J. J. Burke, *Optical Waveguides*, New York: Academic Press, 1972.
12. N. S. Kapany, *Fiber Optics, Principles and Applications*, New York: Academic Press, 1967.
13. N. S. Kapany and J. J. Burke, "Fiber Optics IX, Waveguide Effects," J. Opt. Soc. Am., *51*, No. 10 (October 1961), pp. 1067–1078.
14. N. S. Kapany, "Fiber Optics V, Light Leakage due to Frustrated Total Reflection," J. Opt. Soc. Am., *49*, No. 9 (August 1959), pp. 770–778.
15. L. M. Brekhovskikh, *Waves in Layered Media*, New York: Academic Press, 1960.
16. R. H. Renard, "Total Reflection, A New Evaluation of the Goos-Hänchen Shift," J. Opt. Soc. Am., *54*, No. 10 (October 1964), pp. 1190–1197.
17. W. Culshaw, "Stratified Media and Total Reflection Phenomena," Appl. Opt., *11*, No. 11 (November 1972), pp. 2639–2648.
18. M. Born and E. Wolf, *Principles of Optics*, New York: Pergammon Press, 1970.
19. B. Carnahan, H. A. Luther, and J. O. Wilkes, *Applied Numerical Methods*, New York: John Wiley, 1969.
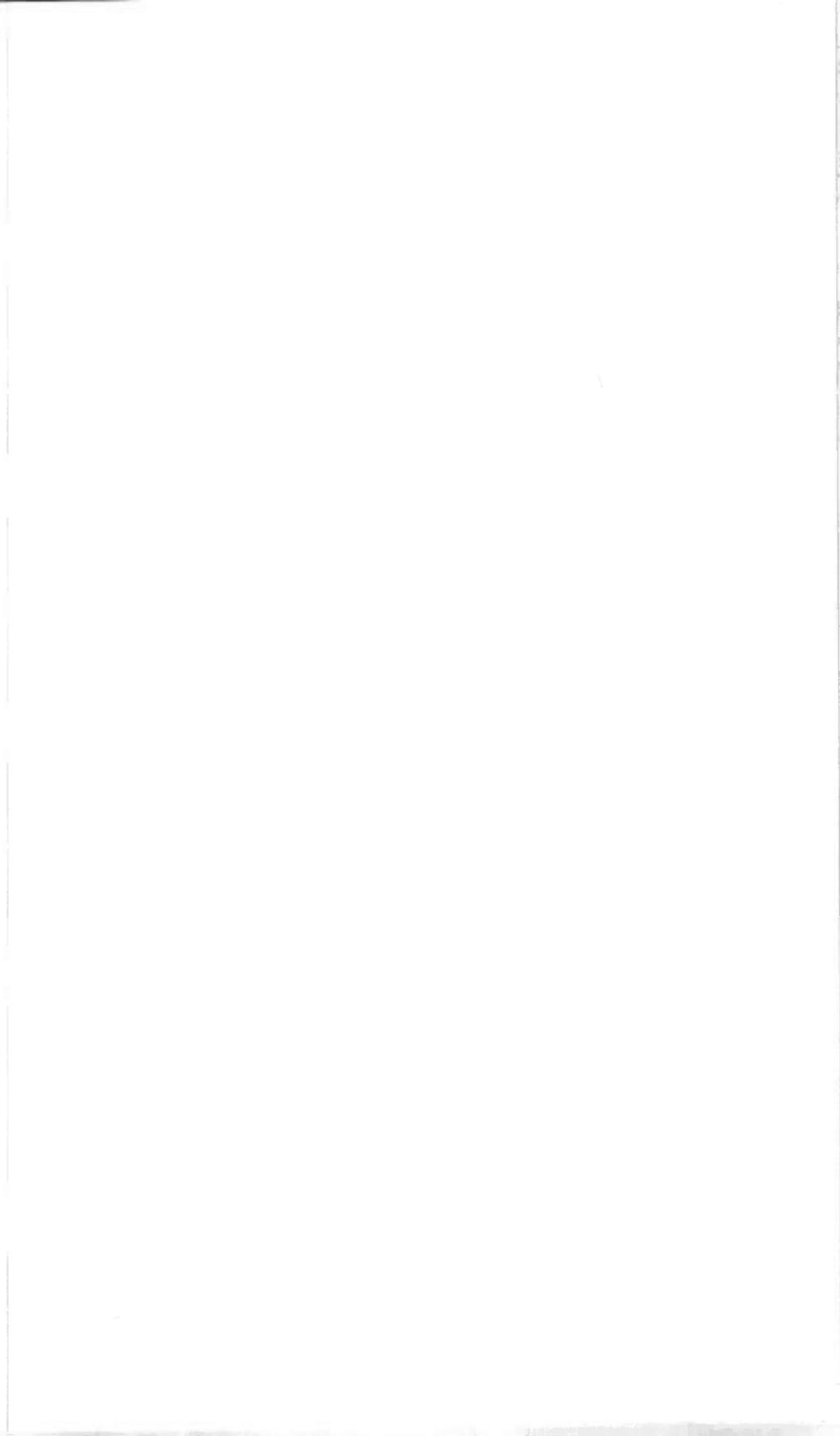
Two Derivations of the Time-Dependent Coupled-Power Equations

By S. D. PERSONICK

In this paper, the time-dependent coupled-power equations originally derived by Marcuse from intuitive arguments are rederived two ways, one using the coupled-line equations with perturbation theory and the other using the Kronecker product approach of Rowe and Young.[†]

I. INTRODUCTION

Suppose a multimode fiber guide with random mode coupling is excited by an optical source at one end at time $t = 0$.

This input excitation will produce, at time t, a response of optical power in each mode v at a position z from the input end given by $s_v(t, z)$.

Marcuse[2] has suggested from intuitive arguments that $s_v(t, z)$ should satisfy the following differential equation

$$\frac{\partial}{\partial z} \langle s_v(t, z) \rangle + \frac{1}{C_v} \frac{\partial}{\partial t} \langle s_v(t, z) \rangle = \sum_\mu |K_{\mu v}|^2 F \langle s_\mu(t, z) \rangle - (\sum_\mu |K_{\mu v}|^2 F) \langle s_v(t, z) \rangle, \quad (1)$$

where $K_{\mu v}$ is the coupling coefficient in the coupled-line equations (described below) between modes μ and v, F is the spectral height of the mechanical perturbation in the fiber geometry responsible for the coupling, $\langle s_v(t, z) \rangle$ is the average[‡] power in mode μ at position z, and C_v is the group velocity in mode v.

In this paper, we derive eq. (1) from the coupled-line equations first using perturbation theory and then using the Kronecker product approach of Rowe and Young.[3]

The importance of this work is to show that the intuitive eq. (1) does in fact follow directly from (i) the coupled-line equations, (ii)

[‡] Average for an ensemble of guides with similar gross properties and similar excitation.

voltage linearity of the guide (from Maxwell's equations), and (*iii*) ensemble averaging over the random coupling perturbation. It is encouraging to note that skillfully framed power-flow arguments such as those used by Marcuse lead to the same results as the more cumbersome approaches that start from the coupled-line equations.

It should be pointed out again that, however it is derived, eq. (1) describes the average of the flow of power for an ensemble of guides. How the flow of power in a particular guide compares to the average flow is still an open subject.[4]

II. ANALYSIS

We start with the coupled-line equations that describe the z evolution of $\mathcal{A}_v(\omega, z)$—the complex amplitude of the voltage in a mode v at position z resulting from a single Fourier component of the optical excitation at frequency ω at position $z = 0$. (We can use Fourier components since the guide is linear in voltage.)

$$\frac{\partial A_v(\omega, z)}{\partial z} = \sum_\mu A_\mu(\omega, z) K_{v\mu} f(z) \exp\left[i(\beta_v - \beta_\mu)z\right], \qquad (2)$$

where $\mathcal{A}_v(\omega, z)$ (the complex amplitude in mode v) $= A_v(\omega, z)$ $\times \exp\left[-i\beta_v(\omega)z\right]$,

$\beta_v =$ propagation constant for mode v (which is a function of ω), and

$K_{\mu v} f(z) =$ coupling coefficient between modes μ and v ($K_{\mu v}^* = -K_{v\mu}$).

If the input *power* excitation is a function of time, then the *average power* response at position z is given by the *average* of the square of the complex envelope, $a_v(z, t)$, at position z, i.e.,

$$\langle s_v(z, t) \rangle = \langle |a_v(z, t)|^2 \rangle = \int \langle \mathcal{A}_v(\omega + \sigma, z) \mathcal{A}_v^*(\omega, z) \rangle \exp\left[i(\omega + \sigma)t\right]$$
$$\times \exp(-i\omega t) d\omega d(\omega + \sigma), \qquad (3)$$

where the complex envelope $a_v(t, z)$ is the Fourier transform of $\mathcal{A}_v(\sigma, z)$.

We shall next derive eq. (1) by obtaining a differential equation for $(\partial/\partial z)\langle s_v(t, z) \rangle$.

2.1 Perturbation theory approach

Following techniques used previously by Marcuse,[2] we first write

$$\frac{\partial}{\partial z} \langle A_v(\omega + \sigma, z) A_v^*(\omega, z) \rangle = \left\langle \left[\frac{\partial}{\partial z} A_v(\omega + \sigma, z) \right] A_v^*(\omega, z) \right\rangle$$
$$+ \left\langle A_v(\omega + \sigma, z) \left[\frac{\partial}{\partial z} A_v^*(\omega, z) \right] \right\rangle. \qquad (4)$$

Using (2) in (4) we obtain

$$\frac{\partial}{\partial z} \langle A_v(\omega + \sigma, z) A_v^*(\omega, z) \rangle$$

$$= \langle \sum_\mu A_\mu(\omega + \sigma, z) K_{v\mu} f(z)$$

$$\cdot \exp \{i[\beta_v(\omega + \sigma) - \beta_\mu(\omega + \sigma)]z\} A_v^*(\omega, z) \rangle$$

$$+ \langle \sum_\mu A_v(\omega + \sigma, z) A_\mu^*(\omega, z) K_{v\mu}^* f(z) \exp \{-i[\beta_v(\omega) - \beta_\mu(\omega)]z\} \rangle. \quad (5)$$

Next we recognize that the perturbation solution to (2) is given by

$$A_v(z) = A_v(z') + \sum_1^N K_{v\mu} A_\mu(z') \int_{z'}^z f(x) \exp[+i(\beta_v - \beta_\mu)x] dx. \quad (6)$$

We shall now substitute (6) into (5) with the following approximations: We assume $f(x)$ is independent of $A_v(z') A_\mu^*(z')$ for $x > z'$. We keep only terms that are second order in $K_{v\mu}$, because first-order terms would include the factor $f(x)$ to first order and, since $f(x)$ has been assumed independent of $A_v(z') A_\mu(z')$ for $x > z'$, the expectation of these terms would vanish [$f(x)$ has zero mean].

We obtain

$$\frac{\partial}{\partial z} \langle A_v(\omega + \sigma, z) A_v^*(\omega, z) \rangle$$

$$= \sum_\mu \sum_\delta \langle A_\delta(\omega + \sigma, z') A_v^*(\omega, z') \rangle K_{\mu\delta} K_{v\mu} \int_{z'}^z \langle f(x) f(z) \rangle$$

$$\cdot \exp i[\beta_\mu(\omega + \sigma) - \beta_\delta(\omega + \sigma)](x - z) dx$$

$$\cdot \exp \{i[\beta_v(\omega + \sigma) - \beta_\delta(\omega + \sigma)]z\}$$

$$+ \sum_\mu \sum_\delta \langle A_\mu(\omega + \sigma, z') A_\delta^*(\omega, z') \rangle K_{v\mu} K_{v\delta}^* \int_{z'}^z \langle f(x) f(z) \rangle$$

$$\cdot \exp \{-i[\beta_v(\omega) - \beta_\delta(\omega)](x - z)\} dx$$

$$\cdot \exp \{i[\beta_v(\omega + \sigma) - \beta_\mu(\omega + \sigma) - \beta_v(\omega) + \beta_\delta(\omega)]z\}, \quad (7)$$

plus two similar terms where ω and $\omega + \sigma$ are interchanged and the conjugate is taken.

In (7) we have again assumed independence between $f(x)$ and $A_\mu(\omega, z')$ for $x > z'$. Now we assume that terms that are rapidly varying in z can be neglected, i.e., we neglect terms proportional to $\exp\{i[\beta_\mu(\omega) - \beta_v(\omega)]z\}$ unless $\mu = v$. We obtain for small $z - z'$ [so that $\langle A(\omega + \sigma, z) A^*(\omega, z) \rangle \approx \langle A(\omega + \sigma, z') A^*(\omega, z') \rangle$].

$$\frac{\partial}{\partial z} \langle A_v(\omega + \sigma, z) A_v^*(\omega, z) \rangle = - \sum_\mu \langle A_v(\omega + \sigma, z) A_v^*(\omega, z) \rangle |K_{\mu v}|^2 F$$

$$+ \sum_\mu \langle A_\mu(\omega + \sigma, z) A_\mu^*(\omega, z) \rangle |K_{\mu v}|^2 F$$

$$\cdot \exp \{i[\beta_v(\omega + \sigma) - \beta_\mu(\omega + \sigma) - \beta_v(\omega) + \beta_\mu(\omega)]z\}, \quad (8)$$

where F is the spectral height of the random process $f(x)$, i.e.,

$$F \cong 2 \int_{x'-x}^{0} \langle f(x)f(z) \rangle$$
$$\cdot \exp \{i[\beta_\mu(\omega + \sigma) - \beta_v(\omega + \sigma)](x - z)\}d(x - z) \quad (9)$$

is assumed independent of $[\beta_u(\omega + \sigma) - \beta_v(\omega + \sigma)]$, which implies $\langle f(x)f(z) \rangle \approx F\delta(x - z)$.

Next we assume that each mode has a well-defined group velocity within the frequency band of interest, that is,

$$\beta_v(\omega + \sigma) - \beta_v(\omega) \cong \frac{\sigma}{C_v}. \quad (10)$$

Substituting (10) into (8) we obtain, using (2),

$$\frac{\partial}{\partial z} (A_v(\omega + \sigma, z)A_v^*(\omega, z))$$

$$= \frac{\partial}{\partial z} (\langle \mathcal{A}_v(\omega + \sigma, z)\mathcal{A}_v^*(\omega, z) \rangle \exp (i\sigma z/C_v))$$

$$= - \sum_\mu |K_{\mu v}|^2 F \langle \mathcal{A}_v(\omega + \sigma, z)\mathcal{A}_v^*(\omega, z) \rangle \exp (i\sigma z/C_v)$$

$$+ \sum_\mu |K_{\mu v}|^2 F (\mathcal{A}_\mu(\omega + \sigma, z)\mathcal{A}_\mu^*(\omega, z)) \exp (i\sigma z/C_v). \quad (11)$$

But

$$\frac{\partial}{\partial z} (\langle \mathcal{A}_v(\omega + \sigma, z)\mathcal{A}_v^*(\omega, z) \rangle \exp (i\sigma z/C_v))$$

$$= \exp (i\sigma z/C_v) \left[\frac{\partial}{\partial z} \langle \mathcal{A}_v(\omega + \sigma, z)\mathcal{A}_v^*(\omega, z) \rangle \right.$$

$$\left. + \frac{i\sigma}{C_v} \langle \mathcal{A}_v(\omega + \sigma, z)\mathcal{A}_v^*(\omega, z) \rangle \right]. \quad (12)$$

Substituting (12) into (11) and Fourier transforming as in (3), we obtain

$$\frac{\partial}{\partial z} (s_v(t, z)) + \frac{1}{C_v} \frac{\partial}{\partial t} \langle s_v(t, z) \rangle$$

$$= - \sum_\mu |K_{\mu v}|^2 F \langle s_v(t, z) \rangle + \sum_\mu |K_{\mu v}|^2 F (s_\mu(t, z)), \quad (13)$$

which is the same as eq. (1).

2.2 Kronecker product approach

Rowe and Young[3] assume two modes and write (in our notation)

$$\frac{\partial A_1(\omega, z)}{\partial z} = K_{12}f(z) \exp [i(\beta_1 - \beta_2)z] A_2(\omega, z)$$
$$\frac{\partial A_2(\omega, z)}{\partial z} = K_{21}f(z) \exp [i(\beta_2 - \beta_1)z] A_1(\omega, z), \quad (14)$$

where

$$K_{12} = K_{21} = iK.$$

They then assume that $f(z)$ is the derivative of an independent increments process satisfying

$$\langle f(z)f(z') \rangle = F\delta(z - z') \text{ (Dirac delta)}.$$

They divide the fiber into intervals (sections) of length Δ and assume that all the coupling takes place at the right end of each interval. They obtain the following difference equation:

$$A_1(\omega, l\Delta) = \cos (Kc_l)A_1[\omega, (l - 1)\Delta] + i \sin (Kc_l)A_2[\omega, (l - 1)\Delta]$$
$$\cdot \exp [i(\beta_1 - \beta_2)l\Delta]$$
$$A_2(\omega, l\Delta) = i \sin (Kc_l)A_1[\omega, (l - 1)\Delta] \exp [i(\beta_2 - \beta_1)l\Delta] \quad (15)$$
$$+ \cos (Kc_l)A_2[\omega, (l - 1)\Delta],$$

where

$$c_l = \int_{\text{interval } l} f(z)dz.$$

From (15) we obtain

$$\langle A_1(\omega + \sigma, l\Delta)A_1^*(\omega, l\Delta) \rangle$$
$$= \langle \cos^2 (Kc_l) \rangle \langle A_1[\omega + \sigma, (l - 1)\Delta]A_1^*[\omega, (l - 1)\Delta] \rangle$$
$$+ \langle \sin^2 (Kc_l) \rangle \langle A_2[\omega + \sigma, (l - 1)\Delta]A_2^*[\omega, (l - 1)\Delta] \rangle$$
$$\cdot \exp \{i[\beta_1(\omega + \sigma) - \beta_1(\omega) - \beta_2(\omega + \sigma) + \beta_2(\omega)]z\}. \quad (16)$$

In (16) we used the fact that c_l is independent of $A_1[\omega, (l - 1)\Delta]$ and $A_2[\omega, (l - 1)\Delta]$ because $f(z)$ has been assumed to be the derivative of an independent increments process. We also used the fact that terms like $\langle \sin (Kc_l) \cos (Kc_l) \rangle$ equal zero because c_l is a symmetrical random variable.

In the limit as $\Delta \to 0$ (intervals $\{l\}$ get small) we have

$$\langle \cos^2 (Kc_l) \rangle \to 1 - FK^2\Delta \quad \langle \sin^2 Kc_l \rangle \to FK^2\Delta.$$

We obtain from (16)

$$\frac{\partial}{\partial z} \langle A_1(\omega + \sigma, z)A_1^*(\omega, z) \rangle$$
$$= -FK^2 \langle A_1(\omega + \sigma, z)A_1^*(\omega, z) \rangle + FK^2 \langle A_2(\omega + \sigma, z)A_2^*(\omega, z) \rangle$$
$$\cdot \exp \{i[\beta_1(\omega + \sigma) - \beta_1(\omega) - \beta_2(\omega + \sigma) + \beta_2(\omega)]z\},$$
$$\frac{\partial}{\partial z} \langle A_2(\omega + \sigma, z)A_2^*(\omega, z) \rangle$$
$$= -FK^2 \langle A_2(\omega + \sigma)A_2^*(\omega) \rangle + FK^2 \langle A_1(\omega + \sigma)A_1^*(\omega) \rangle$$
$$\cdot \exp \{i[\beta_2(\omega + \sigma) - \beta_2(\omega) - \beta_1(\omega + \sigma) + \beta_1(\omega)]z\}.$$

This is the same as eq. (8) with $K_{12} = K_{21} = iK$ and $K_{11} = K_{22} = 0$. Thus, the coupled-power equations follow from (17) by Fourier transforming and making use of approximation (10).

REFERENCES
1. R. Steinberg, to be published in IEEE Trans. on Microwave Theory and Techniques, special issue on integrated optics and optical waveguides, 1975.
2. D. Marcuse, *Theory of Dielectric Optical Waveguides*, New York: Academic Press, 1974, Chapter 5.
3. H. E. Rowe and D. T. Young, "Transmission Distortion in Multimode Random Waveguides," IEEE Trans. on Microwave Theory and Techniques, *MTT20*, No. 6 (June 1972), pp. 349–365.
4. S. D. Personick, "Baseband Linearity and Equalization in Fiber Optic Digital Communication Systems," B.S.T.J., *52*, No. 7 (September 1973), pp. 1175–1193.

Effect of Junction Capacitance on the Rise Time of LED's and on the Turn-on Delay of Injection Lasers

By T. P. LEE

The space-charge capacitance of the forward-biased junction has been found to play a major role in (i) the apparent rise time of the emission from small-area, high-radiance LED's and (ii) the apparent turn-on delay of stripe-geometry DH-structure laser diodes. For a zero-bias capacitance of 200 pF, a typical value for such devices made by oxide-masking techniques, the measured rise time of an LED that is fully turned on and the turn-on delay time of injection lasers may be as much as twice the limitation imposed by the spontaneous carrier recombination time. A proposed method to reduce these delays by preshaping the driving pulse is analyzed, and a reduction of the delay by a factor of 2 or better is predicted. These results are in agreement with experiments.

I. INTRODUCTION

Among the optical sources presently available for optical communication applications, AlGaAs light-emitting diodes[1,2] and injection lasers[3,4] are probably the most compatible with low-loss optical fibers.[5,6] The possibility of direct pulse modulation of the optical output of these devices by varying the driving current is a major advantage, and the intrinsic radiative recombination time suggests that very high modulation rates are possible. Fractional-nanosecond rise times and modulation rates to 100 Mb/s for LED's[7,10] and modulation rates to hundreds of Mb/s for injection lasers[11-17] have been reported. Faster rise times and higher modulation rates are more difficult to attain.

For an ideal LED, if the injected carriers arrive instantaneously at the recombination (diffusion) region, the rise time of the spontaneous emission is governed solely by the spontaneous recombination time of the carriers. However, in driving a practical diode, the junction capacitance and the stray capacitance cause delay in the arrival time of the injected carriers at the recombination region. Thus, the rise time of the spontaneous emission in an LED and the delay time of the stimulated

emission in a laser diode would be either (i) material-limited by the spontaneous recombination time τ_s or (ii) circuit-limited by the time constant τ_c of the driving circuit (including the junction capacitance of the diode). The space-charge capacitance effect has been reported previously in large-area GaAs LED's[18] and in $GaAs_{0.6}P_{0.4}$ LED's.[19]

In a previous report, Dawson and Burrus[20] measured the rise time of small-area GaAs LED's.[1,2] They showed that the rise time decreased with the capacitance of the unit. For instance, a GaAs diode with a junction capacitance of 200 pF had a rise time of 5.5 ns; when the capacitance was reduced to 20 pF by etching away the excess area of the junction, the rise time was reduced to 3.5 ns. They also observed that the rise time decreased with driving current. More recently, in doing the modulation experiment for LED's, Dawson employed a low-impedance driver to improve the speed of modulation. A rate of 48 Mb/s[21] with approximately 1-ns rise time was obtained using a driver with an output impedance of 2 to 4 ohms. At the higher rate of 280 Mb/s[22] a snap-diode driver with an effective impedance of 1 to 2 ohms was used to produce a train of short pulses. The apparent rise time of the observed optical pulses was 0.7 ns. However, the peak power was less than that for dc drive, since the LED was not fully turned on in this case. Also, White and Burrus[8] used a low output-impedance transistor driver that produced a light pulse with 2.5-ns rise time.

In view of the previous experimental results, the present work studies the junction capacitance effect in detail and correlates theory with experiments. The effect on the turn-on delay of injection lasers is studied similarly. The expected reduction of such delays by means of a preshaped driving pulse is calculated.

In practical applications for high modulation rates, a combination of the following techniques can be used: (i) constructing diodes to have low capacitance, (ii) employing drivers with low output impedance (to match the diode impedance) and (iii) properly shaping the driving current pulse.

One additional point in the results of this work is that, in determining the spontaneous recombination lifetime by the measurement of laser turn-on delay,[23] good results can be expected if (i) the junction capacitance is less than about 10 pF and (ii) the laser is operating in a single filament.[24,25] The latter requirement is necessary because the distribution of current among filaments and the determination of the threshold become ambiguous in diodes lasing in multifilaments.

II. DESCRIPTION OF THE DIODES AND THE EQUIVALENT CIRCUIT

The LED used in the study was a Zn-diffused n-type GaAs ($N_d > 3 \times 10^{18} \text{ cm}^{-3}$) diode made by C. A. Burrus.[1,2] In this structure,

p-n junction is formed in the entire wafer of about 1.44×10^{-3} cm^2 area, and the SiO$_2$ layer masks most of the p$^+$ surface except for a small window where the contact is made to the junction. The current is restricted to flow largely in the small contact region, resulting in a small emitting area of $2 - 3 \times 10^{-5}$ cm^2. In our investigation of the rise time, the current density in the primary emission region ranged from 5 kA/cm^2 to 80 kA/cm^2.

The laser diode we used was a GaAs–Al$_x$Ga$_{1-x}$As double-heterostructure device with a stripe contact.[26] This geometry forces the current to flow only in a narrow region, thereby confining the laser actions to the small part of the junction below the contact. The wafer had an area about 2.5×10^{-3} cm^2, whereas the stripe contact area was about 1.25×10^{-4} cm^2 (500 μm $\times$ 25 μm). Room temperature threshold current density was about 3.5 kA/cm^2.

The common feature of both diodes is that the large junction over the entire wafer presents a space-charge capacitance comparable to the diffusion capacitance of the junction.

Figure 1 shows the equivalent circuit for study of the transient behavior of both diodes. A nonlinear resistor represents the $I - V$ characteristic of the diode, described by

$$i_d = I_0 \left[\exp \left(\frac{q v_d}{n k T} \right) - 1 \right],$$ (1)

where v_d is the junction voltage. The factor n is unity for an ideal diode diffusion current but, in our heavily doped p-n junction in GaAs, n is approximately 2. In the transient case, part of the injected carriers will replenish the carriers that recombine, and the rest will build up as stored charges. Thus, the stored charges will increase as $(1 - e^{-t/\tau_s})$,
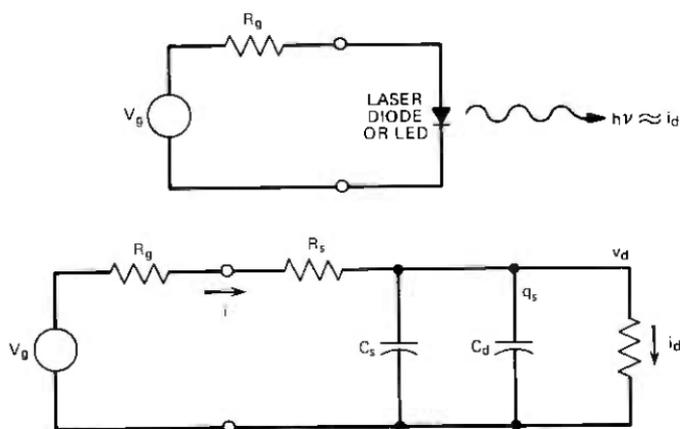

Fig. 1—The diode equivalent circuit.

where the time constant τ_s is the carrier spontaneous recombination time. This is equivalent to a capacitance, C_d, whose charge, $q_d(t)$, is equal to the stored charge and whose instantaneous charge-voltage relationship satisfies

$$q_d = Q_0 \left[\exp \left(\frac{qv_d}{nkT} \right) - 1 \right],$$ (2)

where v_d is the instantaneous junction voltage, and $Q_0 = I_0\tau_s$. To satisfy eq. (2), the diffusion capacitance is defined as

$$C_d = \frac{dq_d}{dv_d} = \tau_s \frac{di_d}{dv_d}.$$ (3)

In addition to the diffusion capacitance, there is the space-charge capacitance of the depletion layer. The space-charge capacitance, in parallel with C_d, is given by the usual form

$$C_s = \frac{(A - A_e)c_o}{(1 - v_{do}/\phi)^m} + \frac{A_e c_o}{(1 - v_d/\phi)^m},$$ (4)

where c_o is the zero-bias capacitance per unit area, ϕ is the barrier voltage which is 1.26 volts from $C - V$ measurement, the exponent m is $\frac{1}{2}$, A is the total area, and A_e is the emitting (contact) area. At small forward voltages, the current spreads over the entire area of the diode chip. The current confinement in the contact area is obtained only at a voltage v_{do}, at which the junction resistance is less than the spreading resistance. Thus, the first term in eq. (4) accounts for the space-charge capacitance of the junction surrounding the emitting area, while the second term is the space-charge capacitance of the emitting area itself. Since $A \gg A_e$ and $v_d \approx v_{do}$, eq. (4) can be approximated by

$$C_s = \frac{C_0}{(1 - v_{do}/\phi)^m},$$ (5)

where $C_0 = Ac_o$ is the total zero-bias capacitance.

Referring to Fig. 1 again, resistance R_s in the equivalent circuit accounts for the series resistance of the bulk material as well as the contact resistances. The driver output impedance is represented by R_g, and V_g is a rectangular voltage pulse amplitude. For completeness, the following circuit equations are included:

$$i = \frac{V_g - v_d}{R_g + R_s}$$ (6)

$$i_c = i - i_d$$ (7)

$$v_d = \int \frac{i_c}{C_s + C_d} \, dt.$$ (8)

Table I — Constants for diode parameters

	I_0	n^*	m	R_s	R_g	C_0
LED	2.3×10^{-12}	2	0.5	5.6 Ω	50 Ω	150–250 pF
Laser	1.2×10^{-13}	2	0.5	0.6 Ω	50 Ω	130–200 pF

* For ideal diodes, $n = 1$ for the diffusion current at low injection and $n \approx 2$ at high injection. In our GaAs with heavily doped n-region ($N_d > 3 \times 10^{18}$ cm^{-3}), n is usually equal to 2.

Numerical solutions to the system of eq. (1) through (8) have been obtained using the experimentally determined constants for both LED's and laser diodes, as tabulated in Table I. Since the spontaneous emission is proportional to the diffusion current, the time dependence of emission can be found from the time response of i_d.

III. EXPERIMENTAL SETUP

The diode was mounted at the end of a 50-ohm microstrip line (in series with either a 50-ohm resistor for the laser diode or a 47-ohm resistor for the LED chosen experimentally to minimize the reflections), and was driven by a Tektronix* 110 pulse generator that produced a rectangular pulse with a rise time less than 250 picoseconds and ripple below 1 percent. The pulse width used was 20 ns, and the repetition rate was 300 p/s. The output was detected by a p-i-n photodiode with a rise time better than 150 picoseconds and displayed on a 90-picosecond rise-time sampling oscilloscope.

IV. NUMERICAL RESULTS AND COMPARISON WITH EXPERIMENTS

4.1 Characteristics of LED response

The diffusion current, as a function of time, resulting from a step input voltage to the LED can be obtained by solving eqs. (1) to (8). Using the measured diode parameters given in Table I, we obtain the dependence of the rise time, between the 10- and 90-percent points of the light pulse, and the delay time between the application of the current pulse and the 10-percent point of the light pulse. These quantities are functions of driving current density and the junction capacitance. Such dependence for various values of zero-bias space-charge capacitance is shown in Fig. 2a. It is seen that, for very large current densities, the rise time is unaffected by C_0 and approaches an asymptotic value of $2\tau_s$, slightly faster than an ideal exponential rise of the form $(1 - e^{-t/\tau_s})$. The delay time is largely due to the space-charge capacitance as clearly shown in Fig. 2b, and it approaches zero at very

---

* Trade name of Tektronix Corporation.
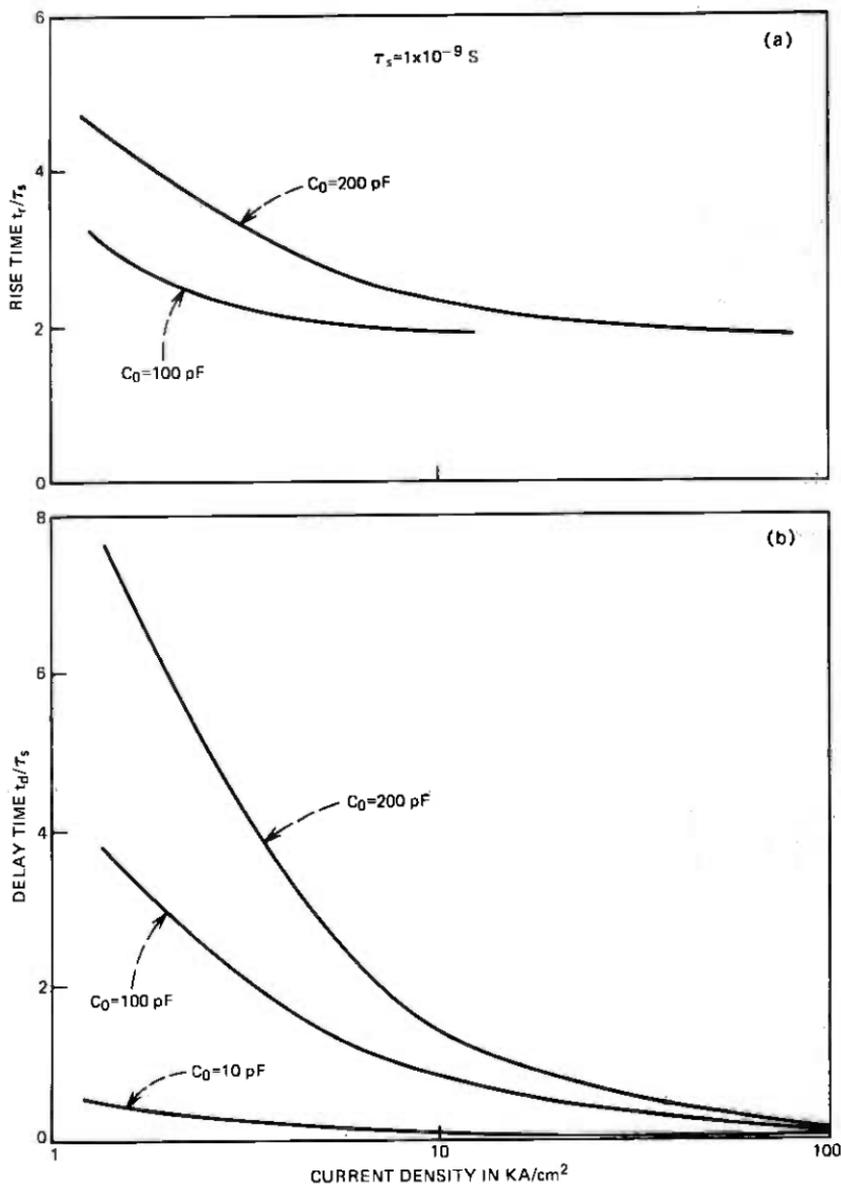
Fig. 2—(a) Rise time of LED output. (b) Delay time of LED output as a function of current density for various zero-bias diode capacitances.

large current densities. This is understandable since, for very large current densities, the space-charge would be completely neutralized. Such neutralization would require a current density on the order of

$$J = \frac{qN_dW_0}{\tau_s}\frac{A}{A_s},$$

(9)
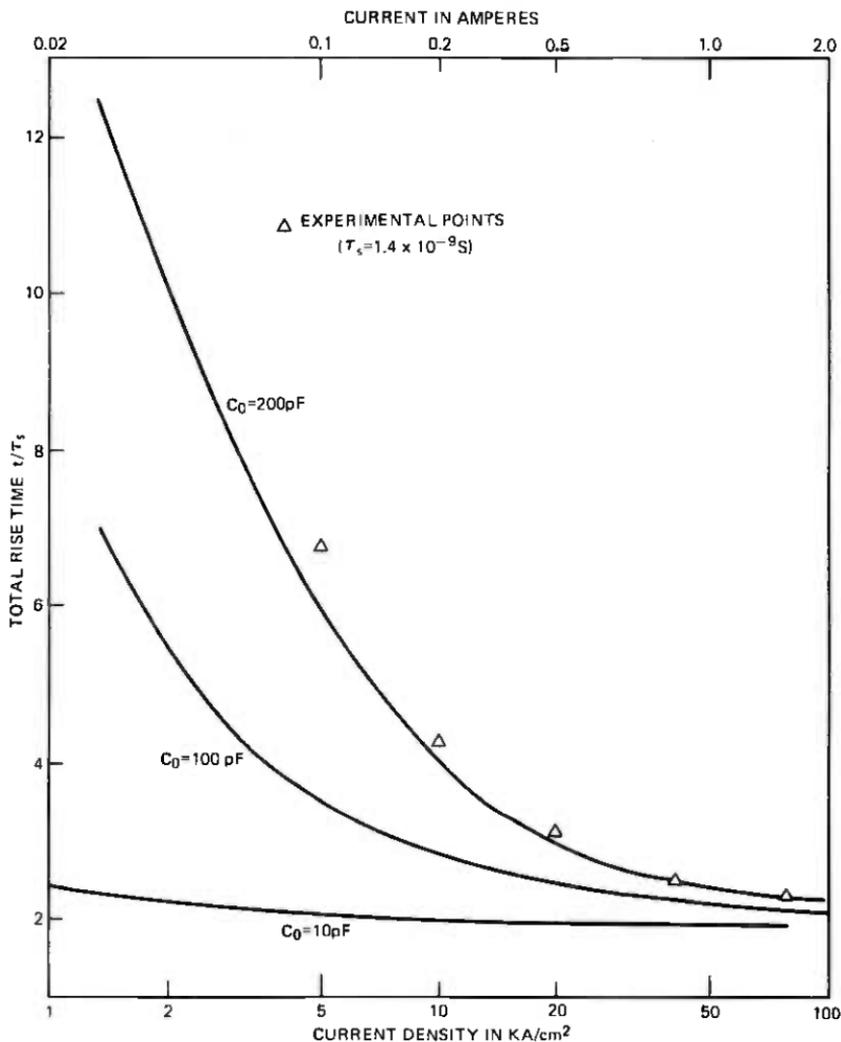
Fig. 3—Comparison of calculated total rise time $(t_d + t_r)$ with measured values.

where W_0 is the space-charge width at zero-bias, N_d is the net donor density, A is the total junction area, A_s is the light-emitting area where the current flows, $W_0 = \epsilon_0 \epsilon_r A/C_0 = 0.66 \times 10^{-5}$ cm (for $A = 1.2 \times 10^{-3}$ cm, $C_0 = 200$ pF, and $\epsilon_r = 12.8$). If we use $N_d = 3 \times 10^{18}$ cm^{-3} and $\tau_s = 1.4 \times 10^{-9}$, we see the estimated current density is 9×10^5 A/cm^2, a value in good agreement with the curves of Figs. 2a and 2b.

The experimentally measured total rise time $(t_d + t_r)$ indeed showed a current dependence similar to those calculated above. The spontaneous recombination time τ_s can be determined by noting that, at large current densities, the rise time approaches $2\tau_s$, independent of the junction capacitance. Figure 3 shows the experimentally observed

current-dependent total rise time normalized to τ_s plotted against the current density. It is clearly seen that the experimental points can be matched to a calculated curve if C_0 is assumed slightly larger than 200 pF. This value of C_0 is in very good agreement with the measured value.

4.2 Laser turn-on delay

Below lasing threshold, the accumulation of injected carriers in a laser is similar to that in an LED. Thus, the spontaneous emission from a laser has a similar response to the emission from an LED. When the current increases to a level at which the laser has gain equal to the total loss of the cavity and mirrors, laser action commences and the light output increases sharply. The stimulated emission reduces the spontaneous lifetime considerably, so that the rise time of the stimulated emission is much shorter than that of the spontaneous emission. Thus, the delay time of the stimulated emission is just the response time required for the diode current to reach the lasing threshold. This delay time can be determined as a function of $I/(I - I_{th})$ by calculating the response of the diffusion current for experimental diode parameters and by determining experimentally the threshold current of the laser.

Curves of delay vs current above threshold for various values of zero-bias diode capacitance are shown in Fig. 3. For $C_0 = 0$, the relationship is a straight line given by the usual delay-time formula[23]

$$t_d = \tau_s \ln \frac{I}{I - I_{th}}. \tag{10}$$

It is clear that the space-charge capacitance will further delay the laser turn-on time. For large C_0, the relationship deviates from the straight line, a phenomenon observed in past experiments* that has not been explained adequately. For $C_0 = 200$ pF, a typical value for the experimental diode, the delay time is about twice that for zero capacitance.

To determine τ_s experimentally from delay measurements, therefore, a correction for the effects of junction capacitance must be made. Using $C_s = 200$ pF, the corrected τ_s would be 1.8×10^{-9} for our particular laser diode. The measured delay time t_d, when normalized to τ_s as a function of $I/(I - I_{th})$, agrees well with the results calculated in this way.

4.3 Effect of driver output impedance

It is obvious that the resistance-capacitance time constant of the charging current can be shortened by either reducing the space-charge

---

* In past experiments, the result that the straight line did not pass the origin was attributed to filamentary lasing or multimoding (Ref. 24). Perhaps the effect of the capacitance is also contributing to the undesirable results (Ref. 25).
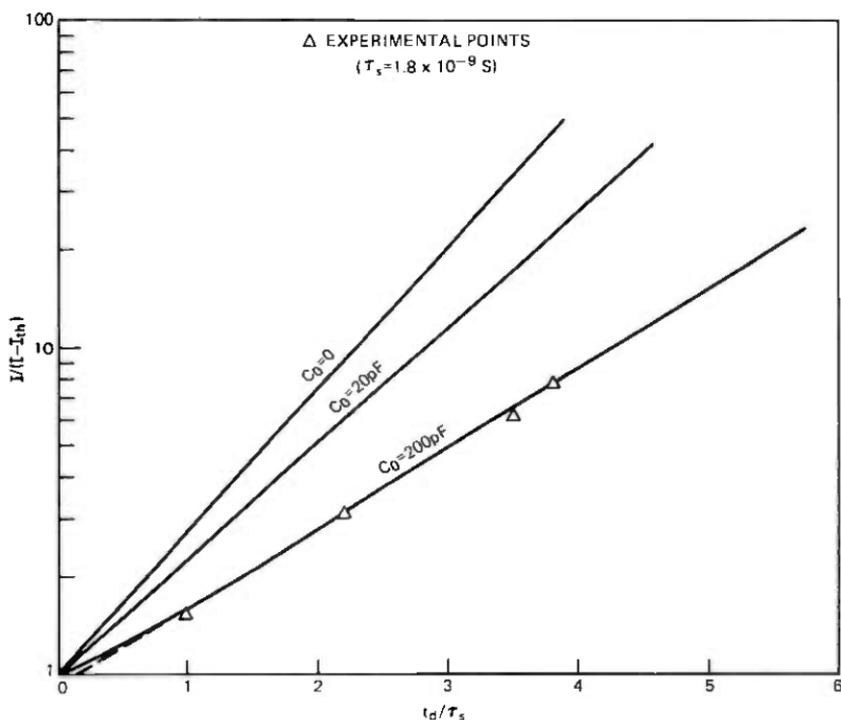
Fig. 4—Laser turn-on delay (normalized to the spontaneous recombination time τ_s) as a function of excitation currents above threshold for various zero-bias diode capacitances. Solid lines are calculated.

capacitance* or lowering the output impedance of the driver.[21,22] Using the same diodes, we varied the impedance of the driver by shunting the diode with a low resistance. Figure 4 shows a typical result of a laser diode. The turn-on delay time of the laser as a function of the shunt resistance for various driving currents and the luminescence rise time when the laser diode was operated below threshold were measured. To assure that the same amount of current was injected into the laser as the shunt resistance was changed, the light output was monitored first for 50 ohms (no shunt). For other shunt resistances, the driving current was increased until the same output was obtained. For a given diode current, the delay was shortened by lowering the driver impedance. The amount of reduction in delay reduced as the diode current increased. This is in line with the results indicated in Section 4.1 that the effect of space-charge capacitance decreased at higher currents. With

---

* J. C. Dyment (Ref. 25) has observed independently that proton-bombarded stripe-geometry lasers with shallow penetration have delay times almost 1.5 times that of units with deep penetration. The unit with shallow penetration had a capacitance of about 75 pF, while those with deep penetration had 15- to 20-pF capacitance.

a 1-ohm shunt resistance, the capacitance effect should be negligible. The results in Section 4.1 showed that the luminescence rise time would be about $2.2\tau_s$. Thus, we can deduce from the rise time $(4.1 \times 10^{-9}$ s) that the spontaneous recombination time $\tau_s = 1.85 \times 10^{-9}$ s. Again, it agreed with the delay time measurement mentioned in Section 4.2.

V. EQUALIZATION OF LASER TURN-ON DELAY

As discussed above, the accumulation of charges in the recombination region is approximately exponential with an effective time constant τ_{eff}. This time constant may be larger than or equal to the spontaneous lifetime τ_s, depending on whether the space-charge capacitance is significant. To reduce the delay from τ_{eff}, a driving current with amplitude many times higher than I_{th} would be necessary, and this would require a driver with high power capability if its output impedance is larger than that of the diode laser, which is about a few ohms. An alternative is to preshape the current pulse to equalize the delay. One of such a pulse shape, the exponential decay shown in Fig. 5a, has been investigated.

Using the exponentially decaying pulse, $I = I_a e^{-t/\tau_c} + I_0$, shown in Fig. 5 as the driving current pulse, and using the same approach as
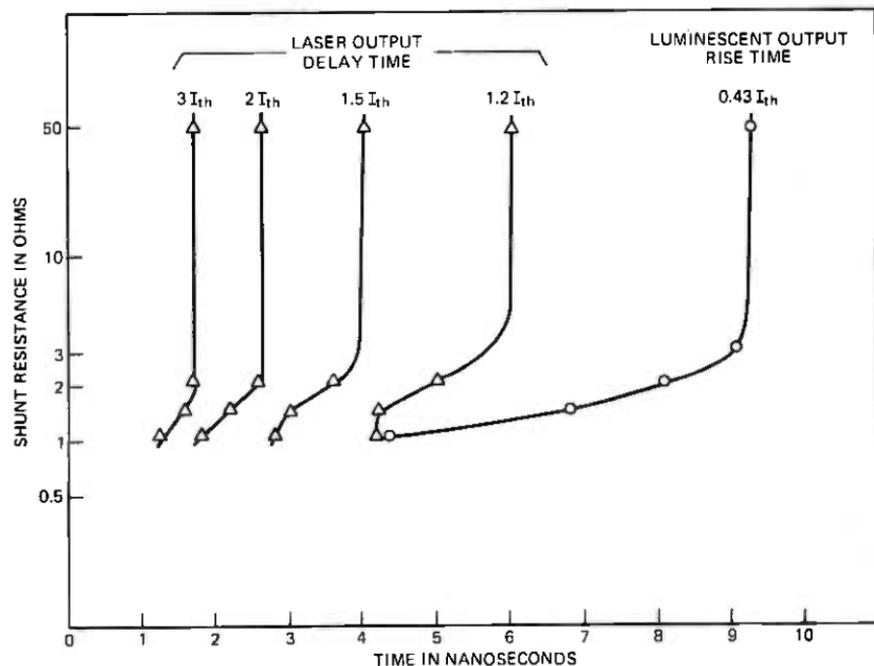
Fig. 5—Laser output delay time and luminescent rise time as a function of the shunt resistance for various excitation currents.

(a)

(b)
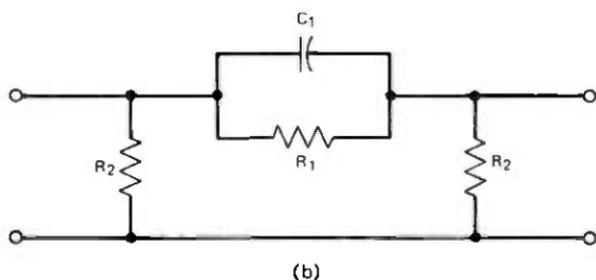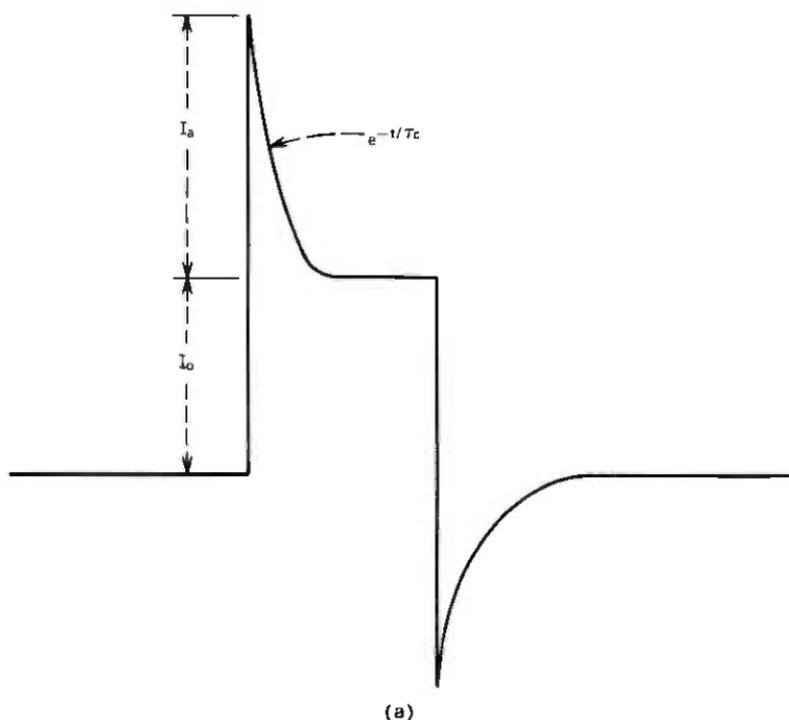
Fig. 6—(a) Current waveform for delay reduction of laser diodes. (b) The pulse shaping network used in the experiment.

in Ref. 23, the laser turn-on delay can be derived as

$$\frac{I_a}{I_{th}}\frac{e^{-T_d/T_c}}{1-T_d/T_c} - \left(\frac{I_a}{I_{th}}\frac{1}{1-1/T_c} + \frac{I_0}{I_{th}}\right)e^{-T_d} + \frac{I_0}{I_{th}} = 1, \quad (11)$$

where I_a is the amplitude of the exponential portion of the current pulse, I_0 is the constant portion of the current pulse, and I_{th} is the threshold current with a rectangular current pulse. $T_d = t_d/\tau_{eff}$ is the normalized delay time, and $T_c = \tau_c/\tau_{eff}$. Since the parameters I_a and τ_c are independent, we considered two cases: (i) Set $\tau_c/\tau_{eff} = 1$, and

find the decrease of t_d with I_a. (ii) Set $I_a = I_0$ and find the decrease of t_d with τ_c. The results are shown in Figs. 6a and 6b, respectively. In Fig. 7a, for each I_0, the curves start on the Y-axis at a value of t_d/τ_{eff} that equals the normalized delay for a rectangular pulse given by (10). As I_a increases, the delay decreases and approaches 0.1 τ_{eff} at $I_a = 8$ I_{th}. This implies that a sharp spike in the driving current pulse would help considerably in reducing the delay.

Figure 7b shows the reduction of t_d as a function of $T_c (= t_c/\tau_{eff})$. For a very small value of T_c, the exponential portion of the current is too fast for the charge to respond, and the reduction of t_d is marginal. For $T_c \gg 1$, t_d asymptotically approaches 0.287 τ_{eff} at $I_a = I_0 = 2 I_{th}$.
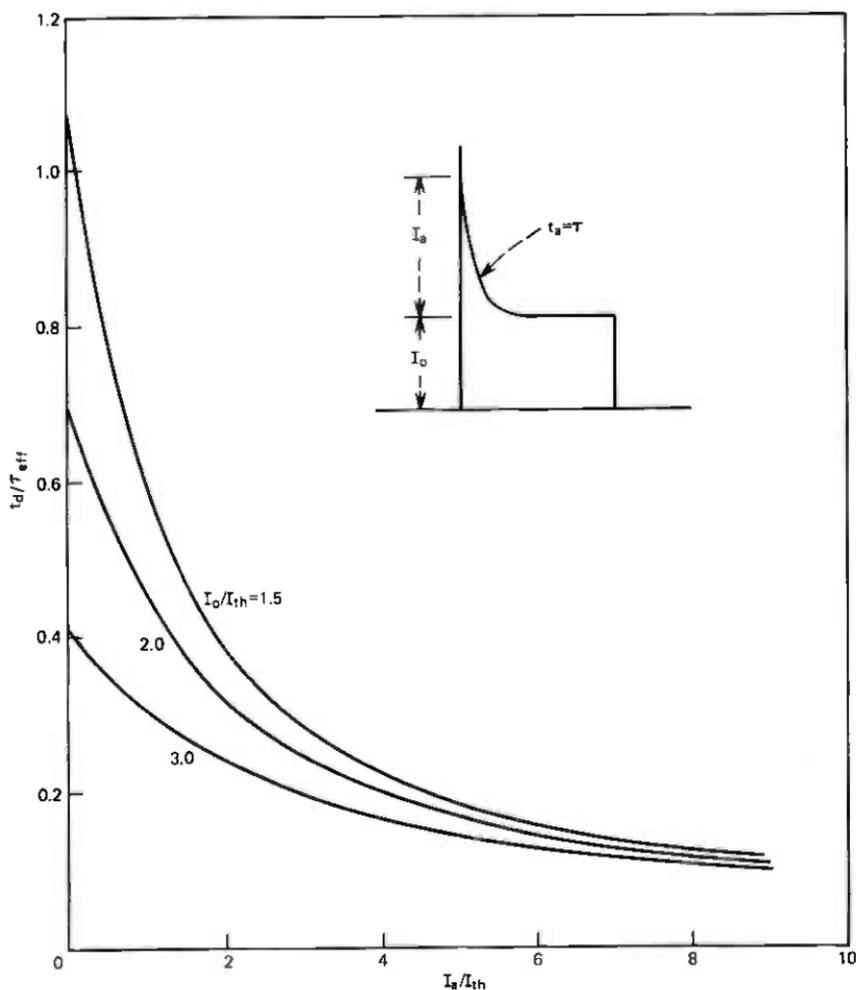
Fig. 7a—Reduction of delay time as a function of the peak pulse amplitude.
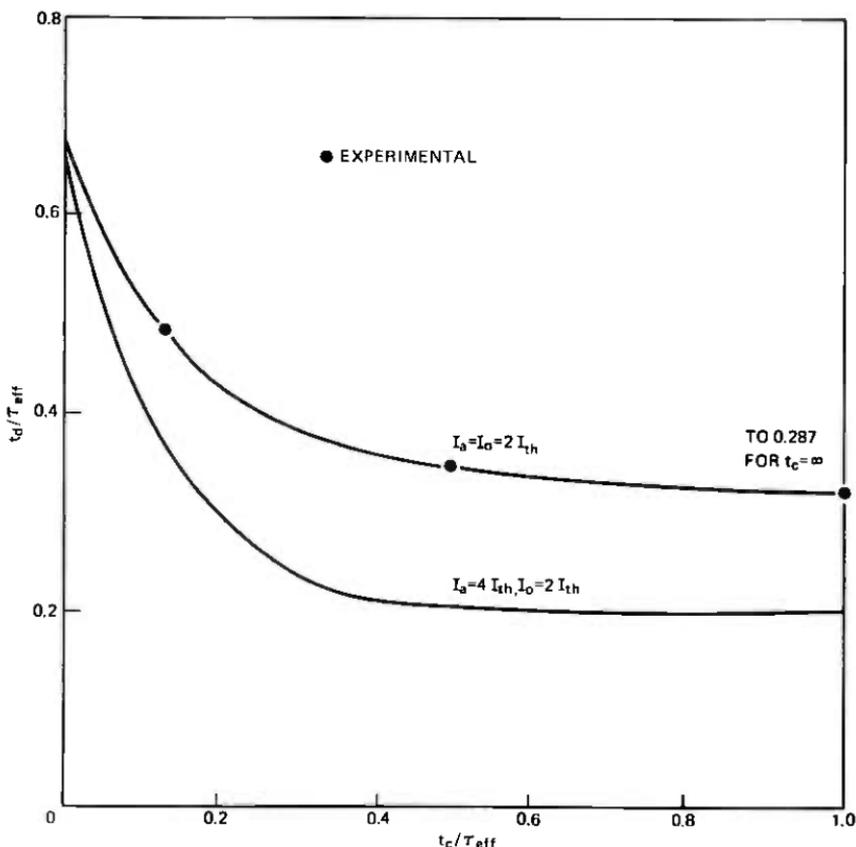
Fig. 7b—Reduction of delay time as a function of the pulse decaying time constant.

At $T_c = 1$, $t_d = 0.316 \, \tau_{eff}$. Thus, a factor of 2.2 reduction in t_d can be realized with this shaped pulse compared to that obtainable with a rectangular pulse.

Experimentally, we have realized the pulse shape in Fig. 6a by shaping a rectangular pulse with the network shown in Fig. 6b as an example. This pulse shape not only serves to reduce delay, but its negative after-pulse serves to turn the laser off faster than a rectangular pulse. The resistors R_1 and R_2 are necessary for matching to our 50-ohm measuring system. The capacitor C_1 may be chosen to provide different time constants. Table II summarizes the experimental and calculated results, and it can be seen that the agreement is excellent.

VI. CONCLUSIONS

Both theoretical and experimental investigations of the effect of junction capacitance on the rise time of LED's and the delay time of

Table II — Comparison of experimental and calculated delay reduction

I/I_{th}	Rectangular Pulse, $I_0 = 2I_{th}$		Exponential Pulse, $I_a = I_0 = 2I_{th}$					
	τ_{eff} (ns)	$t_d(I = 2I_{th})$ (ns)	t_c (ns)	R_1 (Ω)	R_2 (Ω)	C_1 (pF)	t_d Measured (ns)	t_d Calculated (ns)
	3.8	2.4	0.25	100	82	5	2.3	2.2
			0.5	100	82	10	1.6	1.8
			1.95	100	82	39	1.2	1.3
			3.9	100	82	79	1.0	1.1

injection lasers lead to the conclusion that the apparent response delays in these devices can be attributed largely to the shunting capacitance of the junction itself. The effects are greatest in devices operating at relatively low currents. Thus, in the more efficient double-heterostructure devices, the response time is much larger than would be expected from the spontaneous recombination time of the carriers. Homostructure devices, normally operated at much higher current densities, are affected less.[27]

Practically, the effects can be minimized by constructing devices to have low capacitance (by isolation of the active junction through the use of mesa structures or proton bombardment, for example) and by providing the lowest possible impedances in the driving circuits. Furthermore, in certain systems applications where short delay is required, a simple delay equalization (in the form of a shaped driving pulse) can be employed in conjunction with a driver that has low-output impedance.

It should be pointed out, however, that our definition of rise time is the time required for the light output to reach 90 percent of its fully developed steady-state value, when the diode is driven by a step current. In many practical applications, the LED can be modulated at much higher speed than that imposed by the spontaneous recombination time.[28] In such cases, the light output would not reach the steady-state value before the driving pulse is off, resulting in an apparently narrower output pulse but smaller amplitude. The price one pays, of course, is the reduced driving efficiency.[29]

In addition to the effects of the capacitance on the pulse modulation of these devices, it is also pointed out that great care must be exercised in determining the semiconductor carrier recombination time by measurement of the delay time of junction lasers. To assure good results,

the junction capacitance must be less than 10 pF, and the device must be operating in a single filament.

VII. ACKNOWLEDGMENTS

REFERENCES

1. C. A. Burrus, "Radiance of Small-Area High-Current-Density Electroluminescent Diodes," Proc. IEEE 60, No. 2 (February 1972), pp. 231–232.
2. C. A. Burrus and B. I. Miller, "Small-Area, Double-Heterostructure Aluminum-Gallium-Arsenide Electroluminescent Diode Sources for Optical-Fiber Transmission Lines," Opt. Comm., 4, No. 4 (December 1971), pp. 307–309.
3. I. Hayashi, "Progress of Semiconductor Lasers in Japan," Digest of Technical Papers, 1973 IEEE/OSA Conference on Laser Engineering and Applications, Washington, D.C., May 30–June 1, 1973, p. 69.
4. B. C. DeLoach Jr., "Reliability of GaAs Injection Lasers," Digest of Technical Papers, 1973 IEEE/OSA Conference on Laser Engineering and Application, Washington, D.C., May 30–June 1, 1973, p. 70.
5. F. P. Kapron, D. B. Keck, and R. D. Mauer, "Radiation Loss in Glass Optical Waveguides," Appl. Phys. Lett. 17, No. 10 (November 1970), pp. 423–425.
6. Corning Glass Works, Corning, N.Y., press release, August 14, 1972.
7. J. E. Goell, "A Repeater with High Input Impedance for Optical-Fiber Transmission Systems," Digest of Technical Papers, 1973 IEEE/OSA Conference on Laser Engineering and Applications, Washington, D. C., May 30–June 1, 1973, p. 23.
8. G. White and C. A. Burrus, "Efficient 100 Mb/s Driver for Electroluminescent Diodes," Int. J. Elec., 35, No. 6 (December 1973), pp. 751–754.
9. D. M. Henderson and P. K. Runge, "50 Mbits/s Driving Circuits for LED with Improved Efficiencies," unpublished work.
10. P. K. Runge, "A 50-Mbits/s Optical PCM Repeater," International Conference on Communications, Minneapolis, Minn., June 17–19, 1974.
11. T. Ozeki and T. Ito, "Pulse Modulation of DH-(GaAl)As Lasers," IEEE J. Quantum Elec., QE-9, No. 2 (February 1973), pp. 388–391.
12. T. Ozeki and T. Ito, "A 200-Mb/s PCM DH- (GaAl)As Laser Communication Experiment," Digest of Technical Papers, 1973 IEEE/OSA Conference on Laser Engineering and Applications, Washington, D. C., May 30–June 1, 1973, p. 74.
13. W. J. Clemetson and W. O. Schlosser, "An Experimental 300-Mbit GaAs Laser Glass-Fiber Transmitter," unpublished work.
14. J. P. Beccone, W. J. Clemetson, and W. O. Schlosser, "300 M Baud On-Off Modulation of AlGaAs Lasers with a Pseudo Random Word," unpublished work.
15. J. E. Goell, "A 274-Mb/s Optical-Repeater Experiment Employing a GaAs Laser," Proc. IEEE, 61, No. 10 (October 1973), pp. 1504–1505.
16. H. Thim, L. R. Dawson, J. V. DiLorenzo, J. C. Dyment, C. J. Hwang, and D. L. Rode, "Subnanosecond PCM Lasers by Gunn-Effect Switches," Digest of Technical Papers, International Solid State Circuit Conference (February 1973), pp. 92–93.
17. M. Chown, A. R. Goodwin, D. F. Lovelace, G. H. B. Thompson, and P. R. Selway, "Direct Modulation of DH Lasers at Rate Up to 1 Gbit/s," Elec. Lett., 9, No. 2 (January 1973), pp. 34–36.
18. A. G. Dmitriev and B. V. Tsarenkov, "Emission Kinetics of Electroluminescent Diodes," Soviet Phys.—Semiconductors, 5, No. 8 (February 1972), pp. 1307–1313.

19. S. Nakamura, J. Umeda, and O. Nakada, "Response Times of Light-Emitting Diodes," IEEE Trans. on Electron Devices (Correspondence), *19*, No. 8 (August 1972), pp. 995–997.
20. R. W. Dawson and C. A. Burrus, "Pulse Behavior of High-Radiance Small Area Electroluminescent Diodes," Appl. Opt., *10*, No. 10 (October 1971), pp. 2367–2369.
21. R. W. Dawson, private communication.
22. R. W. Dawson, private communication.
23. K. Konnerth and C. Lanza, "Delay Between Current Pulse and Light Emission of a GaAs Injection Laser," Appl. Phys. Letters, *4*, No. 7, 1964, pp. 120–121.
24. C. J. Hwang and J. C. Dyment, "Dependence of Threshold and Electron Lifetime on Acceptor Concentration in GaAs-Al$_x$Ga$_{1-x}$As Lasers," J. Appl. Phys., *44*, No. 7 (July 1973), pp. 3240–3244.
25. J. C. Dyment, private communication.
26. J. E. Ripper, J. C. Dyment, L. A. D'Asaro, and T. L. Paoli, "Stripe-Geometry DH Junction Lasers, Mode Structure and CW Operation Above Room Temperature," Appl. Phys. Lett., *18*, No. 4 (February 1971), pp. 155–157.
27. J. C. Dyment, J. E. Ripper, and T. P. Lee,"Measurement and Interpretation of. Long Spontaneous Lifetime in Double-Heterostructure Lasers," J. Appl. Phys., *43*, No. 2 (February 1972), pp. 452–457.
28. R. W. Dawson, "Pulse Widening in a Multimode Optical Fiber Excited by a Pulsed GaAs LED," J. Appl. Opt., to be published.
29. C. A. Burrus, T. P. Lee, and W. S. Holden, "Direct Modulation Efficiency of LEDs for Optical Fiber Transmission Applications," to be published, Proceeding Letters, IEEE.

A Depletion-Load, p-Channel, Bipolar-IGFET Technology

By G. MARR and G. L. MOWERY

(Manuscript received April 25, 1974)

A high-performance, depletion-load, bipolar-IGFET technology is described. The optimization of device and circuit parameters, the ion-implanted depletion-load fabrication process, and the high-speed input and output circuits, which allow direct interface with the TTL circuit family, are discussed.

I. INTRODUCTION

Since the development of the low-threshold ($V_T = -1$ volt) metal-gate p-channel Al_2O_3-SiO_2 IGFET technology,[1] further technological advances have been made to improve circuit speed and interface flexibility of IGFET circuits. Two significant improvements have been the incorporation of bipolar-IGFET (BIGFET)[2] devices on the same monolithic silicon-integrated circuit (SIC) and the application of ion implantation to fabricate depletion-load IGFET's.[3] These have resulted in an IGFET technology whose circuits are completely compatible with the standard 5-volt bipolar TTL family in terms of input and output levels as well as power supply. Furthermore, the relatively high functional packing density of the IGFET technology makes it an extremely attractive option in MSI/LSI applications. These have been successfully realized in the design and fabrication of several Bell System catalog and custom circuits.

In this paper, the technology characteristics of the depletion-load BIGFET technology are described.

II. DEPLETION-LOAD INVERTER

2.1 Circuit characteristics

The advantages of the depletion-load inverter over a similar all-enhancement or resistor-load inverter become obvious when we consider the various load-line characteristics. In Fig. 1, the load lines of a resistive load, an enhancement-IGFET load, and a depletion-IGFET load are shown over the family of IV curves for the enhancement-
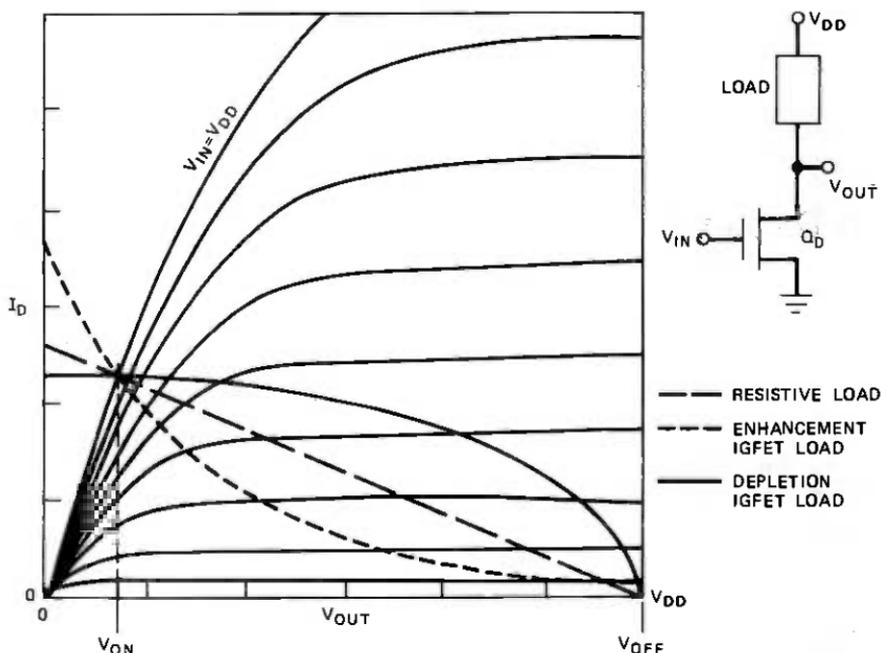
Fig. 1—Inverter load-line characteristics.

IGFET driver (Q_D). The three load lines are chosen to intercept at the same point in the $V_{in} = V_{DD}$ and $V_{in} = 0$ characteristics curves to reflect the fact that they will produce the same output dc levels. The area under the load line represents the transient power dissipated during one logic transition. For a given capacitive loading on the inverter output node, the transient power dissipation is proportional to the frequency of operation. Therefore, the area under the load line is directly proportional to the maximum operating frequency of the inverter. The speed advantage of the depletion load is apparent because it has the largest area between $V_{out} = V_{on}$ and $V_{out} = V_{off}$.

2.2 Circuit equations

In this section, the approximate circuit equations of the depletion-load inverter are reviewed. In spite of their inexactness, they are useful because they provide insight into the functional dependence of circuit parameters on device parameters. However, the final choice of device parameters, discussed in Section 2.3, was based on more exact computer calculations.

Assuming the gradual-channel approximation, the output levels $(V_{on}$ and $V_{off})$ and the corresponding single-stage noise margins $(V_{M1}$ and $V_{M2})$ of a depletion-load inverter with Q_D on and off, respectively,
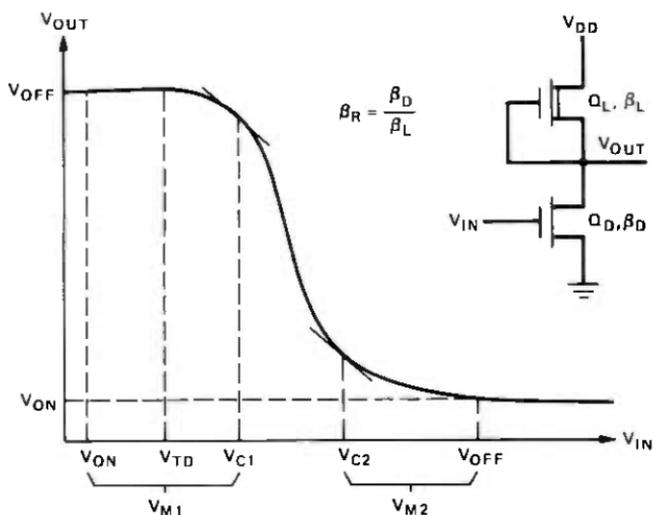
Fig. 2—Depletion load inverter.

as shown in Fig. 2, may be given by[4]

$$V_{\text{on}} \cong \frac{V_a^2}{2(V_{DD} - V_{TD})} \tag{1}$$

$$V_{\text{off}} = V_{DD} \tag{2}$$

$$V_{M1} = |V_{C1} - V_{\text{on}}| \tag{3}$$

$$V_{M2} = |V_{C2} - V_{\text{off}}|, \tag{4}$$

where

$$V_{C1} = \frac{V_a}{(1 + \beta_R)^{\frac{1}{2}}} + V_{TD} \tag{5}$$

$$V_{C2} = \sqrt{\tfrac{4}{3} \cdot V_a} + V_{TD} \tag{6}$$

$$V_a = \sqrt{\frac{V_{TL}^2}{\beta_R}}, \tag{7}$$

and where

V_{TD}, V_{TL} = threshold voltages of Q_D and Q_L, respectively,
$\quad V_{DD}$ = supply voltage,
$\beta_D, \beta_L, \beta_R$ = IGFET gain parameter of driver and load and their ratio.

The propagation delay (t_d) is directly related to the inverter-output rise time (t_r) and fall time (t_f). Since the rise-time charging current is given by $I_r = (\beta_L/2) \cdot (V_{TL})^2$, t_r may be approximated as

$$t_r = \frac{2V_{DD}C}{\beta_L V_{TL}^2}, \tag{8}$$

where C is the output capacitance and, neglecting interconnection capacitance, may be expressed in terms of β_D as

$$C = \frac{l_D^2 \beta_D}{\mu} (1 + F),\qquad(9)$$

where l_D is the channel length of Q_D, μ is the mobility of carriers in the channel, and F is the number of similar inverter gates to be driven (fan-out). If interconnection capacitance is expressed as equivalent fan-out, eq. (9) is still applicable. Substituting (9) into (8) results in

$$t_r = \frac{2V_{DD}\beta_R l_D^2}{\mu \cdot V_{TL}^2} (1 + F).\qquad(10)$$

Likewise, since the fall-time discharging current,

$$I_f = (\beta_D/2)(V_{DD} - V_{TD})^2 - (\beta_L/2)V_{TL}^2,$$

the fall time of the output voltage may be approximated as

$$t_f = \frac{2V_{DD}l_D^2(1 + F)}{\mu[(V_{DD} - V_{TD})^2 - V_{TL}^2/\beta_R]}.\qquad(11)$$

Since t_d is proportional to $(t_r + t_f)$, we observe from (10) and (11) that t_d is directly proportional to l_D^2, F, and V_{DD}. A more important observation, perhaps, is that as V_{TL}^2/β_R increases, t_r decreases and t_f increases. Thus, depletion-load inverters designed for optimal speed have t_r and t_f nearly equal.

A final useful relationship is the power-delay product of the inverter. Using the average power for P_D for small V_{TL}^2/β_R,

$$P_D t_d \sim P_D t_r \sim (V_{DD}I_L)\left(\frac{V_{DD}C}{I_L}\right) \sim V_{DD}^2 C \sim \frac{V_{DD}^2 l_D^2 \beta_D(1 + F)}{\mu},$$

where C_D is the gate capacitance of Q_D, and $P_D t_d$ is completely independent of V_{TL}. However, for larger values of V_{TL}^2/β_R,

$$P_D t_d \sim P_D t_f \sim \frac{V_{DD}^2 \beta_D l_D^2(1 + F)}{\mu\left[\dfrac{\beta_R(V_{DD} - V_{TD})^2}{V_{TL}^2} - 1\right]},$$

and $P_D t_d$ increases quadratically as $|V_{TL}|$ increases until $|V_{TL}| \sim \sqrt{\beta_R}|V_{DD} - V_{TD}|$.

As we stated earlier, the preceding approximate equations provide insight into the dependence of circuit parameters on device parameters for a depletion-load IGFET inverter. This is appropriate for characterizing the properties of any depletion-load IGFET random logic circuit, normally comprising an assortment of n-input NAND and NOR gates,

because both these logic gates can be analytically reduced to an equivalent depletion-load IGFET inverter.

2.3 Selection of V_{TL} and β_R

In this section, the method of determining the parameters of the depletion-load IGFET and the corresponding β_R are discussed. Since the equations developed in Section 2.2 are only approximate, more exact calculations were made using the circuit analysis program SERVICE.[5] In this way, the voltage dependence of V_{TL} was properly accounted for.

The objectives and constraints of the technology were chosen to be:

(*i*) Single-supply voltage $V_{DD} = -5 \pm 0.5$ volt.

(*ii*) $V_{TD} = -1 \pm 0.4$ volt at end of life.

(*iii*) $|V_{\text{on}}| \leqq |V_{TD}|$.

(*iv*) V_{M1} and $V_{M2} \geqq 0.5$ volt.

The problem is then, simply: Given these conditions, determine β_R and V_{TL} such that t_d is nearly a minimum for an assumed typical fan-out of five from each inverter. The net result is a family of t_d vs. V_{TL} curves. Since V_{on}, V_{M1}, and V_{M2} are functions of V_{TL}, a different β_R, the minimum value consistent with conditions (*iii*) and (*iv*) was chosen for every value of V_{TL} used. A nominal t_d vs V_{TL} curve is shown in Fig. 3, where the $P_D t_d$ product is also plotted against the same V_{TL} scale. As can be seen from the curve, the value of t_d reaches a nearly minimum constant value for $V_{TL} \sim 4.5$ volts, when $t_r \sim t_f$. We may further note that the $P_D t_d$ product increases approximately quadratically as V_{TL} increases. Based on this type of plot of all com-
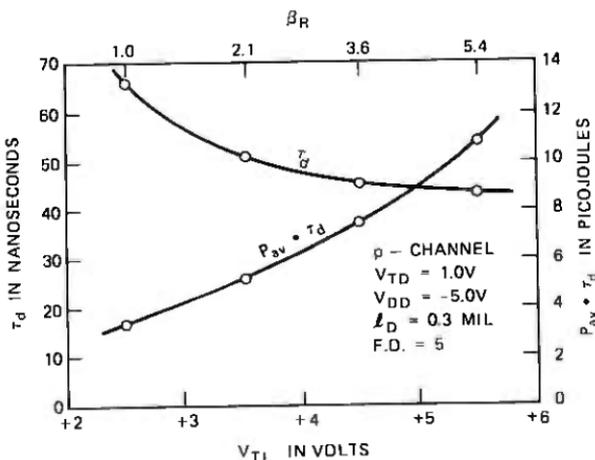
Fig. 3—τ_d and $P_{av} \cdot \tau_d$ vs. V_{TL} curve.
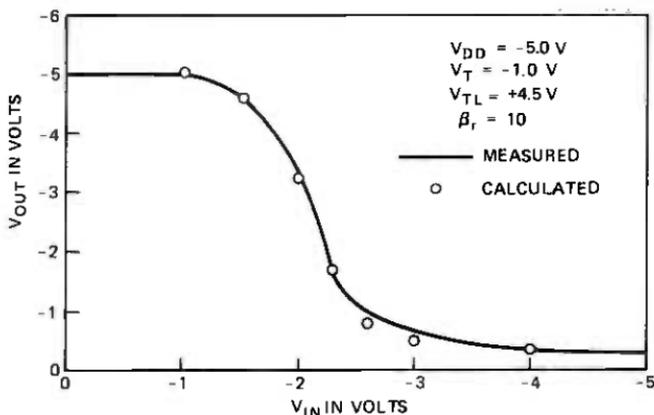
Fig. 4—Depletion-load inverter transfer curve.

binations for different limits of V_{DD} and V_{TD}, an optimal $V_{TL} = 4.5$ $\pm\ 0.5$ volt was chosen, corresponding to a minimum design β_R of 5.

Figure 4 shows a calculated and a measured dc transfer curve. The good agreement between the two is an indication of the validity of the present work.

III. DEPLETION-LOAD FABRICATION AND CHARACTERISTICS

A proven method of fabricating depletion IGFET's is the use of a selective ion implant through the gate dielectric in the IGFET channel regions. To achieve good control of the device properties, it is desirable to have the majority of the implanted ions in the silicon rather than in the gate dielectric, which requires 120 keV for the boron ions used. The resulting device no longer behaves exactly like the normal IGFET
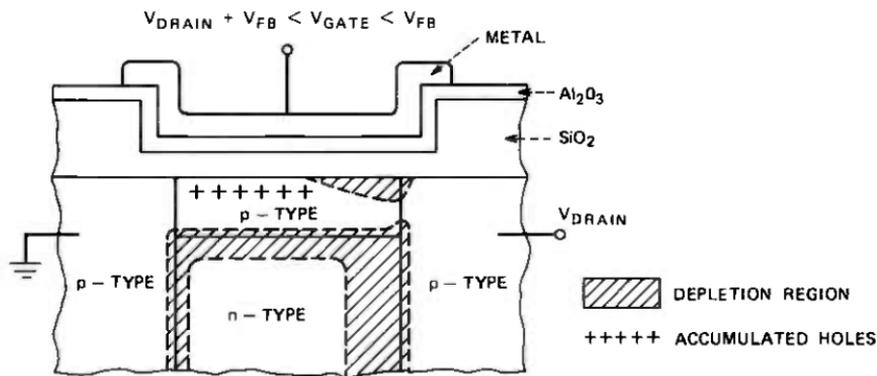
Fig. 5—Depletion IGFET cross-sectional view.
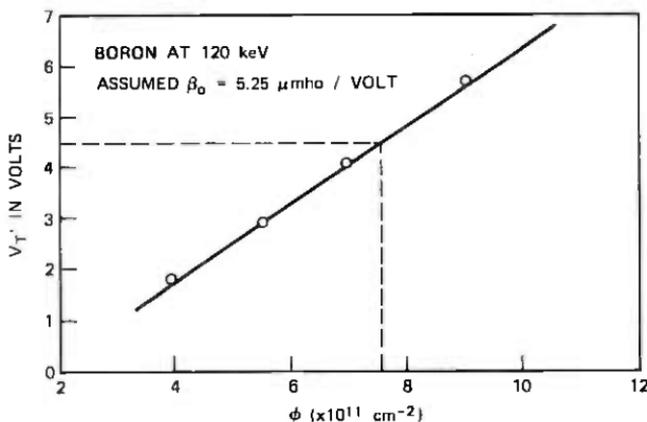
Fig. 6—Effective threshold vs. implant dose.

in which conduction takes place only in the surface inversion. It is actually a deep depletion IGFET with a subsurface channel, whose conductivity is modulated both by surface depletion and by the channel to substrate junction as shown in Fig. 5. The current-voltage characteristics of such a deep depletion IGFET have been previously described.[3]

To achieve a reproducible profile of the ion-implanted dopants in the silicon, it is desirable to minimize the duration and extent of high temperature anneal treatment after the ion implantation. Since all the diffusion processes take place before the gate oxidation, the latter is the last high temperature step. Therefore, ion implantation is performed after the gate oxidation process. The photomask step for the boron implantation utilizes a thick photoresist layer (1.4 μm) that is opaque to the 120-keV boron ions. Furthermore, since the required implantation dose of 7.6×10^{11} cm^{-2} is rather low, no damage to the gate dielectric is either expected or observed in the fabricated samples. The needed anneal[6] of implanted species is achieved in a subsequent process step when the masking SiO_2^* is deposited over the whole wafer at 900°C for 20 minutes.

Figure 6 shows the effective threshold $V'_T = \sqrt{(2I/\beta)}$ as a function of the implantation dose of boron at 120 keV. The current is measured when gate and source are common and the drain biased at -5 volts. We observe that the $V'_T - Q$ curve is essentially linear over a fairly wide range of dose (2×10^{11} to 9×10^{11} cm^{-2}).

---

* The metal-gate IGFET (Ref. 1) has a double layer Al_2O_3-SiO_2 for its gate dielectric. The deposited oxide is needed to serve as a mask when Al_2O_3 is selectively etched.

IV. INTERFACE CIRCUITS

To allow realization of the speed advantages offered by the depletion-load and BIGFET technologies, two buffer circuit configurations have been developed to facilitate efficient interface between inputs and outputs of BIGFET SIC's as well as to inputs and outputs of TTL(L) circuits. These two circuits are the level-shifting input circuit and the low-impedance output driver.

4.1 Input interface

It is necessary that the input voltage to an IGFET inverter be within one device threshold of substrate potential to assure complete turn-off of that inverter. Since this requirement is not met for the 1-volt threshold p-channel technology by either TTL or BIGFET output levels, the input level shifting circuit has been designed to accept these logical 1 levels. The schematic of the input circuit is shown in Fig. 7. The corresponding dc transfer curves are shown in Fig. 8. The internal feedback configuration of the level shifter produces a very sharp gain characteristic that lowers the required level for logic 1 and also provides increased noise margin and input signal wave shaping. In the design of this circuit, both the β ratio of $Q1$ to $Q2$ and the β ratio of $Q3$ to $Q4$ strongly affect the acceptable input voltage range. The β ratio of $Q1$ to $Q3$ is not critical, but may be chosen to produce minimum propagation delay of signals through the input level shifter in a specific application.

A further constraint on the β values of $Q1$ and $Q2$ is the fan-out desired from TTL(L) outputs and from BIGFET outputs into this level shifting input. During a 1 to 0 transition of the input voltage, IGFET's $Q1$ and $Q2$ conduct current from the substrate to the input node. The requirement of conducting this current to ground while maintaining a valid 0 level at the input determines the fan-out from a given output to the level shifting input circuits. Maximum 0 level V_{in} to the level

β (Q1) = 12 μ mhos/V
β (Q2) = 12 μmhos/V
β (Q3) = 6 μmhos/V
β (Q4) = 30 μmhos/V
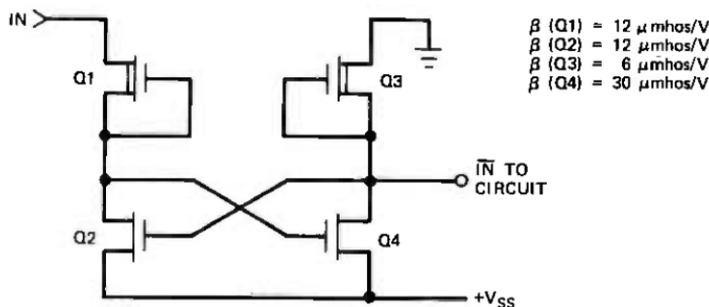
Fig. 7—Inverting level shifting input circuit.
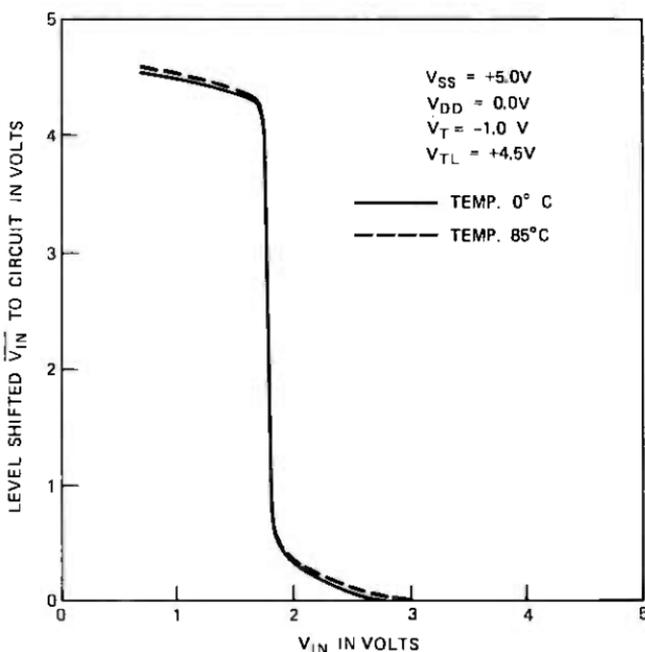
Fig. 8—Inverting level shifter transfer curves.

shifting input inverter is $V_{SS} - 3.4$ volts, which is 1.1 volts for $V_{SS} = 4.5$ volts. Allowing 0.4-volt interface noise margin requires that the input voltage be ≤ 0.7 volt for the 0 level. The TTL(L) output is rated at 28 mA to ground at 0.7 volt. The level shifting input inverter of Fig. 7 can conduct $\beta \times 4.2\ \mu A$ during a 1 to 0 transition, where $\beta = \beta_{Q1}$ or β_{Q2} is expressed in micromhos per volt. Therefore, the fan-out from TTL(L) to level shifting inputs is fan-out = $6667/\beta$, where $\beta = \beta_{Q1}$ or β_{Q2}. Similarly, the BIGFET output with maximum output resistance of 2.1 kilohms can conduct 0.333 mA at 0.7 volt. Therefore, fan-out from BIGFET outputs to level shifting inputs is fan-out = $79/\beta$, where $\beta = \beta_{Q1}$ or β_{Q2}.

The circuit of Fig. 7 has been used in various depletion-load circuits for both custom and general-purpose applications. It provides

	Min.	Max.
$V_{in}\ 0$	0.0	$V_{SS} - 3.4$ volts
$V_{in}\ 1$	$V_{SS} - 2.2$	V_{SS} volts.

4.2 Output Interface

The BIGFET output circuit configuration is shown in Fig. 9. This circuit provides an ideal coupling between the high-impedance IGFET's and the desired low-output impedance. The current gain provided by
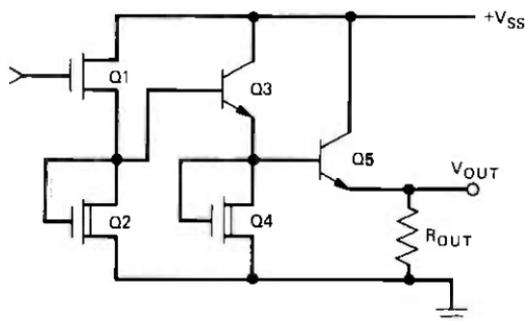
Fig. 9—BIGFET SIC output circuit.

the Darlington BIGFET structure permits the gate capacitance of $Q1$ to be as small as that normally used internally in an IGFET circut. Since average power dissipation of the output circuit is $\propto (R_{out})^{-1}$, it is desirable to make R_{out} large. But the maximum allowed value of R_{out} is determined by the input characteristics of TTL(L), -240 μa at 0.4-volt input logical 0. Consequently, R_{out} is nominally 1.45 kilohms, and the other device values are as shown in Fig. 9 for this BIGFET output configuration with fan-out $= 1$ capability to TTL(L) circuits.

The value of R_{out}, the V_{be} voltage drops of $Q3$ and $Q5$ operating as a Darlington pair, and the BIGFET emitter area of $Q5$ are designed to prevent $Q5$ from entering the saturation region, which would significantly increase the turn-off time of $Q5$.

The circuit of Fig. 9 has been used in various depletion-load circuits for both custom and general-purpose applications. It provides

	Min.	Max.
V_{out} 0	0.0	0.4 volt
V_{out} 1	$V_{SS} - 1.7$	$V_{SS} - 1.2$ volts.

V. SUMMARY

The p-channel depletion-load BIGFET process has been demonstrated to be a successful technology for producing BIGFET SIC's. Circuit behavior has been approximately described using simplified circuit equations, and exact circuit behavior has been characterized using results from computer circuit analysis programs. Good agreement has been achieved between experimental results and circuit simulation. The application of ion implantation to the fabrication of depletion mode IGFET load devices has been described, and a correlation is shown between the effective threshold voltage and implant dose. Finally, the input and output buffer circuits, which allow interface between BIGFET SIC's and TTL(L) circuits, are presented.

REFERENCES

1. H. E. Nigh, J. Stach, and R. M. Jacobs, "A Sealed Gate IGFET," Solid State Dev. Res. Conf., Santa Barbara, California, June 19–21, 1967.
2. G. Marr, G. T. Cheney, E. F. King, and E. G. Parks, "A Fast Bipolar-IGFET Buffer-Driver," B.S.T.J., *51*, No. 2 (February 1972), p. 363.
3. J. R. Edwards and G. Marr, "Ion Implanted Depletion-Mode IGFET," IEEE Trans. Electron Dev., *20*, No. 3 (March 1973), p. 283.
4. T. Masuhara, M. Nagata, and N. Hashimoto, "A High Performance n-channel MOS LSI Using Depletion-Type Load Elements," IEEE J. of Solid State Circuits, *SC-7*, No. 3, 1972, p. 224.
5. J. D. Leggett and E. B. Kozemchak, "The Domain Simulation of Non-Linear Networks by Nodal Analysis," Conf. Record of 5th Asilomar Conference on Circuits and Systems, November 1971, pp. 531–534.
6. T. E. Seidel and A. U. MacRae, Trans. AIME, *245*, 1969, p. 491.

A Speaker-Independent Digit-Recognition System

By M. R. SAMBUR and L. R. RABINER

(Manuscript received July 9, 1974)

This paper describes an implementation of a speaker-independent digit-recognition system. The digit classification scheme is based on segmenting the unknown word into three regions and then making categorical judgments as to which of six broad acoustic classes each segment falls into. The measurements made on the speech waveform include energy, zero crossings, two-pole linear predictive coding analysis, and normalized error of the linear predictive coding analysis. A formal evaluation of the systems showed an error rate of 2.7 percent for a carefully controlled recording environment and a 5.6 percent error rate for on-line recordings in a noisy computer room.

I. INTRODUCTION

With the widespread growth in the use of digital computers, there has been an increasing need for man to be able to communicate with machines in a manner more naturally suited to humans. The realization of this need has motivated a great deal of research in automatic recognition of speech by computer.[1-3] Although only a moderate degree of success has been obtained in solving the problems associated with machine recognition of continuous speech,[4] a greater degree of success has been obtained in recognition of isolated words from a fixed vocabulary. The performance of these systems range from about 92 percent correct decisions for 561 isolated words by an individual for which the system has been carefully trained[5] to nearly error-free performance for the recognition of a limited vocabulary (e.g., the digits) also spoken by a speaker for which the system has been trained.[6] However, performance of many of these word-recognition algorithms is radically degraded when the system has not been tuned to the speech characteristics of the individual user. The subject of this paper is an isolated-word, digit-recognition system that achieves high accuracy without having to be trained every time a different speaker wishes to use the system.

The development of a speaker-independent limited-vocabulary word recognizer is inherently more difficult than a speaker-adaptive system that can use comparatively simple pattern-matching algorithms to recognize the input words. It has been argued that the extended effort needed to design a speaker-independent system is unnecessary in view of the relative ease of training an adaptive scheme to learn to recognize the speech of a new user. There are two major reasons why such arguments are invalid. For small vocabulary systems (e.g., digit recognizers) with a *large number* of potential users, it is not feasible to store training data for every possible user. Furthermore, most systems cannot train themselves on new speakers very rapidly. Thus, the turnaround time for new users is often a major factor limiting the use of speaker-dependent systems. For a large vocabulary (250 words), the time required for a new speaker to form reference patterns for all the words in the vocabulary can be prohibitive. In addition, the variation with time of a speaker's voice characteristics may necessitate frequent updating of his reference patterns. Finally, the design of a speaker-adaptive word-recognition algorithm is so dependent on the uniqueness of each talker that very little insight is gained in the actual problem of recognizing speech. On the other hand, it is hoped that the development of a speaker-independent scheme will contribute to an understanding of the acoustic attributes of speech that reliably distinguish the various sounds. Without such an understanding, it would be difficult to duplicate the human capacity of recognizing the speech of a wide variety of speakers.

This paper discusses a speaker-independent digit-recognition system that was implemented on the computer facility of the acoustics research department at Bell Laboratories. Section II discusses the basic speech parameters that are measured and shows how the digits can be classified from these features in a speaker-independent manner. This section includes a discussion of the various signal-processing techniques that are heavily relied on throughout the classification process. Section III discusses the digit-classification scheme. The classification procedure is a tree-like decision algorithm for which backwards tracing is allowed when one of the parallel-decision algorithms indicates a high probability of error. Section IV gives the results of a formal evaluation of the recognition system. Finally, the paper concludes with a discussion of the strong and weak points of the system and suggestions for how it can be improved.

II. FRAMEWORK OF THE RECOGNITION SYSTEM

Figure 1 is a block diagram of the overall digit recognition system that was implemented. Following endpoint alignment in which the
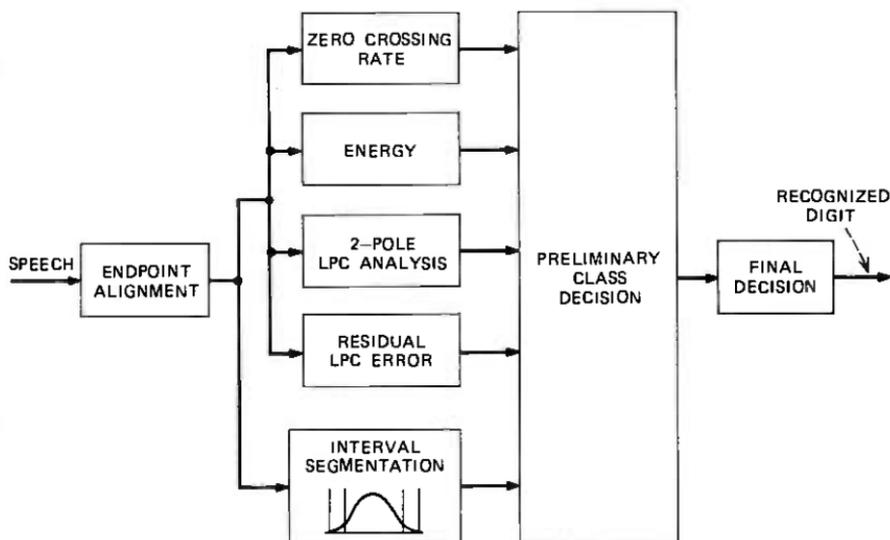
Fig. 1—Block diagram of the overall digit recognition system.

interval containing the word to be recognized is carefully determined, the speech is analyzed every 10 ms to obtain zero-crossing rate, energy, two-pole model linear-predictive-coding (LPC) coefficients, and the residual LPC estimation error. To aid in making preliminary classification decisions, the speech interval is segmented into three well-defined regions. All the speech information is fed in parallel into a preliminary decision-making algorithm that chooses one of several possible digit classes for the input utterance—e.g., one class contains the digits 1 and 9. A final decision is then made based on the presence or absence of certain key features in the input speech.

In this section, we show how the various digits can be characterized in terms of certain acoustic features. Then we discuss some key signal-processing functions that are heavily relied on in the decision algorithms and that contribute strongly to making the system speaker-independent.

2.1 Characterization of the digits

The elemental speech units (phonemes) that comprise English words can be classified into two broad categories, vowels and consonants. The vowels can be further classified into front (/i/, /I/, /e/, /ɛ/, and /ae/), middle (/ɝ/, /ʌ/), and back vowels (/u/, /U/, /ɔ/, and /o/). It is also convenient to subdivide the consonants into the categories noise-like (fricatives, plosives) and vowel-like (nasals, glides). Table I gives a list of the sequence of phoneme categories for each of the ten

Table I — Sound classes characteristic of the digits (from Ref. 6)

Digit	Sequence of Sound Classes
0	VNLC → FV → VLC → BV
1	VLC → MV → VLC
2	UVNLC → FV → BV
3	UVNLC → VLC → FV
4	UVNLC → BV → MV
5	UVNLC → MV → FV → VNLC
6	UVNLC → FV → UVNLC
7	UVNLC → FV → VNLC → FV → VLC
8	FV → UVNLC
9	VLC → MV → FV → VLC

VNLC = Voiced, noise-like consonant.
UVNLC = Unvoiced, noise-like consonant.
VLC = Vowel-like consonant.
FV = Front vowel.
MV = Middle vowel.
BV = Back vowel.

digits, 0 through 9.[6] Our approach toward speaker-independent recognition of the digits is to use a set of robust measurements to classify the phonemes into the six broad categories listed in Table I. By robust measurements, we mean acoustic parameters that give a general indication of the gross nature of each phoneme without being too dependent on the speaker's voice characteristics. Through a combination of parallel processing and self-normalization, the phoneme categories are determined and the spoken digit is recognized. We now discuss the criteria for the selection of the robust measurements that are used, the technique of self-normalization of measurements, and finally the method of parallel-processing of the data to give a speaker independent classification of the digits.

2.2 Robust measurements for digit recognition

The requirements for a recognition parameter to be selected as being a robust measurement are:

(i) The parameter can be simply and unambiguously measured.

(ii) The parameter can be used to grossly characterize a large proportion of speech sounds.

(iii) The parameter can be conveniently interpreted in a speaker-independent manner.

Based on the above requirements, the zero-crossing-rate and spectral-energy parameters are excellent candidates for robust measurements. These parameters can be used to effectively characterize the general acoustic properties of the sound categories listed in Table I. For example, noise-like sounds have a relatively high zero-crossing rate,

relatively low energy, and a relatively high concentration of high-frequency energy. Thus, the noise-like sounds of any speaker can be characterized quite accurately based on these measurements. The term "relatively," in the above classification of noise-like sounds, can be conveniently interpreted for a given speaker by a simple self-normalization technique discussed later in this paper

To measure the distribution of spectral energy, a two-pole LPC analysis has been suggested by Makhoul and Wolf[7] as an excellent means of representing the gross features of the spectrum. Figure 2 (from Makhoul and Wolf[7]) shows the results of applying a two-pole model to a variety of speech sounds. This figure is a comparison of the spectra of several speech sounds obtained directly from FFT spectrum measurements compared with the spectra of the best two-pole LPC fit to the spectrum. For a two-pole LPC analysis, there is either one complex-conjugate pole or two real poles. In Fig. 2a, the spectra for the sound /sh/ as in the word "short" are plotted. For this example, the two-pole LPC analysis gives a complex conjugate pole at about 3000
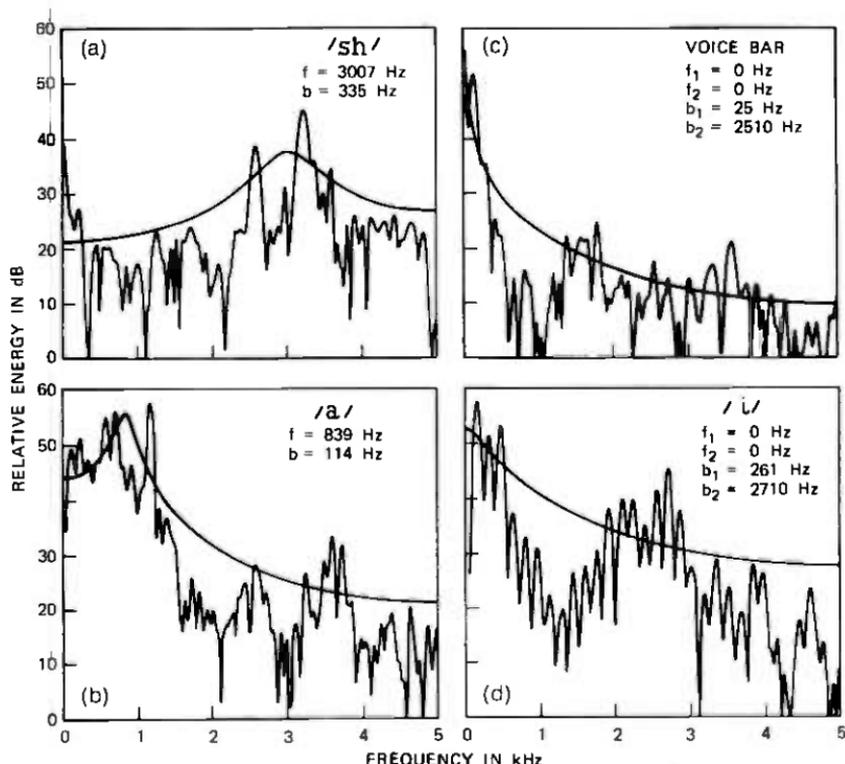
Fig. 2—Comparison of FFT spectra and two-pole LPC spectra for several speech sounds.

Hz—i.e., the region of maximum energy concentration in the spectrum. In Fig. 2b, similar results are shown for the vowel /a/ where the major concentration of energy in the spectrum is around 800 Hz. In the examples of Figs. 2c and 2d (a voice bar and the vowel /i/), the major concentration of spectral energy is around 0 Hz; thus, the two-pole LPC analysis gives two real poles in the right-half z-plane. From Fig. 2, it can be seen that the computed pole frequency gives a good indication of the location of the dominant portion of the spectral energy of the sound and can thus be effectively used to characterize sounds with relatively high-frequency or low-frequency concentrations of energy. For example, noise-like sounds are characterized by a relatively high-frequency spectral concentration of energy, while nasals and vowels generally have a much lower frequency for the energy concentration.

Figure 3 (also from Makhoul and Wolf[7]) illustrates the dynamic behavior of the computed pole frequency of the two-pole model and the corresponding spectrogram of the utterance, "Has anyone measured nickel concentrations · · · ." Examination of Figs. 2 and 3 shows that, for vowel-like sounds, the computed pole frequency is invariably situated somewhere between the first and second formants. In general, when F_1 and F_2 are not too far apart and have comparable amplitudes, the pole frequency falls almost midway between the two resonances. Since F_1 does not usually have as much dynamic movement as F_2, the computed pole frequency tends to follow the motion of the second formant. Since the motion of F_2 is quite important in the characteriza-
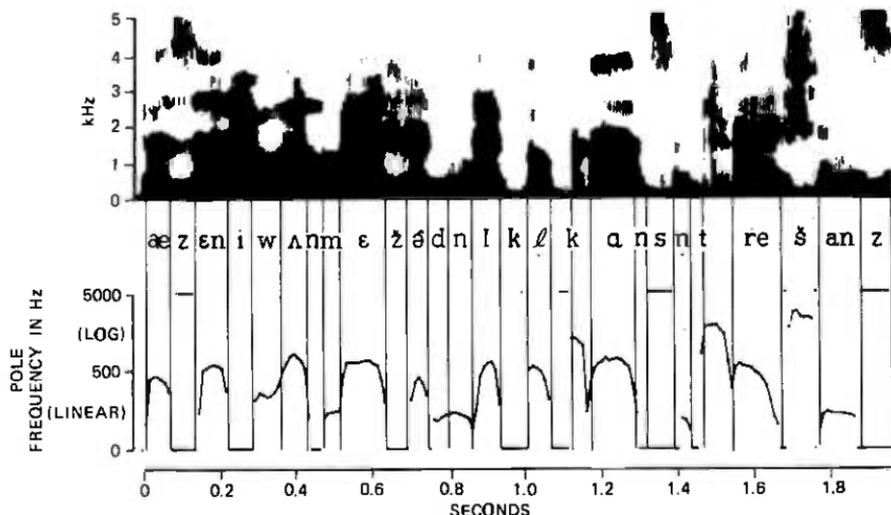
Fig. 3—Spectrogram and computed pole-frequency of two-pole LPC model for the utterance, "Has anyone measured nickel concentrations · · · ."

tion of the digits, the ability of the two-pole LPC analysis to track this motion has been used in the classification phase of the digit-recognition system.

It should be noted, however, that when either F_1 and F_2 are sufficiently far apart or the amplitude of F_1 is significantly greater than the amplitude of F_2, the pole frequency will either follow F_1 or result in two positive real poles. Figure 3 shows that, during the /i/ in the word "anyone," the pole frequency begins to dip sharply as the separation between F_1 and F_2 grows greater and finally results in positive real axis poles as the separation reaches some critical threshold. For nasal sounds, the energy is so highly concentrated near the first resonance that the two-pole model usually results in positive real axis poles, as seen in Fig. 3.

In addition to using the computed pole frequency as a measure of the location of dominant spectral energy and a characterization of the dynamic movement of F_2, the normalized error of the two-pole model contains important information about the spread of spectral energy. The normalized or residual error is defined as

$$V = 1 - a_1r_1 - a_2r_2,$$

where a_1 and a_2 are the two-pole LPC coefficients and r_1 and r_2 are the normalized autocorrelation coefficients. It can be shown that the more concentrated the energy spectrum, the lower the normalized error.[8] For speech sounds, the relative magnitude of the normalized error generally increases from sonorants to vowels and then to fricatives. Within the three vowel types, the back vowels have the lowest relative normalized error and the front vowels have the highest. By observing the relative changes in the pole frequency and normalized error, important information about the structure of the voiced region of the word can be obtained. An example of the usefulness of the pole parameter in specifying the speech sounds comprising the digits is given in the next section.

In summary, a reasonable set of robust measurements that have been implemented for this digit recognition algorithm is as follows:

(i) *Zero crossing rate*, which is defined as the number of zero crossings in a fixed frame length (on the order of 10 ms).

(ii) *Energy*, which is defined as the sum of the squared values of the speech waveform in a given frame.

(ii) *Normalized error* obtained from a two-pole LPC analysis of a given speech frame.

(iv) *Pole frequency* (or frequencies) obtained from a two-pole LPC analysis of a given speech frame.

2.3 Self-normalization of parameters

Almost all classification algorithms use some set of threshold levels in the decision process. Using a fixed set of thresholds leads to a large number of problems for speech recognition in that many of the thresholds are speaker- or time-dependent. To eliminate this difficulty, the technique of self-normalization of parameters was used in which many of the most significant thresholds were obtained from measurements made directly on the speech sample being recognized. Thus, for example, in the case of setting thresholds on zero-crossing rate to determine whether a sound is noise-like or nasal, a statistical description of the zero-crossing rate (zcr) was made for the entire utterance. The statistical description consisted of measuring the mean of the zcr and its standard deviation over the region of strong energy (i.e., the region where the energy exceeded 10 percent of the maximum energy of the utterance). Based on zcr measurements, one criterion for classifying a segment as noise-like was if its zcr exceeded a level one standard deviation *above* the mean during the segment. Figure 4 shows the zcr measurements for the word "seven." Indicated in this figure are the average zcr and a range of one standard deviation around this average. During the initial /s/, the zcr is significantly above the threshold, as anticipated.
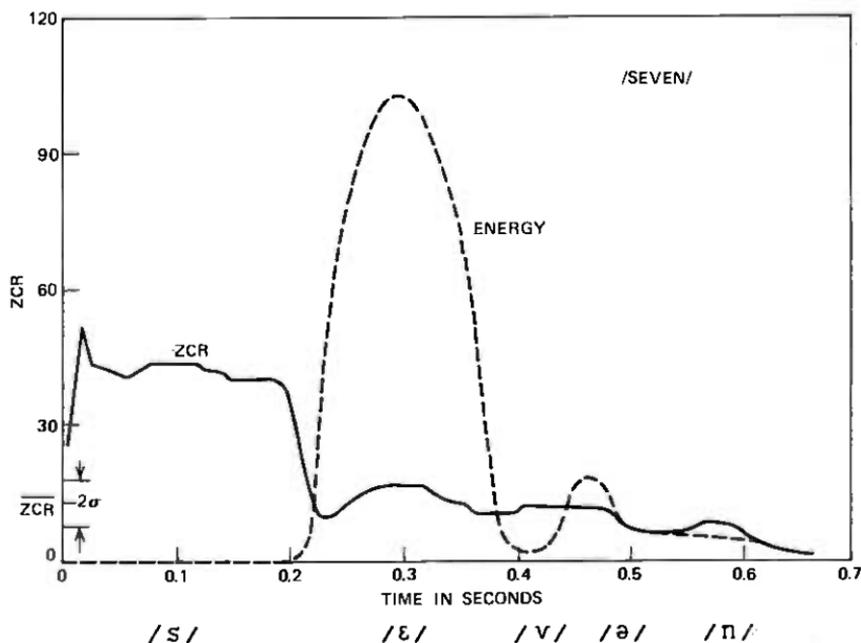
Fig. 4—Energy and zcr for one example of the word "seven."

Fig. 5—Energy and zcʀ for one example of the word "nine."

In much the same manner, a self-normalized zcʀ threshold can be set for classifying a segment as a nasal-like sound. In such cases, the zcʀ generally falls below a level one standard deviation *below* the average zcʀ for the utterance. As an example, Fig. 5 shows the measured zcʀ for the word "nine" and gives the spread of zcʀ around the average value. For the initial and final nasals, the zcʀ is much lower than the average and thus is a good indication of the nasals.

The idea of determining thresholds based on measurements made during the course of the utterance being recognized can be used for any or all measurements described in the preceding section. For example, Fig. 6 shows the two-pole model normalized error and the pole frequency for the word "nine." The nasal sections are clearly characterized by low normalized error and a zero-Hertz pole frequency. In contrast, Fig. 7 shows the same measurements for the word "six." Again, the noise-like sections are clearly depicted by the relatively high values of normalized error, pole frequency, and zcʀ.

The transitional nature of the normalized error and pole frequency can be used to classify the vowels into types high, middle, and back. Figure 8 shows the normalized error and pole frequency throughout the word "two." After the frication region, which is marked by high normalized error and low energy, the normalized error uniformly
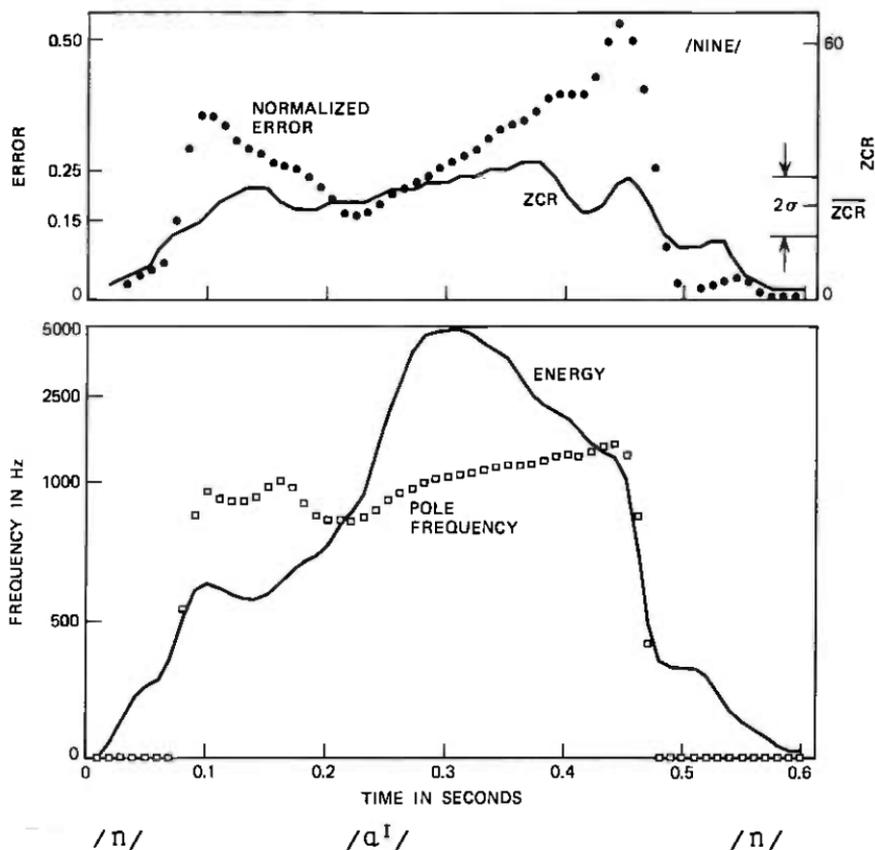
Fig. 6—Complete set of measurements for one example of the word "nine."

decreases. The decrease in the normalized error is due to the fact that
the vowel nature changes from front (because of the /t/) to back. Thus,
without specifying any absolute thresholds, the constituent structure
of the voiced section of the word can be obtained by noting the relative
changes in the normalized error. As described earlier, changes in the
pole frequency can also be used to indicate the constituent vocalic
structure. As seen in Fig. 8, the pole frequency is continually decreasing
throughout the voiced region in the word "two," thereby indicating
a continually decreasing second formant. As another example, Fig. 9
shows the parameters for the word "four." The gradually increasing
normalized error and pole frequency are indicative of a progression
from a back vowel to a middle vowel.

2.4 Parallel processing

Using the self-normalization technique, each robust measurement
can, by itself, classify a speech sound into one of the six broad categories
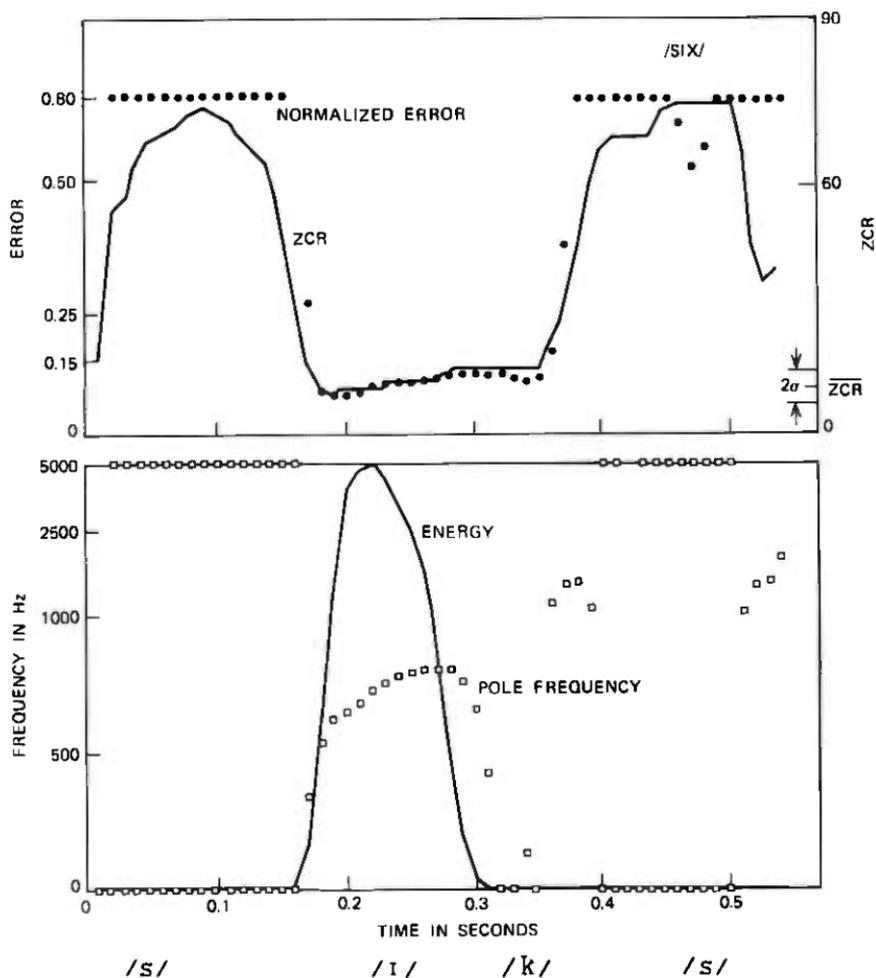
Fig. 7—Complete set of measurements for one example of the word "six."

of interest. Unfortunately, the classification will not be error-free; but if the results of all the measurements are "intelligently" pooled together, then the classification performance can be significantly enhanced. The operation of combining the measurements is termed parallel processing. Parallel-processing ideas have met with good success in other areas of speech processing.[9]

The idea of parallel processing as it is used here not only involves a suitable combination of the robust measurements but also the incorporation of certain structural constraints of the lexicon as additional input. For example, as seen in Fig. 9, the initial section of the word shows only a slight indication of the initial fricative /f/ (i.e., the high zero crossing and normalized error for the initial 10 to 20 ms of the
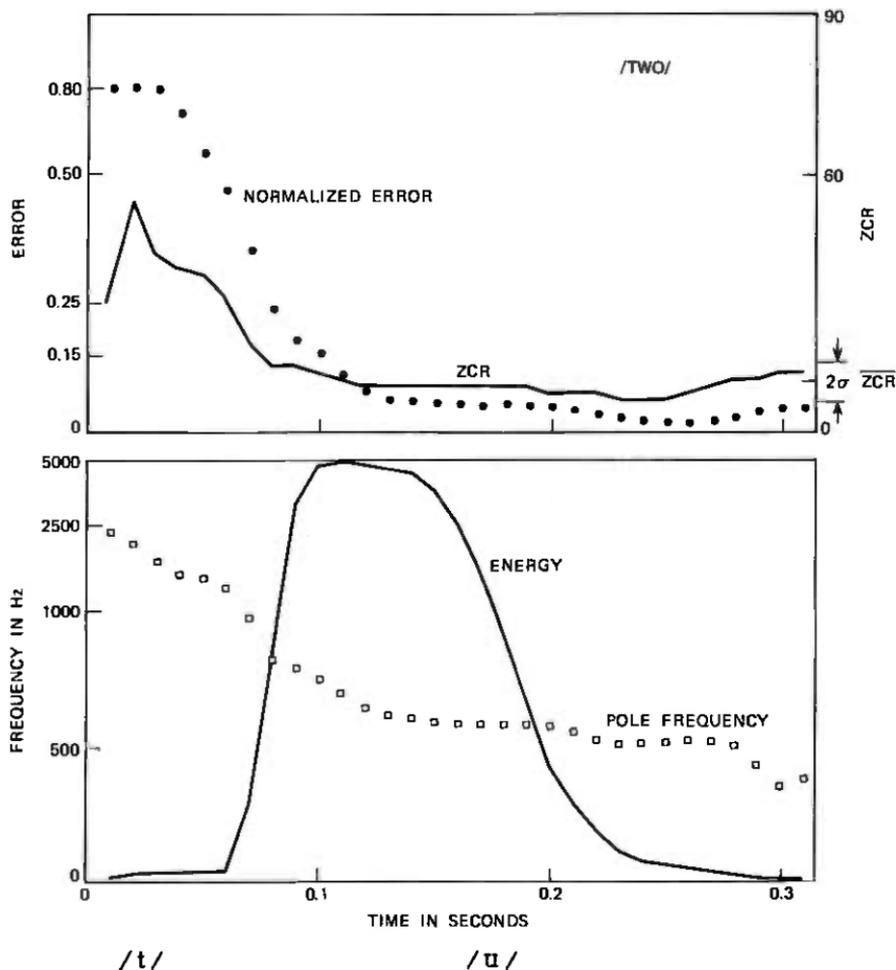
Fig. 8—Complete set of measurements for one example of the word "two."

word), and one might conclude that there is no frication. However, for the digits, it is known that only 1, 9, and 8 do not normally begin with frication. Since the ZCR, normalized error, and pole frequency are relatively too high for the digits 1 or 9, these digits can be safely omitted from consideration. In addition, a combination of the facts that the normalized error is low and increasing and that the pole frequency is increasing indicates that the voiced region in the word is more than likely composed of a back-type voiced sound followed by a middle-type voiced sound. Since the voiced section of the word "eight" is a front vowel sound, the odds are quite high that the word is not "eight." Additional evidence that the word is probably not "eight" can be
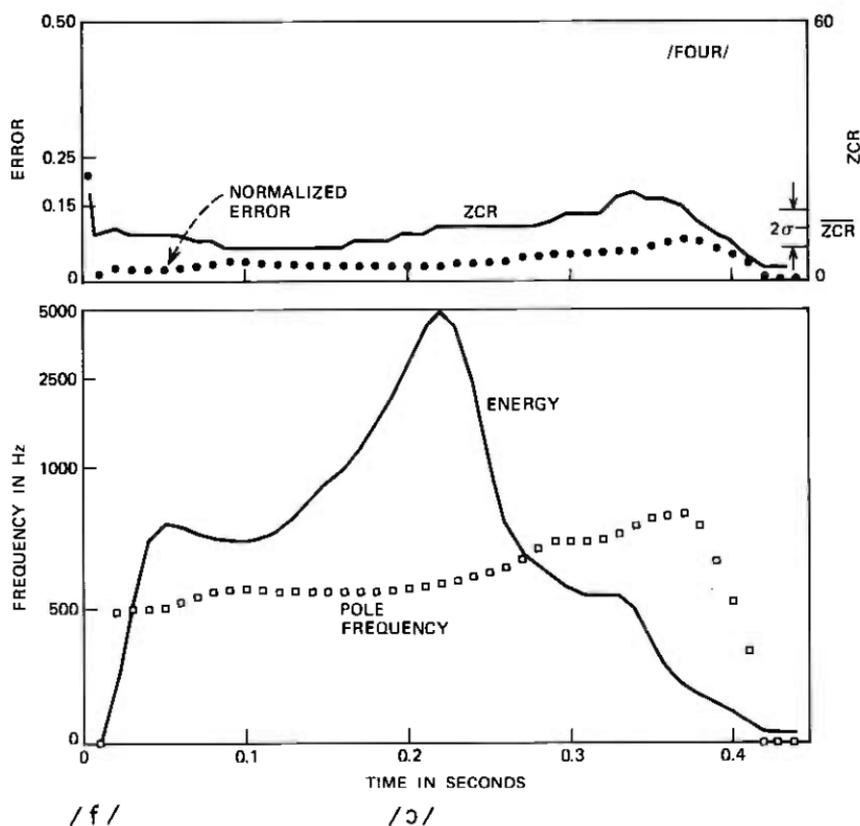
Fig. 9—Complete set of measurements for one example of the word "four."

obtained from the lack of a burst at the end of the spoken word. By pooling knowledge from the individual measurements with information about the structure of the words in the lexicon, the weak frication in the spoken "four" can be recognized. Thus, the major feature of parallel decisions is the ability to arrive at a correct decision even if one or more of the parallel inputs is in error. In the next section, we discuss the organization of the digit recognizer and the specific nature of the logic rules.

III. DIGIT RECOGNIZER

As seen from Fig. 1, the first step in the recognition scheme is the important problem of endpoint alignment (determining the location of the spoken word during the recording interval). The algorithm used in this scheme has been described by Rabiner and Sambur[10] and has been shown to give reasonably good results over a wide variety of speakers and background levels. However, the algorithm sometimes

has trouble finding the end of the word when the speaker sighs or puffs after reciting the word. To compensate for this problem, the decision algorithm does not place too much dependence on the end region of the word. The "end region" is defined as the region from the end of the word to the point at which the energy first exceeds 10 percent of the maximum energy. Equivalently, an "initial region" is defined from the beginning of the word to the point at which the energy first exceeds 10 percent of the maximum. The remaining section is termed the "middle region." The process of determining the three regions of the word is labeled "interval segmentation" in Fig. 1.

Throughout the duration of the detected word, the four robust parameters discussed in Section 2.2 are measured once every 10 ms (i.e., every 100 points for a 10-kHz sampling rate) and smoothed using a nonlinear smoothing algorithm proposed by Tukey.[11] In addition, the first two formant frequencies are computed using a 12-pole LPC analysis at three points during the middle region. These include the point of maximum energy, the beginning of the middle region, and the end of the middle region. Since formant frequencies are quite speaker-dependent, they were used in the decision process only as a *supporting* measurement to discriminate between sounds that were quite dissimilar when viewed in the $F_1 - F_2$ plane (/i/ and /a/ are examples). The supporting nature of the formant measurement is also necessitated by the fact that the extraction of formants is not a simple and unambiguous task, and too great a reliance on these parameters is fraught with danger.

Following the measurement phase, a preliminary class decision is made for the utterance. An expanded view of the preliminary decision box is given in Fig. 10. The decision algorithm is in the form of a tree structure that traces down the most probable branch to arrive at the decided digit. However, it should be noted that there are provisions in the algorithm for back-tracking if some measurement strongly suggests that an error has been made.

The first branch in the tree is to decide whether or not the initial portion is nasal-like. As we discussed previously, this decision is based upon the fact that nasal-like sounds have relatively low ZCR, low normalized error, and low pole frequency. If a nasal-like beginning is detected, the preliminary choice is between 1 and 9. As a further check on this preliminary 1, 9 decision, the ending region is checked for nasal-like characteristics. If there is no evidence of initial nasal-like sounds, the digits 1 and 9 are removed from further consideration. When the initial region is deemed nasal-like, a relatively simple decision can be used to decide between 1 or 9. The digit 9 can be distin-
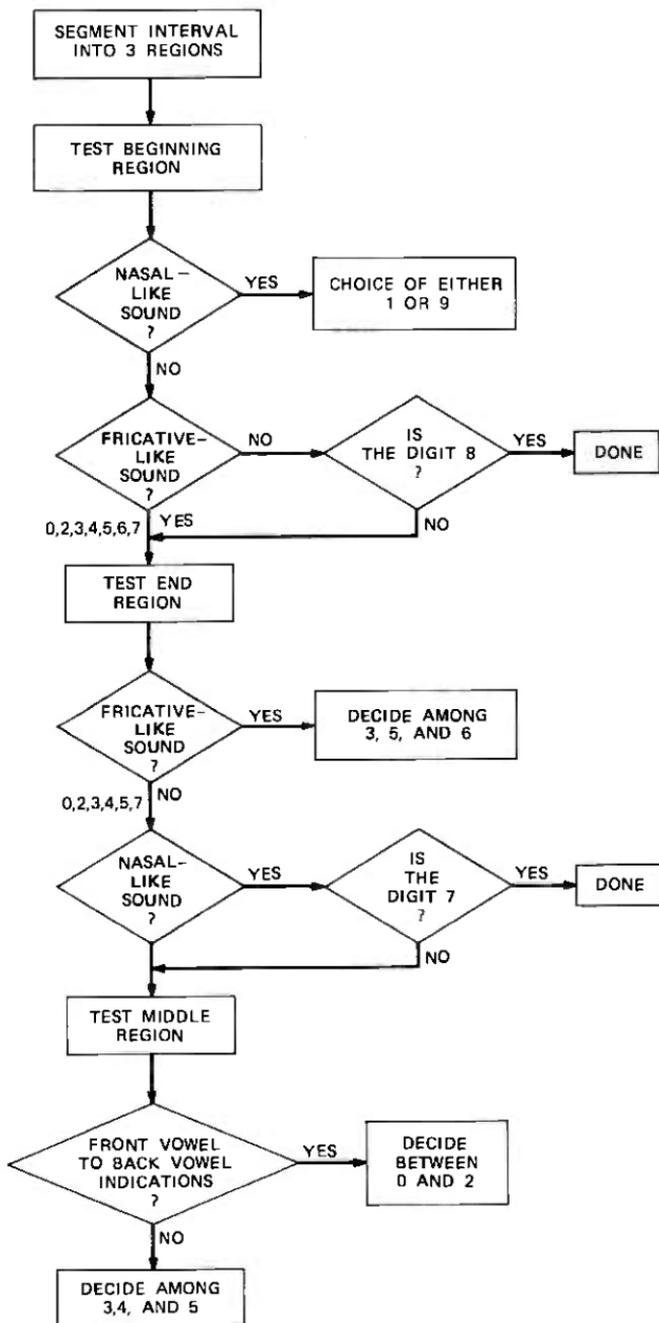
Fig. 10—Preliminary decision tree for digit classification algorithm.

guished by a sharp discontinuity in pole frequency (between the initial nasal and the vowel) and high normalized error during the transition from the nasal /n/ to the diphthong /aI/ (see, for example, Fig. 5). On the other hand, as shown in Fig. 11, there is no discontinuity at the end of the nasal-like section in 1.

Tracing along the decision tree, the next choice is to decide whether there is any definite indication of initial frication as shown by significantly high zcr, normalized error, and pole frequency. The positive detection of frication eliminates the choice of the word "eight." If there is no definite frication, the end region is checked for a burst and the middle region is checked for front vowel-like characteristics. The formant parameters are also used to check if the middle region is composed of only front vowels. A suitable combination of the results of these tests is used to reject or accept the word "eight" as the digit. It should be noted that the decision process is in the form of a hypothesis test. In other words, we assume that the spoken word is "eight" and check to see if the acoustic parameters are consistent with this hypothesis. In fact, the basic structure of the entire digit recognizer is to first hypothesize and then test the acoustic consequences of this hypothesis. The parallel processing aspect of the decision assigns the appropriate weight to a particular test. For example, the detection of a burst at this point in the decision tree is an almost 100-percent indication that the word is "eight." However, the lack of a burst does not necessarily preclude the possibility of "eight," and this result should be weighed accordingly.

Assuming that we reject the spoken word as the digit 8, the remaining possibilities are 0, 2, 3, 4, 5, 6, and 7. The end region is then checked for fricative-like behavior. If frication is indicated, a hypothesis test on the digit 6 is made. The middle region is checked for front vowel characteristics, and the only timing measurement in the entire digit recognition program is performed. This measurement compares the relative duration of the initial frication plus ending frication to the length of the middle region. The frication duration is defined from the beginning (or ending) of the word until the point at which the zcr remains within one standard deviation of the average for three time frames. This definition can be modified when any abrupt discontinuities in normalized error or two-pole frequency indicate a more probable location for frication. In addition, the extent of frication is not allowed to go beyond the 10-percent maximum energy points that form the boundaries of the middle region. For the digit 6, the timing ratio should be less than one and the middle region should be less than 250 ms. A combination of the results of the hypothesis test are used to verify that the spoken word was "six."
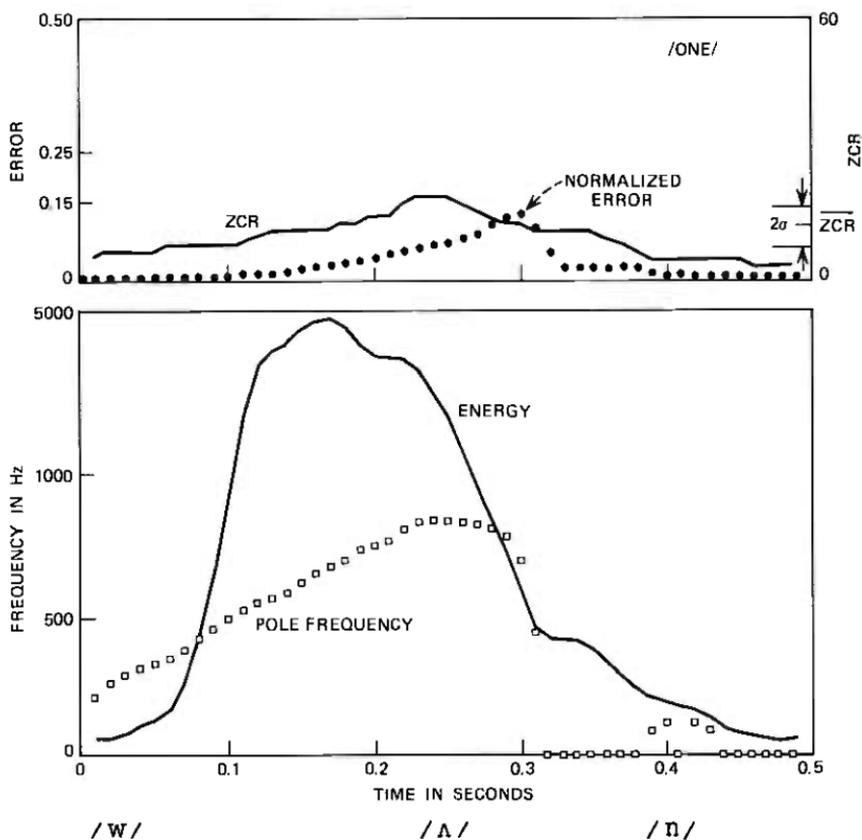
Fig. 11—Complete set of measurements for one example of the word "one."

If no frication is indicated during the ending region, a test for nasal-like behavior is made. If this test is positive, then a hypothesis test of the digit 7 is made. From Fig. 12, we can see that zcr, two-pole frequency, and energy dip sharply during the consonant /v/. The hypothesis test consists of verifying these dips and checking the vowel characteristics on either side of the dip. Again, the combined output of the test determines whether to reject or accept 7.

If the digits 6 and 7 are eliminated from consideration, the middle region is analyzed to ascertain its structure. The preliminary analysis is achieved by noting the relative change in normalized error. If the normalized error increases, then the structure is characterized as an initial back vowel to a middle or front vowel. Thus, for increasing normalized error, the digits 3, 4, and 5 are considered the most likely. The relative change in the three-second formant measurements are used as supporting evidence to confirm the structure. For the digits
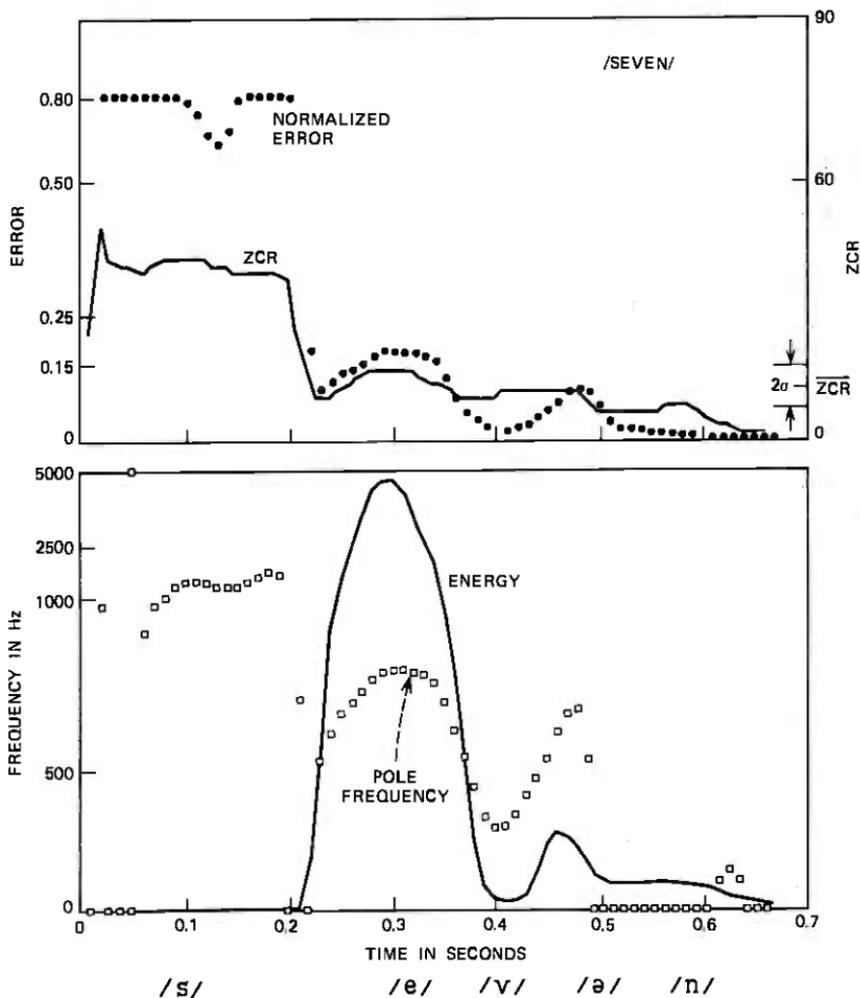
Fig. 12—Complete set of measurements for one example of the word "seven."

3, 4, and 5, F_2 should be increasing. The final decision among the possibilities 3, 4, and 5 is easily achieved on the basis of the robust parameters and formant measurements.

If the normalized error decreases during the middle region, the digits 0 and 2 are then the most probable choices. A decreasing F_2 helps support these choices. To decide between 0 and 2, a dip detector program is used to discover the presence of the sonorant /r/ as depicted by a slight dip in pole frequency, normalized error, and energy. Figure 13 shows the typical dip behavior for these parameters during the spoken 0. The presence (or absence) of a dip is checked with other measurements to verify the final decision.
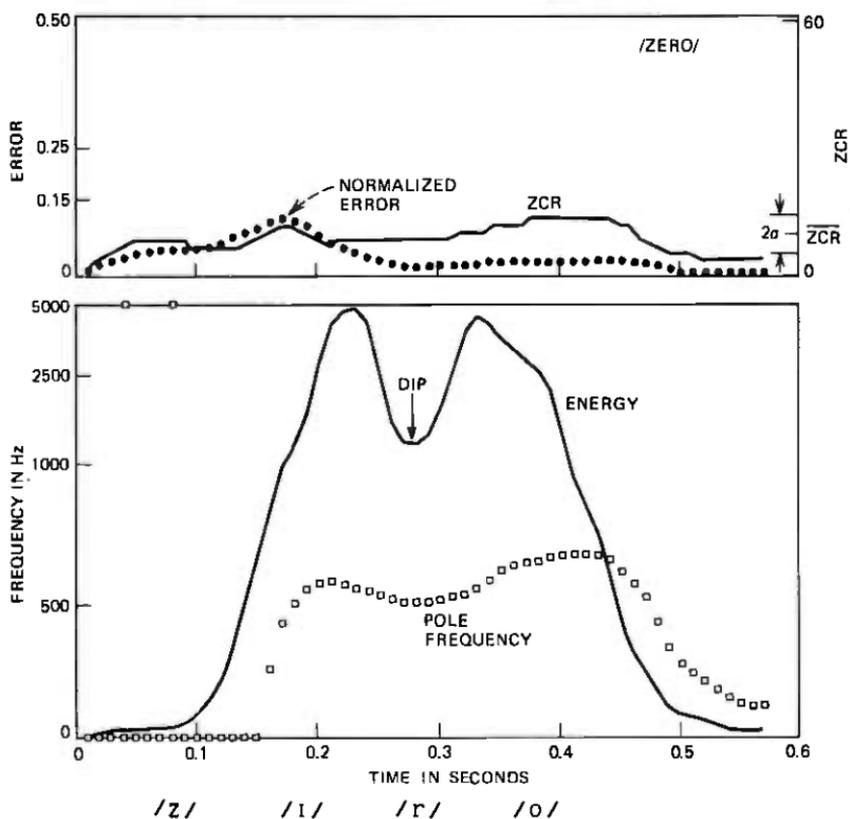
Fig. 13—Complete set of measurements for one example of the word "zero."

IV. EXPERIMENTAL RESULTS

The experimental test of the digit recognizer was conducted in two parts. The first part consisted of 10 speakers (five women and five men) who each made 10 complete recordings of the 10 digits. The recording sessions were spaced over a five-week period to include the effects of time variation in the testing. The recordings were made in a quiet room with a high-quality microphone. The decision algorithm was not designed for the characteristics of each particular speaker, so as to give a true test of the speaker-independent nature of the scheme. The results of this experiment are shown in Table II. The average error rate is 2.7 percent.

A confusion matrix for each of the 100 tests of each digit is presented in Table III. The confusion matrix indicates that all occurrences of initial frication were correctly detected by the decision algorithm. In only 6 out of 200 examples of the digits 1 and 9 was the initial nasal-

Table II — Error scores for first digit recognition experiment

	Correct	Wrong	Percent Correct
Women			
SK	97	3	97
KD	96	4	96
CMcG	96	4	96
BMcD	97	3	97
SP	97	3	97
Total			96.6
Men			
MS*	89	1	98.8
LR	100	0	100
RS	97	3	97
AR*	88	2	97.7
JH	97	3	97
Total			98.1
Sum	954	26	97.3

* Missed one recording session.

like consonant incorrectly determined. The confusion matrix also shows that most errors were made in the final detailed decision. More sophisticated processing would probably enhance the final decision and thereby make the system performance compatible with adaptive schemes.

Table III — Confusion matrix for first digit recognition experiment

		Word Spoken									
		0	1	2	3	4	5	6	7	8	9
	0	93	0	3	0	0	1	0	0	0	1
	1	0	96	0	0	0	0	0	0	0	0
	2	2	0	90	1	0	0	0	1	0	1
	3	2	0	2	97	0	0	1	0	0	2
Word Recognized	4	0	1	2	0	97	1	0	0	0	0
	5	0	0	0	0	0	96	0	0	0	1
	6	0	0	0	0	0	0	97	0	0	0
	7	1	0	0	0	1	0	0	97	0	0
	8	0	0	1	0	0	0	0	0	98	0
	9	0	1	0	0	0	0	0	0	0	93
Individual Errors		5	2	8	1	1	2	1	1	0	5

Total—26 errors out of 980 utterances

Table IV — Distribution of errors for the second digit recognition experiment

	Number of Digits Incorrect			Total Percent Correct
	0	1	2	
Female (30)	11	15	4	92.3
Male (25)	17	8	0	96.8
Total (55)	28	23	4	94.4

To ensure the validity of these experimental results, another more challenging test was conducted. In this experiment, 55 speakers (30 women and 25 men) were selected at random and asked to give one rendition of the 10 digits. Instead of using a high-quality microphone in a quiet environment, the input speech was taken from a close-talking microphone alongside a chattering Teletypewriter. The decision was performed on-line, and the speaker was only instructed when to say each digit. The results for this experiment show an average error rate of 5.6 percent. In addition, no speaker tested did worse than 2 out of 10 wrong. The distribution of errors for this experiment is shown in Table IV. The distribution matrix indicates the generally good performance of the system.

It should be noted that there was no effort in our experimentation to select speakers with good diction. The speakers represent dialects from most of the regions in the U.S. In informal on-line demonstrations of the system, many non-American speakers (French, Japanese, Indian, German) tried having their English-pronounced digits recognized. The informal results were in good agreement with the other experiments. In addition, an informal attempt to "beat" the system by holding one's nose or using falsetto also proved generally unsuccessful.

V. DISCUSSION

The digit-recognition system that has been described in this paper can be considered as a first pass in the direction of speaker-independent speech recognition. Our approach has been to describe a variety of speech sounds in terms of a set of robust measurements. We then devised a tree-structured decision algorithm that used these measurements to characterize the acoustic features of the presented word. The sequence of branches in the tree was designed to resolve the most obvious sounds and then proceed to the more difficult decisions. Thus, the relatively easy problem of distinguishing between noise-like sounds and nasal-like sounds was attacked first, and the determination of the vowel-like constituents was then determined. The output of the pre-

liminary decision tree was a small subset of the 10 possible digits that almost invariably included the spoken word. The major portion of the errors in the system were made in the box labeled "final decision" in Fig. 1.

The preliminary decision tree (Fig. 10) incorporated some original ideas about self-normalization that effectively eliminated the need for tuning the system to the characteristics of a given speaker. Such a decision tree can be extended to prune down a much larger lexicon and arrive at a small list of possible choices. Improvements in the method of selection within the list of possibilities could lead to speaker-independent systems that can truly compete with the performance of adaptive schemes. Such improvements could result by incorporating more sophisticated probabilistic methods into the framework of the "hypothesize-verify" technique proposed in this paper.

Our goal in the development of the digit-recognition system is to show that speaker-independent digit recognition is possible through an intelligent description of broad categories of speech sounds. This description uses what is known about the *necessary* characteristics of each category instead of blindly using pattern-matching algorithms to rigidly quantify the sounds. The later approach is doomed to failure for a large enough speaker population because it overlooks the fact that the information in the patterns contain as much personal information as linguistic information.

REFERENCES

1. Proceedings IEEE Symposium on Speech Recognition, Carnegie-Mellon University, April 1974.
2. Proceedings 1972 Conference on Speech Communication and Processing, April 1972.
3. A. Newell et al., *Speech Understanding Systems*, Springfield, Va.: National Technical Information Service, May 1971.
4. R. Reddy et al., "Working Papers in Speech Recognition," Carnegie-Mellon University, April 1972.
5. P. Vicens, "Aspects of Speech Recognition by Computer," Ph.D. thesis, Stanford University, April 1969.
6. T. B. Martin, "Acoustic Recognition of a Limited Vocabulary in Continuous Speech," Ph.D. thesis, University of Pennsylvania, 1970.
7. J. Makhoul and J. Wolf, "The Use of a Two-Pole Linear Prediction Model in Speech Recognition," Report 2537, Cambridge, Mass.: Bolt, Beranek and Newman, Inc., September 1973.
8. J. Makhoul and J. Wolf, "Linear Prediction and the Spectral Analysis of Speech," Report 2304, Cambridge, Mass.: Bolt, Beranek and Newman, Inc., August 1972.
9. B. Gold and L. R. Rabiner, "Parallel Processing Techniques for Estimating Pitch Periods of Speech in the Time Domain," J. Acoust. Soc. Am., *46*, No. 2 (August 1969), pp. 442–448.
10. L. R. Rabiner and M. R. Sambur, "An Algorithm for Determining the Endpoints of Isolated Utterances," to be published in B.S.T.J., *54*, No. 2 (February 1975).
11. J. W. Tukey, "Nonlinear (Nonsuperposable) Methods for Smoothing Data," 1974 EASCON Record, p. 673.

Space-Diversity Engineering

By A. VIGANTS

Vertically separated antennas are recommended to increase the transmission availability of line-of-sight microwave links by reducing the duration and frequency of multipath fading events. The emphasis is on application to the 4- and 6-GHz bands on links with negligible ground reflections. Necessary signal processing, links utilizing passive repeaters (reflectors), and overwater links are also treated. Some new experimental data are presented.

I. INTRODUCTION

Bell System microwave radio relay routes consist of links (paths, hops) that have an average length of about 26 miles. Transmission on most hops is along the line of sight, with antennas mounted on towers that are typically 250 feet high. In some cases, when line-of-sight transmission between towers is not practical, large reflectors on prominent points of terrain (passive repeaters) are used. Tower locations are selected to avoid ground reflections, but inevitable exceptions, such as transmission across a lake, do occur.

Multipath propagation during anomalous atmospheric conditions can give rise to destructive interferences at the receiving antenna; the resultant signal fluctuates (fades) and may be reduced to practically zero for seconds at a time. The corresponding interruptions to service, if permitted to occur, would be unacceptable. Interruptions to operation can be avoided by switching from an unserviceable radio channel to a protection channel operating at a different radio frequency, since multipath fading, being an interference phenomenon, is frequency selective. Use of such protection (frequency diversity) has been restricted to conserve the frequency spectrum.[1]

Space diversity is an alternative or additional form of protection from the effects of multipath fading.[2-8] Its effectiveness depends upon the fact that multipath propagation results in vertical structure of the electromagnetic fields at the receiving tower. Selection between two vertically separated antennas receiving at the same frequency is com-

parable in effect to a frequency-diversity system with a protection channel for every working channel (switched on a per-hop basis). Unfamiliarity with the principles of operation and unavailability of space-diversity equipment, as well as costs, have inhibited use of space diversity in the past. Currently, use of space diversity is increasing. One reason for this (apart from spectrum conservation) is that, in areas of high fading, frequency diversity alone cannot provide the desired transmission availability.

Space-diversity engineering, as presented here, encompasses estimation of fading, determination from transmission availability objectives of the need for protection, and calculation of the antenna separation needed to obtain the required transmission availability. The effects of signal processing on the improvement available from an antenna pair are discussed, with particular attention given to threshold switching. Space diversity through passive repeaters and on over-water paths is also discussed. Clearance requirements are summarized in the last section; current work on antennas placed so low as to lack clearance under normal propagation conditions is discussed in Appendix A.

The emphasis in this paper is on application in the 4- and 6-GHz frequency bands, although the expressions for the estimation of the amount of fading and the diversity improvement apply to other microwave frequencies. However, transmission in the 11-GHz band is also affected by rain, and this must be taken into account by additional reduction of the effects of multipath fading. This design aspect belongs more properly in a treatment of 11-GHz radio system design and is not discussed here.

II. DESCRIPTION OF FADING

The RF power received after transmission over a microwave radio hop is never absolutely constant; even at noon, when the atmosphere has "stabilized," there can be fractional dB excursions (scintillations recurring a few times per second), as well as slower excursions of a dB or two. In propagation experiments, the normal value of the received signal is determined from the peak in a signal-level histogram obtained over at least one-half hour at or near noon. This so-called free-space value of the received signal is determined repeatedly at least once a week to identify periods during which enhanced or depressed signals have resulted from the relatively steady atmospheric focusing or defocusing. These periods, therefore, are not "normal."

During fading, the received RF power can be practically zero for seconds at a time. The terminology to describe this is introduced in Fig. 1 through an example in which the free-space value is -30 dBm and a single, idealized fade decreases the received power temporarily
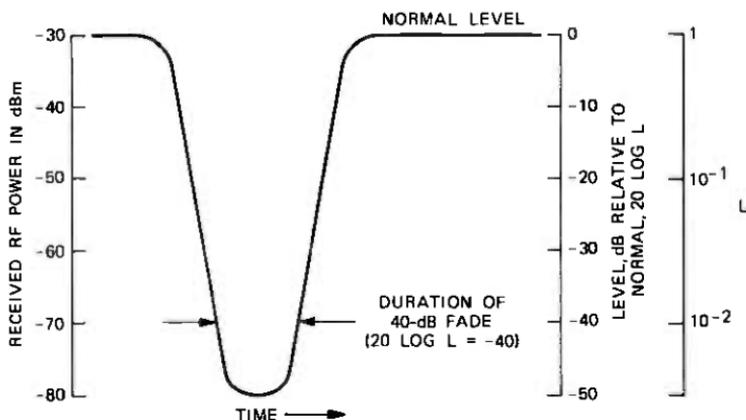
Fig. 1—Definitions of L and fade duration (-30 dBm assumed normal as an example).

to -80 dBm; levels in dB relative to normal are denoted by $20 \log L$. The time during which a signal is below a level is called the duration of fade of that level (the duration of a 40-dB fade is illustrated in Fig. 1). Average durations of fades are independent of microwave frequency[*] and are proportional to L; typical numerical values are given by (see also Fig. 2)[9-11]

$$\langle t \rangle = 410 \, L \text{ seconds}, \quad L < 0.1. \tag{1}$$

As an example, the average duration of a 40-dB fade ($L = 10^{-2}$) is 4.1 seconds, both at 4 and 6 GHz.
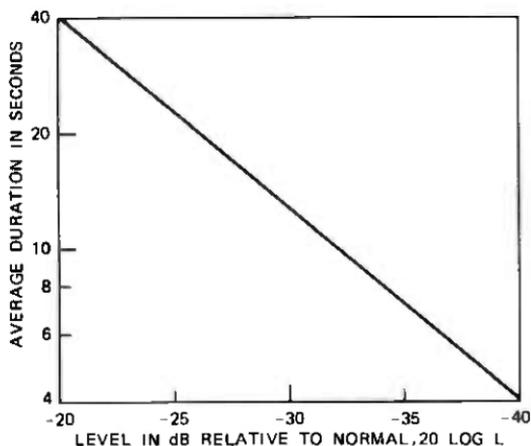
Fig. 2—Average fade durations ($\langle t \rangle = 410L$, $L < 0.1$).

---

[*] An underlying assumption throughout this work is that antenna sizes and path lengths are as encountered most often in practice: antenna diameters are between about 4 and 16 feet, and path lengths are between about 14 and 40 miles.

The sum of the durations of all fades of a particular depth is called "time below level." It is proportional to L^2, since the number of fades is proportional to L, and its numerical values are given by[9]

$$T = rT_0L^2, \quad L < 0.1, \tag{2}$$

where T_0 is the time period over which the summation of fade durations is made (a month, for example); the units of T are those of T_0 (seconds are normally used). The fade occurrence factor r for heavy fading months (see Fig. 3) is[9]

$$r = c(f/4)D^3 10^{-5}, \tag{3}$$

where

$$
\begin{aligned}
c &= 4 \text{ over water and Gulf coast,} \\
&= 1 \text{ average terrain and climate,} \\
&= \tfrac{1}{4} \text{ mountains and dry climate,} \\
f &= \text{frequency in GHz,}
\end{aligned}
\tag{4}
$$

and

D = path length in miles.
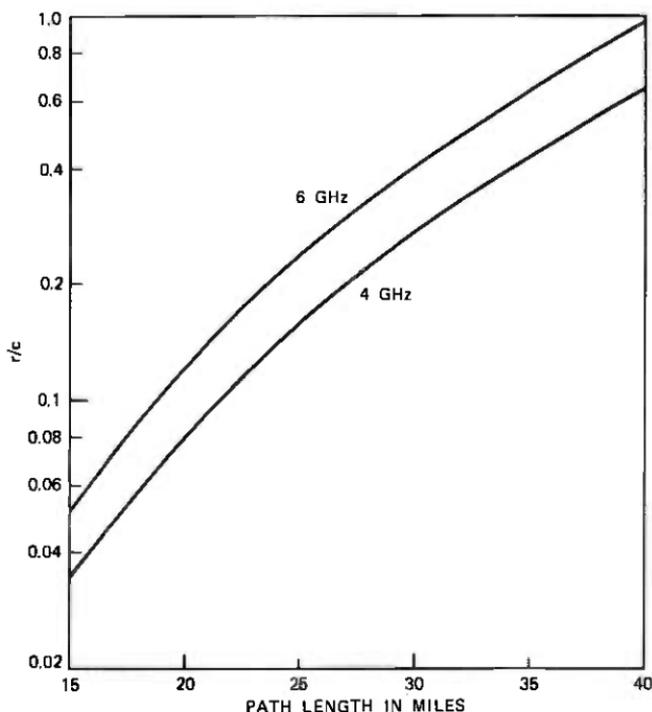
Fig. 3—Multipath occurrence factor.
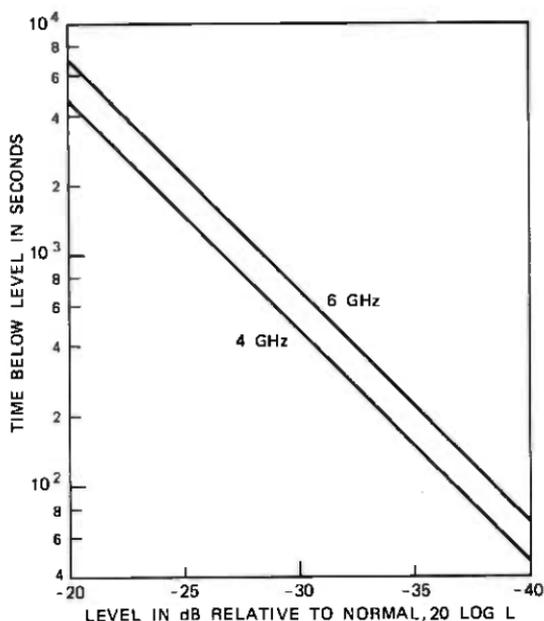
Fig. 4—Time below level in a heavy fading month (D = 26 miles, c = 1, T_0 = 31 days = 2.68×10^6 seconds).

As an example, values of T as a function of fade depth for a 26-mile path (average length) and average terrain and climate are shown in Fig. 4 for a heavy fading month. The lines have the decade of time per 10-dB slope typical of multipath fading, specified by the L^2 functional dependence. The values of T at -40 dB are 47 and 71 seconds at 4 and 6 GHz, respectively. Based on an average duration of 4.1 seconds, this corresponds to 11 fades of 40 dB at 4 GHz.

The coefficient c in (4) incorporates the effects of both terrain and humidity and is adequate for first estimates of expected fading in many cases. Differentiation between paths of identical climate but differing terrain can be done by introducing a terrain roughness parameter[12] quantifying common knowledge that paths over rough terrain fade less than paths over smooth terrain, presumably because stable atmospheric layering is less likely to occur over rough terrain. Terrain roughness is calculated from terrain heights above a reference level (sea level, for example) obtained from the path profile at one-mile intervals, with the ends of the path excluded. The standard deviation of the resulting set of numbers is the terrain roughness, denoted by w (see sample calculation in Appendix B). Applicable values of w range from 20 feet ("smooth") to 140 feet ("rough"); values of 20 and 140 should be used when calculated values of w are less than 20 or larger than 140.

Modified for roughness, the equations for c become:

$$c = 2 \ (w/50)^{-1.3}, \text{ coastal areas,}$$
$$= (w/50)^{-1.3}, \text{ average climate,}$$
$$= 0.5 \ (w/50)^{-1.3}, \text{ dry climate,} \qquad (5)$$

where a roughness of 50 feet has been defined as "normal."

III. PERFORMANCE OBJECTIVES

The time below level of the received signal, at a fade depth equal to the receiver fade margin, represents potential service-failure time. Protection is needed if this exceeds the value set by transmission-availability objectives.

Bell System short-haul objectives limit service failure time to 0.02 percent (two-way) annually on a 250-mile route due to all causes. One-half of this is allocated to causes associated with equipment, maintenance, and plant errors. There are obvious exceptions to this; for example, unavailability on a route where all hops are exceptionally short will be due mainly to equipment outages, and may be so allocated.

The allocation to fading, therefore, is 0.01 percent (two-way) annually. In the past, this allocation had been apportioned between multipath and obstruction fading ("earth-bulge" fading). However, the occurrence of intolerable obstruction fading can be decreased by increasing clearance (higher towers or shorter hops) or by increasing system gain margin; reliability records show that there has been a gradual decrease in the occurrence of obstruction fading over the years.[13] In the construction of new hops or in the upgrading of older ones in locations where obstruction fading is known to occur (or by related experience is expected to occur), increased clearances (or limits on path lengths) are assumed to be used. Consequently, no allocation to obstruction fading is made in space-diversity engineering; the entire 0.01 percent two-way annual fading allocation is applied to multipath fading. The one-way annual multipath fading allocation for 250 miles becomes 0.005 percent or approximately 26 minutes per year (1600 seconds per year). The corresponding allocation to a hop D miles long is $1600 \times D/250$ seconds per year (165 seconds per year for the average 26-mile hop).

Estimated annual time below level (for comparison with the objective) is obtained from the equation for T in the previous section, with T_0 describing the length of the fading season. As a geographic average, the value of T_0 in this estimate is equal to the number of seconds in three months; this assumes that all significant fading is contained in two heavy and two medium fading months, which is equivalent to
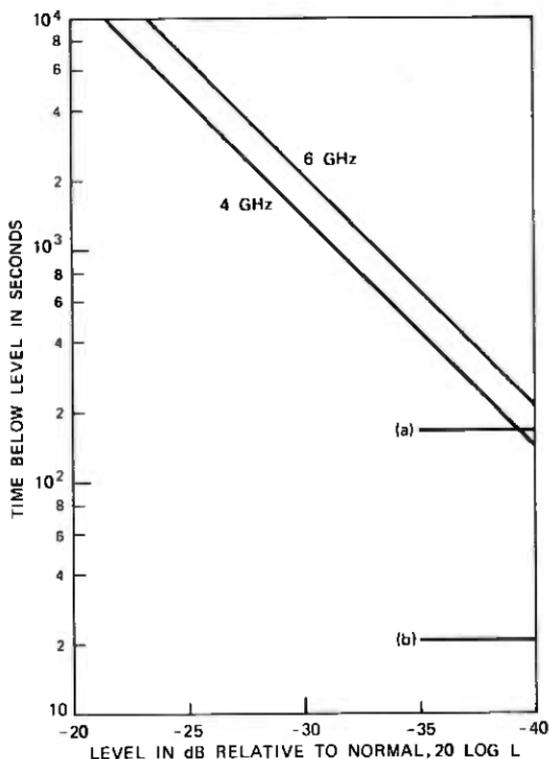
Fig. 5—Annual time below level (average case). Objectives: (a) short haul = 165 seconds/year, (b) long haul = 20 seconds/year.

three heavy fading months.[14] Estimation of the length of the fading season as a function of geographic location is discussed in Appendix C.

Continuation of the example of Fig. 4 provides the annual time-below-level curves in Fig. 5 (142 and 212 seconds below −40 dB at 4 and 6 GHz, respectively). For a fade margin of 35 dB, considered here for the purposes of discussion, the time below level is too large, compared to the 165-second objective, by factors of 2.7 and 4 at 4 and 6 GHz, respectively; protection against multipath fading (space or frequency diversity) is needed when the fade margin is 35 dB. At a fade margin of 40 dB, the 4-GHz channel in the example can get by without protection.

For long-haul radio the overall objective is also 0.02 percent, but the route length is 4000 miles. However, long-haul transmission models assume that half of the hops never experience significant fading.[14] With the addition of this assumption to those previously made in this section, the long-haul multipath allocation to a hop D miles long becomes $1600 \times D/2000$ seconds per year (20 seconds per year for the average 26-mile hop). In practice, long-haul radio has at least one protection

channel, which is required for equipment protection and maintenance and is used to provide frequency-diversity protection. Such use modifies the need for space-diversity protection (see Appendix D). Use of frequency diversity in short-haul radio can be expected to be infrequent because of regulatory restrictions.[1]

IV. THE SPACE-DIVERSITY EFFECT

During periods of multipath fading, deep fades of signals received on two vertically separated receiving antennas rarely overlap in time. The relatively few that do overlap give rise to simultaneous time below level (sum of durations of simultaneous fades, see Fig. 6), which is proportional to L^4 and can be expressed as[7,8,15]

$$T_s = T/I_0, \tag{6}$$

where T is the time below level of the signal received on the main antenna and I_0 is the available improvement, given numerically in practical units by the following (see also nomogram in Fig. 7):[7,8]

$$I_0 = 7 \times 10^{-5} v^2 s^2 f / DL^2, \quad s \leq 50, \tag{7}$$

where

    v = relative gain parameter (gain of secondary antenna relative to main antenna in dB is $20 \log v$),

    s = vertical separation of receiving antennas in feet, center-to-center,

    f = frequency in GHz,

    D = path length in miles,

and

    L = level parameter (level in dB relative to normal is $20 \log L$).
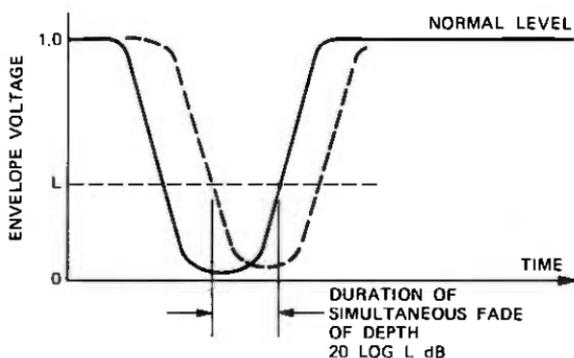
Fig. 6—Definition of simultaneous fade.
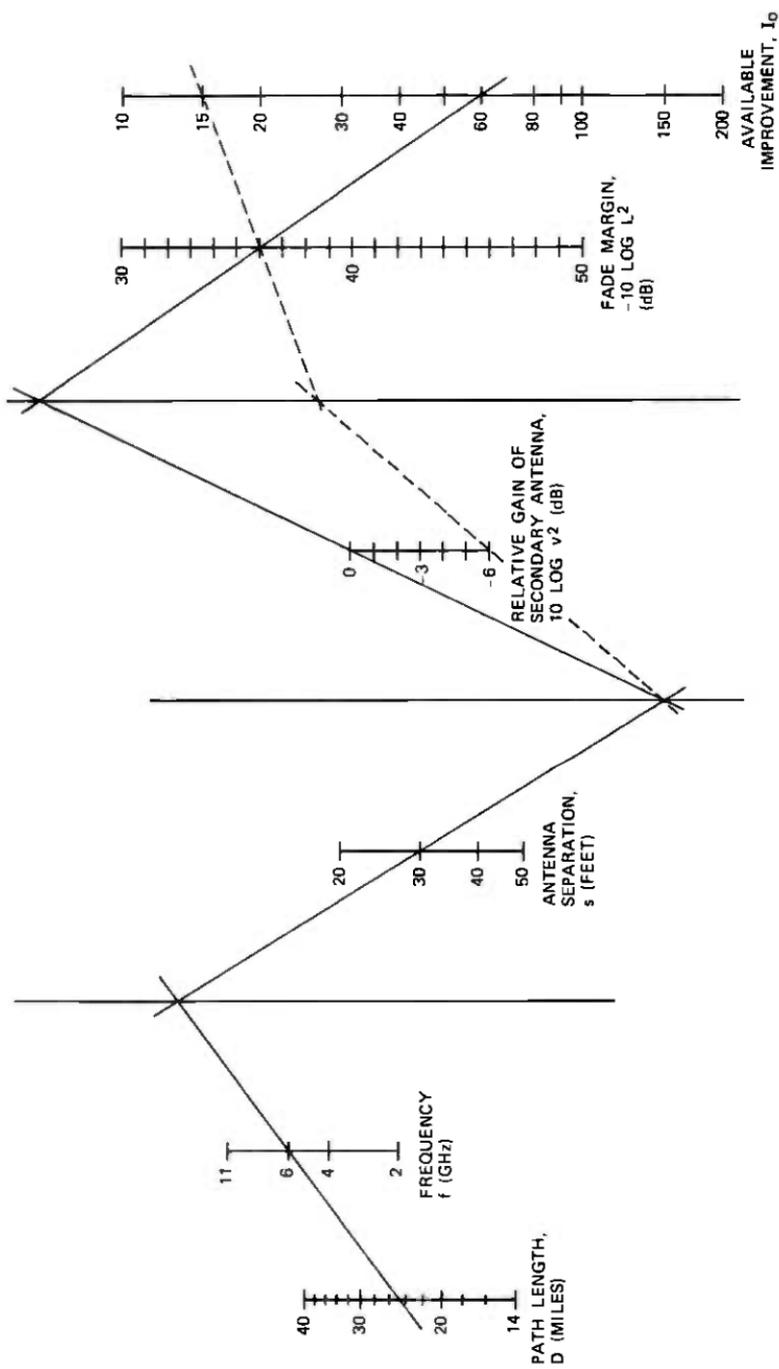
Fig. 7—Nomogram for available improvement I_0.

Equation (7) applies only for the ranges of variables indicated on the nomogram in Fig. 7. Extrapolation of the scales may lead to errors. For example, under some conditions the increase in improvement due to an increase in separation over 50 feet may be small. Separations for which the available improvement is less than 10 should not be used; if possible, separations of at least 30 feet should be used.

The example of Fig. 5 is continued in Fig. 8 to demonstrate drawing of a curve for T_s; the secondary antenna in the example has the same gain as the main antenna ($v^2 = 1$) and the vertical center-to-center separation of the antennas is 30 feet ($s = 30$). The values of I_0 at -40 dB are 145 and 97 for the 6- and 4-GHz channels, respectively. The values of T_s at -40 dB are $212/145 = 1.46$ seconds and $142/97 = 1.46$ seconds, identical for 6 and 4 GHz. This comes about because both T and I_0 are proportional to frequency. Having established one point (1.46 seconds at -40 dB) for T_s, we draw through it a line with a slope of a decade of time per 5 dB, as specified by the L^4 functional dependence. When I_0 becomes less than about 5, terms in addition to that proportional to L^4 are needed to describe T_s, which is why the line for T_s in Fig. 8 is not extended all the way to the left into the region (normally not of practical interest) where T_s approaches T.

The simultaneous time below level in Fig. 8 is smaller than the 165-seconds-per-year short-haul objective for fade margins larger than about 30 dB; it is smaller than the 20-seconds-per-year long-haul objective for fade margins larger than about 34 dB.

V. COMPARISON OF SPACE AND FREQUENCY DIVERSITY

Space diversity in its most common form provides a protection channel for every working channel (1×1 protection) on a per-hop basis. Frequency diversity usually provides one or two protection channels for m working channels ($1 \times m$ or $2 \times m$ protection) on the basis of switching sections that can contain as many as 10 hops in extreme cases. The most effective form of frequency diversity is, of course, 1×1 on a per-hop basis (now restricted in use at 4 and 6 GHz because of spectrum conservation); this can readily be compared to space diversity.

For equal performance the available improvements, I_0, are equated. A convenient form for I_0 is

$$I_0 = v^2 q L^{-2}, \tag{8}$$

where for space diversity (from the previous section)

$$q = 7 \times 10^{-5} s^2 f/D, \quad s \leqq 50 \tag{9}$$
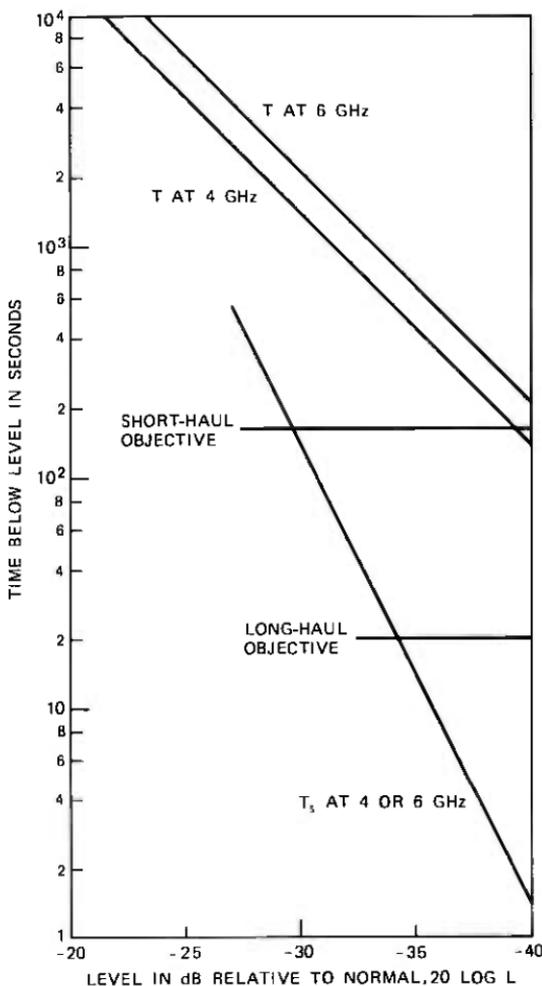
Fig. 8—Annual simultaneous time below level for 30-foot separation of antennas (continuation of average-case example from Fig. 5).

and for frequency diversity[16]

$$q = 50(\Delta f/f)/fD, \quad \Delta f < 0.5 \text{ GHz}, \qquad (10)$$

where f is the frequency in GHz (4 or 6), and Δf is the difference of radio channel center frequencies, also in GHz; D is the path length in miles. Values of separations in space and frequency providing equal performance (for antennas of equal size; $v^2 = 1$) are obtained by eliminating q from (9) and (10):

$$s = 106 \sqrt{\Delta f}, \text{ in the 4-GHz band},$$
$$= 57.5 \sqrt{\Delta f}, \text{ in the 6-GHz band}, \qquad (11)$$
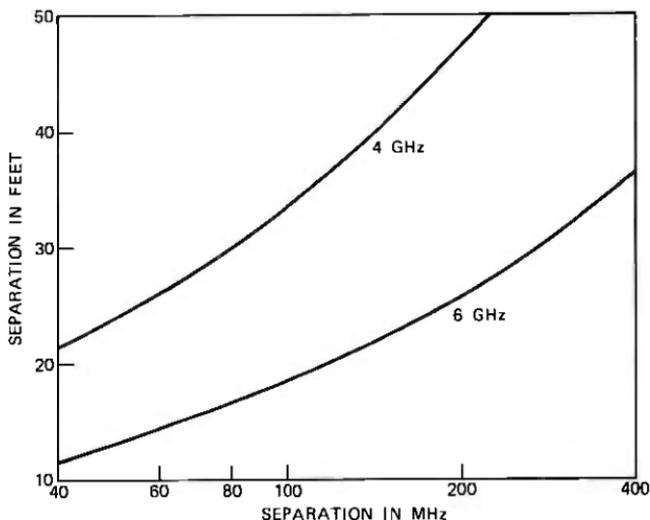
Fig. 9—Separations in space and frequency providing equal protection (one-for-one switching on a per-hop basis; antennas of equal size, $v^2 = 1$).

where s is in feet. A 30-foot separation is equivalent to a Δf of about 0.08 GHz in the 4-GHz band and about 0.27 GHz in the 6-GHz band (Fig. 9).

VI. SIGNAL PROCESSING

The signals received by the two antennas must be processed to obtain a diversity signal. Static addition at RF is not practical, since the duration of one-half cycle at, say, 4 GHz is $\frac{1}{8}$ nanosecond, and delay changes of this magnitude in the relative arrival times of the two signals arise easily because of daily angle-of-arrival variations. Without dynamic phase compensation, such changes would lead to signal cancellations during normal daytime operation, even when multipath fading is not present. The cost of dynamic phase compensation has so far been prohibitive.

Addition at baseband without dynamic phase compensation has been used (duration of a half cycle at 10 MHz is 50 nanoseconds). In some implementations, the weaker signal is dropped from the sum, when the relative signal strength ratio exceeds some 4 dB, to improve the signal-to-noise ratio of the diversity signal. Simultaneous time below level can be used to approximate the time below level of the diversity signal.

Switching, particularly at RF, permits an economically advantageous configuration of equipment. The simultaneous time-below-level curves (Fig. 8) describe the performance of an idealized comparator (switch),

where the diversity signal is, at every instant, the stronger of the two received signals. The switching activity, accompanied by undesirable step phase changes, is high. This can be reduced by introducing hysteresis, but at the cost of reduced performance. Suppose switching occurs only when the ratio of the received powers is larger than a number b^2 (10 log b^2 in dB). The improvement realized becomes

$$I = \eta I_0, \tag{12}$$

where an estimate of the efficiency factor η (see Appendix E) is

$$\eta = 2/(b^2 + b^{-2}). \tag{13}$$

As an example, when 10 log b^2 is 6 dB, η is 0.47, and an available improvement of 100 becomes a realized improvement of 47.

A different approach must be used when only one receiver per radio channel is available, which is the case in long-haul radio. The diversity implementation utilizes an RF waveguide switch activated by the AGC voltage when the receiver input falls below a threshold (one example of switching logic is shown in Table I). Switching of this sort has been referred to as threshold or "blind" switching, since at the instant of switching there is no assurance that a stronger signal will be available at the other antenna. Most of the time, a stronger signal *is* found, since few of the deep fades on the two receiving antennas overlap in time. The time below level T_t for threshold switching is calculated (see Appendix F) in Fig. 10 for a 4-GHz channel (in continuation of the example from Fig. 8); the threshold is at -35 dB relative to normal. The values of T_t and T_s are identical at the threshold. Above threshold, T_t approaches T rapidly, since the switch does not act for fades during which the minimum signal does not drop below -35 dB. The part of T_t of importance in space-diversity engineering is the straight-line section below threshold. This straight-line section is parallel to T, since it represents an average of the two single-antenna statistics by virtue of the continued cycling of the switch due to failure

Table I — Switching logic of a threshold switch

Switch Output Relative to Switch Threshold	Switch Connected to	Action
High	Main antenna	Stay on main antenna
High	Secondary antenna	Switch to main antenna after 15 minutes
Low	Main antenna	Switch to secondary antenna after 0.1 second
Low	Secondary antenna	Switch to main antenna after 0.1 second
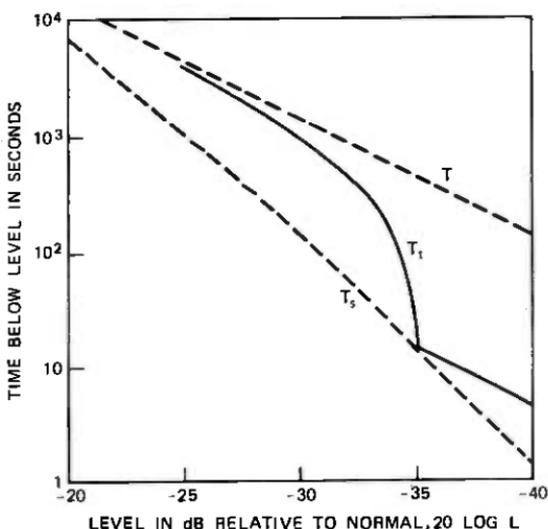
Fig. 10—Time below level in threshold switching—continuation of average-case example (4 GHz, threshold at −35 dB, 30-foot separation of antennas).

of the receiver to locate a signal stronger than the threshold value, which is −35 dB relative to normal in the example.

A suitable value of the threshold is about 2 dB above the fade margin. The −35 dB threshold in the example would be appropriate for a fade margin of 37 dB. The improvement ($I_t = T/T_t$) is constant below threshold (about 30 in the example) and equal to the value of I_0 at
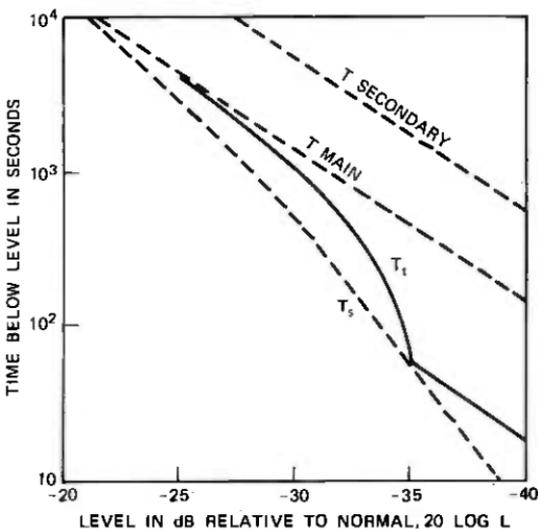
Fig. 11—Example of reduced performance due to smaller secondary antenna ($10 \log v^2 = -6$, other parameters as in Fig. 10).

the threshold. This is one reason for placing the threshold close to the fade margin, since a threshold at -30 dB, for example, would provide an I_t of only about 10. A second reason is to limit the occurrence of the phase steps due to the cycling of the switch to occasions when the signals border on the unusable (actually, the chance is small that the signals remain for prolonged periods in the 2-dB corridor between the threshold and the fade margin, and many of the phase steps will occur, therefore, when the signals are unusable). The cycling should occur at a rate of at least five times per second; at much slower rates, a good signal is not found fast enough and performance deteriorates. Corridors smaller than 2 dB should be avoided, since signal drifts caused by aging of equipment or noise accumulation from hop to hop in a switching section can frustrate design intentions and require frequency diversity action before space diversity.

Use of a smaller secondary antenna reduces performance, as illustrated in Fig. 11, where the only change from Fig. 10 is that a secondary antenna with 6 dB less gain ($v^2 = 0.25$) has been used. The improvement I_t has decreased to about 8 ($\approx 30/4$) from about 30 in the equal-size-antenna case.

VII. SPACE-DIVERSITY TRANSMISSION

Space diversity for microwave radio is usually envisioned as diversity reception; i.e., the system consists of a single antenna which transmits to two vertically separated receiving antennas. Since the path loss from antenna to antenna does not depend on the direction of transmission, an arrangement of two vertically separated transmitting antennas and a single receiving antenna can also be used.

Combined use of transmitting and receiving diversity can reduce tower work; on a long route only alternate towers would be affected. A single bad hop can be corrected by additional installation at one end only. On hops where additional installation at one end is impossible because of economics or zoning, use of both transmitting and receiving diversity at the other end provides full diversity protection in both directions.

A drawback to transmitting diversity is the vulnerability of the control signal. In diversity reception, the signals from both antennas are always available at the switching site, to be processed in any desired manner. In transmitting diversity, only one antenna can transmit at a given frequency at a given time (dynamic phase correction for control of the two transmitters does not appear feasible or desirable). The required control information must be fed back from the receiving to the transmitting end, since control must be based—because of the frequency selectivity of deep fades (see Appendix G)—on fading in

the controlled channel. The feedback link increases both the cost of control and the possibility for failure of control.

VIII. FIELD EXPERIENCE

The basic experimental data on simultaneous fading on vertically separated receiving antennas were obtained in Ohio and Texas in 1966. The variation of the improvement as the square of the vertical separation of the antennas was further verified experimentally in Georgia in 1968 (Appendix H). Field followup in terms of monitoring performance of in-service space diversity was carried out in 1972 in California and Florida. The monitoring was made possible by the development of Portable Propagation Recorders (PPR), the first installation of which was at Brawley, California, in May 1972.

Space diversity in the California case was installed by Pacific Telephone and Telegraph Co. to improve transmission availability of two hops (Salton-Brawley and Brawley-Glamis traversing Imperial Valley) on the Dallas-Los Angeles route. Extensive irrigation and high temperatures with little wind combine to create severe fading on the two hops in question. Measured and calculated values of I_0 are compared in Table II (Fig. 12 is an example of measured data). RF threshold switching was used, and the comparison was made at the threshold value (-35 dB relative to normal) where I_t and I_0 are equal.[17] The agreement of calculated and measured values is good, showing that the equation for I_0 is applicable under diverse climatic conditions.

Two hops monitored in Florida (Andytown South-Andytown North and Andytown North-Okeelanta in Northern Everglades) were equipped with secondary receiving antennas and RF threshold switches as part of a program to improve transmission availability of hops that fade heavily. The measured results and the predictions for Andytown

Table II — Comparison of calculated and measured values of I_0 (Salton-Brawley and Brawley-Glamis, California)*

Antenna Spacing (feet)	Frequency (GHz)	Path Length (miles)	Relative Gain of Secondary Antenna (v^2)	I_0 at 20 log L = -35	
				Calculated from Equation (7)	Measured
30	4	42.6	0.41	7.6	8.0[†]
15	4	42.6	0.41	1.9	1.5[‡]
15	4	37.1	0.41	2.2	2.5[§]

* Antenna configurations were dictated by circumstances; generally, configurations providing an improvement of at least 10 would be used.
  † Brawley to Glamis, 7/7–8/14, 1972.
  ‡ Glamis to Brawley, 5/11–7/3, 7/7–8/14, 1972.
  § Salton to Brawley, 8/17–8/29, 1972.
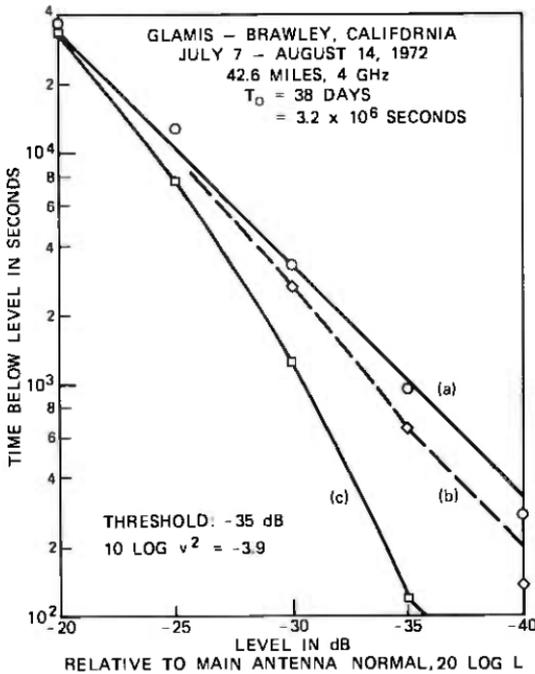
Fig. 12—Threshold switching, Glamis to Brawley, California. Reception at (a) Glamis, main antenna (unprotected), (b) Brawley, 15-foot separation, (c) Glamis, 30-foot separation.

South to Andytown North are summarized in Fig. 13.[18] At −47 dB (bottom level of the PPR for this case), the predicted time below level in the 148-day test period was insignificant (only a few seconds), and none was measured. Diversity performance thus conformed to expectations and the needed improvement in transmission availability was obtained.

IX. PATHS WITH PASSIVE REPEATERS

Use of passive repeaters (large reflectors of, say, 40 by 48 feet) on ridges or hilltops is sometimes dictated by terrain or by the undesirability of active repeaters in remote locations (power and maintenance-access problems). Tests on a Mountain Bell hop (Lusk-Wendover, Wyoming) in 1973 established that space diversity is operative through passive repeaters.[19] A design procedure for practical cases, based on the Wyoming results, is:

(i) Determine reflector size to obtain reasonable fade margin.
(ii) Estimate unprotected time below level (T) as the sum of the values of T for the individual legs (distances from reflector to ends of hop).
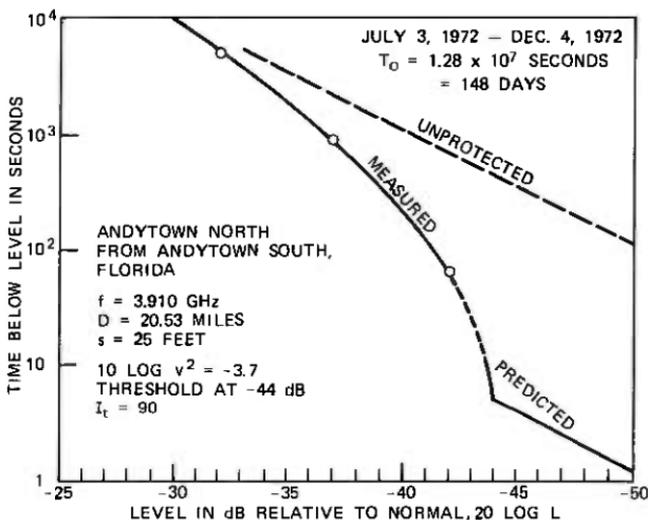
Fig. 13—Threshold switch operation at Andytown North, Florida.

   (iii) Determine improvement needed to meet transmission availability objectives.

   (iv) If improvement is needed, determine vertical spacing based on longest leg (use at least 20 or 30 feet); use this vertical spacing at both ends of hop.

   (v) If an antenna pair is in the near field of a reflector, ensure that projected reflector height is not less than the distance from the bottom edge of the bottom antenna to the top edge of the top antenna.

The experimental data obtained at Wendover and Lusk are summarized in Figs. 14 and 15 (the parameters of the experiment are summarized in Table III).[19] Time-below-level differences for individual antennas in a pair arise from gain differences in instrumentation chains; curves fitted to the points have the standard decade of time per 10-dB multipath slope in the deep-fade region. The curves fitted to the simultaneous fading points have the standard decade of time per 5-dB slope in the deep-fade region. Space-diversity reception at Lusk *is* effective (Fig. 15), which is a result that could not be predicted using present methods. Successful space-diversity operation at Wendover could be predicted beforehand by viewing the transmitting antenna at Lusk as a feed illuminating a large aperture (the double reflectors), since little fading would be generated on the short Lusk-to-reflectors leg.

The measured improvements at −35 dB, 50 at Wendover and 30 at Lusk, are larger than the calculated values of 40 and 16, respectively.
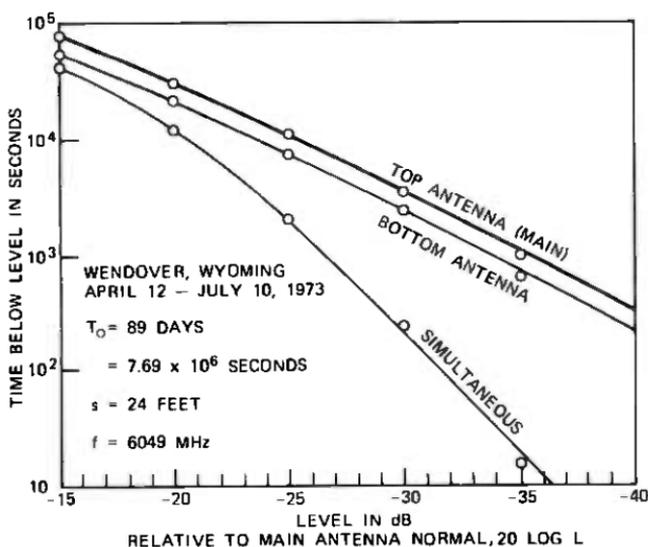
Fig. 14—Space diversity on path containing passive repeater; reception at Wendover, Wyoming.

The difference arises partly because use of the longest leg in the calculations is a simple approximation of a complex situation and partly because some scattering of measured points around the averages predicted by the equation for I_0 can be expected.
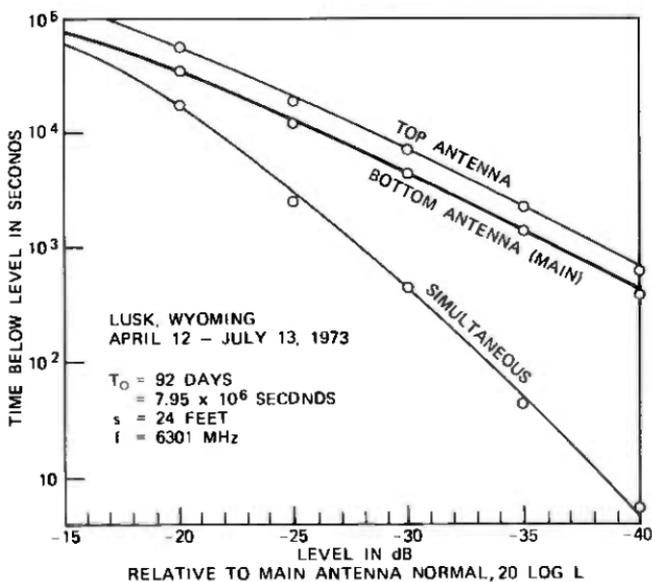
Fig. 15—Space diversity on path containing passive repeater; reception at Lusk, Wyoming.

Table III — Parameters of Lusk Radio test—Lusk to Wendover, Wyoming

Reflectors	Two 40- by 48-ft reflectors separated by 138.5 ft
Path lengths from reflectors: To Wendover To Lusk	31.5 miles 9.5 miles
Antenna separations at Wendover and Lusk	24 ft (nominal) vertical center to center
Frequencies: Wendover Lusk	6049 MHz received 6301 MHz received
Calculations: Values of v^2 arising from gain differences in instrumentation chains: Wendover Lusk Calculated improvement at -35 dB based on 31.5-mile path length Wendover ($v^2 = 1.6$) Lusk ($v^2 = 0.63$) Allocation to multipath fading from performance objectives for 41 miles $(31.5 + 9.5)$	$v^2 = 1.6$ $v^2 = 0.63$ $I_0 = 40$ $I_0 = 16$ $1600 \times 41/250 = 262$ seconds per year

The multipath objective for the Wendover-Lusk link is 262 seconds per year one way (see Table III). Even allowing the annual time below level to increase by a factor of two due to late summer and fall fading, the objective can be met comfortably with receiver fade margins in the 35- to 40-dB range.

X. OVER-WATER PATHS

Over-water paths, undesirable because of reflections, are sometimes unavoidable. When the over-water area is too large to permit shifting the reflection point off it via a high-low antenna combination, space diversity can be used to reduce reflection fading. A procedure for determining antenna spacing is discussed in terms of the geometry shown in Fig. 16, in which the reference plane coincides with the surface of the water during "flat-earth" propagation conditions (i.e., those obtaining when K, the ratio of equivalent to actual earth radius, is infinity). The "bulge" height h of the surface reflection point A is

$$h = d_1 d_2 / 1.5 K, \qquad (14)$$

where h is in feet when d_1 and d_2 are in miles. If the height h_2 of the receiving point were changed, then the relative phase of the rays arriving via the paths TR and TAR would also change. For a fixed
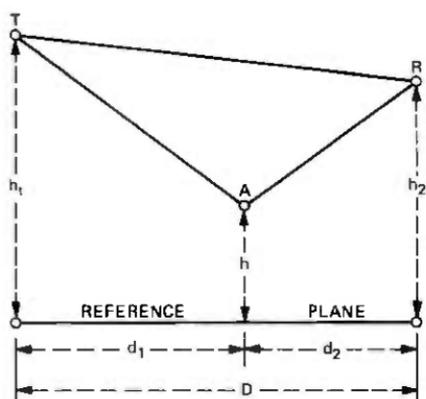
Fig. 16—Geometry for discussion of over-water paths.
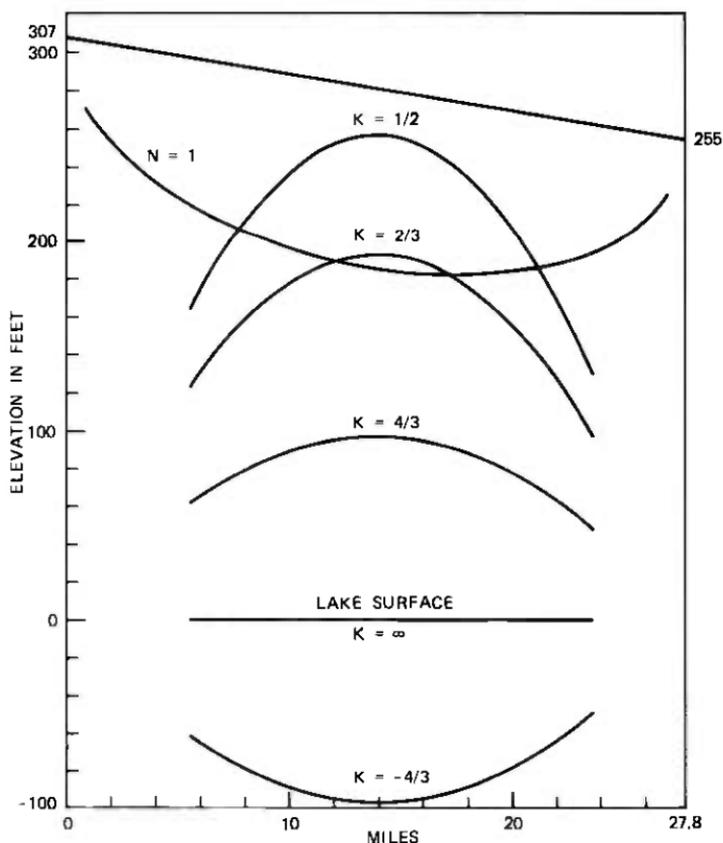
Fig. 17—Lake surface at various K values in over-water path example (N denotes boundary of Nth Fresnel zone at a frequency of 4 GHz).

transmitting antenna height h_t, the increment in h_2 for a change from an in-phase to an out-of-phase condition (half interference fringe spacing) is approximately (for large K), in feet,

$$\Delta h_2 \approx 1300 D / f(h_t - h), \qquad (15)$$

where h_t and h are in feet, D is the path length in miles, and f is the frequency in GHz.

The desired antenna separation for good diversity, therefore, would be equal to Δh_2, but this is possible for one value of K only, since Δh_2 increases as h increases (an increase of h results from a decrease of K). For protection over a large range of K, a suitable initial choice for the antenna separation is

$$s = 1300 D / f h_t, \qquad (16)$$

which corresponds to Δh_2 for $K = \infty$. The appropriateness of this
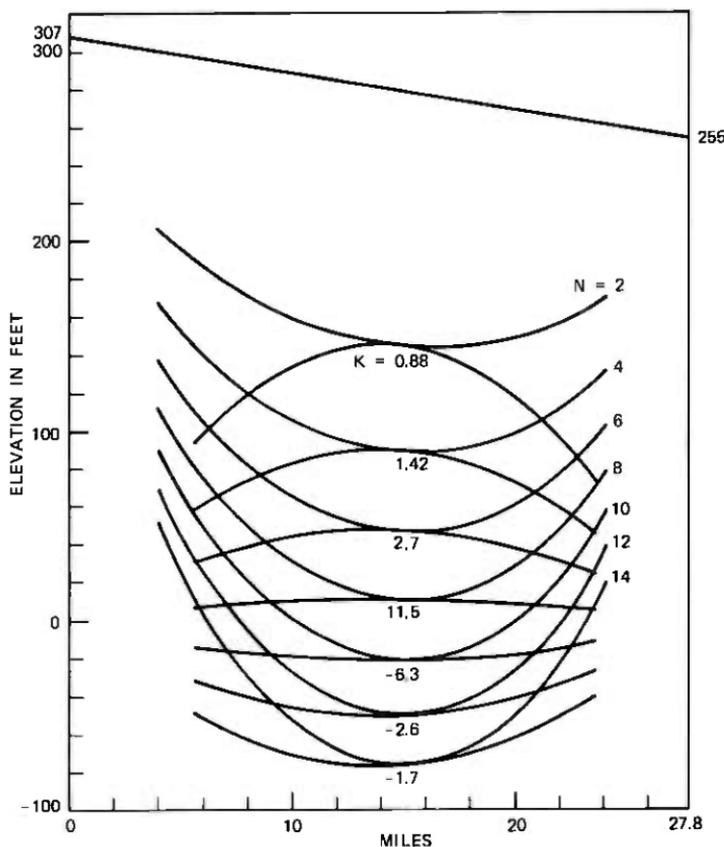
Fig. 18—Values of K at which lake surface is tangent to even Fresnel zone boundaries at a frequency of 4 GHz.

choice must be verified. In a recently examined situation (shown in Fig. 17) the path length D is 27.8 miles, the transmitting antenna height h_t is 307 feet, and the receiving antenna height h_2 is 255 feet; since the planned hop is on the Gulf coast, the antennas are placed high to avoid obstruction fading. Protection from fading created by reflections from the lake surface is desired for values of K ranging from $\frac{1}{2}$ to about $-\frac{4}{3}$ (Fig. 17 and the subsequent drawings were generated very easily on an xy-plotter driven by a programmable desk calculator). Values of K for which the lake surface is tangent to even Fresnel zone boundaries (determined by trial and error) are shown in Fig. 18; at these values of K, signal minima occur at the 255-foot height. For 4-GHz transmission, a first choice for diversity antenna separation would place the second receiving antenna at the 225-foot height $[(1300 \times 27.8)/(4 \times 307) \approx 30 ; 255 - 30 = 225]$. Simultaneous sig-
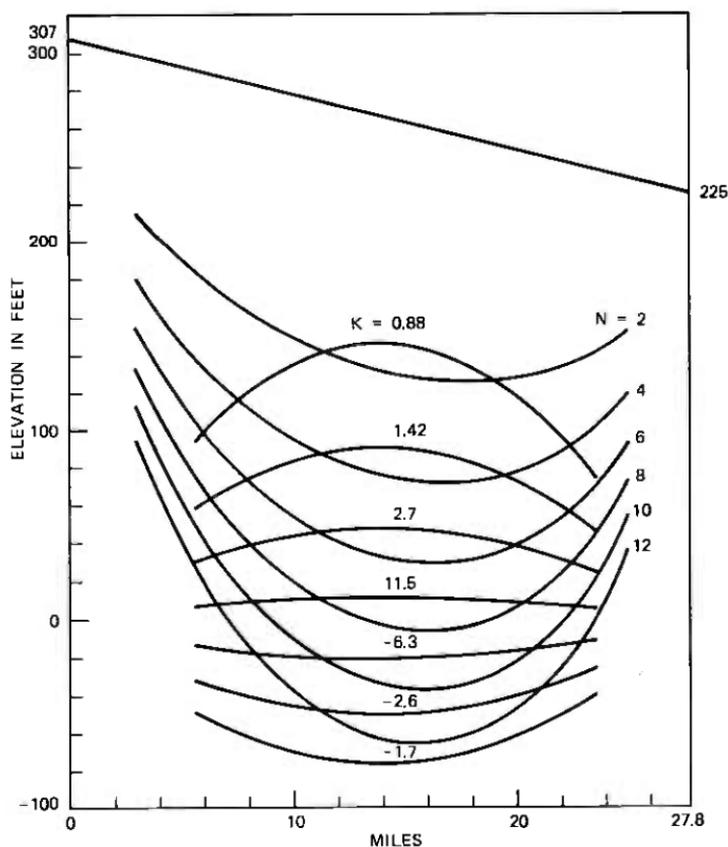
Fig. 19—Even Fresnel zone boundaries at 225-foot height for bottom receiving antenna at a frequency of 4 GHz.

nal minima will not occur on the two antennas because even Fresnel zone boundaries for the 225-foot height are not tangent to the lake surface at K values that correspond to signal minima on the top receiving antenna (see Fig. 19). The question of optimization now arises. It would be desirable to place the bottom antenna somewhat lower to move the boundary of F_2 down, further away from the lake surface at $K = 0.88$; this would bring the boundary of F_{12} closer to the lake surface at $K = -1.7$ and, therefore, is not desirable. If the design requirements were relaxed to cover values of K between $\frac{1}{2}$ and ∞ only, then the bottom antenna could be brought further down to make the separation closer to the value of Δh_2 at $K = \frac{4}{3}$.

A possible alternative location of the bottom antenna, foreground clearance permitting, is the 85-foot height (see Fig. 20). This has some
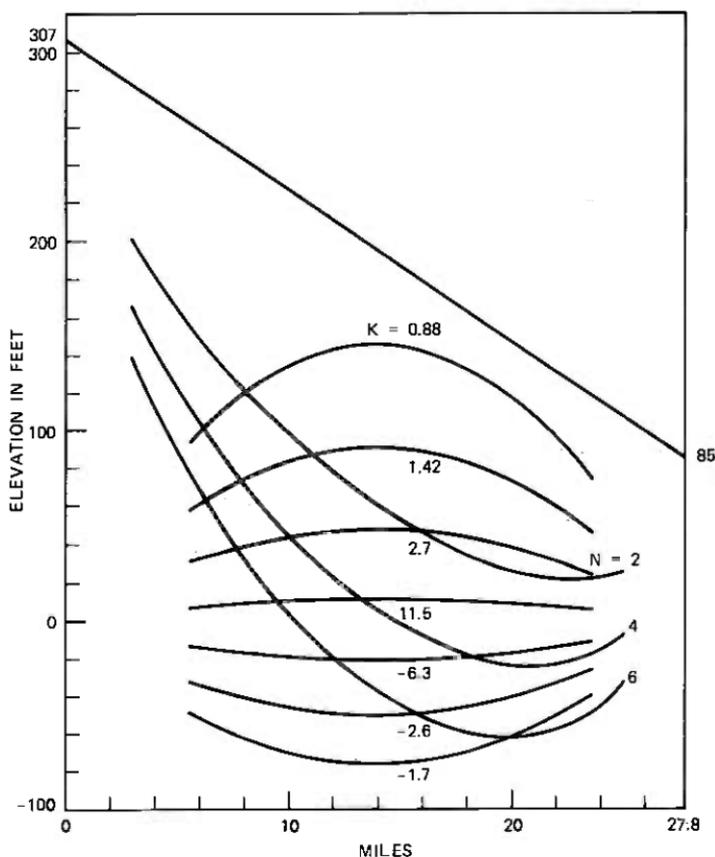
Fig. 20—Even Fresnel zone boundaries at 85-foot height for bottom receiving antenna at a frequency of 4 GHz.

of the properties of a high-low shot, and the even Fresnel zone boundaries are well placed.

A final observation on space-diversity design for protection from reflections is that the antenna spacings often are suitable, fortunately, for protection from atmospheric multipath fading.

XI. CLEARANCE RULES

Operating experience indicates that propagation conditions in various sections of the United States can be classified, in broad terms, as good, average, or difficult as shown on the map in Fig. 21. Desired clearances for transmission between main antennas are summarized in Table IV. In addition, hop lengths in coastal areas of the southern United States are sometimes restricted to 20 miles to reduce obstruction fading.

Clearance requirements for antennas added to provide space-diversity protection from multipath fading are less stringent: $0.6\ F_1$ at $K = \frac{4}{3}$, with a foreground clearance of 10 feet in the first 500 feet from the antenna. In many cases, this permits placement of the secondary antenna below the main antenna.

It appears (from recent results discussed in Appendix A) that it may be possible to relax the secondary-antenna clearance require-

GOOD
DIFFICULT
AVERAGE
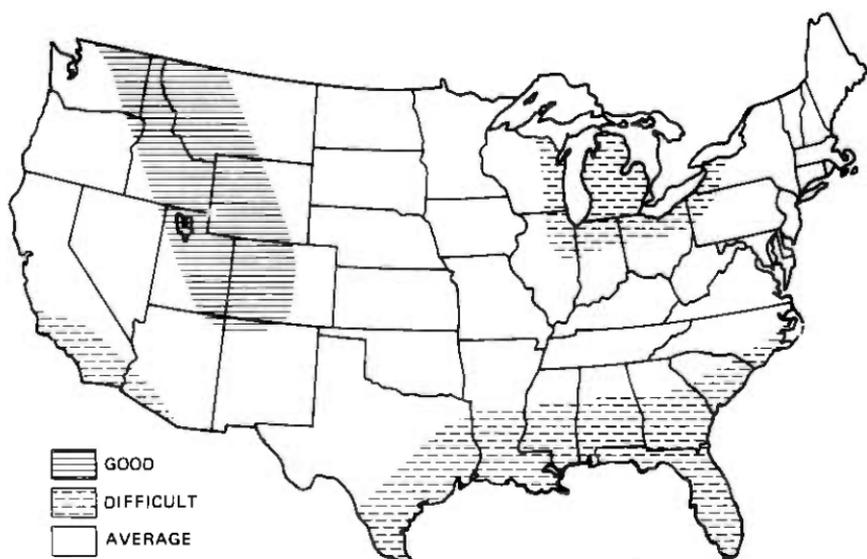
Fig. 21—Geographic occurrence of good, average, and difficult propagation conditions.

Table IV — Clearance rules for top antennas

Propagation Conditions	Clearance
Good	$0.6 \, F_1$ at $K = 1$ (but not less than grazing at $K = \frac{2}{3}$)
Average	The larger of $0.3 \, F_1$ at $K = \frac{2}{3}$ and F_1 at $K = \frac{4}{3}$
Difficult	Grazing at $K = \frac{1}{2}$

ments; however, results of further tests (now in progress) must be evaluated before general recommendations on this can be made.

XII. ACKNOWLEDGMENTS

APPENDIX A

Effects of Reduced Clearance

The purpose of the bottom antenna in a space-diversity pair is to provide protection from multipath fading; its clearance therefore need not be based on requirements for protection from obstruction fading. Placement of secondary antennas lower than currently permitted would reduce tower costs because of reduced wind loading.

An investigation of multipath fading as a function of antenna height was undertaken at Palmetto, Georgia in 1972 at 4.198 GHz on a 26.4-mile path from Atlanta, Georgia.[20] The path profile (Fig. 22; arrows denote trees) shows that the controlling obstruction is located some 6 miles from the receiving antennas at Palmetto. Lines of sight that graze the obstruction are shown in Fig. 22 for three kinds of atmospheric conditions. Under normal conditions the gradient of the index of refraction is $-39 \, N$ units per kilometer,* and the ratio K of the equivalent-to-actual earth radius is $\frac{4}{3}$. When the gradient becomes $-157 \, N$ units per kilometer, the so-called flat-earth condition exists ($K = \infty$) and clearance is substantial. For changes of the gradient in the opposite direction, clearance becomes reduced; in path design the value of interest is $K = \frac{2}{3}$, which corresponds to a positive gradient of $79 \, N$ units per kilometer.

The projections of the grazing lines on the Palmetto tower are given by the tops of the three bars on the right-hand side of Fig. 23, at 30,

---

* The radio refractive index of air, n, is frequently expressed as $n = 1 + N \times 10^6$, where N describes the refractive index in "N units."
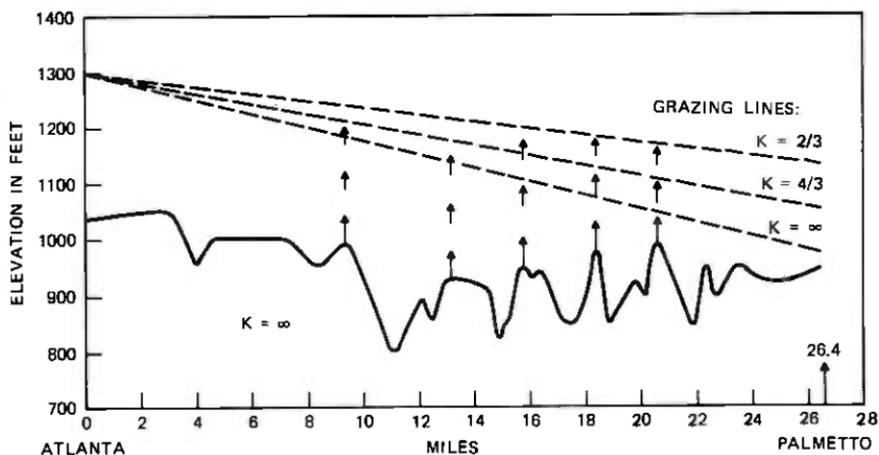
Fig. 22—Path profile, Atlanta to Palmetto, Georgia.

Fig. 23—Receiving antennas at Palmetto, Georgia in 1972.

SPACE-DIVERSITY ENGINEERING **129**
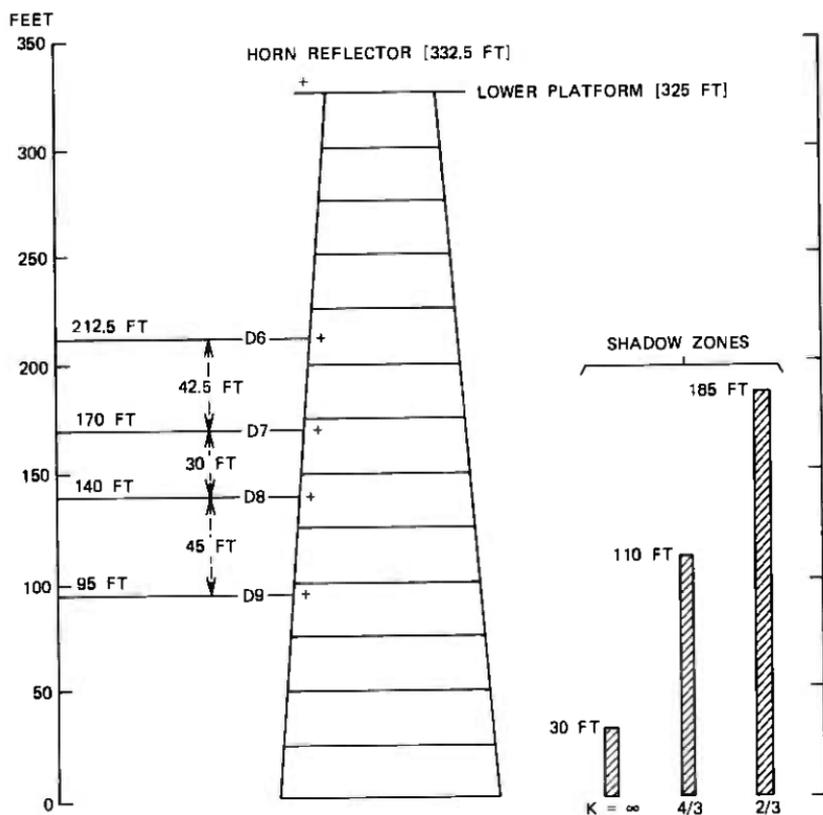
110, and 185 feet from ground. The center line of the horn-reflector receiving antenna is 332.5 feet from ground. The center lines of the antennas added for test purposes are labeled D6 through D9. The antenna at D7 was a small 15-dB-gain horn; the other three are 4-foot-diameter parabolic reflectors. The normal antenna height for reception from Atlanta would be at about D6; the tower is much higher because clearance on another path controls its height.

The antenna of interest is the bottom 4-foot parabolic reflector at D9, with center line 95 feet above ground. The center line is 15 feet below the point where the grazing line under normal conditions ($K = \frac{4}{3}$) projects on the tower. The clearance in Fresnel zones is about $-0.2F_1$. The center line of the bottom antenna falls on the grazing line when K is about 1.72, which corresponds to an index of refraction gradient of about -66 N units per kilometer, which is 27 N units per kilometer higher than the gradient of -39 N units per kilometer under normal $K = \frac{4}{3}$ conditions.

The below-grazing location of the bottom antenna was further verified by noting that there was a signal loss of 14 dB when it was moved from its previous (higher) location on March 14, 1972. A theoretical loss of about 8 dB would occur if the obstruction were a knife edge; diffraction over a tree-covered hilltop accounts for the additional 6-dB loss.

Data were analyzed on received power (recorded by MIDAS[21]) during time intervals containing deep fading (more than 20 dB) on any of the five receiving antennas in April, May, and June of 1972. The total of such time intervals was about 13 hours during which the amount of fading was similar for the top four antennas. The *a priori* expectation was that the bottom antenna would fade more, since its signal level during normal daytime conditions was 14 dB low because of a lack of clearance. The opposite occurred: the signal received by the bottom antenna faded *less* than signals received by the four upper antennas.

The measured time-below-level data (during the 13 hours of deep-fading activity) for the bottom antenna (D9) and the one above it (D8) are shown in Fig. 24 (fading on the other antennas, after adjustment for gain differences, was the same as that on D8). The single antenna curves have the normal slope (decade of time per 10 dB) for deep multipath fades. Time below level for the bottom antenna is smaller than that for D8 by about a factor of three in the deep-fade region.

Diversity performance of the combination of D8 and D9 was excellent. Simultaneous fades on the two antennas were never deeper than about 18 dB from normal (see Fig. 24).

These results relate to aspects of radio propagation not covered by previous experience or theory. Physically, the observations can per-
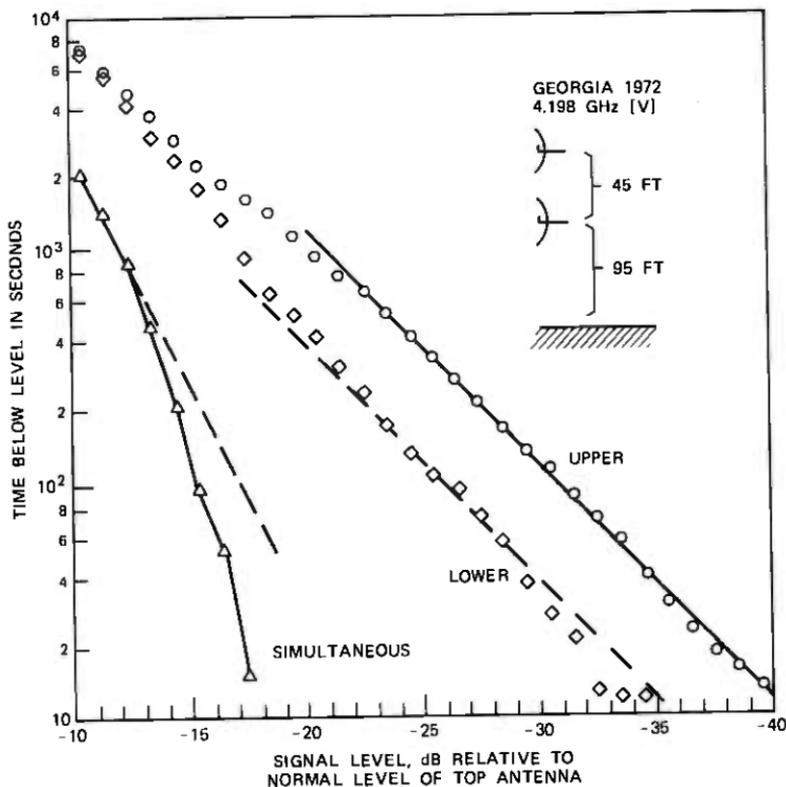
Fig. 24—Fading measured at Palmetto, Georgia.

haps be explained by the presence of layers with strong negative gradients in the index of refraction that create fading but offset the normal lack of clearance for the bottom antenna. Furthermore, since the bottom antenna is relatively close to the ground, the terrain blocks some of the potentially interfering rays, which would also tend to reduce the amount of fading.

A tentative conclusion is that protection from multipath fading can be obtained by placing the bottom antenna somewhat below grazing under normal atmospheric conditions (diffraction loss of perhaps 10 dB). It must be established that antennas placed in such a manner perform well consistently; results from tests at Culver, Indiana (1973–1974) are encouraging.

APPENDIX B

Example of Roughness Calculation

The terrain heights (denoted by x_i, terminal ends omitted) for the 19-mile path in the example of Fig. 25 are provided in Table V. The roughness (standard deviation) is the square root of the average
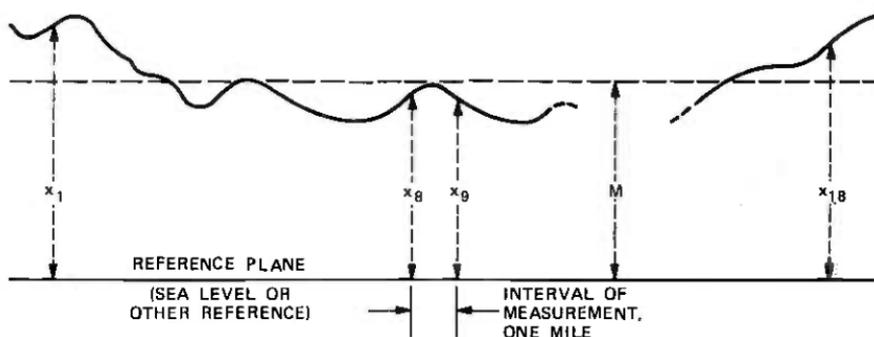
Fig. 25—Determination of terrain roughness.

Table V — Example of terrain heights for 19-mile path (see Fig. 25)

i	x_i (ft)	i	x_i (ft)	i	x_i (ft)
1	600	7	400	13	450
2	625	8	420	14	420
3	515	9	460	15	390
4	440	10	420	16	480
5	480	11	450	17	520
6	450	12	480	18	550

square of the deviation from the mean:

$$w = \sqrt{\frac{1}{18} \sum_{i=1}^{18} (x_i - M)^2}$$

$$= \sqrt{\left[\frac{1}{18} \sum_{i=1}^{18} x_1^2\right] - M^2} \tag{17}$$

$$M = \frac{1}{18} \sum_{i=1}^{18} x_i = 8550/18 = 475 \tag{18}$$

$$w = \sqrt{(4133950/18) - (475)^2} = 63.5 \text{ feet.} \tag{19}$$

APPENDIX C

Length of Fading Season

Multipath fading is a warm-weather phenomenon. Assuming that the length of the warm portion of the year is proportional to the average annual temperature, the value of T_0 to be used in eq. (2) to esti-
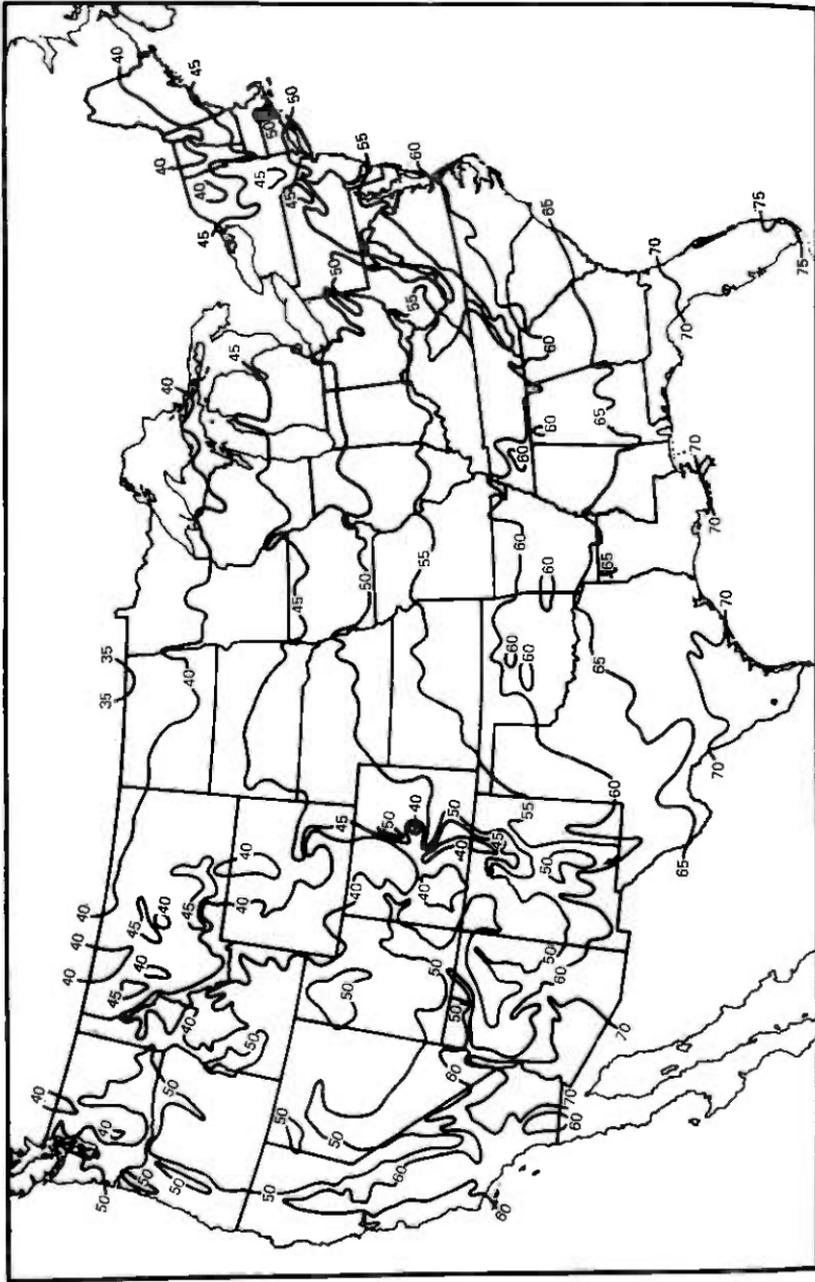
Fig. 26—Average annual temperatures in °F.

mate the annual time below level is

$$T_0 = (t/50)8 \times 10^6 \text{ sec}, \quad 35 \leqq t \leqq 75, \tag{20}$$

where t denotes the average annual temperature of the locality in question in °F as determined from Fig. 26.

In the average case (Figs. 5 and 8), T_0 is equal to the number of seconds in three months; eq. (20) reduces to this when t is 50°F. The temperature contours in Fig. 26 show that 50°F is appropriate for middle latitudes in the United States and for Ohio in particular; data from Ohio have often been used to describe average fading.

The range of average annual temperatures (Fig. 26) is from 35°F in northern North Dakota to 75°F in southern Florida. The corresponding fading season lengths range from 70 percent to 150 percent of average (T_0 ranges from 2 months to $4\frac{1}{2}$ months).

APPENDIX D

Addition of Space Diversity to Frequency Diversity

In the Bell System, fully loaded long-haul routes with radio channels in both the 4- and 6-GHz bands will utilize cross-band frequency-diversity protection with two protection channels—18 working channels in a 2×18 system.[14] Expectations are that space diversity in 2×18 protection systems will be needed only when problem hops are encountered—exceptionally high fading activity or unavoidable ground reflections—that will have to be treated on an individual basis. The need for space-diversity protection will arise on long-haul routes that have radio channels in one frequency band only, 4 or 6 GHz; the single frequency-diversity channel permitted on such routes[1] may not provide adequate protection in geographic areas where fading activity is above average.

The time below level at the fade margin of an average working channel in a frequency-diversity system with one protection channel and m working channels (a $1 \times m$ system) is approximately equal to the simultaneous time below level of channels in an equivalent 1×1 system in which the separation of the center frequencies of the two channels is

$$\Delta f_{eq} = m/[\sum_k (1/\Delta f_k)], \tag{21}$$

where the sum contains $\frac{1}{2}m(m + 1)$ terms, extending the summation over frequency separations of all channel pairs in the $1 \times m$ system. The equivalent 1×1 system (introduced to simplify calculations) is derived from the first term in the alternating series describing transmission unavailability[14] under the assumption that equipment works

perfectly during fading (equipment failures are accounted for via the allocation of transmission availability objectives).

The simultaneous time below level of the two channels in the equivalent 1×1 system is

$$T_{eq} = T/I_{eq}, \tag{22}$$

where T is the time below level in an unprotected channel and

$$I_{eq} = q_{eq}L^{-2}, \tag{23}$$

with q_{eq} determined by using Δf_{eq} in eq. (10). As an example, in a 1×3 system with 60 MHz between channels

$$\begin{aligned}
\Delta f_{eq} &= \frac{3}{(3/60) + (2/120) + (1/180)} \\
&= 41.5 \text{ MHz.}
\end{aligned} \tag{24}$$

Similarly, in a 1×7 system with 30 MHz between channels,

$$\Delta f_{eq} = 15.3 \text{ MHz.} \tag{25}$$

The use of these values of Δf_{eq} to calculate T_{eq} for one hop is illustrated in the example in Table VI and Fig. 27. In a multihop switching section, T_{eq} would be determined by adding values of T_{eq} determined for each hop.
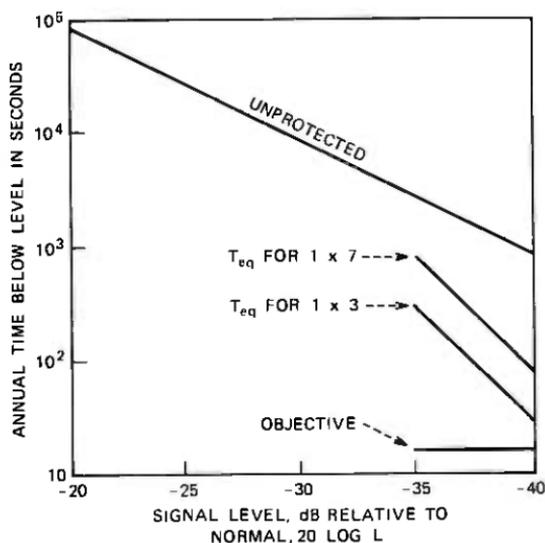
Fig. 27—Example of need for space-diversity protection (see Table VI for path parameters and calculations).

Table VI — Sample calculations

Path parameters	$D = 20$ miles (southern coastal area) $f = 6$ GHz $w = 20$ feet $t = 70°F$
Fading in an unprotected channel	$c = 2(w/50)^{-1.3} = 6.58$ $r = c(f/4)D^3 10^{-5} = 0.79$ $T_0 = (t/50)8 \times 10^6 = 1.12 \times 10^7$ seconds $T = rT_0 L^2 = 8.8 \times 10^6 \, L^2$ seconds/year
Long-haul objective	$1600 \times D/2000 = 16$ seconds/year
Equivalent 1×1 system replacing 1×3	$\Delta f_{eq} = 41.5$ MHz $q_{eq} = (50/6 \times 20) \times (41.5/6000) = 0.0029$ $I_{eq} = 0.0029 \, L^{-2}$
Equivalent 1×1 system replacing 1×7	$\Delta f_{eq} = 15.3$ MHz $q_{eq} = (50/6 \times 20) \times (15.3/6000) = 0.0011$ $I_{eq} = 0.0011 \, L^{-2}$

The long-haul objective (using D_s to denote the length of the switching section) is

$$T_{ob} = 1600 D_s/2000 \text{ sec,} \qquad (26)$$

in terms of which the recommendations for the application of space diversity are:

$$T_{eq} < T_{ob}, \text{ space diversity not needed,} \qquad (27)$$

$$T_{eq} > 1.5 \, T_{ob}, \text{ space diversity needed.} \qquad (28)$$

Use of space diversity is optional in the tolerance band ($T_{ob} \leq T_{eq} \leq 1.5 \, T_{ob}$), which reflects parameter uncertainties and the fact that the approximation used to determine T_{eq} overestimates the unavailability. According to the above criteria, space diversity is needed in the example in Fig. 27 (assuming a one-hop switching section with 40-dB fade margin).

If space-diversity protection is needed, then it should first be applied to the worst hop in a switching section. In subsequent calculations, T_s from eq. (6) should replace T_{eq} for space-diversity-equipped hops, since improvement due to space diversity is normally dominant.

APPENDIX E

Comparative Switching With Hysteresis

The analysis is carried out in terms of the envelope voltages (R_1 and R_2) of the signals from the two receiving antennas. The antennas are

assumed to be of the same size and R_1 and R_2 are normalized to be unity in the absence of fading. The envelope R of the diversity signal is a composite of R_1 and R_2 obtained by switching; R becomes R_1 when $R_1 > bR_2$ and R becomes R_2 when $R_2 > bR_1$. The parameter b ($b \geqq 1$) describes the hysteresis of the switch.

The probability of a fade of $20 \log L$ dB is the probability that $R < L$. Regions in the R_1, R_2-plane over which the joint probability density function $p(R_1, R_2)$ has to be integrated to obtain this probability are shown in Fig. 28. This shows that

$$\Pr\,(R < L) = \int_0^L dR_2 \int_0^L dR_1 p(R_1, R_2) + \frac{1}{2} \int dR_2 \int dR_1 p(R_1, R_2)$$

over triangle 1,

$$+ \frac{1}{2} \int dR_2 \int dR_1 p(R_1, R_2) \tag{29}$$

over triangle 2.

For deep fades and antennas of equal size,[8]

$$p(R_1, R_2) \cong 4q^{-1}R_1 R_2. \tag{30}$$

Substitution and integration gives

$$\Pr\,(R < L) \cong q^{-1}L^4 \frac{b^2 + b^{-2}}{2}. \tag{31}$$

In the absence of hysteresis, b is unity. When hysteresis is introduced, b becomes larger than unity, and the probability of fading is increased by a factor of $0.5(b^2 + b^{-2})$.
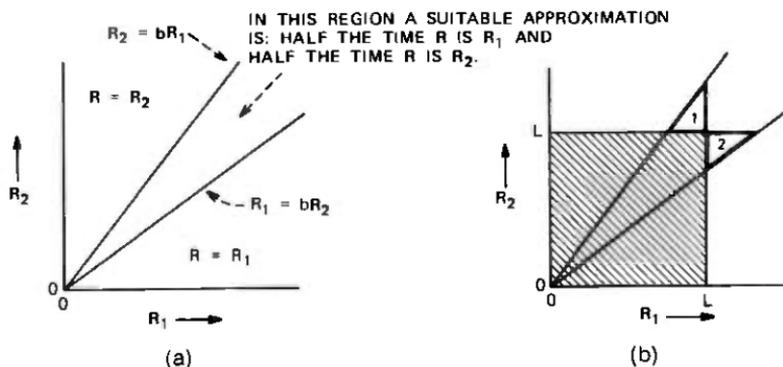
Fig. 28—Comparative switching with hysteresis. (a) Switching diagram. (b) Areas in which $R < L$.

APPENDIX F
Threshold Switching Calculations

Below threshold the switch cycles, and R is R_1 or R_2 in equal amounts of time; the below-level probability for below-threshold operation is, with A denoting the threshold value,

$$\Pr (R < L) = \tfrac{1}{2} \Pr (R_1 < L, R_2 < A)$$
$$+ \tfrac{1}{2} \Pr (R_1 < A, R_2 < L), L < A. \quad (32)$$

Because of symmetry in the joint probability density function, this simplifies to

$$\Pr (R < L) = \int_0^L dR_1 \int_0^A dR_2 p(R_1, R_2), \quad L < A \quad (33)$$

for antennas of equal size. The below-level probability in above-threshold operation is (invoking symmetry as above) the sum of two terms. The first term describes the situation when R is R_1, the second when R has become R_2 upon R_1 dropping below threshold:

$$\Pr (R < L) = \Pr (A < R_1 < L)$$
$$+ \Pr (R_1 < A, R_2 < L), \quad L > A. \quad (34)$$

These equations are based on results obtained for correlated Rayleigh distributed variables.[2] Use of the probability density function from eq. (30) in (33) gives a deep fade approximation:

$$\Pr (R < L) \cong q^{-1}A^2L^2, \quad L < A, \quad (35)$$

which shows the L^2 behavior below threshold.

The curves in Figs. 10 and 11 were calculated (as outlined above) for correlated Rayleigh-distributed variables with q determined from the empirical expression for I_0 (machine computation was necessary because the integrals cannot be evaluated in closed form). The Rayleigh distribution was used in the theoretical calculations because its previous uses have provided insights valuable to the statistical description of space-diversity operation. The theoretically predicted functional forms have been verified by using actual fading records as inputs to computer-simulated threshold switches.[22]

APPENDIX G
Common Control Switching

Space-diversity switching is normally implemented on a per-channel basis: each radio channel has its own switch, and switching decisions for a channel are based, in FM systems, on carrier levels received on
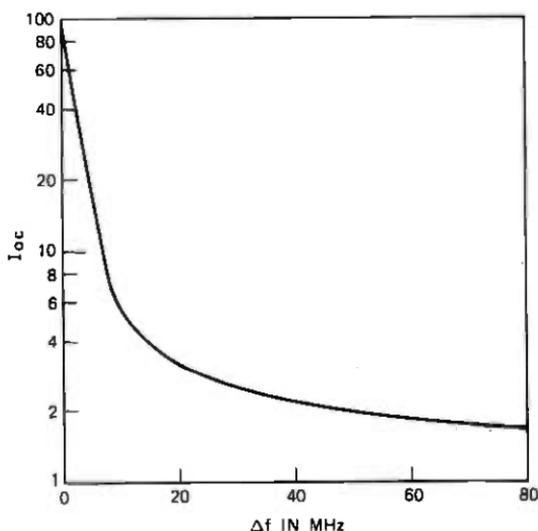
Fig. 29—Improvement in a controlled channel as a function of frequency separation from controlling channel (4-GHz band, 30-foot antenna separation, 26-mile path, 40-dB fade margin).

the upper and lower antennas. Use of the carrier in one channel (controlling channel) to control switching of other channels (controlled channels) would reduce costs. For example, diversity transmission could be controlled by the received carrier in another channel, thus eliminating the feedback link; similarly, a single broadband switch could be used to switch a number of radio channels. Unfortunately, the improvement obtained in such common control schemes is too small to be of practical value.

The available improvement I_{oc} in a controlled channel depends on Δf, the frequency separation of the carriers in the controlling and controlled channels. When Δf is zero, I_{oc} is equal to the available improvement I_0 in the controlling channel, with I_0 given by (7). In the limit, as Δf increases and becomes large, I_{oc} approaches unity as fading at the two frequencies becomes independent. Theoretical estimates[23] show that the initial decrease in I_{oc} as Δf increases from zero is rapid. As an example, the approximate behavior of I_{oc} as a function of Δf in the 4-GHz band is shown in Fig. 29 for a 30-foot antenna separation on a 26-mile path and with a fade margin of 40 dB. The value of I_{oc} at $\Delta f = 0$ is about 100. At one channel separation, $\Delta f = 20$ MHz, I_{oc} has decreased to about 3. Clearly, the improvement in the controlled channels is too small to be of practical value. The lack of performance in common control switching is simply a consequence of the fact that deep multipath fades are highly frequency selective.

Results of Palmetto, Georgia, Experiments in 1968

Power received at 4.198 GHz on five vertically separated antennas (Fig. 30) was measured on a 24.6-mile path from Atlanta to Palmetto, Georgia over a 72-day period in 1968 (August 16–October 27). Five inputs of MIDAS[21] (Multiple Input Data Acquisition System) were used to acquire the data. A power reading from the horn reflector on top was obtained first; about 2 milliseconds later, a reading from the top dish was obtained, and so on down the tower. This scanning was repeated at a rate of five times per second. The readings were converted to a logarithmic scale and recorded in digital form (nominally 1 dB quantizing) on magnetic tape for subsequent computer processing. In the absence of fading, to conserve tape, the recording rate was less than the sampling rate.

Improvement as a function of fade depth for the ten antenna pairs, with separations ranging from 4 to 60 feet center-to-center, was obtained by dividing at a given fade depth the measured time below level for an antenna (T) by the measured simultaneous time below level (T_s). The equation for improvement,

$$I_0 = qL^{-2}, \tag{36}$$

was fitted to the points in the deep-fade region to determine a value of q for each antenna separation ($v^2 = 1$, since gains were set to equalize free-space levels). The values of q so obtained are shown in Fig. 31 as a function of antenna separation; because of random differences in
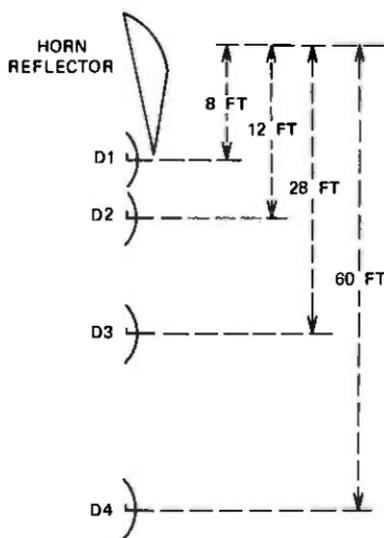
Fig. 30—Receiving antennas at Palmetto, Georgia in 1968. D1 to D4 are 4-foot-diameter parabolic reflectors.
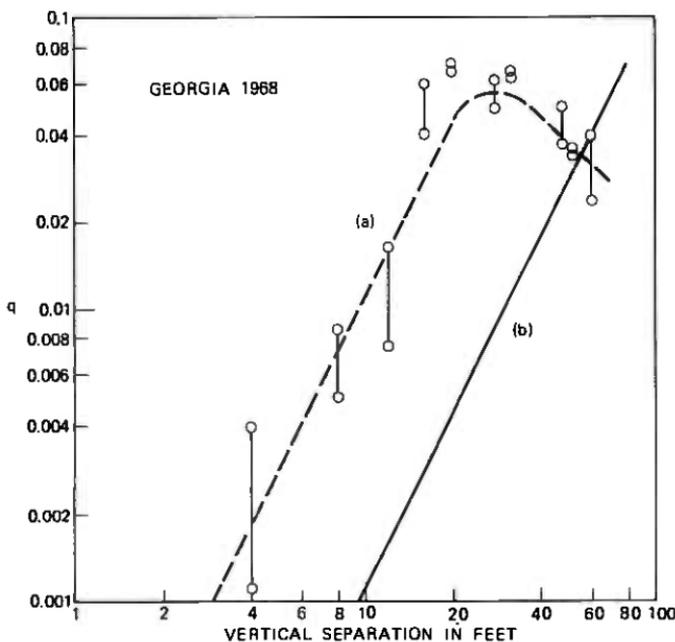
Fig. 31—Variation of correlation parameter q: (a) Georgia 1968, and (b) eq. (9). Increasing values of q denote increasing decorrelation.

T from antenna to antenna, there are two values of q for each antenna separation, depending on whether T for the upper or the lower antenna was used to form the ratio T/T_s.

As a function of antenna separation, the correlation parameter q first increases as the square of the antenna separation, but then levels off and decreases (dotted line in Fig. 31); the presence of a small dominant component in the multipath spectrum is the most likely cause of this. Compared to eq. (9) (see also solid line in Fig. 31), q is enhanced for smaller separations, but is reduced for separations over about 50 feet. Dominant components being random on paths designed for negligible ground reflections, the enhancement cannot be counted upon when choosing an antenna separation; however, the possibility that q may decrease for large separations must be provided for. This is the reason why the applicability (for design purposes) of (7) has been limited to antenna separations of 50 feet or less.

REFERENCES

1. FCC First Report and Order (Docket 18920), June 3, 1971.
2. E. Henze, "Theoretische Untersuchungen über einige Diversity-Verfahren," Archiv Elektrischen Übertragung, *11*, No. 5 (May 1957), pp. 183–194.
3. H. Makino and K. Morita, "The Space Diversity Reception and Transmission Systems for Line-of-Sight Microwave Link," Rev. Elec. Comm. Lab (Japan), *13*, No. 1–2 (January–February 1965), pp. 111–129.

4. H. Makino and K. Morita, "Design of Space Diversity Receiving and Transmitting Systems for Line-of-Sight Microwave Links," IEEE Trans. Commun. Tech., *COM-15*, No. 4 (August 1967), pp. 603–614.
5. R. F. White, "Space Diversity on Line-of-Sight Microwave Systems," IEEE Trans. Commun. Tech., *COM-16*, No. 1 (February 1968), pp. 119–133.
6. R. F. White, *Engineering Considerations for Microwave Communications Systems*, San Carlos, Calif.: Lenkurt Electric Co., June 1970.
7. A. Vigants, "Space-Diversity Performance as a Function of Antenna Separation," IEEE Trans. Commun. Tech., *COM-16*, No. 6 (December 1968), pp. 831–836.
8. A. Vigants, "The Number of Fades in Space-Diversity Reception," B.S.T.J., *49*, No. 7 (September 1970), pp. 1513–1530.
9. W. T. Barnett, "Multipath Propagation at 4, 6, and 11 GHz," B.S.T.J., *51*, No. 2 (February 1972), pp. 321–361.
10. S. H. Lin, "Statistical Behavior of a Fading Signal," B.S.T.J., *50*, No. 10 (December 1971), pp. 3211–3270.
11. A. Vigants, "Number and Duration of Fades at 6 and 4 GHz," B.S.T.J., *50*, No. 3 (March 1971), pp. 815–841.
12. K. W. Pearson, "Method for the Prediction of the Fading Performance of a Multisection Microwave Link," Proc. IEE (London), *112*, No. 7 (July 1965), pp. 1291–1300.
13. A. Vigants, "Observations of 4 GHz Obstruction Fading," Proc. IEEE Int. Conf. on Commun., *8*, 1972, pp. 28-1 to 28-2.
14. W. Y. S. Chen, "Estimated Outage in Long-Haul Radio Relay Systems with Protection Switching," B.S.T.J., *50*, No. 4 (April 1971), pp. 1455–1485.
15. S. H. Lin, "Statistical Behavior of Deep Fade of Diversity Signals," IEEE Trans. Commun., *COM-20*, No. 6 (December 1972), pp. 1100–1107.
16. W. T. Barnett, unpublished work.
17. W. T. Barnett, unpublished work.
18. A. Vigants, "Threshold-Switched LOS Space-Diversity Reception in the Everglades," Conf. Record, pages 12C-1 to 12C-3, IEEE Int. Conf. Commun., June 17–19, 1974 Minneapolis, Minnesota; Institute of Electrical and Electronics Engineers, 1974.
19. A. Vigants, "Space Diversity on a 6-GHz Path with Billboard Reflectors," Conference Record, pages 12B-1 to 12B-4, International Conference on Communications, June 17–19, 1974, Minneapolis, Minnesota; Institute of Electrical and Electronics Engineers, 1974.
20. A. Vigants, "Variation of Fading with Antenna Height and Size," URSI/IEEE G-AP Int. Symp., Williamsburg, Va., December 11–15, 1972.
21. K. Bullington, "Unlocking the Secrets of Microwave Propagation," Bell Laboratories Record, *50*, No. 1 (January 1972), pp. 8–13.
22. M. V. Pursley, unpublished work.
23. A. Vigants, unpublished work.

A New, Fast-Converging Mean-Square Algorithm for Adaptive Equalizers With Partial-Response Signaling

By K. H. MUELLER

A new, generalized mean-square algorithm is presented to adjust the taps of a transversal adaptive equalizer. Any knowledge of the channel or the signaling format can be taken into account and will speed up the convergence process. The main applications are seen in partial-response signaling, where the new algorithm eliminates the interaction between the individual tap increments. This is achieved by decorrelating the components of the gradient in a fixed weighting matrix prior to the adjustments. Convergence is then extremely fast. If the channel has only phase distortion, a single iteration is sufficient to obtain the optimum tap vector (for any timing and carrier phase). This is verified by computer simulations. Finally, the new equalizer is compared with another structure recently proposed, and some advantages of the new system with regard to implementation and flexibility are pointed out.

I. INTRODUCTION

Together with sophisticated modulation techniques and bandwidth-conserving signal designs, automatic transversal equalizers are the prime characteristics of the new generation of high-speed data modems that have appeared on the market in the past few years. These modems allow data rates in the range of 3600 to 9600 b/s to be transmitted over voice-band channels. In addition to point-to-point transmission between computers and data processing centers, a large demand for high-speed data sets exists in multiparty polling systems. Because a message in such a system may consist of only a few hundred bits, it is essential that the start-up time of the modem be very short. Preferably, the modem start-up time should waste fewer bits than are contained in an average message to guarantee a reasonable system throughput. This would basically imply that start-up time should be inversely proportional to the data rate; this is a very contradictory requirement, since more accurate and complex operations are usually

required during start-up of a high-speed modem than with a low-speed version. Timing recovery, carrier recovery (if required), and training of the automatic equalizer are the most important start-up operations. The time that must be allowed for the equalizer training is usually the major delay in start-up; thus, it seems worthwhile to concentrate some of our efforts in this direction.

During the past few years, much attention has been given to partial response signaling.[1-6] Some of these waveforms are particularly attractive for data transmission over band-limited channels, since they have a spectral zero at the Nyquist frequency. These signals can easily be generated with virtually zero excess bandwidth, and yet their time response decays faster than that of a $\sin (x)/x$ pulse. These advantages result from a controlled amount of intersymbol interference in these signals, which increases the number of levels in the baseband eye pattern. In some formats, the impulse response has odd symmetry and the spectrum is zero at dc (which is of particular interest if SSB modulation is used).[5] If a traditional transversal automatic equalizer with an MS (mean-square) gradient algorithm[6,7] is used for partial response signals, the initial convergence will be very slow. Chang has shown that the spread of the eigenvalues of the signal autocorrelation matrix can become very large for Class IV partial response signals, and he proposed a new equalizer structure that eliminates this problem and thus can give fast convergence.[8] Of course, with the appropriate modification of some parameters, his conclusions can be extended to other partial-response signals as well.

In this paper, we propose an alternate solution to the problem of fast partial-response signal equalization. Instead of changing the equalizer structure, we propose to modify the tap updating algorithm. The coefficients associated with the new algorithm are given by a simple matrix inversion.

The advantages of the new equalizer structure are not limited to partial-response signaling. Any knowledge of the channel or the signaling format can be used to speed up the convergence process. The classic transversal equalizer evolves as a special case (Nyquist signaling) of the more general configuration that is proposed in this paper.

II. A GENERAL MEAN-SQUARE ALGORITHM

Let us assume a transversal equalizer structure with N taps, with a tap signal input vector $\mathbf{x}_k$ and a tap coefficient vector $\mathbf{c}$ at time $t_0 + kT$. The equalizer output y_k is given by

$$y_k = \mathbf{x}_k^T \mathbf{c} \tag{1}$$

and will not, in general, be exactly equal to the desired reference value d_k, the error being

$$e_k = y_k - d_k. \tag{2}$$

The MS error that appears at the output will thus be

$$\epsilon^2 = E\{e_k^2\}, \tag{3}$$

and we now choose a tap-updating algorithm that adjusts c in such a way that the MS error (3) is minimized. The gradient of ϵ^2 with respect to c can easily be obtained by combining (1) to (3),

$$\mathbf{g} = \frac{1}{2}\frac{\partial \epsilon^2}{\partial \mathbf{c}} = E\{e_k \mathbf{x}_k\} = A\mathbf{c} - \mathbf{v}, \tag{4}$$

where we have introduced the signal autocorrelation matrix,

$$A = E\{\mathbf{x}_k \mathbf{x}_k^T\}, \tag{5}$$

and the signal correlation vector,

$$\mathbf{v} = E\{\mathbf{x}_k d_k\}. \tag{6}$$

The optimum tap vector in the mean-square sense that minimizes (3) is obtained when we set the gradient to zero,

$$\mathbf{c}_{opt} = A^{-1}\mathbf{v}. \tag{7}$$

We propose a tap-updating algorithm of the form

$$\mathbf{c}_{m+1} = \mathbf{c}_m - Q_m \mathbf{g}_m, \tag{8}$$

where $\mathbf{g}_m$ is the gradient evaluated with tap-vector $\mathbf{c}_m$ after the mth iteration, and Q_m is a nonsingular matrix, which is discussed below. Obviously, an algorithm of the general form (8) will stop updating (turn itself off) when $\mathbf{c} = \mathbf{c}_{opt}$. If the algorithm converges at all, it will thus converge to the MS solution (7). The traditional steepest-descent gradient algorithm is obtained as a special case from (8) if we set $Q_m = \beta I$ (or a more general diagonal matrix instead of the unity matrix I if tapering should be included).

The algorithm (8) can be written in the form

$$\mathbf{c}_{m+1} = (I - Q_m A)\mathbf{c}_m + Q_m \mathbf{v}. \tag{9}$$

If we introduce the tap error vector

$$\boldsymbol{\phi}_m = \mathbf{c}_m - \mathbf{c}_{opt}, \tag{10}$$

it is easy to show that

$$\boldsymbol{\phi}_{m+1} = (I - Q_m A)\boldsymbol{\phi}_m = \prod_{n=1}^{m}(I - Q_m A)\boldsymbol{\phi}_1. \tag{11}$$

Obviously, the convergence can be controlled to some degree by the selection of Q_m, since this choice affects the eigenvalues of $I - Q_m A$. This is discussed in Section III.

III. SELECTION OF THE Q MATRIX

First we will make some general remarks without specifying that the signal format be partial response. If the channel characteristics, demodulating carrier phase, and timing instant were perfectly known, we would select $Q = Q_m = A^{-1}$ and then obtain c_{opt} after the first iteration. Actually, if that knowledge were available, it would, of course, be easier to preset the taps to c_{opt} or to introduce a suitable fixed compensation network. Unfortunately, the transmission channel is not completely known. If, however, the spread within these channels is not extreme, we can reduce the average equalizer training time if we select a good estimate of the optimum tap vector for initial presetting. For example,

$$\mathbf{c}_1 = E\{A_i^{-1}\mathbf{v}_i\}, \tag{12}$$

where the expected value is taken over the ensemble of channels, and a similar estimate for Q,

$$Q_m = \beta_m [E\{A_i\}]^{-1}. \tag{13}$$

For a particular channel with

$$A_i = E\{A_i\} - \Delta_i, \tag{14}$$

the critical matrix $B_i = I - Q_m A_i$ then becomes

$$B_i = (1 - \beta_m)I + \beta_m [E\{A_i\}]^{-1}\Delta_i. \tag{15}$$

If we select $\beta_m = 1$, the average B_i matrix is equal to the zero matrix. For a randomly selected channel, the B_i matrix will, on the average, be much closer to the zero matrix than with a gradient algorithm, and faster convergence may be obtained. Modern digital signal-processing techniques allow quick and easy change (or selection) of (12) and (13) with different read-only memories at a fraction of the costs that are involved when traditional compromise equalizers have to be changed. Furthermore, even different signal formats may be processed efficiently with the same modem.

IV. APPLICATIONS TO PARTIAL-RESPONSE SIGNALING

Our results have so far been quite general, as we have not yet defined a special signaling technique. As was already mentioned in Section I, we see a particular promising application to processing zero excess-bandwidth partial-response signals. For such signals, the A

matrix depends only on the amplitude characteristics of the channel and is independent of the phase parameters[8] (timing phase, carrier phase, and phase characteristic of the channel). Since the amplitude response of voice-grade channels is relatively constant over the transmission band (or at least under much better control than the phase response), we may select an estimate

$$Q_m = \beta_m A_0^{-1}, \tag{16}$$

where A_0 is the ideal partial-response signal matrix. More specifically, for Class IV signaling, the matrix elements of A_0 are

$$[A_0]_{ik} = \delta_{ik} - \tfrac{1}{2}\delta_{i,k+2} - \tfrac{1}{2}\delta_{i,k-2} \tag{17}$$

and for duobinary are

$$[A_0]_{ik} = \delta_{ik} + \tfrac{1}{2}\delta_{i,k+1} + \tfrac{1}{2}\delta_{i,k-1}. \tag{18}$$

The inverse of both (nonsingular) matrices may be expressed as

$$A_0^{-1} = (I-G)^{-1} = I + G + G^2 + \cdots, \tag{19}$$

which shows that the elements of A_0^{-1} are rational numbers, and are symmetric with respect to both matrix diagonals.

Having selected an ideal channel as an estimate for Q, our algorithm takes the form

$$\phi_{m+1} = (I - \beta_m A_0^{-1}A)\phi_m = B_m\phi_m. \tag{20}$$

Obviously, if we have only phase distortion, then $A = A_0$ and we may select $\beta_m = 1$. The equalizer will then reach its optimum tap vector within a single iteration because the matrix B_m is zero.

If the channel has amplitude distortion, then $A \neq A_0$ and thus $B_m \neq 0$. The eigenvalues λ_i and eigenvectors $\mathbf{w}_i$ of $A_0^{-1}A$ are defined by

$$A_0^{-1}A\mathbf{w}_i = \lambda_i\mathbf{w}_i. \tag{21}$$

Following a procedure shown by Chang,[8] we premultiply both sides by $\mathbf{w}_i^\dagger A_0$ and obtain

$$\lambda_i = \frac{\mathbf{w}_i^\dagger A \mathbf{w}_i}{\mathbf{w}_i^\dagger A_0 \mathbf{w}_i}. \tag{22}$$

Since both A_0 and A are symmetric, positive, definite matrices, we conclude that all λ_i are positive. Chang[8] has further shown that the eigenvalues can be bounded by

$$|H(\omega)|^2_{\min} \leqq \lambda_i \leqq |H(\omega)|^2_{\max}, \tag{23}$$

where $H(\omega)$ is the quotient between the actual and the estimated channel transfer function. For convergence, the spectral radius ρ of

B_m must be less than unity; this is guaranteed if we choose

$$0 < \beta_m < \frac{2}{\lambda_{\max}}, \qquad (24)$$

since the eigenvalues μ_i of B_m are given by

$$\mu_i = 1 - \beta_m \lambda_i. \qquad (25)$$

If the amplitude distortion is small, $A_0^{-1}A$ is close to the identity matrix, and the eigenvalues λ_i will be closely scattered around unity. We would then select $\beta_m \approx \text{const} \approx 1$, and the matrix B would be close to the zero matrix with all eigenvalues μ_i close to zero. The tap error vector after m iterations is given, from (20),

$$\phi_m = (I - \beta A_0^{-1}A)^m \phi_0 = B^m \phi_0 \qquad (26)$$

if β is constant (this is, of course, replaced by the usual matrix product if $\beta = \beta_m$). The squared magnitude of ϕ and the excess MS error are given by the expressions

$$|\phi_m|^2 = \phi_0^T B^{2m} \phi_0 \qquad (27)$$

$$\epsilon^2 = \phi_0^T B^m A B^m \phi_0, \qquad (28)$$

and can be used to estimate the convergence speed. The procedure and the results are practically identical to those of Chang[8] and are, therefore, not repeated here.

V. IMPLEMENTATION

Figure 1 is a basic block diagram of the new equalizer. A version with only three taps is shown for clarity, although a much larger and preferably even number of taps would actually be used. The upper part represents a traditional MS equalizer with the well-known circuitry for correlating the error signal with each tap signal to obtain the gradient $\mathbf{g} = A\mathbf{c} - \mathbf{v}$. Because of the particular structure of A for partial response signals, the components of $\mathbf{g}$ are highly correlated. The tap corrections obtained with a direct gradient algorithm would thus not be independent of each other. This interaction (which would cause a very slow convergence) can be eliminated if the gradient is "decorrelated" in the fixed weighting matrix Q shown in the lower part of Fig. 1. The resulting outputs are now decoupled and can be used to adjust the taps free of interaction to obtain rapid convergence.

A somewhat different arrangement is shown in Fig. 2. Here the (same) Q matrix is placed between the tap signals and the correlators. It is obvious that this arrangement is equivalent to Fig. 1 because the error signal is common to all correlator inputs.
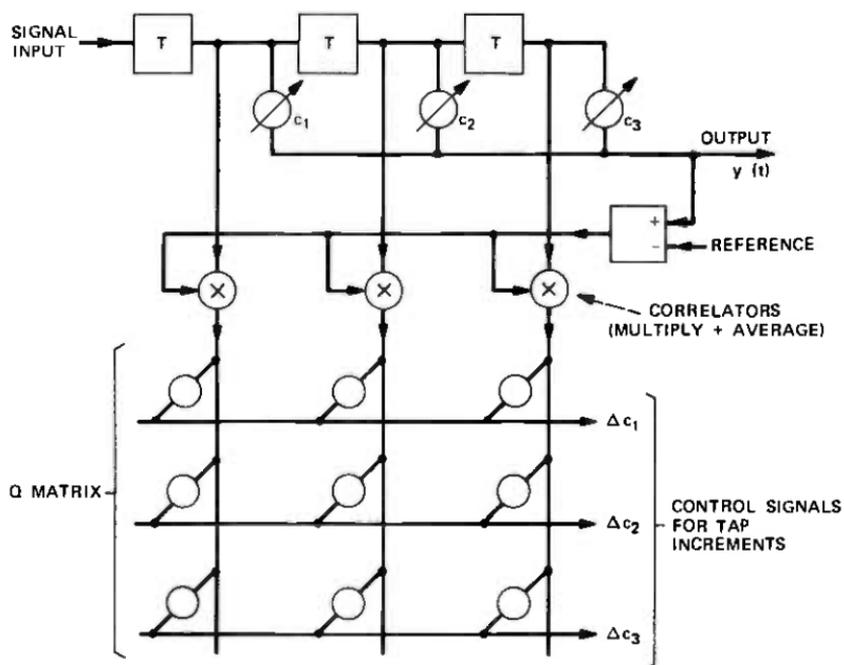
Fig. 1—Use of Q matrix in a first implementation of the new algorithm.
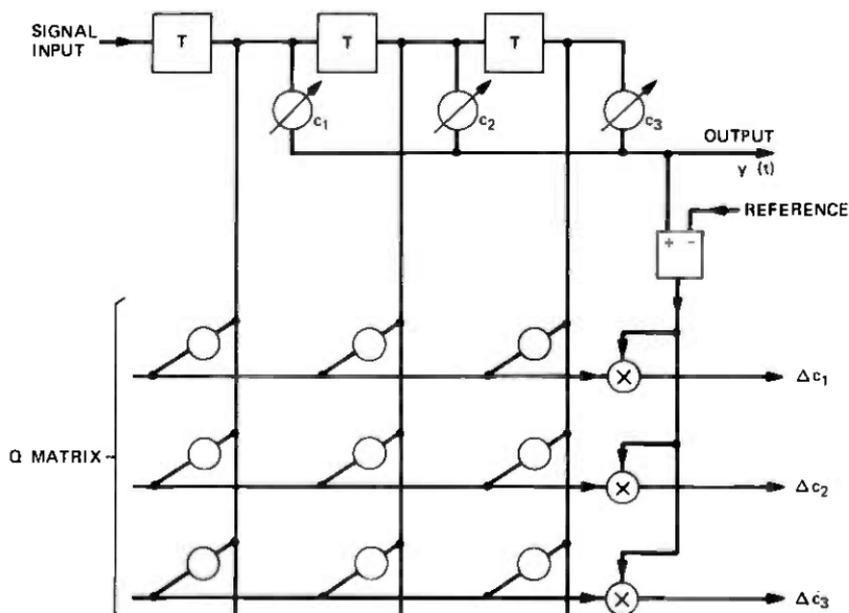
Fig. 2—Use of Q matrix in a second implementation of the new algorithm.
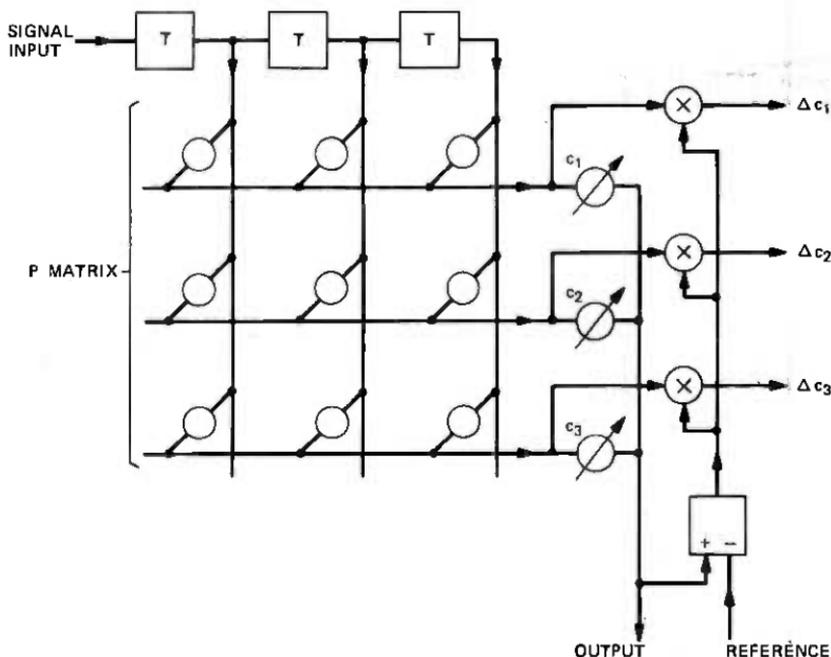
Fig. 3—Use of P matrix to obtain a modified equalizer structure.

As a comparison, we show in Fig. 3 a three-tap equalizer proposed by Chang.[3] Here the input signal is first transformed by a weighting matrix P into a set of orthonormal components that are then processed as in a conventional equalizer. It is, in fact, identical to a conventional equalizer except that the input signal vector $\mathbf{x}$ is replaced by $P\mathbf{x}$, where P is a suitably chosen matrix such that $P^T P = A_0^{-1}$. Initial presetting of the equalizer and constraining the range of the variable gain coefficients can, however, be quite complicated with this structure. These difficulties are avoided in the new proposal, since the equalizer structure remains unchanged and only the updating algorithm is modified.

Although the number of multiplications is the same in all three structures (Figs. 1 to 3), we would like to point out one very important difference: in Chang's structure, the accuracy of the transformation is very critical. The input signals must be represented by their full precision (say, 10 bits). To avoid round-off errors when all the products are summed up, more than 12 accurate bits of these products must be processed and can then (after summation) probably be rounded off to 10 bits for the transformed tap signal. Any errors introduced in this process will result in undesirable output noise that cannot be compensated for by the equalizer. In the new equalizer, however, only the tap increments are affected by errors in the matrix calculation. Obviously,

a much lower accuracy is needed; in the extreme case, the correct polarity will be sufficient. Even with comparatively large errors, the tap-vector noise can be kept low by suitable scaling of the increments. We thus conclude that the new equalizer has inherent advantages as far as complexity, cost, size, speed requirements, and time-sharing possibilities are concerned. Furthermore, simpler algorithms can be derived from this one. If, for example, the structure of Fig. 2 is used for an algorithm with tap signals quantized to one bit, the matrix operation is reduced to controlling the addition and subtraction of the coefficients and may, for a medium-sized equalizer, be cheap and efficiently replaced by a suitably programmed read-only memory. For such nonmean-square algorithms, the optimum Q matrix will, in general, of course, be different from A_0^{-1}.

A few words should be said about the structure of the matrix $Q = A_0^{-1}$ in the special case of partial-response Class IV signaling. It can be shown that every second diagonal array is zero; furthermore, if N is even, every second row (or column) is a shifted version of the previous one. These properties are identical to those of Chang's P matrix (the remaining nonzero coefficients are, however, different from those of the P matrix) and can be used to simplify the signal-processing operations.

Finally, we would like to point out that an all-digital implementation for higher data speeds (group band) would still be extremely complex and expensive. For such applications, an array of operational amplifiers with suitable weighting resistors (probably mounted on a ceramic) would be a far more economic solution with the present state of the art.

VI. SIMULATIONS

To test the proposals made in this paper, a 15-tap MS equalizer was simulated. A 4800-b/s, SSB-modulated, partial-response, Class IV signal was selected. A parabolic delay characteristic was assumed with 1-ms delay difference between the passband channel center and edges. Table I shows the coefficients of the matrix $Q = A_0^{-1}$. In Fig. 4, the equalizer convergence is shown for four different timing phases that are separated by 90 degrees each. After each iteration, the MS distortion of the overall impulse response is plotted. It is seen that the minimum is basically achieved with the first iteration (the small residual error after the first iteration results from the time-domain truncation and round-off errors in the components of the initial impulse response during the simulation). Some additional simulations were made with an MS stochastic approximation algorithm (adjustments at the symbol rate) and with amplitude distortion; they also showed a clearly improved convergence when the decorrelation matrix was used.

Table I — Decorrelation matrix for $N = 15$
(elements multiplied with 36)

64	0	56	0	48	0	40	0	32	0	24	0	16	0	8
0	63	0	54	0	45	0	36	0	27	0	18	0	9	0
56	0	112	0	96	0	80	0	64	0	48	0	32	0	16
0	54	0	108	0	90	0	72	0	54	0	36	0	18	0
48	0	96	0	144	0	120	0	96	0	72	0	48	0	24
0	45	0	90	0	135	0	108	0	81	0	54	0	27	0
40	0	80	0	120	0	160	0	128	0	96	0	64	0	32
0	36	0	72	0	108	0	144	0	108	0	72	0	36	0
32	0	64	0	96	0	128	0	160	0	120	0	80	0	40
0	27	0	54	0	81	0	108	0	135	0	90	0	45	0
24	0	48	0	72	0	96	0	120	0	144	0	96	0	48
0	18	0	36	0	54	0	72	0	90	0	108	0	54	0
16	0	32	0	48	0	64	0	80	0	96	0	112	0	56
0	9	0	18	0	27	0	36	0	45	0	54	0	63	0
8	0	16	0	24	0	32	0	40	0	48	0	56	0	64
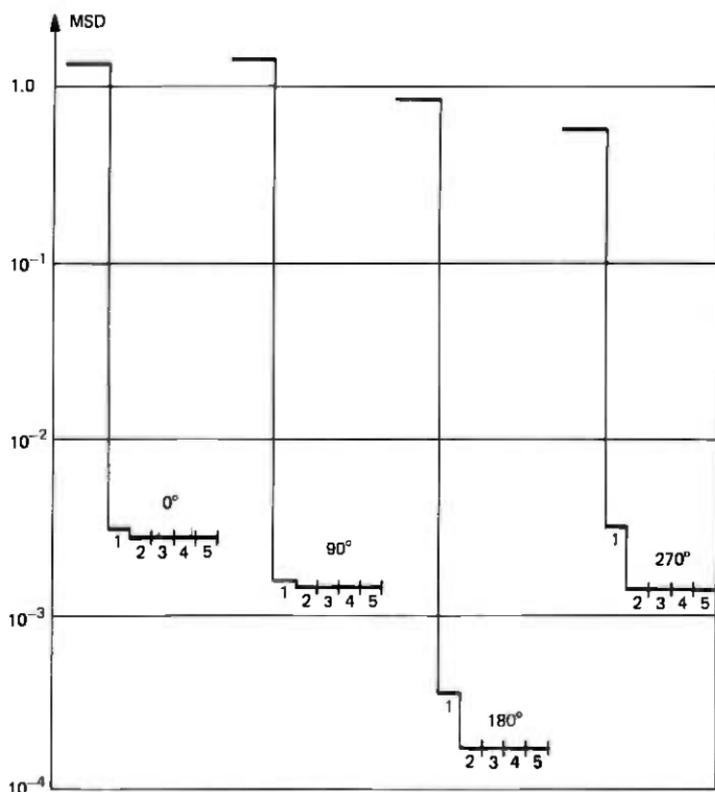
Fig. 4—Convergence of new algorithm for different timing phases.

VII. CONCLUSIONS

The spread of the eigenvalues of the signal autocorrelation matrix can become very large for signals with nonflat amplitude spectra. If the gradient with respect to the tap gains of an MS equalizer is calculated, its components are highly correlated. The resulting interaction of the individual gain adjustments may cause a very slow equalizer convergence. We have proposed a new MS algorithm that eliminates these problems. With this new algorithm, a set of uncorrelated tap increments is calculated so that the individual adjustments are decoupled from each other. This can be achieved in either of two ways: The gradient can first be calculated in the conventional way (correlation of error signal and tap signals) and then be multiplied by a decorrelation matrix Q to produce the final correction vector. On the other hand, tap signals can first be passed through such a matrix and the resulting outputs can then be correlated with the error signal. The matrix Q is chosen as the inverse of the expected (average) signal autocorrelation matrix A. The new algorithm was simulated for partial response signals with zero excess bandwidth. In the case of a flat channel, the optimum tap coefficients can then be determined within a single iteration, independent of delay distortion, carrier phase, and timing instant. This was confirmed by simulation. Compared with other recent proposals for fast start-up, the new algorithm features simpler presetting and requires less accuracy in the matrix transformation. Furthermore, this method can easily be extended to systems with nonmean-square cost functions.

REFERENCES

1. E. R. Kretzmer, "Generalization of a Technique for Binary Data Transmission," IEEE Trans. on Commun. Technology, *COM-14*, No. 1 (February 1966), pp. 67–68.
2. A. Lender, "Correlative Digital Communication Techniques," IEEE Trans. on Commun. Technology, *COM-12*, No. 4 (December 1964), pp. 128–135.
3. A. Lender, "The Duobinary Technique for High Speed Data Transmission," IEEE Trans. on Commun. and Elec., *82*, No. 66 (May 1963), pp. 214–218.
4. P. J. vanGerwen, "On the Generation and Application of Pseudo-Ternary Codes in Pulse Transmission," Philips Research Reports, *20* (August 1965), pp. 469–484.
5. P. Leuthold and F. Tisi, "Ein Einseitenbandsystem fuer Datenuebertragung," Archiv der Elektrischen Uebertragung, *21*, No. 7 (July 1967), pp. 354–362.
6. R. W. Lucky, J. Salz, and E. J. Weldon, *Principles of Data Communication*, New York: McGraw-Hill, 1968.
7. D. Hirsch and W. J. Wolf, "A Simple Adaptive Equalizer for Efficient Data Transmission," IEEE Trans. on Commun. Technology, *COM-18*, No. 1 (February 1970), pp. 5–12.
8. R. W. Chang, "A New Equalizer Structure for Fast Start-Up Digital Communication," B.S.T.J., *50*, No. 6 (July–August 1971), pp. 1969–2014.
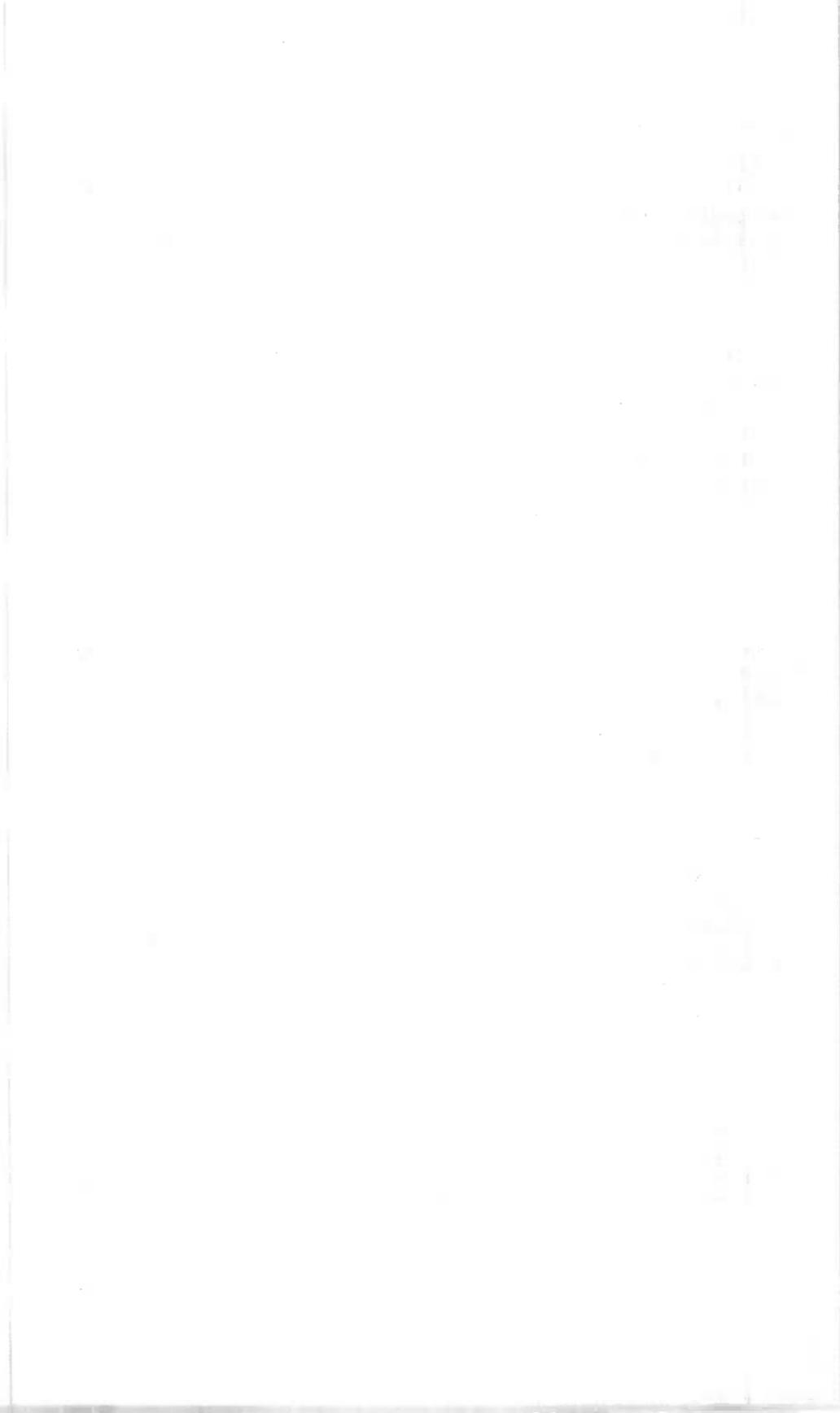
Numerical Evaluation of Integrals with Infinite Limits and Oscillating Integrands

By S. O. RICE

(Manuscript received March 15, 1974)

The numerical evaluation of slowly convergent integrals with infinite limits and regularly oscillating analytic integrands is discussed. Functions are presented that can be subtracted from the integrands to increase the rate of convergence. One of the integrals used as an example is an often-studied Fourier integral related to the power spectrum of a phase-modulated wave. It is pointed out that, when Fourier integrals whose values are known to be positive, e.g., probability densities, are evaluated by the trapezoidal rule, the values given by the trapezoidal rule are never less than the true values.

I. INTRODUCTION

This paper is in the nature of an addendum to an earlier paper in this journal dealing with the numerical evaluation of definite integrals.[1]

Integrals with infinite limits and oscillating integrands often arise in technical problems. In this paper, we assume that the integrand is an analytic function of the variable of integration u (in a suitable region) and tends to oscillate at a regular rate as $u \to \infty$. The main concern here is how to deal with cases in which the rate of convergence is slow.

A number of ways have been proposed to handle the slow convergence. One is to use integration formulas of Filon's type. In particular, E. O. Tuck[2] has given a formula of this type suited to an infinite range of integration. Another method is to evaluate the contributions of positive and negative loops of the integrand and apply Euler's summation formula for slowly converging alternating series. Still another is to (*i*) write the integrand in terms of complex exponentials, (*ii*) deform the path of integration so that it becomes a path (or portions of paths) of steepest descent in the complex u-plane, and then (*iii*) integrate along straight-line segments that approximate the path. A closely related procedure is to tilt the path of integration, if possible, so that the integrand will decrease exponentially.

In this paper most of our attention will be devoted to still another method of improving the convergence, namely, that of subtracting from the integrand functions that behave like the leading terms in the asymptotic expansion of the integrand. This method is quite old. It is related to a method used by Kummer to deal with series, and its application to integrals is described by Krylov.[3]

Sections II through V are concerned with functions useful for subtraction when the limits of integration are 0 and ∞. Section VI deals with the evaluation of

$$J(b, a) = e^{-b}\int_{-\infty}^{\infty} [\exp (bu^{-1}\sin u) - 1 - bu^{-1}\sin u]e^{iau}du, \quad (1)$$

which serves as a typical example of a slowly converging Fourier integral with limits $\pm\infty$.

The integral (1) has been the subject of several papers because of its relation to the problem of calculating the power spectrum of a phase-modulated wave (see Prabhu and Rowe[4] and the references they give). In Section VI, results of trial calculations are presented that show that (1) can be efficiently evaluated by the trapezoidal rule after suitable functions have been subtracted from the integrand.

Sections VII and VIII are concerned with the error introduced when a general Fourier integral with limits $\pm\infty$ is evaluated by the trapezoidal rule.

II. PROPERTIES OF FUNCTIONS SUITABLE FOR SUBTRACTION— LIMITS 0 AND ∞

Let the integral to be evaluated be

$$I = \int_0^{\infty} f(u)du,$$

and let the leading term in the asymptotic expansion of $f(u)$ for large u be $u^{-\nu}\exp(iau)$, where a is real and $\nu > 0$. Quite often, in performing numerical integrations, this leading term is subtracted from the integrand only for values of u that exceed some large value U. We have a slightly different procedure in mind here.

Here we want to use the trapezoidal rule methods described in Ref. 1. Since these methods require the integrand to be analytic, we assume that $f(u)$ is analytic in a region containing the positive real axis. Our aim is to subtract an analytic function (which approaches $u^{-\nu}\exp(iau)$ as $u \to \infty$) from $f(u)$ over the *entire* interval $(0, \infty)$. When this is done, the resulting integral has an integrand analytic over $(0, \infty)$ and can be evaluated by the trapezoidal rule methods described in Ref. 1 (see especially Section IX).

Functions suitable for subtraction when the limits are 0 and ∞ should have the following properties:

(i) As $u \to \infty$, they should be asymptotic to $u^{-\nu} \cos au$ or $u^{-\nu} \sin au$ or, more generally, to

$$u^{-\nu} \exp{(iau)}, \tag{2}$$

where $\nu > 0$ and a is real.

(ii) They should be easy to compute and they should remain finite at $u = 0$.

(iii) Their integrals between the limits $(0, \infty)$ should be easy to compute.

In some cases, moderately simple functions can be used for subtraction. If the leading term in the asymptotic expansion of the integrand is $(\sin au)/u$, then we can subtract and add the integral

$$\int_0^\infty u^{-1} \sin au \, du = \begin{array}{ll} \pi/2, & a > 0 \\ -\pi/2, & a < 0. \end{array} \tag{3}$$

Similarly, for $(\cos au)/u$ we can use

$$\int_0^\infty u^{-1} [\cos au - \exp{(-au)}] du = 0, \tag{4}$$

and, for $(\cos au)/u^2$,

$$\int_0^\infty (\cos au) du/(1 + u^2) = \tfrac{1}{2}\pi \exp{(-|a|)}. \tag{5}$$

The last integral (5) has the disadvantage that its integrand has the asymptotic expansion

$$\frac{\cos au}{u^2} - \frac{\cos au}{u^4} + \cdots, \tag{6}$$

and if we desire to subtract terms of both $O(u^{-2})$ and $O(u^{-4})$ from the original integrand, the term $-(\cos au)/u^4$ in (6) has to be taken into account. In the applications we have in mind, the $\exp{(-au)}$ in (4) decreases so rapidly that it does not have to be taken into account in the way that $-(\cos au)/u^4$ does.

III. GENERAL FUNCTIONS FOR SUBTRACTION—LIMITS 0 AND ∞

Here we present a set of functions that meets the "ease of computation" requirements (ii) and (iii) reasonably well. It is quite likely that there are better sets, but this one is the best I have been able to find. There are two formulas according to whether ν is an integer or not.

The functions are the integrands in

$$\int_0^\infty u^{-n-\epsilon}(e^{-\alpha u} - e^{-\beta u})^n du = \frac{\Gamma(1-\epsilon)}{(\epsilon)_n} \sum_{k=0}^n (-)^{n-k}\binom{n}{k}$$
$$\times [(n-k)\alpha + k\beta]^{n+\epsilon-1}, \quad (7)$$

$$\int_0^\infty \left(\frac{e^{-\alpha u} - e^{-\beta u}}{u}\right)^n du = \frac{1}{(n-1)!} \sum_{k=0}^n (-)^{n-k}\binom{n}{k}$$
$$\times [(n-k)\alpha + k\beta]^{n-1} ln[(n-k)\alpha + k\beta], \quad (8)$$

where n is a positive integer and n, α, β, and ϵ are such that the integrals converge. In (7), ϵ is not an integer, and $(\epsilon)_0 = 1$, $(\epsilon)_n = \epsilon(\epsilon+1)\cdots$ $(\epsilon + n - 1)$. In using (7) and (8), we set $\alpha = -ia/n$, $\beta = |a|/n$. In (7), n and ϵ are chosen so that $\epsilon < -1$ and $n + \epsilon$ is equal to the exponent ν in (2). The choice $\epsilon < -1$ is made to make the integrand and its derivative in (7) approach 0 as $u \to 0$. This tends to reduce the error in the numerical integration. Since a can be either positive or negative in $\alpha = -ia/n$ and $\beta = |a|/n$, we take

$$-\pi/2 \leqq \arg[(n-k)\alpha + k\beta] \leqq \pi/2$$

when calculating the logarithm and the $(n + \epsilon - 1)$th power of $[(n-k)\alpha + k\beta]$.

Equation (8) is obtained by letting $\epsilon \to 0$ in (7), and (7) is based upon (page 243 of Ref. 5)

$$\Gamma(1-n-\epsilon) = \int_0^\infty \frac{dt}{t^{n+\epsilon}}\left\{ e^{-t} - \sum_{l=0}^{n-1} \frac{(-t)^l}{l!} \right\}, \quad (9)$$

where, in contrast to (7), ϵ is restricted to $0 < \epsilon < 1$ to secure convergence. To obtain (7), expand $[\exp(-\alpha u) - \exp(-\beta u)]^n$ by the binomial theorem. Then to each term $\exp[-(n-k)\alpha u - k\beta u]$, add and subtract the series

$$-\sum_{l=0}^{n-1} (-u)^l [(n-k)\alpha + k\beta]^l/l!.$$

Equation (7) then follows from (9) with $t = [(n-k)\alpha + k\beta]u$ and the relation

$$\sum_{k=0}^n (-)^k \binom{n}{k} [(n-k)\alpha + k\beta]^l = 0, \quad (10)$$

where $l = 0, 1, \cdots, n-1$. Equation (10) expresses the fact that u^n is the lowest power of u in the power series for $[\exp(-\alpha u) - \exp(-\beta u)]^n$. After (7) has been established for $0 < \epsilon < 1$, it can be extended by analytic continuation to the values of ϵ needed in the application of (7). Other choices of α and β besides $-ia/n$ and $|a|/n$ can be made in (7) and (8). In some cases, it may be better to take $\alpha = -ia/n$ and

$\beta = \alpha + b$, where b is a positive constant at our disposal. For this choice of α and β, $[\exp(-\alpha u) - \exp(-\beta u)]^n$ becomes $\exp(iau)$ $\times [1 - \exp(-bu)]^n$ and $(n - k)\alpha + k\beta$ becomes $kb - ia$. By using these values in (7), equating imaginary parts, and letting $\epsilon \to 1$, we can obtain

$$\int_0^\infty u^{-n-1}(1 - e^{-bu})^n \sin au \, du$$

$$= \operatorname{Im} \frac{1}{n!} \sum_{k=0}^n (-)^{n-k} \binom{n}{k} (kb + ia)^n ln(kb + ia).$$

Here the integrand tends to $(\sin au)/u^{n+1}$ as $u \to \infty$.

IV. EXAMPLE ILLUSTRATING APPLICATION OF EQ. (8)

Consider the integral

$$I = \int_0^\infty (1 + u^2)^{-\frac{1}{2}} e^{iu} du, \tag{11}$$

which is equal to $K_0(1) + i\pi[I_0(1) - L_0(1)]/2$, where K_0 and I_0 are Bessel functions and L_0 is a Struve function (page 498 of Ref. 6). As $u \to \infty$, the integrand approaches

$$u^{-1} e^{iu} - \tfrac{1}{2} u^{-3} e^{iu} + \cdots.$$

Therefore, if we desire convergence like that of $u^{-5} e^{iu}$, we apply (8) twice. First with $n = 1$, $\alpha = -i$, $\beta = 1$, and then with $n = 3$, $\alpha = -i/3$, $\beta = \frac{1}{3}$:

$$I = \int_0^\infty du \left[(1 + u^2)^{-\frac{1}{2}} e^{iu} - \left(\frac{e^{iu} - e^{-u}}{u} \right) + \frac{1}{2} \left(\frac{e^{iu/3} - e^{-u/3}}{u} \right)^3 \right]$$

$$+ \frac{1}{0!} [-ln(-i) + ln(1)]$$

$$- \frac{1}{2} \frac{1}{2!} \sum_{k=0}^3 \binom{3}{k} (-)^{3-k} [(3 - k)(-i/3) + k/3]^2$$
$$\times ln[(3 - k)(-i/3) + k/3]. \tag{12}$$

The second line reduces to $i\pi/2$ [which, incidentally, gives the combination of (3) and (4)]. Although the last sum appears complicated, it can be readily evaluated by a digital computer when the program is written using complex variables.

The integrand in (12) now decreases as $u^{-5} e^{iu}$ when u is large, and the integral can be evaluated by any suitable method of numerical integration. Some computations were made using the trapezoidal rule as described in Section IX of Ref. 1. In this method, the variable of integration is changed from u to x, where $u = a \, ln[1 + \exp(x/a)]$, and the integral in x is evaluated by the trapezoidal rule. A computation using 52 points with $h = \Delta x = 4.0$ and $a = 6$ gave $I = 0.42102 \cdots + i0.87308 \cdots$, which agrees with tabulated values.

This example brings to mind the question, when does one cease subtracting terms from the integrand and start integrating? That is, what is the trade-off between computing trapezoidal sums and those sums in (7) and (8)? There seems to be no well-defined answer, but experience indicates that there is usually no need to make the integrand decrease faster than the inverse fifth or sixth power of the variable of integration.

Incidentally, the integral for I is quite readily evaluated by tilting the path of integration. Arbitrarily choosing a 45-degree tilt and setting $u = (1 + i)t$ carries (11) into

$$I = \int_0^\infty e^{-t} f(t) dt,$$

$$f(t) = (1 + i)(1 + 2it^2)^{-\frac{1}{4}} e^{it}. \tag{13}$$

The contribution of the 45-degree arc at $|u| = \infty$ is zero by Jordan's lemma (page 115 of Ref. 5). The integral is now in the form discussed in Section VI of Ref. 1, and can be efficiently evaluated by the trapezoidal rule after the change of variable $t = e^v$, $v = x - e^{-x}$. Computations made with this example suggest that when "tilting" can be used, it is preferable to "subtraction." However, tilting cannot always be used, an example being the integral $J(b, a)$ defined by (1).

V. EXAMPLE ILLUSTRATING THE USE OF EQ. (7)

Consider the integral

$$I = \int_0^\infty (1 + u^2)^{-5/6} e^{iu} du, \tag{14}$$

the real part of which is related to the Bessel function $K_{1/3}(1)$. When $u \to \infty$, the integrand becomes

$$u^{-5/3} e^{iu} + O(u^{-11/3}).$$

Therefore, to secure convergence of $O(u^{-11/3})$, we apply (7) once with $n + \epsilon = 5/3$. Since n must be a positive integer, possible choices of ϵ are $2/3$, $-1/3$, $-4/3$, $\cdots$. We also want $\epsilon < -1$. We arbitrarily choose $\epsilon = -4/3$ (and therefore $n = 3$) to make the integrand of (7) tend to 0 as $u^{4/3}$ when $u \to 0$. Because $n = 3$ we take $\alpha = -i/3$ and $\beta = 1/3$. Then

$$I = \int_0^\infty du [(1 + u^2)^{-5/6} e^{iu} - u^{-5/3} (e^{iu/3} - e^{-u/3})^3]$$

$$+ \frac{\Gamma(7/3)}{(-4/3)(-1/3)(2/3)} \sum_{k=0}^3 (-)^{3-k} \binom{3}{k}$$

$$\times [(3 - k)(-i/3) + (k/3)]^{2/3}. \tag{15}$$

This integral can be evaluated by the trapezoidal rule in much the same way as was the integral in (12).

VI. NUMERICAL EVALUATION OF THE INTEGRAL $J(b,a)$

To simplify the discussion of $J(b, a)$ defined by the integral (1), we assume that b is positive. The parameter a is real and $J(b, a)$ is an even function of a.

The integrand in $J(b, a)$ can be expanded as $\exp(iau)$ times

$$\sum_{n=2}^{\infty} (bu^{-1} \sin u)^n/n!,$$

and the term for $n = 2$ shows that the integral converges slowly. Inspection shows that the rate of convergence can be increased by writing $J(b, a)$ as an integral from 0 to ∞ and subtracting terms of the form $u^{-n} \exp(icu)$ as described in Sections III and IV. However, it is more convenient to subtract terms of the form $(u^{-1} \sin u)^n$ and use the known values of

$$B_n = \int_{-\infty}^{\infty} (bu^{-1} \sin u)^n \cos au \, du/n!, \tag{16}$$

where $n \geq 1$ is an integer and B_n is an even function of a.

When $n = 1$, B_1 is equal to $b\pi$ for $|a| < 1$, to $b\pi/2$ for $|a| = 1$, and to 0 for $|a| > 1$. When $n \geq 2$, B_n is a continuous function of a which is 0 when $|a| \geq n$. When $n \geq 2$ and $|a| \leq n$,

$$B_n = \frac{2\pi(b/2)^n}{n!(n-1)!} \sum_{m=0}^{M} (-)^m \binom{n}{m} (n \pm a - 2m)^{n-1}, \tag{17}$$

where M is the integer part of $(n \pm a)/2$. The choice of the sign in $\pm$ is arbitrary, but it must be the same in M and in (17). The minus sign is the more convenient choice when $a > 0$.

When we desire convergence at the rate of, say, $1/u^6$, we subtract and add $e^{-b}(B_2 + B_3 + B_4 + B_5)$:

$$J(b, a) = e^{-b} \int_{-\infty}^{\infty} \left[\exp(bu^{-1} \sin u) - 1 \right.$$
$$\left. - \sum_{n=1}^{5} (bu^{-1} \sin u)^n/n! \right] e^{iau} du + e^{-b}(B_2 + B_3 + B_4 + B_5). \tag{18}$$

Incidentally, $J(b, a)$ is equal to e^{-b} times the infinite sum $B_2 + B_3 + \cdots$, a result useful for computation when b is small.

The integrand in (18) is an analytic function of u, and the results of Ref. 1 suggest direct numerical integration by the trapezoidal rule. Several trial computations were made to obtain an idea of the performance of this method of integration. In the actual work, the integral

Table I — Values of $J(b,a)$ obtained by the trapezoidal rule

b	a	h	N	u_{max}	$J(b, a)$
1	1	0.7	12	7.7	0.413 5433
1	4	0.4	20	7.6	0.000 0428
4	1	0.5	36	17.5	1.341 1671
4	4	0.4	45	17.6	0.011 6253
16	1	0.3	39	11.4	0.997 3179
16	10	0.225	52	11.5	0.000 2046
32	1	0.25	19	4.5	0.736 6452
32	10	0.175	26	4.4	0.007 6251

was rewritten as an integral from 0 to ∞, and exp (iau) was replaced by cos au. Typical results are shown in Table I. Here N is the number of terms in the trapezoidal sum, and h is the spacing between successive values of u. The trapezoidal sum was truncated at $u = u_{max} = (N-1)h$. I believe the values shown for $J(b, a)$ are accurate to the number of places shown. A short table of $J(b, a)$ for $a \leq 1.25$ is given in Ref. 7, where $J(b, a)e^b$ is referred to as "Lewin's integral." When b is large, the work of Prabhu and Rowe[4] gives tight bounds on $J(b, a)$ for all values of a.

The truncation value u_{max} was chosen to be the value of u beyond which the absolute value of the integrand remains less than 10^{-8}. This makes the truncation error less than $10^{-8}u_{max}/5$. It can be shown that u_{max} has a maximum value of about 18 near $b = 6$.

The error E in the computed value of $J(b, a)$ resulting from the finite size of the spacing h, i.e., the "trapezoidal error," is of the order of $J(b, 2\pi h^{-1} - a)$ when h is small. This follows from Section VII below. The trapezoidal errors observed in the construction of Table I agree well with the values of $J(b, 2\pi h^{-1} - a)$ estimated from bounds for $J(b, a)$ given by Prabhu and Rowe.[4] E decreases rapidly as h decreases through the values shown in Table I.

VII. TRAPEZOIDAL ERROR FOR FOURIER INTEGRALS

Let the Fourier integral and the associated trapezoidal error be, respectively,

$$I(a) = \int_{-\infty}^{\infty} F(u)e^{iau}du, \tag{19}$$

$$E = h \sum_{n=-\infty}^{\infty} F(nh) \exp (ianh) - I(a). \tag{20}$$

As in Ref. 1, Poisson's summation formula gives

$$E = \left(\sum_{k=-\infty}^{-1} + \sum_{k=1}^{\infty} \right) I(2\pi k h^{-1} + a). \tag{21}$$

When $F(u)$ is an even function of u, $I(-a)$ is equal to $I(a)$, and (21) gives

$$E = \sum_{k=1}^{\infty} \left[I(2\pi k h^{-1} - a) + I(2\pi k h^{-1} + a) \right]. \tag{22}$$

It is interesting to note that, if $I(a)$ is known to be real and nonnegative for all real values of a, then E is nonnegative and the values of $I(a)$ obtained by evaluating (19) by the trapezoidal rule are always greater than or equal to the true value of $I(a)$, truncation errors aside. This fact is of use in dealing with power spectra and probability densities.

Now let $F(u)$ be analytic in a strip in the u-plane containing the real u-axis. When $F(u)$ is even and analytic, usually only the $k = 1$ term in (22) is important if h is small. In addition, if $a > 0$, E and $I(2\pi h^{-1} - a)$ are usually of the same order of magnitude.

VIII. DISPLACING THE PATH OF INTEGRATION OF A FOURIER INTEGRAL

In some cases, the numerical evaluation of the Fourier integral $I(a)$ defined by (19) can be facilitated by shifting the path of integration from the real u-axis to a parallel path that passes through $u = ic$, where c is some suitably chosen real constant. Changing the variable of integration in the integral (19) for $I(a)$ from u to x, where $u = x + ic$, and applying the trapezoidal rule to the integral in x gives the trapezoidal error

$$E' = h \sum_{n=-\infty}^{\infty} F(nh + ic)e^{ianh-ac} - I(a),$$

which, by Poisson's summation formula, goes into

$$E' = \left(\sum_{k=-\infty}^{-1} + \sum_{k=1}^{\infty} \right) I(2\pi k h^{-1} + a)e^{2\pi kc/h}. \tag{23}$$

When $I(a)$ is known to be real and nonnegative for all real values of a, then E' is nonnegative just as E is in (21), and the trapezoidal values of $I(a)$ obtained from the displaced path are greater than or possibly equal to the true value of $I(a)$, just as before.

The integral in the expression (18) for $J(b, a)$ is of the Fourier type (19), and some calculations were made to test the benefit obtained by shifting the path of integration. Along the path $u = x + ic$, the real part of the integrand is an even function of x, and the integral in (18) can be rewritten as an integral from $x = 0$ to ∞ of twice the real part. For $b = 32$ and $a = 10$, the value of c was chosen to be 0.87. This corresponds to the saddle point $u_0 = i0.87$ of $\exp (bu^{-1} \sin u + iau)$. A trapezoidal rule evaluation with $h = 0.3$ along the line $u = x + i0.87$ gave $J(b, a)$ to the accuracy shown in Table I. The number of terms

required was $N = 16$ compared with $N = 26$ when the integration was taken along the real u-axis (with $h = 0.175$).

REFERENCES

1. S. O. Rice, "Efficient Evaluation of Integrals of Analytic Functions by the Trapezoidal Rule," B.S.T.J., *52*, No. 5 (May–June 1973), pp. 707–722.
2. E. O. Tuck, "A Simple 'Filon-Trapezoidal' Rule," Math. of Computation, *21*, No. 98 (April 1967), pp. 239–241.
3. V. I. Krylov, *Approximate Calculation of Integrals*, trans. A. H. Stroud, New York: Macmillan, 1962.
4. V. K. Prabhu and H. E. Rowe, "Spectral Density Bounds of a PM Wave," B.S.T.J., *48*, No. 3 (March 1969), pp. 789–831.
5. E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, London: Cambridge University Press, 1927.
6. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, National Bureau of Standards, Applied Mathematics Series No. 55, Washington: Government Printing Office, 1964.
7. W. R. Bennett, H. E. Curtis, and S. O. Rice, "Interchannel Interference in FM and PM Systems Under Noise Loading Conditions," B.S.T.J., *34*, No. 3 (May 1955), pp. 601–636.

An Experiment on Propagation of 60-GHz Waves Through Rain

By O. E. DeLANGE, A. F. DIETRICH, and D. C. HOGG

(Manuscript received April 16, 1974)

An experiment was designed to determine the effect of rainfall upon the attenuation and polarization of millimeter waves. Vertically polarized waves at 60 GHz were transmitted over a 1-km path, and the energy received in both the vertical and horizontal polarizations, as affected by rain, was detected and recorded at the receiving site; similarly, a horizontally polarized wave was transmitted and the same components measured. Data were collected over a period of one year (1970) using an all-solid-state measuring set. It was found that attenuation due to rain exceeded 42 dB for 5 minutes, and that a system employing orthogonally polarized signals would have had to tolerate signal-to-crosstalk ratios as poor as 22 dB for that same period of time.

I. INTRODUCTION

Crowding of the frequency spectrum is forcing designers of microwave communication systems to use higher frequencies and to give serious consideration to orthogonally polarized waves for increasing communication capacity. If adequate orthogonality can be maintained, the capacity of a system is double that of a system employing only one polarization. However, it has long been known that imperfections in antennas and other system components limit the degree of orthogonality that can be maintained in a practical system. Rainfall on the transmission path can also affect polarization and thereby limit the degree of discrimination against the unwanted signals at the receiver. The experiment described here was set up to determine the extent of such depolarization as well as the attentuation of 60-GHz waves.

The mechanism of attenuation and depolarization involves the geometry of raindrops. A number of workers have determined that large raindrops are seldom spherical and that they are capable of producing different amounts of attenuation for waves having different polarizations.[1,2] At the time our experiment was in the design stage,

Semplak[3] had already observed, at a frequency of 30.9 GHz, that attenuation for horizontally polarized waves was significantly greater than for vertically polarized waves during rainstorms. In this paper, we report data at about twice that frequency.

II. DESCRIPTION OF THE EXPERIMENT

The propagation path at Holmdel, New Jersey, is about 1.03 km long. The transmitting antenna is located on flat terrain about 30 feet above the ground and the receiver is on the southern slope of Crawford Hill. Propagation is along a path approximately 15 degrees west of true north. The transmitted carrier frequency is 60 GHz, near the maximum of the oxygen absorption line.[4] A horn-reflector antenna with 51-dB gain is employed at the transmitter and a horn-lens antenna with 53-dB gain at the receiver. The net clear-weather path attenuation is 39 dB, which includes a 15-dB loss due to oxygen absorption.

A single transmitter and receiver are employed. The polarization of the transmitted wave is switched from horizontal to vertical 10 times per second and the signal is gated on and off at a 20-Hz rate. The receiver is switched to respond five times per second, first to a horizontally polarized wave, and then to a vertically polarized wave. As a result of the switching we provide four separate outputs corresponding to the following conditions of polarization:

(i) Transmitter and receiver both horizontal, designated T_hR_h.
(ii) Transmitter and receiver both vertical, designated T_vR_v.
(iii) Transmitter horizontal and receiver vertical, T_hR_v.
(iv) Transmitter vertical and receiver horizontal, T_vR_h. *

Output signals corresponding to these four conditions are recorded separately and constitute our data. The effects we measure are the differences in the fading suffered by the various components during rainstorms.

The pulse output of each gate is detected and integrated, applied to a logarithmic amplifier, amplified, and recorded. It was not difficult to obtain a truly logarithmic output over an 80-dB range of input voltage, but it was difficult to obtain a truly linear detector output. By using one detector for the direct components and a separate one for the cross-polarized components, we are able to accommodate 50-dB fades with the recorded output a truly logarithmic function of the input over the complete range.

---

* In this nomenclature, for example, a cross-polarization discrimination is T_vR_v/T_vR_h and a crosstalk ratio is T_vR_v/T_hR_v.

III. EQUIPMENT

The fact that we are dealing with a narrowband measuring set in an application that requires a dynamic range of about 80 dB puts a premium on frequency stability. A considerable part of our effort, therefore, was directed toward obtaining such stability.

3.1 Receiver

We chose to injection-lock the receiving local oscillator to a stable 30-GHz signal obtained by harmonic generation from a 150-MHz source to obtain both a high degree of frequency stability and a low-noise output. To provide means for injecting the locking signal, the oscillator cavity shown in Fig. 1 was constructed. The IMPATT diode is mounted at one end of the center conductor of this coaxial cavity which is one wavelength long at 60 GHz. A probe, one-quarter wavelength from the end, couples out 60-GHz energy into a section of the RG 98/U output waveguide. The 30-GHz energy is coupled from the RG 96/U injection waveguide into the cavity at a point one-half wavelength (at 60 GHz) from the end of the cavity. Since this is at a potential null for the higher frequency, coupling can be accomplished with minimum reaction to the 60-GHz signal. The 30- and 60-GHz waveguide sections are tuned by means of short-circuiting plungers adjacent to the coupling probes. From an inspection of Fig. 1 it is evident that oscillation and locking could take place at 30 GHz and produce a 60-GHz output due to the generation of second harmonic. Alternatively, 30-GHz oscillation could be suppressed by adjustment of the coupling probe, leaving the circuit free to oscillate at 60 GHz
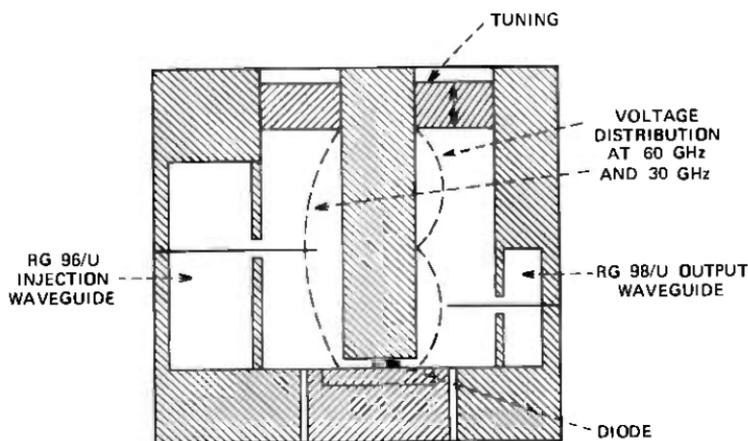
Fig. 1—60-GHz IMPATT oscillator. Diode is mounted at the end of the center conductor of the coaxial cavity.
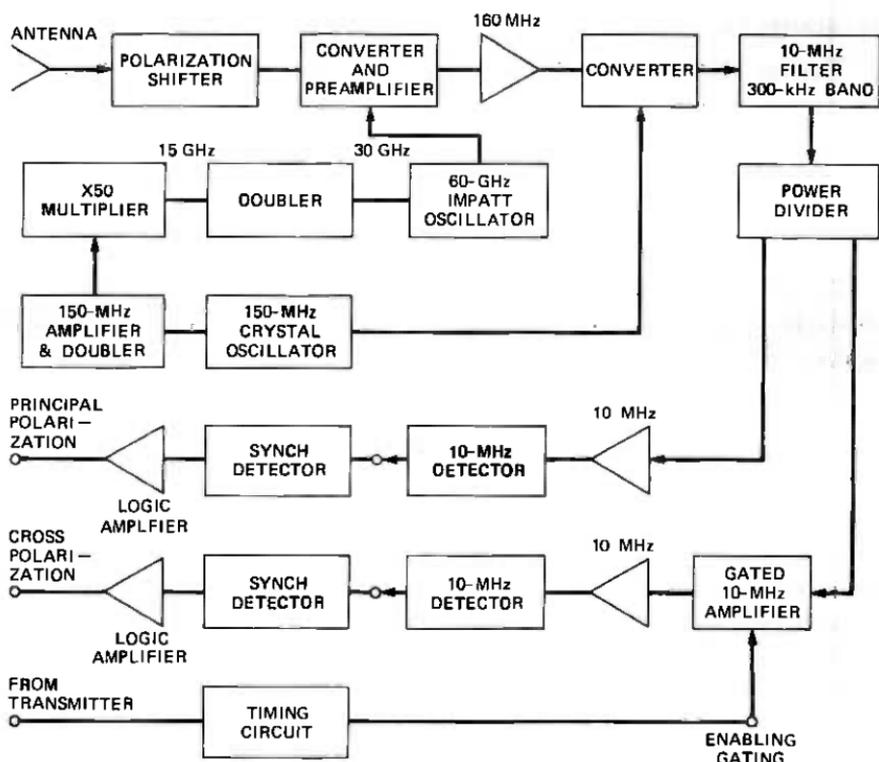
Fig. 2—60-GHz solid-state receiver with injection-locked local oscillator.

and lock on the second harmonic of the 30-GHz signal. The latter mode of operation proved most favorable.*

The IMPATT local oscillator in the receiver is locked to a signal generated from the output of the 150-MHz, crystal-controlled oscillator (see Fig. 2). A high-level 300-MHz voltage is applied to a step-recovery diode which has an output tuned to 15 GHz, thus providing a frequency multiplication of 50. The final locking signal, at a level of 0.2 mW, is obtained by doubling the 15-GHz signal in a diode harmonic generator. The locking range with this power level is known to be a few tens of MHz, which is sufficient to override cavity detuning due to temperature changes. Output power of the local oscillator of the receiver was about 4 mW.

For favorable adjustment, the FM noise output of the oscillator, when injection locked, was found to come largely from the injection

---

* In a preliminary experiment, 3 mW from a tunable Gunn oscillator was applied to the 30-GHz input to the oscillator, which delivered about 15 mW at 60 GHz. The IMPATT output frequency could be varied over a range of plus and minus a few hundred MHz by adjusting the frequency of the Gunn oscillator but leaving the IMPATT tuning fixed, resulting in a very satisfactory frequency lock.

signal. With a frequency multiplication of 400, this noise is 52 dB above that of the 150-MHz oscillator. However, the most frequently observed situation was one in which the IMPATT was noisy and the noise output was reduced by application of the locking signal. The cavity adjustments that resulted in stable, low-noise operation were obviously different from those that gave the widest frequency lock; the former condition was used in operation.

In the receiver the local-oscillator output is applied to a balanced converter where the incoming signal is translated to an intermediate frequency of 160 MHz. The converter employs wafer-type, silicon, point-contact rectifiers of the type described by Sharpless.[5] A transistor preamplifier with 10 dB of gain and a 2-dB noise figure is coupled directly to the converter. Although extensive measurements have not been made, there is evidence that local-oscillator noise is not limiting the performance of our system. The 160-MHz signal is mixed with 150 MHz in a second converter to provide a 10-MHz second intermediate frequency. After amplification, this IF signal is detected by linear detectors.

3.2 Transmitter

With only 0.2 mW of 30-GHz locking signal available, it was not possible to obtain the high output power needed from the transmitter and at the same time have a really satisfactory frequency lock. For this reason, the transmitter is controlled by an AFC circuit using a harmonically-generated 60-GHz signal as reference (see the block schematic in Fig. 3). With a feedback factor of nearly 1000, the transmitter output frequency is determined almost entirely by the crossover point of the 160-MHz discriminator and, therefore, is largely independent of changes in cavity tuning.

3.3 Performance

Before the system was installed in the field, the transmitter output was coupled to the receiver and the resultant 160-MHz IF signal monitored with a frequency counter. The frequency stayed within a few tens of kilohertz over a period of many days.

In the low-frequency section of the receiver, a logarithmic amplifier of the type described by Paterson,[6] simple and easily adjusted, provided a truly logarithmic output over a range of 80 dB. Obtaining a linear detector to drive this amplifier was more of a problem, but by amplifying the 10-MHz signal to a level of about 50 volts at the detector, a linear output was obtained over an input range of more than 50 dB. The required 80-dB dynamic range was obtained by dividing the incoming signal into two branches, one to recover low-
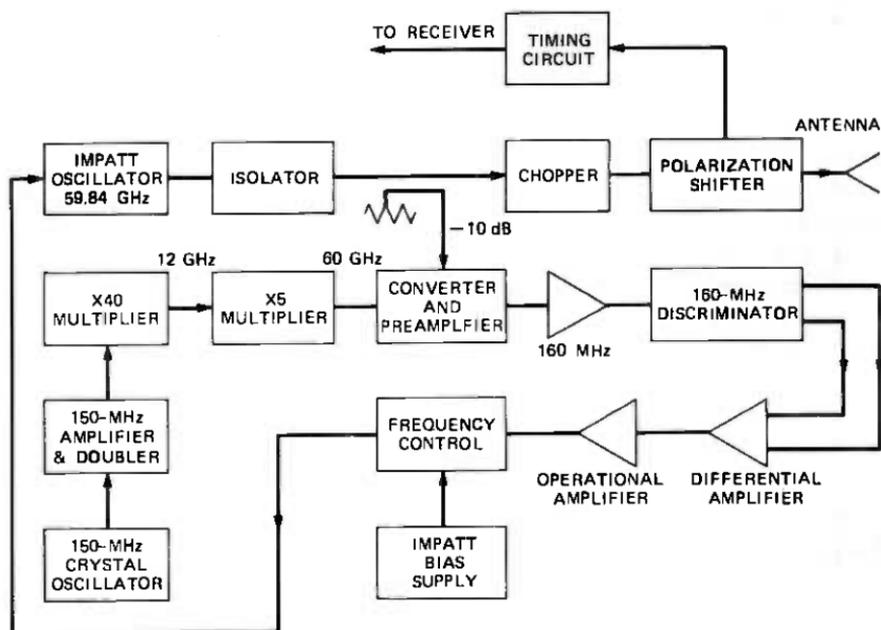
Fig. 3—60-GHz solid-state transmitter with frequency-controlled oscillator.
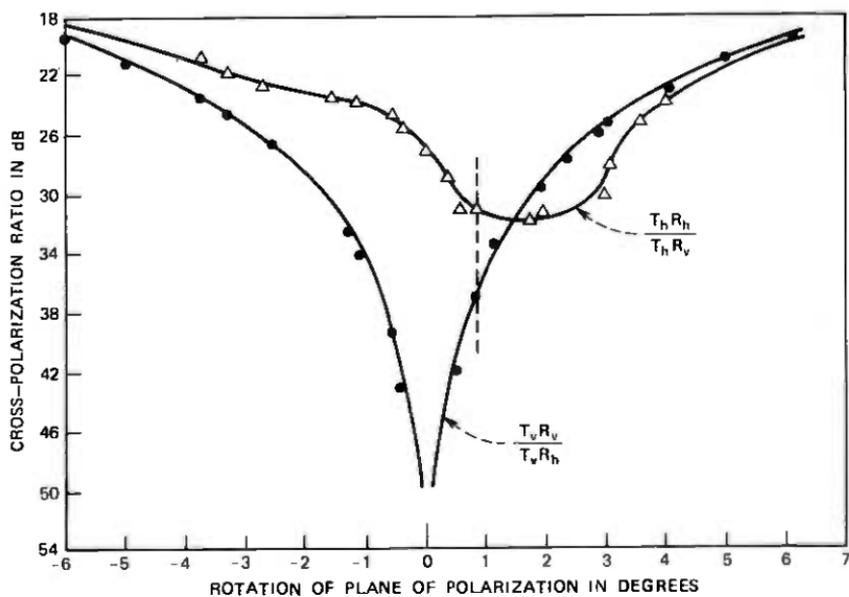
Fig. 4—Polarization discrimination characteristics of the operating system
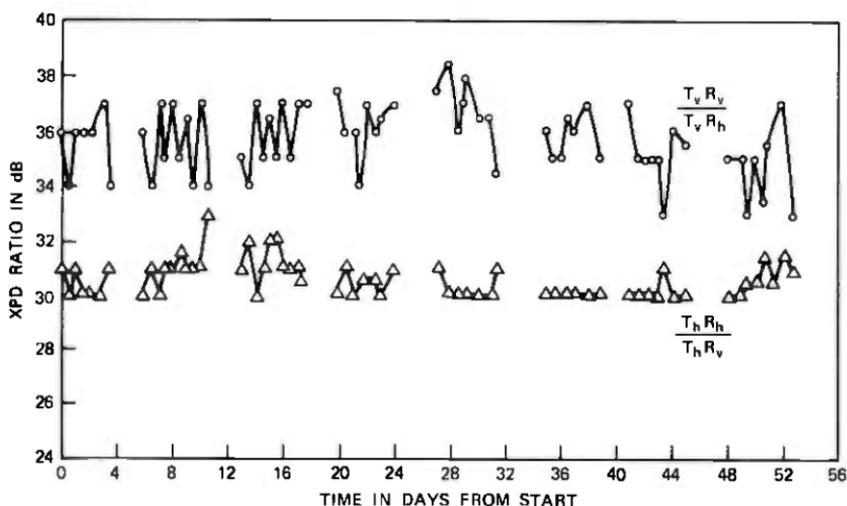
Fig. 5—Polarization discrimination with respect to stability. Not plotted for weekends.

level cross-polarization signals and the other to process the high-level signals.

3.4 Polarization discrimination and stability

After preliminary adjustments of the system on the path, a polarization shifter was placed in the receiver input to rotate the plane of polarization of the incoming waves and thereby determine the polarization discrimination of the system. The resultant outputs, plotted on Fig. 4, show that an excellent balance could be obtained for the condition $T_v R_v / T_v R_h$ but, as a result of system imperfections in the antennas and other system components, the maximum discrimination for the condition $T_h R_h / T_h R_v$ was only 32 dB, which indicates that slight elliptical polarization was present in the latter case. The system is normally adjusted to operate near the minimum of the horizontally cross-polarized component (see dashed line in Fig. 4), which at the same time provides more than 30 dB of discrimination against the vertically polarized component. Due to temperature or voltage changes, the operating point varied somewhat with time, as shown in Fig. 5. The gaps in the plotted data result from readings of stability not being taken during weekends.* Although no particular steps were taken toward stabilization, the discrimination ratio remained better than 30 dB over long periods of time without readjustment.

---

* This remark applies only to stability readings; all other data were accumulated continuously.

IV. EXPERIMENTAL RESULTS

4.1 Distribution of fades

Although the primary purpose of the experiment was to obtain data concerning polarization, it also yielded one year of information on the distribution of rain fades. Figure 6 is a plot showing depth of fade versus the fraction of time this fading level was exceeded. We see, for this year, that for 60-GHz transmission over this 1-km path, a rain-fading margin of 40 to 42 dB would have been needed to meet a reliability objective of not more than 5 minutes of outage time (a probability of 10^{-5}). These data are found to be in good agreement with attenuations predicted from point rainfall rates.[7]

4.2 Polarization crosstalk distribution

The experimental results verify expectations that rainfall does degrade the polarization of 60-GHz waves, although much less than at lower frequencies.[2] The effect for this path is plotted in Fig. 7. The solid dots indicate crosstalk into the vertical channel ($T_v R_v / T_h R_v$), and the open circles indicate crosstalk into the horizontal channel ($T_h R_h / T_v R_h$). It is evident that, for our system, the polarization crosstalk ratio is 22 dB or more for 5 minutes during the year 1970.

Figure 8 shows data for a heavy rainstorm of July 31, 1970. This storm behaved as expected in the sense that all quantities varied in
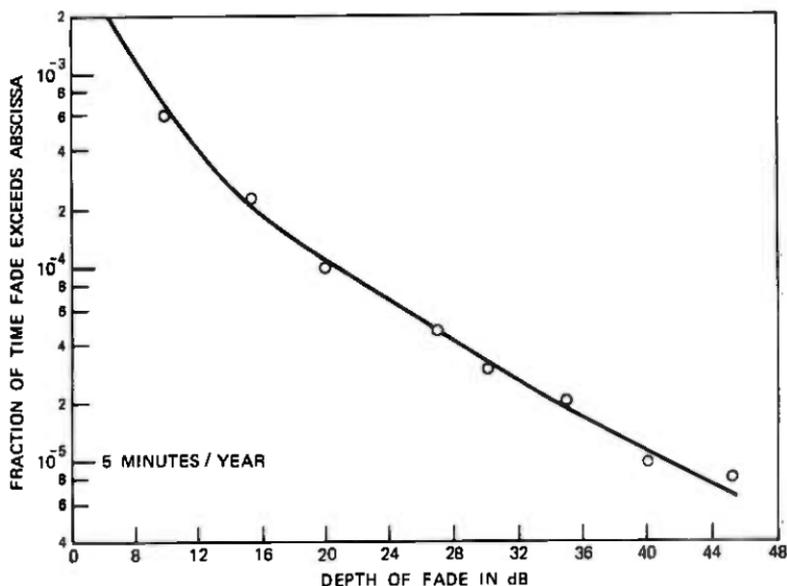

Fig. 6—Distribution of rain attenuation (1970).
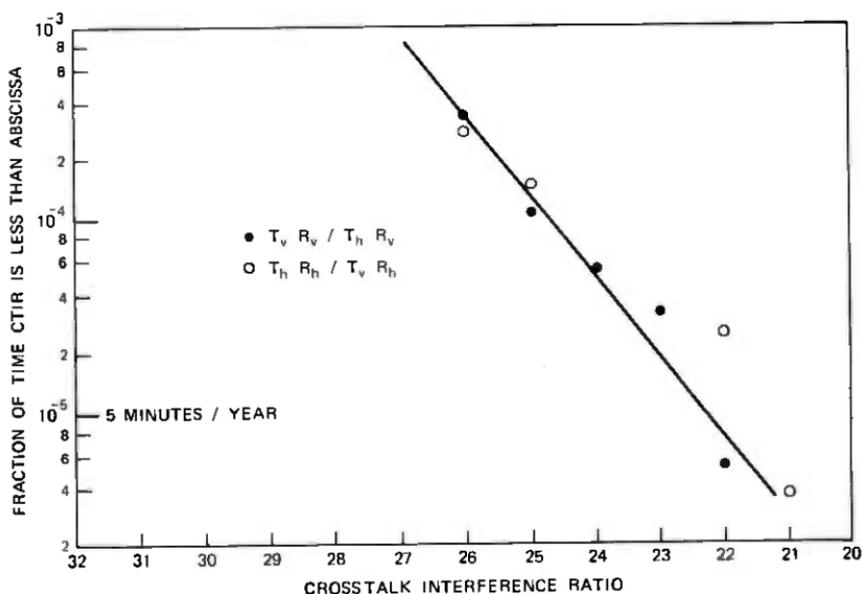
Fig. 7—Distribution of the crosstalk ratio.

the way one might expect from simple theory; i.e., differential attenuation was always positive (fade in horizontal polarization was greater than in vertical). The crosstalk variations were pretty much the same for both polarizations, with the ratio becoming poorer during the deeper fades.

On the other hand, the storm of November 4, 1970, plotted in Fig. 9, did not produce results expected from the simple theory. For a considerable portion of the time the differential attenuation was negative, indicating that attenuation of the vertical component was *greater than* that of the horizontal component. The fact that the crosstalk ratio in both channels (Fig. 9) improved slightly in this case may be explained by referring to Fig. 4, which shows the clear-weather operating point near +0.5 degree; a negative rotation (caused by the rain) of the vertical component from this value would reduce the cross-coupled energy and thereby improve the ratio. Similarly, a positive rotation of the horizontal component would improve the crosstalk ratio in that channel.

One expects differential attenuation at 60 GHz to increase with the absolute value of attenuation (much less than at lower frequencies).[2] This appears to be true on the average, as seen in Fig. 10. For fading levels greater than 25 dB, the differential appears to remain constant, but the limited amount of data available for very large fades makes de-
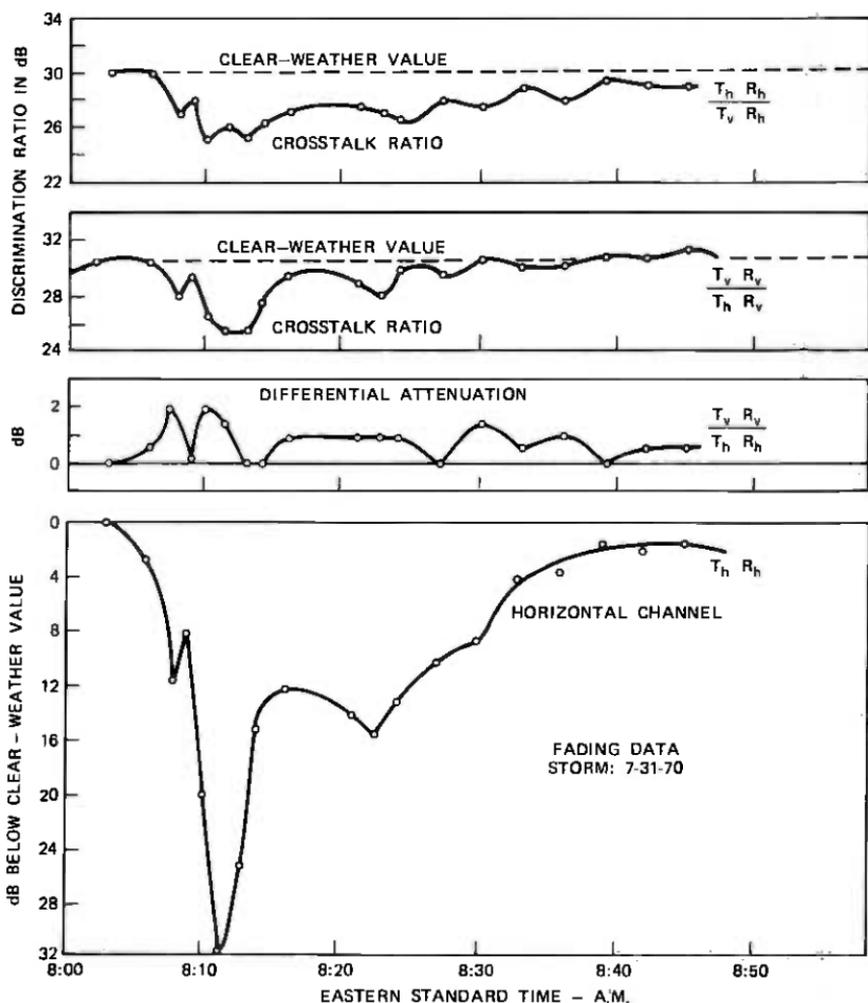
Fig. 8—Behavior of the various polarization components for the shower of July 31, 1970.

termination of the trend uncertain. The data points are very scattered, most likely due to variations of factors such as size, shape, and tilt of the raindrops involved. A number of negative differential attenuations are found for fades up to about 17 dB, but none above that value. The solid curve in Fig. 10 is drawn through the average value of the differential. A more significant quantity might be the maximum values, shown by the dashed curve; the maxima seldom exceed 2 dB, which is much less than the 6 dB observed[3] at 30 GHz for the same magnitude of fading.
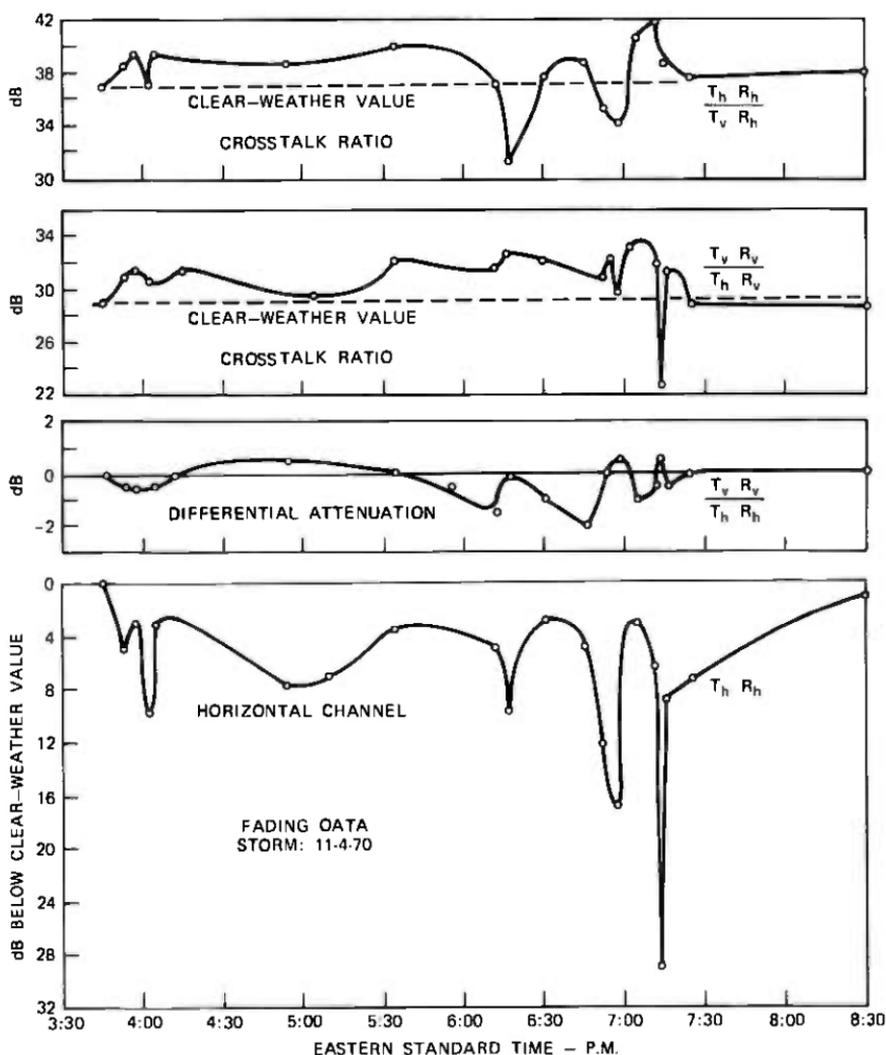
Fig. 9—Behavior of the various polarization components for the shower of November 4, 1970.

V. CONCLUSIONS

Consideration of propagation measured over a period of one year on a 1-km path in New Jersey, at a frequency of 60 GHz, leads to the following conclusions concerning this path and this frequency:

(i) Aside from polarization effects, a rain-fading margin of 40 to 42 dB would have been needed to meet a reliability objective of less than 5 minutes outage per year on a 1-km path.
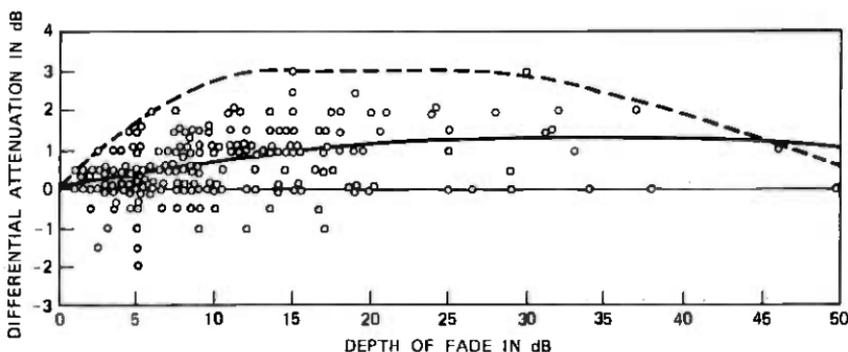
Fig. 10—Plot of measured differential attenuation as a function of fade depth.

   (*ii*) The 60-GHz attenuation produced by rain is usually, but not always, greater for horizontally polarized than for vertically polarized waves; the differential is seldom greater than 2 dB and the average differential is only 1.25 dB, even for fades greater than 30 dB.

   (*iii*) Rainstorms affect the polarization of 60-GHz waves transmitted through them, although not nearly as seriously for a given fade depth as at lower frequencies. A system accommodating a crosstalk interference ratio of 22 dB would operate satisfactorily over this path, suffering no additional outage over that resulting from attenuation.

VI. ACKNOWLEDGMENT

REFERENCES

1. T. Oguchi, "Attention of Electromagnetic Waves Due to Rain with Distorted Raindrops," J. Rad. Res. Labs (Tokyo), Part 1, *7*, No. 33 (1960), pp. 467–485; Part 2, *11*, pp. 19–44.
2. T. S. Chu, "Rain-Induced Cross-Polarization at Centimeter and Millimeter Wavelengths" B.S.T.J., *53*, No. 8 (October 1974), pp. 1557–1579.
3. R. A. Semplak, "Effect of Oblate Raindrops on Attenuation at 30.9 GHz," Radio Science, *5*, No. 3 (March 1970), pp. 559–564.
4. A. B. Crawford and D. C. Hogg, "Measurement of Atmospheric Attenuation at Millimeter Wavelengths," B.S.T.J., *35*, No. 4 (July 1956), pp. 907–916.
5. W. M. Sharpless, "Wafer-Type Millimeter Wave Rectifiers," B.S.T.J., *35*, No. 6 (November 1956), pp. 1385–1402.
6. W. L. Paterson, "Multiplication and Logarithmic Conversion by Operational Amplifier-Transistor Circuits," Rev. Sci. Instr., *34*, No. 12 (December 1963), pp. 1311–1316.
7. W. F. Bodtmann and C. L. Ruthroff, "Rain Attenuation on Radio Paths: Theory, Experiment, and Design," B.S.T.J., *53*, No. 7 (September 1974), pp. 1329–1349.

Rain-Scatter Interference in Terrestrial Microwave Systems

By K. BULLINGTON

The potential interference caused by scattering from raindrops is viewed as foreground reflections that "spoil" the off-beam antenna pattern. This approach greatly simplifies interference calculations and is sufficiently accurate for most engineering purposes.

The principal uncertainty that limits not only this simplified approach but also the "exact" solution is the inability to characterize the variable rainfall distribution along any given radio path by a single parameter, such as the rainfall rate, either at a point or as an average along the path.

I. INTRODUCTION

The potential cochannel interference between two stations, illustrated on Fig. 1a, is usually assumed to occur along the shortest path (dotted line) between stations. The transmission loss along the dotted path includes not only the free-space loss, but also any additional loss caused by obstructions or by site shielding. Transmission is further modified by the antenna patterns in the direction toward the other site. In addition, reflections from nearby trees and terrain may reduce the expected suppression of the antenna in directions away from the main beam.

If the two antenna beams intercept, any reflection from rain, aircraft, or birds in or near their common volume can provide another path by which interference can occur. While the magnitude of such reflections is normally small, this "sneak" path has free-space transmission with full antenna gain at one end and, thus, bypasses the suppression that is normally provided by the antenna patterns and shielding losses.

II. APPROACH

A simplified procedure for estimating the transmission loss by way of reflection from rain can be obtained, in most cases, by replacing Fig. 1a with Fig. 1b. In this case, antenna G_1 is replaced by a phantom
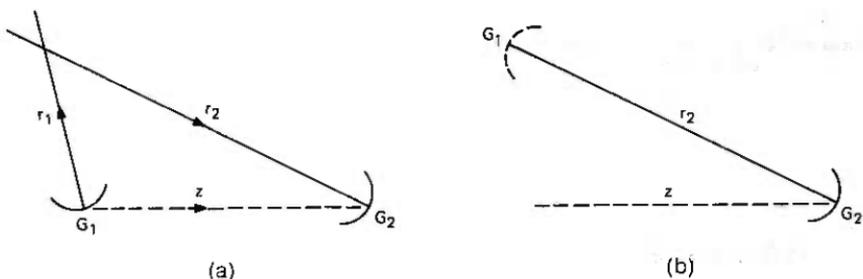
Fig. 1—(a) Plan view of two cochannel stations. (b) Interference path from common volume.

antenna, G_1', whose equivalent gain relative to an isotropic antenna is given on Fig. 2 for various frequencies and rates of rainfall. The phantom antenna, G_1', can be considered to be located at the beam intersection, which is at a distance r_2 from G_2.

In other words, the effect of rain is to fill in the nulls and reduce the normal side- and back-lobe suppression of the antenna, and then to raise it to within line of sight of the distant station. Even when all of the energy from the transmitter is scattered, G_1' cannot exceed the unity "gain" of an isotropic antenna. This means that for comparable antennas, distances, and transmitter power, the unwanted scattered signals will be less than the desired signal by the gain of the antenna nearest the heavy rainstorm.
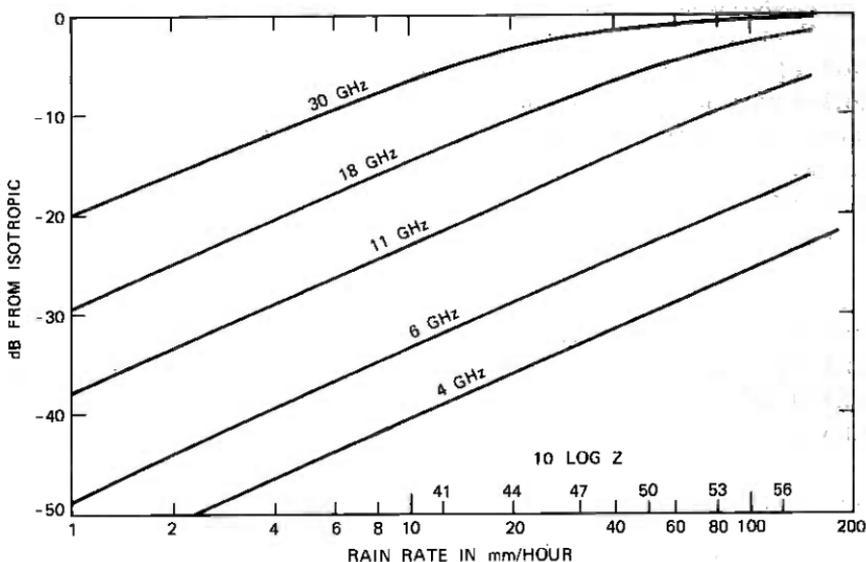
Fig. 2—Effective antenna gain G_1' during rain.

If the two antenna beams do not intersect, the antenna gain G_2 is reduced in accordance with its directivity pattern for the off-beam angle between its axis and the direction needed for full intersection.

The use of Fig. 2 requires an estimate of the rainfall rate R to be expected for a given percentage of the time. Typical values in current use that are expected to be exceeded for no more than 10^{-4} of the time are given in the following table.[1] For a probability of 10^{-5}, the above values of $10 \log Z$ should be increased by about 6 dB.

| Section of USA	R	$10 \log Z^*$
Southeastern and Gulf	80 mm/hr	53 dB
East and Midwest	56	51
Pacific region	33	47.5
Rocky Mountain region	13	41

* Reflectivity factor $Z \approx 10^6 \ (R/50)^{1.5}$.

While Fig. 2 provides a first-order solution that is sufficiently accurate for most engineering purposes, the complete rain-scattering problem is, of course, much more complex. It is necessary to consider the more general problem so that the quantitative significance of the many secondary factors can be evaluated and included whenever it is necessary to do so. Several solutions to the more complete scattering are given in Refs. 2 through 4. In spite of the variety of methods, symbols, and units, all of the papers referenced lead to essentially the same quantitative results for the same set of assumptions, because all use the same meteorological premises. The seeming complexity and diversity result in part from the three-dimensional geometry and in part from the inclusion of several factors that have been neglected in the first-order solution in Fig. 2 because their quantitative effects are small.

Atmospheric absorption is omitted from the first-order engineering solution because it is an unnecessary complication at 4 and 6 GHz; this loss can be included, if needed, at the end of the computation. To some extent, the omission of attenuation caused by rain is compensated for by the asymptotic behavior of the curves shown on Fig. 2. These curves include a factor to ensure that $G_1' \leqq 1$. More complicated methods (that theoretically are more accurate) add the expected absorption loss and then offset it in part by a gain due to "forward scatter." The difference between the two procedures is expected to be less than the effect of the uncertainty in the effective rain rate for a specified location and percentage of the time.

The polarization-mismatch factor is an additional loss of about $20 \log \cos \zeta$ dB, where ζ is the angle between the polarization of the

desired and undesired signal. It is a safety factor that can ordinarily be neglected, because the accuracy of the other factors in the overall computation seldom warrants this precision. Some methods add 3- or 4-dB loss for this factor, but any assumption of a significantly larger amount (say, 20 to 25 dB because the interfering polarizations are expected to be orthogonal) seems unrealistic because the normal cross-polarization discrimination can be significantly reduced during rain.

In addition, some references include a dielectric factor $[(\epsilon - 1)/(\epsilon + 2)]^2$. This factor is negligible for rain but is automatically included in the factor Z used in Fig. 2.

Example

When the two stations illustrated in Fig. 1a are separated by $z = 30$ miles (50 km), the free-space loss between isotropic antennas is 138 dB at 4 GHz. The net transmission loss along the z direction is $138 - G_{1z} - G_{2z} + S$, where S is the shadow and shielding loss in dB and the antenna gains G_{1z} and G_{2z} (along the z direction) are also expressed in decibels. On the other hand, the coupling loss from intersecting beams illustrated in Fig. 1 is $138 - G_1' - G_2 + 20 \log r_2/z$, where G_1' is the equivalent antenna gain shown on Fig. 2 for the station nearest the rainstorm and G_2 is the antenna gain along r_2 of the more distant station. With 40-dB antennas at 4 GHz, $z = 50$ km, $r_2 = 62$ km, and a rainfall rate of $R = 100$ mm/hour, the coupling loss by way of rain scatter is $138 - (-25) - 40 + 2 = 125$ dB.

In calculating the data in Fig. 2 it was assumed that the area of maximum rainfall has an effective diameter $D = 1$ km; if the user prefers some other value, say, D' in km, the rain rate to be used in entering Fig. 2 should be $R' = RD'$. The model also assumes $(G_1/r_1^2) \geqq (G_2/r_2^2)$, but this is no restriction because of reciprocity.

III. DISCUSSION OF THE SCATTER PROBLEM

The remaining sections review the scattering equation that leads to the results given in Fig. 2. In addition to the rain-scatter problem, the general method given below also derives the radar equation, the expression for passive repeaters, and the equation for rough or smooth ground reflections. The free-space transmission loss for a path length r and a wavelength λ is given by

$$\frac{P_R}{P_T} = \frac{G_1 A_2}{4\pi r^2} = \frac{G_1 G_2 \lambda^2}{(4\pi r)^2} = \frac{A_1 A_2}{(\lambda r)^2} , \tag{1}$$

where G is the antenna gain and A is the area of the antenna aperture. The transmission loss P_R/P_T by way of reflection from a large plane surface, illustrated in Fig. 3, is the product of two free-space paths.
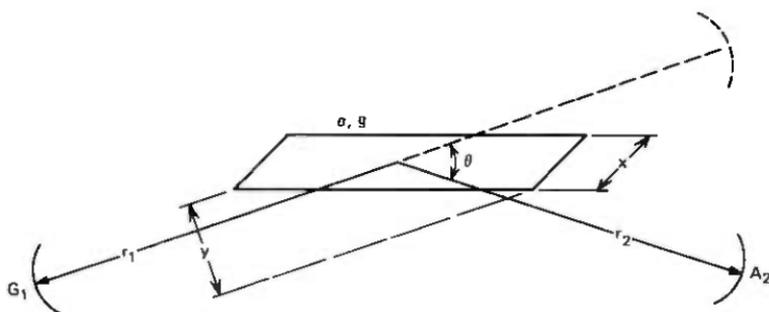
Fig. 3—Reflection from passive repeater.

The first section has an effective receiving area (cross section) denoted by σ and the second section has an effective transmitting antenna gain g that is related to both the cross section σ and the degree of both smoothness and reflectivity of the plane reflector. By algebraic rearrangement, the expression is separated into the free-space loss to the image of the distant antenna and a power reflection coefficient:

$$\frac{P_R}{P_T} = \left(\frac{G_1\sigma}{4\pi r_1^2}\right)\left(\frac{gA_2}{4\pi r_2^2}\right) = \overset{\substack{\text{Free space}\\\text{to image}\\\text{antenna}}}{\left(\frac{G_1 A_2}{4\pi(r_1 + r_2)^2}\right)}\overset{\substack{\text{Reflected}\\\text{power}\\\text{coefficient}}}{\left(\frac{\sigma g}{4\pi d_0^2}\right)}, \qquad (2)$$

where

$$\sigma = \frac{xy}{\left(1 + \frac{x}{\sqrt{\lambda d_0}}\right)\left(1 + \frac{y}{\sqrt{\lambda d_0}}\right)}; \qquad d_0 = \frac{r_1 r_2}{r_1 + r_2},$$

$$g = \sqrt{1 + \left(\frac{4\pi\sigma K}{\lambda^2}\right)^2},$$

$K = 1$ for perfectly smooth surface (mirror)
$\approx \exp(-\phi)$ for irregular surfaces
$\rightarrow 0$ for very rough surface,

and

ϕ = standard deviation of phase variations caused by irregular surface.

By the simplifying assumptions that ordinarily apply to radar ($G_1 = G_2 = 4\pi A_2/\lambda^2$; $r_1 = r_2 = r$; $g = 1$; $K \ll 1$), eq. (2) reduces to the radar equation

$$\frac{P_R}{P_T} = \frac{G^2\lambda^2\sigma}{(4\pi)^3 r^4}. \qquad (3)$$

RAIN SCATTER IN MICROWAVE SYSTEMS   **181**

When both dimensions of the plane reflector in Fig. 3 are large compared with the first Fresnel zone ($\sqrt{\lambda d_0}$), the cross section $\sigma = \lambda d_0$. For this case, the transmission loss is the free-space loss along the image path multiplied by a power-reflection coefficient:

$$\frac{P_R}{P_T} = \left(\frac{G_1 A_2}{4\pi(r_1 + r_2)^2} \right)\left(\frac{\sigma^2 K}{\lambda^2 d_0^2} \right) = \left(\frac{G_1 A_2}{4\pi(r_1 + r_2)^2} \right)(K). \qquad (4a)$$

When the y dimension of the plane reflector is less than the first Fresnel zone, $\sigma = y\sqrt{\lambda d_0}$, the corresponding transmission loss is

$$\frac{P_R}{P_T} = \left(\frac{G_1 A_2}{4\pi(r_1 + r_2)^2} \right)\left(\frac{y^2}{\lambda d_0} \right)(K). \qquad (4b)$$

Finally, when both dimensions of the reflector are small compared with the first Fresnel zone, σ approaches the geometrical cross section of the reflector, and the corresponding transmission loss is

$$\frac{P_R}{P_T} = \left(\frac{G_1 A_2}{4\pi(r_1 + r_2)^2} \right)\left(\frac{xy}{\lambda d_0} \right)^2(K). \qquad (4c)$$

In all of these cases, the first bracketed term is the free-space loss along the image path.

When the cross-sectional area σ of a single reflector is small compared with $\lambda^2/4\pi$ (or its surface is so rough that $K \to 0$), the energy is scattered in all directions. The antenna gain of the scatterer [see (2)] becomes $g = 1$, and this leads to

$$\frac{P_R}{P_T} = \left(\frac{A_2}{4\pi r_2^2} \right)\left(\frac{G_1 \sigma}{4\pi r_1^2} \right) = \left(\frac{A_2}{4\pi r_2^2} \right)\left(\frac{\sigma}{A_1'} \right), \qquad (5)$$

where

$$\frac{G_1}{4\pi} = \frac{1}{\beta_{1V}\beta_{1H}}$$

A_1' = cross-sectional area of beam from antenna 1 at distance r_1

and

$\beta_{V,H}$ = antenna beam widths in radians (assumed uniform).

The assumption of isotropic scattering is not strictly correct, but Setzer[5] has shown that, for frequencies below 30 GHz, the difference between isotropic scattering and the exact value is less than 3 dB at all angles.

The next step toward solution of the rain-scatter problem is to assume a large number (NV) of small particles, each with a cross section $\sigma \ll \lambda^2/4\pi$, randomly spaced with an approximately uniform

$$V = WXY \sin\theta = r_1{}^2\,\beta_{1V}\,\beta_{1H}\,\frac{r_2\,\beta_{2V}}{\sin\theta}$$

$$\frac{V}{A} = \frac{r_2\,\beta_{2V}}{\sin\theta} = D$$
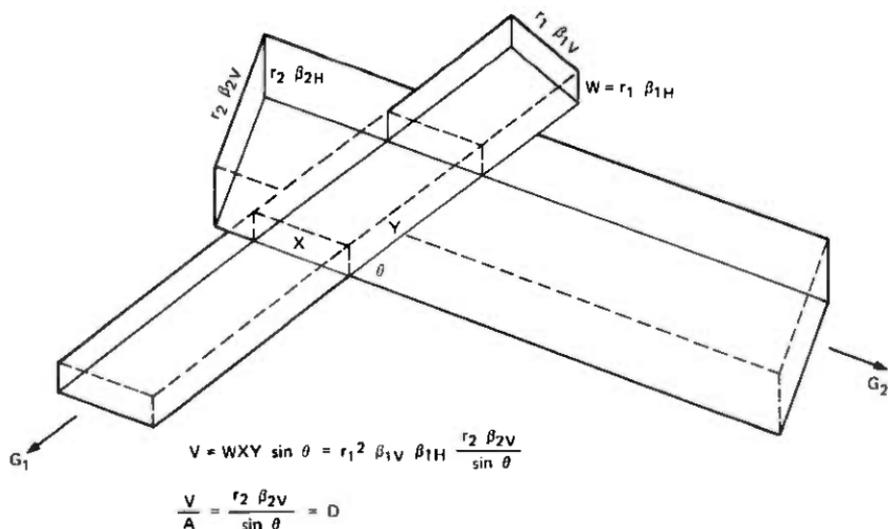
Fig. 4—Geometry of common volume.

density throughout the common volume illustrated in Fig. 4. The resulting transmission loss is

$$\frac{P_R}{P_T} = \left(\frac{A_2}{4\pi r_2^2}\right)\left(\frac{\sigma}{A'}\right)(VN) = \left(\frac{A_2}{4\pi r_2^2}\right)(\sigma N)\left(\frac{D}{1+\sigma ND}\right)$$
$$= \left(\frac{A_2}{4\pi r_2^2}\right)(G_1'), \quad (6)$$

where

N = number of particles per cubic meter,
V = common volume $A'D$,

and

$$D \leqq \frac{r_2\beta_2}{\sin\theta}$$

$$\leqq \left(\begin{array}{c}\text{size of rain cell with}\\ \text{maximum ``uniform'' rain rate}\end{array}\right).$$

The first bracket in (6) is the free-space transmission loss for a distance r_2 between an isotropic antenna and an antenna whose gain is $G_2 = 4\pi A_2/\lambda^2$. The term σN is a meteorological parameter (total cross section per cubic meter) that depends on the rate of rainfall and the wavelength. Essentially, the same assumptions on the distribution of raindrop diameters and on their velocity of fall are used by all the

references with the result that

$$\sigma N = \frac{\pi^5 10^{-18} Z}{\lambda^4} \text{ meters}^{-1},$$ (7)

Z = reflectivity factor,
  = $180\ R^{1.6}$ from Ref. 2,
  = $400\ R^{1.4}$ from table in Ref. 3,
  = $200\ R^{1.6}$ from Ref. 4,

and

R = rainfall rate in mm/hour.

Since the primary interest is in heavy rains ($R > 50$ mm/hr), this work uses

$$Z = 10^5 \left(\frac{R}{50} \right)^{1.5},$$ (8)

which is in good agreement with all references in the vicinity of $R = 50$ mm/hour, and splits the difference for much smaller and much larger values of R.

The equivalent antenna gain given in Fig. 2 is based on $G' = \sigma ND/(1 + \sigma ND)$ as defined in (6). The parameter D represents a somewhat uncertain (or "rubbery") length which, perhaps, can best be defined as the smaller of either V/A' or the diameter of the rain cell associated with the peak rain rate. One of the references uses $D = 4$ km in calculating its principal quantitative result, another uses $D = 1$ km, and a third uses 0.71 km. In this work, a value of $D = 1$ km is assumed in the results shown on Fig. 2.

Finally, the factor $1/(1 + \sigma ND)$ has been included to prevent a calculated value of G' greater than unity, because such a result would be inconsistent with the basic assumption of isotropic scattering.

IV. RAIN ATTENUATION

As previously mentioned, the first-order solution (using Fig. 2) considers only the scattering and omits the attenuation caused by rain. The attenuation computed by Setzer is shown by the solid lines on Fig. 5 and agrees with values given previously by Medhurst.[6,7] The dotted lines are calculated from the following empirical formula, which may be useful for engineering purposes:

$$\alpha = \frac{0.3R}{\left[1 + \left(\frac{\lambda}{2} \right)^2 \right]^2} \text{ dB/km},$$ (9)

where

λ = wavelength in cm.
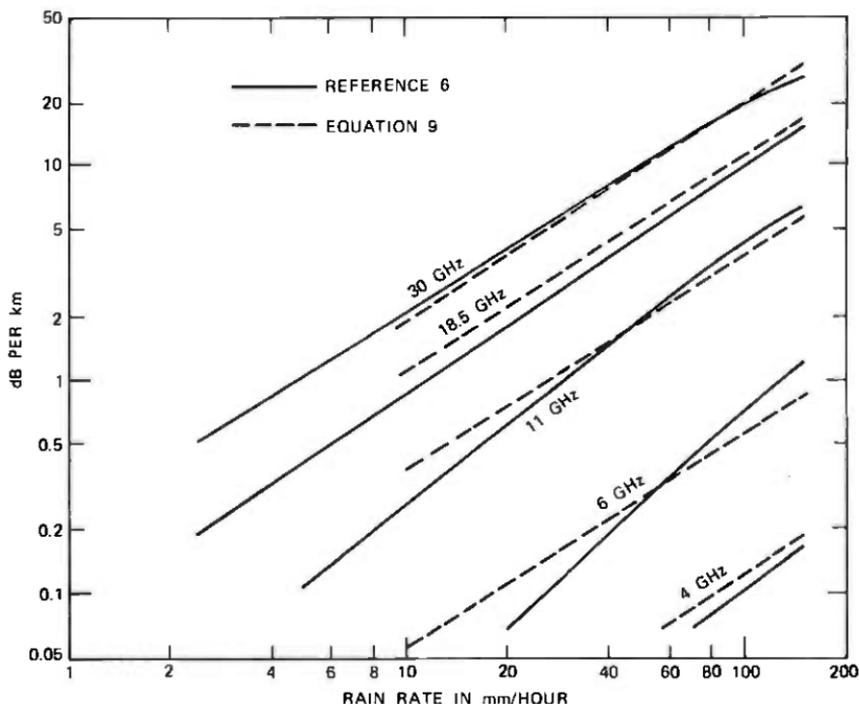
Fig. 5—Rain attenuation.

As in the scattering problem, there is some uncertainty regarding the effective distance to be used with the data in Fig. 5. The distance to be used in computing attenuation is greater than the diameter of a rain cell associated with the maximum rate of rainfall but cannot exceed the path length. Many authors prefer to think in terms of the *average* rate of rainfall over the actual path length, but this merely transfers the uncertainty from an effective distance to the determination of the average R. Based primarily on measurements in New Jersey, Hogg has related the average rainfall rate along a path to the rainfall rate at a point in the same climate.[8] Typical values to be exceeded for less than 10^{-5} of the time are given for path lengths in the 1- to 10-km range and extended to other climates. The results can be represented approximately by the following empirical formula:

$$\frac{\bar{R}}{R} = \left(1 + \frac{Rd}{750}\right)^{-\frac{1}{3}}, \qquad (10)$$

where

$\bar{R}$ = average rainfall rate along path d (in km).

In principle, the rain attenuation can be computed if the number and size of drops are known at every point in the space between the two antennas at every instant in time. In practice, significant variations with time and location are likely because the distribution of the rainstorm is completely independent of the radio path.

Rain attenuation estimates are necessary in calculating the fade margin against set noise, but only the difference in attenuation between signal and interference affects the carrier-to-interference ratio. When the desired signal (C) and the interfering signal (I) are attenuated equally, the C/I ratio is unchanged and the interference-to-noise (I/N) ratio is less than in the absence of attenuation. When the desired signal is attenuated more than the interfering signal (as, for example, by heavy rain at a location to the left of the intersection illustrated on Fig. 1a) both the C/I and C/N are decreased and either one may be controlling. Estimates of outage time caused by rain attenuation (low C/N) can be obtained without reference to rain scatter or other types of interfering signals. Conversely, estimates of the probability that a rain-scatter interference will exceed the set noise (or some lower limit required by the difference in interfering effects between thermal noise and a single frequency tone) can be made independent of the rain attenuation. For an efficient microwave transmission system, it is essential that the probability of service degradation caused by interference should be less than for set noise alone. Any more sophisticated attempt to combine the two separate probabilities of outage requires more information on the distribution of rainfall with geography and time than is available.

V. CONCLUSION

The interference caused by scatter from raindrops is a complex problem, as indicated in the references. Some complexity is unavoidable when all pertinent factors are to be included regardless of their magnitude. Fortunately, a relatively simple engineering method for interference calculation can be obtained by neglecting several factors whose magnitudes are small compared with the fundamental uncertainty in the exact rainfall distribution along the path.

REFERENCES

1. FCC Report No. R-7303, "Precipitation Scatter Interference Computation," September 1, 1973.
2. L. T. Gusler and D. C. Hogg, "Some Calculations on the Interference Between Satellite Communications and Terrestrial Radio Relay Systems Due to Scattering by Rain," B.S.T.J., *49*, No. 7 (September 1970), pp. 1491–1511.
3. R. K. Crane, "Analysis of Data from the Avon-to-Westford Experiment," Lincoln Laboratory Technical Report No. 498, January 8, 1973.

4. CCIR (International Radio Consultative Committee) Report 339-1 (Rev. 72), "Influence of Scattering from Precipitation on Siting of Earth and Terrestrial Stations," Conclusions of the Interim Meeting of Study Group 5, Geneva, April 5–18, 1972.
5. D. E. Setzer, "Anisotropic Scattering Due to Rain at Radio-Relay Frequencies," B.S.T.J., *50*, No. 3 (March 1971), pp. 861–868.
6. D. E. Setzer, "Computed Transmission Through Rain at Microwave and Visible Frequencies," B.S.T.J., *49*, No. 8 (October 1970), pp. 1873–1892.
7. R. G. Medhurst, "Rainfall Attenuation of Centimeter Waves: Comparison of Theory and Measurement," IEEE Trans. on Antennas and Propagation, *13*, No. 4 (July 1965), pp. 550–564.
8. D. C. Hogg, "Statistics on Attenuation of Microwave by Intense Rain," B.S.T.J., *48*, No. 9 (November 1969), pp. 2949–2962.

Optimum Timing Phase for an Infinite Equalizer

By J. E. MAZO

A digital equalizer for data transmission linearly combines a sequence of samples $[b(\tau + kt), k = 0, \pm 1, \cdots]$ of the received data wave to mitigate the effects of intersymbol interference and noise. A natural question is, How will the performance of such a system depend on the timing phase τ, $0 \leqq \tau \leqq T$?

We examine this problem in considerable detail for an infinite equalizer using a mean-square measure of performance. Excess bandwidth results in a significant difference in performance between the best and worst timing phase. Under practical noise conditions, it is estimated that the excess bandwidth must be down to 1 or 2 percent before the timing effect becomes insignificant. With a 10-percent roll-off, a 3-dB penalty can be incurred by choosing a bad timing epoch.

Our main result is that under conditions likely to be encountered on channels similar to voiceband telephone channels, the optimum sampling instants will be accurately approximated by the consecutive maxima and minima of the sine wave that result when an alternating sequence of positive and negative pulses (dotting sequence) is transmitted. There are no additional local minima of the minimum mean-square error as the timing phase is varied.

I. INTRODUCTION

We consider a noisy, random, baseband pulse train,

$$b(t) = \sum a_n x(t - nT) + n(t), \tag{1}$$

where $x(t)$ is the channel (or channel front-end filter) impulse response and the a_n are independent binary data which take values ± 1 with equal probability. The additive zero mean gaussian noise process is denoted by $n(t)$. Under very ideal conditions, the pulse $x(t)$ would be of the form $(\sin \pi t/T)/(\pi t/T)$, and sampling (1) at time $t = kT$ would yield the quantity $a_k + n(kT)$. Under more realistic conditions, the brick-wall shape of the above pulse is difficult to approximate in practice, and a smoother characteristic in the frequency domain is taken as the ideal. Using some excess bandwidth (i.e., the Fourier transform extends beyond π/T rad/s) still, in principle, allows the ideal set of

samples expressed by

$$x(0) = 1, \quad x(kT) = 0, \quad k \neq 0;$$

but this ideal set is never exactly attained in practice because of unknown channel distortion. One possible measure of distortion could be

$$\frac{\sum_{k \neq 0} x^2(kT)}{x^2(0)} .$$

Actually, since there is now nothing special about choosing $t = 0$ as the sampling instant for $x(t)$, an even more appropriate measure would be

$$\min_{\tau} \frac{\sum_{k \neq 0} x^2(\tau + kT)}{x^2(\tau)} .$$

The above measure allows us to choose a best sampling epoch τ before declaring how badly the signal has been distorted. However, any direct consideration of the above type of criterion seems to result in considerable mathematical difficulty. Also, in practice, one is more interested in the situation where the set of digital samples $\{x(\tau + kT)\}$, or rather their noisy versions $\{b(\tau + kT)\}$, are linearly combined by an equalizer attempting to undo the channel distortion. It is the distortion at the output of the equalizer that is of practical interest. Remarkably enough, it is this distortion measured at the output of an infinite equalizer that is most easily studied with regard to its properties as a function of timing phase. A discussion of this dependence follows. We begin by making more precise some of our general remarks.

The received signal $b(t)$ is to be equalized by passing it through a $(2N + 1)$ tap transversal filter to yield for the signal portion of the output a waveform

$$q(t) = \sum a_n h(t - nT), \tag{2}$$

where

$$h(t) = \sum_{-N}^{N} c_n x(t - nT). \tag{3}$$

The coefficients c_n are the tap weights of the equalizer and are adjusted to optimize some measure of performance. Since decisions are based on the sampled output $q(t)$, one usually discusses the problem by fixing a sampling epoch τ and denotes the sampled values of $x(t)$ and $h(t)$ by x_k and h_k, respectively. That is,

$$x(\tau + kT) \equiv x_k$$

$$h(\tau + kT) \equiv h_k = \sum_{n=-N}^{N} c_n x_{k-n} \qquad k = \cdots -1, 0, 1, \cdots. \tag{4}$$

We assume throughout that the tap weights are chosen to minimize the expression

$$\mathcal{E}^2 = \sum_{n=-\infty}^{\infty}{}' h_n^2 + (h_0 - 1)^2 + \sigma_0^2, \tag{5}$$

where σ_0^2 is the sampled noise variance at the output of the equalizer, and the prime on the summation indicates deletion of the $n = 0$ term. Letting $R_n = R_{-n}$ be the (discrete) correlation function of the sampled noise at the input to the equalizer, we see that

$$\sigma_0^2 = \sum_{j,k=-N}^{N} c_j R_{j-k} c_k \tag{6}$$

and

$$\mathcal{E}^2 = \sum_{j,k=-N}^{N} c_j (R_{j-k} + A_{j-k}) c_k - 2 \sum_{n=-N}^{N} c_n x_{-n} + 1, \tag{7}$$

where the channel correlation matrix,

$$A_{n-m} = \sum_{k=-\infty}^{\infty} x_{k-n} x_{k-m}, \tag{8}$$

has been introduced. The optimum tap settings $(c_{\text{opt}})_n$ are determined by differentiating $\mathcal{E}^2$ with respect to the tap gains. This yields a linear system of equations for the best tap gains which, in an obvious matrix-vector notation, reads

$$(A + R)c_{\text{opt}} = x. \tag{9}$$

Using this in (7), we then obtain for the minimum mean-square error

$$\mathcal{E}^2_{\text{min}} = 1 - h_0. \tag{10}$$

Finally, we may state our problem. The above discussion of known results was in terms of a fixed timing epoch τ. Here we shall determine for some special but important cases the value of τ, which will minimize $\mathcal{E}^2_{\text{min}}$, as defined above. Instead of resorting to a direct numerical inversion of (9), we introduce the approximation of treating the equalizer as infinite. This approximation aids the analytical inversion of (9) and thereby directly leads to useful insights concerning the proper timing epoch. It will be seen that in the $N = \infty$ limit, timing recovery is only a problem when the received pulse $x(t)$ has a bandwidth greater than $1/2T$ Hz, a result that could have been guessed from the sampling theorem. With the practical situation of excess bandwidth, there is the possibility that a badly chosen timing phase will considerably degrade performance. Only when there is very little excess bandwidth,

say of the order of one percent, is the timing unimportant in practice. Our main result can be best stated in terms of the pure sine wave that appears at the channel output (sampler input) when the sequence $a_n = (-1)^n$ (dotting sequence) is transmitted. The optimum sampling instants are, under certain conditions, the consecutive maxima and minima of this wave. Furthermore, if the timing phase is varied, no extraneous local minima of the minimum mean-square error (regarded as a function of timing phase) will appear.

II. MATHEMATICAL DETERMINATION OF OPTIMUM TIMING PHASE

On comparing (8) and (9), we see that the infinite matrix equation is in the nature of a convolution and, therefore, is best solved by transform methods. Thus, given a sequence $\{x_n\}$, we define a transform

$$X_{eq}(w) = T \sum_{-\infty}^{\infty} x_n \exp(-iwnT), \quad |w| \leqq \frac{\pi}{T}, \tag{11}$$

with the x_k recoverable from $X_{eq}(w)$ in the obvious way:

$$x_k = \frac{1}{2\pi} \int_{-(\pi/T)}^{\pi/T} X_{eq}(w) e^{iwkT} dw. \tag{12}$$

This formula shows that, from the impulse-response point of view, $X_{eq}(w)$ is, in fact, the Nyquist equivalent spectrum of $x(t)$. That is, if $X(w)$ is the Fourier transform of $x(t)$, then

$$X_{eq}(w) = \sum_{k=-\infty}^{\infty} X\left(w - k\frac{2\pi}{T}\right), \quad |w| \leqq \frac{\pi}{T}. \tag{13}$$

In this transform language, the solution of (9) is*

$$C(w) = \frac{TX_{eq}^*(w)}{R_{eq}(w) + \frac{1}{T}|X_{eq}(w)|^2}. \tag{14}$$

In writing this, we have made use of (8) to write

$$A(w) = \frac{1}{T}|X_{eq}(w)|^2. \tag{15}$$

Equation (4) gives the relation

$$H(w) = \frac{1}{T}C(w)X_{eq}(w) = \frac{|X_{eq}(w)|^2}{R_{eq}(w) + \frac{1}{T}|X_{eq}(w)|^2} \tag{16}$$

---

* We will not always use the subscript "eq" on the transform (11) when there does not exist an associated impulse response in a natural way.

and, hence, from (10) and the recovery formula for h_0, we have

$$\mathcal{E}^2_{\min} = \frac{T}{2\pi} \int_{-(\pi/T)}^{\pi/T} \frac{R(w)}{R_{eq}(w) + \dfrac{1}{T} |X_{eq}(w)|^2} \, dw. \tag{17}$$

Let us now explicitly put timing recovery into this solution for $\mathcal{E}^2_{\min}$. In evaluating the above formula, one would naturally calculate $X(w)$, possibly by choosing $t = 0$ as the peak of the pulse $x(t)$, and calculate $X_{eq}(w)$ by (13). Any other timing epoch τ can be obtained by replacing $X(w)$ by $X(w) \exp(iw\tau)$. This corresponds to sampling the original $x(t)$ at $t = \tau$ instead of $t = 0$; in general, $X_{eq}(w)$ will depend on τ.

For further specifics, we limit our discussion to small but nonzero excess bandwidth, letting α denote the fraction of excess bandwidth. Thus, $(1 + \alpha)\pi/T$ is the total bandwidth (in rad/s) occupied by the baseband pulse, $\alpha < 1$. Further, we take the noise at the input to the equalizer to be independent from sample to sample and to have variance σ_i^2; thus,

$$R_{eq}(w) = T\sigma_i^2. \tag{18}$$

Equation (17) thus becomes

$$\begin{aligned}
\mathcal{E}^2_{\min} &= \sigma_i^2 \frac{T}{2\pi} \int_{-(\pi/T)}^{\pi/T} \frac{dw}{\sigma_i^2 + \dfrac{1}{T^2} |X_{eq}(w)|^2} \\
&= \sigma_i^2 \frac{T}{2\pi} \int_{-(\pi/T)(1-\alpha)}^{\pi/T(1-\alpha)} \frac{dw}{\sigma_i^2 + \dfrac{1}{T^2} |X_{eq}(w)|^2} \\
&\quad + 2\sigma_i^2 \frac{T}{2\pi} \int_{+(\pi/T)(1-\alpha)}^{\pi/T} \frac{dw}{\sigma_i^2 + \dfrac{1}{T^2} |X_{eq}(w)|^2}. \tag{19}
\end{aligned}$$

We note that the first term in the right-hand side of (19) is not affected by timing phase since the portion of $|X_{eq}(w)|$ for $|w| \leq \pi/T(1 - \alpha)$ is not affected, whereas the second term, on account of the foldover from excess bandwidth, is affected.*

To obtain estimates regarding the magnitude of the effects of poor timing phase, we consider the case of an undisturbed pulse $x(t)$ having a transform $X(w)$ that is equal to T for $0 \leq w \leq (1 - \alpha)\pi/T$ and decreases linearly to zero for $(1 - \alpha)\pi/T \leq w \leq (1 + \alpha)\pi/T$, with $X(-w) = X(w)$. If this pulse is sampled at $t = 0$, the "main bang" has value unity, and there is no intersymbol interference. From (19), the minimum mean-square error is $\sigma_i^2(1 + \sigma_i^2) \approx \sigma_i^2$ for small noise.

---

* Of course, if there is no noise, the equalizer inverts $X_{eq}(w)$ (if the latter is not zero) without any penalty, and the question of timing is moot.

If, however, it is sampled at $\tau = T/2$, we have

$$|X_{eq}(w)|^2 = \begin{cases} T^2, & 0 \leq w \leq (1-\alpha)\dfrac{\pi}{T} \\[2mm] \dfrac{T^2}{\alpha^2\pi^2}\left(w - \dfrac{\pi}{T}\right)^2, & (1-\alpha)\dfrac{\pi}{T} \leq w \leq \dfrac{\pi}{T} \end{cases} \quad (20)$$

and, hence, with the infinite equalizer doing the best it can, we obtain from (19) for small noise

$$\mathcal{E}^2_{min} = (1-\alpha)\sigma_i^2 + \alpha\sigma_i \tan^{-1}\frac{1}{\sigma_i} \quad (21)$$

$$\approx (1-\alpha)\sigma_i^2 + \frac{\pi}{2}\alpha\sigma_i.$$

If for binary transmission we assume a 10 or 15 percent roll-off ($\alpha = 0.1$ or 0.15) and $\sigma_i^2 = 0.01$ to 0.04, we see that a degradation of from 2 dB to 4 dB might be encountered. Only if the roll-off is of the order of 1 or 2 percent is the effect of timing negligible for the value of noise power considered here.

We now concentrate further attention on the second term of (19). We introduce a variable

$$\mu = \frac{T}{\pi\alpha}\left[w - \frac{\pi}{T}\right] \quad (22)$$

for $\pi/T(1-\alpha) \leq w \leq \pi/T(1+\alpha)$; thus, $-1 \leq \mu \leq 1$. When w and μ are related by (22), we define a nonnegative amplitude function $R(w)$ and a real phase function $\varphi(\mu)$ by the relation

$$X(w) = R(\mu)e^{i\varphi(\mu)}, \quad R(\mu) \geq 0. \quad (23)$$

Further, we introduce the even and odd parts of amplitude and phase about the Nyquist frequency $w = \pi/T(\mu = 0)$:

$$\begin{aligned} R(\mu) &= R_e(\mu) + R_0(\mu) \\ R_e(\mu) &= R_e(-\mu); R_0(\mu) = -R_0(-\mu); \end{aligned} \quad (24)$$

$$\begin{aligned} \varphi(\mu) &= \varphi_e(\mu) + \varphi_0(\mu) \\ \varphi_0(\mu) &= \varphi_e(-\mu); \varphi_0(\mu) = -\varphi_0(-\mu). \end{aligned} \quad (25)$$

Thus, for $-1 \leq \mu \leq 0$, we calculate

$$|X_{eq}(\mu)|^2 = 4[(R_e^2 - R_0^2)\cos^2\varphi_e + R_0^2], \quad (26)$$

where, for notational simplicity, the μ dependence of the function on the right has not been shown. Using (26), the second term S of (19) is written

$$S = \alpha\sigma_i^2 \int_{-1}^0 \frac{d\mu}{\sigma_i^2 + \dfrac{4}{T^2}\left[(R_e^2 - R_0^2)\cos^2\varphi_e + R_0^2\right]}. \quad (27)$$

A number of remarks concerning (27) are in order.

(i) The odd component of phase distortion about the Nyquist frequency does not influence "foldover" or the timing recovery problem.

(ii) From the positivity of $R(\mu)$, we note that $(R_e^2 - R_0^2) \geqq 0$. We assume that $[R_e^2(\mu) - R_0^2(\mu)] > 0$ on some interval contained in $-1 \leqq \mu \leqq 0$.

(iii) Relative to an arbitrarily selected sampling instant, called zero, the timing phase τ enters φ_e through

$$\varphi_e = \frac{\pi}{T} \tau + \bar{\varphi}_e(\mu), \tag{28}$$

where $\bar{\varphi}_e(\mu)$ is the phase characteristic for $\tau = 0$. It is evident that the mean-square error is a periodic function of τ with period T. Physically, if τ is displaced by an integer (k) multiple of T, the "center tap" will be shifted k units down the equalizer.

(iv) In many situations, $\bar{\varphi}_e$ can, to a first approximation, be treated as a constant. A recent survey[1] shows that an "average" telephone channel has an envelope delay that changes by 312 μs from 2850 Hz to 3000 Hz. Assuming this to be an excess bandwidth region and the change to be due entirely to a quadratic term in the phase curve yields a phase change of 8.4 degrees.*

We now come to the main point of this section, which is to recognize where the optimum timing phase is. This will be done with the simplifying assumption

$$\bar{\varphi}_e = \theta = \text{const.} \tag{29}$$

Since S is a smooth periodic function of τ, we look for a minimum by examining solutions of

$$\frac{\partial S}{\partial \tau} = 0, \tag{30}$$

that is,

$$\sin 2\varphi_e \int_{-1}^{0} \frac{(R_e^2 - R_0^2)d\mu}{\left\{ \sigma_i^2 + \frac{4}{T^2} \left[(R_e^2 - R_0^2) \cos^2 \varphi_e + R_0^2 \right] \right\}^2} = 0. \tag{31}$$

From (ii) we deduce that the integral is strictly positive and, hence, the only solutions are those of

$$\sin 2\varphi_e = 0. \tag{32}$$

---

* This is meant to be a rough estimate only. No attempt was made to separate out even and odd contributions to the change in phase. Low-end envelope delay is more severe and this separation should be carried out so as to obtain a not-too-pessimistic answer.

Thus, because of periodicity the only solutions τ of interest may be written

$$\frac{\pi}{T}\tau + \theta = 0 \tag{33a}$$

or

$$\frac{\pi}{T}\tau + \theta + \frac{\pi}{2} = 0. \tag{33b}$$

By taking second derivatives of S with respect to τ, we see that the first of these, namely (33a), corresponds to the minimum of S while the second solution is a maximum. In addition to giving a description of optimum timing phase, our discussion has also shown that for $\varphi_e = $ const. there are no other local minima of $\mathcal{E}^2_{\min}$ as τ is varied.[*] Thus, a gradient search for the optimum τ is feasible. More precisely, we have shown this for the quadratic Nyquist I criterion (5), and have assumed that the search is slow in the sense that the equalizer "settles down" before the next step in τ is taken.

Equation (33a) provides a mathematical description of the optimum timing phase τ in terms of the phase θ of the channel at the Nyquist frequency π/T. It is useful to give a more physical interpretation of this result. Consider the transmission of the special sequence $a_n = (-1)^n$. From (1), we receive the signal

$$\sum_{n=\infty}^{\infty} (-1)^n x(t - nT),$$

which has as its Fourier transform

$$X(w) \sum_{n=\infty}^{\infty} (-1)^n \exp(-iwnT)$$

$$= X(w) \sum_{k=\infty}^{\infty} \exp(-iw2kT) - X(w) \exp(-iwT) \sum_{k=-\infty}^{\infty} \exp(-iw2kT)$$

$$= \frac{2\pi}{T} X(w) \sum_{-\infty}^{\infty} \delta\left(w - \frac{k\pi}{T}\right) - \frac{2\pi}{T} X(w) \exp(-iwT) \sum_{-\infty}^{\infty} \delta\left(w - \frac{k\pi}{T}\right)$$

$$= \frac{4\pi}{T} X(w) \sum_{k \text{ odd}} \delta\left(w - \frac{k\pi}{T}\right).$$

Hence, if $X(w)$ is band-limited to a bandwidth less than $3\pi/T$, and if we write $X(\pi/L) = A \exp(i\theta)$, the received signal is

$$\frac{4}{T} A \cos\left(\frac{\pi t}{T} + \theta\right).$$

Comparing this with (32) and (33a), we see that the optimum timing

---

[*] See the appendix for a more general condition under which the existence of a unique minimum is guaranteed.

phase is at the successive maxima and minima of the tone that results when a dotting sequence is transmitted.

Finally, we note that corrections for a small, varying, even component of phase distortion about the Nyquist frequency are easy to make. If we define γ by

$$\frac{\pi}{T}\tau + \theta \equiv \gamma, \tag{34}$$

where, again, θ is the phase at the Nyquist frequency, and if we write [recalling (28)],

$$\varphi_e(\mu) = \gamma + \epsilon(\mu), \quad \epsilon(0) = 0,$$

where $\epsilon(\mu)$ is assumed small, then

$$\gamma^* = -\frac{\displaystyle\int_{-1}^0 \frac{(R_e^2 - R_0^2)\epsilon(\mu)}{[\sigma_i^2 + (4/T^2)R_e^2]^2}\, d\mu}{\displaystyle\int_{-1}^0 \frac{(R_e^2 - R_0^2)}{\{\sigma_i^2 + (4/T^2)R_e^2\}^2}\, d\mu}, \tag{35}$$

where γ^* is the approximate value of γ for optimum timing phase τ. We see γ^* is bounded in magnitude by max $\epsilon(\mu)$, but in general it will be a fraction of this. Assuming max $\epsilon(\mu)$ corresponds to 9 degrees, this rough bound suggests that a correction of at most 5 percent of a symbol interval would be required.

For a more refined estimate we take the following

$$R_e(\mu) = 1, \quad R_0(\mu) = \mu, \quad \text{and} \quad \epsilon(\mu) = \epsilon\mu^2. \tag{36}$$

Note that $R_e^2 - R_0^2 = (R_e - R_0)(R_e + R_0) = R(\mu)R(-\mu)$. Since $R(1)$ will vanish by the design of the signal, $R_e^2 - R_0^2$ will be zero at $\mu = -1$ and, thus, no multiplicative constant is included in $R_0(\mu)$. Using (36) in (35) yields

$$\gamma^* \approx -\epsilon\frac{\displaystyle\int_{-1}^0 (1 - \mu^2)\mu^2 d\mu}{\displaystyle\int_{-1}^0 (1 - \mu^2)d\mu} = -\frac{\epsilon}{5}. \tag{37}$$

Thus an estimate of $\frac{1}{5}$ of the even part of the phase variation over the roll-off band results for the correction term to (34). This would appear to be a negligible effect.

III. DISCUSSION FOR PASSBAND TRANSMISSION

We now generalize the foregoing discussion to treat a QAM passband signal format. We assume a transmitted signal

$$[\textstyle\sum a_n p(t - nT)]\cos w_c t - [\textstyle\sum b_n p(t - nT)]\sin w_c t, \tag{38}$$

where (a_n) and (b_n) are independent digits. Let the impulse response of the linear distortion medium $m(t)$ be written as

$$m(t) = 2F_1(t) \cos w_c t - 2F_2(t) \sin w_c t. \tag{39}$$

Further, denote the convolution of $p(t)$ with $F_1(t)$ and $F_2(t)$ by $x(t)$ and $y(t)$, i.e.,

$$p(t){*}F_1(t) = x(t) \quad p(t){*}F_2(t) = y(t). \tag{40}$$

Then the received signal is

$$r(t) = \cos w_c t \left[\sum a_n x(t - nT) - \sum b_n y(t - nT) \right] \\ - \sin w_c t \left[\sum a_n y(t - nT) + \sum b_n x(t - nT) \right] + n(t), \tag{41}$$

where $n(t)$ represents additive gaussian noise. The waveform $r(t)$ is then linearly filtered by a passband equalizer, whose output can be described by replacing $x(t)$ by $g(t)$ and $y(t)$ by $h(t)$, where

$$g(t) = \sum_{-N}^{N} c_n x(t - nT) - \sum_{-N}^{N} d_n y(t - nT)$$

and $\hspace{10cm}$ (42)

$$h(t) = \sum_{-N}^{N} d_n x(t - nT) + \sum_{-N}^{N} c_n y(t - nT).$$

Further, if the input noise $n(t)$ has correlation $R(t)$, then the output noise has variance

$$\sigma_0^2 = c^+ R c + d^+ R d, \tag{43}$$

where $R_{ij} = R[(i - j)T]$. We adopt as our criterion

$$\mathcal{E}^2 = \sum_{i \neq 0} g_i^2 + (1 - g_0)^2 + \sum_{i=-\infty}^{\infty} h_i^2 + \sigma_0^2 \tag{44}$$

and seek to minimize this quantity. Analogously with the baseband case we define two baseband autocorrelation functions [corresponding to $x(t)$ and $y(t)$ in (41)] by

$$A_{nm} = A_{n-m} = \sum_{j=-\infty}^{\infty} x_{j-n} x_{j-m}$$

and $\hspace{10cm}$ (45)

$$B_{nm} = B_{n-m} = \sum_{j=-\infty}^{\infty} y_{j-n} y_{j-m}.$$

Further, a cross-correlation matrix between $x(t)$ and $y(t)$, namely,

$$K_{nm} = K_{n-m} = -K_{mn} = \sum_{j} (x_{j-n} y_{j-m} - x_{j-m} y_{j-n}). \tag{46}$$

If, as before, we write $(x)_n = x_{-n}$, $(y)_n = y_{-n}$, then

$$\mathcal{E}^2 = c^+(A + B)c + d^+(A + B)d - 2c^+Kd - 2c \cdot x + 2d \cdot y \\ + 1 + c^+Rc + d^+Rd. \quad (47)$$

If one defines the matrix $Q \equiv A + B + R$, the equations for the optimum taps are

$$\sum_m (Q_{n-m}c_m - K_{n-m}d_m) = x_{-n}$$

$$\sum_m (Q_{n-m}d_m + K_{n-m}c_m) = -y_{-n} \quad (48)$$

and

$$\mathcal{E}^2_{\text{min}} = 1 - (q_0)_{\text{opt}} = 1 - c_{\text{opt}} \cdot x + d_{\text{opt}} \cdot y. \quad (49)$$

We introduce the same transform as before, noting

$$Q(w) = \frac{|X_{\text{eq}}(w)|^2}{T} + \frac{|Y_{\text{eq}}(w)|^2}{T} + R(w) = Q(-w) \quad (50)$$

and

$$K(w) = \frac{1}{T} \left[X^*_{\text{eq}}(w) Y_{\text{eq}}(w) - X_{\text{eq}}(w) Y^*_{\text{eq}}(w) \right] = -K(-w). \quad (51)$$

Straightforward solution of (48) and (49) yields

$$\mathcal{E}^2_{\text{min}} = \frac{T}{2\pi} \int_{-(\pi/T)}^{\pi/T} \frac{Q(w)R(w)}{Q^2(w) + K^2(w)} \, dw$$

$$= \frac{T}{2\pi} \int_{-(\pi/T)}^{\pi/T} \frac{[Q(w) - iK(w)]R(w)}{Q^2(w) + K^2(w)} \, dw$$

$$= \frac{T}{2\pi} \int_{-(\pi/T)}^{\pi/T} \frac{R(w)}{Q(w) + iK(w)} \, dw, \quad (52)$$

where the oddness of $K(w)$ [and evenness of $Q(w)$ and $R(w)$] has been used. If we formally introduce $S(w)$ by

$$S(w) = X(w) + iY(w), \quad (53)$$

and

$$S_{\text{eq}}(w) = X_{\text{eq}}(w) + iY_{\text{eq}}(w), \quad (54)$$

we see that

$$|S_{\text{eq}}(w)|^2 = [X_{\text{eq}}(w) + iY_{\text{eq}}(w)][X^*_{\text{eq}}(w) - iY^*_{\text{eq}}(w)] \\ = |X_{\text{eq}}(w)|^2 + |Y_{\text{eq}}(w)|^2 \\ + iY_{\text{eq}}(w)X^*_{\text{eq}}(w) - iX_{\text{eq}}(w)Y^*_{\text{eq}}(w) \\ = T[A(w) + B(w) + iK(w)], \quad (55)$$

or, briefly,

$$Q + iK = \frac{|S_{\text{eq}}|^2}{T} + R. \quad (56)$$

Thus,

$$\mathcal{E}_{\min}^2 = \frac{T}{2\pi} \int_{-(\pi/T)}^{\pi/T} \frac{R(w)}{\frac{|S_{eq}(w)|^2}{T} + R(w)} \, dw. \tag{57}$$

This formula has a close analogy to the baseband formula when we realize that the overall channel (transmitting and distorting filter) has frequency characteristic

$$C(w) = \tfrac{1}{2}S(w - w_c) + \tfrac{1}{2}S^*(-w - w_c). \tag{58}$$

Thus, in terms of the overall channel, (57) has a straightforward interpretation. One uses a formula similar to the baseband version (17) except that now one forms an equivalent Nyquist characteristic using twice the positive frequency part of the passband channel and treats the carrier frequency as zero. It is essential to note that the transfer characteristic from which $S_{eq}(w)$ is formed no longer needs to have even amplitude and odd phase about "zero" as would be required in the baseband case. In particular, the envelope delays at the two band edges can have an effect on timing phase. In practical cases, the carrier frequency can be chosen to give equal envelope delays, thus minimizing the effects of odd phase about the band edges. However, in doing this, the amplitude effects may have worsened the situation. The main point to be emphasized here is the interplay of carrier placement and bit timing with an excess bandwidth system.

Regarding the description of optimum timing phase, let $\varphi^\pm$ be the channel phase at $w = w_c \pm \pi/T$, and assume these phases to be constant over their respective roll-off regions. Then, for any reasonable amplitude distortions, the unique best sampling instant τ is given by

$$\frac{\pi}{T}\tau + \tfrac{1}{2}(\varphi^+ - \varphi^-) = 0. \tag{59}$$

Under these conditions, we again have that there exists but one local minimum.

IV. ACKNOWLEDGMENT

APPENDIX

Uniqueness of Local Minimum of Mean-Square Error for Small $\bar{\varphi}_s(x)$

The interesting portion of the mean-square error is given by (27), which is of the form

$$g(\gamma) \equiv \int_0^1 \frac{dx}{a(x) + b(x) \cos^2 \left[\gamma + \varphi(x) \right]}, \qquad (60)$$

with $a(x)$ and $b(x)$ positive. If $\varphi(x)$ is always small, then any local minimum of (60) has to occur when γ is small. Thus, if $|\varphi(x)| < \pi/8$, we see that the derivative

$$g'(\gamma) = \int_0^1 \frac{b(x) \sin \left[2\gamma + 2\varphi(x) \right]}{\{a(x) + b(x) \cos^2 \left[\gamma + \varphi(x) \right]\}^2} \, dx \qquad (61)$$

can only vanish if $|\gamma| < \pi/8$ or $|\pi - \gamma| < \pi/8$, for otherwise the integrand always has the same sign. Let us investigate the solution in $|\gamma| < \pi/8$ in more detail. Taking another derivative gives

$$g''(\gamma) = 2 \int_0^1 dx \, \frac{b \cos (2\gamma + 2\varphi)}{\left[a + b \cos^2 (\gamma + \varphi) \right]^2}$$
$$+ 2 \int_0^1 dx \, \frac{b^2 \sin^2 (2\gamma + 2\varphi)}{\left[a + b \cos^2 (\gamma + \varphi) \right]^3}, \qquad (62)$$

which is positive. Hence, there can be only one solution of $g'(\gamma) = 0$ for $|\gamma| < \pi/8$ (we cannot have two minima without a maximum in between). Likewise, if $|\varphi(x)| < \pi/8$, there is precisely one maximum of $g(\gamma)$, and it occurs in the range $|\pi - \gamma| < \pi/8$.

REFERENCE

1. F. P. Duffy and T. W. Thatcher, "Analog Transmission Performance on the Switched Telecommunications Network," B.S.T.J., 50, No. 4 (April 1971), pp. 1311–1347.

Contributors to This Issue

Kenneth Bullington, B.S., 1936, University of New Mexico; S.M., 1937, Massachusetts Institute of Technology; Bell Laboratories, 1937—. Mr. Bullington has worked on transmission engineering problems on wire, radio, and submarine cable systems. He is now Radio Consultant. In 1956 he received the Morris Liebmann Memorial Prize of the Institute of Radio Engineers and the Franklin Institute's Stuart Ballantine Medal for contributions in tropospheric transmission and its application to practical communications systems. Fellow, IEEE. Member, Phi Kappa Phi, Sigma Tau.

Allen H. Cherin, B.E.E., 1961, City College of New York; M.S.E.E., 1965, University of Vermont; Ph.D., 1971, University of Pennsylvania; Bell Laboratories, 1965—. Mr. Cherin is engaged in studies associated with the characterization, splicing, and packaging of optical fibers. Member, IEEE, OSA.

O. E. DeLange, B.S. (Electrical Engineering), 1930, University of Utah; M.A. (Physics), 1937, Columbia University; Bell Laboratories, 1930–1971. Mr. DeLange was involved in studies of frequency modulation and ultrahigh frequency research, and in the development and design of naval radar. He has studied broadband pulse systems with emphasis on pulse code modulation, and was responsible for the satellite tracking radar for the Echo I experiment. More recently he was involved in studies of optical and millimeter-wave systems. Fellow, IEEE.

A. F. Dietrich, Monmouth College, 1942–1944; Bell Laboratories, 1942–1973. Mr. Dietrich worked on experimental studies on broadband, baseband, and microwave systems, including experimental research on both FM and PCM ultrashort pulse terminals and repeaters for microwave radio and waveguide applications. He also engaged in experimental studies of light transmission systems.

D. C. Hogg, B.Sc., 1949, University of Western Ontario; M.Sc., 1950, Ph.D., 1953, McGill University; Bell Laboratories, 1953—. Mr. Hogg's work has included studies of artificial dielectrics for microwaves,

diffraction of microwaves, and over-the-horizon, millimeter wave, and optical propagation and antenna research. Fellow, IEEE, Union de Radio Scientifique Internationale.

Tien Pei Lee, B.S.E.E., 1957, National Taiwan University, Taiwan, China; M.S.E.E., 1959, Ohio State University; Ph.D., 1963, Stanford University; Bell Laboratories, 1963—. Mr. Lee participated in the research and development of solid-state microwave diodes and photodiodes. He has worked on millimeter-wave IMPATT diodes, varactor diodes, and Schottky barrier diodes for multipliers and mixers. He is presently engaged in semiconductor injection lasers and LED's for optical fiber communication systems. Member, Sigma Xi, IEEE.

Dietrich Marcuse, Diplom Vorpruefung, 1952, Dipl. Phys., 1954, Berlin Free University; D.E.E., 1962, Technische Hochschule, Karlsruhe, Germany; Siemens and Halske (Germany), 1954–1957; Bell Laboratories, 1957—. At Siemens and Halske, Mr. Marcuse was engaged in transmission research and studying coaxial cable and circular waveguide transmission. At Bell 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 Laboratories at the University of Utah. He is presently working on the transmission aspect of a light communications system. Mr. Marcuse is the author of three books. Fellow, IEEE; member, Optical Society of America.

George Marr, B.A., 1963, Hope College; M.S., 1965, Miami University; Ph.D. (Physics), 1968, The Ohio State University; Bell Laboratories, 1968–1974. While at Bell Laboratories, Mr. Marr worked on the design and development of custom and catalog MOS integrated circuits, ion-implanted depletion-mode devices, BIGFET output drivers, and CMOS RAM design. Member, APS, IEEE.

J. E. Mazo, B.S. (Physics), 1958, Massachusetts Institute of Technology; M.S. (Physics), 1960, and Ph.D. (Physics), 1963, Syracuse University; Research Associate, Department of Physics, University of Indiana, 1963–1964; Bell Laboratories, 1964—. At the University of Indiana, Mr. Mazo worked on studies of scattering theory. At Bell Laboratories, he has been concerned with problems in data transmission and is now working in the Mathematical Research Center. Member, American Physical Society, IEEE.

Gilbert L. Mowery, S.B., 1965, Massachusetts Institute of Technology; M.S., 1966, and Ph.D. (E.E.), 1971, Carnegie-Mellon University; Bell Laboratories, 1970—. Mr. Mowery has worked on the design and development of custom and catalog MOS integrated circuits, ion-implanted depletion-mode devices, and BIGFET output drivers, and is presently engaged in the design of custom MOS LSI circuits. Member, IEEE.

Kurt H. Mueller, E.E. Diploma, 1961, Ph.D., 1967, Swiss Federal Institute of Technology; Bell Laboratories, 1969—. Mr. Mueller has worked on various problems in the fields of high-speed data communication. During 1972–1973, he was on leave of absence at the Swiss Federal Institute of Technology, and was a member of the Executive Body of the European Informatics Network. He is presently involved in digital signal processing techniques for data transmission.

Elizabeth J. Murphy, B.S., 1964, Spring Hill College; M.S., 1966, Auburn University; Bell Laboratories, 1967—. Ms. Murphy has worked on computation and numerical analysis associated with the characterization of transmission lines and optical fibers. Member, ACM.

S. D. Personick, B.E.E., 1967, City College of New York; S. M., 1968, E.E., 1969, and Sc.D., 1969, Massachusetts Institute of Technology; Bell Laboratories, 1967—. Mr. Personick is engaged in studies of optical communication systems.

Herman M. Presby, B.A., 1962, and Ph.D., 1966, Yeshiva University; Research Scientist, Columbia University, 1966–1968; Asst. Prof. Physics, Belfer Graduate School of Science, Yeshiva University, 1968–1972; Bell Laboratories, 1972—. Mr. Presby is engaged in studies on the properties of optical fiber waveguides.

Lawrence R. Rabiner, S.B., S.M., 1964, Ph.D., 1967, Massachusetts Institute of Technology; Bell Laboratories, 1962—. Mr. Rabiner has worked on digital circuitry, military communications problems, and problems in binaural hearing. Presently he is engaged in research on speech communications and digital signal processing techniques. Member, Eta Kappa Nu, Sigma Xi, Tau Beta Pi; Fellow, Acoustical Society of America; President, IEEE G-ASSP Ad Com; member, G-ASSP Technical Committee on Digital Signal Processing, G-ASSP

Technical Committeee on Speech Communication, IEEE Proceedings Editorial Board, Technical Committee on Speech Communication of the Acoustical Society; former Associate Editor of the G-ASSP Transactions.

Stephen O. Rice, B.S. (Electrical Engineering), 1929, and D.Sc. (Hon.), 1961, Oregon State College; Bell Laboratories, 1930—1972. Mr. Rice has been concerned with theoretical problems related to electromagnetic wave propagation, signal modulation, and noise. At the time of his retirement from Bell Laboratories, he was head of the Communications Analysis Research Department. In 1965, Mr. Rice received the Mervin J. Kelly Award from the Institute of Electrical and Electronic Engineers. Fellow, IEEE.

Marvin R. Sambur, B.E.E., 1968, City College of New York; S.M., 1969, Ph.D., 1972, Massachusetts Institute of Technology; Bell Laboratories, 1968—. At present, Mr. Sambur is engaged in automatic speaker verification and automatic speech recognition research in the Acoustic Research group at Bell Laboratories. Member, Eta Kappa Nu, Sigma Xi, Tau Beta Pi.

Arvids Vigants, B.E.E., 1956, City College of New York; M.S. (E.E.), 1957, and Eng. Sc.D. (E.E.), 1962, Columbia University; Bell Laboratories, 1962—. Mr. Vigants has worked on various electromagnetic wave propagation topics and is currently working on problems in line-of-sight microwave propagation and microwave radio systems. Member, Eta Kappa Nu, Tau Beta Pi, Sigma Xi, Commission 2 of URSI, IEEE, AAAS.

Erratum

Equation (35) on page 183 should read as follows:

$$\sigma = \frac{n_1}{c} \frac{\Delta(u-2)}{\sqrt{u+2}} \left(\frac{1}{3u+2} - \frac{(u+2)}{(2u+2)^2} \right)^{\frac{1}{2}} \tag{35}$$