THE BELL SYSTEM
TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING
ASPECTS OF ELECTRICAL COMMUNICATION

Volume 54 September 1975 Number 7

Pulse Broadening in Multimode Optical Fibers

By J. A. ARNAUD

(Manuscript received February 21, 1975)

Closed-form expressions are obtained for the impulse response of graded-index fibers whose relative permittivity is a homogeneous function of the two transverse coordinates x, y, and for the impulse width in graded-index fibers whose profile departs slightly, but otherwise arbitrarily, from a square law. The inhomogeneous dispersion of the material is taken into account. Pulse broadening can be reduced by a factor of 12 from the value obtained for square-law fibers. Simple expressions are found for the acceptance of highly oversized fibers.

I. INTRODUCTION

Light-emitting diodes supply their optical power in a time and space incoherent form. The line width is typically of the order of 200 Å, and the radiation is approximately lambertian with an emissive area of the order of 50 × 50 μm. Time and space incoherent optical pulses can be transmitted by oversized optical fibers. However, optical pulses propagating in such fibers tend to broaden as they travel. This is in part due to the nonzero line width of the source and the dispersion ($d^2k/d\omega^2$) of the fiber material. The other cause of pulse broadening is associated with the fact that the time of flight of a pulse along a ray depends on the ray trajectory. Pulses traveling along axial rays usually go faster than pulses traveling along rays of large amplitude. Because both types of rays are excited by spatially incoherent sources, the difference in axial group velocity causes a broadening of the input pulse. In the main text of this paper, we assume that the carrier is

monochromatic and that the spatial distribution of the rays is time-invariant. This is the case, for instance, when the source is an injection laser that oscillates simultaneously on many transverse modes. The difference in frequency between these various transverse modes can usually be neglected.

It was first pointed out by Kompfner[1] that pulse broadening in step-index fibers could be drastically reduced by introducing ray equalizers at various locations along the fiber. The role of ray equalizers is to exchange fast and slow rays. A possible implementation of this idea is shown in Fig. 1 together with the calculated impulse response for uncorrected and corrected step-index fibers.[2] Because natural mode mixing appears to be very small in the most recently made optical fibers, ray converters may be practical. They have not been experimented with, however, and we shall therefore restrict ourselves to uniform, uncorrected fibers.

Important results concerning the broadening of spatially incoherent optical pulses in graded-index fibers have already been reported. In Refs. 3 to 9, the difference in group velocity between the various modes that can propagate in step-index and graded-index fibers has

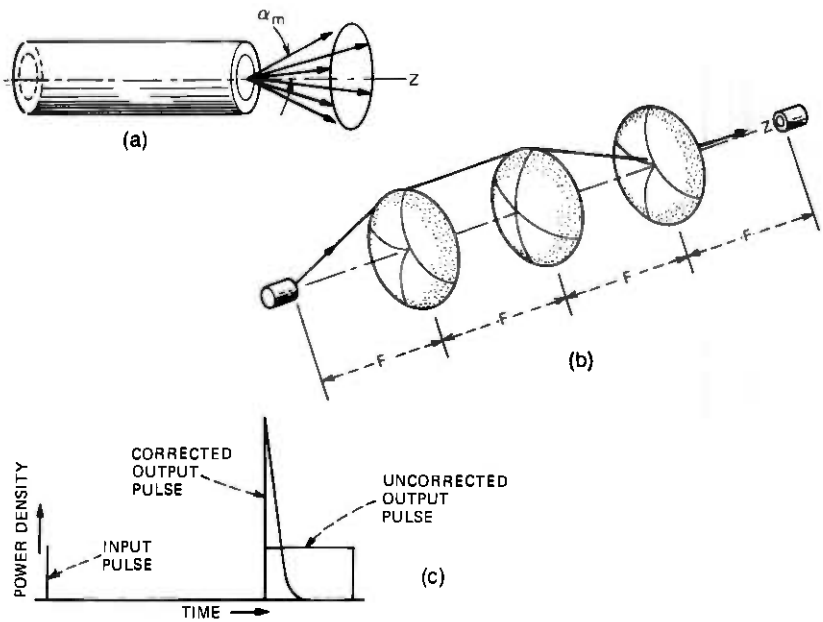


Fig. 1—Ray converter that minimizes pulse broadening in step-index fibers. (a) Angular spread of a step-index fiber. (b) Optical arrangement with confocal lenses. The first and last lenses are unconventional. (c) Calculated impulse response for uncorrected [$P(t) = 1$ within the pulse] and corrected step-index fibers (from Ref. 2).

been evaluated. The impulse response is obtained by adding the contribution of each mode, under the assumption that all modes are equally excited by the source. The calculation of the group velocities can be simplified with the help of the W.K.B. approximation (see, in particular, Ref. 4).

Let us now describe an alternative ray-optics method. The time of flight of a pulse along a ray is first evaluated according to the laws of geometrical optics. A ray can be defined by the point x, y where it intersects the input plane of the fiber (plane $z = 0$), and by the transverse components, k_x, k_y of the wave vector **k**. **k** is, by definition, directed along the ray and has magnitude $(2\pi/\lambda_0)n$, where λ_0 denotes the free-space wavelength and n the refractive index of the fiber material, usually a function of x and y. Thus, the time of flight of a pulse (at a fixed carrier frequency) is, in general, a function of the four parameters x, y, k_x, k_y. These four parameters can be considered the components of a four-vector **p**, in the so-called phase space. The impulse response is subsequently obtained by assuming that the density of rays is equal to $(2\pi)^{-2}$ in the phase space. In other words, we assume that the number of rays whose points of intersection with the input plane are between x, $x + dx$ and y, $y + dy$, and whose direction is defined by values of k_x, k_y between k_x, $k_x + dk_x$ and k_y, $k_y + dk_y$, is equal to $dxdy\ dk_xdk_y/(2\pi)^2$. The total power transmitted is the acceptance (or number of modes) of the fiber. This is the power transmitted for a source of luminance unity (see, for example, Ref. 10).

The approach used in Refs. 11 to 13 is based on the conventional ray equations. We have shown in Refs. 14 and 15 that it brings a considerable simplification to write the ray equations in the Hamiltonian form. The relationship between the ray-optics method and the W.K.B. method becomes more obvious with the Hamiltonian form. It can be shown that the W.K.B. method and the ray-optics method are essentially identical.[14]

An important difference, however, should be noted. In the W.K.B. method, modes whose axial wave number k_z is less than the free wave number k_s in the surrounding medium (or cladding) are assumed to leak out so rapidly that they can be ignored. On that basis, the acceptance of a step-index round fiber with radius a, for example, is found to be

$$N^2 = (k^2 - k_s^2)a^2/2 \equiv V^2/2.$$

The radiation loss of leaky modes can be small in the case of highly oversized fibers, however, as was pointed out by Snyder.[16] The ray-optics condition is distinctly different: Only those rays are ignored whose tangential component of the wave vector at the core-cladding

interface $[(k_z^2 + k_\varphi^2)^{\frac{1}{2}}$, where k_φ denotes the azimuthal wave number] is less than the free wave number k_s in the surrounding medium. According to ray optics, the acceptance of a step-index fiber is $N^2 = V^2$ instead of $V^2/2$. The influence of the slightly leaky rays on the impulse response of fibers has not been observed. This is perhaps because high-order modes are more sensitive to irregularities than low-order modes. Slightly leaky rays may become important when highly oversized fibers of good quality are fabricated. This is even more so for graded-index (e.g., near-square-law) fibers, because the field decays exponentially beyond the caustic line, which bounds the ray trajectories.

In most previous works, the effect of inhomogeneous dispersion* on quasi-monochromatic pulse broadening was neglected. This effect, however, was taken into account for square-law and linear-law graded-index fibers in Appendix B of Ref. 14, and by Gambling and Matsuhara[9] for circularly symmetric modes in square-law fibers perturbed by an r^4 term. The result for arbitrary small deviations from square-law was given by Arnaud in Ref. 15. Olshansky and Keck[9] first pointed out that inhomogeneous dispersion is of great practical importance, at least for fibers doped with TiO_2. Dispersion for the promising GeO_2 doped fibers is not known at the time of this writing. The variation of the loss of that material as a function of doping is likewise unknown. If we consider further that the sources used in pulse broadening experiments are not fully characterized in terms of their distribution in phase space, it appears that a precise comparison between theory and experiment is difficult at the moment. We shall therefore restrict ourselves to the theoretical evaluation of pulse broadening.

II. GENERAL RESULTS

The derivations of the general results given in this section appear in Appendix A. They follow in a straightforward manner from the Hamilton equations for pulse trajectories in space-time.

Fibers are most often characterized by a refractive-index profile: $n(x, y, \omega)$. However, the quantity that enters directly in the expressions for pulse broadening is the square of the wave number $k^2(x, y, \omega) \equiv (2\pi/\lambda_0)^2 n^2(x, y, \omega)$, where λ_0 denotes the wavelength in free space. We shall therefore deal directly with $k^2(x, y, \omega)$.

Let $x(z)$, $y(z)$ denote a ray trajectory. Assuming that the fiber is time-invariant and uniform and that the material is isotropic, we ob-

* Inhomogeneous dispersion refers to the spatial variations of the ratio of the local phase to group velocities in the material. This parameter should not be confused with the parameter $d^2k/d\omega^2$, usually refered to as "material dispersion." The latter is important only for broadband carriers.

tain from the ray equations the following differential equation (see Appendix A):

$$\tfrac{1}{2}k_z^2 d^2(X + Y)/dz^2 = k^2 - k_z^2 + X\partial k^2/\partial X + Y\partial k^2/\partial Y, \qquad (1)$$

where we have set, for convenience, $X \equiv x^2(z)$, $Y \equiv y^2(z)$. The quantity k_z in (1) denotes the axial (z) component of the wave vector $\mathbf{k}$ and is a constant of motion. In other words, k_z remains the same along any given ray. In a wave theory, k_z corresponds to the propagation constant of a mode (sometimes denoted β). Note that, in spite of the fact that we are using the language of wave optics, the theory given in this paper is based strictly on ray optics, except when we impose the condition $k_z > k_s$ to make contact with previous results.

It follows from the space-time ray equations that the ratio of the time of flight of a pulse along a ray to the corresponding time on axis is (see Appendix A)

$$t = (k_0/k_z)\langle \partial k^2/\partial \omega^2 \rangle / (dk_0^2/d\omega^2), \qquad (2a)$$

where $k_0 \equiv k(0, 0, \omega)$. The sign $\langle\ \rangle$ denotes an average over a ray period. For any function $a(x, y, \omega)$, we have defined

$$\langle a(x, y, \omega) \rangle \equiv Z^{-1} \int_0^Z a[x(z), y(z), \omega] dz, \qquad (2b)$$

where $x(z)$, $y(z)$ denotes a particular ray trajectory and Z the ray period. If the ray trajectory is not periodic, $\langle a \rangle$ should be understood as the limit of the right-hand side of (2b) when $Z \to \infty$. In the special case where the inhomogeneous dispersion of the material can be neglected, k is proportional to ω and, consequently, $\partial k^2/\partial \omega^2 = k^2/\omega^2$. Thus, (2a) reduces to

$$t = \langle k^2 \rangle / (k_0 k_z). \qquad (2c)$$

Finally, if the source of rays has a distribution $f(\mathbf{p})$ in the phase space $\mathbf{p} \equiv \{x, y, k_x, k_y\}$, the response of the fiber to an input $P'(t)$ is (see Appendix A)

$$P(t) = \int P'[t - t(\mathbf{p})] f(\mathbf{p}) T(\mathbf{p}) (d\mathbf{p}). \qquad (3)$$

The quantity $T(\mathbf{p})$ is the transmission of a ray (usually $T < 1$), and $(d\mathbf{p}) \equiv dx dy\, dk_x dk_y$. In the special case of a uniform lambertian source of luminance unity, we have $f(\mathbf{p}) = 1/(2\pi)^2$. For simplicity, we can assume that $T(\mathbf{p})$ is unity when the point x, y is within the core cross section and the components k_x, k_y of $\mathbf{p}$ are within some area to be defined later for specific examples and zero outside that area. All the subsequent results follow from (1), (2), and (3).

III. IMPULSE RESPONSE WHEN $k^2(x, y) - k_0^2$ IS A HOMOGENEOUS FUNCTION OF x AND y

Let the differential equation (1) be averaged over a ray period. The left-hand side of (1) vanishes because $d(X + Y)/dz$ assumes the same values at the ends of the integration interval. (In this integration, k_z can be considered a constant.) Thus, we have

$$\langle k^2 - k_z^2 + X\partial k^2/\partial X + Y\partial k^2/\partial Y\rangle = 0. \tag{4}$$

Let us further assume that $h(X, Y) \equiv k^2(X, Y) - k_0^2$ is a homogeneous function of degree κ in $X \equiv x^2$ and $Y \equiv y^2$. This means that, for any λ,

$$h(\lambda X, \lambda Y) = \lambda^\kappa h(X, Y). \tag{5}$$

If we differentiate (5) with respect to λ and set $\lambda = 1$, we obtain

$$X\partial h/\partial X + Y\partial h/\partial Y = \kappa h(X, Y). \tag{6}$$

Thus, going back to $k^2(x, y, \omega)$,

$$X\partial k^2/\partial X + Y\partial k^2/\partial Y = \kappa(k^2 - k_0^2). \tag{7}$$

In that case, a simple and general expression for the relative delay in the absence of material dispersion is readily obtained from (2c), (4), and (7),

$$t = [(k_z/k_0) + \kappa(k_0/k_z)]/(1 + \kappa). \tag{8}$$

The relative delay t is plotted in Fig. 2 as a function of k_z/k_0 with κ as a parameter. This result is applicable, for example, to the index profile

$$k^2(x, y) = k_0^2 - \alpha|x| - \beta|y|, \tag{9}$$

where α and β denote constants. In that example, $\kappa = \frac{1}{2}$. Note that the fiber described by (9) is not circularly symmetric, even if $\alpha = \beta$. Examples of circularly symmetric fibers that satisfy (5) will be given in the uext section.

In almost any z-invariant focusing system, any initial distribution eventually reaches a steady state. This steady state in general differs from the initial distribution. A lambertian distribution f = constant, however, remains lambertian because it is a (trivial) solution of the Liouville equation (see Appendix A). Note that the distribution f in (3) represents a ray (or pulse) density. If the medium introduces a nonuniform attenuation on the rays, the power density $T(\mathbf{p})f(\mathbf{p})$ in phase space needs to be distinguished from the distribution f.

A fiber is usually surrounded by a homogeneous material, called the cladding, with wave number k_s. For fibers that are not highly

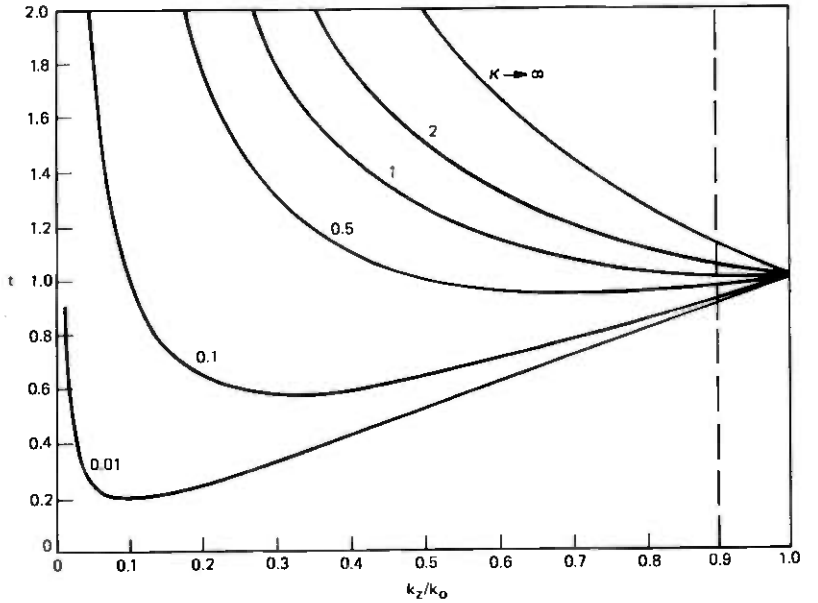


Fig. 2—Relative time of flight in a fiber where $k^2(x, y) - k_0^2$ is a homogeneous function of degree κ in x and y $[k_0 \equiv k(0, 0)]$. For most fibers, k_z/k_0 is close to unity.

overmoded, the transmission law

$$T(x, y, k_x, k_y) = \begin{cases} 1 & \text{if} \quad k_z > k_s \\ 0 & \text{if} \quad k_z \leqq k_s \end{cases} \tag{10}$$

is often acceptable. Equation (10) says that rays whose axial wave number is less than the free wave number in the surrounding medium are leaking sufficiently rapidly to be ignored. The distribution f of the lambertian source is set equal to $1/(2\pi)^2$ so that the luminance is unity. In that case, the total power transmitted is the acceptance of the fiber. The relative time of flight is, within the present assumptions, solely a function of k_z. The upper and lower bounds on k_z are $k(x, y)$ and k_s, respectively. It remains to express the volume element $dk_z dk_y dx dy$ in (3) as a function of dk_z, dx, dy. For given x, y, a constant value of k_z corresponds to a circle of radius squared $k^2(x, y) - k_z^2$ in the k_x, k_y space because $k_x^2 + k_y^2 = k^2(x, y) - k_z^2$. Thus,

$$dk_x dk_y = \pi dk_z^2. \tag{11}$$

Let us evaluate the acceptance of the fiber. The light acceptance of any optical system is, as we have seen, the volume in phase space of the accepted rays divided by $(2\pi)^2$. It is also equal to the effective number of modes that the system can transmit. If we integrate $P(t)$

from $t = -\infty$ to $t = +\infty$ in (3), the integral over P' in the integrand is unity, and we obtain

$$N^2 = (1/4\pi) \int\int [k^2(x, y) - k_s^2] dx dy, \tag{12}$$

where we have used (10) and (11). Thus, $4\pi N^2$ is the volume enclosed by the profile: $k^2(x, y)$. For a step-index fiber of any shape with cross-section area A, for example, we have from (12)

$$N^2 = (A/4\pi)(k^2 - k_s^2). \tag{13}$$

This expression should be multiplied by 2 to take into account the two states of polarization.

The pulse transformation in (3) becomes, using (11),

$$P(t) = (1/4\pi) \int dx dy \int_{k_s^2}^{k^2(x,y)} P'[t - t(k_z)] dk_z^2. \tag{14}$$

Let the input pulse $P'(t)$ be a symbolic δ function (e.g., a rectangular pulse of width Δt and height Δt^{-1} in the limit $\Delta t \to 0$). The output pulse in (14) becomes

$$P(t) = (1/4\pi) |dk_z^2/dt| A(k_z), \qquad k_z > k_s, \tag{15}$$

where $|dk_z^2/dt|$ denotes the absolute value of dk_z^2/dt and $A(k_z)$ denotes the cross-section area that satisfies $k(x, y) > k_z$. k_z can be expressed as a function of the delay t by inverting the relation $t(k_z)$ given earlier. We obtain, from (8),

$$dk_z'^2/dt = 2(1 + \kappa)k_z'[1 - (\kappa/k_z'^2)], \tag{16}$$

where

$$k_z' = k_z/k_0 = (1 + \kappa)t/2 \pm \{[(1 + \kappa)t/2]^2 - \kappa\}^{\frac{1}{2}}. \tag{17}$$

If $\kappa > 1$, there is only one value of k_z' between $k_s' \equiv k_s/k_0$ and 1, for any k_s'. If

$$k_s'^2 < \kappa < 1, \tag{18}$$

there are two values of k_z' that need be considered. Their contributions to P should be added. If $\kappa < k_s'^2$, there is again only one relevant value of k_z'.

Let us consider as an example a (noncircularly symmetric) square-law medium

$$k^2(x, y) = k_0^2(1 - \Omega_x^2 x^2 - \Omega_y^2 y^2), \tag{19}$$

where Ω_x, Ω_y denote arbitrary constants. $2\pi/\Omega_x$ and $2\pi/\Omega_y$, for small x, y, are the periods of ray oscillation in the xz and yz planes, respectively. The area $A(k_z)$ defined earlier is the interior of an ellipse

$$A(k_z) = \pi(1 - k_z'^2)/\Omega_x\Omega_y. \tag{20}$$

The impulse response is obtained from (15) and (16) with $\kappa = 1$, and (20),

$$P(t) = \begin{cases} k_0^2 k_z'^3/(\Omega_x \Omega_y), & k_z > k_s \\ 0 \quad , & k_z \leqq k_s, \end{cases} \tag{21}$$

where, from (17), $k_z' = t - (t^2 - 1)^{\frac{1}{2}}$. Because, in most fibers, k_z remains close to k_0, the variation of k_z' can be neglected, and the pulse response is almost rectangular.

For a step-index fiber, the area A is the area of the core cross section, and $t = k_0/k_z$. Thus, the impulse response of a step-index fiber with cross-section area A is simply

$$P(t) = k_0^2 A/2\pi t^3, \qquad 1 < t < k_0/k_s. \tag{22}$$

Because, in most fibers, t remains close to unity, the pulse response is almost rectangular.[2] The pulse width, however, is considerably larger than for square-law fibers, as we shall see in more detail later.

IV. CIRCULARLY SYMMETRIC FIBERS WITH $k^2 - k_0^2$ A POWER OF THE RADIUS

Let the wave-number profile be of the form

$$k^2(R, \omega) = k_0^2(\omega) - k_\kappa^2(\omega) R^\kappa, \tag{23}$$

where $R \equiv X + Y \equiv r^2$ denotes the square of the radius. The relative time of flight is, substituting (23) in (2a),

$$\begin{aligned} t &= (k_0/k_z)\langle \partial k^2/\partial \omega^2 \rangle / (dk_0^2/d\omega^2) \\ &= (k_0/k_z)(1 - \epsilon_\kappa D_\kappa \langle R^\kappa \rangle), \end{aligned} \tag{24}$$

where we have defined

$$\epsilon_\kappa \equiv k_\kappa^2/k_0^2 \tag{25}$$

$$D_\kappa \equiv k_0^2(dk_\kappa^2/d\omega^2)/k_\kappa^2(dk_0^2/d\omega^2). \tag{26}$$

D_κ is a dispersion factor equal to unity in the absence of dispersion. Thus, we need to evaluate $\langle R^\kappa \rangle$. It is interesting that we can do that without solving the ray equations. The quantity $\langle R^\kappa \rangle$ is, of course, independent of dispersion, so we may omit the ω arguments.

For circularly symmetric fibers, (1) can be written

$$\tfrac{1}{2} k_z^2 d^2R/dz^2 = d(k^2R)/dR - k_z^2. \tag{27}$$

Averaging (27) over a ray period, we obtain

$$k_z^2 = \langle d(k^2R)/dR \rangle. \tag{28}$$

We have also, directly from (23),

$$\langle k^2 \rangle = k_0^2 - k_\kappa^2 \langle R^\kappa \rangle \tag{29a}$$

and, from (28) and (23),

$$\langle k^2 \rangle = (k_z^2 + \kappa k_0^2)/(1 + \kappa). \tag{29b}$$

Equating the two expressions (29a) and (29b) for $\langle k^2 \rangle$, we obtain

$$\epsilon_\kappa \langle R^\kappa \rangle = (1 - k_z'^2)/(1 + \kappa), \tag{30}$$

where $k_z' \equiv k_z/k_0$. Thus, substituting $\langle R^\kappa \rangle$ from (30) into (24), the relative time of flight is

$$t = k_z'^{-1} - D_\kappa(k_z'^{-1} - k_z')/(1 + \kappa). \tag{31}$$

In applications, we need k_z' as a function of t. Solving (31) for k_z' and setting $D_\kappa' \equiv D_\kappa/(1 + \kappa)$, we obtain

$$k_z' = (t/2D_\kappa') \pm [(t/2D_\kappa')^2 + 1 - D_\kappa'^{-1}]^{\frac{1}{2}}. \tag{32}$$

By differentiating (32), we further obtain

$$dk_z'^2/dt = 2k_z'[D_\kappa' - (1 - D_\kappa')/k_z'^2]^{-1}. \tag{33}$$

To obtain explicitly the impulse response in (15), we need the area $A(k_z)$ defined by $k_z < k(R)$. For $k(R)$ in (23), this area is

$$A(k_z) = \pi R(k_z) = \pi[(1 - k_z'^2)/\epsilon_\kappa]^{1/\kappa}. \tag{34}$$

If ϵ_κ were kept a constant as the parameter κ varies, the core radius a, defined by $k(a) = k_s$, would vary. Thus, it is preferable to express ϵ_κ as a function of the core radius a. We have

$$\epsilon_\kappa^{1/\kappa} = (1 - k_s'^2)^{1/\kappa}/a^2, \tag{35}$$

where $k_s' \equiv k_s/k_0$. The impulse response is finally obtained from (15), (33), (34), and (35);

$$P(t) = (k_0^2 a^2/2)k_z'[(1 - k_z'^2)/(1 - k_s'^2)]^{1/\kappa}/[D_\kappa' - (1 - D_\kappa')k_z'^{-2}]. \tag{36}$$

The possibly doubled value k_z' is expressed as a function of t by (32). Thus, a closed-form expression has been obtained for the impulse response of a fiber with $k^2 - k_0^2$ a power of r, that takes inhomogeneous dispersion into account.

In the absence of dispersion, we have $D_\kappa' = 1/(1 + \kappa)$, and (36) reduces to

$$P(t) = (k_0^2 a^2/2)k_z'[(1 - k_z'^2)/(1 - k_s^2)]^{1/\kappa}(1 + \kappa)/(1 - \kappa k_z'^{-2}). \tag{37}$$

As indicated in the previous section, there are in general two values of k_z' that contribute to P. Note that the shape of the impulse response does not depend on the core radius a.

The impulse response $P(t)$ in (37) is shown in Figs. 3 and 4 for various values of the parameter κ. These curves are essentially the

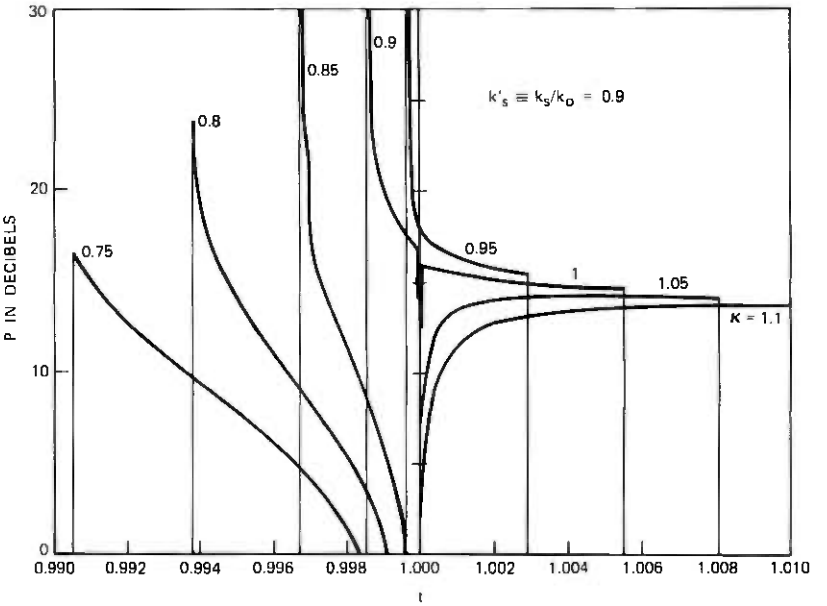


Fig. 3—Impulse response of a circularly symmetric fiber with $k^2(r) = k_0^2 - k_s^2 r^{2\kappa}$ for a lambertian source and various values of κ. The optimum impulse response is for $\kappa \approx k_s' = 0.9$.

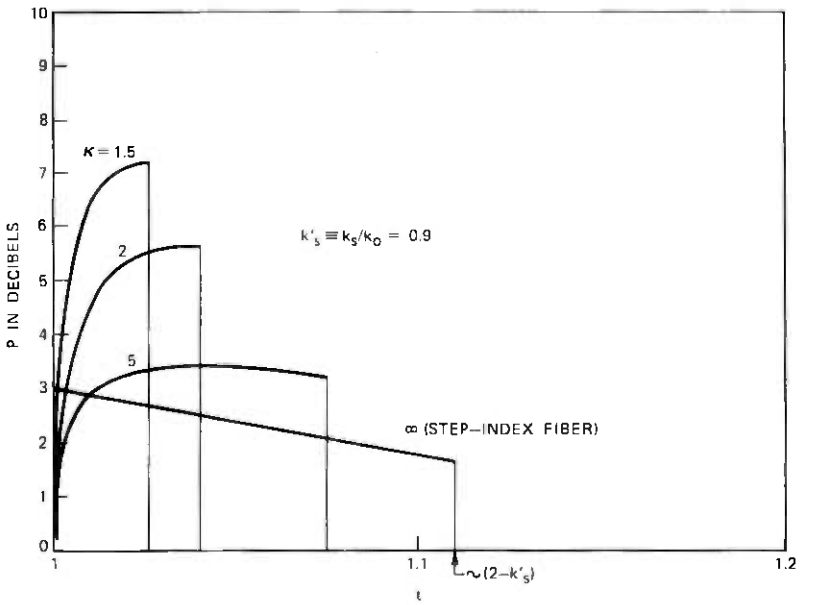


Fig. 4—Continuation of Fig. 3 for larger values of κ.

same as those shown in Ref. 4. Figures 3 and 4, however, are much more detailed. We have assumed that $k_s/k_0 = 0.9$, that is, $\Delta n/n = 10\%$, a rather large value. For $\kappa = 1$ (square-law fiber), the pulse width τ is 0.0054. For example, if $n = 1.45$ and the fiber length is 1 km, the pulse width is 26 ns. For $\kappa = 0.9$, however, the corresponding pulse width is only 7 ns. We find, in agreement with Ref. 4, that the minimum pulse width occurs when $\kappa = k_z'$. For a step-index fiber ($\kappa \to \infty$), the pulse width would be as large as 630 ns. Note the following detailed features on the curves in Figs. 3 and 4. For $(0.9)^2 < \kappa < 1$, the response starts from infinity because of the minimum in the $t(k_z)$ curve. For $\kappa = 0.85$, P drops suddenly for $t \approx 0.998$. This is because, at that time, the smaller of the two k_z' becomes less than 0.9, and is rejected. For $\kappa = 0.95$, the response crosses the $t = 1$ axis.

Figure 4 is applicable to larger values of κ. We note that, for a very large κ (step-index fiber), the response is almost rectangular. The slow decay in power shown in Fig. 4 would be almost negligible for small $\Delta n/n$.

The effect of inhomogeneous dispersion is shown in Fig. 5. The parameter κ is kept equal to 0.9 (this is the optimum value in the absence of inhomogeneous dispersion), but D_κ is made to vary in the

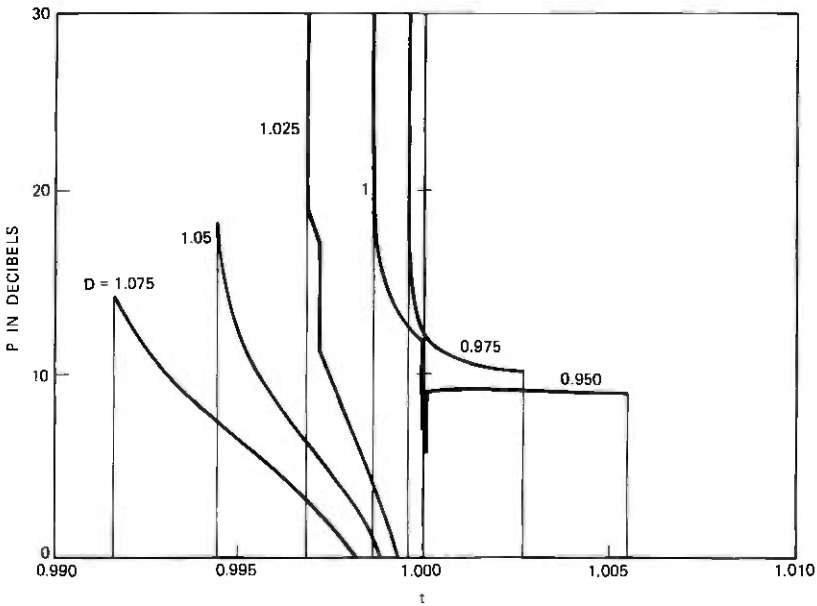


Fig. 5—Impulse response for a fiber with $k^2(r) = k_0^2 + k_\kappa^2(r^2)^{0.9}$ for various values of the parameter D that expresses the inhomogeneous dispersion of the material. $D = 1$ corresponds to the absence of dispersion. $D \neq 1$ merely introduces a shift in the optimum value of κ.

neighborhood of unity. These curves have a striking resemblance to those in Fig. 3. This means that the effect of inhomogeneous dispersion merely consists in shifting the optimum value of κ. The impulse response remains essentially the same, at least for $\kappa \approx 1$.

The total pulse power is the acceptance of the fiber, a function of κ. The acceptance is, in the present case,

$$\begin{aligned} N^2 &= \int_{-\infty}^{+\infty} P(t)dt = (\tfrac{1}{4}) \int_0^a [k^2(R) - k_s^2]dR \\ &= (\tfrac{1}{4})[(k_0^2 - k_s^2)a^2 - k_\kappa^2 a^2 (a^2)^\kappa/(\kappa + 1)] \\ &= [\kappa/4(\kappa + 1)](k_0^2 - k_s^2)a^2. \end{aligned} \tag{38}$$

The coefficient in the last expression in (38) is $\frac{1}{4}$ for step-index fibers ($\kappa \to \infty$) and $\frac{1}{8}$ for square-law fibers. The acceptance given in (38) should be multiplied by 2 to account for the two states of polarization. The same rule applies to all the expressions given in this paper. It is more difficult to obtain the ray-optics acceptance of fibers. The result is derived in Appendix B.

In the next section, we consider fibers whose profile departs slightly, but otherwise arbitrarily, from a square law.

V. NEAR-SQUARE-LAW FIBERS

Let us rewrite the differential equation (1) for circularly symmetric fibers

$$\tfrac{1}{2}k_z^2 d^2R/dz^2 = d(k^2R)/dR - k_z^2. \tag{39}$$

For square-law fibers

$$k^2(R) = k_0^2 - k_1^2 R, \tag{40}$$

the solution of (39) is

$$R(z) = R_0 + (R_0^2 - l_z^2/k_1^2)^{\frac{1}{2}} \cos (2\Omega z/k_z'), \tag{41}$$

where

$$R_0 \equiv \tfrac{1}{2}(k_0^2 - k_z^2)/k_1^2 \equiv \tfrac{1}{2}(1 - k_z'^2)/\Omega^2 \tag{42}$$

and $\Omega \equiv k_1/k_0$. We have introduced in (41) the axial component of the angular momentum (or Bouguer invariant)

$$l_z = xk_y - yk_x, \tag{43}$$

which is the second constant of motion. Let us set

$$\theta \equiv (l_z/k_1 R_M)^2, \tag{44}$$

where R_M denotes the maximum radius squared. Note that, for meridional rays, $\theta = 0$ and, for helical rays, $\theta = 1$. Equation (41) can be written in the convenient form

$$R = \tfrac{1}{2}R_M(1 + \theta) + \tfrac{1}{2}R_M(1 - \theta) \cos (2\Omega z/k_z'). \tag{45}$$

For later use let us evaluate $\langle R^n \rangle$, the average of R^n over a ray period. Using the binomial expansion and the result

$$\langle \cos^m \rangle = m!2^{-m}[(m/2)!]^{-2} \tag{46}$$

for m even and 0 for m odd, we obtain

$$\langle R^n \rangle = n!2^{-n}R_M^n \sum_{m=0,2\ldots}^{n} \frac{(1+\theta)^{n-m}(1-\theta)^m}{2^m(n-m)![(m/2)!]^2}. \tag{47}$$

In particular,

$$\langle R^2 \rangle = R_M^2(3\theta^2 + 2\theta + 3)/8 \tag{48a}$$

$$\langle R^3 \rangle = R_M^3(1+\theta)(5\theta^2 - 2\theta + 5)/16 \tag{48b}$$

$$\langle R^4 \rangle = R_M^4(35\theta^4 + 20\theta^3 + 18\theta^2 + 20\theta + 35)/128. \tag{48c}$$

Let us now show that a closed-form expression can be obtained for the times of flight in fibers whose permittivity profiles depart slightly from a square law. Inhomogeneous dispersion is taken into account. Let the profile be of the form

$$k^2(R) = k_0^2 - k_1^2 R + \sum_{n=2}^{N} k_n^2 R^n. \tag{49}$$

We assume that $\epsilon_n R^{n-1}$, $n \geqq 2$, is of the order $\epsilon \ll 1$, where $\epsilon_n \equiv k_n^2/k_0^2$.

Substituting (49) in (2a), we obtain (with $\Omega^2 \equiv \epsilon_1 \equiv k_1^2/k_0^2$)

$$t = k_z'^{-1}(1 - D_1\Omega^2\langle R \rangle + \sum_{n=2}^{N} \epsilon_n D_n \langle R^n \rangle), \tag{50}$$

where we have defined inhomogeneous dispersion factors

$$D_n = (k_0^2 dk_n^2/d\omega^2)/(k_n^2 dk_0^2/d\omega^2). \tag{51}$$

The D_n are unity in the absence of inhomogeneous dispersion. Because the perturbation is small, $\langle R^n \rangle$ in the sum (50) can be replaced by the expression (47) applicable to square-law fibers. This approximation is not permissible, however, for the term $\langle R \rangle$ in (50) because this term is not small. We need an exact expression for $\langle R \rangle$. We proceed as in the previous section. We first observe that, for k^2 in (49),

$$d(k^2R)/dR = 2k^2 - k_0^2 + \sum_{n=2}^{N} (n-1)k_n^2 R^n. \tag{52}$$

Integrating (39) over a ray period, the left-hand side vanishes and, using (52), we obtain an expression for $\langle k^2 \rangle$ that does not involve $\langle R \rangle$

$$\langle k^2 \rangle = \tfrac{1}{2}(k_z^2 + k_0^2) + \tfrac{1}{2} \sum_{n=2}^{N} (1-n)k_n^2 \langle R^n \rangle. \tag{53}$$

We also have, directly from (44),

$$\langle k^2 \rangle = k_0^2 - k_1^2 \langle R \rangle + \sum_{n=2}^{N} k_n^2 \langle R^n \rangle. \tag{54}$$

Thus, by comparing (53) and (54),

$$k_1^2 \langle R \rangle = \tfrac{1}{2}(k_0^2 - k_z^2) + \tfrac{1}{2} \sum_{n=2}^{N} (n+1) k_n^2 \langle R^n \rangle. \tag{55}$$

Substituting this expression for $\langle R \rangle$ in (50), we obtain the relative time of flight for circularly symmetric near-square-law fibers

$$t = k_z'^{-1} \{1 - \tfrac{1}{2}(1 - k_z'^2) D_1 + \sum_{n=2}^{N} [D_n - \tfrac{1}{2}(n+1) D_1] \epsilon_n \langle R^n \rangle\}. \tag{56}$$

Alternatively, t can be expressed in terms of the azimuthal and radial mode numbers. The result is given in Appendix C.

In the absence of inhomogeneous dispersion, (56) reduces to

$$t = \tfrac{1}{2} k_z'^{-1} [k_z'^2 + 1 + \sum_{n=2}^{N} (1 - n) \epsilon_n \langle R^n \rangle]. \tag{57}$$

Limiting ourselves to an r^4 correction to the square-law profile, $\epsilon_3 = \epsilon_4 = \cdots = 0$, and setting $\epsilon_2 \equiv \epsilon$, (57) becomes, using (48),

$$\begin{aligned} t &= \tfrac{1}{2}[1 - \rho_M(1+\theta)]^{-\frac{1}{2}}[2 - \rho_M(1+\theta) - \epsilon\rho_M(3\theta^2 + 2\theta + 3)/8] \\ &\approx 1 + \rho_M^2[(2 - 3\epsilon) + (4 - 2\epsilon)\theta + (2 - 3\epsilon)\theta^2]/16 + 0(\rho_M^3) \\ \rho_M &\equiv \Omega^2 R_M. \end{aligned} \tag{58}$$

The first two terms in (58) give sufficient accuracy when $\rho_M \lesssim 0.01$, that is, when the total relative change in refractive index $\Delta n/n \approx \rho_a/2$ is less than 0.005 ($\rho_a \equiv \Omega^2 a^2$).

The total pulse width τ is the maximum variation of t for $0 < \theta < 1$ and $0 < \rho < \rho_a$. For the square-law fiber [$\epsilon = 0$ in (58)], we obtain

$$\tau = \rho_a^2/2 \qquad \text{(ray optics)}. \tag{59}$$

It should be noted that, in defining τ in (59), we have specified that the maximum radius of the ray be less than a for any θ. This condition is different from the condition used earlier that k_z be larger than k_s. The ray-optics condition $\rho_M < \rho_a$ is applicable to highly oversized fibers.

If we now consider the expression in (58) with a correction term in r^4, we find that $t = 1$ for meridional rays ($\theta = 0$) when $\epsilon_2 = \frac{2}{3}$ in agreement with Ref. 17, where it is shown that all the rays have exactly the same optical length when $k^2(x) = [\cosh(x)]^{-2} \approx 1 - x^2$

$+(\frac{2}{3})x^4 + \cdots$. We also find that $t = 1$ for helical rays ($\theta = 1$) when $\epsilon_2 = 1$, in agreement with Ref. 3, where it is shown that helical rays have exactly the same optical length when $k^2(r) = (1 + r^2)^{-1} \approx 1 - r^2 + r^4 + \cdots$. By considering all rays whose maximum radius is less than a, we find that the minimum τ is obtained for $\epsilon_2 = 0.91$. In that case, $\tau = 0.046\rho_a^2$. The improvement compared with square-law media is therefore as large as 11. If we had imposed instead the wave-optics condition $k_z > k_s$, the vertical scale in Fig. 6 would be divided by $(1 + \theta)^2$. For $\epsilon_2 = 0$, for example, the wave-optics pulse width is $\rho_a^2/8$ instead of $\rho_a^2/2$ as in (59). With the wave-optics limit, the optimum value of ϵ_2 turns out to be $\frac{2}{3}$ instead of 0.91. The improvement over the square-law case is only 4, instead of 11.

Let us now consider the effect of r^6 terms. Figure 7 shows the variation of the pulse width τ, defined as the maximum variation of t for any $0 < \theta < 1$ and any $0 < \rho_M < 0.002$, as a function of ϵ_2 for various values of ϵ_3. The effect of ϵ_3 is essentially to shift the optimum value of ϵ_2 to lower values. The reduction in pulse width is rather modest. Nevertheless, a small improvement is obtained, compared to the case

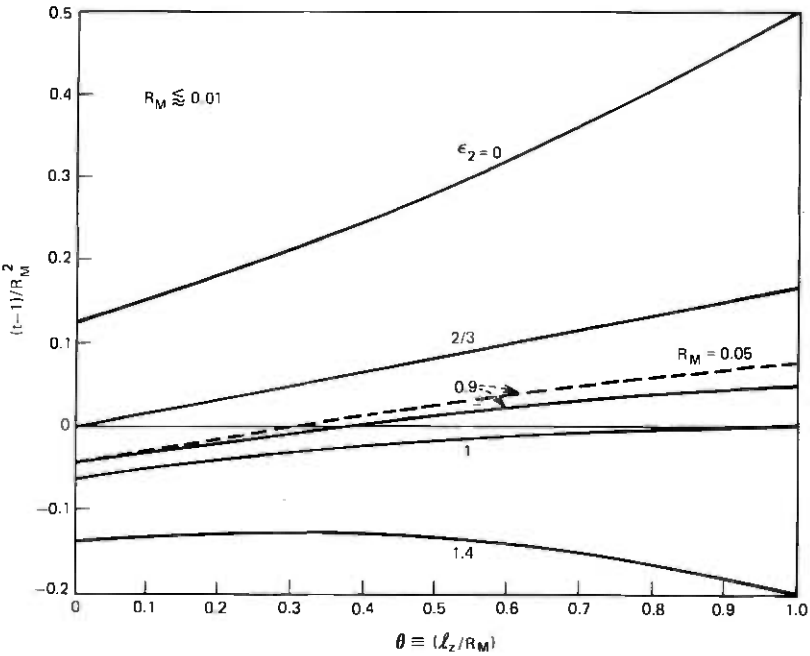


Fig. 6—Variation of the normalized time of flight for a fiber with $k^2(r) = k_0^2 - k_1^2 r^2 + \epsilon_2(k_1^4/k_0^2)r^4$ in the absence of material dispersion for various values of the parameter ϵ_2. $\theta = 0$ corresponds to meridional rays and $\theta = 1$ to helical rays. ϵ_2 has been redefined to be dimensionless.

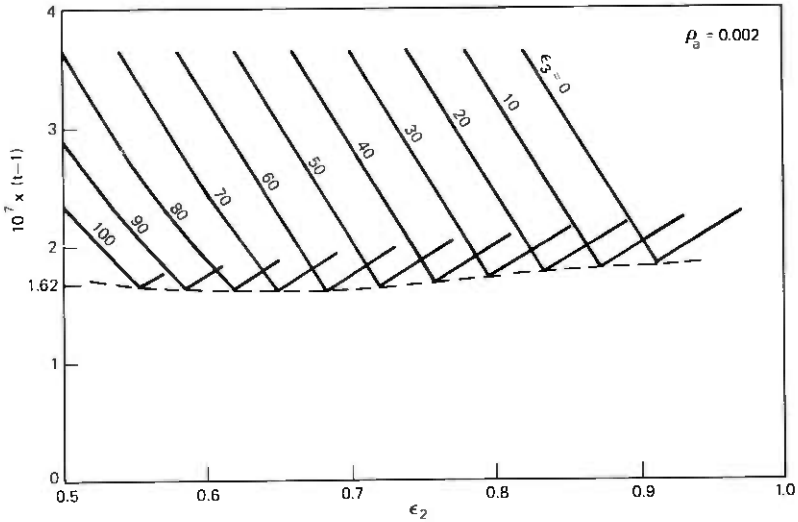


Fig. 7—Variation of the width of the impulse response with ϵ_2 for various values of ϵ_3 for a fiber with $k^2(r) = k_0^2 - k_1^2 r^2 + \epsilon_2(k_1^4/k_0^2)r^4 + \epsilon_3(k_1^6/k_0^4)r^6$.

where $\epsilon_3 = 0$ when

$$k_0^{-2}k^2(r) = 1 - \rho + 0.615\rho^2 + 70\rho^3, \qquad \rho_a = 0.002 \qquad (60)$$
$$\rho \equiv \Omega^2 r^2.$$

We give only the result when the departures from a square-law profile are not circularly symmetric. The free wave number in the fiber is now in the form

$$k^2(x, y, \omega) = k_0^2(\omega) - k_1^2(\omega)R + \sum_{n=1}^{N} \sum_{l=0}^{n} k_{nl}^2(\omega) X^l Y^{n-l}, \qquad (61)$$

where, as before, $X \equiv x^2$, $Y \equiv y^2$, $R \equiv x^2 + y^2 \equiv r^2$. The ratio, t, of the time of flight along a ray to the corresponding time on-axis is found to be

$$t = k_z'^{-1}\{1 - \tfrac{1}{2}(1 - k_z'^2)D_1 + \sum_{n=1}^{N} \sum_{l=0}^{n} [D_{nl} - \tfrac{1}{2}(n + 1)D_1] \times \epsilon_{nl}\langle X^l Y^{n-l}\rangle\}, \qquad (62)$$

where $\epsilon_{nl} \equiv k_{nl}^2/k_0^2$ and D_{nl} is defined as D_n in (51) with k_n replaced by k_{nl}. Let us assume that it is permissible to use the sinusoidal rays of the square-law medium to evaluate the quantity $\langle X^l Y^{n-l}\rangle$. Because the average over one cycle of the product of powers of sinusoidal functions is known, the relative delay t can be written in closed form.

Let the ray trajectory be written

$$x(z) = x_0 \cos(\alpha z + \phi_x) \tag{63a}$$

$$y(z) = y_0 \cos(\alpha z + \phi_y). \tag{63b}$$

The coefficient α does not enter in the final result and is henceforth omitted. We evaluate

$$\begin{aligned}\langle X^l Y^{n-l} \rangle &\equiv \langle x^{2l} y^{2(n-l)} \rangle \\ &= x_0^{2l} y_0^{2(n-l)} \langle [\cos(z + \phi_x)]^{2l} [\cos(z + \phi_y)]^{2(n-l)} \rangle. \end{aligned} \tag{64}$$

It can be shown that[18]

$$\begin{aligned}&\langle [\cos(z + \phi_x)]^{2l} [\cos(z + \phi_y)]^{2(n-l)} \rangle \\ &\quad = 2^{-2n} \left\{ 2 \sum_{s=1}^{n-l} \binom{2(n-l)}{n-l-s} \binom{2l}{l-s} \cos[2s(\phi_x - \phi_y)] \right. \\ &\qquad \left. + \binom{2(n-l)}{n-l} \binom{2l}{l} \right\}, \end{aligned} \tag{65}$$

where

$$\binom{a}{b} \equiv \frac{a!}{(a-b)!b!}. \tag{66}$$

Thus, given a ray trajectory, defined by the parameters x_0, y_0, ϕ_x, and ϕ_y (or, equivalently, by the values of x, y, k_x, and k_y at the input of the fiber), we can evaluate in closed form the quantity $\langle X^l Y^{n-l} \rangle$ that enters in formula (62) for the relative time of flight, from (64) to (66).

The above calculation is incomplete for the following reasons. When the power law $n^2(r)$ of a fiber departs from the exact square law, projected ray trajectories in the (xy) transverse plane are *precessing* ellipses.* That is, the principal axes of the near-elliptical trajectories slowly rotate as a function of z. This precession is unimportant for circularly symmetric fibers. For noncircularly symmetric fibers, however, the ellipse precession introduces an averaging effect. Furthermore, the noncircularly symmetric components of $n^2(r, \varphi)$ change the *eccentricity* of the precessing ellipse. The axial component k_z of the wave vector remains a constant, but the axial component l_z of the angular momentum varies. Finally, in real fibers, slow (adiabatic) changes of the refractive index law along the fiber axis are likely to occur that must be taken into account. The twists of the fiber axis must also be taken into account. Thus, a realistic assessment of the effect of small noncircularly symmetric departures of the index law

* It is well known in mechanics that the only $r^{2\epsilon}$ potentials (potential $U \sim n^2$) that give closed trajectories are the harmonic potential $U(r) \sim n^2(r) = 1 - r^2$ and the Newton potential $U(r) \sim n^2(r) = 1/r$.

from square law on pulse broadening requires a deeper and more intricate analysis than the one given in the present section. However, the result in (62) and (65) can be used as a basis for more complete analyses.

VI. CONCLUSION

From a rather straightforward application of the Hamilton ray equations, we have obtained closed-form expressions for the pulse width in graded-index fibers when $k^2(x, y) - k_0^2$ is a homogeneous function of x and y, and for fibers whose profile departs slightly, but otherwise arbitrarily, from a square law. Inhomogeneous dispersion was taken into account. The expressions obtained are exact. The small angle (or weakly guiding) approximation need not be made. We have also given simple expressions for the wave optics and ray optics acceptance of weakly guiding graded-index fibers.

The algebraic results given should prove more accurate and require much less computer time than the straightforward numerical integration of time along ray trajectories. We have carried the perturbation only to first order in the small parameter ϵ. To obtain more accurate results, up to order ϵ^2, we need a more accurate expression of the ray trajectory. This expression can be obtained, for example, by the method of strained coordinates.[19] These more accurate expressions are probably not needed, however, in most practical cases.

VII. ACKNOWLEDGMENT

The author expresses his thanks to E. A. J. Marcatili for useful comments.

APPENDIX A

The Hamilton Equations of Ray Optics

The Hamiltonian form of the ray equations are well known in mechanics and wave dynamics,[20] and they have also been used frequently in optics (e.g., Refs. 14, 15, 17, 21, and 22). However, their simplicity and power is not always appreciated. The physical difficulty is that it is not always recognized that ray momenta and wave vectors (or photon momenta) are identical concepts. On the other hand, ray momenta (proportional to the wave vectors) need be carefully distinguished from mass-carrying momenta (proportional to the group velocities).[23] On the mathematical side, we need distinguish a function such as $k_z(x, y)$ and the value k_z assumed by that function. We must also be aware that da/dz denotes a total derivative, that is, in the present context, the variation of the quantity a along some given ray.

If a is a known function of x and y, and $x = x(z)$, $y = y(z)$ denote a known ray trajectory, then $da/dz = (\partial a/\partial x)(dx/dz) + (\partial a/\partial y)(dy/dz)$ can be evaluated explicitly as a function of z. Here again, an arbitrary point in space x, y should not be confused with a specific ray trajectory $x = x(z)$, $y = y(z)$. Unfortunately, it is not possible to go into more details here. An excellent reference is Lighthill's paper.[20] A comparison between the W.K.B. method and the Hamilton equations is given in Ref. 14.

Let $\mathbf{X} \equiv (x, y, z, ict)$ denote a point in space-time (t is time) and $\mathbf{K} \equiv (k_x, k_y, k_z, i\omega/c)$ denote the four-wave vector, with ω the angular frequency. An arbitrary medium is characterized by a function of $\mathbf{K}$ and $\mathbf{X}$ that we denote

$$H(\mathbf{K}, \mathbf{X}) = 0. \tag{67}$$

The Hamilton equations for light pulses $\mathbf{X}(\sigma)$, $\mathbf{K}(\sigma)$ are

$$d\mathbf{X}/d\sigma = \partial H/\partial \mathbf{K} \tag{68a}$$

$$d\mathbf{K}/d\sigma = -\partial H/\partial \mathbf{X}, \tag{68b}$$

where σ denotes an arbitrary parameter.

Equations (68a) and (68b) can be considered the basic postulates of geometrical optics. From a wave-optics point of view, (68a) follows from the requirement that the wave lengths and periods of the waves that constitute a wave packet be the same in the direction of a ray. Equation (68b) follows from (67), (68a), and the fact that $\mathbf{K}$ is the gradient of an eikonal function. Thus, in wave optics, the Hamilton equations (68) are derived from first principles and need not be postulated.

Let ξ denote a point in phase space $(k_x, k_y, \omega, x, y, t)$ at the input plane, and ξ' a point in phase space at the output plane. The optical system maps the input phase space into the output phase space, that is,

$$\xi = \xi(\xi'). \tag{69}$$

It follows from (67) and (68) that the Jacobian of the transformation (69) is unity, a result often used in photometry. Equivalently, we can say that the determinant of paraxial ray matrices is unity or that the ray density in phase space is a constant of motion (Liouville theorem). These three statements are obviously equivalent, provided the rays are not reflected.

A source of light that is time and space incoherent is described by a distribution $S(\xi)$ in phase space. Each small volume in phase space can be pictured as an optical pulse, provided the dimensions of the volume are larger than unity. More precisely, this picture requires that $\Delta\omega\Delta t \gg 1$, $\Delta k_x \Delta x \gg 1$, and $\Delta k_y \Delta y \gg 1$. The detailed structure

of the pulse is ignored in ray optics. Only the motion of the center is considered.

The transmission T_1 of an optical pulse through the optical system is a presumably known function of ξ that we denote as

$$T_1 = T_1(\xi). \tag{70}$$

For lossy optical systems, $T_1 < 1$. Because the Jacobian of the transformation $\xi \to \xi'$ is unity, the output distribution is simply

$$S'(\xi') = S(\xi)T_1(\xi). \tag{71}$$

The power emitted by the source and the power that can be collected at the output of the optical system are obtained by integrating S (or S') over all variables, except t (or t'). Thus,

$$P(t) = \int S(\xi)(d\zeta) \tag{72a}$$

$$P'(t') = \int S'(\xi')(d\zeta'), \tag{72b}$$

where $\zeta \equiv (k_x, k_y, \omega, x, y)$ and ζ' is similarly defined. The terms $(d\zeta)$ and $(d\zeta')$ denote elementary volumes in ζ and ζ' spaces, respectively. The response of the detector could be described by a function $D(\xi')$. For simplicity, we do not take the detector response into consideration. All subsequent results follow in a rather straightforward manner from the above results, through a succession of approximations.

Let us assume that the properties of the fiber do not vary with time. This means that the Hamiltonian in (67), the transmission T_1, the mapping $\xi \to \xi'$, and the pulse delay do not depend on time. In particular,

$$t' = t + t_1(\zeta). \tag{73}$$

Sources that are t-separable, on the other hand, have the property that

$$S(\xi) = P(t)F(\zeta). \tag{74}$$

That is, the distribution in ζ-space does not vary with time. For a hot tungsten wire whose temperature varies with time, the spatial phase-space distribution is almost lambertian at all times, but the frequency spectrum (approximately given by the Plank law of radiation) varies with time. Thus, (74) is not applicable to that source. For consistency with (72), we assume that $F(\zeta)$ is normalized to unity.

For most sources, we can further assume that

$$F(\zeta) = \Omega(\omega)f(\mathbf{p}), \tag{75}$$

where $\mathbf{p} \equiv (k_x, k_y, x, y)$ denotes a point in spatial phase-space. That is, we assume that the spatial distribution does not depend on what part of the frequency spectrum we are considering. Both Ω and f are assumed normalized to unity. This ensures that F is normalized to unity. When the spectral width of the source is small (e.g., less than 1 percent, as is the case for light-emitting diodes) and the fiber material absorption does not exhibit sharp resonances in that band, we can assume that

$$T_1(\zeta) = T_0(\omega)T(\mathbf{p}) \tag{76}$$

and

$$t_1(\zeta) = t_0(\omega) + t(\mathbf{p}). \tag{77}$$

For definiteness, we assume that the maximum value of $T_0(\omega)$ is unity, and we define $t_0(\omega)$ as the delay experienced by axial pulses. We evaluate in the main text $t(\mathbf{p})/t_0$ at a fixed angular frequency.

The pulse response is obtained from (71) to (77),

$$\begin{aligned} P'(t') &= \int P[t' - t_0(\omega) - t(\mathbf{p})]\Omega(\omega)T_0(\omega)f(\mathbf{p})T(\mathbf{p})(d\mathbf{p})d\omega \\ &= \int P''[t' - t_0(\omega)]\Omega(\omega)T_0(\omega)d\omega, \end{aligned} \tag{78}$$

where

$$P''(t'') = \int P[t'' - t(\mathbf{p})]f(\mathbf{p})T(\mathbf{p})(d\mathbf{p}). \tag{79}$$

In writing (78) we have used the fact that the Jacobian of the transformation $\xi \to \xi'$ is unity and that $d\omega = d\omega'$. The pulse response is the convolution of the pulse response in (79), which we may call the quasi-monochromatic pulse response, and the spectral width of the source. In most cases, $T_0(\omega)$ is a constant. For injection lasers, the quasi-monochromatic pulse response is the most important contribution.

In what follows, we assume that the fiber is uniform and long compared with the period of ray oscillation and therefore approximately z-invariant. Let the Hamiltonian in (67) be written

$$H = k_z - k_z(k_x, k_y, \omega, x, y) = 0. \tag{80}$$

The Hamilton equations (68) are

$$dx/dz = -\partial k_z/\partial k_x \tag{81a}$$

$$dy/dz = -\partial k_z/\partial k_y \tag{81b}$$

$$dt/dz = \partial k_z/\partial \omega \tag{81c}$$

$$dk_x/dz = \partial k_z/\partial x \tag{81d}$$

$$dk_y/dz = \partial k_z/\partial y. \tag{81e}$$

Let us assume further that the medium is isotropic, that is,

$$k_z^2 = k^2(\omega, x, y) - k_x^2 - k_y^2. \tag{82}$$

Thus, (81a) to (81e) are

$$dx/dz = k_x/k_z \tag{83a}$$

$$dy/dz = k_y/k_z \tag{83b}$$

$$dt/dz = (\partial k^2/\partial\omega)/2k_z \tag{83c}$$

$$dk_x/dz = (\partial k^2/\partial x)/2k_z \tag{83d}$$

$$dk_y/dz = (\partial k^2/\partial y)/2k_z. \tag{83e}$$

According to (83c), the time of flight of a pulse along a ray for a period (period $\equiv Z$) is obtained by integrating $(\partial k^2/\partial\omega)/2k_z$ from $z = 0$ to $z = Z$. If $k_0(\omega) \equiv k(\omega, 0, 0)$ denotes the wave number on axis, the time of flight of a pulse along the z axis is similarly obtained by integrating $(\partial k_0^2/\partial\omega)/2k_0$. Thus, the ratio of the time of flight of a pulse along a ray to the corresponding time on axis is

$$t = (k_0/k_z)\langle\partial k^2/\partial\omega^2\rangle/(dk_0^2/d\omega^2), \tag{84}$$

where $\langle\ \rangle$ denotes an average over a ray period. If the trajectory is not periodic, $\langle\ \rangle$ is understood as a limit for $z \to \infty$. When k is proportional to ω, (84) reduces to

$$t = \langle k^2\rangle/k_0k_z. \tag{85}$$

Let us now observe that, from (83a), (83b), (83d), and (83e),

$$\tfrac{1}{2}k_z^2d^2(X + Y)/dz^2 = k^2 - k_z^2 + X\partial k^2/\partial X + Y\partial k^2/\partial Y, \tag{86}$$

where $X \equiv x^2$, $Y \equiv y^2$. This is easily verified by carrying out the differentiations. Equations (86), (84), and (79) (with a slightly different notation) are those used in the main text.

APPENDIX B

Acceptance of Highly Oversized Fibers

The acceptance, or effective number of modes transmitted by the optical system, is the volume of the accepted rays in phase space divided by $(2\pi)^2$. We have said earlier that, if the fiber is very long, all leaky rays are eliminated and the acceptance is simply the volume enclosed by the profile $k^2(x, y)$ divided by 4π. If the fiber is highly oversized, however, many leaky rays ($k_z < k_s$) are not significantly attenuated.[16] We need then consider the ray-optics condition that the tangential (rather than axial) component of the wave vector be larger than k_s at the core-cladding interface. The ray-optics acceptance is

now evaluated for circularly symmetric fibers. We specify that

$$k_z^2 + k_\varphi^2 > k_s^2, \qquad \text{at} \quad r = a, \tag{87a}$$

where k_φ denotes the azimuthal wave number at the interface. We also have the condition

$$k_z^2 > 0, \tag{87b}$$

which is not implied by (87a). In this appendix, we restrict ourselves to small differences in refractive index, in which case condition (87b) can be ignored. Because of the conservation of l_z (the axial component of the angular momentum), we have

$$rk_y = ak_\varphi \tag{88}$$

for a ray with $x = r$, $y = 0$, k_x, k_y, at the input plane, that can reach the interface $r = a$. Thus, condition (87a) is

$$k^2(r) - k_x^2 - k_y^2 + (r^2/a^2)k_y^2 > k_s^2. \tag{89}$$

Equation (89) defines an area in the k_x, k_y plane bounded by an ellipse. We have to make sure, however, that rays outside that area do in fact reach the interface. This is not necessarily the case. The maximum ray radius r_M is defined implicitly by

$$k_x^2 + (1 - r^2/r_M^2)k_y^2 = k^2(r) - k^2(r_M), \tag{90}$$

where r_M is the largest real number that satisfies (90). (The initial radius r is considered a constant in the present discussion.) Equation (90) shows that the k_x, k_y that correspond to r_M are contained in an ellipse with semi-axes squared $k_{x0}^2 = k^2(r) - k^2(r_M)$ and $k_{y0}^2 = [k^2(r) - k^2(r_M)]/(1 - r^2/r_M^2)$, respectively. If $k^2(r)$ is never increasing, we are sure that k_{x0} keeps increasing as r_M increases from r to a. We do not have any such assurance for k_{y0}, however. When r_M reaches a, there may be acceptable values of k_x, k_y that are located outside the ellipse defined above. For each profile, we need therefore verify that $k_{y0}^2(r_M)$ never exceeds $k_{y0}^2(a)$. We easily verify that this is the case for square-law fibers, because

$$k_{y0}^2 = k_1^2(r_M^2 - r^2)/(1 - r^2/r_M^2) = k_1^2 r_M^2 \tag{91}$$

increases with r_M for any r.

Thus, for square-law fibers at least, we can proceed with the calculation of the area of the ellipse defined by (89). This area is

$$\pi[k^2(r) - k_s^2](1 - r^2/a^2)^{-\frac{1}{2}}. \tag{92}$$

Substituting this result in the general expression for the acceptance

factor, we obtain

$$N^2 = (\tfrac{1}{4}) \int_0^{a^2} [k^2(r) - k_s^2](1 - r^2/a^2)^{-\frac{1}{2}} dr^2. \tag{93}$$

This expression simplifies if we introduce the variable $u \equiv (1 - r^2/a^2)^{\frac{1}{2}}$. Equation (93) becomes

$$N^2 = (a^2/2) \int_0^1 [k^2(u) - k_s^2] du. \tag{94}$$

Thus, the ray-optics acceptance of most circularly symmetric fibers is half the area enclosed by the curve $k^2(u)a^2$. For a step-index fiber, we obtain from (94)

$$N^2 = (k_0^2 - k_s^2)a^2/2 \quad \text{(step-index, ray optics)}. \tag{95}$$

This is twice the wave-optics acceptance. Thus, for step-index fibers, the slightly leaky rays carry half the power. Our result agrees with that in Ref. 16 for weakly guiding fibers. For a square-law fiber, with $k(a) = k_s$, we obtain

$$N^2 = (k_0^2 - k_s^2)a^2/6 \quad \text{(square-law, ray optics)}. \tag{96}$$

In square-law fibers, 25 percent of the total power is carried by slightly leaky rays.*

APPENDIX C

Impulse response width of near-square law fibers

When the source distribution is lambertian, all propagating modes are equally excited. It is convenient in that case to express the relative time of flight t for near-square-law fibers given in (56) as a function of the mode numbers (azimuthal number: $\mu = \cdots -2, -1, 0, 1, 2 \cdots$ and radial number: $\alpha = 0, 1, 2 \cdots$) rather than k_z and l_z. This can be done by quantizing the ray trajectories. [If the W.K.B. method is used, it is essential to first remove the singularity of the Helmholtz equation at $r = 0$. This is achieved by changing the independent variable from r to $\log(r)$.] One easily finds that the axial component of the ray angular momentum l_z is equal to μ. Furthermore, we can use for k_z the well-known expression applicable to square-law media. The result (56) is written below as a function of α, μ, for the reader's convenience. We have

$$t(\alpha, \mu) = (1 - B)^{-\frac{1}{2}} \left(1 - \tfrac{1}{2}BD_1 + \sum_{\gamma=2}^{\infty} F_\gamma N_\gamma\right), \tag{97}$$

* This is in agreement with a recent result by D. N. Payne.

where

$$B \equiv 2gK_1^{1/2}/K_0$$
$$g \equiv 2\alpha + |\mu| + 1; \quad |\mu| \equiv \text{abs. val. } (\mu)$$
$$F_\gamma \equiv \gamma!2^{-\gamma}[D_\gamma - \tfrac{1}{2}(\gamma + 1)D_1]K_\gamma/(K_0K_1^{\gamma/2})$$

$$N_\gamma \equiv (2g)^\gamma \sum_{m=0,2}^{\gamma} \{2^m(\gamma - m)![(m/2)!]^2\}^{-1}[1 - (\mu/g)^2]^{m/2}. \quad (98)$$

The parameters $K_\gamma \equiv k_\gamma^2$, $\gamma = 0, 1 \cdots$ and D_γ, $\gamma = 1, 2 \cdots$ are obtained from the square of the wave number: $K(R) \equiv k^2(R) \equiv (\omega/c)^2 n^2(R)$ of the fiber as a function of $R \equiv r^2$, measured at the nominal wavelength λ_0 and at a slightly different wavelength, λ_0', expanded in power series of R as follows

$$K(R) = K_0 - K_1R + K_2R^2 + \cdots \quad (\lambda_0) \quad (99)$$
$$K'(R) = K_0' - K_1'R + K_2'R^2 + \cdots \quad (\lambda_0').$$

The D_γ are obtained from (99)

$$D_\gamma = K_0(K_\gamma' - K_\gamma)/K_\gamma(K_0' - K_0). \quad (100)$$

If we can neglect the power in the leaky modes, the mode numbers α, μ are restricted by the condition $k_z > k_s$, that is,

$$B < 1 - K_s/K_0 \approx 2\Delta n/n, \quad (101)$$

where $K_s \equiv k_s^2$ is the square of the cladding wave number. The root-mean-square impulse response width is defined as

$$\sigma = 5{,}000[\langle t^2\rangle - \langle t\rangle^2]^{\frac{1}{2}} \text{ ns/km}, \quad (102)$$

where ( ) denotes an average taken over all the modes permitted by (101). Thus, it is a straightforward matter to evaluate from our expression in (56) the root-mean-square width of the impulse response of any circularly symmetric near-square law fiber, provided the wave-number profile can be measured with sufficient accuracy at two wavelengths.

REFERENCES

1. R. Kompfner, private communication, 1971.
2. J. A. Arnaud, "Delay Equalizer for Multimode Fibers Using Ordinary Lenses," September 8, 1971, unpublished work; "Ray Inverters for Minimizing Delay Distortion in Multimode Optical Fibers," U. S. Patent 3,759,590, September 18, 1973.
3. S. Kawakami and J. Nishisawa, "An Optical Waveguide With the Optimum Distribution of the Refractive Index With Reference to Waveform Distortion," IEEE Trans. on Microwave Theory and Techniques, *MTT 16*, No. 10 (October 1968), p. 814.
4. D. Gloge and E. A. J. Marcatili, "Multimode Theory of Graded-Core Fibers," B.S.T.J., *52*, No. 9 (November 1973), p. 1563. Note that the exponent denoted κ in the present paper and α in Ref. 15 was denoted p by W. Streifer

and C. N. Kurtz in "Scalar Analysis of Radially Inhomogeneous Guiding Media," J. Opt. Soc. of Amer., *57*, No. 6 (June 1967), pp. 779–786, and $\alpha/2$ by Gloge and Marcatili.
5. S. E. Miller, "Delay Distortion in Generalized Lens-Like Media," B.S.T.J., *53*, No. 2 (February 1974), p. 177.
6. M. Ikeda, "Propagating Characteristics of Multimode Fibers With Graded Core Index," IEEE J. of Quantum Electronics, *QE-10*, No. 3 (March 1974), p. 362.
7. C. C. Timmermann, "Influence of Deviation From the Square-Law Refractory Index Profile of Gradient Core Fiber on Mode Dispersion," AEU, *28*, No. 7–8 (1974), pp. 344–346.
8. J. S. Cook, "Minimum Impulse Response in Ideal Graded-Index Fibers," unpublished work.
9. W. A. Gambling and H. Matsuhara, "Pulse Dispersion in a Lens-like Medium," Optoelectronics, *5*, 1973, pp. 429–437. R. Olshansky and D. B. Keck, "Material Effects on Minimizing Pulse Broadening," Topical Meeting on Opt. Fiber Trans., Williamsburg, Va., January 1975. Note that our parameter D is denoted $1 + y/4$ by Olshansky and Keck.
10. J. A. Arnaud, "Degenerate Optical Cavities, Part III, Effect of Aberrations," Appl. Opt., *9*, No. 5 (May 1970), p. 1192.
11. A. Cozannet, M. Tréheux, and R. Bouillie, "Etude de la Propagation de la Lumière dans les Fibres Optiques à Gradient d'Indice," CNET Annales des Télécommunications, *29*, No. 5–6 (May–June 1974), p. 219.
12. K. H. Steiner, "A Delay Formula for Arbitrary Ray Paths in Graded Index Media," Nachrichtentechn, Z., *27*, 1974, p. 250.
13. F. W. Ostermayer, "On the Geometrical Optical Analysis of the Impulse Response of Graded Index Fibers," Topical Meeting on Opt. Fiber Trans., Williamsburg, Va., January 1975.
14. J. A. Arnaud, "Pulse Spreading in Multimode, Planar Optical Fibers," B.S.T.J., *53*, No. 8 (October 1974), p. 1599.
15. J. A. Arnaud, "Pulse Broadening in Multimode Graded-Index Fibers," Electronics Letter, *11*, No. 1 (January 1975), pp. 8–9. *Beam and Fiber Optics* (To be published by Academic Press, New York).
16. A. W. Snyder and D. J. Mitchell, "Leaky Rays Cause Failure of Geometric Optics on Optical Fibers," Electronics Letters, *9*, No. 19 (September 20, 1973), pp. 437–438. See also A. W. Snyder, "Leaky-Ray Theory of Optical Waveguides of Circular Cross Section," Appl. Phys., *4*, 1974, p. 273, and C. Pask, A. W. Snyder, and D. J. Mitchell "Number of Modes on Optical Waveguides," J. Opt. Soc. of Amer., *65*, No. 3 (March 1975), p. 356.
17. R. K. Luneburg, *Mathematical Theory of Optics*, Berkeley: University of California Press, 1964.
18. A. A. M. Saleh, private communication.
19. A. H. Nayfeh, *Perturbation Methods*, New York: John Wiley, 1973.
20. M. J. Lighthill, "Contributions of the Theory of Waves in Non-linear Dispersive Systems," J. Inst. Maths. Applics., *1*, No. 3, 1965, pp. 269–306.
21. W. B. Joyce, "Classical Particle Description of Photons and Phonons," Phys. Rev. D, *9*, No. 12 (June 15, 1974), pp. 3234–3256.
22. J. A. Arnaud, "Hamiltonian Theory of Beam Mode Propagation," in *Progress in Optics*, Vol. 11, E. Wolf ed., Amsterdam: North-Holland, 1973.
23. J. A. Arnaud, "Application of the Mechanical Theory of Light to Fiber Optics," J. Opt. Soc. Amer., *65*, No. 2 (February 1975), pp. 174–181.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 54, No. 7, September 1975
Printed in U.S.A.

Effects of Environmental Nuclear Radiation on Optical Fibers

By J. SHAH

(Manuscript received March 11, 1975)

A comparison of the available data on environmental radiation and on the radiation damage in fibers and glasses under controlled laboratory conditions leads us to the conclusion that long-term exposure to gamma rays and neutrons in the environment does not pose a serious problem for the optical fibers.

Resistance of optical fibers to damage from long-term exposure to environmental nuclear radiation is an important factor to be considered when planning a communication system using these fibers. In this report, we first summarize the nature and intensity of the natural environmental radiation to which the fibers will be exposed, and then discuss the available data on radiation damage in optical fibers and glasses under controlled laboratory conditions. A comparison of these data leads us to conclude that long-term exposure to gamma rays and neutrons in the natural environment does not pose a serious problem for the optical fibers.

I. NATURAL ENVIRONMENTAL RADIATION

The total background radiation at sea level is divided approximately equally between extraterrestrial and terrestrial components.[1] The extraterrestrial component results from the secondary radiations induced by cosmic rays, solar radiation, and Van Allen belt radiation. The terrestrial component is due to the radiation from naturally occurring radionuclides in the earth. Gamma rays (γ rays) and neutrons (n) are important constituents of this radiation[2] and we will concentrate on them for the purposes of this report.

1.1 Gamma rays

A number of measurements of the intensity of the environmental radiation as a function of location, altitude, and latitude have been made. According to Hollaender,[3] the worldwide average exposure is approximately 0.5 R/year (R stands for roentgen, a unit of exposure

dose. A brief discussion of units relevant to this report is given in the appendix). More recent studies (e.g., Ref. 4) indicate that values range from 0.1 to 0.2 rad/year in normal regions (rad is a unit of absorbed dose—see the appendix). An average value of 0.13 rad/year appears to be generally acceptable[5] for *normal regions.** However, for the purposes of our discussion of radiation damage, we will deliberately overestimate the γ-ray dose and assume a value of 0.5 rad/year.

1.2 Neutrons

Hess et al[6] have measured the extraterrestrial neutron flux as a function of neutron energy and found that the total neutron flux $[\int_0^\infty \phi(E)dE]$ at sea level is $\approx 1.5 \times 10^6$ n/cm^2-year.† Measurements by Herbst[8] indicate that the additional neutron flux from terrestrial sources is $\lesssim 10^6$ n/cm^2-year in open air. However, in tunnels or above rocks containing a high density of radioactive nuclides, or in regions with high radioactivity, Herbst obtained a flux of up to 4×10^7 n/cm^2-year. For the purposes of estimating neutron-induced damage, we will assume a rather high value of 1×10^8 n/cm^2-year to provide us with an extra margin of safety.

II. RADIATION DAMAGE IN FIBERS AND GLASSES

2.1 Gamma rays

The γ rays interact with glasses principally by forcing the electrons to leave their normal positions and move through the glass network. The primary consequence of this is an increase in the absorption coefficient in the UV-visible-near-IR range. A detailed study of γ-induced damage in fibers has been made by G. H. Sigel and co-workers[9] at the Naval Research Laboratory. They find that the γ-induced change in the refractive index is small ($<10^{-3}$) at doses as high as 10^9 rads. They also find that the γ-induced losses in optical fibers depend strongly on the fiber composition and vary from 10^{-4} dB/km-rad for bulk Suprasil SiO_2 to 5 dB/km-rad for Corning fiber No. 5010 at 8000 Å. Thus, pure fused silica is extremely resistant to radiation, while the Corning 5010 is quite susceptible to it.

A 20-year exposure to natural environmental γ radiation (assumed to be 0.5 rad/year) would lead to an increase of 50 dB/km for Corning

* There are regions with exceptionally high level of natural background radiation. In some special areas such as Kerala in India or the Santo Spirito province in Brazil, values of up to 14 R/year have been reported (Ref. 2). Certain regions of the Black Forest (Germany) have shown dose rates up to 1.8 R/year. However, these regions are rare and will not concern us in this report.

† Later reports (see Ref. 7) indicate that the values reported by Hess et al may be too high by a factor of 2 to 4.

5010 fiber. However, even in this case, the normal bleaching of the damage[9] would probably reduce the total (20-year) loss to $\lesssim 15$ dB/km. Although this number appears large, it is only a small fraction of the loss of 1000 dB/km present in the Corning 5010 fiber before exposure to any radiation. Furthermore, this is the worst case reported by Sigel et al.[9] The γ-induced losses are generally smaller in fibers with smaller initial losses.* For example, the Corning low-loss fiber (type B), having germanium-doped silica core and pure silica cladding, has an initial loss of 10 dB/km and a γ-induced loss of 0.01 dB/km-rad between 8000 Å and 12,000 Å (1.2 μm). Thus, even if we neglect bleaching, the γ-induced loss in 20 years would amount to only 0.1 dB/km. Since fibers with small initial losses are precisely the ones that will be used in communication systems, it seems reasonable to conclude that long-term exposure to environmental γ radiation will not seriously affect the fiber performance.

2.2 Neutrons

Neutrons interact principally with the nuclei rather than electrons in solids. Neutron radiation, therefore, results not only in increased absorption losses but also in structural changes that lead to changes in density, refractive index, rotary power, birefringence, thermal conductivity etc. Since small differences in refractive indices of the core and the cladding are essential to fiber performance, we will pay particular attention to refractive index changes as well as to increased losses caused by n-irradiation.

To our knowledge, the only study of n-induced losses in optical fibers is by Maurer et al.[10] They irradiated high-silica-glass multimode fiber waveguides with 14-MeV neutrons, using doses of as high as 1.4×10^{12} n/cm^2. They concluded that the n-induced loss varies roughly linearly with the total dose and is less than 1.5×10^{-11} (dB/km)/(n/cm^2) in the 8000-Å to 12,000-Å region. This number, which is obtained from the figure given by Maurer et al,[10] is in fact an overestimate of n-induced damage, because we have disregarded the fact that the n-irradiated samples also received a simultaneous dose ≈ 1000 rads of γ radiation. However, even if we assume this to be the true value, a 20-year exposure to environmental n-irradiation (2×10^9 n/cm^2) would increase the loss by only about 3×10^{-2} dB/km. It should be emphasized that this extrapolation is only approximately valid because neutrons in the environment have a wide range of energies (from 0.01 eV to 10^{10} eV), whereas the neutrons in the con-

* While there is no evidence that the correlation between low radiation damage and low initial losses is *universally* valid, such a correlation definitely exists in the presently available data.

trolled experiment were monoenergetic (14 MeV). However, even after a *thirtyfold* increase, the n-induced losses would still be less than 1 dB/km. Therefore, it seems reasonable to conclude that absorption losses induced by long-term exposure to environmental n-radiation will not seriously affect fiber performance.

Neutron-induced changes in the refractive index of the fibers can be a potential source of problems. We know of no measurements on fibers which can shed light on this problem. However, an extensive literature exists on the effects of n-irradiation on various forms of silica and other commonly used glasses (a good summary is given in Ref. 11). The refractive index of vitreous silica changes by 0.67 percent under a flux of 2×10^{20} n/cm^2 of thermal (<0.1 eV) neutrons.[12] From the measurement by Primak,[12] we deduce that the rate of increase of the refractive index of vitreous silica is approximately 5×10^{-22} per (n/cm^2) for doses less than 1×10^{19} n/cm^2. This suggests that the changes in refractive index induced by environmental neutrons (2×10^9 n/cm^2 in twenty years) will be less than 1×10^{-12}, a truly negligible effect when we consider the fact that the difference in the refractive index of the core and the cladding is typically larger than 10^{-3}.

No data are available on the n-induced changes in refractive indices of other glasses. However, density changes have been investigated for many glasses.[11] For vitreous silica,[12] the density increases approximately linearly (10^{-19} percent per n/cm^2) up to 2.5×10^{19} n/cm^2 and then saturates. Other glasses (except borosilicate glasses) are also quite resistant to neutrons and show very few changes up to about 10^{17}–10^{18} n/cm^2.* The borosilicate glasses are more susceptible because boron, like other light elements, has high neutron cross section. However, even these glasses show damage only when flux levels exceed 10^{14} n/cm^2,* which is some five orders of magnitude larger than the accumulated (20 years) flux of $\approx 2 \times 10^9$ n/cm^2 encountered in the environment.

III. CONCLUSIONS

We have summarized the available data on environmental nuclear radiation and also the data on radiation damage in glasses under controlled laboratory conditions. Unfortunately, the laboratory experiments were not performed with the exact γ ray or neutron energy distributions that one encounters in environmental radiation. It is difficult, therefore, to make accurate predictions about the radiation damage in fibers caused by environmental radiation. However, we

* See Table 6.12 in Ref. 10.

have made some approximate estimates from which it is reasonable to conclude that damage induced by environmental γ or neutron radiation should not pose any serious problems to optical fibers so far as their optical loss or refractive index are concerned. More recent experimental works by Evans and Sigel[13] and Mattern et al.[14] do not affect this conclusion.

Some general comments seem to be appropriate in conclusion. Pure fused silica seems to be extremely resistant to radiation damage. It is also useful to remember that the addition of small amounts (0.1 to 0.2 percent) of Cerium[9,11] makes most glasses more resistant to radiation. We have not discussed damage by α particles, but it is appropriate to mention here that α particles have very short ranges in air as well as in most other materials. Therefore, it seems unlikely that α particles will pose any problems for the optical fibers if the fibers are enclosed in a conduit. Finally, the background luminescence induced by environmental ionizing radiation has been considered by Cohen and Lanzerotti[15] and found to be not significant for fiber optic communications systems.

IV. ACKNOWLEDGMENTS

We acknowledge helpful discussions with G. H. Sigel, Jr. and M. Weiss.

APPENDIX

Units of Dosimetry

The most useful units in the study of radiation damage in solids are the particle or photon fluxes as a function of their energy. Thus, $\phi(E)dE$, expressing the number of particles/cm^2-sec in the energy range E to $E + dE$, completely specifies the incident radiation field. However, many special units are frequently used in specifying the radiation. Roentgen (R) is a unit of exposure dose used for X rays and γ rays and is defined as follows. Roentgen is that exposure of X or γ radiation which gives a dose of 87.7 ergs/g to air.

A special unit of absorbed dose is called a "rad." One rad = 100 ergs/g.

Unlike the roentgen, the rad is independent of the irradiated material. This means that a given beam of radiation acting for the same time will deliver different doses, expressed in rads, according to whether it is absorbed in air, tissue, or other materials. The rad in Section I refers to air as the reference material. The rad as used here is indirectly a measure of the radiation field rather than the absorbed dose in the sample because it refers to energy absorbed by air rather than the sample under study. Under these conditions, rad and roentgen are

numerically equal within about 20 percent and can be used interchangeably.

In the work reported by Sigel et al.,[9] Si is used as the reference absorbing material. The differences in using air or Si as the reference material are small (less than a factor of two) and are inconsequential for the purposes of this report.

The conversion between rad and n/cm^2 and photons/cm^2 as a function of energy are given by H. Stern.[16] (See also the report by J. Moteff.[17]) For example, for 1 MeV γ-ray photons, 1 rad $\approx 2 \times 10^9$ photons/cm^2. For 1 MeV neutrons, 1 rad $\approx 2.6 \times 10^8$ n/cm^2. For γ rays with energy E between 0.07 and 2 MeV, 1 rad (air) $\approx 2 \times 10^9/E$ photon/cm^2. Conversion factors at other energies may be obtained from the above references. See also the American Institute of Physics Handbook.[18]

REFERENCES

1. C. A. Sondhaus and R. D. Evans, "Dosimetry of Radiation in Space Flight," in *Radiation Dosimetry*, vol. III, eds. F. H. Attix, W. C. Roesch, and E. Tochilin, New York: Academic Press, 1969, pp. 453–521.
2. Report of the United Nations Scientific Committee on the Effects of Atomic Radiation, General Assembly Official Records, Suppl. No. 14 (A/6314), New York: United Nations, 1966.
   W. V. Mayneord and C. R. Hill, "Natural and Man-Made Background Radiation," in *Radiation Dosimetry*, vol. III, eds. F. H. Attix, W. C. Roesch, and E. Tochilin, New York: Academic Press, 1969, pp. 401–451.
3. A. Hollaender, ed. *Radiation Biology*, vol. I, *High Energy Radiation*, New York: McGraw–Hill, 1964, p. 582.
4. K. Z. Morgan and J. E. Turner, eds. *Principles of Radiation Protection*, New York: John Wiley, 1967, Chap. 1.
5. Report of the United Nations Scientific Committee on the Effects of Atomic Radiation, General Assembly Official Records, Suppl. No. 16 (A/5216), New York: United Nations, 1962.
   J. A. S. Adams and W. M. Lowder, eds., *The Natural Radiation Environment*, Chicago: The University of Chicago Press, 1964.
   Proceedings of United Nations International Conference on Peaceful Uses of Atomic Energy (1971), vols. 1–14, New York: United Nations, 1972.
6. W. N. Hess, H. W. Patterson, and R. W. Wallace, "Cosmic-Ray Neutron Energy Spectrum," Phys. Rev., *116*, No. 2 (October 15, 1959), pp. 445–457.
7. J. Kastner, B. G. Oltman, and L. D. Marinelli, "Progress Report on Flux and Spectrum Measurement of the Cosmic-Ray Neutron Background," *The Natural Radiation Environment*, eds. J. A. S. Adams and W. M. Lowder, Chicago: The University of Chicago Press, 1964, pp. 441–448.
8. W. Herbst, "Investigations of Environmental Radiation and Its Variability," *The Natural Radiation Environment*, eds. J. A. S. Adams and W. M. Lowder, Chicago: The University of Chicago Press, 1964, pp. 781–796.
9. G. H. Sigel, Jr., "Radiation Effects in Fiber Optical Waveguides," NRL Memorandum Report 2704, Washington, D. C.: Naval Research Laboratory, 1973. Also, NRL Memorandum Report 2934, November 1974.
   G. H. Siegel, Jr. and B. D. Evans, "Effects of Ionizing Radiation on Transmission of Optical Fibers," Appl. Phys. Letters, *24*, No. 8 (April 15, 1974), pp. 410–412.
   E. J. Friebele, R. J. Ginther, and G. H. Sigel, Jr., "Radiation Protection of Fiber Optic Materials: Effects of Oxidation and Reduction," Appl. Phys. Letters, *24*, No. 8 (April 15, 1974), pp. 412–414.
10. R. D. Maurer et al., "Effect of Neutron- and Gamma-Radiation on Glass Optical Waveguides," Applied Optics, *12*, No. 9 (September 1973), pp. 2024–2026.

11. R. A. Wullaert et al, "Effects of Radiation on Ceramic Materials," *Effects of Radiation on Materials and Components*, eds. J. F. Kircher and R. E. Bowman, New York: Reinhold, 1964, pp. 277–402.
W. Primak and M. Bohmann, "Radiation Damage," *Progress in Ceramic Science*, vol. 2, ed. J. Burke, Oxford, England: Pergamon Press, 1963, distributed by MacMillan, Chap. 3.
E. Lell, N. J. Kreidl, and J. R. Hensler, "Radiation Effects in Quartz, Silica, and Glasses," *Progress in Ceramic Science*, vol. 4, ed. J. E. Burke, Oxford, England: Pergamon Press, 1966, Chap. 1.
H. M. Presby and W. L. Brown, "Refractive Index Variations in Proton-Bombarded Fused Silica," Appl. Phys. Letters, *24*, No. 10 (May 15, 1974), pp. 511–513.
12. J. Lukesh, "Neutron Damage to the Structure of Vitreous Silicon," Phys. Rev., *97*, No. 2 (January 15, 1955), pp. 345–346.
W. Primak, "Fast-Neutron-Induced Changes in Quartz and Vitreous Silicon," Phys. Ref., *110*, No. 6 (June 15, 1958), pp. 1240–1254.
13. B. D. Evans and G. H. Sigel, Jr., "Permanent and Transient Radiation-Induced Losses in Optical Fibers," IEEE Trans. Nuc. Sci., *NS-21*, No. 6 (December 1974), pp. 113–118.
14. P. L. Mattern et al., "The Effects of Radiation on the Absorption and Luminescence of Fiber Optic Waveguides and Materials," IEEE Trans. Nuc. Sci., *NS-21*, No. 6 (December 1974), pp. 81–95.
15. R. L. Cohen and L. J. Lanzerotti, "Noise in Fiber Optics Communications Systems Induced by Ionizing Radiation," Appl. Optics, *13*, No. 10 (October 1974), pp. 2190–2192.
16. H. E. Stern, "Introduction: Shield Design Criteria and Procedures," in *Reactor Handbook*, 2nd ed., E. P. Blizard, ed., New York: Interscience Publishers, 1962, pp. 4–15.
17. J. Moteff, "Miscellaneous Data for Shielding Calculations," U. S. Atomic Energy Commission, Report APEX 176, December 1, 1954.
18. D. W. Gray, ed. American Institute of Physics Handbook, 2nd ed., New York: McGraw-Hill, 1963, pp. 8–321.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 54, No. 7, September 1975
Printed in U.S.A.

Loose Tube Splices for Optical Fibers

By C. M. MILLER

(Manuscript received March 14, 1975)

A technique for splicing optical fibers has been developed that uses a self-aligning square cross-section tube, with inner dimensions slightly larger than the optical fiber. A total loss of 0.58 dB was obtained for eight splices connected in series using a graded-index fiber with a 68-μm core diameter. The splices were made one at a time without the use of microscopes or micromanipulators; however, the fabrication process could be mechanized and extended to groups of fibers. A holding fixture could be added to adapt this technique to a connect-disconnect type splice. The size of the splice is presently 0.012 in. square, making it suitable for use within cables. Measurement set refinements that were needed to measure individual splice losses as low as 0.05 dB include an improved detector and means for better control of launching conditions.

I. INTRODUCTION

The basic requirements for low-loss splices are (*i*) accurate alignment, (*ii*) good fiber ends, and (*iii*) accurate diameter control. Transverse alignment accuracy of approximately ±0.1-fiber-core radius (typically, ±0.0001 in.) is required to achieve a splice loss of 0.1 dB. Good fiber ends may be produced by scoring and breaking,[1] grinding and polishing, or disc sawing. Accurate fiber diameter control is also needed; however, significant progress is being made in this area. Of these three requirements, accurate transverse alignment may be the most difficult problem to solve, especially when the field environment and variability of craftsmen's skill are considered.

Single-fiber splicing has been accomplished by Bisbee[2] and Dyott et al[3] using heat fusion. Someda[4] suggested using embossed plastic to obtain transverse alignment. This paper describes a splicing technique that uses a loose-fitting, square, cross-section tube to align the fibers. The splices produced are small, exhibit very low losses, and are simple and inexpensive.

Previously, snug-fitting sleeves have been suggested,[5] but three problems are usually encountered.

(*i*) If the glass sleeve is to support the fiber with the required alignment accuracy, it must be less than typically 0.0001 in. larger than the fiber. Both fiber and sleeve must be highly circular, and the fiber diameter must be controlled to at least the same tolerances. These tolerance requirements have discouraged efforts to use a snug tube as an alignment mechanism.

(*ii*) Given a snug tube of the proper dimensions, the initial insertion of a fiber into that tube is difficult. Pinnow[5] has described a method of flaring the inner diameter of capillary tubes, which reduces the initial insertion difficulty.

(*iii*) Contaminants that are scraped off the inside wall of snug-fitting sleeves during fiber insertion are trapped between the fiber ends where the effect of contamination is worst.

The "loose"-fitting square-tube splice described below reduces these difficulties substantially and appears to have potential application in several places in a fiber-optic communication system.

II. SPLICE CONFIGURATION AND ASSEMBLY

The loose-tube splice combines the alignment accuracy obtainable by using a groove for alignment[4] with the small size and simplicity of glass sleeves.[5] The fiber ends are biased to one corner of the square cross section by bending the fiber outside the tube. Figure 1 is a pictorial layout of the square tube with two fiber ends in position within the tube. The tube has nearly flat interior walls and a small radius in the interior corners, as shown in the cross section in Fig. 2. One corner of the square is used as a groove for aligning the fibers.

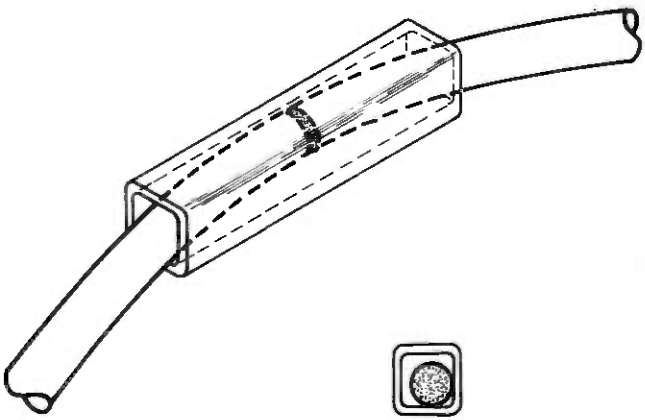


Fig. 1—Splice configuration.

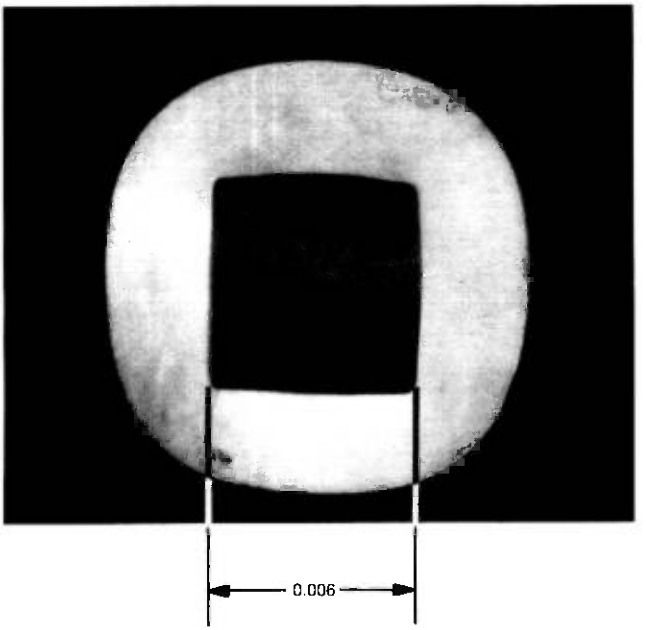


Fig. 2—Square tube cross section.

Epoxy is forced into the square tube prior to insertion of the fibers and serves several useful functions.

(*i*) The epoxy serves as an adhesive after curing to hold the splice together.
(*ii*) The epoxy also serves as an index-matching material with good glass-wetting characteristics.
(*iii*) Contamination on the fiber ends is washed away by the flow of epoxy around the fiber ends during insertion of the fibers.

Assembly of a splice involves inserting two fibers with good ends approximately halfway into each end of a square cross-section tube filled with uncured epoxy. No particular orientation of the square-tube cross section is required. The fibers are placed on a flat surface and bent in a curved pattern. This causes forces to be generated that rotate the tube so that a diagonal of the square cross section is in the same plane as the bent fibers. The tube is therefore self-aligned and the fibers biased to one corner by action of the fiber stiffness. After the bends are made, the fibers are taped to a flat surface in the bent configuration and the fibers pushed into the tube until they touch each other. Figure 3 is a cross-section photograph of a splice showing a

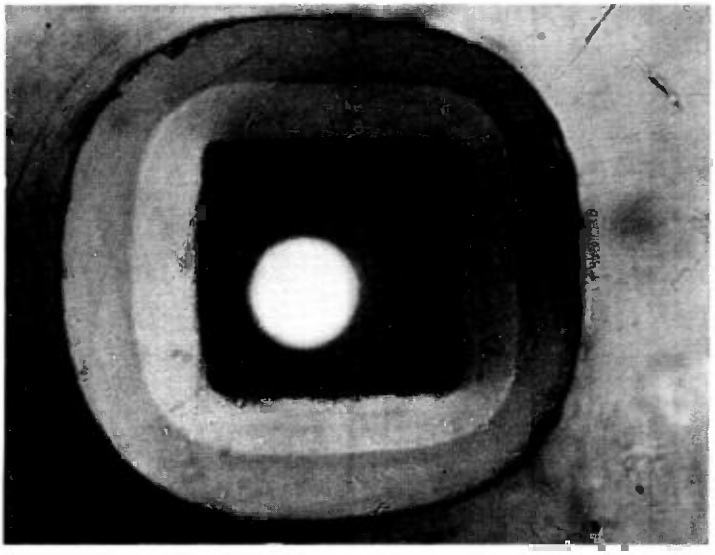


Fig. 3—Splice cross section showing position of fiber in vicinity of joint.

fiber in a corner of the square tube. Figure 4 is a magnified view of one splice, and a longitudinal section is shown in Fig. 5. In spite of the small angle between the fiber ends caused by one end not being broken

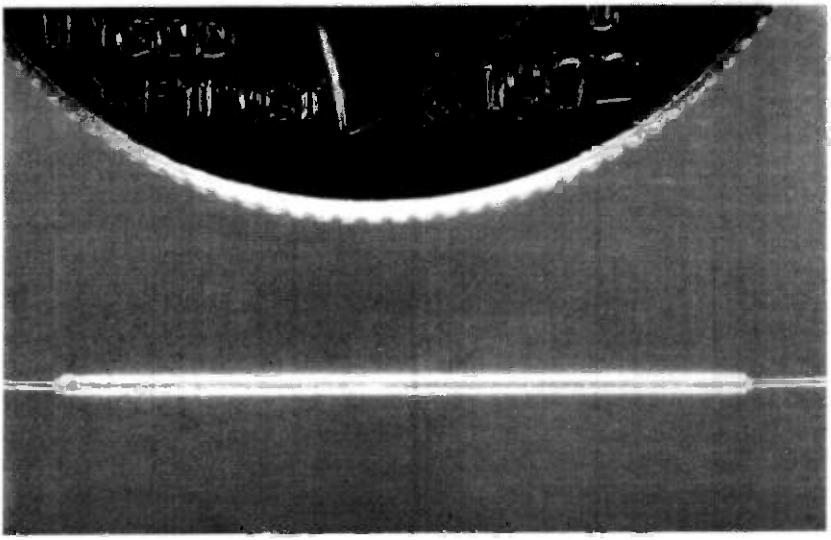


Fig. 4—Single loose tube splice, with tube approximately 0.5-inch long and 0.012-inch wide.

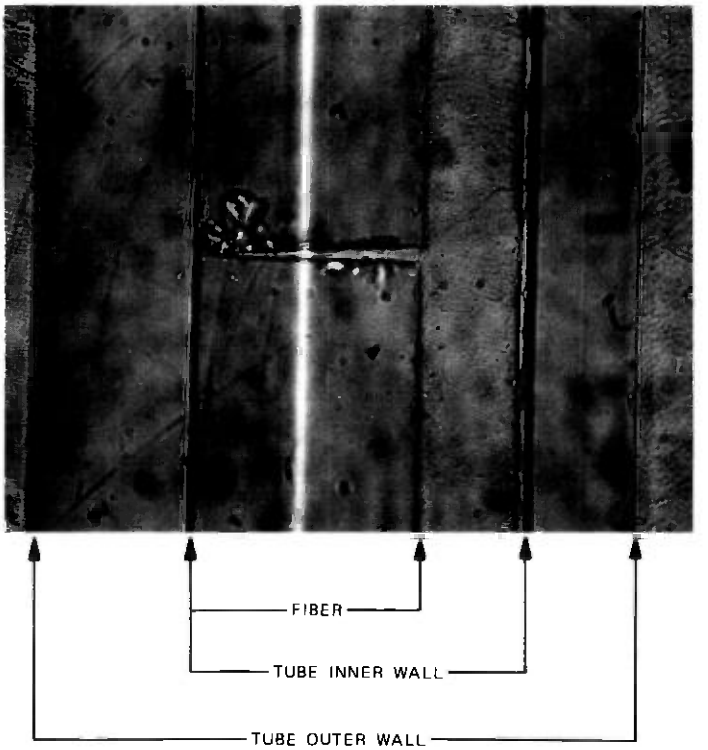


Fig. 5—Longitudinal section, 250X.

perpendicular to the fiber axis, the splice loss for the splice in Fig. 5 was only 0.07 dB.

III. END PREPARATION

Good fiber ends are necessary for the fabrication of low-loss splices. As mentioned earlier, several techniques exist for preparing suitable ends. A score-and-break technique[1] was used for end preparation on all splices reported in this paper. A single fiber is clamped in the apparatus shown in Fig. 6 with approximately 100-g load applied to the fiber by a spring. The fiber rests in a groove along a 2-in. radius arc and is scored lightly with a hand-held diamond knife edge. Fiber ends prepared by this method are nearly perfect, as shown in Fig. 7. A very small amount of edge chipping is present where the fiber was scored.

IV. SPLICE-LOSS MEASUREMENTS

After constructing just a few square-tube splices, it became evident that the measurement set-up being used was not adequate for losses

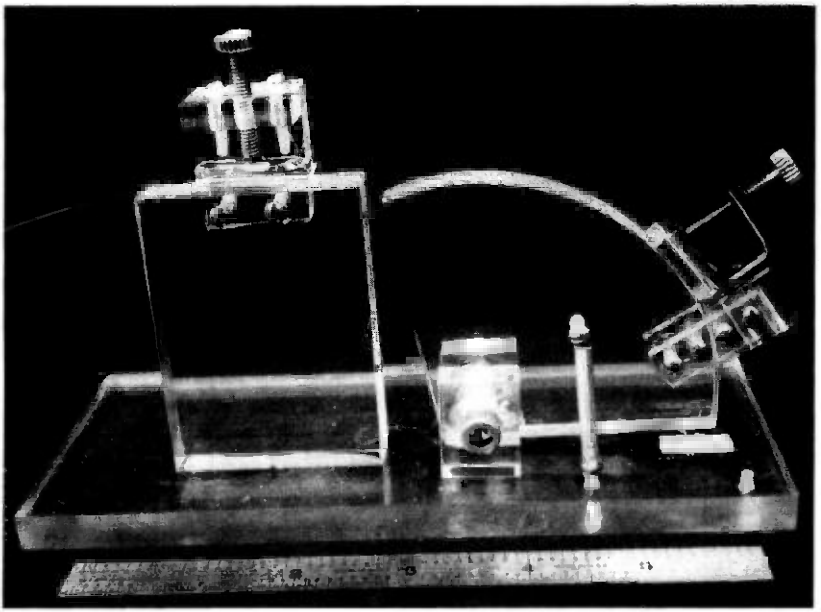


Fig. 6—End preparation apparatus.

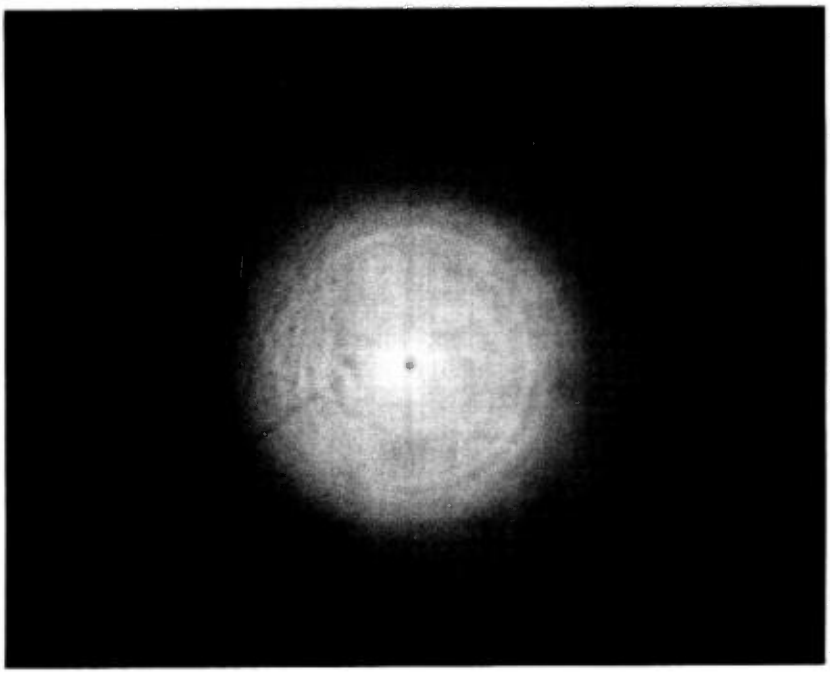


Fig. 7—Good fiber end.

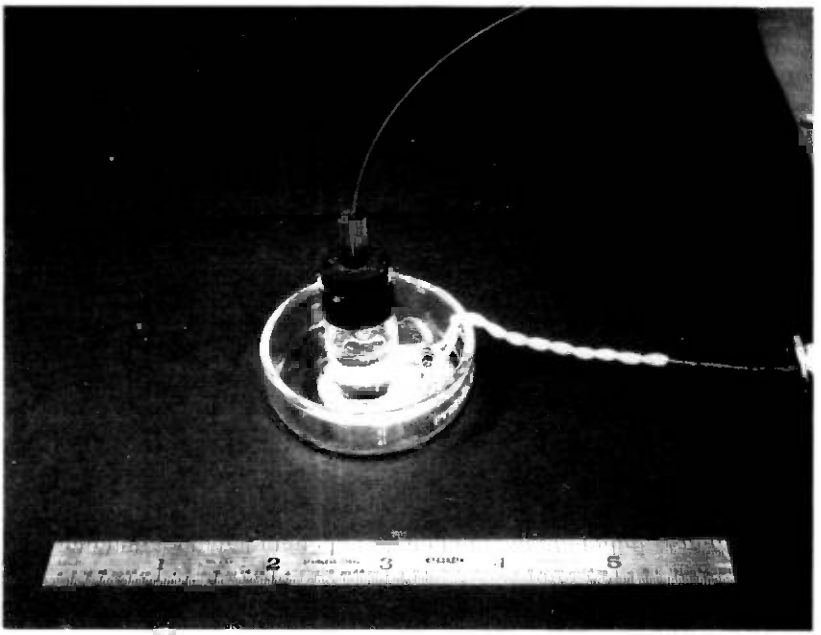


Fig. 8—Single fiber detector.

below 0.1 dB. A new detector was built, which repeated to within ±0.015 dB. Accurate positional repeatability was accomplished by using a 0.05-in. inner-diameter capillary tube which was tapered to 0.006 in. inner diameter by heating and pulling. The fiber is easy to insert in this detector, and the positional repeatability is excellent. The solar cell sensor was immersed in index-matching liquid to reduce reflections, and a microscope cover slip was used to protect the cell from damage by the fiber being measured. A photograph of the detector is shown in Fig. 8. At the input end of the measurement set, a vacuum chuck was added to ensure repeatability of launching conditions. This chuck positions the fiber accurately along the optical axis of the 30X launching objective lens and the laser. The overall repeatability of splice-loss measurements is within ±0.03 dB.

Care was taken to place the fiber in the same coiled configuration after splicing so that bending losses before and after splicing would be similar. Fiber loss of approximately 0.01 dB/m was subtracted from the total loss measurements so that losses stated are for the splices only. All loss measurements were made at a wavelength of 0.6328 μm.

V. RESULTS

Initially, a fixture was used to hold the square tube and control the fiber bending. Although the maximum loss measured on 25 consecutive

splices fabricated with this fixture was 0.21 dB, it was found that the tube would align itself if allowed to rotate. Losses were lower and the assembly of the splice, as previously mentioned, was much easier.

Eight epoxied square-tube splices fabricated in series using no fixtures except the end-making apparatus produced a total splice loss of 0.58 dB or 0.073 dB per splice. The splices had approximately $1\frac{1}{2}$ m of fiber between each splice and on each end and were put together in series to increase the total loss to an accurately measurable quantity.

It has been found that splices measured in series have higher losses than when measured individually. Ten earlier splices were fabricated in series and measured 1.37 dB. The ten-series splices were measured separately and gave the distribution shown in Fig. 9. The average loss per splice was 0.077 dB compared to 0.137 dB in series, or a nearly 2-to-1 increase for splices in series. The process of peaking up the power through a single splice probably selects the launching conditions and therefore the mode structure best suited to the particular imperfections of that splice. A loss measurement made in this way gives a value that is too low. That is, the loss of a splice with long lengths on either side or with other splices nearby is apt to be considerably higher than when measured separately with short fibers on each side of the splice. The eight-series splices mentioned earlier were not measured separately because losses as low as 0.03 dB, which would be expected based on the series loss, could not be accurately measured.

A slight longitudinal separation of fiber ends within the tube occurred during epoxy cure for the 10 splices in series. Loss measured before epoxy cure was 0.69 dB and 1.37 dB after epoxy cure as stated above. Fibers were bent in a 90-degree arc and taped to an optical table while the epoxy cured. This configuration applied very little, if any, force component to hold the fibers in place during cure. The eight-series splices were bent through an arc of approximately 45 degrees and

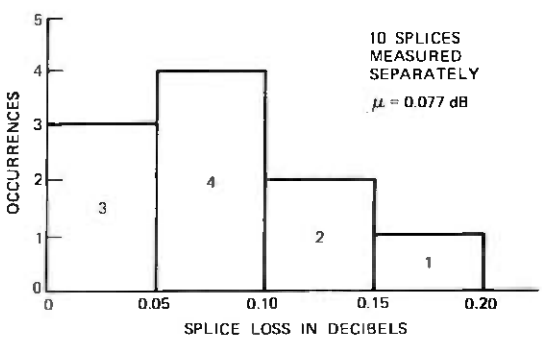


Fig. 9—Splice loss histogram.

taped to the table with a definite force applied to keep the fiber ends in contact. Although this force (a result of the fiber stiffness) is small, end separation did not occur and a loss of 0.58 dB was measured before and after epoxy cure.

These extremely encouraging results have stimulated thinking as to how loose square-tube splices could be applied to other types of splices, e.g., connect-disconnect configuration.

VI. CONNECT-DISCONNECT SPLICE FOR SINGLE FIBERS

Several configurations based on the square tube can be envisioned for a connect-disconnect splice, that is, a splice that can be reassembled many times and used as a connector. Figure 10 is a photograph of a simple fixture that supports a single square-tube splice by the fibers on each side of the splice. The splice itself is suspended in air. This fixture is not intended to be a finished design, but it does produce losses of 0.1 dB or less. The clamps are lined with a thick, soft EVA layer that grips the fibers and holds them in position. More practical designs are sketched in Fig. 11. An index-matching material is necessary to achieve 0.1 dB, although a liquid index-matching material may be feasible.

Polymethylmethacrylate (PMMA) was suggested as an alternative index-matching material by Pinnow.[5] This thermoplastic could be drawn into a fiber and inserted into the square tube. Heat could then be applied to melt the PMMA, the fibers inserted, and the PMMA allowed to cool. The splice could be disassembled by again heating the tube and removing the fibers. These connect-disconnect splices may

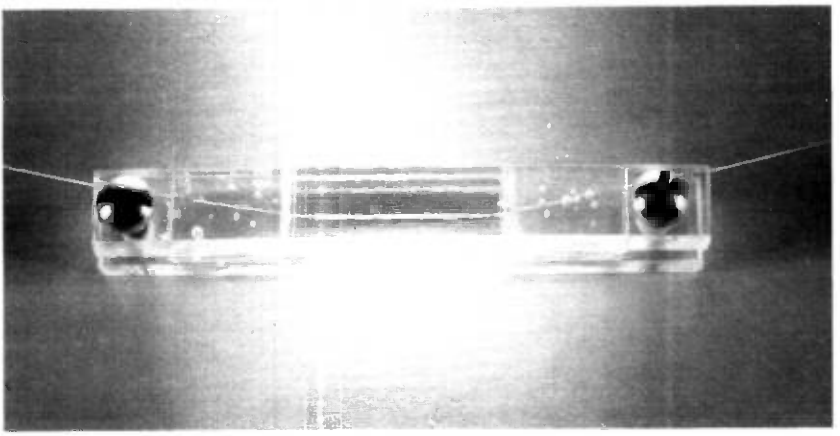


Fig. 10—Quick-connect holding fixture.

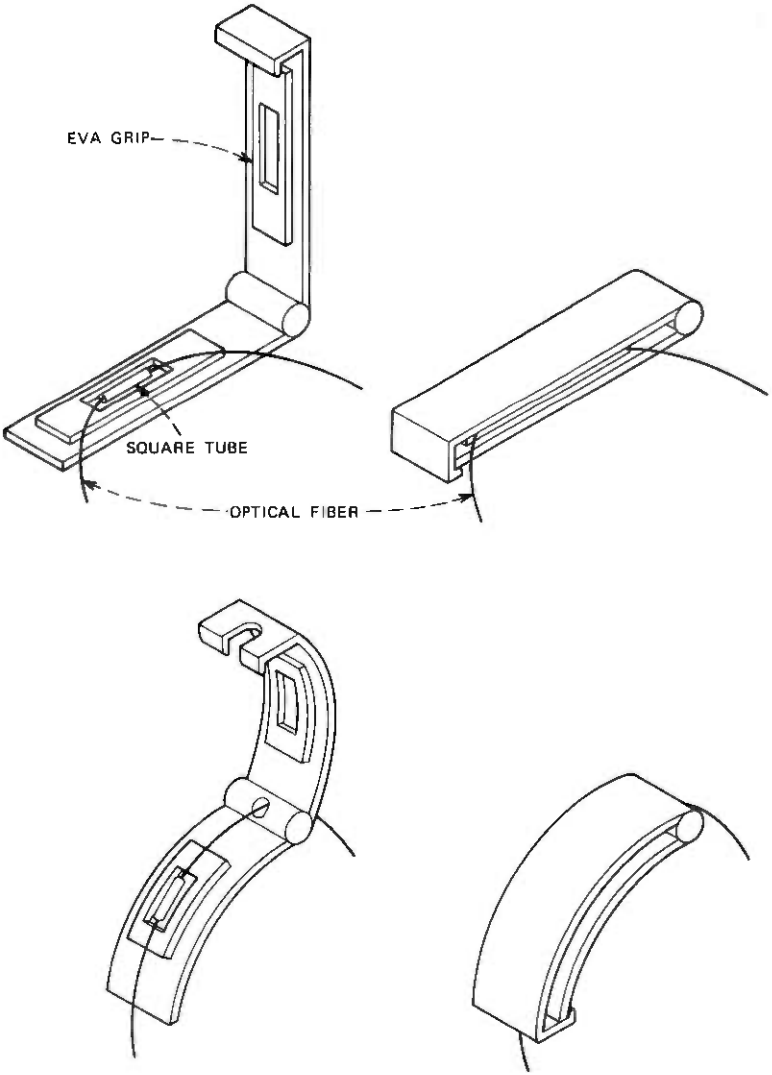


Fig. 11—Connect-disconnect holders for single-fiber splices.

be useful as methods for connecting sources, detectors, and line regenerators in a fiber-optic communication system.

VII. ACKNOWLEDGMENTS

The author appreciates the assistance of A. W. Miller in splice fabrication and the discussions with M. I. Schwartz, during which time some of the single-fiber connect-disconnect ideas evolved.

REFERENCES

1. E. L. Chinnock, D. Gloge, P. W. Smith, and D. L. Bisbee, "Optical Fiber End Preparation for Low Loss Splices," B.S.T.J., *52*, No. 9 (November 1973), pp. 1579–1588.
2. D. L. Bisbee, "Optical Fiber Joining Experiment," B.S.T.J., *50*, No. 10 (December 1971), pp. 3153–3158.
3. R. B. Dyott, J. R. Stern, and J. H. Stewart, "Fusion Junction for Glass-Fiber Waveguides," Elect. Lett., *8*, No. 11 (June 1, 1972), pp. 290–292.
4. C. G. Someda, "Simple, Low-Loss Joints Between Single-Mode Optical Fibers," B.S.T.J., *52*, No. 4 (April 1973), pp. 583–596.
5. D. A. Pinnow, U. S. Patent 3,810,802, May 14, 1974.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 54, No. 7, September 1975
Printed in U.S.A.

The Effect of Longitudinal Imbalance on Crosstalk

By G. MILLER

(Manuscript received December 30, 1974)

Multipair cables are carefully constructed to control the coupling between wire pairs to prevent excessive crosstalk. Several types of coupling modes exist between pairs, but the principal effort is devoted to controlling the "metallic-to-metallic" coupling mode because the coupling loss in this path is the most important in ordinary telephonic use of the cable. Over a half-century ago, Campbell showed that crosstalk behavior of this mode could be characterized by measuring a simple function called capacitance unbalance. This paper shows that at voice frequencies the crosstalk characteristics of the "longitudinal-to-metallic" mode can be predicted by measuring a second similar function of the same parameters that are contained in the capacitance unbalance. With the aid of these two functions, it is shown how the longitudinal balance of terminal equipment connected to a cable pair affects crosstalk. It is further shown that a longitudinal balance of approximately 40 dB or more is necessary for any station or terminal equipment used in the telephone network so that it will not significantly increase the small amounts of crosstalk inherent in the careful cable design. Also, a limitation is established for the maximum longitudinal voltages at voice frequency that can be applied without noticeably increasing crosstalk and noise in other cable pairs. This limitation is approximately 40 dB more restrictive than the tariff limitations for metallic voltages.

I. INTRODUCTION

A multipair cable consists of many insulated but unshielded conductors within a protective conducting sheath. The individual conductors are used to form circuits. In one configuration, called a metallic circuit, two conductors are paired and form the circuit. Signals are applied between them. This is called metallic excitation of the circuit, and the signal is said to propagate in the metallic mode. In another configuration, called a longitudinal circuit, two conductors are paral-

leled and these, plus the conducting cable sheath, form the circuit. Signals are applied between the paralleled conductors and the conducting cable sheath, which is grounded. This is called longitudinal excitation of the circuit, and the signal is said to propagate in the longitudinal mode. It is also possible for one wire pair to be used for both circuits simultaneously and, consequently, for one wire pair to be simultaneously excited in the metallic and longitudinal modes of propagation. This happens when the terminal equipment is longitudinally unbalanced, as will be explained.

Because the conductors are not shielded and are in close proximity to each other, electromagnetic fields generated by current flowing through the conductors cause energy to be coupled from one circuit to another. This is called crosstalk and is undesirable, since it may cause noise in other circuits that can impair the performance of digital and analog systems, or even be intelligible speech that is overheard and leads to loss of privacy.

Crosstalk cannot be eliminated, but several things can be done to reduce it, that is, to increase the crosstalk loss between circuits. First, metallic circuits are used rather than longitudinal circuits, because it was found by experience that the crosstalk loss between two metallic circuits is generally greater than the loss between two longitudinal circuits or between a longitudinal and a metallic circuit. Second, adjacent conductors are paired and often twisted and are used for the metallic circuits because they are less susceptible to inductive noise and the crosstalk loss between twisted pairs is generally greater than between nontwisted pairs. Twisting reduces crosstalk by assuring that each pair of the cable is exposed to opposing couplings by transposing its conductors relative to the disturbing pair. Third, the terminal equipment at both ends of a pair should be longitudinally balanced, i.e., have impedance symmetry with respect to ground, because longitudinal imbalance has the effect of producing longitudinal excitation which consequently can increase crosstalk. Finally, the cable pairs are also constructed to have longitudinal impedance symmetry for the same reason.

Since metallic circuits are usually used, and both cables and terminal equipment are usually constructed to be longitudinally balanced, most crosstalk studies to date have concentrated on what is called metallic-to-metallic crosstalk, i.e., crosstalk between balanced metallic circuits. Much less is known analytically about the crosstalk loss between longitudinal circuits or between a longitudinal and a metallic circuit. For example, to explain crosstalk between balanced metallic circuits, Campbell[1] assumed that all circuits within a cable were longitudinally

symmetrical, that the pairs were excited metallically, and that, consequently, the applied metallic signal would not excite any longitudinal voltage in the disturbing or disturbed pairs. Thus, crosstalk would be due to cable characteristics alone. Campbell was then able to show that crosstalk at low or voice frequencies, where inductive coupling is negligible, was very nearly proportional to the capacitance unbalance, which is a function of the four interwire capacitances between two cable pairs, and is now used as a measure of quality of a cable with regard to crosstalk performance.

In another study, Foschini[2] developed an accurate transmission model of cable systems for computing crosstalk which is an extension of Campbell's work. He too assumed longitudinal symmetry and showed that crosstalk coupling losses between metallic circuits can be predicted quite accurately from Campbell's capacitance unbalance. Although his results are valuable in the study of crosstalk for the metallic mode of propagation, they, as well as Campbell's results, do not consider the effects of terminal imbalance on crosstalk loss.

The objectives of this paper are to extend the results of Campbell and Foschini by first removing the constraints of metallic circuits, terminal balance, and pair symmetry; and to construct a model to permit calculating the crosstalk loss between pairs as a function of terminal balance and pair symmetry. The model is used to show why terminal imbalance can greatly increase crosstalk by causing longitudinal excitation of a cable pair and, consequently, why limitations must be imposed on the longitudinal balance of terminal equipment and on the direct application of longitudinal signals. These objectives are accomplished by showing, through numerical solutions and experimental results, that longitudinal excitation couples energy into adjacent wire pairs with much less loss than does metallic excitation.

The paper is divided into four sections. First, the important results on longitudinal balance and longitudinal voltage restrictions are given. Next, the model of crosstalk between two wire pairs in a cable is analyzed using transmission line equations. This model is used to derive a new set of crosstalk coupling coefficients that can be used to relate the crosstalk loss between two metallic circuits, a longitudinal and a metallic circuit, and two longitudinal circuits. Third, average values for these coupling coefficients for a typical cable are obtained, derived from measured characteristics. Using the coupling coefficients the predicted increase in crosstalk resulting from longitudinal excitation is compared with direct measurements of the increase made on another cable. Finally, restrictions on longitudinal balance and longitudinal voltages are established.

II. RESULTS

2.1 Requirement on longitudinal balance

For application to crosstalk performance, it is appropriate to define a longitudinal balance* of terminal equipment as

$$\mathrm{BAL}_{\mathrm{M-L}}(f) = 20 \log_{10} \left| \frac{e_M(f)}{e_L(f)} \right| ,$$

where e_L is the longitudinal voltage produced when a *metallic* voltage e_M is applied at any frequency f. The subscript "M − L" means the conversion from a metallically applied voltage to a longitudinal voltage. This paper shows that a balance of approximately 40 dB or more in the voice frequency region is required for any terminal device to ensure that the level of crosstalk that already exists in the network will not be significantly increased. This requirement is based on measurements of the near-end crosstalk at 1000 Hz of cable with a balanced and unbalanced termination. It is assumed that any metallic signal applied to the telephone network does not exceed the power level specified in Ref. 3. Longitudinal and metallic voltages are defined in Section 3.3.

2.2 Restriction of longitudinally applied voltages

Crosstalk coupling losses decrease with increasing frequency and hence voltage restrictions are frequency dependent. Figure 1 shows the limitations on applied longitudinal voltages established so as to increase the crosstalk energy already present in the telephone network by no more than about 1 dB.

2.3 Derivation of crosstalk coupling coefficients

Three new capacitive coupling coefficients have been derived that can be used with a simple but reliable computation method to predict the degradation in crosstalk performance for a particular cable when any of its terminations are unbalanced. These coefficients are given in Table I. The coefficients are defined in eqs. (10) to (13), and the interpair capacitances given in the formulas are the capacitances between the pairs shown in Fig. 2.

III. COUPLING BETWEEN TWO WIRE PAIRS IN A CABLE

3.1 Transmission line model

Figure 2 models two wire pairs within a cable of length l. The following assumptions about a cable are made to construct this model:

---

* A second type of balance for noise immunity purposes is a separate but important consideration for good telephone network performance. It is defined in Section 4.2. However, crosstalk does not enter into establishing its restrictions.

(*i*) The impedance and admittance per unit length of each wire pair, the admittances to ground per unit length, etc., are constant.

(*ii*) The conductance to ground and between wire pairs is negligible, i.e., the admittances are purely capacitive.

(*iii*) The inductive coupling between pairs is negligible at voice frequencies.

(*iv*) The impedances per unit length of all wire pairs are equal.

Let the admittances per unit length between the two circuits be Y_{13}, Y_{23}, Y_{24}, and Y_{14} connected between conductors 1 and 3, 3 and 2, 2 and 4, and 4 and 1, respectively, where conductors 1–2 form one twisted wire pair and 3–4 form the other pair. The impedances per unit length of the four wires are Z_1, Z_2, Z_3, and Z_4 and the admittances of wires to ground are Y_{1g}, Y_{2g}, Y_{3g}, and Y_{4g}. The admittances per unit length of the wire pairs are Y_{12} and Y_{34}, and the voltages and currents are labeled in the figure.

Consider a differential section of the model of length Δx. It is readily seen that the following eight current-voltage relationships hold for this differential section:

$$V_1(x + \Delta x) = V_1(x) - I_1(x)Z_1\Delta x \tag{1a}$$

$$V_2(x + \Delta x) = V_2(x) - I_2(x)Z_2\Delta x \tag{1b}$$

$$V_3(x + \Delta x) = V_3(x) - I_3(x)Z_3\Delta x \tag{1c}$$

$$V_4(x + \Delta x) = V_4(x) - I_4(x)Z_4\Delta x \tag{1d}$$

$$\begin{aligned} I_1(x + \Delta x) = I_1(x) &- \{V_1(x) - V_2(x)\}Y_{12}\Delta x \\ &- V_1(x)Y_{1g}\Delta x - \{V_1(x) - V_3(x)\}Y_{13}\Delta x \\ &- \{V_1(x) - V_4(x)\}Y_{14}\Delta x \end{aligned} \tag{1e}$$

$$\begin{aligned} I_2(x + \Delta x) = I_2(x) &- \{V_2(x) - V_1(x)\}Y_{12}\Delta x \\ &- V_2(x)Y_{2g}\Delta x - \{V_2(x) - V_3(x)\}Y_{23}\Delta x \\ &- \{V_2(x) - V_4(x)\}Y_{24}\Delta x \end{aligned} \tag{1f}$$

$$\begin{aligned} I_3(x + \Delta x) = I_3(x) &- \{V_3(x) - V_4(x)\}Y_{34}\Delta x \\ &- V_3(x)Y_{3g}\Delta x - \{V_3(x) - V_2(x)\}Y_{23}\Delta x \\ &- \{V_3(x) - V_1(x)\}Y_{13}\Delta x \end{aligned} \tag{1g}$$

$$\begin{aligned} I_4(x + \Delta x) = I_4(x) &- \{V_4(x) - V_3(x)\}Y_{34}\Delta x \\ &- V_4(x)Y_{4g}\Delta x - \{V_4(x) - V_2(x)\}Y_{24}\Delta x \\ &- \{V_4(x) - V_1(x)\}Y_{14}\Delta x. \end{aligned} \tag{1h}$$

Dividing through by Δx, taking the limit as Δx approaches zero and recognizing the definition of the derivative, using assumption 3, and writing the resulting eight equations in matrix form, we obtain eq. (2).

$$
\begin{bmatrix} dV_1/dx \\ dV_2/dx \\ dV_3/dx \\ dV_4/dx \\ dI_1/dx \\ dI_2/dx \\ dI_3/dx \\ dI_4/dx \end{bmatrix}
=
\begin{bmatrix}
0 & 0 & 0 & 0 & -Z_1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -Z_2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -Z_3 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -Z_4 \\
-(Y_{1g} + Y_{12} + Y_{13} + Y_{14}) & Y_{12} & Y_{13} & Y_{14} & 0 & 0 & 0 & 0 \\
Y_{12} & -(Y_{2g} + Y_{12} + Y_{23} + Y_{24}) & Y_{23} & Y_{24} & 0 & 0 & 0 & 0 \\
Y_{13} & Y_{33} & -(Y_{3g} + Y_{34} + Y_{13} + Y_{23}) & Y_{34} & 0 & 0 & 0 & 0 \\
Y_{14} & Y_{24} & Y_{34} & -(Y_{4g} + Y_{34} + Y_{14} + Y_{24}) & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ I_1 \\ I_2 \\ I_2 \\ I_4 \end{bmatrix}
\tag{2}
$$

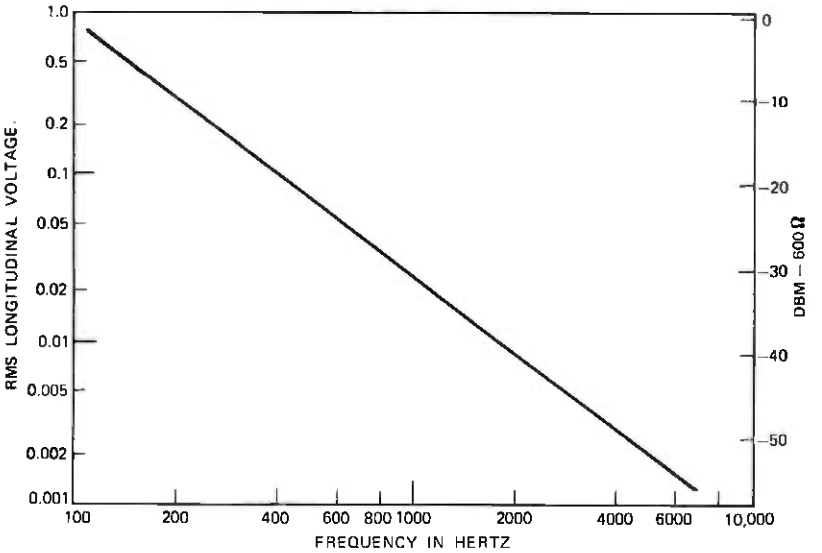


Fig. 1—Longitudinal signal limitations.

3.2 Transmission line equations

Equation (2) above can be written in matrix notation as

$$\frac{dV}{dx} = -\mathbf{Z}I \tag{3a}$$

$$\frac{dI}{dx} = -\mathbf{Y}V, \tag{3b}$$

where

$$V = \begin{bmatrix} V_1(x) \\ V_2(x) \\ V_3(x) \\ V_4(x) \end{bmatrix}, \quad \mathbf{Z} = \begin{bmatrix} Z_1 & 0 & 0 & 0 \\ 0 & Z_2 & 0 & 0 \\ 0 & 0 & Z_3 & 0 \\ 0 & 0 & 0 & Z_4 \end{bmatrix}, \quad I = \begin{bmatrix} I_1(x) \\ I_2(x) \\ I_3(x) \\ I_4(x) \end{bmatrix}$$

Table I — Average crosstalk coupling coefficients for a multipair cable

Coefficient	Average Magnitude (picofarads)	Formula
$C_{M_2M_1}$	7.5	$C_{13} - C_{14} - C_{23} + C_{24}$
$C_{L_2M_1}$	65.2	$2\{C_{13} + C_{14} - C_{23} - C_{24}\}$
$C_{L_1M_2}$	64.0	$2\{C_{13} - C_{14} + C_{23} - C_{24}\}$
$C_{L_2L_1}$	9076.0	$4\{C_{13} + C_{14} + C_{23} + C_{24}\}$

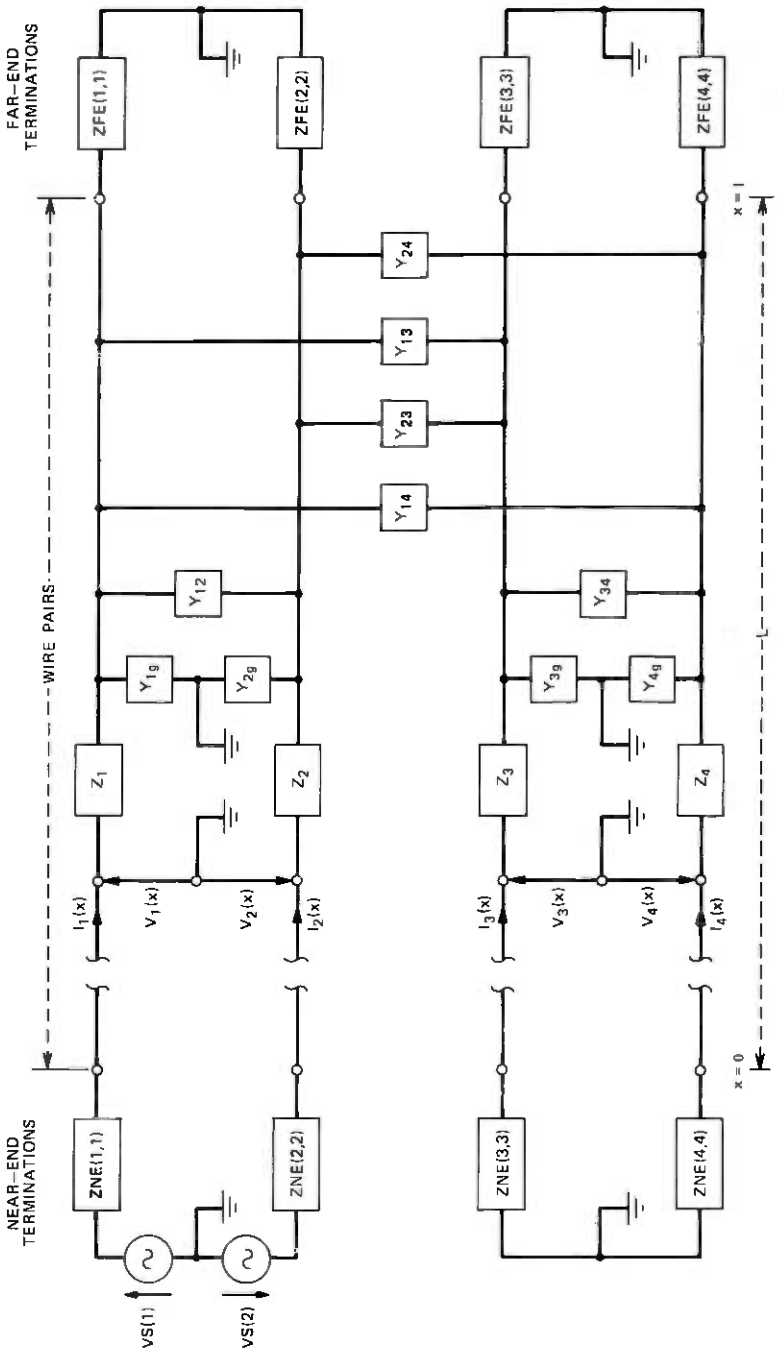


Fig. 2—Model of wire pairs and terminations used to analyze crosstalk coupling.

are the voltage, impedance, and current matrices, respectively, and, using assumption (*ii*),

$$\mathbf{Y} = j\omega \begin{bmatrix} C_{1g} + C_{12} + C_{13} + C_{14} & -C_{12} & -C_{13} & -C_{14} \\ -C_{12} & C_{2g} + C_{12} + C_{23} + C_{24} & -C_{23} & -C_{24} \\ -C_{13} & -C_{23} & C_{3g} + C_{34} + C_{13} + C_{23} & -C_{34} \\ -C_{14} & -C_{24} & & C_{4g} + C_{34} + C_{14} + C_{24} \end{bmatrix}$$

is the admittance matrix. Equations (3a) and (3b) are basic transmission line equations describing the voltage-current relationships between wire pairs. We will use them to calculate crosstalk coupling between wire pairs. They are more conveniently written as a matrix differential equation:

$$\frac{d}{dx}\begin{bmatrix} V \\ I \end{bmatrix} = -\begin{bmatrix} \mathbf{O} & \mathbf{Z} \\ \mathbf{Y} & \mathbf{O} \end{bmatrix}\begin{bmatrix} V \\ I \end{bmatrix}, \tag{4}$$

where $\mathbf{O}$ is a 4×4 null matrix. The solution of eq. (4) is straightforward and is discussed in Appendix A.

3.3 Crosstalk coefficients for various coupling modes

Since there is negligible inductive coupling at voice frequencies, the insight to crosstalk coupling can be obtained from the transmission line equations involving the admittance matrix only. Rewriting eq. (3b) explicitly, we have

$$\begin{bmatrix} dI_1/dx \\ dI_2/dx \\ dI_3/dx \\ dI_4/dx \end{bmatrix} = -j\omega \begin{bmatrix} C_{1g} + C_{12} + C_{13} + C_{14} & -C_{12} & -C_{13} & -C_{14} \\ -C_{12} & C_{2g} + C_{12} + C_{23} + C_{24} & -C_{23} & -C_{24} \\ -C_{13} & -C_{23} & C_{3g} + C_{34} + C_{13} + C_{23} & -C_{34} \\ -C_{14} & -C_{24} & -C_{34} & C_{4g} + C_{34} + C_{14} + C_{24} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \end{bmatrix}. \tag{5}$$

It is clear from eq. (5) that coupling between wire pairs 1–2 and 3–4 could not possibly occur if the four interpair capacitances C_{13}, C_{14}, C_{23}, and C_{24} were all zero regardless of the longitudinal imbalance at the terminations. Furthermore, inspection of eq. (5) shows that the

coupling between the wire pairs is some function of these four interwire capacitances. This is seen by observing the contributions to dI_1/dx and dI_2/dx from V_3 and V_4, and the contributions to dI_3/dx and dI_4/dx from V_1 and V_2.

The insight needed to understand why longitudinal excitation and longitudinally unbalanced terminations increase crosstalk is obtained when eq. (5) is transformed and expressed in terms of the longitudinal and metallic voltages and currents, rather than in terms of the conductor currents and conductor-to-ground voltages. This transformation is easily made because the longitudinal and metallic voltages and currents are linearly related to the conductor voltages and currents. If wire pair 1–2 is now denoted as circuit one and wire pair 3–4 is denoted as circuit two, then the metallic voltages and currents on the two circuits are defined to be

$$V_{1m} = V_1 - V_2, \qquad I_{1m} = \frac{I_1 - I_2}{2}$$

and

$$V_{2m} = V_3 - V_4, \qquad I_{2m} = \frac{I_3 - I_4}{2}.$$

The longitudinal voltages and currents on the two circuits are

$$V_{1L} = \frac{V_1 + V_2}{2}, \qquad I_{1L} = I_1 + I_2$$

and

$$V_{2L} = \frac{V_3 + V_4}{2}, \qquad I_{2L} = I_3 + I_4.$$

Expressed in matrix form, these eight equations become

$$\begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 1 \\ 0 & 0 & -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} V_{1M} \\ V_{1L} \\ V_{2M} \\ V_{2L} \end{bmatrix} \tag{6}$$

and

$$\begin{bmatrix} I_{1M} \\ I_{1L} \\ I_{2M} \\ I_{2L} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix}. \tag{7}$$

Now, by using eqs. (6) and (7), eq. (5) can be expressed in terms of

the longitudinal and metallic voltages and currents as

$$\begin{bmatrix} dI_{1M}/dx \\ dI_{1L}/dx \\ dI_{2M}/dx \\ dI_{2L}/dx \end{bmatrix} = -j\omega \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$\times \begin{bmatrix} C_{1g}+C_{12}+C_{13}+C_{14} & -C_{12} & -C_{13} & -C_{14} \\ -C_{12} & C_{2g}+C_{12}+C_{23}+C_{24} & -C_{23} & -C_{24} \\ -C_{13} & -C_{23} & C_{3g}+C_{34}+C_{13}+C_{23} & -C_{34} \\ -C_{14} & -C_{24} & -C_{34} & C_{4g}+C_{34}+C_{14}+C_{24} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 1 \\ 0 & 0 & \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} V_{1M} \\ V_{1L} \\ V_{2M} \\ V_{2L} \end{bmatrix}, \tag{8}$$

or, performing the matrix multiplications,

$$\begin{bmatrix} dI_{1M}/dx \\ dI_{1L}/dx \\ dI_{2M}/dx \\ dI_{2L}/dx \end{bmatrix} = -\frac{j\omega}{4} \begin{bmatrix} q_{11} & q_{12} & q_{13} & q_{14} \\ q_{21} & q_{22} & q_{23} & q_{24} \\ q_{31} & q_{32} & q_{33} & q_{34} \\ q_{41} & q_{42} & q_{43} & q_{44} \end{bmatrix} \begin{bmatrix} V_{1M} \\ V_{1L} \\ V_{2M} \\ V_{2L} \end{bmatrix}, \tag{9}$$

where the elements of the 4×4 symmetric matrix $\mathbf{Q}$ in eq. (9) are given in Appendix B.

Much useful information can be obtained by simple inspection of some elements of $\mathbf{Q}$. First, the derivative of the metallic current in circuit one due to the metallic voltage in circuit two is proportional to q_{13}, i.e., the crosstalk coupling loss between the metallic circuits is directly related to q_{13}. Thus, the coupling between two metallic circuits, i.e., the metallic-to-metallic coupling, is proportional to

$$C_{M_2M_1} = -q_{13} = C_{13} - C_{14} - C_{23} + C_{24}. \tag{10}$$

This is the capacitance unbalance term first derived by Campbell[1] and used today as one measure of cable quality. Referring again to eq. (9), we see that the derivative of the metallic current in circuit one due to the longitudinal voltage in circuit two is proportional to q_{14} and that the derivative of the metallic current in circuit two due to the longitudinal voltage in circuit one is proportional to q_{32}. In other words, the crosstalk coupling from a longitudinal to a metallic circuit is proportional to

$$C_{L_2M_1} = -q_{14} = 2(C_{13} + C_{14} - C_{23} - C_{24}) \tag{11}$$

or

$$C_{L_1M_2} = -q_{32} = 2(C_{13} - C_{14} + C_{23} - C_{24}). \tag{12}$$

The subscript L_2M_1 means "from the longitudinal mode in circuit two to the metallic mode in circuit one." Also, we can readily see that the derivative of longitudinal current in circuit one resulting from the longitudinal voltage in circuit two is proportional to q_{24}. In other words, the crosstalk coupling between two longitudinal circuits is proportional to

$$C_{L_2L_1} = -q_{24} = 4(C_{13} + C_{14} + C_{23} + C_{24}). \tag{13}$$

Using these four coupling coefficients, it is now possible to compare the difference in crosstalk loss between two metallic circuits, a longitudinal and a metallic circuit, and two longitudinal circuits. This comparison was made for one cable and the results are discussed in the next section.

3.4 Comparison of crosstalk using the coupling coefficients

One good feature of the four coupling coefficients given in eqs. (10) through (13) is that they are easily measured. Hence, they provide a simple method for comparing the difference in crosstalk between two metallic circuits, a longitudinal and a metallic circuit, and two longitudinal circuits. To make such a comparison, it is necessary to have data on the interwire capacitances, C_{13}, C_{14}, C_{23}, and C_{24}, for real cable. Such measurements were made in 1968 on a 22-gauge, pulp-insulated cable manufactured by Western Electric. These measurements were made on many different 50-pair binder groups.* The data on interwire capacitances were taken for random samples out of the 1225 possible sets† of interwire pair combinations within each binder group.

Using these data, the average value of the four coupling coefficients were calculated and are given in Table I. These show that, on the average, the coupling between two metallic circuits is significantly less than the coupling between a longitudinal and a metallic circuit, and that the coupling between two longitudinal circuits is by far the greatest. *Hence, the fundamental reason why terminal longitudinal imbalance increases crosstalk is that longitudinal imbalance causes excitation of the longitudinal circuit.*

Comparison of the values of the coupling coefficients made so far does not provide any quantitative estimate of the amount of the differences in crosstalk losses to be expected. Such an estimate can be obtained by using the coupling coefficients for individual wire-pair combinations to construct distributions of 1000-Hz near-end crosstalk

---

* A binder group is a unit of 12, 16, 20, 25, 50, or 100 twisted wire pairs bound together within a cable.

† For a 50-wire pair cable there are $n(n-1)/2 = 50(49)/2 = 1225$ possible two-wire pair combinations. The sample sizes ranged from 200 to 600 pair combinations.

loss. This was done for the metallic-to-metallic, longitudinal-to-metallic, and longitudinal-to-longitudinal crosstalk loss distributions by using the formula given in Ref. 4,

$$N_l = 20 \log_{10} \left[\frac{j\omega C_u Z_0}{8} \right],$$

where C_u is the capacitance unbalance, i.e., $C_{M_2M_1}$, $C_{L_2M_1}$, or $C_{L_2L_1}$, ω is the radian frequency in Hertz, and Z_0 is the characteristic impedance of 22-gauge pulp. The inductive contribution is neglected. The distributions are shown in Fig. 3. The rms crosstalk loss corresponds to that loss which would result in the average crosstalk power in watts. Consequently, crosstalk power transferred between two circuits with crosstalk loss equal to the rms value would be the average crosstalk power. The rms values are 105.2-dB loss between metallic circuits and

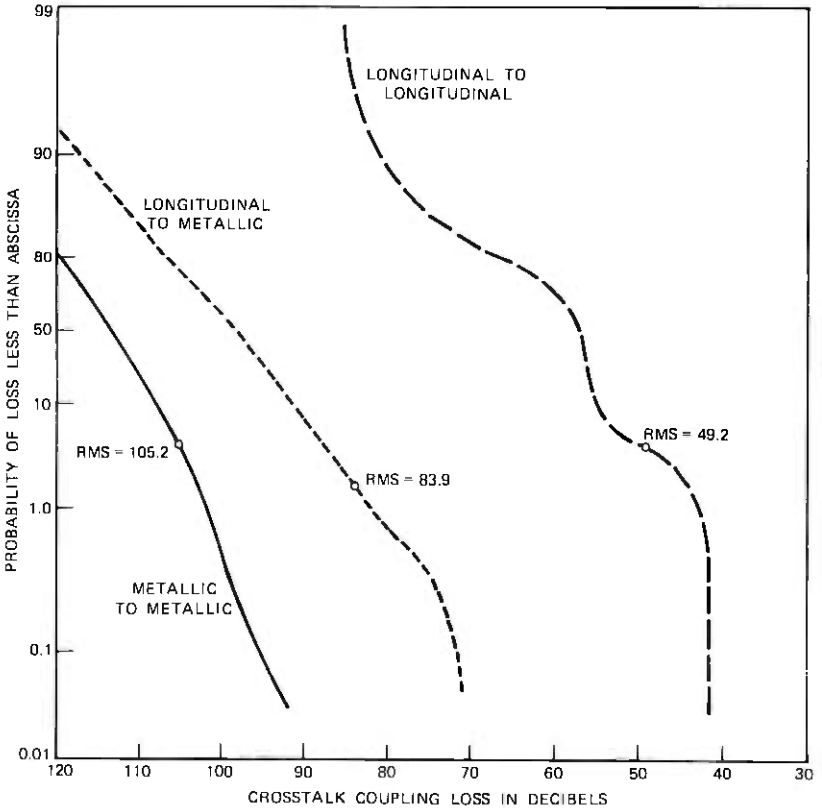


Fig. 3—Computed 1000-Hz near-end crosstalk based on the interwire capacitances of 416 pair combinations of 1319 ft of 22-gauge pulp cable.

only 83.9-dB loss between a longitudinal and a metallic circuit. Hence, the new coefficient $C_{L_2M_1}$ predicts that longitudinal excitation of a wire pair in the cable measured causes 21.3 dB (105.2 − 83.9) more average crosstalk power in a metallic circuit than metallic excitation. Similarly, the rms loss between two longitudinal circuits, predicted by $C_{L_2L_1}$, is 49.2 dB, which is 56 dB (105.2—49.2) less than the metallic-to-metallic loss. These results are compared in Section 3.6 to crosstalk loss measurements made on another cable. In Section 3.5, these results are compared to results obtained from computer solution of the transmission line equations (9), i.e., by simulation of the cable.

3.5 Numerical solutions

A second, more difficult method of calculating the crosstalk between the various modes is direct solution of the transmission line equations on a computer with an appropriate set of boundary conditions.

A computer program has been written to solve these equations that simulates a cable of the same length and identical characteristics of the Western Electric cable used to obtain the coupling coefficients. Two conditions of interest were simulated on the computer. First, metallic excitation by a balanced 1000-Hz signal generator in series with a 600-ohm resistance was applied to a pair, denoted the disturbing pair, and all other pairs were terminated metallically with 600-ohm resistors from tip to ring. Second, the same conditions applied except one wire of the disturbing pair was grounded. This resulted in a degradation of 14.5 dB in the rms value of the balanced near-end crosstalk loss distribution.

To compare the degradation in crosstalk obtained by the two methods, i.e., coupling coefficients versus numerical solutions, it is necessary to note that grounding a wire connected to a signal generator produces a longitudinal voltage that is one-half the value of the applied metallic voltage. This follows directly from the definitions of longitudinal and metallic voltages in terms of the voltage from each wire pair conductor to ground, eq. (6) with $V_2 = 0$. Hence, an approximate 6-dB adjustment must be made when using the longitudinal-to-metallic coupling coefficients $C_{L_1M_2}$ and $C_{L_2M_1}$, which predict a 15.3-dB degradation in the near-end rms crosstalk loss at 1000 Hz due to grounding, as compared to 14.5 dB predicted by the numerical computation. This good agreement suggests that the new capacitive coupling coefficients do provide a simple but reliable method of predicting the degradation in crosstalk performance for a particular cable when its terminations are unbalanced.

3.6 Comparison with measured data

In 1962, measurements of the degradation of near-end crosstalk loss caused by grounding one conductor of the disturbing or disturbed pairs on randomly selected pair combinations in a 7500-ft length of a 22-gauge multipair pulp-insulated exchange grade trunk cable were made by Bell Laboratories. This was a real working cable between Oceanside and Vista, California. All load coils in the section under test here were first removed and the cable ends spliced. The results of these measurements are shown in Fig. 4. The results reveal, at 940 Hz,* a degradation of about 19.4 dB in the near-end rms crosstalk loss when a ground was applied to one wire of either the disturbing or the disturbed pairs.

Since grounding one conductor causes a longitudinal voltage excitation that is one-half the metallic voltage, a 6-dB numerical adjustment was made on the measurements to predict that the rms crosstalk loss between a longitudinal and a metallic circuit is 25.4 dB worse than the rms loss between metallic circuits. This is compared to 21.3 dB obtained using the coupling coefficients for the cable discussed in the previous section. This 4.1-dB difference may be due to the fact that the two cables were not the same, each having different value parameters characterizing them as well as different lengths.

When one conductor of both disturbing and disturbed pairs were grounded, the measured rms crosstalk loss was 32.6 dB, as shown in Fig. 4. This is the loss between the two longitudinal circuits† and, as can be seen, it is 61.5 dB less than this rms loss between the metallic circuits. This measured difference compares favorably to the calculated difference of 56 dB as shown in Fig. 3. The 5.5-dB difference may be due to cable differences. In conclusion, direct crosstalk measurements on another cable substantially support the analytical method for calculating crosstalk using the coupling coefficients or computer simulations.

3.7 Metallic-to-longitudinal conversion because of wire pair imbalance

So far, we have analyzed the effect of direct longitudinal excitation of wire pairs on crosstalk between pairs. This excitation results when longitudinally unbalanced terminations are used. However, now we discuss how longitudinal excitation can also result because of "pair longitudinal imbalance," which is defined as any lack of symmetry

* 940 Hz is close enough to 1000 Hz to permit direct comparison with calculated results.

† Referring to the definitions of longitudinal and metallic voltages, it is simple to show that the coupling loss for the longitudinal-to-longitudinal mode is the same as for both pairs grounded.

in the wire pairs with respect to ground or with respect to each other. Such asymmetry can cause part of the metallic signal to be converted to a longitudinal excitation even when there is perfect longitudinal balance at the terminations.

To understand the causes of wire pair longitudinal imbalance, again refer to eq. (9). Perfect pair balance is the condition that exists whenever a metallic signal does not excite the longitudinal modes in either the disturbing or disturbed wire pair. This requirement can be met if and only if

$$q_{21} = q_{23} = q_{41} = q_{43} = 0.$$

These last four conditions are satisfied if

$$C_{1g} = C_{2g} \tag{14a}$$

$$C_{3g} = C_{4g} \tag{14b}$$

$$C_{14} = C_{23} \tag{14c}$$

$$C_{13} = C_{24}. \tag{14d}$$

Equations (14a) and (14b) are necessary since, for example, if C_{1g} were not equal to C_{2g}, there would be a lack of longitudinal symmetry in wire pair one even if the terminations were all perfectly balanced. Equations (14c) and (14d) imply that equal and opposite currents are coupled (metallic-to-metallic crosstalk) from each of the wires in the disturbing pair to the disturbed pair preserving the pair symmetry.

In other words, if the conditions of eqs. (14) are met and all the terminations are balanced, then all the currents are strictly confined to the metallic circuits. This is not to say that crosstalk cannot occur. It means that only one of the three kinds of coupling can occur, i.e., from metallic circuit to metallic circuit. In fact, the crosstalk will then be proportional to Campbell's capacitance unbalance expression which simplifies to

$$C_{M_2M_1} = C_{13} + C_{24} - C_{14} - C_{23} = 2(C_{13} - C_{14}). \tag{15}$$

Cable data reveal that the capacitances to ground for wire pairs are nearly equal, their differences on the average being less than 2 percent of their magnitude. The percent differences in the interwire capacitances are larger (e.g., 10 percent), but they are much smaller than the capacitances to ground. This suggests that metallic-to-longitudinal conversion of signals due to the cable characteristics alone is small. Computer simulation of wire pairs, using eq. (9) and assuming balanced terminations, supports this suggestion. To put it another way, the high quality of manufactured multipair cable used in the Bell System ensures excellent pair longitudinal balance. The small imbalance in

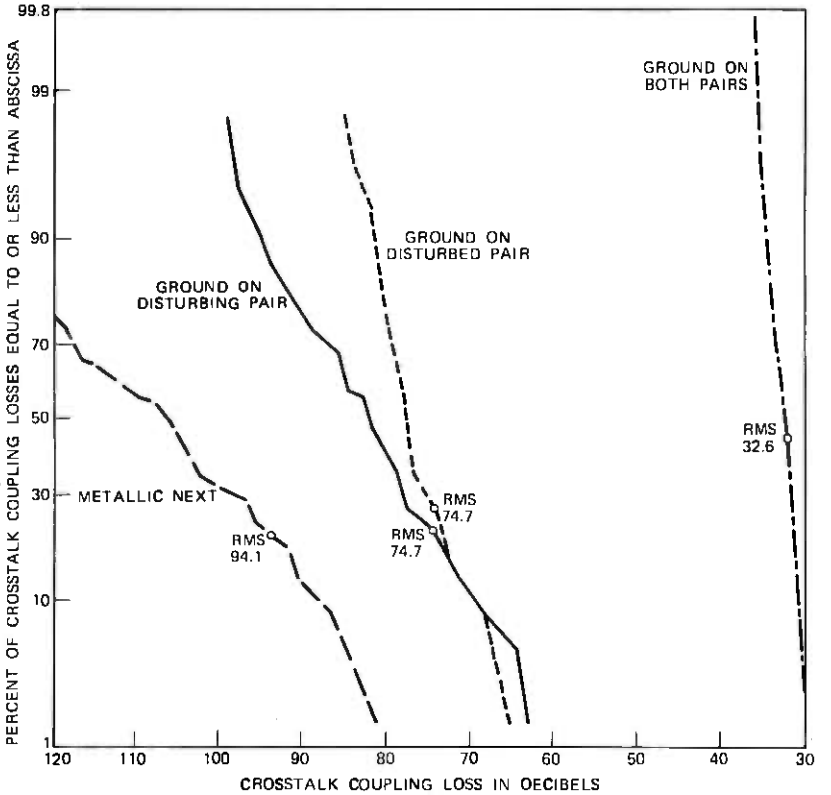


Fig. 4—940 cycles near-end crosstalk coupling loss measured on 60 pair combinations within a 51-pair unit of a DSAC-202 cable (cable length 7500 ft). All grounds are applied at measure end.

the pairs is rarely a significant factor in the contribution to longitudinal voltages that degrade crosstalk. Substantial conversion to longitudinal modes does occur when there is imbalance at the terminations, as revealed by our analysis and, consequently, it is necessary to place limits on permissible terminal longitudinal imbalance.

IV. LONGITUDINAL BALANCE REQUIREMENTS

Crosstalk energy can reach the metallic mode in the disturbed circuit, circuit two, from an applied metallic signal in the disturbing pair, circuit one, in three different ways.

(*i*) Direct coupling from a metallic signal in circuit one to a metallic disturbance in circuit two.

(*ii*) Conversion of the metallic signal of circuit one to a longitudinal signal in circuit one because of an unbalanced

termination* in that circuit, then coupling of the longitudinal signal in circuit one to a metallic disturbance in circuit two.

(*iii*) Conversion of the metallic signal in circuit one to a longitudinal signal in circuit one because of an unbalanced termination, coupling of the longitudinal signal in circuit one to a longitudinal signal in circuit two, and, finally, conversion of the longitudinal signal of circuit two back to a metallic disturbance in circuit two due to an unbalanced termination on circuit two.

The crosstalk described in (*i*) above is independent of the imbalance at the terminations. It is the result of the capacitance unbalance $C_{M_2M_1}$ between the individual wire pairs and there is little more that can practically be done to circuits to reduce it. *The important thing is to make sure that any equipment that is connected at the terminations of the cable does not degrade the low levels of crosstalk that currently exist by introducing longitudinal excitations.*

4.1 Longitudinal balance requirement

The data on the vulnerability of cable to longitudinal imbalance have been obtained by measurements made on two different cables.† Measurements on the cable in California, with the 6-dB numerical adjustment, revealed that longitudinal signals, on the average, crosstalk into adjacent wire pairs with 25.4-dB less coupling loss for that cable than do metallic signals. The data on the Western Electric reel of cable, used in the newly derived capacitance unbalance formulas, showed 21.3-dB less coupling loss for longitudinal signals.

The definition of longitudinal balance, for application to crosstalk performance, is repeated:

$$\mathrm{BAL}_{\mathrm{M-L}}(f) = 20 \log_{10} \left| \frac{e_M(f)}{e_L(f)} \right| ,$$

where e_L is the longitudinal voltage produced when a *metallic* voltage e_M is applied at any frequency f. The measurements made on the cable in California establish the more stringent longitudinal balance requirement, and it shall therefore be assumed that rms longitudinal-to-metallic crosstalk loss is 25 dB less than metallic-to-metallic cross-

* Conversion because of imbalance in the cable itself can be neglected, as discussed in Section 3.7.

† Subsequent to the beginning of this investigation, measurements on the vulnerability of one other cable to crosstalk because of longitudinal imbalance have been made. These measurements do not alter the conclusions reached by using the data on the first two cables only.

talk loss. This implies that, if a network which is sending a metallic signal has a balance of about 25 dB, the longitudinal signal developed because of this imbalance may contribute the same amount of crosstalk power to nearby cable pairs as the direct metallic signal applied to it. It is not known exactly how the two components of the crosstalk produced by metallic and longitudinal signals will add, that is, on a voltage basis, a power basis, or somewhere in between. However, a longitudinal signal developed because of imbalance is likely to be correlated to the metallic signal causing it. Hence, it will be assumed that the signals add approximately on a voltage basis.

What is needed is a balance such that the contribution to crosstalk power because of imbalance is small compared to the crosstalk that exists when a metallic signal is applied. For illustrative purposes, it is assumed that an increase of 1.0 dB is not too noticeable and is thus a permissible contribution. In Fig. 5, which shows how two voltages expressed in dB are added, it is seen that, in order for the power in a signal to be increased by no more than 1.0 dB because of the presence of a second signal, the voltage difference must be over 17 dB. Thus, a longitudinal balance of approximately 42 dB is required (we will use 40 dB) to ensure that crosstalk is increased by no more than this amount, due to the longitudinal-to-metallic coupling path, type (*ii*), described at the beginning of Section IV.

We now show that the crosstalk resulting from the coupling path described as type (*iii*) is less severe and has no bearing in determining the balance requirement. To do this requires discussing a second measure of balance, that known as longitudinal-to-metallic balance.

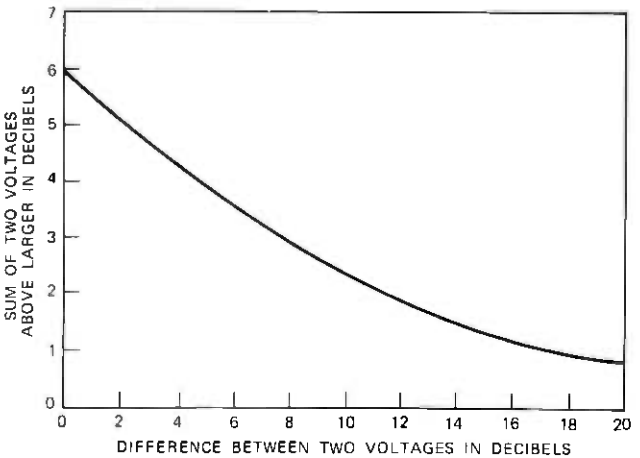


Fig. 5—Sum of two voltages expressed in decibels.

4.2 Longitudinal-to-metallic balance

In addition to the possibility of severely degrading the crosstalk levels already occurring in the telephone network, there is a second important reason why high longitudinal balance at the terminations must be maintained. This is to reduce the metallic noise resulting from power line induction. Here, the disturbing signals are longitudinal in nature and, to assure good performance of the user's circuit, the conversion loss from a longitudinal noise signal on his circuit to a metallic signal on his circuit must be large. A measure of this conversion loss, defined as longitudinal-to-metallic balance, is:

$$\mathrm{BAL}_{\mathrm{L-M}}(f) = 20 \log_{10} \left| \frac{e_L(f)}{e_M(f)} \right| ,$$

where e_L is the applied longitudinal voltage source and e_M is the resulting metallic signal. The balance subscript "L − M" means a conversion from an applied longitudinal signal to a metallic signal. It is important to note that the two measures of balance, $\mathrm{BAL}_{\mathrm{M-L}}$ and $\mathrm{BAL}_{\mathrm{L-M}}$, are not equal, i.e., reciprocity does not necessarily apply. Moreover, they are not necessarily correlated.

Generally, to assure good performance, the minimum balance $\mathrm{BAL}_{\mathrm{L-M}}$ of a termination is well in excess of 40 dB over the voiceband. Consequently, we use 40 dB as a lower bound on the $\mathrm{BAL}_{\mathrm{L-M}}$ of terminations, keeping in mind that this in no way establishes 40 dB as the necessary performance minimum. Discussion of that topic is outside the scope of this paper.

Using the assumed minimum bound on $\mathrm{BAL}_{\mathrm{L-M}} = 40$ dB of terminations on the disturbed circuit, it is now shown that the crosstalk coupling loss path (*iii*) is at least 20 dB less than the metallic-to-metallic path and, hence, is not a factor. It is also assumed that the balance $\mathrm{BAL}_{\mathrm{M-L}}$ of the terminations on the disturbing circuit is 40 dB, determined from the crosstalk requirement because of the coupling path of type (*ii*). Now, since there is approximately 60 dB less crosstalk loss between two longitudinal circuits than between two metallic circuits, the difference in the losses between type (*i*) and type (*iii*) paths is $\mathrm{BAL}_{\mathrm{M-L}}$ (disturbing circuit) -60 dB $+$ $\mathrm{BAL}_{\mathrm{L-M}}$ (disturbed circuit), or at least 20 dB.

V. REQUIREMENTS ON LONGITUDINALLY APPLIED SIGNALS

So far, we have considered how longitudinally unbalanced terminations can cause increased crosstalk, because they longitudinally excite a wire pair. We have recognized that it is this longitudinal excitation that is the fundamental cause of the increased crosstalk, and we have

recommended a metallic-to-longitudinal balance, BAL_{M-L}, limit for termination of 40-dB minimum to restrict the amount of longitudinal excitation produced. However, it is also possible to longitudinally excite wire pairs directly from a voltage source connected between the tip and ring of a wire pair and the cable sheath, or ground. Such direct excitation must also be limited, because it too causes crosstalk in a disturbed metallic circuit in two ways:

(*i*) Direct coupling from the longitudinal mode in circuit one to the metallic mode in circuit two.

(*ii*) Coupling of the longitudinal mode in circuit one to the longitudinal mode in circuit two and conversion of energy in the longitudinal mode of circuit two to the metallic mode in circuit two because of an unbalanced termination in circuit two.

Since the effect of directly applying longitudinal signals is the same as longitudinal signals arising from metallic-to-longitudinal imbalance, and since the rms crosstalk loss for this type of signal is on the average 25 dB less than metallic signals, longitudinal voltage limits should be 40 dB more restrictive than metallic voltage limits. Figure 1 shows the restriction on longitudinally applied voltages as a function of frequency. It is based on the restrictions already placed on metallic voltages determined by a previous study at Bell Laboratories* and the 40-dB restriction determined here.

VI. SUMMARY

The following has been accomplished in this paper:

(*i*) Three new capacitive coupling coefficients have been derived that provide a simple but reliable method of predicting the degradation in crosstalk performance for a particular cable when its terminations are unbalanced.

(*ii*) It has been demonstrated that a metallic-to-longitudinal balance requirement of 40 dB or more for any terminations connected into network should not noticeably increase the low levels of crosstalk that are already present.

(*iii*) A requirement has been established on longitudinally applied signals that if met should not degrade crosstalk performance.

VII. ACKNOWLEDGMENTS

The author would like to thank J. R. Rosenberger and the reviewers for their helpful comments.

* Restrictions on metallic voltages are given in terms of permissible power into 600 ohms in Ref. 3.

APPENDIX A

In eq. (4), let $T(x)$ be an 8×1 column matrix where

$$T = \begin{bmatrix} V(x) \\ I(x) \end{bmatrix} \tag{16}$$

and let $\mathbf{A}$ be an 8×8 constant matrix (not a function of x),

$$\mathbf{A} = \begin{bmatrix} \mathbf{O} & \mathbf{Z} \\ \mathbf{Y} & \mathbf{O} \end{bmatrix}. \tag{17}$$

Then

$$\frac{dT(x)}{dx} = -\mathbf{A}T. \tag{18}$$

The solution to this matrix differential equation[6] is known to be

$$T(x) = \exp(-\mathbf{A}x)\,T(0).$$

Since the parameters that characterize the line are independent of x, it is readily seen by solving eq. (4) that

$$\begin{bmatrix} VFE \\ CFE \end{bmatrix} = \exp\left\{-\begin{bmatrix} \mathbf{O} & \mathbf{Z} \\ \mathbf{Y} & \mathbf{O} \end{bmatrix} l\right\} \begin{bmatrix} VNE \\ CNE \end{bmatrix}, \tag{19}$$

where the far-end voltages and currents where $x = l$ are

$$VFE = \begin{bmatrix} V_1(l) \\ V_2(l) \\ V_3(l) \\ V_4(l) \end{bmatrix}, \qquad CFE = \begin{bmatrix} I_1(l) \\ I_2(l) \\ I_3(l) \\ I_4(l) \end{bmatrix}$$

and the near-end voltages and currents are

$$VNE = \begin{bmatrix} V_1(0) \\ V_2(0) \\ V_3(0) \\ V_4(0) \end{bmatrix}, \qquad CNE = \begin{bmatrix} I_1(0) \\ I_2(0) \\ I_3(0) \\ I_4(0) \end{bmatrix}.$$

With the eight equations given in (19) and a knowledge of the boundary relations at the terminations, we can characterize the model of the system at each point in space (x) by a vector pair of voltages $V(x)$ and pair currents $I(x)$.

The matrix exponential,

$$\exp\left\{-\begin{bmatrix} \mathbf{O} & \mathbf{Z} \\ \mathbf{Y} & \mathbf{O} \end{bmatrix} l\right\},$$

may be evaluated in closed form. With a closed-form representation, the voltages and currents can be expressed in closed form and more complex structures such as spliced cable systems can be simulated. The way to explicitly determine the matrix exponential is to use the fact that

$$\mathbf{Q} = -\begin{bmatrix} \mathbf{O} & \mathbf{Z} \\ \mathbf{Y} & \mathbf{O} \end{bmatrix}$$

satisfies its own characteristic equation (Cayley-Hamilton Theorem[5]). Then

$$\exp \mathbf{Q} = \sum_{k=0}^{7} \alpha_k \mathbf{Q}^k.$$

Replacing Q by a diagonal matrix consisting of the eight eigenvalues of Q enables us to solve for α_k. However, for fairly short unspliced cable systems we may use the first few terms in a power series, i.e.,

$$\exp[-\mathbf{Q}l] = \mathbf{I} - l\mathbf{Q} + \frac{l^2}{2}\mathbf{Q}^2 - \cdots,$$

where I is the identity matrix, to accurately approximate the matrix exponential as was done for the cable in the numerical solutions section.

A.1 Boundary conditions

For any two wire pairs within a cable, four sets of current-voltage relationships exist at the wire terminations. Referring back to Fig. 2, we define the near end to be the subscriber side of the loop with its termination where $x = 0$, and the far end, where $x = l$, to be the other termination, possibly a central office. The disturbing pair will always be designated wire pair 1–2 with a generator of some kind at the near end, and the disturbed pair will be designated 3–4. Suppose the generator is two voltage sources each grounded at one end and in series with an impedance and the remaining terminations consist each of two complex impedances to ground shown in Fig. 2. Then we have the eight relations at the boundaries

$$V_1(0) = VS(1) - ZNE(1, 1)I_1(0) \tag{20a}$$

$$V_2(0) = VS(2) - ZNE(2, 2)I_2(0) \tag{20b}$$

$$V_3(0) = -ZNE(3, 3)I_3(0) \tag{20c}$$

$$V_4(0) = -ZNE(4, 4)I_4(0) \tag{20d}$$

$$V_1(l) = ZFE(1, 1)I_1(l) \tag{20e}$$

$$V_2(l) = ZFE(2, 2)I_2(l) \tag{20f}$$

$$V_3(l) = ZFE(3, 3)I_3(l) \tag{20g}$$

$$V_4(l) = ZFE(4, 4)I_4(l), \tag{20h}$$

which may be written in matrix form as

$$\boldsymbol{VNE} = -\mathbf{ZNE}\ \boldsymbol{CNE} + \boldsymbol{VS} \tag{21a}$$

$$\boldsymbol{VFE} = \mathbf{ZFE}\ \boldsymbol{CFE}, \tag{21b}$$

where

$$\mathbf{ZNE} = \begin{bmatrix} ZNE(1, 1) & 0 & 0 & 0 \\ 0 & ZNE(2, 2) & 0 & 0 \\ 0 & 0 & ZNE(3, 3) & 0 \\ 0 & 0 & 0 & ZNE(4, 4) \end{bmatrix},$$

$$\mathbf{ZFE} = \begin{bmatrix} ZFE(1,1) & 0 & 0 & 0 \\ 0 & ZFE(2, 2) & 0 & 0 \\ 0 & 0 & ZFE(3, 3) & 0 \\ 0 & 0 & 0 & ZFE(4, 4) \end{bmatrix}.$$

and*

$$\boldsymbol{VS}^T = [VS(1), VS(2), 0, 0].$$

We can solve the 16 equations (19) and (21) and determine $\boldsymbol{VNE}$ and $\boldsymbol{CNE}$. Now we have the model completely characterized by the vector pair of voltages $\boldsymbol{V}(x)$ and $\boldsymbol{I}(x)$ via the equation

$$\begin{bmatrix} \boldsymbol{V} \\ \boldsymbol{I} \end{bmatrix} = \exp\left\{-\begin{bmatrix} \mathbf{O} & \mathbf{Z} \\ \mathbf{Y} & \mathbf{O} \end{bmatrix} x\right\} \begin{bmatrix} \boldsymbol{VNE} \\ \boldsymbol{CNE} \end{bmatrix}.$$

It should be pointed out that the terminations are not always simple impedances to ground. For instance, for a second type of termination, where an ordinary telephone set is connected to a wire pair, tip, and ring, there is no direct conducting path to ground. If the impedance to ground from the tip and ring is assumed to be infinite, then we cannot write a simple impedance matrix relating the current to voltage as in eqs. (21a) or (21b). As a result, there is a rather tedious but straightforward rearrangement of eqs. (19) and (21). Finally, a third type of termination could be a central office that will also require modification of the impedance matrix. All three of these types of terminations have been simulated in computer programs.

* The superscript T means transpose.

APPENDIX B

The elements of the admittance matrix in eq. (13) are given below.

$$
\begin{aligned}
q_{11} &= \{C_{1g} + C_{2g} + 4C_{12} + C_{13} + C_{14} + C_{23} + C_{24}\} \\
q_{12} &= q_{21} = 2\{(C_{1g} - C_{2g}) + (C_{13} + C_{14} - C_{23} - C_{24})\} \\
q_{13} &= q_{31} = -(C_{13} - C_{14} - C_{23} + C_{24}) \\
q_{14} &= q_{41} = -2(C_{13} + C_{14} - C_{23} - C_{24}) \\
q_{22} &= 4\{(C_{1g} + C_{2g}) + (C_{13} + C_{14} + C_{23} + C_{24})\} \\
q_{23} &= q_{32} = -2(C_{13} - C_{14} + C_{23} - C_{24}) \\
q_{24} &= q_{42} = -4(C_{13} + C_{14} + C_{23} + C_{24}) \\
q_{33} &= \{C_{3g} + C_{4g} + 4C_{34} + C_{13} + C_{14} + C_{23} + C_{24}\} \\
q_{34} &= q_{43} = 2\{(C_{3g} - C_{4g}) + (C_{13} - C_{14} + C_{23} - C_{24})\} \\
q_{44} &= 4\{(C_{3g} + C_{4g}) + (C_{13} + C_{14} + C_{23} + C_{24})\}.
\end{aligned}
$$

REFERENCES

1. "Dr. Campbell's Memoranda of 1907 and 1912," B.S.T.J., *14*, No. 4 (October 1935), pp. 558–572.
2. G. J. Foschini, "Crosstalk in Outside Plant Cable Systems," B.S.T.J., *50*, No. 7 (September 1971), pp. 2421–2448.
3. "Transmission Specifications for Private Line Metallic Circuits," Bell System Technical Reference, PUB 43401, December 1971.
4. G. S. Eager, Jr., L. Jachimowicz, I. Kolodny, and D. E. Robinson, "Transmission Properties of Polyethylene-Insulated Telephone Cables at Voice and Carrier Frequencies," Transaction of the AIEE, Communications and Electronics, *78*, Part 1, November 1959, pp. 618–640.
5. L. A. Zadeh and C. A. Desoer, *Linear System Theory: The State Space Approach*, Appendix C, New York: McGraw-Hill, 1963.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 54, No. 7, September 1975
Printed in U.S.A.

Wideband Amplifier Design Using Major Multiloop Feedback Techniques

By T. J. APRILLE, JR.

(Manuscript received February 11, 1974)

Multiloop feedback has heretofore been ignored as a means of obtaining shaped gain amplifiers. In this paper, a theoretical basis is developed for using dual major loop feedback amplifiers to obtain shaped power gain with input and output reflection coefficient constraints. From the theoretical results, practical design procedures can easily be developed and one such procedure is discussed.

The aim of this study was an alternative to the "brute force" termination technique of realizing input and output impedance matches. The development is otherwise unique in that it uses no hybrid transformers for beta circuit coupling or for realization of the reflection coefficient constraints.

I. INTRODUCTION

Wideband feedback amplifier design has heretofore mainly been accomplished by the use of single major loop feedback techniques.[1–4] Major loop feedback implies that the current or voltage on the input to the basic amplifying element is manipulated by the current or voltage that appears on the output of the basic amplifying element. The design concept follows the classical feedback design procedure of assuming a unilateral forward amplifying element of voltage gain μ and a feedback path with voltage gain β. Existing multiloop feedback techniques have been primarily concerned with stability considerations of "tandem" [5] and minor multiloop[2,3] feedback arrangements.

In many applications, input and output impedance matching of the amplifier is necessary. The communications amplifier is one such example, since it requires very low levels of signal interference due to input or output impedance mismatch. The classical single-loop feedback techniques offer little help in designing for the impedance matching constraint. This is due to the fact that the more loop gain in a single-loop feedback circuit, the more extreme (zero or infinite) the input and output impedance becomes.[1–3] Two techniques that are

used to circumnavigate this problem are "brute force" terminations and bridge couplings.

The brute force approach obtains the impedance match by placing a resistor in series with the input (or output) for a feedback amplifier with zero input (or output) impedance. A parallel resistor is used for the infinite input or output impedance case.

The use of a balanced resistor bridge is also useful in obtaining an impedance match. This is accomplished by balancing the bridge components with respect to the input (or output) impedance of the feedback amplifier. The use of a resistive bridge is limited, though, due to the excessive resistive losses associated with such a bridge. A useful four-port device, which exhibits the same qualities as a resistive bridge but with much less through loss, is the hybrid transformer.[6] The impedance match with this device is obtained by manipulation of the two unused port impedances.[7]

Since the hybrid transformer is similar to a bridge, one of the two unused ports can be used for the β return path. This technique is theoretically the best alternative mentioned since a property of such a connection is that the impedance match is improved with the amount of loop gain.[2] This technique has been used to advantage on several communications amplifiers.[8,9]

The limitations of the above alternatives of obtaining an impedance match become evident when other design constraints are investigated. For example, the noise figure of an amplifier is degraded by any loss that exists on the input to the amplifier.[10] Thus, the use of brute force or hybrid transformer coupling causes an increase in noise figure. On the output side, a loss increases the power requirement on the last stage of the amplifier. Even if this is no problem, the resultant increase in the distortion may be. This is due to the fact that second-order distortion power increases twice as fast as fundamental power and third-order distortion power three times as fast.[11] Thus, the losses associated with the matching techniques will increase the power requirement and reduce the linearity of the overall amplifier.

The use of the hybrid transformer in the β path may also cause a stability problem. Since the transformer introduces phase shift, due both to the physical length and techniques of construction, their use is limited at very high frequencies.

Investigation into alternative methods of design is therefore desirable. To this end, this paper presents fundamental concepts on the techniques of using major multiloop feedback in amplifier design. The objective is the design of wideband-frequency-dependent gain amplifiers with input and output match constraints. The design procedure does not use hybrid transformers and attempts to minimize brute force termination techniques.

In Section II, the basic amplifier element is introduced. The analysis that follows is applicable to configurations of active devices that can be modeled by this basic amplifier. A circuit form using this basic amplifier is then introduced. Matched impedances and gain relationships are developed for this circuit form in such a way as to make the open loop gain characteristics evident. This paves the way for an initial design approach that is independent of the loop gain characteristics.

In Section III, a second circuit form of shunt-series feedback using the same basic amplifier is introduced. Matched impedances and gain relationships are again developed. The derivations in this section exactly parallel those of Section II.

In Section IV is given the results of the two previous sections to demonstrate the procedure used to obtain an initial circuit design for a practical amplifier configuration. The configuration treated is that of a cascade of N common emitter transistor stages. It is shown that for N odd, the results of Section II can be used, and for N even, the results of Section III apply. One numerical example is supplied for each case. Two appendices provide the calculations used to derive the results in Sections II and III.

II. SHUNT TRANSADMITTANCE: SERIES TRANSIMPEDANCE FEEDBACK

Each dual-loop feedback amplifier discussed in this paper contains three major components: two feedback networks and one amplifying element. Each major component is assumed to be made up of any number of passive and active elements. Characteristics of importance for the amplifying element component are given in Fig. 1; this abbreviated model is designated a basic amplifier. In this figure, z_x is the input impedance and I_s is a current-controlled current source. I_s is given by the product of a frequency-dependent variable k and the current through z_x.

In Fig. 2, the first multiloop feedback circuit form is given. Series feedback voltage source aI_o sums up the most important characteristic

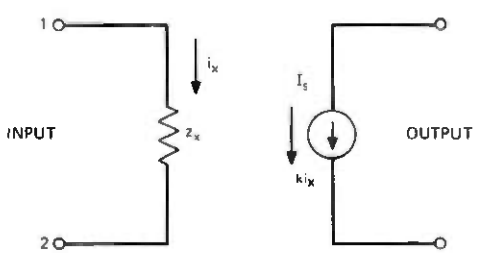


Fig. 1—Basic amplifier.

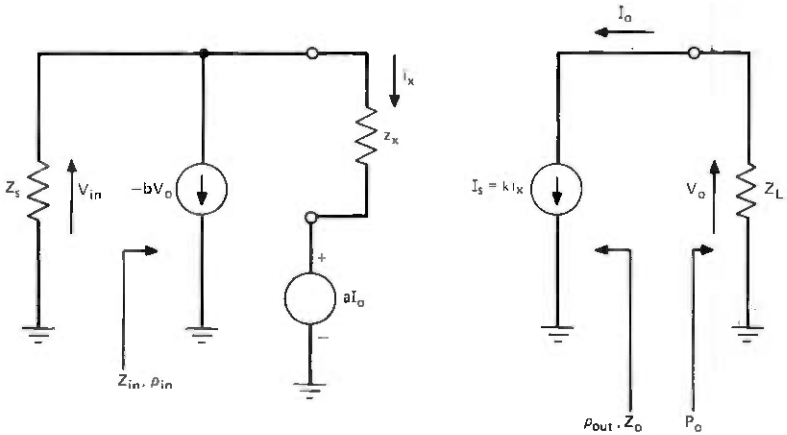


Fig. 2—First feedback form.

of one feedback network. Shunt current source $-bV_o$ likewise is the important characteristic of the second feedback network. Since the series feedback voltage source is dependent upon output current, it represents a transimpedance feedback. Similarly, the shunt current source is dependent on the output voltage yielding a transadmittance feedback. Source and load impedances, Z_s and Z_L, summarize amplifier interaction with the driving circuitry and the loading circuitry, respectively.

2.1 Input and output impedance

Z_{in} and Z_{out}, the input and output impedance, are desired to be matched to Z_s and Z_L, respectively. Thus, Z_{in} and Z_{out} are needed and are given by

$$Z_{in} = \frac{z_x + ka}{1 + kbZ_L}, \tag{1}$$

$$Z_{out} = \frac{z_x + Z_s + ka}{kbZ_s}. \tag{2}$$

If the amplifier gain k is large, then Z_{in} and Z_{out} become

$$Z_{in} = \frac{ak}{bkZ_L} = \frac{a}{bZ_L}, \tag{3}$$

$$Z_{out} = \frac{ak}{bkZ_s} = \frac{a}{bZ_s}. \tag{4}$$

For the matched condition, $Z_{in} = Z_s^*$ and $Z_{out} = Z_L^*$. Using these condi-

tions in eqs. (3) and (4) yields

$$Z_s^* = Z_{in} = \frac{a}{bZ_L}, \tag{5}$$

$$Z_L^* = Z_{out} = \frac{a}{bZ_s}. \tag{6}$$

Substituting $(Z_L^*)^* = Z_L$ from eq. (6) into (5) yields

$$Z_s^* = \frac{a}{b(a^*/b^*Z_s^*)}, \tag{7}$$

which implies

$$\frac{a^*}{b^*} = \left(\frac{a}{b}\right)^* = \frac{a}{b}. \tag{8}$$

Thus, conjugate matching yields the requirement that the ratio of a to b (or more generally ka to kb) must be real. Given this fact, eqs. (5) and (6) are identical, i.e.,

$$Z_s^* Z_L = Z_s Z_L^* = \frac{a}{b}. \tag{9}$$

The imaginary part of $Z_L Z_s^*$ is therefore constrained by

$$\text{Im}\,\{Z_L\}\ \text{Re}\,\{Z_s\} - \text{Re}\,\{Z_L\}\ \text{Im}\,\{Z_s\} = 0. \tag{10}$$

A necessary condition for an amplifier to be absolutely stable is that Z_{in} and Z_{out} be passive.[12] This is satisfied when the real parts of Z_s and Z_L are positive. Thus, the imaginary part of Z_s and Z_L have the same sign, implying that if the matched load impedance is capacitive (inductive), then the matched source impedance is capacitive (inductive).

If the reflection coefficient [reflection coefficient ρ is defined as $(Z - Z_{ref}^*)/(Z + Z_{ref})$][13,14] at the input is evaluated (assuming $Z_s = a/bZ_L^*$), the following is obtained:

$$\rho_{in} = \rho_{in_0}\left(1 + \frac{2ab\ \text{Re}\ (Z_L)}{z_x bZ_L^* + a}k\right)^{-1}. \tag{11}$$

In eq. (11) ρ_{in_0} is the input reflection coefficient when $k = 0$.

Evaluating the return ratio T (Ref. 2) of the circuit in Fig. 2 with respect to the output dependent current source gives

$$T = \frac{-2ab\ \text{Re}\ (Z_L)}{z_x bZ_L^* + a}k. \tag{12}$$

Return difference F (Ref. 2) is defined as $1 - T$; thus, eq. (11) can be rewritten as

$$\rho_{in} = \rho_{in_0}\frac{1}{F}. \tag{13}$$

Therefore, the return loss ($20 \log 1/|\rho|$) is improved with increasing return difference F. Since the output reflection coefficient is given by

$$\rho_{out} = \rho_{out_0} \frac{1}{F}, \tag{14}$$

it also realizes the same improvement with increased return difference.

2.2 Gain equations

The transducer gain for the circuit in Fig. 2 can be calculated when it is assumed that $Z_s = a/bZ_L^*$:

$$|S_{21}|^2 = \frac{1}{|ab|} \frac{|T|^2}{|1 - T|^2}. \tag{15}$$

Again, T is the return ratio and is given by eq. (12).

In eq. (15), T is proportional to k. Thus, for large k, $|S_{21}|^2$ goes to $1/|ab|$. Therefore, eq. (15) can be rewritten as

$$|S_{21}|^2 = |S_{21\infty}|^2 \frac{|T|^2}{|1 - T|^2}, \tag{16}$$

where

$$|S_{21\infty}|^2 = \frac{1}{|ab|}. \tag{17}$$

2.3 Design procedure

In the derivations given thus far, a definite effort has been made to separate the dependence of k. This was done for two reasons: (*i*) to allow an initial design to be effected with k not a variable, and (*ii*) to allow definitive statements to be easily made concerning the effects of k. The former can easily be implemented by assuming $k = \infty$.

In this case of $k = \infty$, eqs. (8), (9), and (17) are relevant. These equations are repeated for convenience:

$$\left(\frac{a}{b}\right)^* = \frac{a}{b} \tag{8}$$

$$Z_s Z_L^* = \frac{a}{b} \tag{9}$$

$$|S_{21\infty}|^2 = \frac{1}{|ab|}. \tag{17}$$

It should be noted that eq. (8) implies that a/b is real, but a and b can be complex. $|S_{21\infty}|^2$ in eq. (17) is the maximum available gain since it is obtained with the input and output matched.

In summary, the design procedure given below could be used when the desired gain g and impedance matches are known.

(*i*) Choose an arbitrary starting Z_s and Z_L such that $Z_s Z_L^*$ is real.
(*ii*) Substitute a from eq. (9) into eq. (17), yielding

$$g = |S_{21\infty}|^2 = \frac{1}{Z_s Z_L^* |b|^2}. \tag{18}$$

Synthesize b such that

$$|b| = \frac{1}{\sqrt{Z_s Z_L^* |g|}}. \tag{19}$$

There are no constraints upon the phase of b except those that may result from stability considerations.

(*iii*) Synthesize a such that

$$a = bZ_s Z_L^*. \tag{20}$$

(*iv*) The value of k is now obtained by considering the practical active devices used to simulate the ideal amplifying element. With k known, the return ratio T [eq. (12)] can be calculated; this yields the obtainable impedance match and gain deviation, eqs. (13), (14), and (16). If the design objectives are not met, the previous calculations should make the necessary changes evident, e.g., lower Z_s or a higher value of k.

III. CURRENT TRANSFER SHUNT; VOLTAGE TRANSFER SERIES FEEDBACK

The last multiloop feedback circuit to be considered is shown in Fig. 3. In this case, the series feedback voltage source is dependent upon the output voltage and thus represents a voltage transfer feedback. Similarly, the shunt current source is a current transfer feedback. The voltage source is given by aV_0 and the current source by $-bI_0$, otherwise Figs. 2 and 3 are identical.

3.1 Input and output impedance

The input and output impedances, when evaluated, are given by

$$Z_{\text{in}} = \frac{z_x - akZ_L}{1 - bk}, \tag{21}$$

$$Z_{\text{out}} = \frac{Z_s + z_x - kZ_s b}{-ka}. \tag{22}$$

For large k, eqs. (21) and (22) become

$$Z_{\text{in}} = \frac{a}{b} Z_L, \tag{23}$$

$$Z_{\text{out}} = \frac{b}{a} Z_s. \tag{24}$$

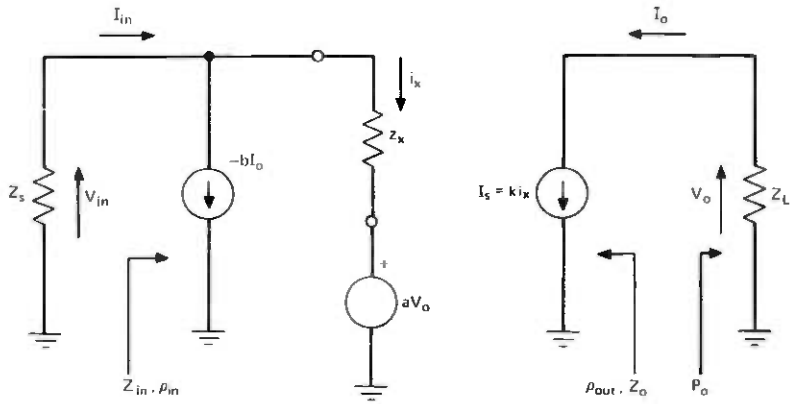


Fig. 3—Second feedback form.

The conditions of input and output matches yield

$$Z_s^* = Z_{in} = \frac{a}{b} Z_L = \frac{a}{b} (Z_{out})^* = \frac{a}{b} \left(\frac{b}{a}\right)^* Z_s^*. \tag{25}$$

Thus,

$$\left(\frac{a}{b}\right)^* = \left(\frac{a}{b}\right), \tag{26}$$

and

$$\frac{Z_s^*}{Z_L} = \frac{Z_s}{Z_L^*} = \frac{a}{b}. \tag{27}$$

Since a/b is real and the real part of Z_s and Z_L are nonnegative, then eq. (27) implies that if Z_s is capacitive (inductive) then Z_L must be inductive (capacitive).

Input and output reflection coefficients can be evaluated along with the return ratio and return difference. The results are shown below for Z_s and Z_L, satisfying eq. (27).

$$T = \frac{2kab \operatorname{Re}(Z_L)}{bz_x + aZ_L^*}. \tag{28}$$

$$F = 1 - T. \tag{29}$$

$$\rho_{in} = \rho_{in_0} \frac{1}{F}; \qquad \rho_{in_0} = \frac{bz_x - aZ_L}{bz_x + aZ_L^*}. \tag{30}$$

$$\rho_{out} = \rho_{out_0} \frac{1}{F} = \frac{1}{F}; \qquad \rho_{out_0} = 1. \tag{31}$$

Thus, as in the case of the first circuit form, the reflection coefficients are improved by the return difference.

3.2 Gain equation

When the load and source impedances satisfy eq. (27), the transducer gain for the circuit in Fig. 3 is given by

$$|S_{21}|^2 = |S_{21\infty}|^2 \frac{|T|^2}{|1 - T|^2}, \tag{32}$$

$$|S_{21\infty}|^2 = \frac{1}{|ab|}. \tag{33}$$

This is the same form as was given in eqs. (16) and (17); thus, the same statements apply to the above equations concerning improvement with feedback.

3.3 Design procedure

Initial circuit design can proceed in a manner similar to the first case. The term k again is assumed equal to infinity; this yields the germane equations summarized below.

$$\left(\frac{a}{b}\right)^* = \frac{a}{b}. \tag{26}$$

$$\frac{Z_s^*}{Z_L} = \frac{a}{b}. \tag{27}$$

$$|S_{21\infty}|^2 = \frac{1}{|ab|}. \tag{33}$$

The four design steps outlined previously apply except as follows.

(*i*) Choose Z_s and Z_L such that Z_s^*/Z_L is real.

(*ii*) Substitute eq. (27) into (33) so that

$$g = |S_{21\infty}|^2 = \frac{1}{\frac{Z_s^*}{Z_L}|b|^2}. \tag{34}$$

Synthesize b such that

$$|b| = \frac{1}{\sqrt{\frac{Z_s^*}{Z_L}|g|}}. \tag{35}$$

(*iii*) Synthesize a such that

$$a = \frac{Z_s^*}{Z_L} b. \tag{36}$$

(*iv*) With k known, the return ratio is obtained from eq. (28). Equations (30), (31), and (32) then yield the obtainable impedance matches and gain deviation, respectively.

IV. DESIGN EXAMPLES

Results obtained in the last two sections will now be applied to a basic amplifier consisting of a cascade of N common emitter transistor stages. Transistors will be assumed to be used in a frequency range well below cutoff. The first case to be treated is for N odd.

4.1 N odd

Consider the circuit given in Fig. 4a. In this circuit, the transistor will be modeled by the circuit given in Fig. 4b. The circuit given in Fig. 4a will now be converted to the form given in Fig. 2. Z_s and Z_L have their obvious counterparts. z_x is given by the impedance from base to ground with aI_0 equal to zero; this is obtained when $I_0 = 0$, which can be obtained by setting α_1 (first stage α) to zero. From the

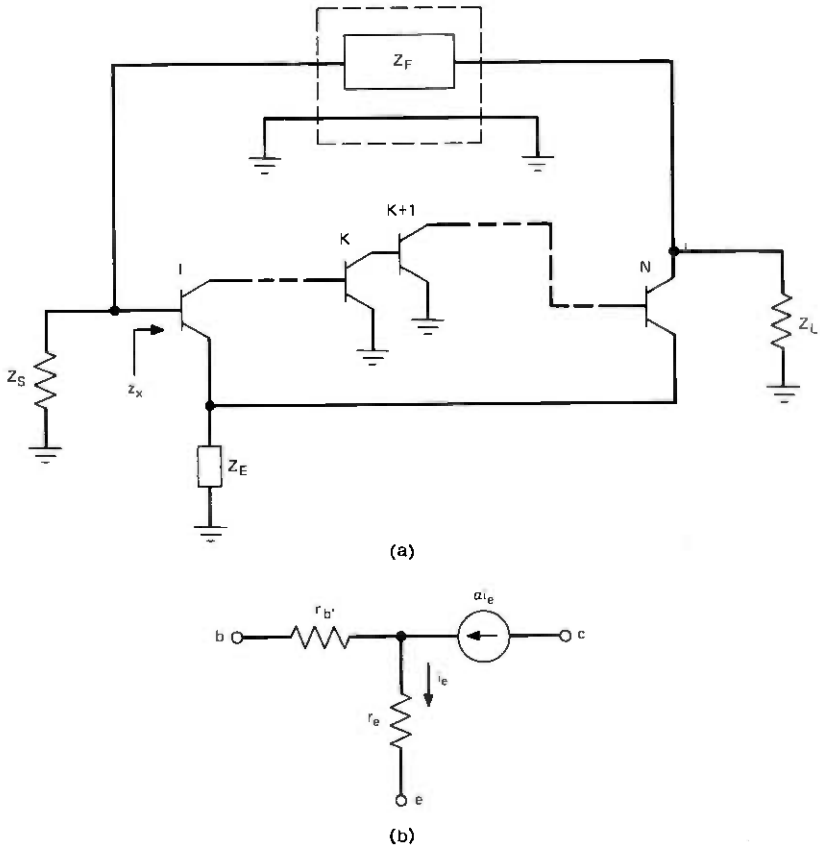


Fig. 4—Design example N odd.

transistor model, this is evidently

$$z_x = r_b' + r_e + Z_E. \tag{37}$$

I_s is given by the collector current on the Nth transistor when the base current on the first is unity. With the cascade of N transistors, this yields

$$I_s = \beta_1, \cdots, \beta_N i_x; \qquad \beta_i = \frac{\alpha_i}{1 - \alpha_i}; \qquad k = \beta_1, \cdots, \beta_N. \tag{38}$$

The term a is given as the value of open circuit input voltage (Z_s and Z_F removed) that exists when I_0 equals unity. This is given by

$$a = Z_E + \frac{Z_E + r_e}{\beta_2, \cdots, \beta_N}, \qquad N > 1 \tag{39}$$

$$a = Z_E + r_e, \qquad N = 1. \tag{40}$$

For $|\beta_2, \cdots, \beta_N| \gg |Z_E + r_e|$, eq. (39) can be approximated by

$$a = Z_E, \qquad N > 1. \tag{41}$$

The last remaining parameter b can be obtained by evaluating the y_{12} parameter of the network Y, yielding

$$b = -(y_{12}) = -\left(\frac{-1}{Z_F}\right) = \frac{1}{Z_F}. \tag{42}$$

Loading effects of the Y networks, i.e., y_{11}, y_{22}, can be ignored if they are sufficiently small.

As a numerical example, the value of k, Z_F, and Z_E are calculated to yield an input and output reflection coefficient of 0.18 (return loss of 15 dB) and a gain to within a factor of 1.26 (1 dB) of $f^2/400$, f in MHz, in the band from 80 to 140 MHz. The remaining parameters of the transistors are $r_b' = 1$, and $r_e = 0.173$.

Following the four-step design procedure yields

(*i*) Let $Z_s = Z_L = 20$, $Z_s Z_L^* = 400$.

(*ii*) $g = \dfrac{f^2}{400} = \dfrac{1}{400|b|^2}$

$|b| = \dfrac{1}{f} = \left|\dfrac{1}{Z_f}\right|$

$|Z_f| = f$, f in MHz.

If Z_f is chosen as an inductor, then

$|Z_f| = 2\pi f L = f$; thus,

$L = 1/2\pi\ \mu$H, and

$b = \dfrac{1}{2\pi f L j} = \dfrac{1}{fj}$.

(*iii*) $a = Z_s Z_L^* b = 400b = 400\,\frac{1}{Z_f} = \frac{400}{2\pi L f j}$

$a = \frac{400}{fj}$; thus, a can be realized when

$Z_E (\simeq a)$ is a capacitor of value

$C = \frac{1}{2\pi(400)}\,\mu\text{F}.$

(*iv*) Using eq. (12) and $a \approx Z_E$, the following is obtained:

$$T = \frac{-2ab\ \text{Re}\ (X_L)}{z_a b Z_L^* + a}\,k = \left(-2\,\frac{400}{fj}\,\frac{1}{fj}\,20\right)k \Big/ \left[\left(1.173 + \frac{400}{fj}\right)\frac{1}{fj}\,20 + \frac{400}{fj}\right]$$

$$T = \frac{-16000k}{8000 + fj423.4} = \begin{cases} -0.46e^{-j77°}k \text{ at 80 MHz} \\ -0.27e^{-j82°}k \text{ at 140 MHz.} \end{cases}$$

For $k = 20$, $|1 - T| = |F|$ at 140 MHz (the worst case point) is given by $|1 + 5.4e^{-j82°}| = |1.75 - j5.35| = 5.63$. This reduces the reflection coefficient by $1/5.03 = 0.18$. Thus, the input and output reflection coefficient specification is initially satisfied.

The gain deviation at 140 MHz is calculated from eq. (15) and is

$$\frac{|T|^2}{|1 - T|^2} = \frac{|5.4|^2}{|5.63|^2} = 0.919.$$

This implies a gain deviation from nominal of 0.37 dB, and initially satisfies the design requirements.

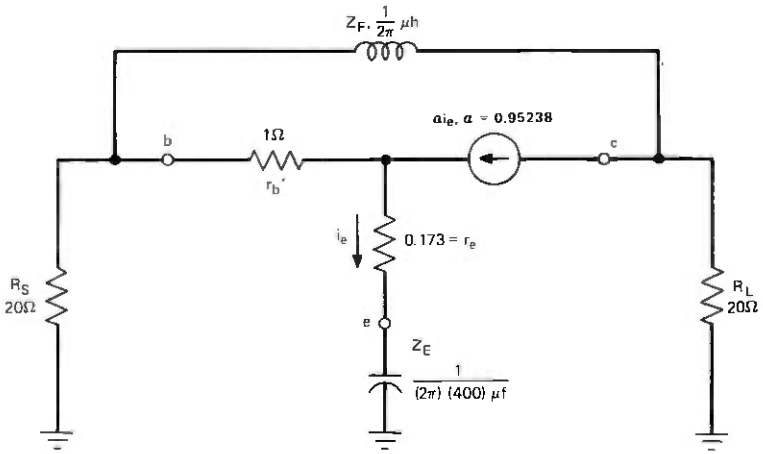


Fig. 5—Numerical example N odd.

The resulting circuit for N odd is given in Fig. 5. The actual transducer gain, input impedance, and output impedance values for this circuit were obtained by a computer-aided design program.[15] A summary of the results is given in Fig. 6. It can be seen from this figure that this procedure yields a practical first iteration in the design procedure.

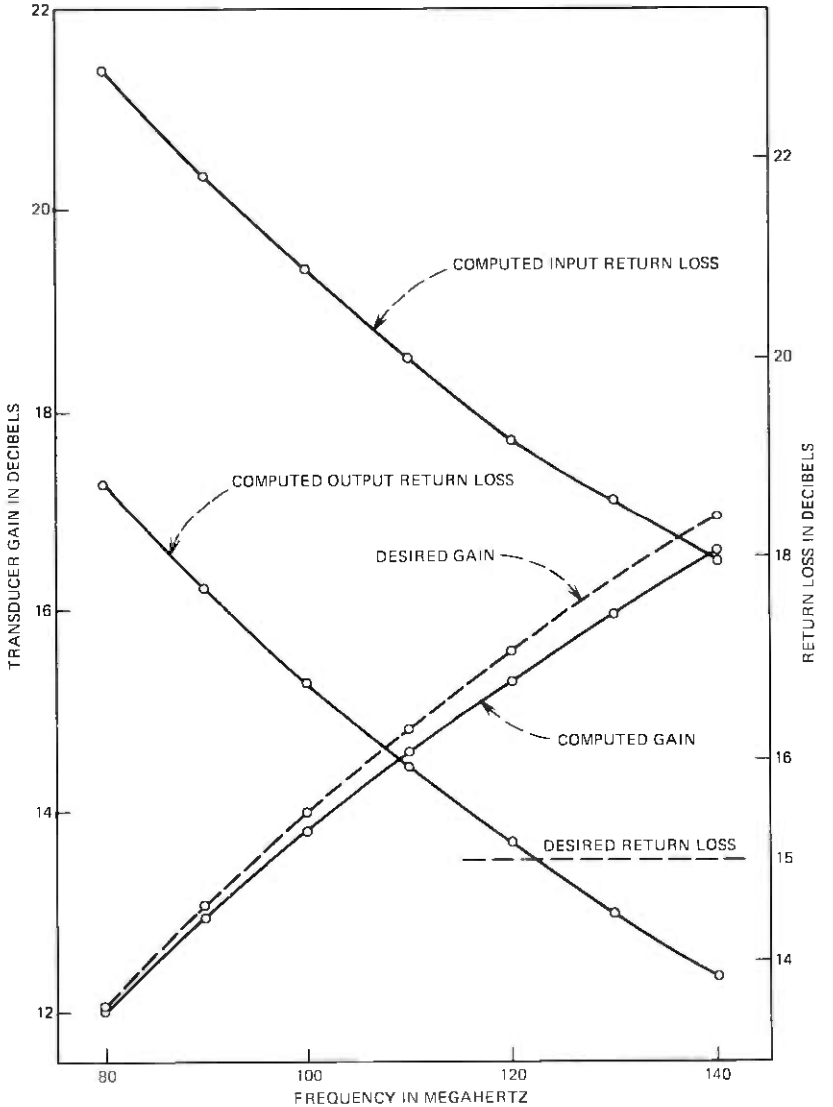


Fig. 6—N odd results.

4.2 N even

For an even number of transistors, the circuit in Fig. 7 is used. The transistor model of Fig. 4b is again used. The evaluation of the parameters of Fig. 3 follows on the same basis as in the N odd case. Using Fig. 7, the calculations are summarized below.

$$z_x \approx r_b + r_e + Z_E, \tag{43}$$

$$I_s = \beta_1, \cdots, \beta_N i_x; \qquad k = -\beta_1, \cdots, \beta_N, \tag{44}$$

$$a = \frac{Z_E}{200 + Z_E} + \frac{r_e + (200Z_E)/(Z_E + 200)}{Z_L \beta_2, \cdots, \beta_N}$$

$$\approx \frac{Z_E}{200}; \qquad \beta_i \gg 1, \qquad i = 2, \cdots, N, \qquad |Z_E| \ll 200 \tag{45}$$

$$b = \frac{2}{Z_F + 2} \approx \frac{2}{Z_F}, \qquad |Z_F| \gg 2. \tag{46}$$

A numerical example is given to show the initial design steps for obtaining a maximum input and output reflection coefficient of 0.18 and a gain to within 1.26 of f^2 (f in MHz), from 80 to 140 MHz. The transistor parameters are again $r_b' = 1$ ohm and $r_e = 0.173$.

The four design steps yield

(*i*) Let $Z_s = Z_L = 20$, $\frac{Z_s}{Z_L^*} = 1$.

(*ii*) $g = f^2 = \frac{1}{|b|^2}$

$|b| = \frac{1}{f} = \frac{2}{|Z_f|}$

$|Z_F| = 2f$.

Let $Z_F = j2f = j2\pi fL$, L in μH

$L = \frac{1}{\pi}\,\mu$H.

(*iii*) $a = b = \frac{2}{j2\pi fL} = \frac{Z_E}{200} = \frac{1}{jf}$.

Thus, $Z_E = \frac{200}{jf}$.

This implies that Z_E is a capacitor of value

$C = \frac{1}{2\pi(200)}\,\mu$F.

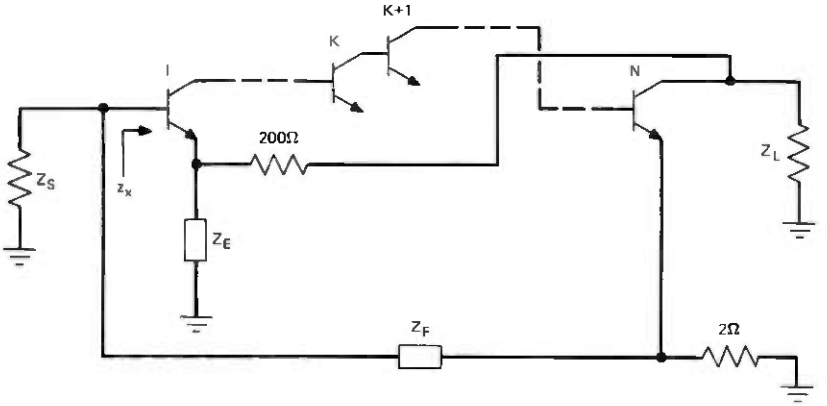


Fig. 7—Design example N even.

(*iv*) Using eq. (32),

$$T = \frac{2ab \operatorname{Re}(Z_L)}{bz_x + aZ_L^*} k = \frac{2(1/jf)(1/jf)20}{(1/jf)(r_b' + r_e + 200/jf) + 20/jf}$$

$$T = \frac{40}{200 + 21.173jf} k = \begin{cases} 0.023e^{-83°j}k \text{ at 80 MHz} \\ 0.015e^{-86°j}k \text{ at 140 MHz.} \end{cases}$$

For $k = -400$, $|1 - T| = |F|$ at 140 MHz is given by $|1 + 6e^{-86°j}| = 6.1$.

The reflection coefficient is reduced by a factor of 1/6.1 (15.7 dB). Gain deviation can be calculated and is equal to 0.95 (0.22 dB); thus, the initial specifications are satisfied.

Figure 8 gives the resulting circuit. The results of the computer analysis of this circuit are given in Fig. 9. Again the data show that the approach yields good results.

It can be seen in Fig. 9 that the difference in gain is greater than the computed 0.22 dB. This is due to the fact that a was taken as $Z_E/200$, rather than the term given in eq. (45). A more accurate evaluation (denoted by the hatted variables) of a is given as

$$\hat{a} = \frac{Z_E}{200 + Z_E} + \frac{r_e + (200Z_E)/(Z_E + 200)}{Z_L\beta_2} \approx \frac{Z_E}{200} + \frac{Z_E}{Z_L\beta_2}$$

$$\hat{a} \approx \frac{Z_E}{200} + \frac{Z_E}{(20)(80)} = \frac{Z_E}{200}(1 + 0.125) = 1.125a,$$

where a was the numerical value previously obtained. Using eq. (33) yields

$$|\hat{S}_{21\infty}|^2 = \frac{1}{|\hat{a}\hat{b}|} = \frac{1}{|1.125ab|} = \frac{1}{1.125}\frac{1}{|ab|} = \frac{1}{1.125}|S_{21\infty}|^2.$$

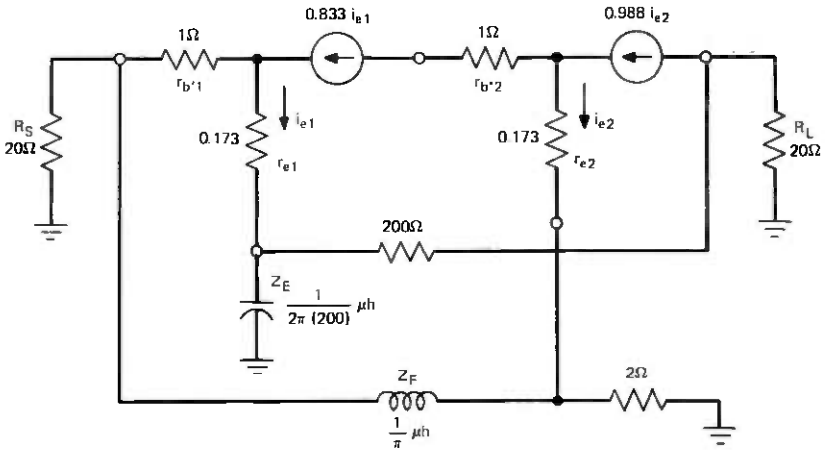


Fig. 8—Numerical example N even.

Again the unhatted quantities were the ones used in the five design steps. The factor of 1.125 accounts for an additional 0.51 dB of the gain difference.

In this example, the gain difference can easily be reduced by increasing the β of the second-stage transistor. This was not done since it was desired to keep the $\beta_1\beta_2$ product at 400. Since β_2 is 80, this forces β_1 to be 5; any high value of β_2 results in unrealistic values of β_1. Nonetheless it is evident that a high $\beta_2, \cdots, \beta_n$ product is needed for an even number of cascade stages.

V. CONCLUSIONS

In this paper, the basic characteristics of two forms of major multiloop feedback have been investigated. The design characteristics treated have been input and output impedance and frequency-dependent power gain. It has been shown that, with sufficient open loop gain, the equations that describe the gain and impedance quantities are very simple in nature. An initial circuit-design iteration can easily be performed since many complicating variables are eliminated.

This initial circuit-design concept would be extremely useful in a computer circuit analysis-optimization program. Well known is the major practical limitation of optimization programs: the obtaining of a convergent starting point. For dual-loop amplifiers, this paper offers the designer a method of easily finding a good starting point.

Although not reported here, several frequency-shaped amplifiers were actually built using multiloop feedback. The excellent perform-

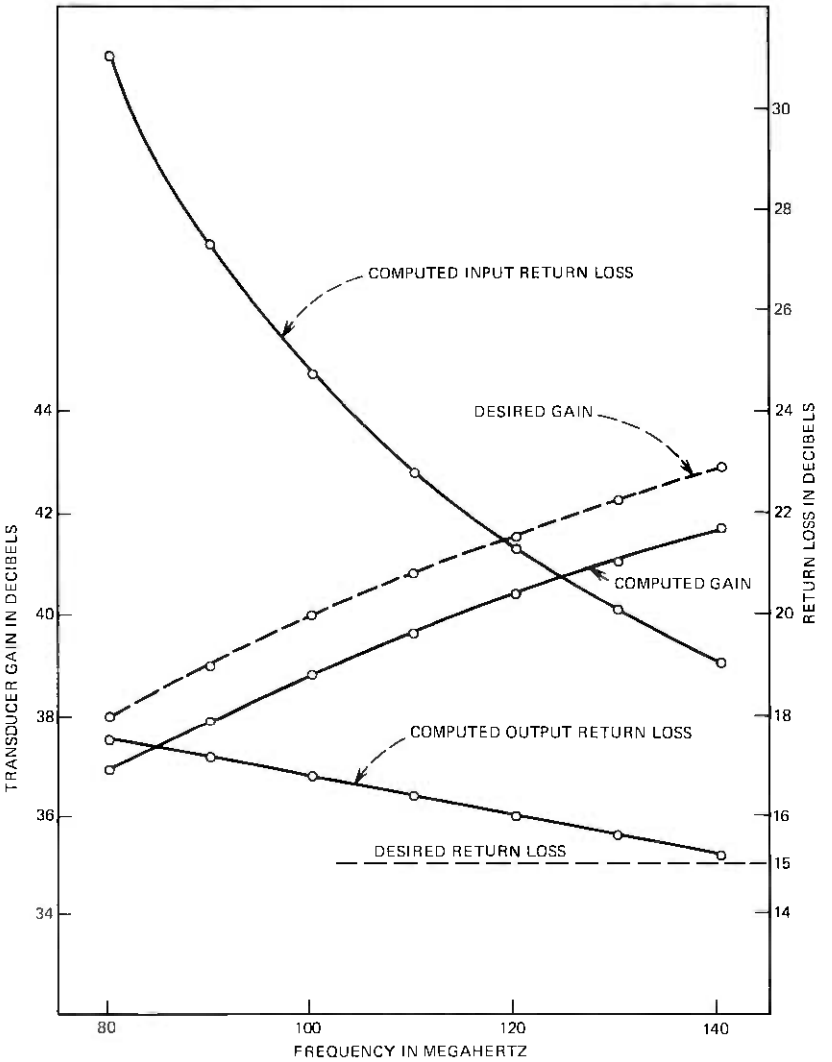


Fig. 9—N even results.

ance of these amplifiers, with respect to input and output matching and gain shaping, has precipitated the work reported in this paper. We anticipate that future papers will discuss more complicated active devices, feedback network loading effects, and feedback network synthesis.

APPENDIX A

Shunt Transadmittance: Series Transimpedance Feedback Calculations

A.1 Calculation of Z_{in}

$$i_x = \frac{V_{in} = aki_x}{z_x};$$

$$i_x(z_x + ak) = V_{in};$$

$$V_0 = -ki_xZ_L.$$

Thus,

$$I_{in} = i_x - bV_0 = i_x + kbi_xZ_L;$$

$$I_{in} = (1 + kbZ_L)i_x = \frac{(1 + kbZ_L)}{z_x + ak} V_{in};$$

$$Z_{in} = \frac{V_{in}}{I_{in}} = \frac{z_x + ka}{1 + kbZ_L}.$$

A.2 Calculation of Z_{out}

$$i_x = bV_0 \frac{Z_s}{z_x + Z_s} - \frac{aI_0}{z_x + Z_s};$$

$$ki_x = I_0 = k\left(bV_0 \frac{Z_s}{Z_s + z_x}\right) - k\left(\frac{aI_0}{z_x + Z_s}\right);$$

$$I_0(z_x + Z_s + ka) = kbZ_sZ_0;$$

$$Z_0 = \frac{V_0}{I_0} = \frac{(z_x + Z_s + ka)}{kbZ_s}.$$

A.3 Input reflection coefficient calculation

$$\rho_{in} = \frac{Z_{in} - Z_s^*}{Z_{in} + Z_s}, \qquad Z_{in} = \frac{z_x + ka}{1 + kbZ_L}, \qquad Z_s = \frac{a}{bZ_L^*},$$

$$\rho_{in} = \left(\frac{z_x + ka}{1 + bkZ_L} - \frac{a}{bZ_L}\right) \Big/ \left(\frac{z_x + ka}{1 + kbZ_L} + \frac{a}{bZ_L^*}\right)$$

$$= \frac{z_xbZ_L + abkZ_L - a - abkZ_L}{z_xbZ_L^* + abkZ_L^* + a + abkZ_L},$$

$$\rho_{in} = \frac{z_xbZ_L - a}{z_xbZ_L^* + a + abk[2\,\mathrm{Re}\,(Z_L)]},$$

$$\rho_{in_0} = \rho_{in|\,k=0} = \frac{z_xbZ_L - a}{z_xbZ_L^* + a}.$$

Therefore,

$$\rho_{in} = \frac{z_xbZ_L - a}{z_xbZ_L^* + a} \times \left(1 + \frac{2ab\,\mathrm{Re}\,(Z_L)}{z_xbZ_L^* + a} k\right)^{-1}$$

$$= \rho_{in_0}\left(1 + \frac{2ab\,\mathrm{Re}\,(Z_L)}{z_xbZ_L^* + a} k\right)^{-1}.$$

A.4 Return ratio calculation

Assuming the output current source I_s is disconnected and replaced by a 1-ampere current source, T is given by the current that flows through the disconnected current source.

$$i_x = bV_0 \frac{Z_s}{z_x + Z_s} - \frac{aI_0}{z_s + Z_s}; \qquad I_s = 1, \qquad V_0 = -Z_L,$$

$$i_x = -bZ_L \frac{Z_s}{z_x + Z_s} - \frac{a}{z_x + Z_s} = \frac{-bZ_L Z_s - a}{z_x + Z_s}; \qquad Z_s = \frac{a}{bZ_L^*};$$

$$T = ki_x = k\,\frac{-bZ_L(a)/(bZ_L^*) - a}{z_x + (a)/(bZ_L^*)} = -k\,\frac{abZ_L + abZ_L^*}{bz_xZ_L^* + a}$$

$$= \frac{-2ab\,\mathrm{Re}\,(Z_L)}{bz_xZ_L^* + a}\,k.$$

A.5 Output reflection coefficient

$$\rho_{\text{out}} = \frac{Z_{\text{out}} - Z_L^*}{Z_{\text{out}} + Z_L}; \qquad Z_{\text{out}} = \frac{z_x + Z_s + ka}{kbZ_s}; \qquad Z_L = \frac{a}{bZ_s^*};$$

$$\rho_{\text{out}} = \left(\frac{z_x + Z_s + ka}{kbZ_s} - \frac{a}{bZ_s}\right)\Big/\left(\frac{z_x + Z_s + ka}{kbZ_s} + \frac{a}{bZ_s^*}\right),$$

$$\rho_{\text{out}} = \left(\frac{z_x + Z_s + ak - ak}{kbZ_s}\right)\Big/\left(\frac{z_xZ_s^* + Z_sZ_s^* + akZ_s^* + akZ_s}{kbZ_sZ_s^*}\right),$$

$$\rho_{\text{out}} = \frac{z_xZ_s^* + Z_sZ_s^*}{z_xZ_s^* + Z_sZ_s^* + ak(Z_s + Z_s^*)},$$

$$\rho_{\text{out}_0} = \rho_{\text{out}}\big|_{k=0} = \frac{z_xZ_s^* + Z_sZ_s^*}{z_xZ_s^* + Z_sZ_s^*} = 1.$$

Therefore,

$$\rho_{\text{out}} = \rho_{\text{out}_0}\left(1 + \frac{ak(Z_s + Z_s^*)}{z_xZ_s^* + Z_sZ_s^*}\right)^{-1}; \qquad Z_s = \frac{a}{bZ_L^*};$$

$$\rho_{\text{out}} = \rho_{\text{out}_0}\left(1 + \frac{ak(a/bZ_L^* + a/bZ_L)}{z_x(a/bZ_L) + (a/bZ_L^*)(a/bZ_L)}\right)^{-1}$$

$$= \rho_{\text{out}_0}\left(1 + \frac{kabZ_L + kabZ_L^*}{z_xbZ_L^* + a}\right)^{-1},$$

$$\rho_{\text{out}} = \rho_{\text{out}_0}\left(1 + \frac{2ab\,\mathrm{Re}\,(Z_L)}{bz_xZ_L^* + a}\,k\right)^{-1}.$$

A.6 Transducer gain calculation

Assume a voltage source of value V_s is inserted in series with the source impedance Z_s in Fig. 2. Let $Z_s = a/bZ_L^*$. P_{AS} will denote the

available power from the source and P_{out} the real power delivered to the load Z_L.

$$P_{AS} = \frac{|V_s|^2}{4\,\mathrm{Re}\,(Z_s)} = \frac{|V_s|^2|Z_L|^2|b|}{4\,\mathrm{Re}\,(Z_L)|a|};$$

$$V_{in} = \frac{V_s Z_{in}}{Z_s + Z_{in}} = \left(V_s \frac{z_x + ka}{1 + kbZ_L}\right) \Big/ \left(\frac{a}{bZ_L^*} + \frac{Z_x + ka}{1 + kbZ_L}\right)$$

$$= \frac{(z_x + ka)bZ_L^*}{a + kabZ_L + (z_x + ka)bZ_L^*} V_s,$$

$$V_{in} = \frac{(z_x + ka)bZ_L^* V_s}{a + z_x bZ_L^* + 2kab\,\mathrm{Re}\,(Z_L)} = \frac{(z_x + ka)bZ_L^*}{a + z_x bZ_L^*} \frac{1}{1 - T} V_s$$

$$\frac{V_{in} - aI_0}{z_x} = i_x = \frac{ki_x}{k} = \frac{I_0}{k}.$$

Thus,

$$I_0 = \frac{k}{z_x + ka} V_{in};$$

$$P_{out} = I_0 I_0^* \,\mathrm{Re}\,(Z_L) = \frac{|k|^2\,\mathrm{Re}\,(Z_L)}{|z_x + ka|^2} |V_{in}|^2,$$

$$P_{out} = \frac{|k|^2\,\mathrm{Re}\,(Z_L)}{|z_x + ka|^2} \frac{|z_x + ka|^2|b|^2|Z_L|^2}{|a + z_x bZ_L^*|^2} \frac{|V_s|^2}{|1 - T|^2}.$$

Therefore,

$$|S_{21}|^2 = \frac{P_{out}}{P_{AS}} = \frac{|k|^2\,\mathrm{Re}\,(Z_L)|b|^2|Z_L|^2|V_s|^2 4\,\mathrm{Re}\,(Z_L)|a|}{|a + z_x bZ_L^*|^2|1 - T|^2|V_s|^2|Z_L|^2|b|}$$

$$= \frac{|k|^2\,\mathrm{Re}^2\,(R_L)|ab|4}{|a + z_x bZ_L^*|^2|1 - T|^2} = \frac{1}{|ab|} \frac{|T|^2}{|1 - T|^2}.$$

APPENDIX B

Current Transfer Shunt: Voltage Transfer Series Feedback Calculations

B.1 Calculation of Z_{in}

$$i_x = \frac{V_{in} - aV_0}{z_x} = \frac{V_{in} + aki_x Z_L}{z_x},$$

$$i_x = \frac{V_{in}}{z_x - akZ_L};$$

$$I_{in} = -bI_0 + i_x = (-bk + 1)i_x = \frac{V_{in}(1 - bk)}{z_x - akZ_L};$$

$$Z_{in} = \frac{V_{in}}{I_{in}} = \frac{z_x - akZ_L}{1 - bk}.$$

B.2 Calculation of Z_{out}

$$i_x = \frac{bI_0Z_s}{Z_s + z_x} - \frac{aV_0}{Z_s + z_x} = \frac{Z_s bI_0 - aV_0}{Z_s + z_x};$$

$$ki_x = I_0 = \frac{kZ_s bI_0 - kaV_0}{Z_s + z_x};$$

$$I_0(Z_s + z_x - kZ_s b) = -kaV_0;$$

$$Z_{out} = \frac{V_0}{I_0} = \frac{Z_s + z_x - kZ_s b}{-ka}.$$

B.3 Return ratio calculation

$$i_x = bI_0\frac{Z_s}{z_x + Z_s} - aV_0\frac{1}{z_x + Z_s}, \qquad I_0 = 1, \qquad V_0 = -Z_L;$$

$$i_x = \frac{bZ_s}{z_x + Z_s} + \frac{aZ_L}{z_x + Z_s}; \qquad Z_s = \frac{a}{b}Z_L^*;$$

$$T = ki = \frac{kb(a/b)Z_L^* + aZ_Lk}{z_x + (a/b)Z_L^*} = \frac{2kab\,\mathrm{Re}\,(Z_L)}{bz_x + aZ_L^*}.$$

B.4 Input reflection coefficient calculation

$$\rho_{in} = \frac{Z_{in} - Z_s^*}{Z_{in} + Z_s}; \qquad Z_{in} = \frac{z_x - akZ_L}{1 - bk}, \qquad Z_s = (a/b)Z_L^*, \qquad (a/b)\ \text{real};$$

$$\rho_{in} = \left(\frac{z_x - akZ_L}{1 - bk} - \frac{aZ_L}{b}\right)\Big/\left(\frac{z_x - akZ_L}{1 - bk} + \frac{a}{b}Z_L^*\right)$$

$$= \frac{bz_x - abkZ_L - aZ_L + abkZ_L}{bz_x - abkZ_L + aZ_L^* - abkZ_L^*},$$

$$\rho_{in} = \frac{bz_x - aZ_L}{bz_x + aZ_L^* - abk\,\mathrm{Re}\,(Z_L)},$$

$$\rho_{in_0} = \rho_{in}\big|_{k=0} = \frac{bz_x - aZ_L}{bz_x + aZ_L^*};$$

therefore,

$$\rho_{in} = \rho_{in_0}\left(1 - \frac{2abk\,\mathrm{Re}\,(Z_L)}{bz_x + aZ_L^*}\right)^{-1}.$$

B.5 Output reflection coefficient calculation

$$\rho_{out} = \frac{Z_{out} - Z_L^*}{Z_{out} + Z_L}; \qquad Z_{out} = \frac{Z_s + z_x - kZ_s b}{-ka},$$

$$Z_s = (a/b)Z_L^*, \qquad (a/b)\ \text{real};$$

$$\rho_{\text{out}} = \left(\frac{Z_s + z_x - kZ_s b}{-ka} - Z_L^*\right) \Big/ \left(\frac{Z_s + z_x - kZ_s b}{-ka} + Z_L\right)$$

$$= \frac{Z_s + z_x - kZ_s b + kaZ_L^*}{Z_s + z_x - kZ_s b - kaZ_L} = \frac{Z_s + z_x}{Z_s + z_x - ka2 \operatorname{Re}(Z_L)},$$

$$\rho_{\text{out}_0} = \rho_{\text{out}}|_{k=0} = 1,$$

$$\rho_{\text{out}} = \rho_{\text{out}_0}\left(1 - \frac{2ka \operatorname{Re}(Z_L)}{(a/b)Z_L^* + z_x}\right)^{-1} = \rho_{\text{out}_0}\left(1 - \frac{2kab \operatorname{Re}(Z_L)}{aZ_L^* + bz_x}\right)^{-1}.$$

B.6 Transducer gain calculation

Assume a voltage source of value V_s is inserted in series with the source impedance Z_s in Fig. 3. Let $Z_s = (a/b)Z_L^*$, a/b real. P_{AS} will denote the available power from the source and P_{out} the real power delivered to the load Z_L.

$$P_{as} = \frac{|V_s|^2}{4 \operatorname{Re}(Z_s)} = \frac{|V_s|^2 |b|}{4 \operatorname{Re}(Z_L)|a|};$$

$$V_{\text{in}} = \frac{V_s Z_{\text{in}}}{Z_s + Z_{\text{in}}} = \left(V_s \frac{z_x - akZ_L}{1 - bk}\right) \Big/ \left(\frac{a}{b} Z_L^* + \frac{z_x - akZ_L}{1 - bk}\right)$$

$$= \frac{V_s(bz_x - abkZ_L)}{aZ_L^* - abkZ_L^* + bz_x - abkZ_L},$$

$$V_{\text{in}} = V_s \frac{bz_x - abkZ_L}{aZ_L^* + bz_x - 2abk \operatorname{Re}(Z_L)} = V_s \frac{bz_x - abkZ_L}{aZ_L^* + bz_x} \times \frac{1}{1 - T},$$

$$\frac{V_{\text{in}} - aV_0}{z_x} = i_x = \frac{ki_x}{k} = \frac{I_0}{k} = \frac{V_{\text{in}} + aI_0 Z_L}{z_x};$$

$$I_0 = V_{\text{in}} \Big/ \left(\frac{z_x}{k} - aZ_L\right) = \frac{k}{z_x - akZ_L} V_{\text{in}};$$

$$P_{\text{out}} = I_0 I_0^* \operatorname{Re}(Z_L) = \frac{|k|^2 |V_{\text{in}}|^2}{|z_x - akZ_L|^2} \operatorname{Re}(Z_L),$$

$$P_{\text{out}} = \frac{|k|^2 \operatorname{Re}(Z_L)}{|z_x - akZ_L|^2} \frac{|bz_x - abkZ_L|^2}{|aZ_L^* + bz_x|^2} \frac{1}{|1 - T|^2} |V_s|^2,$$

$$P_{\text{out}} = \frac{|k|^2 |b|^2 \operatorname{Re}(Z_L)|V_s|^2}{|aZ_L^* + bz_x|^2 |1 - T|^2};$$

therefore,

$$|S_{21}|^2 = \frac{P_{\text{out}}}{P_{AS}} = \frac{|k|^2 |b|^2 \operatorname{Re}(Z_L)|a| 4 \operatorname{Re}(Z_L)|V_s|^2}{|aZ_L^* + bz_x|^2 |1 - T|^2 |V_s|^2 |b|}$$

$$= \frac{1}{|ab|} \frac{|T|^2}{|1 - T|^2}.$$

REFERENCES

1. P. E. Gray and C. L. Searle, *Electronic Principles*, New York: John Wiley, 1969.
2. H. W. Bode, *Network Analysis and Feedback Amplifier Design*, New York: Van Nostrand, 1945.
3. I. M. Horowitz, *Synthesis of Feedback Systems*, New York: Academic Press, 1963.
4. F. H. Blecher, "Design Principles for Single Loop Transistor Feedback Amplifiers," IRE Transactions on Circuit Theory, *CT-4*, No. 3 (September 1957), pp. 145–156.
5. F. H. Blecher, "Transistor Multiple Loop Feedback Amplifiers," Proc. of the National Elec. Conf. (Sponsored by AIEE and IRE), *13*, 1957, pp. 19–34.
6. H. O. Friedheim, "On Hybrid Transformers," A.T.E. J., *14*, No. 3 (July 1958), pp. 218–228.
7. L. H. Morris, G. H. Lovell, and F. R. Dickinson, "L3 Coaxial System-Amplifiers," Trans. of AIEE, PT 1, Commun. and Elec., *72* (November 1953), pp. 505–517.
8. J. L. Garrison, L. P. Labbe, and C. C. Rock, "L-4 System: Basic and Regulating Repeaters," B.S.T.J., *48*, No. 4 (April 1969), pp. 841–888.
9. E. H. Angell and M. M. Luniewicz, "Low Noise, Ultra Linear Line Repeaters for the L5 Coaxial System," IEEE International Conference on Communications, June 1972, pp. 3.10–3.16.
10. W. W. Mumford and E. H. Scheibe, *Noise Performance Factor in Communication System*, Dedham, Mass.: Horizon House—Microwave, 1968.
11. S. Narayanan, "Transistor Distortion Analysis Using Volterra Series Representation," B.S.T.J., *46*, No. 5 (May 1967).
12. M. E. Van Valkenberg, *Introduction to Modern Network Synthesis*, New York: John Wiley, 1960.
13. K. Kurokawa, "Power Waves and the Scattering Matrix," IEEE Trans. on Microwave Theory and Techniques, *MTT-13*, No. 2 (March 1965), pp. 194–202.
14. E. S. Kuh and R. A. Rohrer, *Theory of Linear Active Networks*, San Francisco: Holden-Day, 1967.
15. R. M.-M. Chen et al., "Role of Computing and Precision Measurements," B.S.T.J., *53*, No. 10 (December 1974), pp. 2249–2267.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 54, No. 7, September 1975
Printed in U.S.A.

Variance of Load Measurements in Markovian Service Systems

By A. DESCLOUX

(Manuscript received November 25, 1974)

The load carried by a queuing system under equilibrium conditions is the average amount of server usage per unit of time. In telephony, this parameter is often evaluated by recording the number of busy servers at regular time intervals; these readings are then cumulated and their sum, after division by the number of observations, is an unbiased estimate of the carried load. The purpose of this paper is to derive exact formulas for the computation of the variance of this measurement in systems with arbitrary input and departure rates. The results obtained here thus apply to a wide class of teletraffic models which includes, in particular, the delay-and-loss systems with finite- or infinite-source inputs, exponential service times, and arbitrary defection rates from the queue. Problems related to computations are also considered, special attention being paid to the reduction of both computer time and storage when the number of states is large.

I. INTRODUCTION

Analysis of the stochastic behavior of traffic measurements is of considerable practical relevance, as it provides means for appraising field data as well as guidelines for selecting performance standards. Load measurements play a central role in this effort, and determination of their accuracy is therefore of particular interest. The present investigation yields an answer to this problem for a broad class of teletraffic models.

Whenever statistical equilibrium prevails (and it is assumed to throughout this paper), the load carried by a service system is the average amount of server usage per unit of time or, equivalently, the average number of busy servers at an arbitrary instant. In telephony, an estimate of this parameter is often obtained by "switch-counting." [1] This statistic, which is determined by recording the number of busy servers at regular intervals and then by taking the arithmetic mean of these discrete observations, is an unbiased estimate of the carried load.

The variance of this measurement, called hereafter the switch-count load to distinguish it from the estimate obtained by continuous observation, was first determined approximately by Palm[2] and Hayward[1] in the case of an infinite server group with Poisson input and exponential holding times. This result was later extended by Beneš,[3] who obtained the exact variance of the switch-count load for groups of finite sizes without waiting positions (loss systems). A further generalization to loss systems with recurrent input and exponential service is due to Neal and Kuczura.[4] Their formal analysis stops, however, with a derivation of the Laplace transform of the covariance function of the underlying carried-load process. From this point on, they proceed numerically, since explicit inversion of the transform appears to be difficult in general.

In this paper we are concerned with derivations of exact formulas for the variance of the switch-count load in finite systems with arbitrary state-dependent input and departure rates. The results presented here, therefore, fill a rather large gap, since they apply to a broad class of teletraffic models that includes, in particular, the (finite) delay systems with exponential holding-time distributions, arbitrary defection rates from the queue (if one is allowed to form) and either Poisson or quasi-random input (in the latter case, the traffic is generated by a finite number of sources that place demands for service at the same constant probability rate when free but that do not submit requests while being either served or waiting).

Let $N(t)$, the state of the system at time t, be defined as the number of busy devices at that instant (by device, we mean here either a server or a waiting position). Let c and d be, respectively, the number of servers and the number of devices.

Unless stated otherwise, we make the following assumptions:

(*i*) When $N(t) = n$ and $0 \leqq n < d$, the probability that a request originates during $(t, t + h)$, $h > 0$, is of the form $\lambda_n h + o(h)$, with $\lambda_n > 0$.

(*ii*) The requests which are submitted when all the devices are occupied are dismissed and, accordingly, λ_d is set equal to zero.

(*iii*) When $N(t) = n$ and $0 < n \leqq c$, the probability that a service time terminates during $(t, t + h)$ is of the form $\mu_n h + o(h)$, where $\mu_n > 0$.

(*iv*) When $N(t) = n > c$, the probability that either a service time terminates or a waiting request defects from the queue is of the form $\mu_n h + o(h)$ where $\mu_n > 0$ and $n \leqq d$.

(*v*) When a server becomes free, it is immediately reseized by one of the waiting requests if any are present in the system at that time.

Let $N_c(t)$ be the number of busy servers at time t and let $c \wedge n$ be the smaller of the two integers c and n. Then

$$N_c(t) = c \wedge N(t) = \begin{cases} N(t) & \text{if } N(t) \leqq c, \\ c & \text{if } N(t) > c, \end{cases}$$

and the switch-count load, $L_n(T)$, based on n observations (scans) made over $[0, T]$ at times $\tau, 2\tau, \cdots, n\tau$, is, by definition, equal to

$$n^{-1} \sum_{j=1}^{n} N_c(j\tau),$$

where $\tau \equiv T/n$.

Let $\text{Cov}\,[N_c(t_1), N_c(t_2)]$ be the covariance between $N_c(t_1)$ and $N_c(t_2)$. Under equilibrium conditions, this covariance depends only on $|t_1 - t_2|$ so that

$$\text{Cov}\,[N_c(t_1), N_c(t_2)] = \text{Cov}\,[N_c(0), N_c(|t_1 - t_2|)].$$

Hence, the variance of $L_n(\tau)$, cast in a form that will be convenient later, is given by the formula (Ref. 3, p. 137):

$$\text{Var}\, L_n(T) = n^{-2} \sum_{k=-n}^{n} (n - |k|) R_c(k\tau), \tag{1}$$

where

$$\begin{aligned} R_c(k\tau) &\equiv \text{Cov}\,[N_c(0), N_c(k\tau)] \\ &= \text{Cov}\,[N_c(0), N_c(|k|\tau)]. \end{aligned}$$

It is clear from (1) that the variance of the switch-count load is completely determined by the covariance function $R_c(\cdot)$ of the carried-load process $\{N_c(t), -\infty < t < \infty\}$, and therefore much of what follows is concerned with expressing $R_c(\cdot)$ in the most convenient form.

The covariance function can be stated at first in terms of the transition probabilities, and the resulting expression can then be reduced by taking the structural properties of the process into account. But alternate forms can also be obtained by making use of the fact that the conditional expectations, $E\{N_c(t) \mid N(0) = m\}$, $m = 0, 1, \cdots, d$, satisfy simple linear differential equations. The covariance formulas obtained by these diverse procedures exhibit distinct features that may be exploited in the computations. In all cases, however, $R_c(t)$ is expressed as a diagonal, positive-definite quadratic form which reveals that $R_c(\cdot)$ is completely monotonic.[5]

Expressions for the transition probabilities, the covariance function, and the variance of the switch-count load are derived in Sections II, III, and IV, respectively. The variance of load measurements based on

continuous observations can be found in Section IV. Extensions of the results of Sections II to IV to reversible Markov processes are also considered. Questions of a computational nature are dealt with in Section V, while Section VI is concerned with some asymptotic properties (d large) of the spectrum of the underlying process.

The formulas presented here have been programmed and used to explore the effects of parameter changes on the variance of the switch-count load. The result of that investigation will appear in another paper.

II. TRANSITION PROBABILITIES

In this section, we express the transition-probability matrix as a symmetric product of vectors and matrices. As becomes apparent later, this representation makes it possible to write the variance of the switch-count load in a way that greatly simplifies its evaluation.

Let $p_{mn}(t)$ be the probability of a transition from state m to state n in time t:

$$p_{mn}(t) \equiv \Pr\,[N(t) = n \mid N(0) = m], \qquad m, n = 0, 1, \cdots, d.$$

These transition probabilities satisfy the following system of differential equations:

$$\begin{aligned}
\frac{d}{dt}\, p_{m0} &= \mu_1 p_{m1} - \lambda_0 p_{m0}, \\
\frac{d}{dt}\, p_{mn} &= \mu_{n+1} p_{m,n+1} + \lambda_{n-1} p_{m,n-1} - (\lambda_n + \mu_n) p_{mn}, \qquad 1 \leqq n < d, \\
\frac{d}{dt}\, p_{md} &= \lambda_{d-1} p_{m,d-1} - \mu_d p_{md}.
\end{aligned} \tag{2}$$

Let

$$\mathbf{A}_d \equiv \begin{bmatrix}
-\lambda_0 & \lambda_0 & & & & & \\
\mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & & & & \\
 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & & & \\
 & & \cdots & \cdots & \cdots & & \\
 & & & \mu_{d-1} & -(\lambda_{d-1}+\mu_{d-1}) & \lambda_{d-1} \\
 & & & & \mu_d & -\mu_d
\end{bmatrix}$$

and

$$\mathbf{P}_d(t) \equiv [p_{mn}(t)], \qquad m, n = 0, 1, \cdots, d,$$

be the transition-probability matrix. Capital and lower-case bold-face letters are used exclusively to designate matrices and vectors, respectively.

With this notation, the system of differential equations (2) becomes

$$\frac{d}{dt}\mathbf{P}_d(t) = \mathbf{P}_d(t)\cdot\mathbf{A}_d, \qquad t \geqq 0, \tag{3}$$

so that, for $k = 1, 2, \cdots$,

$$\frac{d^k}{dt^k}\mathbf{P}_d(t) = \frac{d^{k-1}}{dt^{k-1}}\mathbf{P}_d(t)\cdot\mathbf{A}_d, \qquad t \geqq 0. \tag{4}$$

It follows from our assumptions that if the system is in state m at time zero $[N(0) = m]$, then $\lim_{t\downarrow 0} p_{mm}(t) = 1$ and $\lim_{t\downarrow 0} p_{mn}(t) = 0$ for $n \neq m$. Hence, with $\mathbf{I}_d$ the identity matrix of order $d + 1$, the initial conditions take the following form:

$$\mathbf{P}_d(0) \equiv \lim_{t\downarrow 0}\mathbf{P}_d(t) = \mathbf{I}_d,$$

and by (3) and (4) we therefore have

$$\lim_{t\downarrow 0}\frac{d^k}{dt^k}\mathbf{P}_d(t) = \mathbf{A}_d^k. \tag{5}$$

The initial conditions state that $\mathbf{P}_d(\cdot)$ is right-continuous at $t = 0$ and imply that $\mathbf{P}_d(\cdot)$ is continuous for all $t > 0$. By (3) and (4), all the derivatives of $\mathbf{P}_d(\cdot)$ exist for $t > 0$, and by (5) they are also right-continuous at $t = 0$. An application of Taylor's theorem then yields (Ref. 6, pp. 240 ff.)

$$\mathbf{P}_d(t) = \exp(\mathbf{A}t) = \sum_{k=0}^{\infty}\frac{1}{k!}\mathbf{A}_d^k t^k, \qquad t \geqq 0. \tag{6}$$

The elements of $\mathbf{A}_d$ situated immediately either above or below the diagonal are all strictly positive and so $\mathbf{A}_d$ can be symmetrized. Indeed, let

$$\mathbf{D}_d \equiv \operatorname{diag}[\delta_0, \delta_1, \cdots, \delta_d]$$

with

$$\delta_0 = \zeta \quad \text{and} \quad \delta_m \equiv \zeta\left(\frac{\mu_1\mu_2 \cdots \mu_m}{\lambda_0\lambda_1 \cdots \lambda_{m-1}}\right)^{\frac{1}{2}} = \xi p_m^{-\frac{1}{2}}, \qquad m = 1, \cdots, d,$$

where (*i*) ζ is a nonvanishing but otherwise arbitrary constant, (*ii*) the p_m are the equilibrium state probabilities, and (*iii*) $\xi = \zeta p_0^{\frac{1}{2}}$. Without loss of generality, we can—and shall—set $\zeta = p_0^{\frac{1}{2}}$ so that $\xi = 1$ and

$$\mathbf{D}_d^{-1} = \operatorname{diag}[p_0^{\frac{1}{2}}, p_1^{\frac{1}{2}}, \cdots, p_d^{\frac{1}{2}}]. \tag{7}$$

It is easy to verify that

$$\mathbf{S}_d = \mathbf{D}_d^{-1}\cdot\mathbf{A}_d\cdot\mathbf{D}_d \tag{8}$$

is symmetric, its nonvanishing elements being

$$s_{mm} = -(\lambda_m + \mu_m), \qquad m = 0, 1, \cdots, d, \qquad (\lambda_d = 0),$$
$$s_{m,m+1} = s_{m+1,m} = (\lambda_m \mu_{m+1})^{\frac{1}{2}}, \qquad m = 0, 1, \cdots, d-1.$$

Hence, by (8) we have

$$\mathbf{A}_d^k = \mathbf{D}_d \cdot \mathbf{S}_d^k \cdot \mathbf{D}_d^{-1}, \qquad k = 0, 1, 2, \cdots, \tag{9}$$

and, by (6),

$$\mathbf{P}_d(t) = \mathbf{D}_d \cdot \exp(\mathbf{S}_d t) \cdot \mathbf{D}_d^{-1} = \sum_{k=0}^{\infty} \frac{1}{k!} (\mathbf{D}_d \cdot \mathbf{S}_d^k \cdot \mathbf{D}_d^{-1}) t^k. \tag{10}$$

The representation of $\mathbf{A}_d$ in terms of the symmetric tridiagonal matrix $\mathbf{S}_d$ entails substantial formal simplification of the final results. And it is also particularly convenient computationally, since the determination of the characteristic values of $\mathbf{A}_d$ (which are needed for an exact solution) is best carried out after symmetrization.

The matrices $\mathbf{A}_d$ and $\mathbf{S}_d$ clearly have the same characteristic values, $r_0, r_1, \cdots, r_d$. But $\mathbf{S}_d$ is symmetric and is therefore unitarily similar to the diagonal matrix

$$\mathbf{C}_d \equiv \operatorname{diag}[r_0, r_1, \cdots, r_d].$$

This means that an orthogonal matrix $\mathbf{B}_d$ exists such that

$$\mathbf{S}_d = \mathbf{B}_d' \cdot \mathbf{C}_d \cdot \mathbf{B}_d, \qquad \mathbf{B}_d \cdot \mathbf{B}_d' = \mathbf{B}_d' \cdot \mathbf{B}_d = \mathbf{I}_d, \tag{11}$$

where $\mathbf{B}_d'$ is the transpose of $\mathbf{B}_d$.

But $\mathbf{S}_d$ is also tridiagonal, and its off-diagonal elements never vanish. Hence, $\mathbf{S}_d$ is nonderogatory and its characteristic values are necessarily distinct (Ref. 7, p. 26). The elements in the nth column of $\mathbf{B}_d'$ are then the components of the (uniquely defined) normalized characteristic vector associated with the nth characteristic value $r_n (n = 0, 1, \cdots, d)$.

We now substitute (11) into (10). This yields

$$\mathbf{P}_d(t) = \sum_{k=0}^{\infty} \frac{1}{k!} (\mathbf{D}_d \cdot \mathbf{B}_d' \cdot \mathbf{C}_d^k \cdot \mathbf{B}_d \cdot \mathbf{D}_d^{-1}) t^k,$$

so that

$$\begin{aligned} \mathbf{P}_d(t) &= \mathbf{D}_d \cdot \mathbf{B}_d' \cdot \exp(\mathbf{C}_d \cdot t) \cdot \mathbf{B}_d \cdot \mathbf{D}_d^{-1} \\ &= \mathbf{D}_d \cdot \mathbf{B}_d' \cdot \operatorname{diag}[e^{r_0 t}, e^{r_1 t}, \cdots, e^{r_d t}] \cdot \mathbf{B}_d \cdot \mathbf{D}_d^{-1}. \end{aligned} \tag{12}$$

We note now that all the row sums of $\mathbf{A}_d$ vanish and one of the characteristic roots, r_0, say, must therefore be equal to zero. Furthermore, known extremal properties of the characteristic values can be used to show that $r_1, r_2, \cdots, r_d$ are negative. It is also readily seen that

$$\mathbf{p}_d^{(\frac{1}{2})} \equiv (p_0^{\frac{1}{2}}, p_1^{\frac{1}{2}}, \cdots, p_d^{\frac{1}{2}})'$$

is the characteristic vector of $\mathbf{S}_d$ that corresponds to the vanishing

characteristic root r_0. Indeed, let $\mathbf{e}_d$ and $\mathbf{o}_d$ be the $(d+1)$ dimensional (column) vectors whose components are all equal to 1 and 0, respectively. Then, since $\mathbf{A}_d \cdot \mathbf{e}_d = \mathbf{o}_d$, we have, by (8) and (7),

$$\mathbf{D}_d \cdot \mathbf{S}_d \cdot \mathbf{D}_d^{-1} \cdot \mathbf{e}_d = \mathbf{D}_d \cdot \mathbf{S}_d \cdot \mathbf{p}_d^{(\frac{1}{2})} = \mathbf{o}_d.$$

But none of the diagonal elements of $\mathbf{D}_d$ vanishes and the relation

$$\mathbf{D}_d \cdot \mathbf{S}_d \cdot \mathbf{p}_d^{(\frac{1}{2})} = \mathbf{o}_d$$

can hold if and only if $\mathbf{S}_d \cdot \mathbf{p}_d^{(\frac{1}{2})} = \mathbf{o}_d$. Thus, $\mathbf{p}_d^{(\frac{1}{2})}$ is the characteristic vector associated with $r_0(=0)$, a fact that may be of relevance in the computations, as a comparison of $\mathbf{p}_d^{(\frac{1}{2})}$ with $\mathbf{D}_d^{-1}$ provides an accuracy check for the method used to determine the characteristic vectors.

In the derivation of formula (12), advantage was taken of the fact that the transition-rate matrix $\mathbf{A}_d$ is symmetrizable. It is worth noting that this relatively simple expression for $\mathbf{P}_d$ is a consequence of this property, and therefore holds for all (and actually only for) reversible Markovian processes with finite state spaces. Indeed, by definition, the class of these processes—which includes those of the birth-and-death type—is fully characterized by the following conditions (Refs. 8 and 9):

$$p_m p_{mn}(t) = p_n p_{nm}(t), \qquad m, n = 0, 1, \cdots, d, \tag{13}$$

or, equivalently, by the single relation:

$$\mathbf{D}_d^{-2} \cdot \mathbf{P}_d = \mathbf{P}_d' \cdot \mathbf{D}_d^{-2}. \tag{14}$$

Hence (12), written in terms of $\mathbf{S}_d$, implies that

$$\begin{aligned} \mathbf{D}_d^{-2} \cdot \mathbf{P}_d &= \mathbf{D}_d^{-1} \cdot \exp(\mathbf{S}_d t) \cdot \mathbf{D}_d^{-1} \\ &= (\mathbf{D}_d^{-1} \cdot \exp(\mathbf{S}_d t) \cdot \mathbf{D}_d^{-1})' = \mathbf{P}_d' \cdot \mathbf{D}_d^{-2}, \end{aligned}$$

and (14) is therefore satisfied.

Conversely, we show next that (14) is a sufficient condition for (12) to hold.

Pre- and post-multiplication of (14) by $\mathbf{D}_d$ yield

$$\mathbf{D}_d^{-1} \cdot \mathbf{P}_d \cdot \mathbf{D}_d = \mathbf{D}_d \cdot \mathbf{P}_d' \cdot \mathbf{D}_d^{-1}. \tag{15}$$

Substituting the expansion of $\mathbf{P}_d$ as given by (6) into (15), and performing the multiplications by $\mathbf{D}_d$ and $\mathbf{D}_d^{-1}$ under the summation sign (which is clearly legitimate), we obtain:

$$\sum_{k=0}^{\infty} \frac{1}{k!} (\mathbf{D}_d^{-1} \cdot \mathbf{A}_d^k \cdot \mathbf{D}_d) t^k = \sum_{k=0}^{\infty} \frac{1}{k!} (\mathbf{D}_d \cdot (\mathbf{A}_d^k)' \cdot \mathbf{D}_k^{-1}) t^k, \qquad t \geqq 0.$$

However, this relation cannot be satisfied unless

$$\mathbf{D}_d^{-1} \cdot \mathbf{A}_d \cdot \mathbf{D}_d = \mathbf{D}_d \cdot \mathbf{A}_d' \cdot \mathbf{D}_d^{-1},$$

so that, by transposition,

$$(\mathbf{D}_d^{-1}\cdot\mathbf{A}_d\cdot\mathbf{D}_d)' = \mathbf{D}_d^{-1}\cdot\mathbf{A}_d\cdot\mathbf{D}_d.$$

This means that $\mathbf{A}_d$ is symmetrizable by pre- and post-multiplication by $\mathbf{D}_d^{-1}$ and $\mathbf{D}_d$, respectively, and (12) then follows as shown earlier.

Under the assumption that the process is reversible and that all the states communicate with each other[10] (i.e., $p_{mn}(t) > 0$, $m, n = 0, \cdots, d$, $t > 0$), the characteristic roots of $\mathbf{A}_d$ are necessarily simple. (Note that $\mathbf{A}_d$, and hence $\mathbf{S}_d = \mathbf{D}_d^{-1}\cdot\mathbf{A}_d\cdot\mathbf{D}_d$, need no longer be tridiagonal.) This can be proved as follows.

The matrix $\mathbf{S}_d$ is symmetric and can therefore be tridiagonalized by a method from Householder (Ref. 7, pp. 152, 153, 290–293, and 343). According to this procedure, the tridiagonalization of $\mathbf{S}_d$ is achieved by successive right and left multiplications by symmetric orthogonal matrices, $\mathbf{U}_1, \mathbf{U}_2, \cdots, \mathbf{U}_{d-1}$, of the form

$$\mathbf{U}_r = \mathbf{I}_d - 2\mathbf{w}_r\cdot\mathbf{w}_r',$$

where $\mathbf{w}_r$ is a suitably chosen $d + 1$ dimensional (column) vector whose first r components are zero. (All the $\mathbf{U}_r$ are of order $(d + 1)$ and $\mathbf{U}_r^2 = \mathbf{I}_d$, $r = 1, 2, \cdots, d - 1$.) A direct application of the results derived in Ref. 7, above, shows that $\mathbf{S}_d$ admits of the following representation:

$$\mathbf{S}_d = \mathbf{U}_1\cdot\mathbf{U}_2\cdot\ \cdots\ \cdot\mathbf{U}_{d-1}\cdot\mathbf{T}_d\cdot\mathbf{U}_{d-1}\cdot\ \cdots\ \cdot\mathbf{U}_2\cdot\mathbf{U}_1,$$

where $\mathbf{T}_d$ is a symmetric tridiagonal matrix of order $d + 1$.

Let θ_{ij} be the elements of $\mathbf{T}_d$.

We are now faced with two possibilities. Either $\theta_{i,i+1} = \theta_{i+1,i} \neq 0$ for $i = 0, 1, \cdots, d - 1$, or there is an index $j(<d)$ such that $\theta_{j,j+1} = \theta_{j+1,j} = 0$ so that

$$\mathbf{T}_d = \begin{bmatrix} \mathbf{T}_j & 0 \\ 0 & \mathbf{T}_{d-j-1} \end{bmatrix}. \tag{16}$$

In the first instance, all the characteristic roots of $\mathbf{S}_d$, and hence of $\mathbf{A}_d$, are distinct (Ref. 7, p. 26). To complete the proof, it is therefore sufficient to show that the second contingency cannot occur when all states communicate with each other. To this end, we proceed indirectly. We assume that (16) is satisfied for some $j < d$ and show that some states then do not communicate with others.

When (16) holds for $j < d$, we have, for any $k \geqq 0$,

$$\begin{aligned}\mathbf{A}_d^k &= \mathbf{D}_d\cdot\mathbf{S}_d^k\cdot\mathbf{D}_d^{-1} \\ &= \mathbf{D}_d\cdot\mathbf{U}_1\cdot\ \cdots\ \cdot\mathbf{U}_{d-1}\cdot\begin{bmatrix} \mathbf{T}_j^k & 0 \\ 0 & \mathbf{T}_{d-j-1}^k \end{bmatrix}\cdot\mathbf{U}_{d-1}\cdot\ \cdots\ \cdot\mathbf{U}_1\cdot\mathbf{D}_d^{-1}.\end{aligned} \tag{17}$$

The first row in $\mathbf{D}_d$ is $(p_0^{-\frac{1}{2}}, 0, \cdots, 0)$ and the elements in the first column and first row of each of the $\mathbf{U}_r$'s are zero except for their first component, which is always equal to 1. Hence,

$$(1, 0, \cdots, 0)\cdot\mathbf{D}_d\cdot\mathbf{U}_1\cdot\cdots\cdot\mathbf{U}_{d-1} = (p_0^{-\frac{1}{2}}, 0, \cdots, 0). \tag{18}$$

Similarly, since the $\mathbf{U}_r$'s are symmetric, we have

$$\mathbf{U}_{d-1}\cdot\cdots\cdot\mathbf{U}_1\cdot\mathbf{D}_d^{-1}\cdot(1, 0, \cdots, 0)' = (p_0^{\frac{1}{2}}, 0, \cdots, 0)'. \tag{19}$$

Hence, by (17) to (19),

$$\begin{aligned} P_{00}(t) &= (1, 0, \cdots, 0)\left(\sum_{k=0}^{\infty}\frac{1}{k!}\mathbf{A}_d^k\cdot t^k\right)(1, 0, \cdots, 0)' \\ &= \sum_{k=0}^{\infty}\frac{1}{k!}\theta_0^{(k)}t^k, \end{aligned} \tag{20}$$

where $\theta_0^{(k)}$ is the element belonging to the first row and first column of $\mathbf{T}_j^k$.

Let a_{mn} and s_{mn}, $m, n = 0, 1, \cdots, d$, be the elements of $\mathbf{A}_d$ and $\mathbf{S}_d$, respectively. Under the present assumptions, $a_{mn}\cdot a_{nm} \geqq 0$ and $s_{mn} = (a_{mn}\cdot a_{nm})^{\frac{1}{2}}$, $m, n = 0, 1, \cdots, d$, $m \neq n$. The elements in the first r rows and columns of $\mathbf{S}_d$ are therefore uniquely determined by the elements in the first r rows and columns of $\mathbf{A}_d$. Similarly, the vector $\mathbf{w}_r$ depends only on the components, s_{mn}, of $\mathbf{S}_d$ for which either $m \leqq r - 1$ and $n = r - 1, \cdots, d$ or, by symmetry, $n \leqq r - 1$ and $m = r - 1, \cdots, d$ (Ref. 7, pp. 290 ff). Consequently, the elements of $\mathbf{T}_j^k$ (which are all obtained after $j - 1$ steps) depend only on the elements of the first $j + 1$ rows and columns of $\mathbf{A}_d$. This implies that the transition probability $P_{00}(t)$, as given by (20), is independent of the rates a_{mn}, $m, n > j$. However, the process being Markovian, this can only be true if $P_{0m}(t) = 0$ for $m > j$ which means (since, by assumption, $j < d$) that state 0 does not communicate with states $j + 1, \cdots, d$, as was to be proved.

III. COVARIANCE FUNCTION

3.1 First version

The covariance function of the carried-load process is, by definition,

$$\begin{aligned} R_c(t) &\equiv \sum_{m,n=0}^{d}(c \wedge n)\cdot(c \wedge m)p_n p_{nm}(t) - M_{c1}^2 \\ &= \sum_{m,n=0}^{d}(c \wedge n)\cdot(c \wedge m)p_n[p_{nm}(t) - p_m], \end{aligned}$$

where

$$M_{c1} \equiv EN_c(0) = \sum_{n=0}^{d}(c \wedge n)p_n.$$

However, if γ is an arbitrary constant, the covariance of the process $\{N_c(t) + \gamma, t \geqq 0\}$ is also R_c. Hence, with the notation

$$\rho_n = (c \wedge n) + \gamma, \qquad n = 0, 1, \cdots, d,$$

we also have

$$R_c(t) = \sum_{m,n=0}^{d} \rho_n \rho_m p_n [p_{nm}(t) - p_m]. \tag{21}$$

Let

$$\hat{\mathbf{P}}_d \equiv \begin{bmatrix} p_0 & p_1 & \cdots & p_d \\ p_0 & p_1 & & p_d \\ \vdots & \vdots & & \vdots \\ p_0 & p_1 & \cdots & p_d \end{bmatrix}$$

and

$$\mathbf{G}_d(t) \equiv [p_{nm}(t) - p_m].$$

The matrix $\hat{\mathbf{P}}_d$ can be obtained by letting $t \to \infty$ in (13). Hence,

$$\hat{\mathbf{P}}_d = \mathbf{D}_d \cdot \mathbf{B}_d' \cdot \text{diag}\,[1, 0, 0, \cdots, 0] \cdot \mathbf{B}_d \cdot \mathbf{D}_d^{-1}$$

and

$$\begin{aligned} \mathbf{G}_d(t) &= \mathbf{P}_d(t) - \hat{\mathbf{P}}_d \\ &= \mathbf{D}_d \cdot \mathbf{B}_d' \cdot \text{diag}\,[0, e^{r_1 t}, \cdots, e^{r_d t}] \cdot \mathbf{B}_d \cdot \mathbf{D}_d^{-1}. \end{aligned}$$

We now introduce two auxiliary row vectors:

$$\mathbf{r}_d' \equiv (\rho_0, \rho_1, \cdots, \rho_d), \qquad \mathbf{s}_d' \equiv (p_0\rho_0, p_1\rho_1, \cdots, p_d\rho_d).$$

Then the coefficient of $e^{r_i t}$ in the linear form

$$\mathbf{s}_d' \cdot \mathbf{G}_d(t) \cdot \mathbf{r}_d = \mathbf{s}_d' \cdot \mathbf{D}_d \cdot \mathbf{B}_d' \cdot \text{diag}\,[0, e^{r_1 t}, \cdots, e^{r_d t}] \cdot \mathbf{B}_d \cdot \mathbf{D}_d^{-1} \mathbf{r}_d$$

is the same as the coefficient of $e^{r_i t}$ in (21), and we may conclude that

$$R_c(t) = \mathbf{s}_d' \cdot \mathbf{D}_d \cdot \mathbf{B}_d' \cdot \text{diag}\,[0, e^{r_1 t}, \cdots, e^{r_d t}] \cdot \mathbf{B}_d \cdot \mathbf{D}_d^{-1} \cdot \mathbf{r}_d.$$

With the notation

$$\mathbf{q}_d' \equiv (\rho_0 p_0^{\frac{1}{2}}, \cdots, \rho_d p_d^{\frac{1}{2}}),$$

we have

$$\mathbf{q}_d = \mathbf{s}_d' \cdot \mathbf{D}_d \qquad \text{and} \qquad \mathbf{q}_d = \mathbf{D}_d^{-1} \cdot \mathbf{r}_d,$$

so that

$$R_c(t) = \mathbf{q}_d' \cdot \mathbf{B}_d \cdot \text{diag}\,[0, e^{r_1 t}, \cdots, e^{r_d t}] \cdot \mathbf{B}_d \cdot \mathbf{q}_d \tag{22}$$

or, alternatively,

$$R_c(t) = \sum_{i=1}^{d} b_i^2 e^{r_i t} \tag{23}$$

with b_i the ith component of the row vector $\mathbf{q}_d' \cdot \mathbf{B}_d'$. This last expression shows that the coefficient of $e^{r_i t}$ in either (22) or (23) is necessarily

nonnegative. Furthermore, since all the r_i's $(i = 1, \cdots, d)$ are negative, we also have

$$(-1)^k \frac{d^k}{dt^k} R_c(t) \geqq 0, \qquad t \geqq 0, \qquad k = 0, 1, \cdots,$$

and R_c is therefore completely monotonic over $[0, \infty)$.[5]

If we now set $\gamma = -c$, the last $d - c$ components of $\mathbf{q}_d$ are equal to zero, and so the determination of R_c by means of (22) necessitates only the computation of the first c components of the characteristic vectors. Formula (22) is therefore often well-suited to the case of delay systems. But unless the number of waiting positions exceeds the number of servers, greater reduction of the computations can be achieved by means of the formulas derived below.

In the preceding derivation, the ρ's are independent of the arrival and departure rates, and the formulas of this subsection therefore hold for arbitrary, reversible, Markov processes with finite state spaces. In contrast, the results of the next subsection are restricted to birth-and-death processes.

3.2 Alternative forms

Multiplying the nth equation in (2) by $(c \wedge n)$ and then summing with respect to $n (0 \leqq n \leqq d)$, we obtain, after rearranging and canceling terms,

$$\sum_{n=0}^{d} (c \wedge n) \cdot \frac{d}{dt} p_{mn} = \sum_{n=0}^{c-1} \lambda_n p_{mn} - \sum_{n=1}^{c} \mu_n p_{mn}. \tag{24}$$

But

$$\sum_{n=0}^{d} (c \wedge n) p_{mn}(t) = E\{N_c(t) \mid N(0) = m\},$$

so that, by (24),

$$\frac{d}{dt} E\{N_c(t) \mid N(0) = m\} = \sum_{n=0}^{c-1} \lambda_n p_{mn} - \sum_{n=1}^{c} \mu_n p_{mn}. \tag{25}$$

Adding and subtracting $\kappa E\{N_c(t) \mid N(0) = m\}$ on the right-hand side of (25), we obtain

$$\frac{d}{dt} E\{N_c(t) \mid N(0) = m\} = \kappa E\{N_c(t) \mid N(0) = m\} + \sum_{n=0}^{d} \rho_n^*(\kappa) p_{mn}, \qquad m = 0, 1, \cdots, d, \tag{26}$$

where

$$\rho_n^*(\kappa) = \begin{cases} (\lambda_n - \mu_n - \kappa n) & \text{if } n = 0, 1, \cdots, c-1, \qquad (\mu_0 = 0), \\ -(\kappa c + \mu_c) & \text{if } n = c, \\ -\kappa c & \text{if } n = c+1, \cdots, d. \end{cases} \tag{27}$$

In the preceding formulas, κ is an arbitrary real number that may be positive, negative, or null. We see later that the covariance formulas can occasionally be simplified by appropriate choices of κ.

Taking the initial conditions,

$$E\{N_c(0)\,|\,N(0) = m\} = (c \wedge m), \qquad 0 \leqq m \leqq d,$$

into account, the solution of (26) is

$$E\{N_c(t)\,|\,N(0) = m\} = (c \wedge m)e^{\kappa t} + \int_0^t e^{\kappa(t-u)} \left[\sum_{n=0}^{d} \rho_n^*(\kappa) p_{mn}(u) \right] \cdot du \tag{28}$$

so that

$$\begin{aligned} R_c(t) &\equiv \sum_{m=0}^{d} (c \wedge m) p_m E\{N_c(t)\,|\,N(0) = m\} - M_{c1}^2 \\ &= M_{c2}e^{\kappa t} - M_{c1}^2 \\ &\qquad + \int_0^t e^{\kappa(t-u)} \sum_{m=0}^{d} (c \wedge m) p_m \left[\sum_{n=0}^{d} \rho_n^*(\kappa) p_{mn}(u) \right] \cdot du, \end{aligned}$$

where $M_{c2} \equiv EN_c^2(0)$. By means of (13), the preceding relation can be expressed in a much more convenient form:

$$\begin{aligned} R_c(t) &\equiv M_{c2}e^{\kappa t} - M_{c1}^2 \\ &\qquad + \int_0^t e^{\kappa(t-u)} \sum_{n=0}^{d} \rho_n^*(\kappa) p_n \left[\sum_{m=0}^{d} (c \wedge m) p_{nm}(u) \right] \cdot du \\ &= M_{c2}e^{\kappa t} - M_{c1}^2 \\ &\qquad + \int_0^t e^{\kappa(t-u)} \sum_{n=0}^{d} \rho_n^*(\kappa) p_n E\{N_c(u)\,|\,N(0) = n\} \cdot du. \end{aligned} \tag{29}$$

Next, substituting (28) into (29), we obtain

$$\begin{aligned} R_c(t) = M_{c2}e^{\kappa t} - M_{c1}^2 + te^{\kappa t} \sum_{n=0}^{d} (c \wedge n) \cdot \rho_n^*(\kappa) \cdot p_n \\ + \int_0^t e^{\kappa(t-u)} \sum_{n,m=0}^{d} \rho_m^*(\kappa) \cdot p_n \int_0^u e^{\kappa(u-v)} p_{nm}(v) \cdot dv \cdot du. \end{aligned}$$

Let R_c^* be the Laplace transform of R_c and p_{nm}^* that of p_{nm}. The preceding relation then yields

$$\begin{aligned} R_c^*(s) = \frac{M_{c2}}{s - \kappa} - \frac{M_{c1}^2}{s} + \frac{1}{(s - \kappa)^2} \sum_{n=0}^{d} (c \wedge n) \cdot \rho_n^*(\kappa) \cdot p_n \\ + \sum_{n,m=0}^{d} \rho_n^*(\kappa) \cdot \rho_m^*(\kappa) \cdot p_n \frac{p_{nm}^*(s)}{(s - \kappa)^2}. \end{aligned} \tag{30}$$

We know, however, that $p_{nm}(t)$ is of the form

$$p_{nm}(t) = p_m + \sum_{i=1}^{d} \gamma_{nmi} e^{r_i t},$$

so that

$$R_c(t) = \sum_{m,n=0}^{d} (c \wedge n)\cdot(c \wedge m)p_n \sum_{i=1}^{d} \gamma_{nmi} e^{r_i t}.$$

This implies that the only poles of R_c^* are r_i, $i = 1, \cdots, d$, and that $\lim_{t\to\infty} R_c(t) = 0$. Taking these two facts into consideration, we see at once that (30) reduces to

$$R_c^*(s) = \sum_{n,m=0}^{d} \rho_n^*(\kappa)\cdot\rho_m^*(\kappa)\cdot p_n \sum_{i=1}^{d} \frac{\gamma_{nmi}}{(r_i - \kappa)^2}\cdot\frac{1}{s - r_i},$$

provided $\kappa \neq r_i$, $i = 1, 2, \cdots, d$. And referring back to the derivation of (22), it is readily seen that

$$R_c(t) = [\mathbf{q}_d^*(\kappa)]'\cdot\mathbf{B}_d'\cdot\operatorname{diag}\left[0, \frac{e^{r_1 t}}{(r_1 - \kappa)^2}, \cdots, \frac{e^{r_d t}}{(r_d - \kappa)^2}\right]\cdot\mathbf{B}_d\cdot\mathbf{q}_d^*(\kappa),$$
$$\kappa \neq r_i, \qquad i = 1, \cdots, d, \quad (31)$$

where

$$[\mathbf{q}_d^*(\kappa)]' \equiv [\rho_0^*(\kappa)\cdot p_0^{\frac{1}{2}}, \cdots, \rho_d^*(\kappa)\cdot p_d^{\frac{1}{2}}].$$

The modifications needed when κ is equal to one of the characteristic roots are immediate. Let

$$\operatorname{diag}^{(j)}[a_0, a_1, \cdots, a_d]$$

be the diagonal matrix obtained by setting the jth diagonal element of $\operatorname{diag}[a_0, a_1, \cdots, a_d]$ equal to zero. Then if $\kappa = r_j$ we must have, with some as-yet-undetermined constant a and σ_c^2 the variance of $N_c(0)$,

$$R_c(t) = (\sigma_c^2 + a)e^{\kappa t}$$
$$+[\mathbf{q}_d^*(\kappa)]'\cdot\mathbf{B}_d'\cdot\operatorname{diag}^{(j)}\left[0, \frac{e^{r_i t}}{(r_i - \kappa)^2}, i = 1, \cdots, d\right]\cdot\mathbf{B}_d\cdot\mathbf{q}_d^*(\kappa).$$

But $R_c(0) = \sigma_c^2$, so that

$$a = -[\mathbf{q}_d^*(\kappa)]'\cdot\mathbf{B}_d'\cdot\operatorname{diag}^{(j)}[0, (r_i - \kappa)^{-2}, i = 1, \cdots, d]\cdot\mathbf{B}_d\cdot\mathbf{q}_d^*(\kappa).$$

Hence,

$$R_c(t) = \sigma_c^2 e^{\kappa t} + [\mathbf{q}_d^*(\kappa)]'\cdot\mathbf{B}_d'\cdot\operatorname{diag}^{(j)}\left[0, \frac{e^{r_i t} - e^{\kappa t}}{(r_i - \kappa)^2}, i = 1, \cdots, d\right]$$
$$\cdot\mathbf{B}_d\cdot\mathbf{q}_d^*(\kappa). \quad (32)$$

It should be noted that (32) is valid even if $\kappa \neq r_i$, $i = 1, \cdots, d$, and that it should be used in the computations [rather than (31)] whenever κ is "close" to one of the characteristic roots so as to avoid accuracy losses (see below). Since κ is arbitrary, one could always choose it so that it is not "close" to any of the characteristic roots. But, as shown next, it is often preferable to select it in such a way as to reduce the amount of computation, and this, in turn, may dictate

the use of (32). As we have seen,

$$R_c(l) = [\mathbf{q}_d^*(\kappa)]' \cdot \mathbf{B}_d' \cdot \operatorname{diag}\left[0, \frac{e^{r_1 t}}{(r_1 - \kappa)^2}, \cdots, \frac{e^{r_d t}}{(r_d - \kappa)^2}\right] \cdot \mathbf{B}_d \cdot \mathbf{q}_d^*(\kappa).$$

But, as noted previously, $\mathbf{p}_d^{(\frac{1}{2})}$ is the characteristic vector associated with the vanishing root (r_0), so that

$$(\mathbf{p}_d^{(\frac{1}{2})})' \cdot \mathbf{B}_d' \cdot \operatorname{diag}\left[0, \frac{e^{r_1 t}}{(r_1 - \kappa)^2}, \cdots, \frac{e^{r_d t}}{(r_d - \kappa)^2}\right] \cdot \mathbf{B}_d \cdot \mathbf{p}_d^{(\frac{1}{2})} = 0. \quad (33)$$

Consequently, (31) remains valid for all $[\mathbf{q}_d^*(\kappa)]'$ of the form

$$\{[\rho_0^*(\kappa) + \gamma] \cdot p_0^{\frac{1}{2}}, \cdots, [\rho_d^*(\kappa) + \gamma] \cdot p_d^{\frac{1}{2}}\},$$

where γ is an arbitrary constant. [The same remark, of course, also applies to (32).]

We are therefore always at liberty to add the same constant to all the $\rho_i^*(\kappa)$'s. Under some circumstances, this degree of freedom, together with the one provided by the introduction of κ, can be used to reduce the dimension of $\mathbf{B}'$ and $\mathbf{B}$: entire rows in $\mathbf{B}'$ and the corresponding columns in $\mathbf{B}$ can be set equal to zero without affecting the computation either of (31) or (32), or of the variance of the switch-count load. It is relevant to note here that this reduction would be largely illusory were it not for the fact that the normalized components of any of the characteristic vectors can be obtained without having to compute other components of that vector (see below).

According to the result of Section 3.1, the covariance can always be cast in a form that involves only the first c components of the characteristic vectors. But when the input and departure rates for $0 \leqq n < c$ are linear in n, the covariance can also be expressed in terms of the last $d - c + 1$ components of these vectors. Indeed, the rates are then of the form

$$\begin{aligned} \lambda_n &= \lambda n + \lambda', \\ \mu_n &= \mu n + \mu', \qquad n = 0, 1, \cdots, c - 1, \end{aligned}$$

so that

$$\lambda_n - \mu_n = (\lambda - \mu)n + (\lambda' - \mu').$$

Hence, with $\kappa = \lambda - \mu$ and $\gamma = \mu' - \lambda'$, (27) yields

$$\rho_n^*(\lambda - \mu) = \begin{cases} 0 & \text{if } n = 0, 1, \cdots, c - 1, \\ (\mu - \lambda)c - \mu_c + \mu' - \lambda' & \text{if } n = c, \\ (\mu - \lambda)c + \mu' - \lambda' & \text{if } n = c + 1, \cdots, d. \end{cases}$$

For the random (Poisson) and the quasi-random inputs, the ρ's take the

following simple form whenever the service time is exponential with mean 1.

(*i*) Random input ($\lambda_n = \alpha$, $n = 0, 1, \cdots$):

$$\rho_n^*(-1) = \begin{cases} 0 & \text{if } n = 0, 1, \cdots, c-1, \\ -\alpha & \text{if } n = c, \\ c - \alpha & \text{if } n = c+1, \cdots, d. \end{cases}$$

(*ii*) Quasi-random input [N sources, $\lambda_n = (N - n)\lambda$, $n = 0, 1, \cdots, N$]:

$$\rho_n^*[-(1+\lambda)] = \begin{cases} 0 & \text{if } n = 0, 1, \cdots, c-1, \\ (c - N)\lambda & \text{if } n = c, \\ c + (c - N)\lambda & \text{if } n = c+1, \cdots, d. \end{cases}$$

From the preceding developments, we see that the ρ's can be chosen in such a way that the number of components of the characteristic vectors needed to express R_c is the smaller of the two integers c and $d - c + 1$. In particular, in the case of loss systems, only the $(c+1)$st component of each vector is needed.

The parameters κ and γ can also be chosen so that only the first $c + 1$ components of the characteristic vectors actually enter in the expression of R_c. This will be the case if we set $\kappa = \mu_{c-1} - \lambda_{c-1}$ and $\gamma = c(\mu_{c-1} - \lambda_{c-1})$.

In Ref. 3, the derivation of the covariance function for loss systems [$d = c$, $N_c(t) \equiv N(t)$] with Poisson input and exponential service time makes use of the differential equations

$$\frac{d}{dt} E\{N(t) \mid N(0) = m\} = -E\{N(t) \mid N(0) = m\} + \alpha[1 - p_{mc}(t)], \qquad m = 0, 1, \cdots, c.$$

These equations appear here as that particular instance of (26) for which $\kappa = -1$, $\lambda_n = \alpha$, $n = 0, 1, \cdots, c-1$, and $\mu_n = n$, $n = 1, \cdots, c$. Note also that now $\sum_{n=0}^{c-1} p_{mn}(t) = 1 - p_{mc}(t)$. But we stress that, in Ref. 3, the determination of the covariance relies on known recurrence relations between the so-called "sigma" functions (Ref. 3, pp. 129 and 143 ff.); the more general problem considered in the present paper is not as readily amenable to such a treatment because of the greater complexity of the expressions that would now have to be used instead of the sigma functions. As we have seen, however, relatively simple formulas for R_c can be obtained without extensive algebraic developments as long as the underlying process is reversible.

IV. VARIANCE OF THE SWITCH-COUNT LOAD

The variance of the switch-count load is now readily obtained. Depending on which expression we select for R_c, we have either

(*i*) Var $L_n(T)$

$$= n^{-2} \sum_{k=-n}^{n} (n - |k|) R_c(k\tau)$$

$$= n^{-2} \cdot \mathbf{q}_d' \cdot \mathbf{B}_d' \cdot \mathrm{diag} \left[0, \sum_{k=-n}^{n} (n - |k|) e^{r_i |k| t}, i = 1, \cdots, d \right] \cdot \mathbf{B}_d \cdot \mathbf{q}_d^*(\kappa), \quad (34)$$

with

$$\mathbf{q} = (\rho_0 p_0^{\frac{1}{2}}, \cdots, \rho_{c-1} p_{c-1}^{\frac{1}{2}}, 0, \cdots, 0), \text{ or}$$

(*ii*) Var $L_n(T) = n^{-2} [\mathbf{q}_d^*(\kappa)]' \cdot \mathbf{B}_d'$

$$\cdot \mathrm{diag} \left[0, \sum_{k=-n}^{n} (n - |k|) \frac{e^{r_i |k| t}}{(r_i - \kappa)^2}, i = 1, \cdots, d \right] \cdot \mathbf{B}_d \cdot \mathbf{q}_d^*(\kappa), \quad (35)$$

with $\kappa \neq r_i$, $i = 1, \cdots, d$, or

(*iii*) Var $L_n(T) = n^{-2} \{ \sigma_c^2 - [\mathbf{q}_d^*(\kappa)]' \cdot \mathbf{B}_d$

$$\cdot \mathrm{diag}^{(j)} [0, (r_i - \kappa)^{-2}, i = 1, \cdots, d] \cdot \mathbf{B}_d \cdot \mathbf{q}_d \}$$

$$\cdot \sum_{k=-n}^{n} (n - |k|) e^{\kappa |k| \tau} + n^{-2} [\mathbf{q}_d^*(\kappa)]' \cdot \mathbf{B}_d'$$

$$\cdot \mathrm{diag}^{(j)} \left[0, \sum_{k=-n}^{n} (n - |k|) \frac{e^{r_i |k| \tau}}{(r_i - \kappa)^2}, i = 1, \cdots, d \right] \cdot \mathbf{B}_d \cdot \mathbf{q}_d^*(\kappa), \quad (36)$$

where $\kappa \neq r_i$ for $i \neq j$.

We now make use of the following identity (Ref. 3, p. 137):

$$\sum_{n=-k}^{k} (n - |k|) e^{-2|k| u} = n \cdot \coth u - \frac{1 - e^{-2nu}}{2} \cdot \mathrm{csch}^2 u.$$

By means of this relation, (34) to (36) can also be written as

$$\mathrm{Var}\, L_n(T) = n^{-1} \cdot \mathbf{q}_d' \cdot \mathbf{B}_d' \cdot \mathrm{diag} \left[0, \coth \left(\frac{-\tau r_i}{2} \right) - \frac{1 - e^{n \tau r_i}}{2n} \cdot \mathrm{csch}^2 \left(\frac{-\tau r_i}{2} \right), i = 1, \cdots, d \right] \cdot \mathbf{B}_d \mathbf{q}_d; \quad (34a)$$

$$\operatorname{Var} L_n(T) = n^{-1}\cdot[\mathbf{q}_d^*(\kappa)]'\cdot\mathbf{B}_d'$$
$$\cdot\operatorname{diag}\left[0, \frac{1}{(r_i-\kappa)^2}\left\{\coth\left(\frac{-\tau r_i}{2}\right) - \frac{1-e^{n\tau r_i}}{2n}\right.\right.$$
$$\left.\left.\cdot\operatorname{csch}^2\left(\frac{-\tau r_i}{2}\right)\right\}, i = 1, \cdots, d\right]\cdot\mathbf{B}_d\cdot\mathbf{q}_d^*(\kappa),$$
$$\kappa \neq r_i, \qquad i = 1, \cdots, d; \quad (35a)$$

and

$$\operatorname{Var} L_n(T)$$
$$= n^{-1}\{\sigma_c^2 - [\mathbf{q}_d^*(\kappa)]'\cdot\mathbf{B}_d'\cdot\operatorname{diag}^{(j)}[0, (r_i-\kappa)^{-2}, i=1, \cdots, d]\cdot\mathbf{B}_d\cdot\mathbf{q}_d^*(\kappa)\}$$
$$\cdot\left\{\coth\left(\frac{-\tau\kappa}{2}\right) - \frac{1-e^{n\tau\kappa}}{2n}\cdot\operatorname{csch}^2\left(\frac{-\tau\kappa}{2}\right)\right\} + n^{-1}\cdot[\mathbf{q}_d^*(\kappa)]'\cdot\mathbf{B}_d'$$
$$\cdot\operatorname{diag}^{(j)}\left[0, \frac{1}{(r_i-\kappa)^2}\left\{\coth\left(\frac{-\tau r_i}{2}\right) - \frac{1-e^{n\tau r_i}}{2n}\right.\right.$$
$$\left.\left.\cdot\operatorname{csch}^2\left(\frac{-\tau r_i}{2}\right)\right\}, i = 1, \cdots, d\right]\cdot\mathbf{B}_d\cdot\mathbf{q}_d^*(\kappa), \quad (36a)$$

where $\kappa \neq r_i$ for $i \neq j$.

Let $\operatorname{Var} L_\infty(T) \equiv \lim_{n\to\infty} \operatorname{Var} L_n(T)$ be the variance of the load measurement obtained by continuous observation of the number of busy servers. If we replace τ by T/n in (34a) to (36a) and then let n tend to infinity while keeping T fixed, we obtain the following formulas:

$$\operatorname{Var} L_\infty(T) = -\frac{2}{T}\cdot\mathbf{q}_d'\cdot\mathbf{B}_d'$$
$$\cdot\operatorname{diag}\left[0, \frac{1}{r_i}\left(1 + \frac{1-e^{r_iT}}{Tr_i}\right), i = 1, \cdots, d\right]\cdot\mathbf{B}_d\cdot\mathbf{q}_d, \quad (34b)$$

$$\operatorname{Var} L_\infty(T) = -\frac{2}{T}[\mathbf{q}_d^*(\kappa)]'\cdot\mathbf{B}_d'$$
$$\cdot\operatorname{diag}\left[0, \frac{1}{r_i(r_i-\kappa)^2}\left(1 + \frac{1-e^{r_iT}}{Tr_i}\right), i = 1, \cdots, d\right]\cdot\mathbf{B}_d\cdot\mathbf{q}_d^*(\kappa),$$
$$\kappa \neq r_i, \qquad i = 1, \cdots, d, \quad (35b)$$

and

$$\operatorname{Var} L_\infty(T) = \frac{2}{T}\{\sigma_c^2 - [\mathbf{q}_d^*(\kappa)]'\cdot\mathbf{B}_d'$$
$$\cdot\operatorname{diag}^{(j)}[0, (r_i-\kappa)^{-2}, i = 1, \cdots, d]$$
$$\cdot\mathbf{B}_d\cdot\mathbf{q}_d^*(\kappa)\}\frac{1}{\kappa}\left(1 + \frac{1-e^{\kappa T}}{T\kappa}\right) - \frac{2}{T}[\mathbf{q}_d^*(\kappa)]'\cdot\mathbf{B}_d'$$
$$\cdot\operatorname{diag}^{(j)}\left[0, \frac{1}{r_i(r_i-\kappa)^2}\left(1 + \frac{1-e^{r_iT}}{Tr_i}\right), i = 1, \cdots, d\right]\cdot\mathbf{B}_d\cdot\mathbf{q}_d^*(\kappa),$$
$$\kappa \neq r_i \quad \text{for} \quad i \neq j. \quad (36b)$$

We note that the formula for the variance of sums of dependent random variables makes it possible to compute the covariance between load measurements performed over distinct time intervals. Indeed, consider for instance a sequence of load measurements over the intervals $(0, T]$, $(T, 2T]$, $(2T, 3T]$, $\cdots$. Let $L_n^{(i)}(T)$ be the switch-count load over the ith interval $(i = 0, 1, \cdots)$, $S_n(t) = n^2 L_n(t)$, and $\Gamma_n^{(i)} \equiv \mathrm{Cov}\,[L_n^{(0)}(T), L_n^{(i)}(T)]$. Then

$$\mathrm{Var}\, S_{n(k+1)}[(k+1)T] = (k+1)\,\mathrm{Var}\, S_n(T) + 2\sum_{i=1}^{k} (k+1-i) n^2 \Gamma_n^{(i)}(T)$$

and

$$\Gamma_n^{(k)}(T) = \frac{(k+1)^2}{2}\,\mathrm{Var}\, L_{n(k+1)}[(k+1)T] - \frac{k+1}{2}\,\mathrm{Var}\, L_n(T) - \sum_{i=1}^{k-1} \Gamma_n^{(i)}(T).$$

The preceding formulas may be used to determine the $\Gamma_n^{(k)}(T)$ recurrently. But the results of such computations shall be exact only if, for some choice of the time origin, all the scanning instants are multiples of τ.

We conclude this section with the remark that the variance formulas (34), (34a), and (34b) are valid for arbitrary reversible Markov processes with finite state spaces.

V. NUMERICAL CONSIDERATIONS

The exact variance formulas of the preceding section are very well suited to electronic computation and are easily programmed since, apart from straightforward evaluations of hyperbolic functions and simple products of matrices and vectors, they only involve the determination of characteristic values and vectors for which powerful subroutines are readily available. The fact that $\mathbf{S}_d$ is symmetric and tridiagonal (or reducible to tridiagonal form by an orthogonal similarity transformation) allows us to use the subprogram TQL2, which is particularly efficient under the present circumstances (Ref. 11, pp. 227–240). Without going into details, we mention here only that this subprogram is based on the so-called QR-algorithm and relies on the construction of a sequence of symmetric tridiagonal matrices, $\mathbf{S}_d^{(n)}$, $n = 1, 2, \cdots$, unitarily similar to $\mathbf{S}_d$, which converges to diag $[0, r_1, r_2, \cdots, r_d]$. At the nth iteration $\mathbf{S}_d^{(n)}$ is expressed as a product of an

orthogonal matrix $\mathbf{Q}_d^{(n)}$ and a lower-triangular matrix $\mathbf{L}_d^{(n)}$:

$$\mathbf{S}_d^{(n)} = \mathbf{Q}_d^{(n)} \cdot \mathbf{L}_d^{(n)}.$$

In Ref. 11, this decomposition is carried out by the Givens' triangularization in which the $\mathbf{Q}_d^{(n)}$'s are expressed as products of simple plane rotations (Ref. 7, pp. 239–240).

The $(n + 1)$st iterate of $\mathbf{S}_d$ is then given by

$$\mathbf{S}_d^{(n+1)} = \mathbf{L}_d^{(n)} \cdot \mathbf{Q}_d^{(n)},$$

whose unitary similarity to $\mathbf{S}_d^{(n)}$ (and hence to $\mathbf{S}_d$) follows from the relation

$$\mathbf{L}_d^{(n)} \cdot \mathbf{Q}_d^{(n)} = (\mathbf{Q}_d^{(n)})' \cdot \mathbf{S}_d^{(n)} \cdot \mathbf{Q}_d^{(n)}.$$

This method avoids the numerical difficulties frequently associated with the computation of the zeros of the characteristic polynomial.

As shown in Ref. 11, p. 228,

$$\mathbf{Q}_d^{(n)} = \mathbf{R}_d^{(n)} \cdot \cdots \cdot \mathbf{R}_d^{(1)}, \tag{37}$$

where $\mathbf{R}_d^{(i)}$ is of the form:

$$\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & C_i & S_i & & & \\ & & -S_i & C_i & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \leftarrow \text{row } i. \tag{38}$$

According to Ref. 11, p. 231, the matrix $\mathbf{B}_d'$ (whose columns are the characteristic vectors of $\mathbf{S}_d$) is given "almost to working accuracy" by

$$\mathbf{Q}_d^{(n)} \cdot \cdots \cdot \mathbf{Q}_d^{(1)}, \tag{39}$$

where n ($\leqq 30$) is the number of iterations needed for the (numerical) symmetrization of $\mathbf{S}_d$. Taking (37) to (39) into account, it is then readily seen that the elements of $\mathbf{B}_d'$ can be determined row by row which, as we have remarked earlier, is a desirable feature in the present context.

Computations have been carried out for systems having as many as 400 devices (and hence transition-rate matrices of order 401) to determine the numerical accuracy of the approach described in the preceding sections. Checks were performed by comparing the value of Var $L_1(T)$ obtained by means of (35a) or (36a) with the corresponding σ_0^2, which can be calculated directly and independently from the equilibrium state probabilities. These two quantities, which are theoretically equal,

turned out in all cases to agree to at least 10 decimal places with the greatest difference occurring when d was largest. Hence, our procedure indeed yields very accurate results for the type of systems that are likely to occur in practice. But when d is large, the storage requirements and the amount of computations become critical. It is therefore always important to select κ and γ in such a way as to minimize the number of $\mathbf{B}'$ rows that actually enter into the computations. (It follows from earlier remarks that this number, for proper choice of κ and γ, never exceeds the integral part of $(d + 1)/2$.) Further reduction can also be achieved by excluding the states whose probabilities of occurrences are so small that neglecting them will not materially affect the final results. In this connection, we make the following remarks.

The variance of the switch-count load is perturbed by at most

$$[\rho_j^*(\kappa)]^2 \cdot p_j \cdot \sigma_c^2$$

if p_j is set equal to zero in the particular formula used to evaluate Var $L_n(T)$. Hence, since

$$\text{Var } L_n(T) \geqq \sigma_c^2/n,$$

we always have the following upper bound for the relative error, ϵ_j, induced by setting p_j equal to zero:

$$\epsilon_j \leqq [\rho_j^*(\kappa)]^2 \cdot p_j \cdot n, \qquad j = 0, 1, \cdots, d.$$

For a given relative accuracy of Var $L_n(T)$, these inequalities make it possible to determine ahead of time whether some components of the characteristic vectors can be "safely" eliminated from the computations. In large systems, the gains achieved by such a reduction may be quite substantial, as either low occupation states [$N(t)$ small] and/or high occupation states [$N(t)$ large] have then frequently very small probabilities of occurrences.

Computations could be arranged to determine only those characteristic roots that are required to reach a given degree of accuracy [plus those needed to compute $\mathbf{B}_d \cdot \mathbf{q}_d^*(\kappa)$]. This is rather readily achieved in loss systems with Poisson input and exponential service times since, in this case, the coefficients b_i^2 of

$$\coth\left(\frac{-\tau r_i}{2}\right) - \frac{1 - e^{n\tau r_i}}{2n} \cdot \text{csch}^2\left(\frac{-\tau r_i}{2}\right)$$

in the variance formulas of Section 3.1 are then monotonically decreasing as $|r_i|$ increases:

$$b_i^2 < b_j^2 \quad \text{if} \quad |r_i| > |r_j|, \qquad i, j = 1, \cdots, d.$$

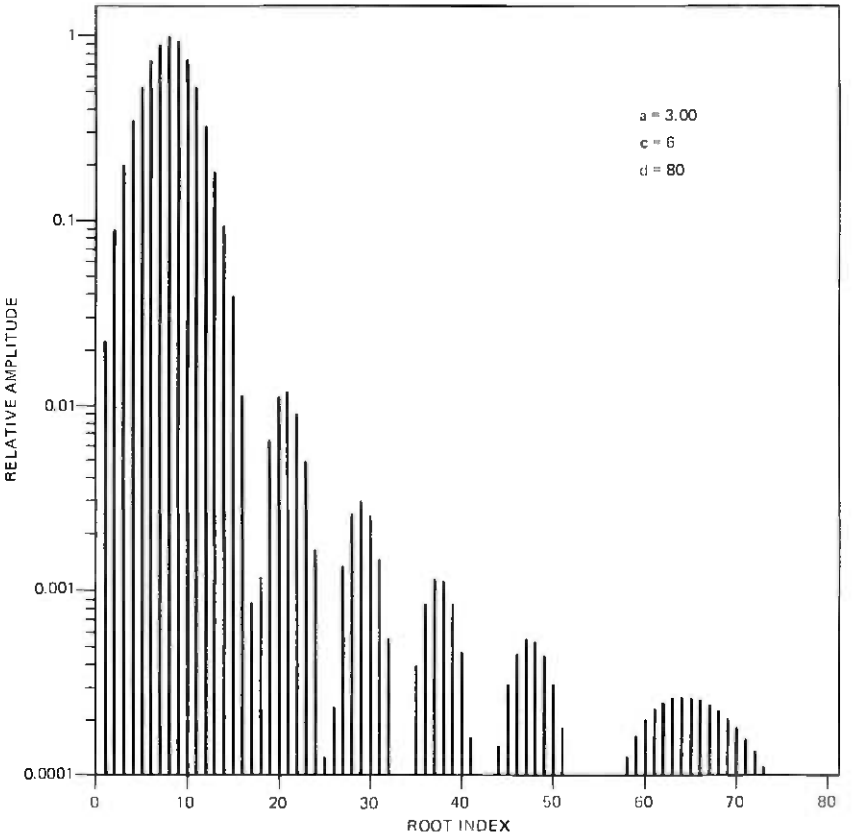


Fig. 1—Spectral measure of the carried load process.

But, in general, one would encounter an additional difficulty, namely, that the b_i^2's do not have the monotonicity property alluded to above and may actually fluctuate widely. This is illustrated in Fig. 1, where the roots are assumed to be indexed in order of increasing magnitude and the ordinates are the corresponding b_i^2's, normalized in such a way that $\max_i b_i^2 = 1$.

The computations should be based on (35a)–(36a) or on (35b)–(36b) in the case of continuous measurements—as these formulas provide us with all the flexibility needed to cut down both storage space and computation time. When choosing between (35a) and (36a) or between (35b) and (36b), one should keep in mind that, for κ close to r_i, the difference $r_i - \kappa$ may not be determinable with enough precision to allow accurate computation of Var $L_n(T)$. This is shown in Table I where $\kappa = -1$ and r_1 is the root of smallest positive absolute value.

Table I — Loss system, 80 servers, Poisson input, exponential service

Offered Load in Erlangs	$1 + r_1$	σ_c^2	Var $L_1(T)$	
			Formula (36a)	Formula (35a)
10	1.09×10^{-13}	10.000000	10.000000	0.034614
20	-9.90×10^{-13}	20.000000	20.000000	0.013173
30	-1.14×10^{-11}	30.000000	30.000000	20.891365
40	-3.38×10^{-6}	39.999986	39.999986	39.999926

(Note that the last two columns of this table should be equal and that errors of the same magnitude would arise if one were to use (18) of Ref. 3.) In all our computations, we have made use of (35a) and (36a) whenever $|\kappa - r_i| < 10^{-4}$ for some i. This bound for $|\kappa - r_i|$ is both large enough to ward off appreciable accuracy losses and small enough, under prevailing conditions, to be satisfied by only one root.

VI. REMARKS ON INFINITE SYSTEMS

It is known that infinite systems can be regarded as limits of finite ones,[12] and it is therefore of practical interest to have information concerning the spacing of the characteristic values as the dimension, d, of these approximating systems becomes large. Indeed, as d increases, computational difficulties may arise because of a lack of separation between these roots. Such problems would certainly come up sooner or later if the spectrum of $\mathbf{A} \equiv \lim_{d\to\infty}\mathbf{A}_d$ happens to be dense over some interval as, for instance, in the case of a single-server queue with Poisson input, exponential service time, and unlimited waiting room (Ref. 12, pp. 365–366). Infinite systems with well-separated roots do, of course, also occur. As an example of this type, we mention the systems with an infinite number of servers, Poisson input, and exponential service which often provide useful idealizations. (In this case, as is well known, the nonvanishing characteristic roots are the negative integers, $-1, -2, -3, \cdots$.) Other examples of systems with discrete spectra are given in Ref. 12, where sufficient conditions for this to occur are discussed in some details; but in all these instances the λ_n and μ_n both increase as n^ν for some $\nu > 0$. This condition is unlikely to be satisfied in queuing systems; generally, in this particular area, the arrival and the departure rates remain bounded:

$$0 \leqq \lambda_n \leqq \Lambda < \infty, \qquad 0 \leqq \mu_n \leqq M < \infty, \qquad n = 0, 1, \cdots. \tag{40}$$

As briefly described below, these inequalities imply the existence of definite bounds for the spectrum of $\mathbf{A}$.

Consider an infinite system, and let $\mathbf{A}$ be its (infinite) transition-rate matrix. Let $\mathbf{A}_d$ be the matrix obtained by retaining only the elements belonging to the first $(d + 1)$ rows and columns of $\mathbf{A}$ and then setting λ_d equal to 0. Let $r_{d0}(=0) > r_{d1} > \cdots$ be the characteristic roots of $\mathbf{A}_d$. Then, under conditions (40) it can be shown that, for any $k \geqq 0$:

(*i*) $\quad |r_{dk}| < \Lambda + M \quad$ for d sufficiently large,

(*ii*) $\quad |r_{d,d-k}| < 2(\Lambda + M) \quad$ for $d \geqq k$.

Either of these two inequalities implies that the characteristic roots do not remain separated as $d \to \infty$ whenever (40) is satisfied. Under the more stringent requirements that (40) holds and that

$$\lim_{n\to\infty} \lambda_n = \Lambda, \qquad \lim_{n\to\infty} \mu_n = M,$$

more precise statements can be made, namely, that, for all k's and d's,

$$|r_{dk}| < (\sqrt{\Lambda} + \sqrt{M})^2$$

and that the spectrum of $\mathbf{A}$ always comprises a closed interval, viz.,

$$\Omega = [-(\sqrt{\Lambda} + \sqrt{M})^2, -(\sqrt{\Lambda} - \sqrt{M})^2].$$

(In addition to Ω, the spectrum of $\mathbf{A}$ may also include a finite number of roots in $[-(\sqrt{\Lambda} - \sqrt{M})^2, 0]$.) But it turns out (as will be shown elsewhere) that, as d increases, the characteristic roots of $\mathbf{A}_d$ fill Ω rather "evenly"; furthermore, for practical accuracy levels, large values of d are needed only when the length of Ω tends to be relatively large (a circumstance corroborated by extensive computations). Hence, within the present framework, it appears that root-spacing is not likely to be critical except in the improbable event that extreme precision is required.

VII. ACKNOWLEDGMENT

I wish to thank Peter Businger for very helpful suggestions.

REFERENCES

1. W. S. Hayward, Jr., "The Reliability of Telephone Traffic Load Measurements by Switch Counts," B.S.T.J., *31*, No. 2 (March 1952), pp. 357–377.
2. C. Palm, "Accuracy of Measurements in Determining Traffic Volumes by the Scanning Method," Tekniska Medelanden fran Kungl. Telegrafstyrelen, No. 7–9, 1941.
3. V. E. Beneš, "The Covariance Function of a Simple Trunk Group, with Applications to Traffic Measurement," B.S.T.J., *40*, No. 1 (January 1961), pp. 117–148.
4. S. R. Neal and A. Kuczura, "A Theory of Traffic-Measurement Errors for Loss Systems with Renewal Input," B.S.T.J., *52*, No. 6 (July–August 1973), pp. 967–990.
5. D. V. Widder, *The Laplace Transform*, Princeton: Princeton University Press, 1946.

6. J. L. Doob, *Stochastic Processes*, New York: John Wiley, 1953.
7. J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford: Clarendon Press, 1965.
8. E. Reich, "Waiting Times When Queues are in Tandem," Ann. Math. Stat., *28*, No. 3 (1957), pp. 768–773.
9. S. Karlin and J. McGregor, "The Classification of Birth and Death Processes," Trans. Amer. Math. Soc., *86*, 1957, pp. 366–400.
10. K. L. Chung, *Markov Chains*, second edition, New York: Springer-Verlag, 1967.
11. H. Bowdler, R. S. Martin, C. Reinsch, and J. H. Wilkinson, "The QR and QL Algorithms for Symmetric Matrices," Contribution II/3, pp. 227–240, in *Handbook for Automatic Computation, Vol. II, Linear Algebra*, edited by J. H. Wilkinson and C. Reinsch, New York, Heidelberg, Berlin: Springer-Verlag, 1971.
12. W. Ledermann and G. E. H. Reuter, "Spectral Theory for the Differential Equations of Simple Birth and Death Processes," Phil. Trans. Roy. Soc. of London, Series A, *246*, 1954, pp. 321–369.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 54, No. 7, September 1975
Printed in U.S.A.

T1 Carbon Transmitter Model for Use in Computer-Aided Analysis of Telephone Set Transmission Characteristics

By D. R. MEANS

(Manuscript received March 5, 1975)

A carbon transmitter model is presented, the purpose of which is to serve as a tool for computer-aided analysis of telephone set transmission characteristics. The derivation of the model is based upon the physical theory of the device. The parameters in the model are evaluated by comparing the analytically derived expressions for the device characteristics to the measured characteristics of a typical device. Because these parameters are related to the physical theory, the model not only serves its desired practical end, but also serves as a vehicle whereby an understanding is obtained of the relationship between device characteristics and physical theory.

I. INTRODUCTION

Computer-aided optimization of the transmission characteristics of telephone sets requires that accurate models be obtained for all transmission-related telephone set components. A carbon transmitter model has been derived for this purpose. This model has been used in a telephone set transmission analysis computer program, and good agreement between computed and measured transmission characteristics was obtained.

The dc V-I characteristic of the carbon transmitter is nonlinear. This nonlinearity must be taken into account in the dc model so that, in the transmission analysis program, the operating point of the transmitter can be determined, as well as the operating points of any nonlinear telephone speech network components, e.g., silicon carbide varistors. Thus, the dc model is a voltage-dependent resistance.

The ac model is similar to that of a vacuum-tube triode, consisting of a Thevenin-equivalent resistance and voltage source. The voltage source is dependent on the amplitude of the force acting on the carbon granules because of the acoustic excitation of the transmitter. This

force is in turn the output of a filter which represents the transmitter structure itself. The input to the filter is the acoustic sound pressure. This filter can be represented by an electrical equivalent circuit,[1] the derivation of which is straightforward. However, it is the characterization of the effects taking place within the carbon chamber which is of primary interest here. The filter is represented simply by its measured frequency response.

Values for the various parameters of the model were determined by comparing the expressions derived analytically for the characteristics of the transmitter to the measured characteristics of a typical device. All measurements were made with the transmitter face in a vertical plane, in a telephone handset, and in a position relative to the artificial mouth as specified in IEEE Standard 269-1971.[2] Also, the transmitter was mechanically and acoustically conditioned prior to the measurements. The acoustic conditioning signal was swept between 300 and 3300 Hz at a rate of six sweeps per second and frequency-weighted corresponding to the average sound pressure spectrum of continuous speech, and had an average sound pressure level of 94 dB (re 0.0002 dyn/cm^2). The conditioning signal was applied for 3 s.

II. DC MODEL

As shown in Fig. 1, the dc V-I characteristic of the carbon transmitter is nonlinear. Goucher[3] attributed the nonlinearity to the effect of joule heating on the contact resistance between carbon granules. Later, Mol,[4] disputing Goucher, attributed the nonlinearity to the

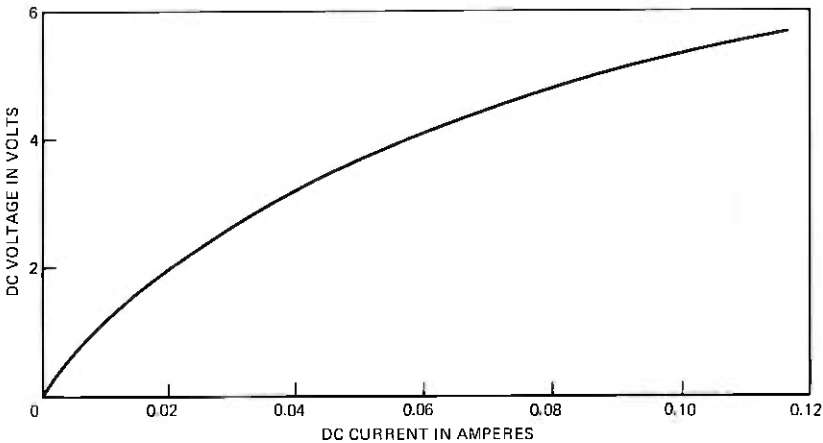


Fig. 1—Direct-current V-I characteristic.

effect of electrostatic forces between carbon granules. On this basis, he derived an expression for the dc transmitter resistance as a function of voltage and found it to agree fairly well with measured transmitter resistance. However, Mol's expression for the dc voltage dependence of the resistance variation when an acoustic signal acts on the transmitter agrees rather poorly with measurement. This casts doubt on the electrostatic force theory. In fact, experimental results indicate that the effect of electrostatic forces between carbon granules is negligible. These experiments are described in Appendix A. On the other hand, a more recent study of the theory of electric contacts tends to support Goucher's theory. Holm[5] treats the subject of the effect of joule heating on contact resistance extensively, and his results will be applied to the derivation of the transmitter model.

As will be seen, the nonlinearity of the V-I characteristic cannot be accounted for entirely by the effect of joule heating on contact resistance. The effect of the thermal expansion of the carbon chamber due to joule heating must also be considered. This effect is readily demonstrated experimentally because of the relatively long time constant involved. If the transmitter current is changed abruptly, a slowly decaying voltage transient is observed, owing to the hysteresis associated with the expansion of the carbon chamber. The time constant is approximately 1 s.

A cross section of the T1 transmitter is shown in Fig. 2. The carbon chamber consists of a movable dome electrode connected to a fixed conical back electrode by a flexible, nonconducting chamber closure. According to the results of Fritsch's analysis[6] of the thermal response of the transmitter structure, the transient effect is due primarily to the expansion of the dome electrode. As the dome electrode expands, it compresses the carbon granules, lowering their resistance. Fritsch called this effect "thermal packing." Of course, after the transient has decayed, the transmitter can be reconditioned to unpack the granules. However, following the reconditioning, a new thermal equilibrium will be established so that some degree of thermal packing will still occur. This effect, as well as the effect of joule heating on contact resistance, must be included in the model. An understanding of these effects is based on a consideration of the factors affecting the resistance of a single carbon contact.

2.1 Effect of contact force on contact resistance

The contact resistance between two carbon granules is related to the magnitude of the force pressing the granules together by the expression:

$$r_k = KP^{-\gamma}, \tag{1}$$

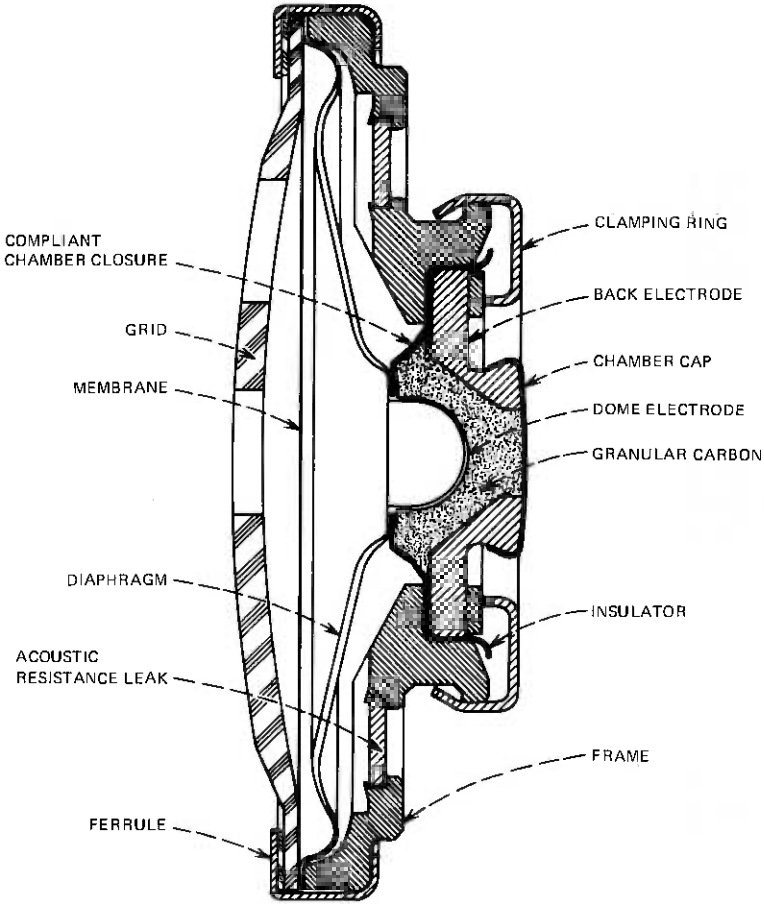


Fig. 2—Cross section of T1 transmitter.

where P is the contact force and K and γ are constants. Equation (1), which is generally valid for electric contacts, was shown by Goucher[3] to be valid for carbon granules, based on direct measurement of contact resistance and force. Later, this equation was indirectly shown by Joscheck[7] to be valid based on measurements of the bulk resistance of carbon granules as a function of the filling height of the granules in the measurement chamber.

In the carbon chamber in the T1 transmitter, the significant forces acting on the carbon granules are the gravitational force resulting from the weight of the granules themselves, the force resulting from the thermal expansion of the dome electrode, and, when the transmitter is acoustically excited, the force resulting from the acoustic

pressure acting on the diaphragm and dome electrode. Because of the random orientation of the contacts throughout the chamber with respect to the directions of the applied forces, the contact force and, hence, the contact resistance will also vary randomly. However, because of the large number of contacts involved, the random variation in contact resistance can be ignored. Only the larger overall gradients in bulk resistivity will be considered.

For simplicity, an approximate chamber geometry is assumed in which the two electrodes are concentric hemispheres. A spherical coordinate system is defined such that the hemispherical chamber walls lie at constant radial distances from the origin, the inner and outer radii being designated a and b, respectively. The component of contact force resulting from the weight of the granules themselves is represented by $P_g(r, \theta, \phi)$. The component resulting from the thermal expansion of the dome electrode, being proportional to the power dissipated by the transmitter, is represented by (V^2/R) $P_d(r, \theta, \phi)$, where V and R are the transmitter dc voltage and resistance, respectively. The component resulting from the acoustic excitation of the transmitter, being a function of time as well as of the spatial coordinates, is represented by $\Delta P(r, \theta, \phi, t)$.

Equation (1), therefore, becomes

$$r_k(r, \theta, \phi, t) = K[P_g(r, \theta, \phi) + (V^2/R)P_d(r, \theta, \phi) + \Delta P(r, \theta, \phi, t)]^{-\gamma}. \quad (2)$$

It will subsequently be seen that, for normal speech signal levels, ΔP is small enough compared to the static components of contact force that its effect on the dc component of contact resistance is negligible. Thus, the relationship between the dc contact resistance and the contact force is

$$r_k(r, \theta, \phi) = K[P_g(r, \theta, \phi) + (V^2/R)P_d(r, \theta, \phi)]^{-\gamma}. \quad (3)$$

Although a change in dc transmitter resistance is observed when acoustic excitation is applied to the transmitter, this is judged to be due to the effect of the acoustic excitation on the state of compactness of the carbon granules rather than to the effect of the nonlinearity of the contact resistance–contact force characteristic. The effect of acoustic excitation on the state of compactness of the granules will be discussed further.

2.2 *Effect of joule heating on contact resistance*

According to Holm's analysis[5] of the effect of joule heating on contact resistance, if certain assumptions regarding the temperature

dependence of the electrical resistivity and the thermal conductivity of the contact members are satisfied, then the effect of joule heating can be accounted for by multiplying the contact resistance by the factor

$$\eta(V_k) = [B + (1 - B)(A_k/V_k) \tan^{-1}(V_k/A_k)]^{-1}, \tag{4}$$

where V_k is the contact voltage and A_k and B are constants. Since the random variation in contact voltage is of no concern in the model, V_k is considered to be the average contact voltage, which is proportional to the total transmitter voltage. Then

$$\eta(V) = [B + (1 - B)(A/V) \tan^{-1}(V/A)]^{-1}, \tag{5}$$

where A is also a constant.

The assumptions upon which the derivation of eq. (4) is based are that the thermal conductivity satisfies

$$\lambda = \lambda_0(1 + \beta\Delta T), \tag{6}$$

and the electrical resistivity satisfies

$$\rho = \rho_0(1 + \epsilon\Delta T)/(1 + \beta\Delta T), \tag{7}$$

where ρ_0, λ_0, ϵ, and β are constants, and ΔT is the change in temperature because of joule heating. Apparently, these assumptions are valid in this case because of the excellent agreement between eq. (5), using the values for A and B, listed subsequently, and the measured data presented by Hufstutler and Kerns[8] for the resistivity of granular carbon contained in a quartz test chamber having a negligible thermal expansion coefficient. This implies that, aside from the effect of chamber expansion, the effect of joule heating on contact resistance is alone sufficient to account for the nonlinearity of the V-I characteristic.

Now, if eqs. (5) and (3) are combined, the expression for the dc contact resistance becomes

$$r_k(r, \theta, \phi) = K\eta(V)[P_a(r, \theta, \phi) + (V^2/R)P_d(r, \theta, \phi)]^{-\gamma}. \tag{8}$$

2.3 Total dc resistance

If $r_k(r, \theta, \phi)$ is the contact resistance and there are n contacts per unit length, then the resistivity of the carbon granules is $r_k(r, \theta, \phi)/n$, where $r_k(r, \theta, \phi)$ is given by eq. (8). Then, for the approximate chamber geometry which has been assumed, the total dc transmitter resistance is

$$R = \int_a^b \frac{dr}{r^2 \int_0^\pi \int_0^\pi [n \sin\phi / r_k(r, \theta, \phi)] d\phi d\theta}. \tag{9}$$

The mean value theorem can be applied to perform the integrations

with the result

$$R = R_0\eta(V)[1 + \alpha V^2/R]^{-\gamma}, \tag{10}$$

where

$$R_0 \equiv K(b - a)/[n\pi^2\bar{r}^2 \sin\bar{\phi}P_g^{\gamma}(\bar{r}, \bar{\theta}, \bar{\phi})], \tag{11}$$

$$\alpha \equiv P_d(\bar{r}, \bar{\theta}, \bar{\phi})/P_g(\bar{r}, \bar{\theta}, \bar{\phi}), \tag{12}$$

and where $\bar{r}$, $\bar{\theta}$, and $\bar{\phi}$ are constants, being the coordinates of some point within the chamber. Note that R_0 is the limiting value of R as V approaches zero, and that $\alpha V^2/R$ is the average ratio of the component of contact force due to thermal expansion of the carbon chamber to the component of contact force due to the weight of the granules themselves. Although R is not expressed as an explicit function of V, a solution to eq. (10) can be obtained using iterative techniques easily implemented on the computer. Values for the parameters R_0, α, A, B, and γ will be determined by fitting eq. (10) simultaneously with equations for the transmitter ac resistance and open circuit output voltage to measured data. The ac resistance and open circuit output voltage will now be considered.

III. SMALL-SIGNAL AC RESISTANCE

Over the range of frequencies of interest for speech transmission, the transmitter ac impedance is purely resistive. However, as is obvious from Fig. 3, the ac resistance is not the slope of the dc V-I characteristic except in the limit as V approaches zero. This is explained by the fact that, because of the large hysteresis effect, the thermal expansion of the dome electrode cannot follow the ac signal, at least not at frequencies above a few hertz. Thus, the difference between the ac resistance, which is not affected by the thermal expansion of the dome electrode, and the slope of the dc V-I characteristic, which is affected, increases as the power dissipated by the transmitter increases.

Because the thermal expansion of the dome electrode has no effect on the ac resistance, the ac resistance is the slope, not of the actual V-I characteristic, but of the V-I curve defined by setting the term equal to zero which accounts for the expansion of the dome electrode. This is the curve defined by

$$I = V/[R_0\eta(V)], \tag{13}$$

where R_0 and $\eta(V)$ are defined by eqs. (11) and (5), respectively. Thus, the ac resistance is given by

$$\begin{aligned} r_{ac} &= R_0\eta^2(V) \Big/ \left[\eta(V) - V\frac{d}{dV}\eta(V)\right] \\ &= R_0(A^2 + V^2)/(A^2 + BV^2). \end{aligned} \tag{14}$$

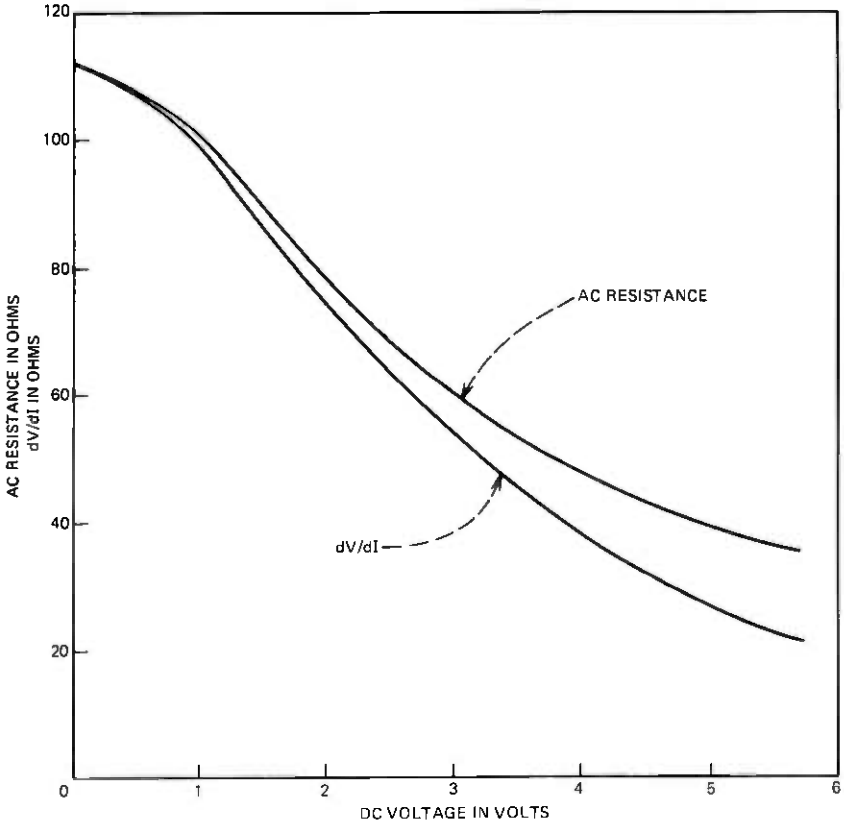


Fig. 3—Alternating-current resistance characteristic and slope of dc V-I characteristic.

It has been assumed that the ac signal level is low enough so that there is no significant joule heating effect due to the ac signal. This assumption is valid for signal levels typical of speech transmission. For higher signal levels, the ac joule heating effect will cause the ac as well as the dc resistance to decrease, as is easily verified experimentally.

At frequencies low enough that the period of the ac signal becomes significant compared to the time constant associated with the thermal response of the dome electrode, the transmitter ac impedance exhibits a reactive component due to the effect of the thermal hysteresis. Figure 4, drawn from a photograph of a storage oscilloscope trace, shows the effect of the thermal hysteresis in the response to a sinusoidal driving voltage at a frequency of 0.2 Hz for four different operating points. The effect becomes greater as the dc bias increases, as would

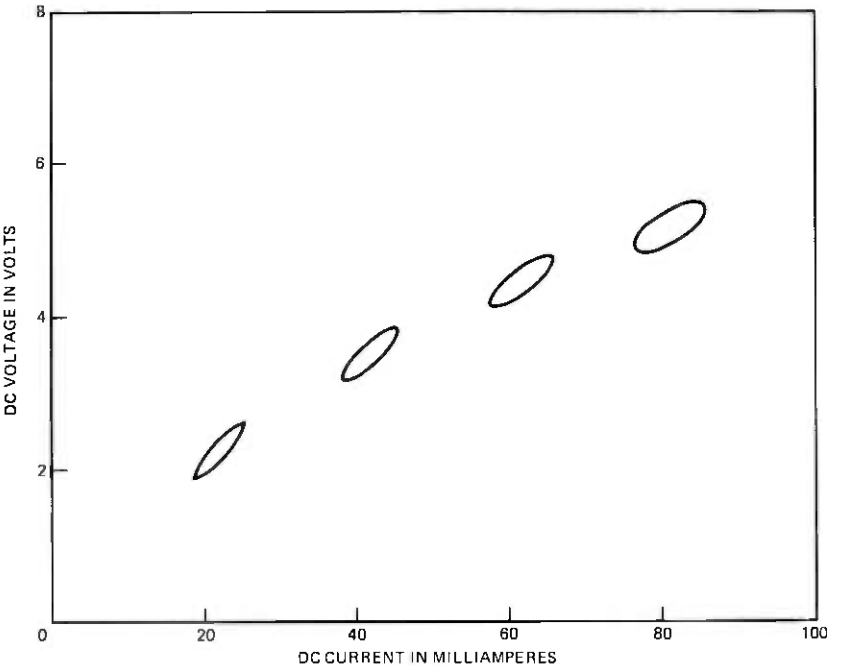


Fig. 4—Thermal hysteresis effect at 0.2 Hz.

be expected since the expansion of the dome electrode is proportional to the power dissipated by the transmitter. An expression for the low-frequency transmitter impedance that accounts for this effect is derived in Appendix B. For frequencies greater than approximately 10 Hz, the reactive component of the transmitter impedance becomes negligible, and the expression derived in Appendix B reduces to the expression given by eq. (14).

IV. OPEN-CIRCUIT OUTPUT VOLTAGE

When the transmitter is acoustically excited, the contact force will vary owing to the effect of the acoustic pressure acting on the diaphragm and dome electrode. The variation in the contact force at the point (r, θ, ϕ) is designated $\Delta P(r, \theta, \phi, t)$. Then the resistance will vary by an amount $\Delta R(t)$ such that

$$R + \Delta R(t) = R_0\eta(V)[1 + \alpha V^2/R + \overline{\Delta P(t)}]^{-\gamma}, \tag{15}$$

where

$$\overline{\Delta P}(t) \equiv \Delta P(\bar{r}, \bar{\theta}, \bar{\phi}, t)/P_g(\bar{r}, \bar{\theta}, \bar{\phi}). \tag{16}$$

The transmitter voltage will change by an amount $\Delta V(t)$ where, if the transmitter current I is held constant,

$$V + \Delta V(t) = I[R + \Delta R(t)]. \tag{17}$$

Then

$$\begin{aligned} \Delta V(t) &= I\Delta R(t) \\ &= V\Delta R(t)/R, \end{aligned} \tag{18}$$

which is the ac open-circuit output voltage. From eq. (15),

$$\Delta R(t)/R = [1 + \overline{\Delta P(t)}/(1 + \alpha V^2/R)]^{-\gamma} - 1, \tag{19}$$

so that

$$\Delta V(t) = V\{[1 + \overline{\Delta P(t)}/(1 + \alpha V^2/R)]^{-\gamma} - 1\}. \tag{20}$$

Although this is a nonlinear relationship, $\overline{\Delta P}(t)$ will be found to be small enough compared to $1 + \alpha V^2/R$ at normal speech levels so that a linear approximation is justified. Thus,

$$\Delta V(t) \approx -\gamma V\overline{\Delta P}(t)/(1 + \alpha V^2/R). \tag{21}$$

Note that eq. (21) is $-\gamma V$ multiplied by the ratio of the dynamic to the static forces acting on the carbon granules.

V. EVALUATION OF MODEL PARAMETERS

The parameters R_0, α, γ, A, B, and $\overline{\Delta P}(t)$ were evaluated using an iterative optimization computer program to fit eqs. (10), (14), and (21) simultaneously to measured dc resistance, ac resistance, and open-circuit output voltage, respectively, the latter being measured at a frequency of 1 kHz with a sound pressure level of 94 dB (re 0.0002 dyn/cm^2). The measurements were performed on a T1 transmitter considered to be a typical unit. The resulting parameter values are listed in Table I. Of course, transmitter characteristics are subject to such factors as aging, temperature, conditioning, orientation, and manufacturing variations. The values of the parameters in eqs. (10), (14), and (21) will vary accordingly.

Table I — Parameter values

Parameter	Value
R_0	111.0
α	0.94
γ	0.43
A	7.12
B	6.65
$\overline{\Delta P}$(rms)	0.20

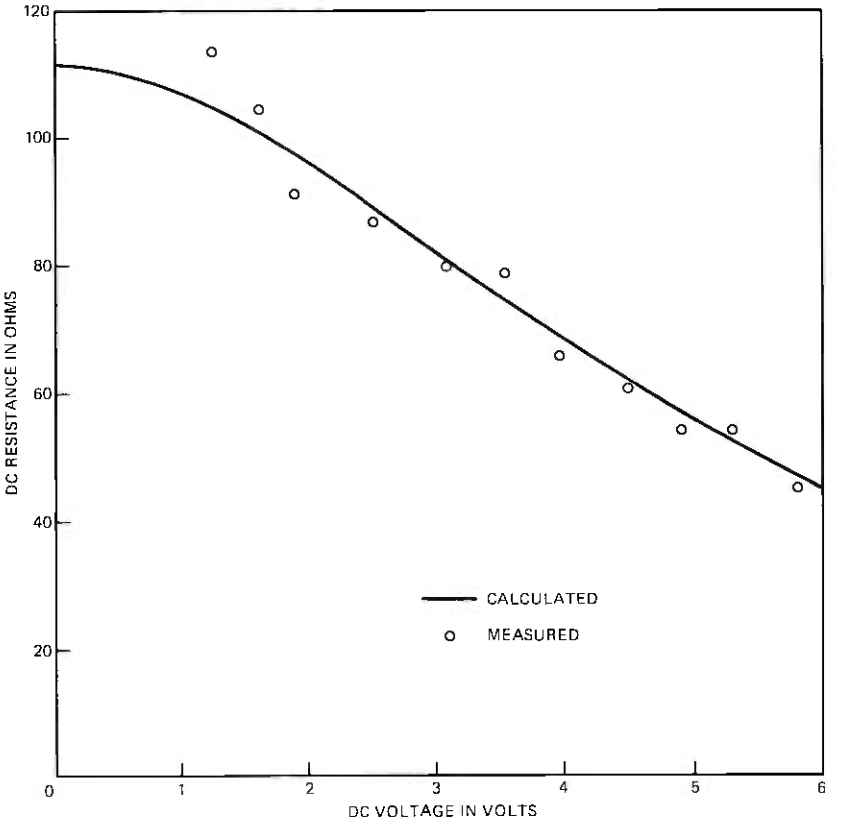


Fig. 5—Direct-current resistance characteristic.

Equations (10), (14), and (21), with the parameter values listed in Table I, are plotted in Figs. 5, 6, and 7, respectively, along with measured data points. The agreement between calculated and measured data is judged to be within the limits of measurement error.

VI. INPUT-OUTPUT AND FREQUENCY RESPONSE CHARACTERISTICS

To complete the model, the input-output and frequency response characteristics of the transmitter must be specified. The input-output characteristic is nonlinear owing to the effect of the acoustic excitation on the compactness of the carbon granules. As the acoustic signal level increases, the carbon granules are agitated into a less compact state and the mechanical impedance of the granules decreases. Therefore, the transmitter efficiency increases as the acoustic signal level increases, resulting in an input-output characteristic having a greater-

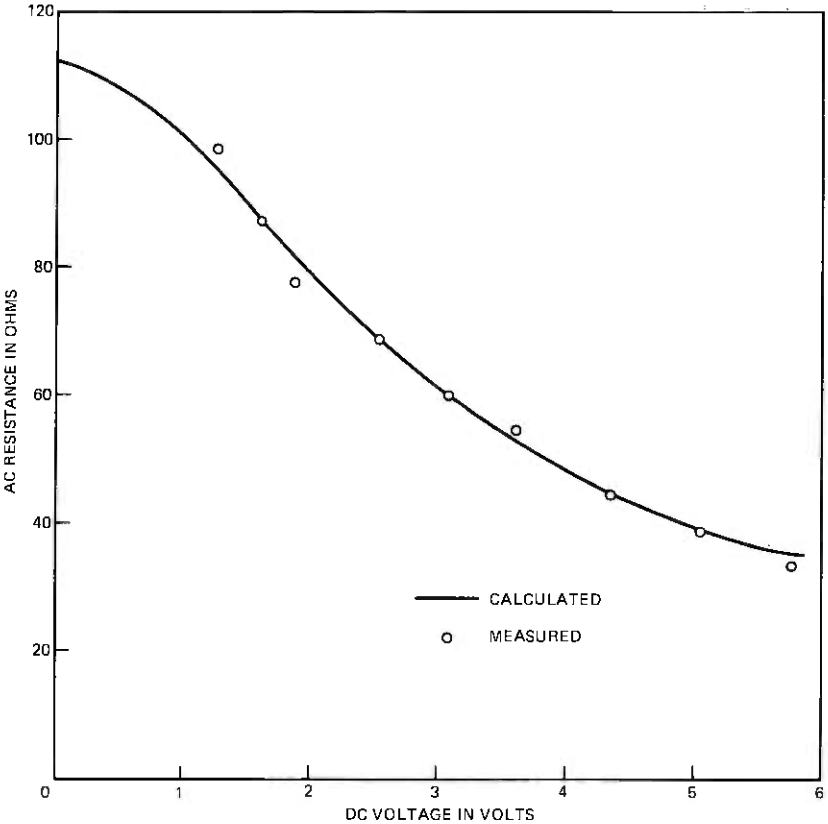


Fig. 6—Alternating-current resistance characteristic.

than-unity slope. Because of this effect, the transmitter is able to discriminate against distant sounds, and thereby reduce interference resulting from background noise. However, the weak components in a composite signal such as speech are not discriminated against, since the compactness of the carbon granules, which is controlled by the strong components, is the same for all components of the signal. Therefore, the nonlinearity of the input-output characteristic does not affect the components of individual speech sounds, and the compactness of the carbon granules varies only as the overall energy content of the speech signal varies.

As discussed by Bryant,[9] the frequency response of the transmitter is related to the nonlinearity of the input-output characteristic, since it also depends on the mechanical impedance of the carbon granules. This implies that the frequency response depends on the nature of the input signal. The response to a swept frequency sinusoidal signal is

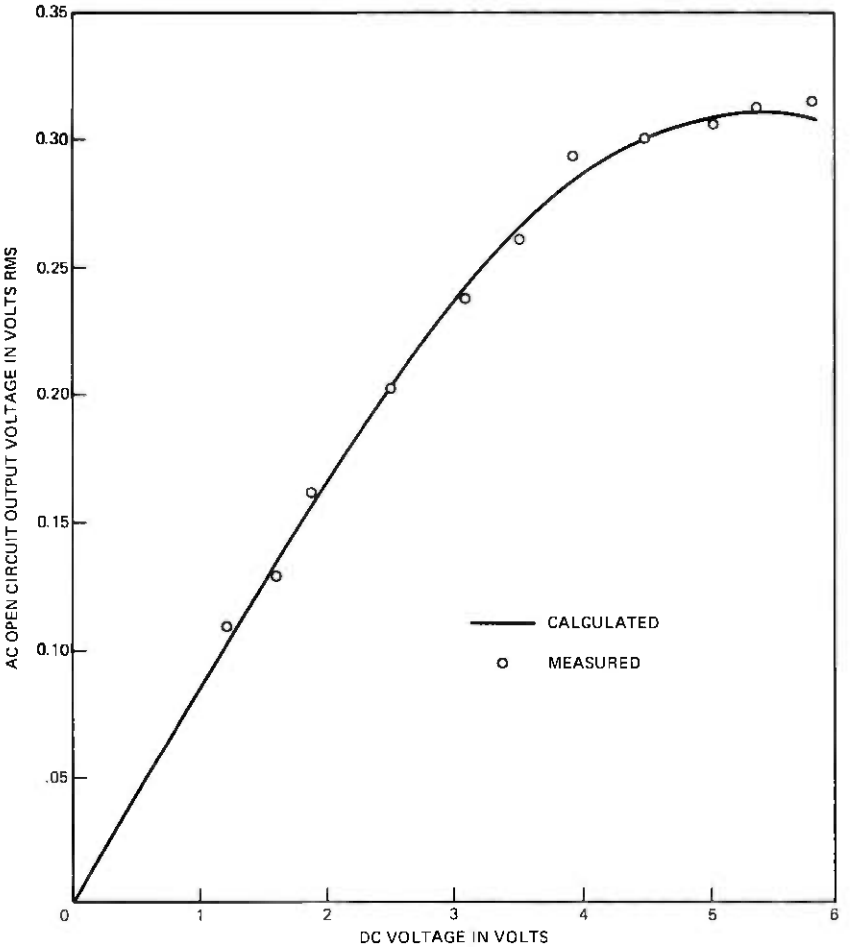


Fig. 7—Alternating-current open-circuit output voltage characteristic at 1 kHz for a 94-dB (re 0.0002 dyn/cm²) acoustic input level.

somewhat different from the response to a speech signal. (For further discussion, see Ref. 9.)

For the purposes of the model, a continuous, random, speech input signal is assumed. Accordingly, the nonlinearity of the input-output characteristic is ignored, and the frequency response is measured as suggested by Bryant. The response is shown in Fig. 8 plotted relative to the 1-kHz output level. This is the response of the transmitter structure itself to the acoustic input signal, or, in the model, the response of the input filter. In the computer program, the filter response was stored as a table of values at discrete values of frequency.

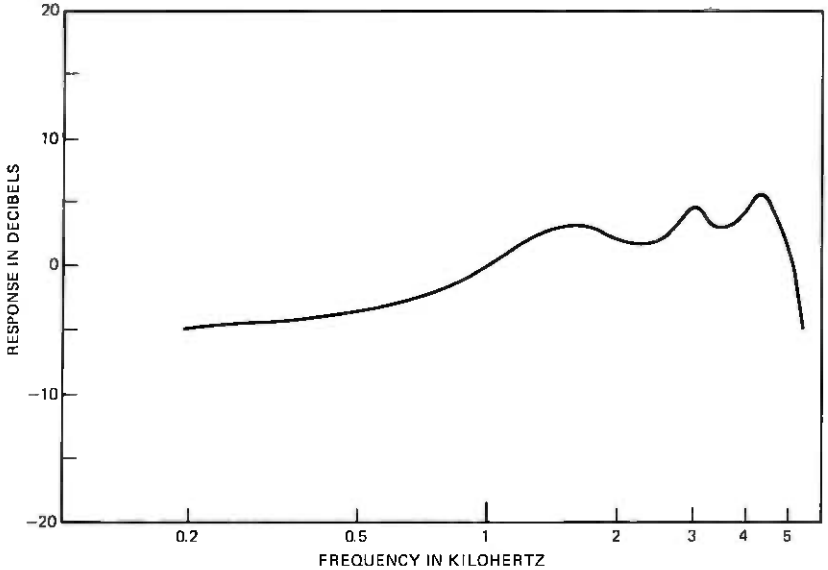


Fig. 8—Transmitter gain (relative) vs frequency.

VII. CONCLUSION

A carbon transmitter model has been presented. The physical theory upon which the model is based is supported by the close agreement between the calculated characteristics of the model and the measured characteristics of an actual device. Thus, it can be concluded that the nonlinearity of the dc V-I characteristic is due primarily to the effect of joule heating on contact resistance and to the effect of the thermal expansion of the dome electrode due to joule heating. The effect of electrostatic forces is negligible. Furthermore, the difference between the ac resistance and the slope of the dc V-I characteristic is due to the hysteresis associated with the thermal expansion of the dome electrode. Finally, the relative resistance change due to the acoustic excitation of the transmitter decreases as the dc voltage increases due also to the effect of the thermal expansion of the dome electrode.

VIII. ACKNOWLEDGMENT

The author would like to thank F. W. Hewlett, Jr., with whom he collaborated during the initial phase of the transmitter modeling project.

APPENDIX A

In this appendix, the question is considered of whether electrostatic forces between carbon granules have a significant effect on transmitter

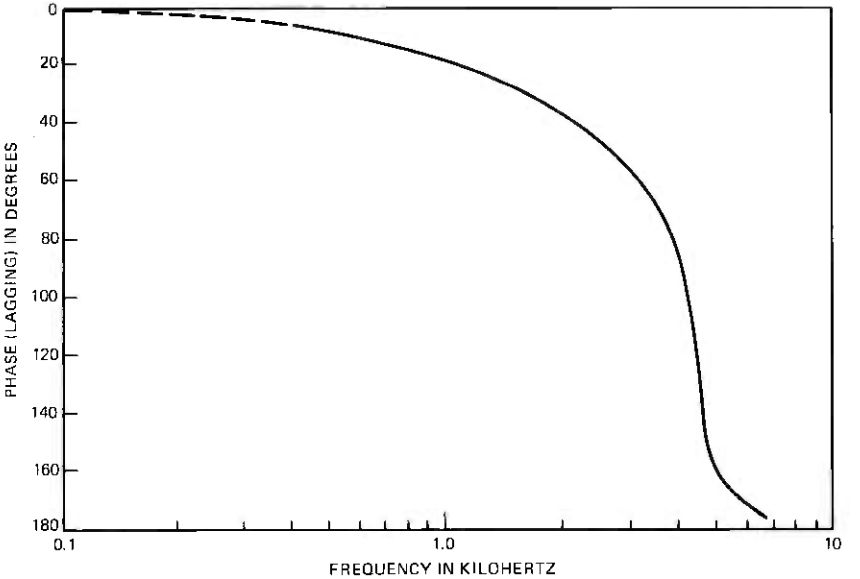


Fig. 9—Phase of diaphragm displacement resulting when transmitter output is driven by sinusoidal ac voltage.

characteristics. Two experiments, the results of which indicate that they do not, are described. In both experiments, the transmitter output is driven with a sinusoidal voltage, i.e., as if the transmitter were a receiver. It is observed that, if the transmitter is driven hard enough, an audible signal is generated. Approximately 2-V rms is required for the signal to be audible at 1 kHz. This phenomenon could be the result of attractive forces between carbon granules owing to electric fields or the result of thermal expansion, presumably of the carbon granules, since the thermal inertia of the dome electrode and diaphragm is too large for their thermal response to follow the instantaneous voltage at frequencies above a few hertz. The results of the two experiments which are now described indicate that the forces are due to thermal expansion of the carbon granules. This effect is insignificant compared to the predominant thermal effects that are accounted for in the model.

In the first experiment, the phase of the dome electrode displacement was measured. An outward displacement, in phase with the square of the driving voltage at frequencies far enough below resonance so that the mass of the system can be ignored, would indicate that the force was due to thermal expansion, while an inward displacement would indicate that the force was due to electric fields. The displacement was measured using an optical proximity detector. A small mirror was mounted on the dome electrode to provide a flat reflecting surface

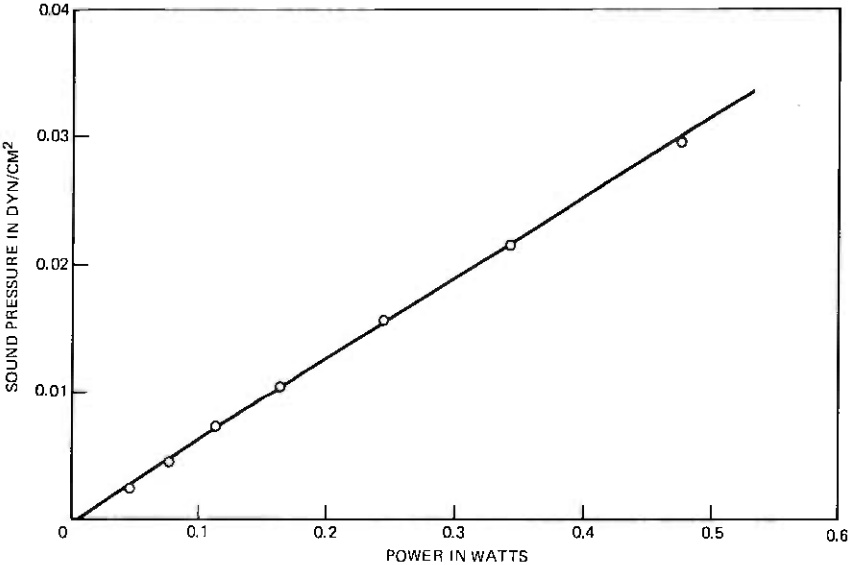


Fig. 10—Output sound pressure—input power characteristic resulting when transmitter output is driven by sinusoidal ac voltage.

for the detector. This was necessary to obtain measurable detector output, the displacement being very small. The phase of the outward displacement relative to that of the square of the driving voltage is plotted as a function of frequency in Fig. 9. The phase angle approaches zero at low frequencies, indicating that the force acting on

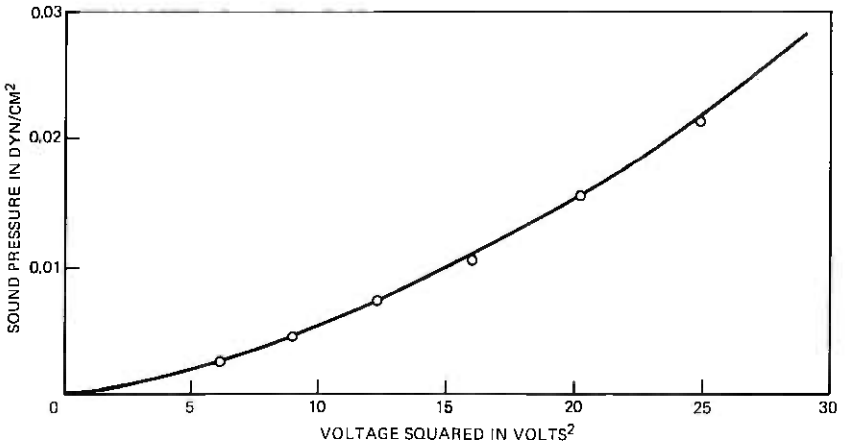


Fig. 11—Output sound pressure—input voltage squared characteristic when transmitter output is driven by sinusoidal ac voltage.

the dome electrode is due to the thermal expansion of the carbon granules.

Since this is the case, the acoustic sound pressure level should be proportional to the power dissipated by the transmitter rather than the square of the voltage, as would be the case if the forces were due to electric fields. (Power is not proportional to the square of the voltage, since the resistance is a function of voltage.) This was verified by the second experiment, the results of which are plotted in Figs. 10 and 11.

APPENDIX B

In this appendix, an expression is derived for the low-frequency impedance of the transmitter, taking into account the effect of the thermal expansion of the carbon chamber.

If the transmitter current is changed abruptly by an incremental amount ΔI, the displacement of the dome electrode due to the additional power dissipation will lag behind the change in current due to the thermal hysteresis. According to Fritsch's analysis,[6] the transient can be expressed as an infinite sum of decaying exponentials. Thus,

$$\Delta V(t) = (m + \sum_{i=1}^{\infty} k_i e^{-t/\tau_i})\Delta I u(t), \tag{22}$$

where m is the slope of the dc V-I characteristic and where k_i and τ_i are constants. Because the initial change in voltage must be $r_{ac}\Delta I$,

$$\sum_{i=1}^{\infty} k_i = r_{ac} - m. \tag{23}$$

In the frequency domain,

$$\Delta V(s) = [m + \sum_{i=1}^{\infty} k_i s/(s + 1/\tau_i)]\Delta I/s, \tag{24}$$

from which

$$Z(s) = m + \sum_{i=1}^{\infty} k_i s/(s + 1/\tau_i). \tag{25}$$

As pointed out by Fritsch,[6] the infinite series solution converges too slowly for practical evaluation. A practical expression for ac impedance can be obtained by assuming a single time-constant approximation for the transient response. Then

$$Z(s) \approx m + ks/(s + 1/\tau), \tag{26}$$

where τ is the effective time constant and

$$k = r_{ac} - m. \tag{27}$$

If eq. (27) is substituted into eq. (26), then

$$Z(s) \approx r_{ac}(s + m/r_{ac}\tau)/(s + 1/\tau). \tag{28}$$

REFERENCES

1. A. H. Inglis and W. L. Tuffnell, "An Improved Telephone Set," B.S.T.J., *30*, No. 2 (April 1951), pp. 239–270.
2. IEEE Standard 269-1971, "IEEE Standard Method for Measuring Transmission Performance of Telephone Sets."
3. F. S. Goucher, "The Carbon Microphone: An Account of Some Researches Bearing on Its Action," B.S.T.J., *13*, No. 2 (April 1934), pp. 163–194.
4. H. Mol, "Theorie van de Koolmicrofoon," PTT-Bedriff, *3*, No. 4 (June 1951), pp. 128–134.
5. R. Holm, *Electric Contacts—Theory and Applications*, 4th ed., New York: Springer-Verlag, 1967.
6. C. A. Fritsch, "Effects Associated With the Thermal Response of the T1 Telephone Transmitter," B.S.T.J., *47*, No. 8 (October 1968), pp. 1615–1636.
7. R. Joscheck, "Electrishe and Mechanishe Eigenschaften des Kohlengrieszes von Mikrophonen," Wiss. Veroff, a.d. Siemens-Werken, *16*, No. 1 (1937), pp. 105–119.
8. M. C. Huffstutler Jr., and B. T. Kerns, unpublished work.
9. H. W. Bryant, "Comparable Tests on Linear- and Carbon-Type Microphones," J. Acoust. Soc. Amer., *53*, No. 3 (March 1973), pp. 695–698.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 54, No. 7, September 1975
Printed in U.S.A.

Some Far-Field Studies of an Offset Launcher

By M. J. GANS and R. A. SEMPLAK

(Manuscript received January 24, 1975)

An offset paraboloidal reflector illuminated by a balanced feed horn constitutes an efficient launcher for coupling microwaves into quasi-optical beams. Measurements on a launcher with low blockage show low cross polarization. The amplitude, phase, and polarization characteristics are predicted by two gaussian beam modes, and the resulting formulas are found to agree well with measurements at 19 and 28 GHz. For example, with increasing offset angles, the ratio of the maximum cross-polarized signal in the radiation pattern to the on-axis co-polarized signal is observed to vary from −44 to −37 dB, within 1 dB of the predicted variation.

I. INTRODUCTION

At millimeter wavelengths, normal waveguide losses become too large in many applications; for example, long lengths of waveguide are required in satellite earth stations between the transceiver and the reflector antenna focus. To reduce these losses, one may use quasi-optical beams that employ reflectors or lenses for refocussing at various intervals, thereby confining the beam within a geometric tube with no (lossy) guiding walls. To couple the circuit components to these beams, it is desirable to provide a beam launcher that has quasi-gaussian amplitude over the aperture, low loss, good polarization purity, and high return loss.

Offset reflectors provide high return loss, i.e., they are well matched, because the radiation field of the illuminated aperture bypasses and, therefore, does not reenter the feed horn.[1] If the reflector is made large so that the level of the feed-horn illumination at the edge of the reflector is low, spillover and diffraction losses are small. Also, since the feed horn does not block the aperture of the reflector, blockage losses are negligible, and the radiation patterns are unaffected. But cross polarization can be serious with offset reflectors, as pointed out in Ref. 2; however, by choosing a small offset angle the cross polarization can be held to acceptable levels throughout the beam.

Here, the far-field properties of an offset paraboloidal reflector (Fig. 1) are investigated. Section II describes the radiation charac-

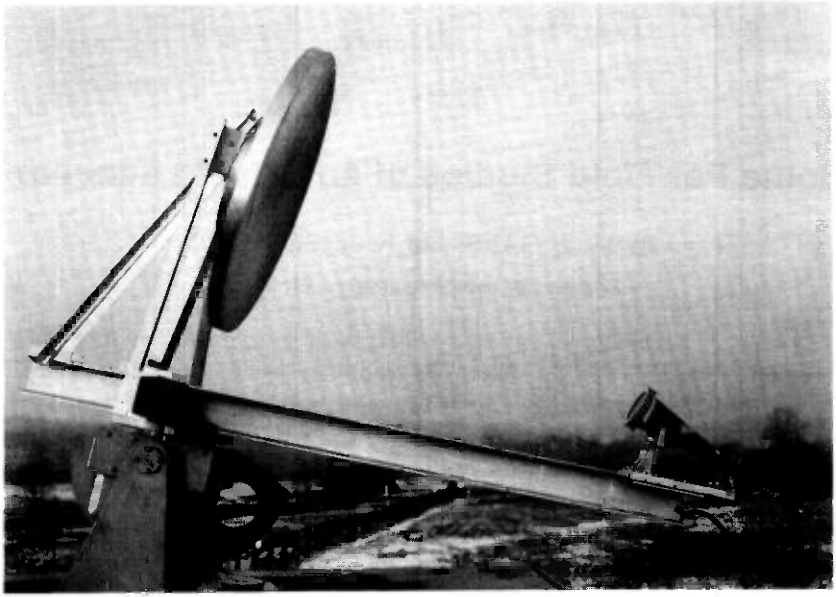


Fig. 1—Offset launcher. The 76-cm (30-in.) diameter reflector is a numerically machined section of a paraboloid. Reflector focal length is 115.7 cm and it is fed, in this instance, by a 28.5-GHz dual-mode horn with a polarizer.

teristics of the dual-mode feed horns, the experimental setups, and the measurements of the far-field properties of the complete launcher. Section III provides the theoretical formulas showing that the radiation performance of offset launchers can be well characterized in terms of gaussian modes.

Specific applications of this type of launcher are in feeding Cassegrainian antennas of the type discussed in Ref. 1, and in launching and collecting beams on Hertzian cable transmission lines.[3]

II. MEASUREMENTS

2.1 Dual-mode feed horns

The dual-mode feed horn designed to feed the offset launcher of Fig. 1 is shown in Fig. 2a. The input section generates the TE_{11} and TM_{11} modes in a circular waveguide by means of a conical step in the waveguide.[4] This section slides in the input waveguide so that the length from the step to the horn aperture, the "drift space," may be adjusted to co-phase the TE_{11} and TM_{11} modes to provide zero current at the edge of the aperture (minimizing the side lobes and symmetrizing the pattern). The small horn taper angle of 7.121 degrees was chosen to prevent disturbing the TE_{11} and TM_{11} modes, and to provide a

FREQUENCY, F IN GHz	DIAMETER, D IN cm	LENGTH, L IN cm
19.04	12.70	59.06
28.56	8.38	39.37

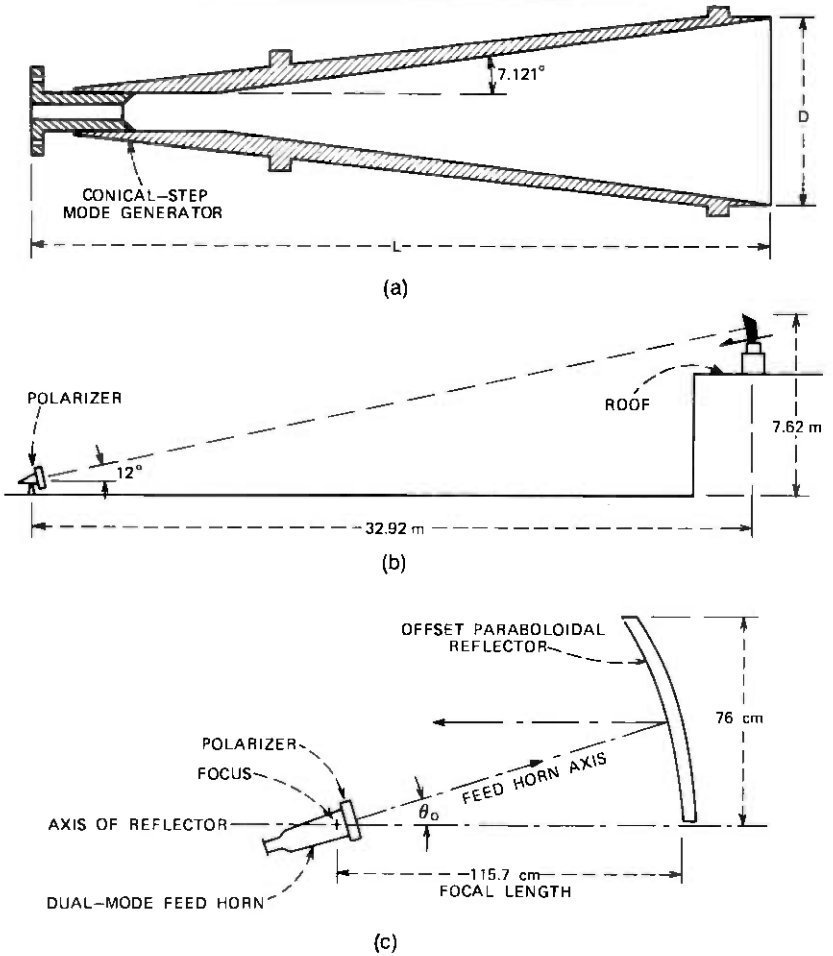


Fig. 2—(a) Cross-sectional view of the dual-mode feed-horn design used for the offset launcher of Fig. 1. (b) Profile of measuring range. (c) Launcher schematic.

small aperture phase error [$\epsilon = (2\pi/\lambda)(D^2/8L) = \pi/2$ radians] without making the horn too long.

The azimuth radiation patterns of 28.5-GHz dual-mode feed horns for horizontal, vertical, and 45-degree polarizations, along with the associated cross polarizations for each case, as measured in an anechoic chamber, are shown in Fig. 3. Figure 3a shows the results obtained

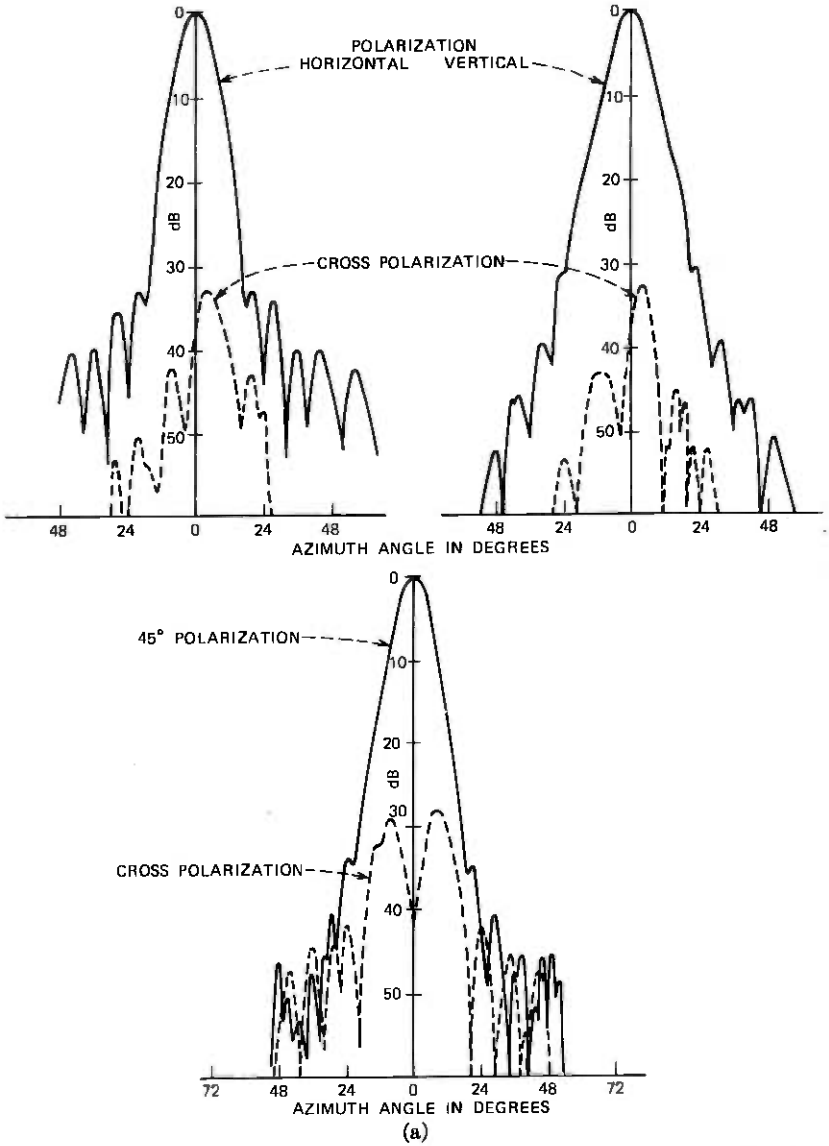
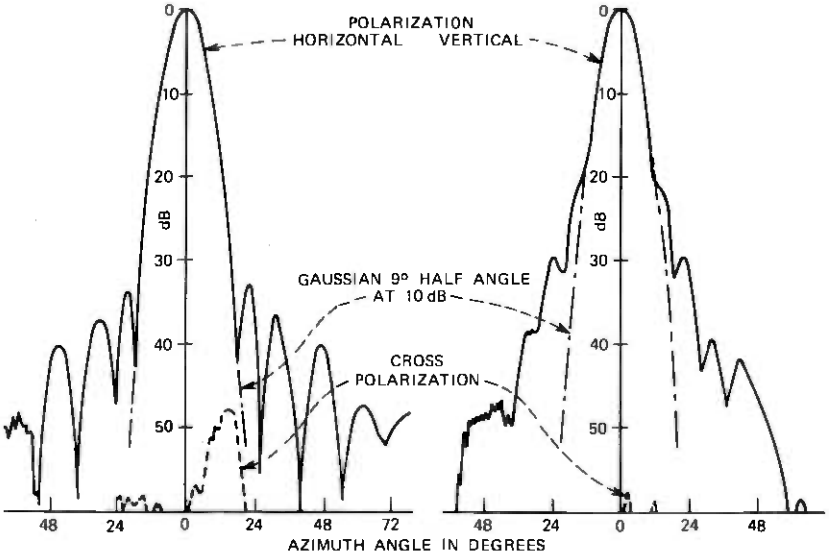
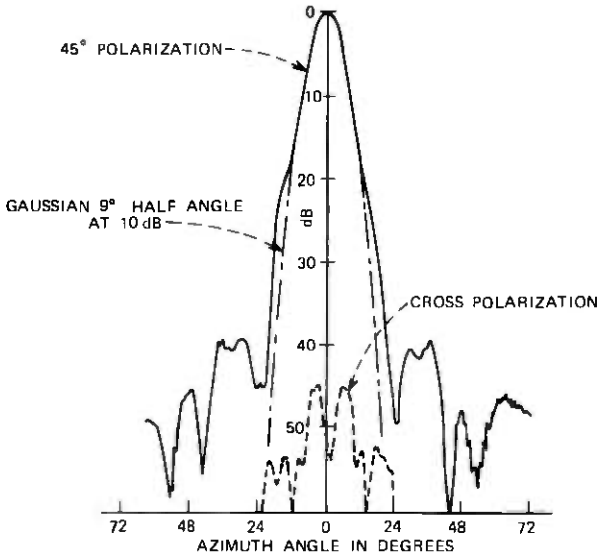


Fig. 3—The 28.5-GHz dual-mode feed-horn radiation patterns for the principal and 45-degree polarizations. Included are the corresponding cross polarizations: (a) without polarizer on dual-mode feed; (b) with polarizer on dual-mode feed.

without a polarizer on a horn. Note that the cross polarization for the 45-degree polarization condition has peaks of about −28 dB at angles of about ±8 degrees. (This behavior of dual-mode horns is predictable; superior cross-polarization performance is obtainable from hybrid-





(b)

Fig. 3 (continued).

mode horns.) One can also see that, even though the quarter-wavelength aperture error fills in the first null in the feed-horn pattern, a shoulder appears at about −20 dB relative to the on-axis value for the case of horizontal polarization. At least 30 degrees of the feed pattern illuminates the reflector of the launcher (resulting in about

a 20-dB taper at the edge of the reflector in the principal polarization). Thus, the cross-polarization maxima in Fig. 3a also illuminate the reflector. However, as shown in Fig. 3b, when a grid polarizer is used on the dual-mode feed horn, the cross polarization for the 45-degree condition is reduced to a very acceptable −45-dB level. Also, with a polarizer on the horn, cross polarization for polarization in the principal planes is essentially nonexistent. Since one of the main purposes of the experiment is to measure the cross polarization generated by the offset reflector, as discussed in the next section, the feed horn per se must therefore be devoid of cross-polarized components. For that reason, the feed horn was equipped with a polarizer (patterns of Fig. 3b) for all ensuing measurements.

The dot-dash curves in Fig. 3b show that the feed-horn pattern is well approximated down to about −20 dB by a gaussian beam [eqs. (8) and (9)]. The 10-dB half angle, θ_c, used in the gaussian beam approximation, is 9 degrees.

The radiation patterns for the 19-GHz dual-mode feed horn are essentially the same as in Fig. 3. Significant cross-polarization levels were also observed at this frequency, but, using the polarizer, the cross polarization is reduced to a very acceptable level (≤ -50 dB).

2.2 Antenna measuring range

A profile of the antenna range used for measuring the offset launcher is shown in Fig. 2b. To determine the cross-polarization characteristics at the range, a gently tapered pyramidal horn with a 15- by 15-cm aperture equipped with a wire grid polarizer (to eliminate any cross polarization caused by the horn itself) was used as the source. A standard gain horn with a similar polarizer was used as the receiver on the antenna azimuth-elevation positioner. These measurements show the cross polarization introduced by the range to be very small; in the range of interest, i.e., within ±3 degrees of the axial direction, it is of the order −47 dB.

2.3 Offset launcher measurements

Figure 1 is a photograph of one of the offset reflectors, along with its supporting structure. The reflector is illuminated by a 28.5-GHz dual-mode feed fitted with an etched grid polarizer as discussed in Section 2.1.

At 28.5 GHz, the far-field patterns in the principal and 45-degree polarizations, along with the associated cross polarizations for the offset launcher fed by a dual-mode horn with polarizer, are as shown in Fig. 4. Note the shoulders rather than sidelobe structure for the co-polar patterns. Although the shoulders in the launcher pattern are

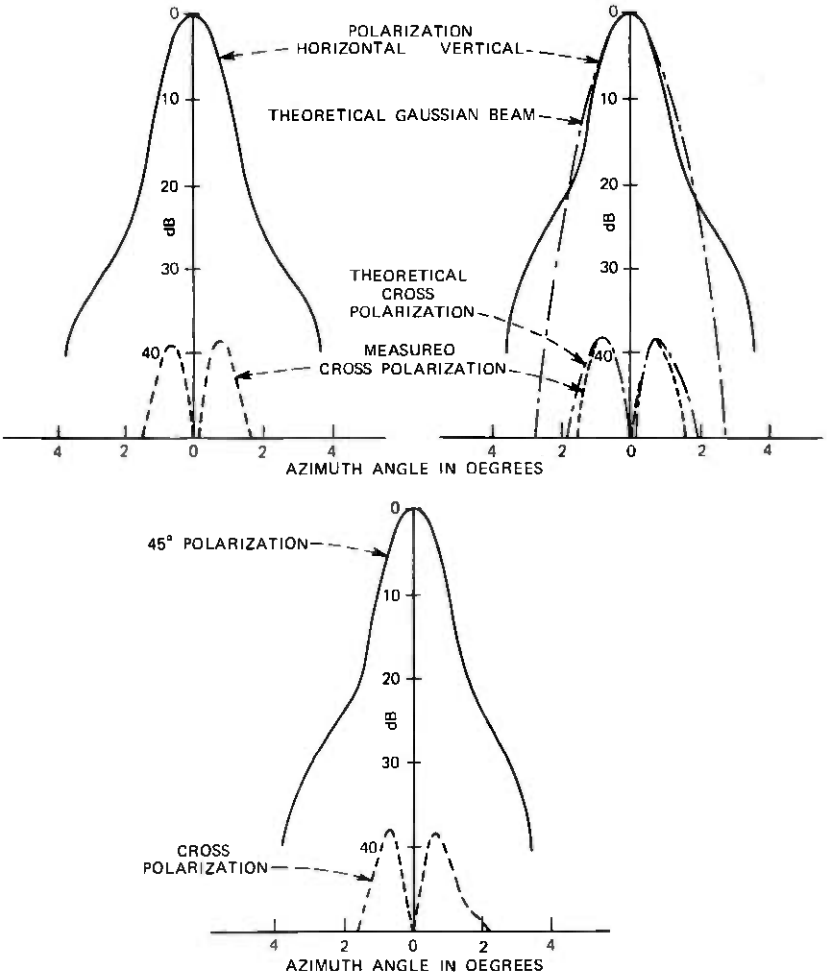


Fig. 4—Far-field radiation patterns of the principal and 45-degree planes for the offset launcher at 28.5 GHz with polarizer on horn. Theoretical (gaussian-beam) calculations are also included.

also apparent in the feed-horn patterns (compare the vertical polarization patterns of Figs. 3b and 4), this is not always the case. For example, the launcher pattern for horizontal polarization exhibits prominent shoulders (Fig. 4), whereas the corresponding feed-horn pattern (Fig. 3b) does not. Furthermore, the shoulders in the launcher patterns are at about the −24 dB level, while those in the feed-horn pattern are at about −20 dB. Most likely, the shoulders in the launcher patterns are due to the phase errors in the illumination which are caused by the finite taper length of the conical feed horn.

At 19.04 GHz, the co-polar and cross-polarization radiation patterns for the offset launcher are as shown in Fig. 5. Excellent symmetry is observed in the cross-polarization patterns even at this −40-dB level. Both Figs. 4 and 5 include the theoretical curves discussed in the next section.

One can see from the configuration of Fig. 1 that there is a possibility of a small amount of blockage by the feed and its mount (and subsequent cross-polarization effects) when the launcher is scanned upward in elevation. To examine this, a set of azimuth scans for various elevation settings was made at both 19.05 GHz and 28.5

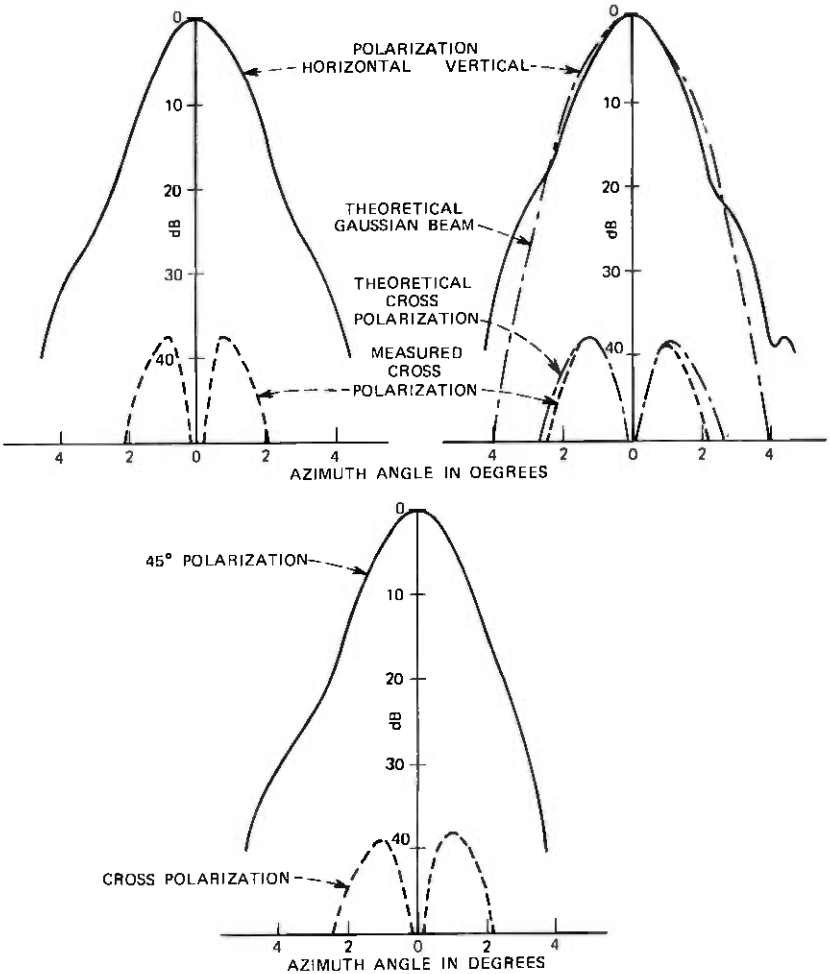


Fig. 5—Far-field radiation patterns of the principal and 45-degree planes for the offset launcher at 19.05 GHz. Dual-mode feed is equipped with polarizer on horn. Theoretical (gaussian-beam) calculations are also included.

GHz; the increase in cross polarization due to blockage was found to be negligible in both cases.

One may with justification raise the question as to why the cross polarization of the offset launcher, shown in Figs. 4 and 5, evidences values of the order of 38 dB even with a feed that has negligible cross polarization. There is an inherent depolarization introduced by an offset reflection surface,[2] which increases if θ_0, the angle between the feed axis and the reflector axis, is increased. Figure 6 shows the experimental results obtained by varying θ_0 between 12 and 26 degrees; the ratio of the maximum cross-polarized signal in the radiation pattern to the on-axis co-polarized signal correspondingly varies from −44 to −37 dB at both frequencies. Note that the ordinate of Fig. 6 is the average of the *peaks* of cross polarization obtained from an azimuth scan of the launcher; they should not be misinterpreted as on-axis values which, of course, are much lower. In the following section, we show that calculations based on gaussian-mode theory provide good agreement with the measured data; the theoretical result is shown by a solid curve in Fig. 6.

III. THEORY

3.1 Cross polarization in the aperture

An approximate method for computing the cross polarization due to the offset angle θ_0 consists of applying geometric optics to compute

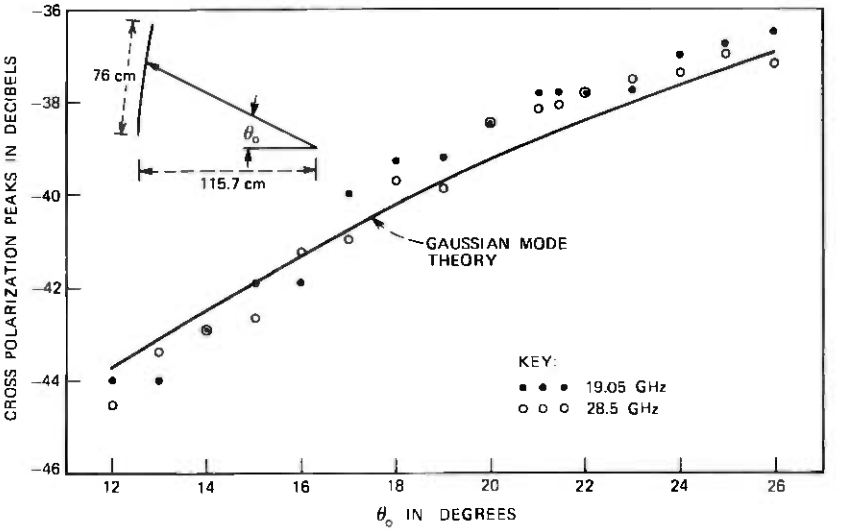


Fig. 6—Plot of cross-polarization peaks at 19.05 and 28.5 GHz introduced by the reflector itself as a function of offset angle, θ_0. Data are obtained by scanning the launcher in azimuth. Gaussian-mode theory is shown by the solid line. Incident polarization is horizontal.

the reflected field in the aperture from that radiated by the feed horn. The aperture field is then decomposed into two gaussian-beam modes to predict the far field of the offset launcher; this is a logical procedure because dual-mode horns produce an illumination that is approximated by a gaussian beam (see Fig. 3b).

The geometry of the offset launcher is shown in Fig. 7a. The dual-mode feed horn approximately provides "balanced feed" polarization[4] with respect to the coordinates (x', y', and z') aligned with the horn axis which is tilted at angle θ_0 from the reflector axis. The field radiated by the feed horn is given by

$$\mathbf{E}_{f'} = (\hat{\theta}' \sin \phi' + \hat{\phi}' \cos \phi') \frac{f'(\theta', \phi')}{r'}, \tag{1}$$

where $f'(\theta', \phi')$ is an arbitrary function of θ' and ϕ', and r', θ', and ϕ' are the usual spherical coordinates associated with the feed (see Fig. 7b). The caret indicates a unit vector. In eq. (1), the expressions corresponding to "vertical" polarization are used; identical results are obtained for "horizontal" polarization. ["Vertical" and "horizontal" are used in the sense that the polarization of the field in the aperture of an axisymmetric paraboloidal reflector coaxial with the feed-horn axis would be vertical or horizontal when the feed-horn polarization is as given in eq. (1).][5]

The axis of the paraboloidal reflector shown in Fig. 7a is co-linear with the $\hat{z}$ axis. An aperture field with no cross polarization would, therefore, result if the feed illumination were given by

$$\mathbf{E}_f = (\hat{\theta} \sin \phi + \hat{\phi} \cos \phi) \frac{f(\theta, \phi)}{r}, \tag{2}$$

where r, θ, and ϕ are the usual spherical coordinates associated with the x, y, z coordinates of Fig. 7a, i.e., the coordinates of a feed horn whose axis is aligned with the reflector axis. Theoretically, it is possible to hypothesize an asymmetric "balanced" feed whose axis is aligned with the reflector axis, giving the polarization of eq. (2), but whose amplitude distribution is offset to illuminate the reflector as would the amplitude distribution of a tilted symmetric "balanced" feed. Simple means (excluding multiple reflectors, etc.) are not known for the construction of such an asymmetric "balanced" feed. Therefore, in applications, one must approximate an asymmetric "balanced" feed horn with a tilted "balanced" feed horn; the cross polarization thereby introduced is calculated below.

Using geometric optics, we assert that if the polarization of a ray incident on the reflector from the feed is rotated by a given angle

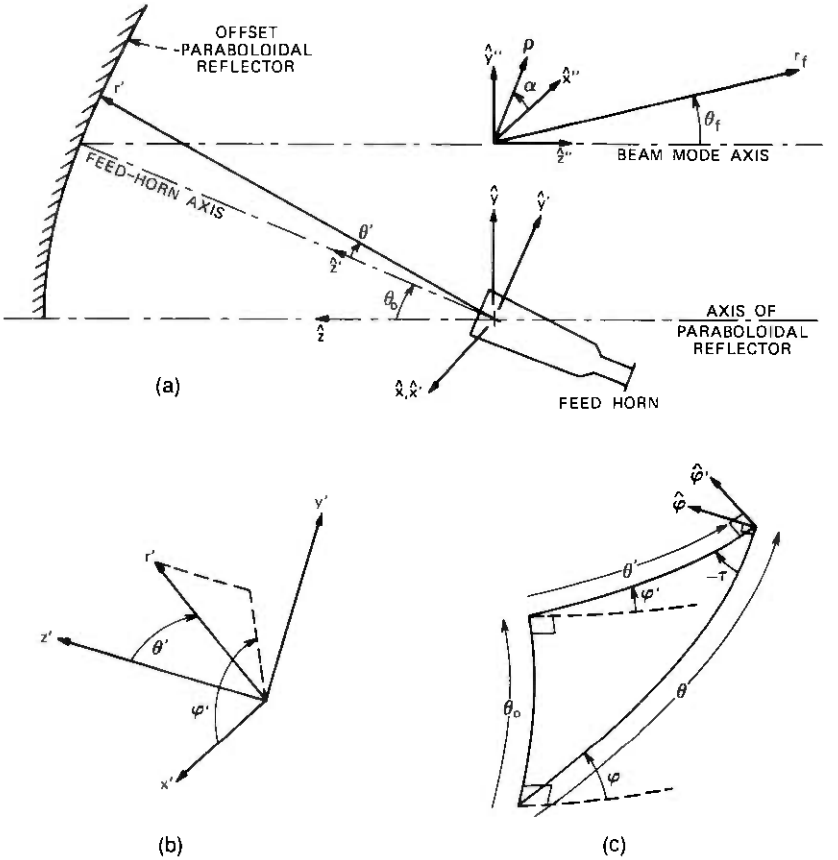


Fig. 7—(a) Launcher geometry. (b) Spherical coordinates. (c) Spherical triangle.

around the ray vector, then the polarization of the field of the corresponding ray in the reflector aperture is rotated by that same angle. By geometric optics, the intensity of the field in the aperture along a ray is the same as the intensity of the field incident on the reflector from the feed at its focus. Thus, the cross polarization in the aperture relative to the peak in-line polarization in the aperture can be computed by projecting the field of a tilted "balanced" feed horn, incident at any point on the reflector, onto the cross-polarized asymmetric "balanced" field at the same point and dividing by the peak in-line field of the tilted "balanced" feed horn; i.e.,

$$C(\theta', \phi') = \frac{\mathbf{E}_{f'} \cdot [\hat{\theta}(-\cos\phi) + \hat{\phi}\sin\phi]}{\mathbf{E}_{f'} \cdot [\hat{\theta}\sin\phi + \hat{\phi}\cos\phi]|_{\text{peak}}}. \tag{3}$$

Performing the scalar products indicated in eq. (3) yields

$$C(\theta', \phi') = \sin(\phi - \phi' + \tau)\frac{f'(\theta', \phi')}{r'} \cdot \left[\cos(\phi - \phi' + \tau)\frac{f'(\theta', \phi')}{r'}\bigg|_{\text{peak}}\right]^{-1}, \quad (4)$$

where τ is the angle between the primed and unprimed spherical angle coordinates shown in Fig. 7c. The angle, $\phi' - \phi - \tau$, is equal to the area, A (called spherical excess), of the spherical triangle shown in Fig. 7c and is related to the offset angle, θ_0, and the primed (feed horn) coordinates by the formula[6]

$$A = \phi' - \phi - \tau = 2 \arctan\left[\frac{\cos\left(\frac{\theta_0 - \theta'}{2}\right)}{\left(\frac{\theta_0 + \theta'}{2}\right)} \cos\left(\frac{\phi'}{2} + \frac{\pi}{4}\right)\right] + \phi' - \frac{\pi}{2}. \quad (5)$$

The feed-horn amplitude pattern, f', is approximately uniform in ϕ' and maximum on axis (as for a corrugated or a dual-mode feed horn), and r', the distance from the feed to the reflector, is relatively constant over the region in which the cross-polarized field is a significant fraction of the peak in-line field. In this case, the cross-polarization amplitude can be approximated by

$$C(\theta', \phi') \doteq -\sin A \frac{f'(\theta')}{f'(0)}. \quad (6)$$

For a given θ', the ϕ' which maximizes A and the cross polarization is

$$\phi_m' = \arcsin\left(\tan\frac{\theta_0}{2}\tan\frac{\theta'}{2}\right). \quad (7)$$

For offset angles, θ_0, less than or equal to 90 degrees and θ' less than 30 degrees, which covers the case of interest, ϕ_m' is less than 16 degrees. This leads us to approximate ϕ_m' by 0 degrees, which results in a particularly simple yet accurate formula for the peak cross-polarization amplitude ratio, **C**, as will be shown. **C** is the ratio of the maximum cross-polarized amplitude to the maximum in-line polarized amplitude.

For dual-mode and corrugated horns, the pattern is approximated by a gaussian beam,[7]

$$f'(\theta') = e^{-a\theta'^2}, \quad (8)$$

where a is related to the 10-dB half-angle beamwidth, θ_c, by

$$a = \frac{\ln 10}{2\theta_c^2}. \quad (9)$$

Approximating ϕ_m' by 0 degrees, we have the following relation[6] for a right spherical triangle (Fig. 7c):

$$\sin A = \frac{\sin \theta_0 \sin \theta'}{1 + \cos \theta_0 \cos \theta'}. \tag{10}$$

Since we are interested in small θ',

$$\sin A = \frac{\theta' \sin \theta_0}{1 + \cos \theta_0}, \tag{11}$$

and eq. (6) becomes

$$C(\theta', \phi_m) \doteq -\frac{\sin \theta_0}{1 + \cos \theta_0} \theta' e^{-a\theta'^2}. \tag{12}$$

By differentiating eq. (12) to locate the angle which maximizes $C(\theta', \phi_m')$, we find

$$\theta_m' \doteq \sqrt{\frac{1}{2a}} = \theta_c/\sqrt{\ln 10}, \tag{13}$$

and

$$\mathbf{C}_a \doteq -\frac{\theta_c \tan \frac{\theta_0}{2}}{\sqrt{e \ln 10}}. \tag{14}$$

As mentioned above, **C** is the ratio of the maximum cross-polarized amplitude to the maximum in-line polarized amplitude. The subscript a indicates that the ratio is of the maxima found in the reflector aperture. If we denote by θ_T the half-angle of the gaussian beam approximation to the feed-horn pattern, where the power is T dB below that on axis, then eq. (14) becomes

$$\mathbf{C}_a \doteq -\theta_T \tan \frac{\theta_0}{2} \sqrt{\frac{10}{eT \ln 10}}. \tag{15}$$

By comparing eq. (14) with the exact formula for gaussian beams, eqs. (6), (7), and (8), eq. (14) is found accurate to within 0.1 dB for all offset angles, θ_0, less than or equal to 90 degrees and all 10-dB half-beamwidth angles, θ_c, less than or equal to 45 degrees. The maximum value of C in the exact formula was found by trial and error with the aid of a calculator.

The above calculations give the cross polarization in the reflector aperture; as will be shown in the next section, eq. (14) is also a good approximation for the far field in most cases.

3.2 Two-mode approximation to the aperture field

After reflection from the offset reflector, the gaussian beam from the horn is converted into two gaussian beam modes in the aperture:

a fundamental mode with the in-line polarization (denoted $\mathbf{E}_{00}$) and a higher-order gaussian beam mode which includes the cross polarization (denoted $\mathbf{E}_{01}$). Depending on the polarization of the feed horn, the polarization of the fundamental and higher-order modes vary as shown in Fig. 8. For an arbitrary balanced feed polarization, a superposition of the two polarization cases shown in Fig. 8 can be made.

The expressions for the gaussian beam modes are[7]

(*i*) Fundamental mode:

$$\mathbf{E}_{00} = (H_{00}\hat{x}'' + V_{00}\hat{y}'')\frac{\mathbf{w}_{00}}{w_{00}} \exp\left\{-jkz'' - \frac{\rho^2}{w_{00}^2} + j\left[\arctan\left(\frac{2z''}{k\mathbf{w}_{00}^2}\right) - \frac{k\rho^2}{2R_{00}}\right]\right\}, \quad (16)$$

(*ii*) Higher-order mode:

$$\mathbf{E}_{01} = |V_{01}(\hat{x}'' \cos\alpha - \hat{y}'' \sin\alpha) - H_{01}(\hat{x}'' \sin\alpha + \hat{y}'' \cos\alpha)| \frac{\sqrt{2}\rho\mathbf{w}_{01}}{w_{01}^2} \cdot \exp\left\{-jkz'' - \frac{\rho^2}{w_{01}^2} + j\left[2\arctan\left(\frac{2z''}{k\mathbf{w}_{01}^2}\right) - \frac{k\rho^2}{2R_{01}}\right]\right\}. \quad (17)$$

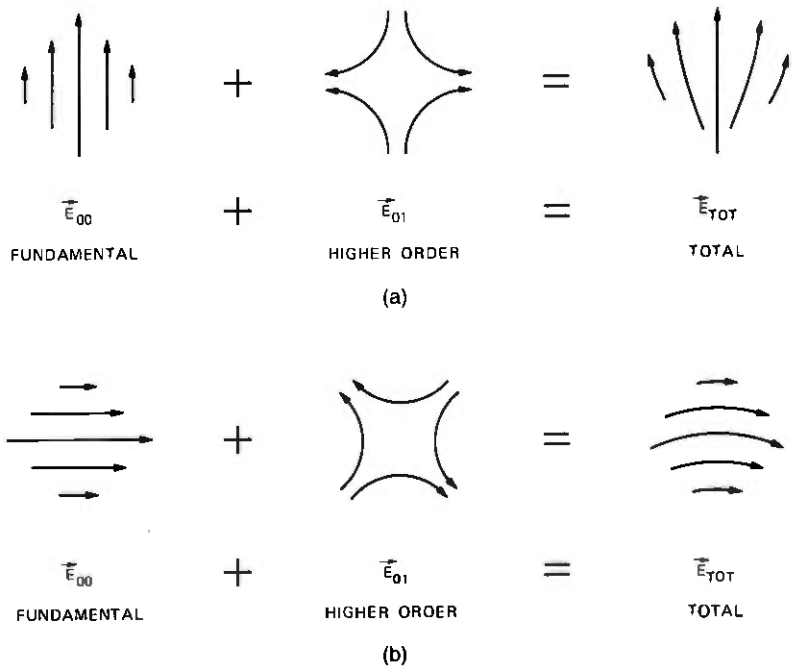


Fig. 8—Two-mode decomposition of aperture field (polarization looking in positive z direction, i.e., looking at the reflector). (a) Feed horn vertically polarized. (b) Feed horn horizontally polarized.

H_{00}, V_{00} and H_{01}, V_{01} are the phasor coefficients for horizontally and vertically polarized feeds. The subscripts refer to the standard TEM_{00} and TEM_{01} mode notations of Ref. 8; ρ, α, and z'' are cylindrical coordinates, with z'' denoting the distance along the beam axis from the beam waist. At the beam waist, the radius of curvature of the phase front, R, is infinite, and the field varies with increasing distance, ρ, from the axis at a rate determined by $\mathbf{w}$. For the fundamental mode, the field decreases to the $1/e$ value at $\rho = \mathbf{w}_{00}$. For the higher-order mode the field is maximum at $\rho = \mathbf{w}_{01}/\sqrt{2}$ and decreases to $\sqrt{2/e}$ of that value at $\rho = \mathbf{w}_{01}$. Away from the beam waist, $z'' \neq 0$, the field amplitude varies with ρ at a rate determined by w instead of $\mathbf{w}$, and the phase front has a finite radius of curvature, R. w and R are given by[8]

$$w = \mathbf{w}\sqrt{1 + \left(\frac{2z''}{k\mathbf{w}^2}\right)^2}, \tag{18}$$

and

$$R = z''\left[1 + \left(\frac{k\mathbf{w}^2}{2z''}\right)^2\right]. \tag{19}$$

Both modes have a characteristic exponential attenuation with distance from axis, $e^{-\rho^2/w^2}$, and a spherical wave front near the axis at constant z'', denoted by the term, $e^{-jk\rho^2/2R}$. Passing through a beam waist, the on-axis phase advances by π for the fundamental mode and 2π for the higher-order mode (relative to the plane-wave retardation, $e^{-jkz''}$). Thus, if the cross polarization and the in-line polarization are in phase at the beam waist (normally near the reflector aperture), they will be in phase quadrature at large distances from the beam waist (the far field). This phase quadrature relation gives rise to a beam shift with circular polarization as described in Ref. 2.

The choice of eq. (17) as the appropriate higher-order mode is based on its ability to approximate simultaneously both the cross polarization and the "space" taper (amplitude asymmetry from top to bottom of dish) of offset reflectors.

The in-line and cross-polarized fields in the aperture of the offset launcher of Fig. 1 were computed exactly by means of eq. (3); the resulting field-amplitude contours are shown in Fig. 9. In the aperture plane, the in-line and cross-polarized fields are in phase. Thus, the corresponding gaussian beam modes have their beam waists at the aperture and are in phase. This implies that the total field is linearly polarized everywhere in the aperture, and the direction of polarization varies in a manner determined by the ratio of the in-line and cross-polarized fields.

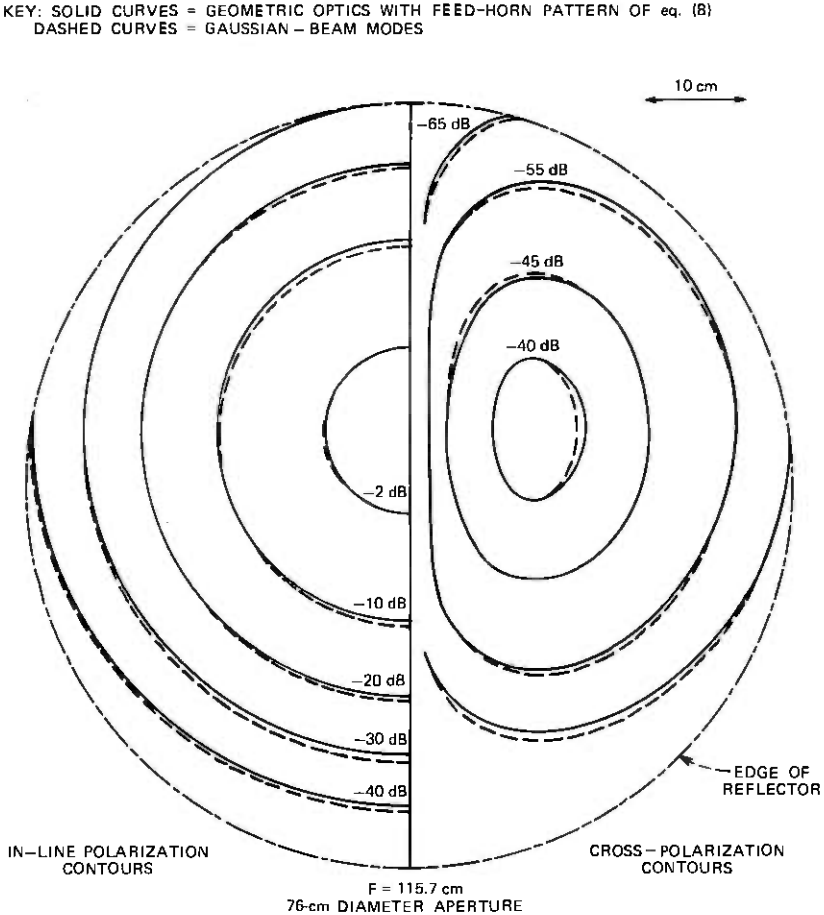


Fig. 9—Amplitude contours for the example reflector shown in Fig. 1.

The gaussian beam fields required to match the exactly computed aperture fields are found by first choosing a fundamental mode centered on the aperture with beam waist radius w_{00} such that it decreases 8.686 dB in power at the same radius as does the computed in-line polarized field, both being normalized to unit amplitude on axis.

With these criteria and the approximation that r' is nearly constant over the reflector,

$$r' \doteq r_0 = F \sec^2\left(\frac{\theta_0}{2}\right), \tag{20}$$

the parameters of the fundamental mode are determined as follows. The radius, $\rho_{1/e}$, at which the field drops to $1/e$ times its on-axis value

is found in the direction $\alpha = 0$. (Note that $\rho_{1/e}$ is larger in the $\alpha = -\pi/2$ direction and smaller in the $\alpha = \pi/2$ direction due to the space taper in these directions. Thus, $\rho_{1/e}$, as determined from the $\alpha = 0$ direction, approximates the average $\rho_{1/e}$ over all directions.)

$$\mathbf{w}_{00} \triangleq \rho_{1/e} = 2r_0 \sin\left(\tfrac{1}{2}\theta'_{1/e}\right), \tag{21}$$

where, from the gaussian approximation for the feed-horn pattern,

$$\theta'_{1/e} = \theta_c\sqrt{\frac{2}{\ln 10}}. \tag{22}$$

For unit amplitude on-axis, eq. (16) requires (we restrict our discussion here to the vertically polarized case; the horizontally polarized case yields identical expressions)

$$V_{00} = 1. \tag{23}$$

The higher-order mode parameters are found from the cross-polarization characteristics. Since the cross polarization is maximum at θ'_m, we have, from eq. (13),

$$\mathbf{w}_{01} = 2\sqrt{2}r_0 \sin\left(\tfrac{1}{2}\theta'_m\right) = 2\sqrt{2}r_0 \sin\left(\frac{\theta_c}{2\sqrt{\ln 10}}\right). \tag{24}$$

By comparing eqs. (24) and (21), it is seen that, for small θ_c, $\mathbf{w}_{01} \doteq \mathbf{w}_{00}$.

The amplitude of the higher-order gaussian beam mode is given by the maximum cross-polarization amplitude ratio ($\mathbf{C}_a$). From eqs. (14) and (17):

$$V_{01} = \frac{\theta_c \tan\dfrac{\theta_0}{2}}{\sqrt{\ln 10}}. \tag{25}$$

The phase of the higher-order mode follows from the fact that at the beam waist the in-line and cross-polarized fields are in phase.

Using eqs. (21) through (25), the gaussian beam mode approximations to the aperture fields, plotted as dashed contours in Fig. 9, compare favorably with those obtained by geometrical optics (solid contours).

3.3 The far fields

The parameters of the gaussian beam modes being thus determined, it is possible to compute the in-line and cross-polarized fields at any position in the main beam of the field radiated from the reflector. The far field is of particular interest,

$$z'' \gg \frac{k\mathbf{w}^2}{2}, \tag{26}$$

allowing eqs. (18) and (19) to be approximated by

$$w \doteq \frac{2z''}{k\mathbf{w}} \quad \text{and} \quad R \doteq z''. \tag{27}$$

Also, if the beamwidth is small,

$$r_f \doteq \sqrt{(z'')^2 + \rho^2} = z'' + \frac{\rho^2}{2z''}; \quad \theta_f \doteq \frac{\rho}{z''}. \tag{28}$$

From eqs. (16) and (17), the in-line field is therefore given by

$$E_V \doteq \frac{jke^{-jkr_f}}{2r_f}\left[V_{00}\mathbf{w}_{00}^2 \exp\left(-\frac{\theta_f^2}{\psi_{00}^2}\right) - j\sqrt{2}V_{01}\mathbf{w}_{01}^2 \frac{\theta_f}{\psi_{01}} (\sin\alpha) \exp\left(-\frac{\theta_f^2}{\psi_{01}^2}\right)\right] \tag{29}$$

and the cross-polarized field by

$$E_H = -\frac{ke^{-jkr_f}}{2r_f}\sqrt{2}V_{01}\mathbf{w}_{01}^2 \frac{\theta_f}{\psi_{01}} (\cos\alpha) \exp\left(-\frac{\theta_f^2}{\psi_{01}^2}\right), \tag{30}$$

for the "vertically" polarized feed. We have defined

$$\psi_{00} \triangleq \frac{2}{k\mathbf{w}_{00}} \quad \text{and} \quad \psi_{01} \triangleq \frac{2}{k\mathbf{w}_{01}}, \tag{31}$$

the angular beam radii in the far field.

The $\mathbf{C}_f$ from eqs. (14), (23), (25), (29), and (30), where the subscript, f, indicates the far-field maximum cross-polarization amplitude ratio, is

$$\mathbf{C}_f = j\frac{V_{01}}{V_{00}\sqrt{e}}\left(\frac{\mathbf{w}_{01}}{\mathbf{w}_{00}}\right)^2 = \mathbf{C}_a\left(\frac{\mathbf{w}_{01}}{\mathbf{w}_{00}}\right)^2. \tag{32}$$

Thus, the far-field $\mathbf{C}$ is $j(\mathbf{w}_{01}/\mathbf{w}_{00})^2$ times that in the aperture and occurs in the azimuth plane at an angle

$$\frac{\psi_{01}}{\sqrt{2}} = \frac{\sqrt{2}}{k\mathbf{w}_{01}} \text{ radians.} \tag{33}$$

As mentioned above, for small feed beamwidths, θ_c,

$$\mathbf{w}_{01} \doteq \mathbf{w}_{00}, \tag{34}$$

so the far-field $\mathbf{C}$ is approximately equal to that of the aperture, and the peak cross-polarization lobe occurs at approximately the -4.34-dB level of the main beam.

A comparison of the experimental and theoretical far-field patterns for in-line and cross-polarized fields from the offset launcher of Fig. 1 at 28.56 GHz and 19.04 GHz are shown in Figs. 4 and 5. The main discrepancy between theory and experiment are the shoulders on the

sides of the experimental pattern at a level of about −24 dB not present on the sides of the theoretical main beam. As mentioned in Section 2.3, the shoulders are probably due to phase error in the horn aperture which is not in the gaussian feed pattern assumed in the theory.

3.4 Truncation effects

The effect of truncating the aperture of the launcher at various circular contours of the fundamental mode can be computed. Let the radius of the aperture at the truncation be c. Then the taper at the truncation is

$$T = \left(\frac{20}{\ln 10}\right)\left(\frac{c}{\mathrm{w}_{00}}\right)^2. \tag{35}$$

The radiation integrals in the x'', z'' plane (where the cross-polarization is largest) for the fundamental mode and higher-order mode are

$$E_{00_f} = \int_0^c \rho d\rho \int_0^{2\pi} d\alpha V_{00} \exp\left[-\left(\frac{\rho}{\mathrm{w}_{00}}\right)^2 - jk\rho \cos\alpha \sin\theta_f\right], \tag{36}$$

and

$$E_{01_f} = -\int_0^c \rho d\rho \int_0^{2\pi} d\alpha V_{01} \frac{\sqrt{2}\rho}{\mathrm{w}_{01}} \cdot \exp\left[-\left(\frac{\rho}{\mathrm{w}_{01}}\right)^2 - jk\rho \cos\alpha \sin\theta_f\right] \cos\alpha \cos\theta_f. \tag{37}$$

From eq. (36), the in-line polarization far-field on-axis is

$$E_{00_f}(\theta_f = 0) = \pi V_{00}\mathrm{w}_{00}^2\left[1 - \exp\left(-\frac{c^2}{\mathrm{w}_{00}^2}\right)\right]. \tag{38}$$

The cross-polarized far-field pattern from (37) is, letting $\cos\theta_f \doteq 1$,

$$E_{01_f} \doteq \frac{j\pi V_{01}\mathrm{w}_{01}^2}{\sqrt{e}} I(\theta_f), \tag{39}$$

where

$$I(\theta_f) \triangleq 2\sqrt{2e}\int_0^{c/\mathrm{w}_{01}} t^2 e^{-t^2} J_1(tk\mathrm{w}_{01} \sin\theta_f) dt, \tag{40}$$

and $J_1(x)$ is the first-order Bessel function of the first kind. From eqs. (34) and (35), choosing the taper, T, determines $I(\theta_f)$. Numerically integrating eq. (40), the peak value, I_P, of $I(\theta_f)$ and the location, θ_{f_p}, at which $I(\theta_f)$ is maximum were determined for various tapers, T. The resulting far field C_f relative to that of an infinite aperture,

$$-20 \log_{10}\left(\frac{\mathrm{C}_f}{\mathrm{C}_{f\infty}}\right) = -20 \log_{10}\left(\frac{I_P}{1 - e^{-c^2/\mathrm{w}^2_{00}}}\right), \tag{41}$$

is plotted in Fig. 10 along with the angular position, θ_{fP}, of the peak in the cross-polarization lobe relative to that, $\theta_{fp\infty}$, of an infinite aperture. Notice that the effect of truncation is to reduce the cross-polarized signal relative to the co-polarized signal (higher $\mathbf{C}_f$) and to move the cross-polarization lobes out to larger angles off-axis. Furthermore, these truncation effects depend only on the truncation taper and are essentially independent of the offset geometry.

Figure 11 shows $\mathbf{C}_f$ as determined from the gaussian beam formulas for a wide range of offset reflector geometries indicating an infinite aperture (no truncation) and the aperture truncated at a 10-dB taper. With the infinite aperture, the $\mathbf{C}_f$ from the gaussian beam theory are within 0.2 dB of the cross polarization in the aperture, $\mathbf{C}_a$ (not shown), for all geometries on the figure. When the aperture is truncated at the 10-dB level, $\mathbf{C}_f$ is, from Fig. 10, 2.2 dB smaller than that for the infinite aperture, as is also seen from the dashed curve in Fig. 11. These dashed curves agree with cross polarization obtained by numerically computing[2] the radiation integral over the aperture field found by geometrical optics projection of the radiation pattern of a dual-mode feed horn.

Using eqs. (14) and (32), the maximum cross-polarization levels as a function of offset angle were computed for the precise geometry

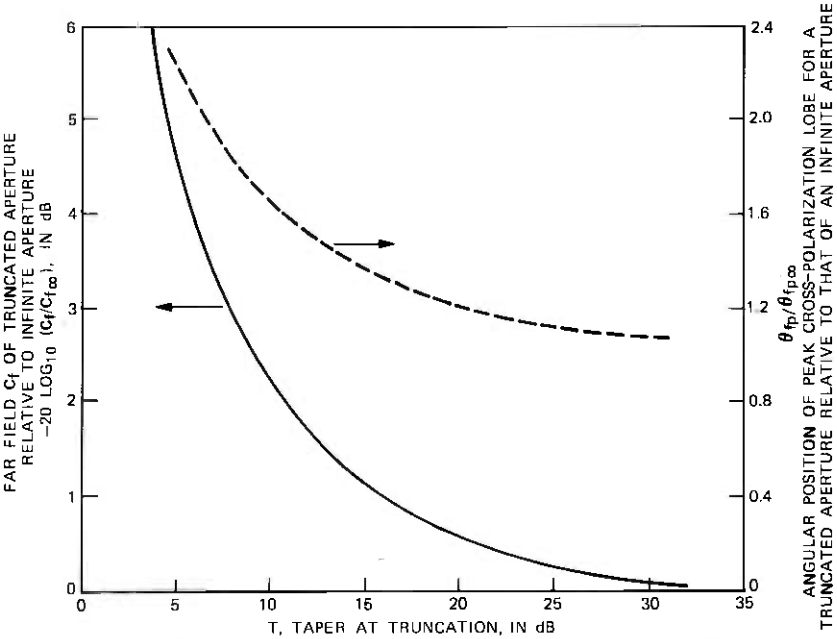


Fig. 10—Effect of aperture truncation on cross-polarization pattern.

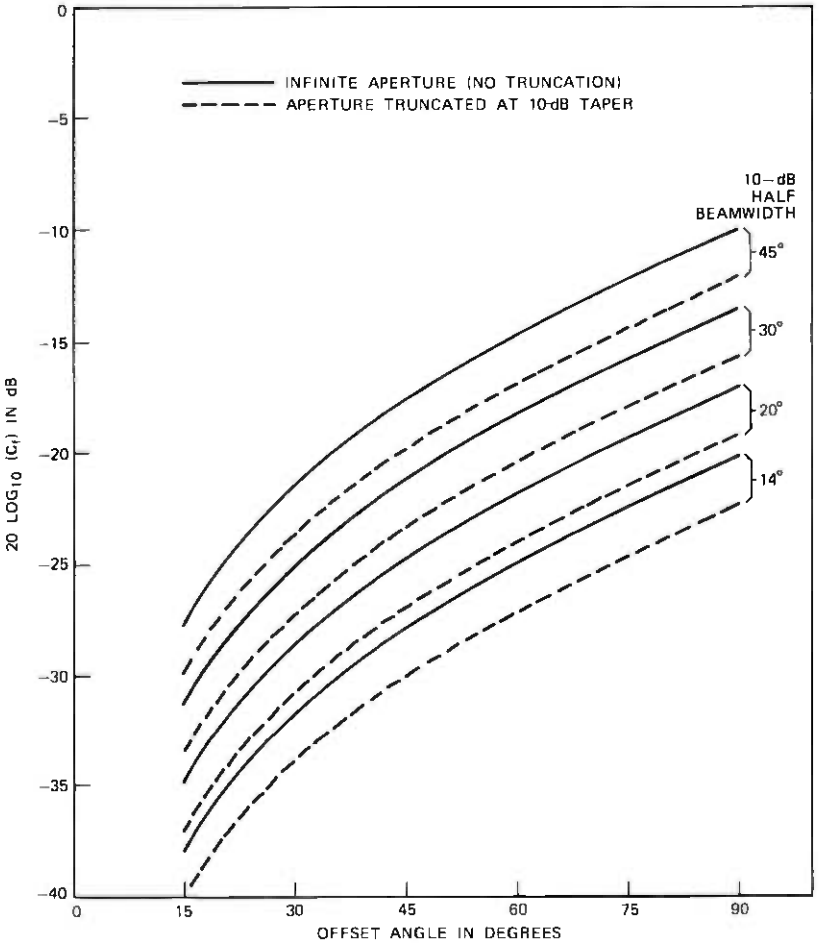


Fig. 11—Maximum cross-polarization amplitude ratio for a range of offset geometries.

of the offset launcher of Fig. 1. The factor $(w_{01}/w_{00})^2$ which converts aperture cross polarization to far-field cross polarization is only 1.0004 (i.e., 0.008 dB). The 10-dB half angle, θ_c, of the feed-horn gaussian-beam approximation (discussed in Section 2.1) is 9 degrees. The calculated cross polarization is compared with the measured cross polarization as a function of offset angle in Fig. 6. The theory appears to be in good agreement with the measurements at both frequencies.

IV. CONCLUSIONS

It has been demonstrated that an offset launcher can provide low cross polarization and a low-sidelobe symmetrical beam when fed

with a suitable horn at a small enough offset angle. Simple formulas for the far-field performance of the launcher are derived in terms of two gaussian modes; comparison with measurements at 19 and 28 GHz shows good agreement. The maximum cross-polarization amplitude ratio is found to change little from aperture to far field. Offset reflector geometries have also been useful for multiple-beam applications.[9,10]

V. ACKNOWLEDGMENTS

It is a pleasure to acknowledge D. C. Hogg for many helpful suggestions, and W. E. Legg and R. H. Turrin for measurements of the dual-mode feed horn.

REFERENCES

1. C. Dragone and D. C. Hogg, "The Radiation Pattern and Impedance of Offset and Symmetrical Near-Field Cassegrainian and Gregorian Antennas," IEEE Trans. Antennas & Propagation, *AP-22*, No. 3 (May 1974), p. 472.
2. T. S. Chu and R. H. Turrin, "Depolarization Properties of Offset Reflector Mirrors," IEEE Trans. Microwave Theory & Techniques, *MTT-23*, No. 4 (April 1975), pp. 377–379.
3. J. A. Arnaud and J. T. Ruscio, "Guidance of 100-GHz Beams by Cylindrical Mirrors," IEEE Trans. Microwave Theory & Techniques, *MTT-23*, No. 4 (April 1975), pp. 377–379.
4. R. H. Turrin, "Dual Mode Small-Aperture Antennas," IEEE Trans. Antennas & Propagation, *AP-15*, No. 2 (March 1967), pp. 307–308.
5. E. M. T. Jones, "Paraboloid Reflector and Hyperboloid Lens Antennas," IRE Trans. Antennas and Propagation, *AP-2* (July 1954), pp. 119–127.
6. *Reference Data for Radio Engineers*, Fifth Edition, ITT, New York: H. W. Sams, 1968, pp. 44–9—44–10.
7. G. Gaubau and F. Schwering, "On the Guided Propagation of Electromagnetic Wave Beams," IRE Trans. Antennas & Propagation, *AP-9*, No. 3 (May 1961), pp. 248–256.
8. H. Kogelnik and T. Li, "Laser Beams and Resonators," Applied Optics, *5*, No. 10 (October 1966), pp. 1550–1567.
9. C. J. Sletten et al., "Corrective Line Sources for Paraboloids," IRE Trans. Antennas & Propagation, *AP-6* (July 1958), pp. 239–251.
10. E. A. Ohm, "A Proposed Multiple-Beam Microwave Antenna for Earth Stations and Satellites," B.S.T.J., *53*, No. 8 (October 1974), pp. 1657–1667.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 54, No. 7, September 1975
Printed in U.S.A.

Upper Bound on Error Probability for Detection With Unbounded Intersymbol Interference

By A. D. WYNER

(Manuscript received February 27, 1975)

Forney's asymptotic upper bound for per-bit error probability in the detection of pulse-amplitude-modulated digital data in the presence of additive white gaussian noise was obtained for the case where the duration of the intersymbol interference is bounded. In this paper, we show the validity of Forney's bound under much weaker assumptions that allow unbounded intersymbol interference.

I. INTRODUCTION

We consider the situation where a data sequence $a_0, \cdots, a_{N-1}$ of ± 1's is transmitted via pulse amplitude modulation as $\sum_{k=0}^{N-1} h(t - kT)a_k$ and received in the presence of additive white gaussian noise with one-sided spectral density σ^2. In a recent series of papers, Forney,[1] Foschini,[2] and Mazo[3] developed an asymptotic (as $\sigma^2 \to 0$) upper bound on the error probability per data bit P_e:

$$P_e \leqq \exp\left\{-\frac{d^2(h)}{4\sigma^2}[1 + o(1)]\right\}, \tag{1}$$

where $d(h)$ is the minimum $\mathcal{L}_2$ distance between distinct modulated pulse sequences. This bound holds under the strong assumption that the pulse $h(t)$ is supported on finite interval.

In this paper, we show that (1) is valid for a considerably wider class of $h(t)$. Roughly speaking, our assumptions are little more than that $h(t)$ is in $\mathcal{L}_1(-\infty, \infty)$ and $\mathcal{L}_2(-\infty, \infty)$, and that $H(f)$, the Fourier transform of $h(t)$, does not vanish on an interval. The precise conditions on $h(t)$ under which (1) holds are given below. In particular, (1) is valid when $H(f)$ is a rational function.

In Section II we give a precise statement of our results, and the proof follows in Section III.

II. FORMAL STATEMENT OF PROBLEM AND RESULTS

In this section, we give a precise statement of the problem and the results that were stated informally in Section I.

We begin with some definitions. We denote N vectors by boldface superscripted letters, and components by subscripted letters, e.g.

$\mathbf{u}^N = (u_0, \cdots, u_{N-1})$. When the dimension N is clear from the context, we omit the superscript. Define the sets $\mathcal{A}_N$, $\mathcal{A}_{N,k}^+$, $\mathcal{A}_{N,k}^-$ by

$$\begin{aligned} \mathcal{A}_N &= \{\mathbf{u}^N : u_j = \pm 1, 0 \leqq j \leqq N-1\}, \\ \mathcal{A}_{N,k}^+ &= \{\mathbf{u}^N : u_k = +1, u_j = \pm 1, j \neq k\}, \\ \mathcal{A}_{N,k}^- &= \{\mathbf{u}^N : u_k = -1, u_j = \pm 1, j \neq k\}. \end{aligned} \tag{2}$$

Of course, $\mathcal{A}_N = \mathcal{A}_{N,k}^+ \cup \mathcal{A}_{N,k}^-$. Again, when N is clear from the context, we write $\mathcal{A}_N = \mathcal{A}$, $\mathcal{A}_{N,k}^+ = \mathcal{A}_k^+$, $\mathcal{A}_{N,k}^- = \mathcal{A}_k^-$.

Next, let $f(t)$, $g(t)$, and $-\infty < t < \infty$ be real-valued measurable functions. The inner product of f and g is denoted by

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(t)g(t)dt, \tag{3a}$$

and the norm of f is

$$\|f\| = (\langle f, f \rangle)^{\frac{1}{2}} = \left(\int_{-\infty}^{\infty} f^2(t)dt \right)^{\frac{1}{2}}. \tag{3b}$$

For a vector $\mathbf{u}^N \in \mathcal{A}_N$, and $f(t)$, $-\infty < t < \infty$, a real-valued function, let the function $h*\mathbf{u} = s$ be defined by

$$s(t) = \sum_{k=0}^{N-1} f(t - kT)u_k,$$

where $T > 0$ is a fixed parameter.

We are concerned with the following modulation scheme. Let $\mathbf{a}^N = (a_0, \cdots, a_{N-1}) \in \mathcal{A}_N$ denote the data to be transmitted. Assume that all the 2^N vectors in $\mathcal{A}_N$ are equally likely. The transmitted signal is the function $h*\mathbf{a}^N$, where the pulse $h(t)$ is a fixed function for which $\|h\| < \infty$. The received signal is

$$y(t) = (h*\mathbf{a}^N)(t) + z(t), \qquad -\infty < t < \infty, \tag{4}$$

where $z(t)$ is a sample from a white gaussian noise process with zero mean and one-sided spectral density σ^2.

The decoder associates with the received signal y, a vector $D(y) = \hat{\mathbf{a}}^N \in \mathcal{A}_N$. Corresponding to a given decoder function D, let the bit error probability be

$$P_{eN}(D) = \frac{1}{N} \sum_{k=0}^{N-1} \Pr\{\hat{a}_k \neq a_k\}. \tag{5}$$

Also, define the optimum error probability

$$P_{eN}^* = P_{eN}^*(h, \sigma^2) = \inf_D P_{eN}(D). \tag{6}$$

We are concerned here with the asymptotics of $P_{eN}^*(h, \sigma^2)$, as $\sigma^2 \to 0$, i.e., as the signal-to-noise ratio approaches infinity. Accordingly, define

$$E_N(h) = -\liminf_{\sigma^2 \to 0} \sigma^2 \log P_{eN}^*(h, \sigma^2), \tag{7a}$$

so that, as $\sigma^2 \to 0$,

$$P_{eN}^*(h, \sigma^2) \leqq \exp\left\{-\frac{E_N(h)}{\sigma^2}[1 + o(1)]\right\}. \tag{7b}$$

Next, consider a particular decoder that is of special interest here—the maximum-likelihood decoder, denoted D_h. In the present problem, $D_h(y)$ can be taken to be that $\hat{\mathbf{u}} \in \mathfrak{A}_N$ such that for all $\mathbf{u} \in \mathfrak{A}_N$, $\mathbf{u} \neq \hat{\mathbf{u}}$,

$$\|y_1 - h*\hat{\mathbf{u}}\| < \|y_1 - h*\mathbf{u}\|, \tag{8}$$

where y_1 is the projection of y onto the subspace of $\mathcal{L}_2(-\infty, \infty)$ spanned by the signals $h*\mathbf{u}$, $\mathbf{u} \in \mathfrak{A}_N$. With probability 1, (8) will be satisfied for some $\hat{\mathbf{u}} \in \mathfrak{A}_N$.

Now, subject to the condition that $h(t)$ has finite support, i.e., there exists a $t_0 > 0$ such that

$$h(t) = 0, \qquad \text{for } |t| > t_0, \tag{9}$$

Forney,[1] Foschini,[2] and Mazo[3] have shown that $E(h) \geqq d^2(h)/4$, where the "minimum distance" $d(h)$ is defined by

$$d(h) = \liminf_{N \to \infty} \min_{\substack{\mathbf{u},\mathbf{v} \in \mathfrak{A}_N \\ \mathbf{u} \neq \mathbf{v}}} \|h*\mathbf{u} - h*\mathbf{v}\|. \tag{10}$$

Thus, as $\sigma^2 \to 0$,

$$P_{eN}^*(h, \sigma^2) \leqq \exp\left\{-\frac{d^2(h)}{4\sigma^2}[1 + o(1)]\right\}. \tag{11}$$

Inequality (11) is established by showing that the error probability for the maximum likelihood decoder, $P_{eN}(D_h)$, is overbounded by the right member of (11). This is done by writing (this is not as difficult as it looks)

$$\begin{aligned}
P_{eN}(D_h) &= \frac{1}{N}\sum_{k=1}^{N} \sum_{\mathbf{u} \in \mathfrak{A}_{N,k}^+} 2^{-(N-1)} \Pr\{D_h(y) \in \mathfrak{A}_{N,k}^- | y = h*\mathbf{u} + z\} \\
&= \frac{1}{N}\sum_k \sum_{\mathbf{u} \in \mathfrak{A}_k^+} 2^{-(N-1)} \Pr\{\bigcup_{\mathbf{v} \in \mathfrak{A}_k^-} \{D_h(y) = \mathbf{v}\} | y = h*\mathbf{u} + z\} \\
&\leqq \frac{1}{N}\sum_k \sum_{\mathbf{u} \in \mathfrak{A}_k^+} 2^{-(N-1)} \\
&\quad \cdot \Pr\{\bigcup_{\mathbf{v} \in \mathfrak{A}_k^-} \{\|y_1 - h*\mathbf{u}\| \geqq \|y_1 - h*\mathbf{v}\|\} | y = h*\mathbf{u} + z\} \\
&= \frac{1}{N}\sum_k \sum_{\mathbf{u} \in \mathfrak{A}_k^+} 2^{-(N-1)} \Pr \bigcup_{\mathbf{v} \in \mathfrak{A}_k^-} \{\langle z, h*(\mathbf{v} - \mathbf{u})\rangle \\
&\qquad\qquad \geqq \tfrac{1}{2}\|h*(\mathbf{v} - \mathbf{u})\|\} \\
&\triangleq \psi_N(h, \sigma^2).
\end{aligned} \tag{12}$$

Relation (12) is valid for any $h(t)$. Subject to condition (9), it is then

shown that, as $\sigma^2 \to 0$,

$$\psi_N(h, \sigma^2) \leqq \exp\left\{-\frac{d^2(h)}{4\sigma^2}[1 + o_1(1)]\right\}, \tag{13}$$

where $o_1(1)$ does not depend on N. Thus, since $P_{eN}^* \leqq P_{eN}(D_h)$, (11) holds. Further, the $o(1)$ term in (11) does not depend on N. An interesting by-product of these results is that the performance indicated in (11) is achievable via the decoder D_h. This decoder can be instrumented (using the Viterbi algorithm) with a complexity which remains bounded as $N \to \infty$.

We now drop the assumption that $h(t)$ has finite support. Instead, we assume that $h(t)$ satisfies the following conditions:

(*i*) There exists a nonnegative $\mathcal{L}_1$ function $g_0(t)$, i.e.,

$$\int_{-\infty}^{\infty} g_0(t)dt < \infty,$$

such that

$$|h(t)| \leqq g_0(t), \qquad -\infty < t < \infty, \tag{14}$$

and such that g_0 is monotone in $|t|$.

(*ii*) Let

$$H(f) = \int_{-\infty}^{\infty} h(t)e^{-i2\pi ft}dt, \qquad -\infty < f < \infty \tag{15}$$

be the Fourier transform of $h(t)$. By (*i*), $\int |h(t)|dt < \infty$, so that $H(f)$ is well defined for all f. We assume that there exists a nonnegative $\mathcal{L}_1$ function $G_1(f)$ which is monotone in $|f|$, such that

$$|H(f)|^2 \leqq G_1(f), \qquad -\infty < f < \infty. \tag{16}$$

(*iii*) Let the "folded spectrum" of h be

$$S(f) = \sum_{n=-\infty}^{\infty} \left|H\left(f + \frac{n}{T}\right)\right|^2, \qquad 0 \leqq f \leqq \frac{1}{T}. \tag{17}$$

We show in Appendix A that $S(f)$, $0 \leqq f \leqq 1/T$, is finite and continuous. We assume that $S(f) > 0$, $0 \leqq f \leqq 1/T$. Let

$$m = \min_{0 \leqq f \leqq 1/T} S(f) > 0, \tag{18}$$

where the existence of the minimum follows from the continuity of $S(f)$ on the compact interval $[0, 1/T]$.

Remarks:

(1) Condition (*i*) is just slightly stronger than simply requiring h to be in $\mathcal{L}_1(-\infty, \infty)$. Condition (14) forces $h(t)$ to go to zero as $|t| \to \infty$ in a "well-behaved" manner. Condition (*ii*) imposes a similar condition on $|H(f)|^2$.

(2) For the very important special case where $H(f)$ is a rational function, i.e., $H(f) = P(i2\pi f)/Q(i2\pi f)$, and P, Q are polynomials

with the degree of $P <$ degree of Q, then conditions (*i*) and (*ii*) are satisfied. Since $H(f)$ has only a finite number of zeros, condition (*iii*) is also satisfied.

(3) Suppose that $H(f)$ has no more than a countable number of zeros, but that $S(f) = 0$ for some $f \in [0, 1/T]$. It is easy to see that some arbitrarily small change in T will cause $S(f)$ to be strictly positive for all $f \in [0, 1/T]$. Thus, condition (*iii*) is not especially restrictive.

We now state our main result, the proof of which is in Section III.

Theorem 1: *Let h satisfy conditions (i), (ii), and (iii) above. Then, for all $\epsilon > 0$, there exists a $\tau_0 = \tau_0(\epsilon)$ sufficiently large so that, for all $\tau > \tau_0$,*

$$P_{eN}(D_{h_\tau}) \leqq \psi_N[h_\tau, \sigma^2(1 + \epsilon)^2],$$

where

$$h_\tau(t) = \begin{cases} h(t), & |t| \leqq \tau, \\ 0, & |t| > \tau, \end{cases} \tag{19}$$

is the truncated version of $h(t)$. The quantity τ_0 does not depend on N.

Since h_τ has finite support, we conclude from Theorem 1 and (13) that, for all $\epsilon > 0$ and τ sufficiently large,

$$\begin{aligned} P^*_{eN}(h, \sigma^2) &\leqq P_{eN}(D_{h_\tau}) \\ &\leqq \exp\left\{-\frac{d^2(h_\tau)}{4\sigma^2(1+\epsilon)^2}[1 + o_2(1)]\right\}, \end{aligned} \tag{20a}$$

[where $o_2(1)$ is independent of N] so that

$$E_N(h) = -\liminf_{\sigma^2 \to 0} \sigma^2 \log P^*_e(h, \sigma^2) \geqq \frac{d^2(h_\tau)}{4(1+\epsilon)^2}. \tag{20b}$$

We show in Appendix B that

$$d(h_\tau) \to d(h), \qquad \text{as } \tau \to \infty, \tag{21}$$

so that letting $\epsilon \to 0$ and $\tau \to \infty$ in (20b) yields

$$E_N(h) \geqq \frac{d^2(h)}{4}. \tag{22}$$

We state this as

Corollary 2: *Let h satisfy conditions (i) to (iii) above. Then, as $\sigma^2 \to 0$,*

$$P^*_{eN}(h, \sigma^2) \leqq \exp\left\{-\frac{d^2(h)}{4\sigma^2}[1 + o_3(1)]\right\},$$

where $o_3(1)$ is independent of N.

We conclude this section with a remark concerning the relationship of the bound of Forney et al. (11) with the result of Corollary 2. We

can rewrite (11) as

$$P^*_{eN}(h, \sigma^2) = K_1(t_0, \sigma^2)e^{-d^2(h)/2\sigma^2}, \tag{23}$$

and the bound of Corollary 2 as

$$P^*_{eN}(h, \sigma^2) \leqq K_2(\|h\|_1, m, \sigma^2)e^{-d^2(h)/4\sigma^2}. \tag{24}$$

Here, we made explicit the dependence of K_1 on the support interval t_0 of $h(t)$ [see (9)], and the dependence of K_2 on $\|h\|_1$, the $\mathcal{L}_1$ norm of h, and on $m = \min S(f)$. Both K_1 and K_2 increase in $1/\sigma^2$ slower than $e^{d^2(h)/4\sigma^2}$. But $K_1(t_0, \sigma^2) \to \infty$ as $t_0 \to \infty$, and $K_2(\|h\|_1, m, \sigma^2) \to \infty$, as $\|h\|_1 \to \infty$ or as $m \to 0$. Thus, although it might seem reasonable to assume that all $h(t)$ satisfy (9) for some t_0, the bound of (23) depends on that t_0 and becomes meaningless as $t_0 \to \infty$. Similarly, although it might be reasonable to assume for any $h(t)$ that $\|h\|_1 < \infty$ and $m = \min S(f) > 0$, the bound of (24) depends on these quantities and also becomes meaningless as $\|h\|_1 \to \infty$ or $m \to 0$. Therefore, both bounds have their limitations; the new one, however, is considerably less limited.

III. PROOF OF THEOREM 1

Let h satisfy (*i*) to (*iii*). Let $h_\tau(t)$ be as defined in (19), and let $\tilde{h}_\tau(t) = h(t) - h_\tau(t)$, i.e.,

$$\tilde{h}_\tau(t) = \begin{cases} 0, & |t| \leqq \tau, \\ h(t), & |t| > \tau. \end{cases} \tag{25}$$

Then, if the data sequence is $\mathbf{u} \in \mathcal{A}_N$, the received sequence is $y = h*\mathbf{u} + z = h_\tau*\mathbf{u} + \hat{z}$, where

$$\hat{z} = z + \tilde{h}_\tau*\mathbf{u}. \tag{26}$$

Following the same steps as in (12), we obtain

$$P_{eN}(D_{h_\tau}) \leqq \frac{1}{N} \sum_k \sum_{\mathbf{u} \in \mathcal{A}_k^+} 2^{-(N-1)} \Pr \bigcup_{\mathbf{v} \in \mathcal{A}_k^-} \{\langle \hat{z}, h_\tau*(\mathbf{v} - \mathbf{u})\rangle \geqq \tfrac{1}{2}\|h_\tau*(\mathbf{v} - \mathbf{u})\|\}, \tag{27}$$

where $\hat{z}$ is given in (26).

We will show that, for arbitrary $\epsilon > 0$, there exists a $\tau_0 = \tau_0(\epsilon, h)$ (τ_0 independent of N), such that for $\tau \geqq \tau_0$, the event

$$\begin{aligned} &\{\langle \hat{z}, h_\tau*(\mathbf{v} - \mathbf{u})\rangle \geqq \tfrac{1}{2}\|h_\tau*(\mathbf{v} - \mathbf{u})\|^2\} \\ &\subseteq \{\langle z, h_\tau*(\mathbf{v} - \mathbf{u})\rangle \geqq \frac{(1 + \epsilon)}{2}\|h_\tau*(\mathbf{v} - \mathbf{u})\|^2\}, \end{aligned} \tag{28}$$

for all $\mathbf{u} \in \mathcal{A}_k^+$, $\mathbf{v} \in \mathcal{A}_k^-$. Substituting (28) into (27) yields, on comparison with (12),

$$P_{eN}(D_{h_\tau}) \leqq \psi_N[h_\tau, \sigma^2(1 + \epsilon)^2],$$

which is Theorem 1. It remains to establish (28).

Relation (28) will follow immediately when we show the existence of a $\tau_0(\epsilon, h)$ such that, for $\tau \geqq \tau_0$ and all $\mathbf{u}$, $\mathbf{v}$,

$$|\langle \tilde{h}_\tau * \mathbf{u}, h_\tau * (\mathbf{v} - \mathbf{u})\rangle| \leqq \frac{\epsilon}{2}\|h_\tau * (\mathbf{v} - \mathbf{u})\|^2. \tag{29}$$

If (29) holds, the event in the left member of (28)

$$\begin{aligned}\{\langle \hat{z}, h_\tau * (\mathbf{v} - \mathbf{u})\rangle \geqq \tfrac{1}{2}\|h_\tau * (\mathbf{v} - \mathbf{u})\|^2\} \\ = \{\langle z, h_\tau * (\mathbf{v} - \mathbf{u})\rangle \geqq \tfrac{1}{2}\|h_\tau * (\mathbf{v} - \mathbf{u})\|^2 - \langle \bar{h}_\tau * \mathbf{u}, h_\tau * (\mathbf{v} - \mathbf{u})\rangle\} \\ \subseteq \left\{\langle z, h_\tau * (\mathbf{v} - \mathbf{u})\rangle \geqq \frac{1+\epsilon}{2}\|h_\tau * (\mathbf{v} - \mathbf{u})\|^2\right\},\end{aligned} \tag{30}$$

which is the right member of (28). Thus, it remains to establish (29).

Let $\mathbf{w} = (w_0, \cdots, w_{N-1}) = \mathbf{v} - \mathbf{u}$. The entries of $\mathbf{w}$ are $0, \pm 2$. Also set $q = \bar{h}_\tau * \mathbf{u}$, and $r = h_\tau * \mathbf{w}$. Then

$$\begin{aligned}|\langle \tilde{h}_\tau * \mathbf{u}, h_\tau * (\mathbf{v} - \mathbf{u})\rangle| = |(q, r)| &\leqq \int_{-\infty}^{\infty} |q(t)|\,|r(t)|\,dt \\ &\leqq [\sup_{-\infty < t < \infty} |q(t)|]\int_{-\infty}^{\infty} |r(t)|\,dt.\end{aligned} \tag{31}$$

Consider

$$\begin{aligned}\int_{-\infty}^{\infty} |r(t)|\,dt &= \int_{-\infty}^{\infty} \left|\sum_{k=0}^{N-1} h_\tau(t - kT)w_k\right| dt \\ &\leqq \sum_k |w_k| \int_{-\infty}^{\infty} |h_\tau(t - kT)|\,dt \leqq \|h\|_1 \sum_{k=0}^{N-1} |w_k|,\end{aligned} \tag{32}$$

where $\|h\|_1 = \int |h(t)|\,dt < \infty$, by condition (i).

We obtain an upper bound on $\sum |w_k|$ as follows. Since $w_k = 0, \pm 2$, we have $\sum |w_k| = \frac{1}{2}\sum w_k^2$. Now, let $H_\tau(f) = \int_{-\infty}^{\infty} h_\tau(t)e^{i2\pi f t}dt$ be the Fourier transform of $h_\tau(t)$, and let

$$S_\tau(f) = \sum_{n=-\infty}^{\infty} \left|H_\tau\left(f + \frac{n}{T}\right)\right|^2, \qquad 0 \leqq f \leqq \frac{1}{T} \tag{33}$$

be the corresponding folded spectrum. Then, from Parseval's theorem,

$$\begin{aligned}\|r\|^2 = \|h_\tau * \mathbf{w}\|^2 &= \int_{-\infty}^{\infty} |H_\tau(f)|^2 |\sum_k w_k e^{i2\pi kTf}|^2 df \\ &= \int_0^{1/T} S_\tau(f)\,|\sum_k w_k e^{i2\pi kTf}|^2 df \\ &\geqq [\inf_{0 \leqq f \leqq 1/T} S_\tau(f)]\int_0^{1/T} |\sum w_k e^{i2\pi kTf}|^2 df \\ &= [\inf_f S_\tau(f)]\sum_k w_k^2 = [\inf_f S_\tau(f)]\sum |w_k| \cdot 2.\end{aligned} \tag{34}$$

Therefore,

$$\sum_k |w_k| \leqq \tfrac{1}{2}\|r\|^2[\inf_{0 \leqq f \leqq 1/T} S_\tau(f)]^{-1}. \tag{35}$$

Combining (31), (32), and (35), we have

$$|\langle q, r\rangle| \leqq [\sup_{-\infty<t<\infty} |q(t)|][\inf_{0\leqq f\leqq 1/T} S_\tau(f)]^{-1} \tfrac{1}{2}\|h\|_1\|r\|^2. \quad (36)$$

Now, we show in Appendix A that

$$\liminf_{\tau\to\infty} [\inf_{0\leqq f\leqq 1/T} S_\tau(f)] \geqq m > 0, \quad (37)$$

where $m = \min S(f) > 0$ [see condition (*iii*)]. Further, using condition (*i*) [particularly the monotonicity of $g_0(t)$], we have

$$\begin{aligned}
|q(t)| &= |\sum_k \tilde{h}_\tau(t-kT)u_k| \leqq \sum_{k=0}^{N-1} |\tilde{h}_\tau(t-kT)| \\
&\leqq \sum_{k=-\infty}^{\infty} |\tilde{h}_\tau(t-kT)| = \sum_{k:|t-kT|\geqq\tau} |h(t-kT)| \\
&\leqq \sum_{k:|t-kT|\geqq\tau} g_0(t-kT) \leqq \sum_{j=0}^{\infty} [g_0(\tau+jT)+g_0(-\tau-jT)] \\
&\leqq \frac{1}{T}\int_{\tau-T}^{\infty} g_0(t)dt + \frac{1}{T}\int_{-\tau+T}^{\infty} g_0(t)dt \to 0, \quad \text{as } \tau\to\infty. \quad (38)
\end{aligned}$$

Combining (36), (37), and (38), we obtain

$$\frac{|\langle q, r\rangle|}{\|r\|^2} = \frac{|\langle \tilde{h}_\tau * \mathbf{u}, h_\tau * (\mathbf{v}-\mathbf{u})\rangle|}{\|h_\tau * (\mathbf{v}-\mathbf{u}\|^2} \to 0,$$

as $\tau\to\infty$. This is equivalent to (29), so that the proof of Theorem 1 is complete.

IV. ACKNOWLEDGMENTS

I would like to acknowledge with thanks the help of J. E. Mazo and H. J. Landau.

APPENDIX A

The Folded Spectrum S(f)

We first show that $S(f)$ as given in (17) is always finite, i.e., the series in (17) converges for all $f \in [0, 1/T]$. From condition (*ii*), using the monotonicity of G_1,

$$\begin{aligned}
\xi(n_0, f) &\triangleq \sum_{|n|\geqq n_0} \left|H\left(f+\frac{n}{T}\right)\right|^2 \leqq \sum_{|n|\geqq n_0} G_1\left(f+\frac{n}{T}\right) \\
&= \sum_{n\leqq -n_0} G_1\left(f+\frac{n}{T}\right) + \sum_{n\geqq n_0} G_1\left(f+\frac{n}{T}\right), \\
&\leqq \sum_{n\leqq -n_0} G_1\left(\frac{n+1}{T}\right) + \sum_{n\geqq n_0} G_1\left(\frac{n}{T}\right) \leqq T\int_{-(n_0-2)/T}^{\infty} G_1(x)dx \\
&\qquad + T\int_{n_0-1/T}^{\infty} G_1(x)dx \to 0, \quad \text{as } n_0\to\infty, \quad (39)
\end{aligned}$$

so that the series in (17) converges.

To establish the continuity of $S(f)$, write

$$S(f) = \sum_{|n| \leqq n_0} \left| H\left(f + \frac{n}{T}\right) \right|^2 + \xi(n_0, f), \quad 0 \leqq f \leqq 1/T. \tag{40}$$

For arbitrary δ, $0 \leqq f \leqq 1/T$,

$$|S(f) - S(f + \delta)| \leqq \left| \sum_{n \leqq n_0} \left[\left| H\left(f + \frac{n}{T}\right) \right|^2 - \left| H\left(f + \delta + \frac{n}{T}\right) \right|^2 \right] \right. \\ + |\xi(n_0, f)| + |\xi(n_0, f + \delta)|. \tag{41}$$

Now since $h(t) \in \mathfrak{L}_1(-\infty, \infty)$, $H(f)$ is continuous. To make the right member of ineq. (41) $\leqq \epsilon$, first let n_0 be sufficiently large so that the last two terms of the right member of ineq. (41) $\leqq \epsilon/2$; then choose $|\delta|$ sufficiently small so that the first term of the right member of inequality (41) $\leqq \epsilon/2$. This establishes the continuity of $S(f)$.

We next verify (37), which concerns $S_\tau(f)$. Since h_τ is in $\mathfrak{L}_1(-\infty, \infty)$, $H_\tau(f)$ exists for all $f \in (-\infty, \infty)$. Thus, $S_\tau(f)$ as defined in (33) is meaningful, though perhaps infinite on a set of measure zero. With $\xi(n_0 f)$ as in (39), write

$$S_\tau(f) - S(f) = \sum_{n=-\infty}^{\infty} \left| H_\tau\left(f + \frac{n}{T}\right) \right|^2 - \sum_{n=-\infty}^{\infty} \left| H\left(f + \frac{n}{T}\right) \right|^2 \\ \geqq \sum_{|n| \leqq n_0} \left[\left| H_\tau\left(f + \frac{n}{T}\right) \right|^2 - \left| H\left(f + \frac{n}{T}\right) \right|^2 \right] - \xi(n_0, f). \tag{42}$$

Now let $\epsilon > 0$ be arbitrary. From (39) we can choose n_0 sufficiently large such that $\xi(n_0, f) \leqq \epsilon/2$, for $f \in [0, 1/T]$. With n_0 so chosen,

$$S_\tau(f) - S(f) \geqq \sum_{|n| \leqq n_0} \left[\left| H_\tau\left(f + \frac{n}{T}\right) \right|^2 - \left| H\left(f + \frac{n}{T}\right) \right|^2 \right] - \frac{\epsilon}{2}. \tag{43}$$

Now let $\tilde{H}_\tau(f)$ be the Fourier transform of $\tilde{h}_\tau$. Then

$$|H(f)| = |H_\tau(f) + \tilde{H}_\tau(f)| \leqq |H_\tau(f)| + |\tilde{H}_\tau(f)|.$$

Therefore

$$|H(f)|^2 \leqq |H_\tau(f)|^2 + 2|H_\tau(f)||\tilde{H}(f)| + |\tilde{H}_\tau(f)|^2 \\ \leqq |H_\tau(f)|^2 + 2\|h_\tau\|_1 \|\tilde{h}_\tau\|_1 + \|\tilde{h}_\tau\|_1^2,$$

where $\| \ \|_1$ denotes $\mathfrak{L}_1$ norm. Since $\|\tilde{h}_\tau\|_1 \to 0$, as $\tau \to \infty$, if τ is sufficiently large, then

$$|H_\tau(f)| - |H(f)| \geqq -\frac{\epsilon}{2} \cdot \frac{1}{(2n_0 + 1)}, \quad f \in (-\infty, \infty). \tag{44}$$

Inequalities (43) and (44) imply that, for all $\epsilon > 0$, there exists a $\tau_0(\epsilon)$ such that for all $\tau \geqq \tau_0(\epsilon)$,

$$S_\tau(f) \geqq S(f) - \epsilon, \quad 0 \leqq f \leqq \frac{1}{T}. \tag{45}$$

Thus, for $\tau \geqq \tau_0(\epsilon)$,

$$\inf_{0 \leqq f \leqq 1/T} S_\tau(f) \geqq \inf_{0 \leqq f \leqq 1/T} S(f) - \epsilon = m - \epsilon. \tag{46}$$

Letting $\tau \to \infty$ and $\epsilon \to 0$ in (46) yields (37).

APPENDIX B

Convergence of the Minimum Distance

In this appendix, we shall verify (21), i.e.,

$$d(h_\tau) \to d(h), \quad \text{as } \tau \to \infty. \tag{47}$$

From the definition of $d(h_\tau)$ (10), for arbitrary $\epsilon > 0$, we are assured of the existence of a $\mathbf{w} = \mathbf{u} - \mathbf{v}$ such that $\mathbf{u}, \mathbf{v} \in \mathcal{A}_N$, and

$$\|h_\tau * \mathbf{w}\| \leqq d(h_\tau) + \epsilon. \tag{48}$$

Repeating the steps in (34), we obtain

$$\|h_\tau * w\|^2 = \int_0^{1/T} S_\tau(f) \,|\sum_k w_k e^{i2\pi kTf}|^2 df \geqq 2[\inf_f S_\tau(f)] \sum_{k=0}^{N-1} |w_k|. \tag{49}$$

From (37) we can choose τ sufficiently large so that

$$\inf_f S_\tau(f) \geqq m. \tag{50}$$

Hence, for such a choice of τ,

$$\sum_{k=0}^{N-1} |w_k| \leqq \frac{[d(h_\tau) + \epsilon]^2}{2m}. \tag{51}$$

Now

$$\begin{aligned} d(h) \leqq \|h * \mathbf{w}\| = \|h_\tau * \mathbf{w} - \tilde{h}_\tau * \mathbf{w}\| &\leqq \|h_\tau * \mathbf{w}\| + \|\tilde{h}_\tau * \mathbf{w}\| \\ &\leqq d(h_\tau) + \epsilon + \|\tilde{h}_\tau * \mathbf{w}\|. \end{aligned} \tag{52}$$

Since

$$(\tilde{h}_\tau * \mathbf{w})(t) = \sum_k h_\tau(t - kT) \cdot w_k,$$

we have, with τ large enough to satisfy (51),

$$\|\tilde{h}_\tau * \mathbf{w}\| \leqq \sum_k |w_k| \, \|\tilde{h}_\tau\| = \|\tilde{h}_\tau\| \sum_k |w_k| \leqq \frac{\|\tilde{h}_\tau\|}{2m} [d(h_\tau) + \epsilon]^2. \tag{53}$$

Combining (52) and (53) yields for τ sufficiently large (and $\epsilon > 0$ arbitrary)

$$d(h) \leqq d(h_\tau) + \epsilon + \frac{\|\tilde{h}_\tau\|}{2m} [d(h_\tau) + \epsilon]^2. \tag{54}$$

Letting $\tau \to \infty$ and $\epsilon \to 0$ yields

$$d(h) \leqq \liminf_{\tau \to \infty} d(h_\tau). \tag{55}$$

The identical argument with h and h_τ reversed yields for all $\tau > 0$, $\epsilon > 0$,

$$d(h_\tau) \leqq d(h) + \epsilon + \frac{\|\tilde{h}_\tau\|}{2m} [d(h) + \epsilon]^2,$$

so that (letting $\tau \to \infty$, $\epsilon \to 0$)

$$\limsup_{\tau \to \infty} d(h_\tau) \leqq d(h). \tag{56}$$

Inequalities (55) and (56) yield (47) or (21), completing the proof.

REFERENCES

1. G. David Forney, "Maximum Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference," IEEE Trans. on Information Theory, *IT-18* (May 1972), pp. 363–378.
2. G. J. Foschini, "Performance Bound for Maximum Likelihood Reception of Digital Data," IEEE Trans. Information Theory, *IT-21* (January 1975), pp. 47–50.
3. J. E. Mazo, "A Geometric Derivation of Forney's Upper Bound," B.S.T.J., *54*, No. 6 (July–August 1975), pp. 1087–1094.

Contributors to This Issue

Thomas J. Aprille, Jr., B.S., 1967, Northeastern University; M.S., 1968, Ph.D., 1972 (Electrical Engineering), University of Illinois; The Magnavox Company, 1968–1970, Bell Laboratories, 1972—. Mr. Aprille is presently involved in the design of new multiplex equipment for the L5 long-haul coaxial system. Member, IEEE, Sigma Xi, Phi Kappa Phi, Tau Beta Pi, Eta Kappa Nu.

Jacques A. Arnaud, Dipl. Ing., 1953, Ecole Supérieure d'Electricité, Paris, France; Docteur Ing., 1963, University of Paris; Docteur es Science, 1972, University of Paris; Assistant at E.S.E., 1953–1955; CSF, Centre de Recherche de Corbeville, Orsay, France, 1955–1966; Warnecke Elec. Tubes, Des Plaines, Illinois, 1966–1967; Bell Laboratories, 1967—. At CSF, Mr. Arnaud was engaged in research on high-power traveling-wave tubes and supervised a group working on noise generators. He is presently studying optical wave propagation. Senior Member, IEEE; member, Optical Society of America.

A. Descloux, Math. Dipl., 1948, Swiss Federal Institute of Technology, Zürich; Ph.D. (Mathematical Statistics), 1961, University of North Carolina; Bell Laboratories, 1956—. Mr. Descloux spent 1955–1956 on the staff of the University of Washington where he taught mathematics and statistics. At Bell Laboratories, he has been concerned chiefly with the application of probability theory to traffic problems. Member, Institute of Mathematical Statistics, American Mathematical Society.

Michael J. Gans, B.S. (E.E.), 1957, Notre Dame University, M.S., 1961, Ph.D. (E.E.), 1965, University of California, Berkeley; Bell Laboratories, 1966—. At Bell Laboratories, Mr. Gans has been engaged in research on antennas for mobile radio and satellite communications.

Donald R. Means, B.E.E., 1961, University of Akron; M.E.E., 1963, and Ph.D., 1972, New York University; Bell Laboratories, 1961–1967, 1968–1969, 1972—. Mr. Means is currently engaged in computer-aided analysis and design of speech networks for telephone sets. Member, IEEE, Sigma Tau, Eta Kappa Nu.

Calvin M. Miller, B.S.E.E., 1963, North Carolina State University; M.S.E., 1966, Akron University; Bell Laboratories, 1967—. Mr. Miller has developed equipment and methods for transmission line charac-

terization. His present interests are in the area of fiber optics as a practical transmission medium. Member, Eta Kappa Nu, OSA.

Gerry Miller, B.E.S., 1963, Ph.D., 1969, Johns Hopkins University; Bell Laboratories, 1969—. Mr. Miller is a member of the Customer Interconnection Planning Department and is concerned with the effects of unbalanced terminal equipment on crosstalk performance and out-of-band signal power requirements. Member, Tau Beta Pi, Eta Kappa Nu, Sigma Xi.

R. A. Semplak, B.S. (Physics), 1961, Monmouth College, N.J.; Bell Laboratories, 1955—. Mr. Semplak, a member of the Antennas and Propagation Research Department, has done research on microwave antennas and propagation. He participated in the Project Echo and *Telstar*® communications satellite experiments. He currently is concerned with the attenuation effects of rain on propagation at 18.5 and 30.9 GHz. Member, Sigma Xi and Commission II of International Scientific Radio Union (URSI).

Jagdeep Shah, B.Sc. (Physics), 1962, University of Bombay, Bombay, India; Ph.D. (Physics), 1967, M.I.T.; Bell Laboratories, 1967—. Mr. Shah's interests include optical properties of semiconductors, high-intensity effects in semiconductors, properties of organic dyes and dye lasers, and photoexcitation of superconductors. Member, American Physical Society.

Aaron D. Wyner, B.S., 1960, Queens College; B.S.E.E., 1960, M.S., 1961, and Ph.D., 1963, Columbia University; Bell Laboratories, 1963—. Mr. Wyner has been doing research in various aspects of information and communication theory and related mathematical problems. He is presently Head of the Communications Analysis Research Department. He spent the year 1969–1970 visiting the Department of Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel, and the Faculty of Electrical Engineering, the Technion, Haifa, Israel on a Guggenheim Foundation Fellowship. He has also been a full- and part-time faculty member at Columbia University and the Polytechnic Institute of Brooklyn. He has been chairman of the Metropolitan New York Chapter of the IEEE Information Theory Group, has served as an associate editor of the Group's *Transactions*, and has served as cochairman of two international symposia. He is presently first vice-president of the IEEE Information Theory Group. Fellow, IEEE, member, AAAS, Tau Beta Pi, Eta Kappa Nu, Sigma Xi.