THE BELL SYSTEM
TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING
ASPECTS OF ELECTRICAL COMMUNICATION

Volume 56 April 1977 Number 4

Information in the Zero Crossings of Bandpass Signals

By B. F. LOGAN, JR.

(Manuscript received October 4, 1976)

An interesting subclass of bandpass signals $\{h\}$ is described wherein the zero crossings of h determine h within a multiplicative constant. The members may have complex zeros, but it is necessary that h should have no zeros in common with its Hilbert transform $\hat{h}$ other than real simple zeros. It is then sufficient that the band be less than an octave in width. The subclass is shown to include full-carrier upper-sideband signals (of less than an octave bandwidth). Also it is shown that full-carrier lower-sideband signals have only real simple zeros (for any ratio of upper and lower frequencies) and, hence, are readily identified by their zero crossings. However, under the most general conditions for uniqueness, the problem of actually recovering h from its sign changes appears to be very difficult and impractical.

I. INTRODUCTION

Voelcker and Requicha[1] raised the question, among others, as to when a bandpass signal $h(t)$ might be recovered (within a multiplicative constant) from sgn $\{h(t)\}$, that is, from its zero crossings. There are really two questions here that should be treated separately: the question of uniqueness and the question of recoverability. Recoverability implies that there is an effective (stable) way of recovering the signal from the data. Uniqueness does not always imply recoverability. For example,

a band-limited signal is uniquely determined from its samples (at slightly greater than the Nyquist rate) just on a half line, say $t < 0$, but there is no stable way of recovering the signal from the half-line samples. However, to demonstrate recoverability we must first establish uniqueness. Here we examine the question of uniqueness.

There are countless ad hoc ways of choosing a subset Z of bandpass signals such that

$$\operatorname{sgn}\{h_1(t)\} \equiv \operatorname{sgn}\{h_2(t)\}, \quad h_1 \in Z, \quad h_2 \in Z \tag{1}$$

implies

$$h_1(t) \equiv Ah_2(t), \tag{2}$$

e.g., by choosing the first member in an arbitrary way and then choosing successive members that have distinct signum functions. However, the subset Z could be considered interesting only if it reveals basic constraints on the sign changes of members of the whole class. Our objective is to illuminate the structure of bandpass functions (signals) having the same signum function.

In connection with (1), we are going to assume that the function $\operatorname{sgn}\{h(t)\}$ has no removable discontinuities,* and, hence, does not mark the location of zeros of even multiplicity. Also, in the context of the problem here we say two functions are *distinct* only if one is not a constant multiple of the other.

We first focus on the problem of constructing distinct bandpass functions having the same signum function. This leads to the concept of the "free" zeros of a bandpass function h.

The free zeros of h are those zeros that may be removed or moved around (by replacing the removed zero with another) without destroying the bandpass property of h. Removing (or moving) any zero of h does not destroy the overall low-pass property of h but may destroy the bandpass property. The simple examples $\sin t$ and $t^{-1} \sin t$ illustrate this fact.

We show that the free zeros of h are simply the common zeros of h and its Hilbert transform $\hat{h}$. These are further identified as common zeros of certain low-pass functions in the representation of h. In case of real-valued $h(t)$, the free zeros of h are conveniently identified in the representation

$$h(t) = \operatorname{Re}\{f(t)e^{i\mu t}\}$$

as the real zeros of f and those complex zeros of f that occur in conjugate

* According to the usual convention, $\operatorname{sgn} 0 = 0$, the function $s(t) = \operatorname{sgn}|\sin^2 t|$ would have removable discontinuities at the zeros of $\sin t$. Here we assume that the function of t, $\operatorname{sgn}|h(t)|$, takes the value 0 only at points where $h(t)$ changes sign.

pairs. Here $f(t)$ is an arbitrary *complex-valued* band-limited function which need not have complex zeros occurring in conjugate pairs.

It follows readily that if a real-valued bandpass $h(t)$ has free zeros other than real simple (free) zeros, then there is a distinct function in the same class having the same signum function as $h(t)$. So, in the absence of some meaningless ad hoc rule, we must restrict our attention to functions that have no free zeros other than real simple zeros if we require sgn $\{h(t)\}$ to determine $h(t)$ within a constant multiplier.

It is possible, however, as shown by an example, for *distinct* bandpass functions to have the same signum function when *neither* has any free zeros. This is possible only when the passband spans an octave or more.

Our main result is that (1) implies (2) when $Z = Z(\alpha,\beta)$ consists of those real-valued $h(t)$ having no free zeros other than real simple free zeros and having spectrum confined to $[\alpha,\beta]$ (and $[-\beta,-\alpha]$), where $0 < \alpha < \beta < 2\alpha$.

The key to this result is the simple identity (37)

$$\hat{h}_1(t)h_2(t) - \hat{h}_2(t)h_1(t) = g(t),$$

where in terms of the representation (9),

$$h_i(t) = p_i(t)\cos \mu t - q_i(t)\sin \mu t \quad (i = 1,2),$$

g is given by

$$g(t) = p_2(t)q_1(t) - p_1(t)q_2(t).$$

Here p_i and q_i are band limited to $[-\lambda/2,\lambda/2]$ and, hence, g is band limited to $[-\lambda,\lambda]$, where $\lambda = \beta - \alpha$. Then, if h_1 and h_2 have enough common zeros $\{t_k\} = S$, we can conclude from

$$g(t_k) = 0 \quad \text{all } t_k \text{ in } S$$

that

$$g(t) \equiv 0,$$

and, hence, that

$$\frac{h_1(t)}{\hat{h}_1(t)} \equiv \frac{h_2(t)}{\hat{h}_2(t)}.$$

Then, if h_1 and h_2 have *no* free zeros, i.e., *no* zeros in common with their Hilbert transforms, we can conclude that

$$h_1(t) \equiv Ah_2(t).$$

The same conclusion can be obtained with some additional argument when

$$h_1(t_k) = h_2(t_k) = 0, \quad \text{all } t_k \text{ in } S$$

is replaced by

$$\operatorname{sgn} h_1(t) \equiv \operatorname{sgn} h_2(t)$$

and h_1 and h_2 are allowed free zeros that are only real and simple.

It is well known that g can have no more zeros, roughly speaking, than $\cos \lambda t$ without vanishing identically. It is also known that h_i must have, roughly speaking, at least as many sign changes as $\cos \alpha t$. In any particular case, all we really need in addition to the free-zero constraint is that h_i has, roughly speaking, more sign changes than $\cos \lambda t$, where λ is the width of the passband. This is always assured, then, when $\alpha > \lambda$, i.e., when $\beta < 2\alpha$, but of course may obtain in other cases.

For the rigorous development of our results we first require some basic definitions.

II. BAND-LIMITED FUNCTIONS

These are restrictions to the real line of entire functions of exponential type, which are bounded on the real line. The standard reference on the subject is the book by Boas.[2] It is convenient to introduce a notation for subclasses of band-limited functions.

Definition: $B_p(\lambda)$, $(1 \leqq p \leqq \infty)$ denotes the collection of functions $f(t)$, $-\infty < t < \infty$, which belong to L_p on the real line and extend as entire functions $f(\tau)$, $\tau = t + iu$, of exponential type λ, $\lambda \geqq 0$. [$B_p(0)$ is empty except for $p = \infty$.]

For $1 \leqq p \leqq 2$, the functions in $B_p(\lambda)$ have ordinary Fourier transforms that vanish outside $[-\lambda,\lambda]$. This follows from the Paley-Wiener theorem[3] for B_2 and the fact that*

$$B_p(\lambda) \subset B_{p'}(\lambda), \quad p' > p. \tag{3}$$

Thus, $B_\infty(\lambda)$ contains $B_p(\lambda)$ for all $p \geqq 1$ and it has been shown[4] that for f in $B_\infty(\lambda)$ (see Appendix),

$$\lim_{T\to\infty} \int_{-T}^{T} \left(1 - \frac{|t|}{T}\right) f(t) e^{-i\omega t} dt = 0, \quad |\omega| > \lambda. \tag{4}$$

So, in a very real sense, the Fourier transforms of functions in $B_\infty(\lambda)$ can be said to vanish outside $[-\lambda,\lambda]$.

III. BANDPASS FUNCTIONS

These are bounded functions whose spectra are confined to the intervals $[\alpha,\beta]$ and $[-\beta,-\alpha]$ where $0 < \alpha < \beta < \infty$.

* See Ref. 2, Theorem 6.7.18, page 102.

Definition: $B_p(\alpha,\beta)$ denotes the class of functions of the form

$$h(t) = f_1(t)e^{i\mu t} + f_2(t)e^{-i\mu t}, \tag{5}$$

where f_1 and f_2 belong to $B_p(\lambda/2)$,

$$\lambda = \beta - \alpha \quad (0 < \alpha < \beta < \infty) \tag{5a}$$

$$\mu = \frac{\alpha + \beta}{2}. \tag{5b}$$

It follows from (3) that

$$B_\infty(\alpha,\beta) \supset B_p(\alpha,\beta), \quad 1 \leqq p < \infty,$$

so we focus on the more general class $B_\infty(\alpha,\beta)$.

Functions of the form (5) have Hilbert transforms (see Ref. 5) $\hat{h}(t)$ given by

$$\hat{h}(t) = -if_1(t)e^{i\mu t} + if_2(t)e^{-i\mu t}. \tag{6}$$

(We could take (6) as the definition of the Hilbert transform of a bounded bandpass function and show that it agrees with the usual definition.) We have

$$h(t) + i\hat{h}(t) = 2f_1(t)e^{i\mu t} \tag{7}$$

$$h(t) - i\hat{h}(t) = 2f_2(t)e^{-i\mu t} \tag{8}$$

We may write (5) and (6) in the forms

$$h(t) = p(t)\cos\mu t - q(t)\sin\mu t \tag{9}$$

$$\hat{h}(t) = p(t)\sin\mu t + q(t)\cos\mu t, \tag{10}$$

where

$$p(t) = f_1(t) + f_2(t); \quad q(t) = if_2(t) - if_1(t). \tag{11}$$

Then for real-valued $h(t)$, we must have p and q real and, therefore,

$$f_2(t) = \overline{f_1(t)}. \tag{12}$$

That is, a real-valued function in $B_\infty(\alpha,\beta)$ is completely described by one complex-valued function f in $B_\infty(\lambda/2)$, $\lambda = \beta - \alpha$, or equivalently by two real-valued functions p and q in $B_\infty(\lambda/2)$; i.e.,

$$h(t) = \text{Re}\,\{f(t)e^{i\mu t}\}, \tag{13}$$

where

$$f(t) = p(t) + iq(t), \qquad p,q \in B_\infty(\lambda/2). \tag{13a}$$

It is sometimes convenient to exhibit one of the end points of the interval $[\alpha,\beta]$ in the exponential factor by writing for (13)

$$h(t) = \text{Re}\,\{f_+(t)e^{i\alpha t}\}, \tag{14}$$

where

$$f_+(t) = f(t) \exp\left\{i\frac{\lambda t}{2}\right\} \equiv x(t) + iy(t) \tag{14a}$$

or

$$h(t) = \text{Re}\,\{f_-(t)e^{i\beta t}\}, \tag{15}$$

where

$$f_-(t) = f(t) \exp\left\{-\frac{i\lambda t}{2}\right\} \equiv r(t) - is(t). \tag{15a}$$

In (14a) f_+ is a function whose spectrum is confined to $[0,\lambda]$ and whose real and imaginary parts x and y belong to $B_\infty(\lambda)$. In (15a) f_- is a function whose spectrum is confined to $[-\lambda,0]$ and whose real and imaginary parts r and $-s$ belong to $B_\infty(\lambda)$. In the form (14), $h(t)$ may be interpreted as the upper single-sideband signal associated with $x(t)$ and carrier frequency α, whereas in (15), $h(t)$ may be interpreted as the lower single-sideband signal associated with $r(t)$ and carrier frequency β. Usually one thinks of y as the Hilbert transform of x and s as the Hilbert transform of r. However, such a relation does not follow without further restrictions on f; e.g., $f \in B_p(\lambda/2)$, $p < \infty$. Because x and y (r and s) are interdependent through p and q, the representation (13) is usually more convenient to work with.

IV. FREE ZEROS OF BANDPASS FUNCTIONS

If h belongs to $B_\infty(\alpha,\beta)$ and $h(\xi) = 0$, then the function

$$g(t) \equiv \frac{at + b}{t - \xi} h(t) \tag{16}$$

for arbitrary (a,b) certainly belongs to $B_\infty(\beta)$, since $g(\tau)$ is an entire function of exponential type β bounded on the real line. However, it does not follow that g belongs to $B_\infty(\alpha,\beta)$. For this reason, it is not so easy to construct distinct bandpass functions having the same signum function.

Definition: A complex (or real) number ξ is said to be a free zero of h if the function g defined in (16) belongs to $B_\infty(\alpha,\beta)$ whenever h belongs to $B_\infty(\alpha,\beta)$.

Theorem 1: *A complex (or real) number ξ is a free zero of a function h in $B_\infty(\alpha,\beta)$ if and only if*

$$h(\xi) = 0$$

and

$$\hat{h}(\xi) = 0.$$

(In other words, the free zeros of h are the common zeros of h and its Hilbert transform).

Proof: In order for the function g in (16) to belong to $B_\infty(\alpha,\beta)$, it must be of the form (5); i.e.,

$$g(t) = g_1(t)e^{i\mu t} + g_2(t)e^{-i\mu t} \tag{17}$$

$$g_1, g_2 \quad \text{in } B_\infty(\lambda/2).$$

And since

$$h(t) = f_1(t)e^{i\mu t} + f_2(t)e^{-i\mu t} \tag{18}$$

$$f_1, f_2 \quad \text{in } B_\infty(\lambda/2),$$

we must have

$$g_1(t) \equiv \frac{at+b}{t-\xi} f_1(t) \quad \text{in } B_\infty(\lambda/2) \tag{19}$$

$$g_2(t) \equiv \frac{at+b}{t-\xi} f_2(t) \quad \text{in } B_\infty(\lambda/2) \tag{20}$$

and, therefore, we must have

$$f_1(\xi) = 0 \tag{21}$$

$$f_2(\xi) = 0; \tag{22}$$

and, hence, from (6)

$$\hat{h}(\xi) = 0. \tag{23}$$

So $h(\xi) = \hat{h}(\xi) = 0$ is a necessary condition for ξ to be a free zero. On the other hand, if

$$h(\xi) = f_1(\xi)e^{i\mu\xi} + f_2(\xi)e^{-i\mu\xi} = 0 \tag{24}$$

and

$$\hat{h}(\xi) = -if_1(\xi)e^{i\mu\xi} + if_2(\xi)e^{-i\mu\xi} = 0, \tag{25}$$

it follows that

$$f_1(\xi) = f_2(\xi) = 0, \tag{26}$$

and, hence, that g_1 and g_2 defined in (19) and (20) belong to $B_\infty(\lambda/2)$ and, therefore, that g defined in (16) belongs to $B_\infty(\alpha,\beta)$. Hence, $h(\xi) = \hat{h}(\xi) = 0$ is a necessary and sufficient condition for ξ to be a free zero of h. In the course of the proof, we have established the following results which we label for future reference.

Theorem 2: If h belongs to $B_\infty(\alpha,\beta)$ and $h(\xi) = \hat{h}(\xi) = 0$, then $g(t) = (at + b)/(t - \xi)h(t)$ belongs to $B_\infty(\alpha,\beta)$ and has the Hilbert transform

$$\hat{g}(t) = \frac{at+b}{t-\xi}\hat{h}(t) \quad \textit{also in } B_\infty(\alpha,\beta).$$

Theorem 3: *The free zeros (if any) of a function of the form (5) are the common zeros of f_1 and f_2, or equivalently, the common zeros of p and q in the representation (9).*

Corollary 3.1: *A real-valued function h of the form (13) has a free zero ξ if and only if*

$$f(\xi) = f(\bar{\xi}) = 0.$$

Corollary 3.2: *If the function $f(\tau)$, $\tau = t + iu$, in (13) is zero-free in either half-plane $u \geqq 0$ or $u \leqq 0$ then h has no free zeros.*

In connection with Corollary 3.2, we note that for f to be zero-free in the (closed) upper half-plane $u \geqq 0$, it is sufficient that $x(t)$ defined in (14a) satisfy

$$x(t) > 0, \quad -\infty < t < \infty. \tag{27}$$

Also, for f to be zero-free in the (closed) lower half-plane $u \leqq 0$, it is sufficient that $r(t)$ defined in (15a) satisfy

$$r(t) > 0, \quad -\infty < t < \infty. \tag{28}$$

This follows from the fact that a function $f_+(\tau)$ bounded and analytic in the upper half-plane may be represented by the Poisson integral[6]

$$f_+(t + iu) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{u}{(t - \xi)^2 + u^2} f_+(\xi) d\xi \quad u > 0.$$

Hence, if $\operatorname{Re}\{f_+(t)\} = x(t) > 0$, then $\operatorname{Re}\{f_+(t + iu)\} > 0$ for $u > 0$. A similar statement holds for functions $f_-(\tau)$ bounded and analytic in the lower half-plane.

Now the role of free zeros in the problem under consideration is made clear by the following:

Theorem 4: *If h_1 is a real-valued function in $B_\infty(\alpha,\beta)$ having a complex free zero $\xi = a + ib$, $b > 0$, or a multiple real free zero ξ, then there is a function h_2 in $B_\infty(\alpha,\beta)$ such that*

$$\operatorname{sgn}\{h_1(t)\} = \operatorname{sgn}\{h_2(t)\}, \quad -\infty < t < \infty$$

and

$$h_2(t) \not\equiv A h_1(t), \quad -\infty < t < \infty.$$

Proof: It follows from Corollary 3.1 that if $\xi = a + ib$, $b > 0$, is a free zero of h_1, then $\bar{\xi} = a - ib$ is also a free zero of h_1. Hence, we may take h_2 to be

$$h_2(t) = \frac{P_2(t) h_1(t)}{(t - \xi)(t - \bar{\xi})} \quad \text{in } B_\infty(\alpha,\beta), \tag{29}$$

where $P_2(t)$ is any polynomial of degree 2 satisfying

$$P_2(t) > 0, \quad -\infty < t < \infty, \quad P_2(t) \not\equiv A(t-\xi)(t-\bar{\xi}). \tag{30}$$

In case ξ is a multiple real free zero of h_1 (i.e., of multiplicity ≥ 2), then (29) is still valid with $\xi = \bar{\xi}$.

The converse of Theorem 4 is not true. We need a condition on how often h_1 and h_2 vanish together.

V. BANDPASS FUNCTIONS WHICH VANISH TOGETHER ON LARGE SETS

Here we would like to investigate the implications of

$$h_1(\tau_k) = h_2(\tau_k) = 0, \quad \text{all } \tau_k \in S, \tag{31}$$

where h_1, h_2 belong to $B_\infty(\alpha,\beta)$ and S is a set of uniqueness for $B_\infty(\lambda)$, $\lambda = \beta - \alpha$. We suppose that (31) does not imply that h_1 or h_2 vanish identically.

Definition: $S = \{\tau_k\}$ is said to be a set of uniqueness for $B_\infty(\lambda)$ if

$$g(t) \quad \text{in } B_\infty(\lambda)$$

and

$$g(\tau_k) = 0 \quad \text{all } \tau_k \in S$$

imply

$$g(t) \equiv 0.$$

We do not assume that h_1 and h_2 are real-valued (on the real axis) and write, using (7), (8), and (11),

$$h_1(t) \pm i\hat{h}_1(t) = \{p_1(t) \pm iq_1(t)\}e^{\pm i\mu t} \tag{32}$$

$$h_2(t) \pm i\hat{h}_2(t) = \{p_2(t) \pm iq_2(t)\}e^{\pm i\mu t}, \tag{33}$$

where p_1, q_1, p_2, q_2 are arbitrary functions in $B_\infty(\lambda/2)$ and $\mu = (\alpha+\beta)/2 > \lambda/2$. Then,

$$\{h_1(t) + i\hat{h}_1(t)\}\{h_2(t) - i\hat{h}_2(t)\} = \{p_1(t) + iq_1(t)\}\{p_2(t) - iq_2(t)\} \tag{34}$$

$$\{h_1(t) - i\hat{h}_1(t)\}\{h_2(t) + i\hat{h}_2(t)\} = \{p_1(t) - iq_1(t)\}\{p_2(t) + iq_2(t)\}. \tag{35}$$

It follows from (34) and (35) that

$$h_1(t)h_2(t) + \hat{h}_1(t)\hat{h}_2(t) = p_1(t)p_2(t) + q_1(t)q_2(t) \in B_\infty(\lambda) \tag{36}$$

$$\hat{h}_1(t)h_2(t) - \hat{h}_2(t)h_1(t) = q_1(t)p_2(t) - p_1(t)q_2(t) \in B_\infty(\lambda). \tag{37}$$

Thus, the functions on the left in (36) and (37), apparently of type $\leq 2\beta$, in fact are of type $\leq \lambda$ as the functions on the right show. Then, from (37), if $h_1(\tau)$ and $h_2(\tau)$ vanish together on a set of uniqueness for $B_\infty(\lambda)$, the functions on the right and left vanish identically. We state this result as a theorem for future reference, using the representation (9).

Theorem 5: *Let h_1 and h_2 belong to $B_\infty(\alpha,\beta)$*

$$h_1(t) = p_1(t)\cos \mu t - q_1(t)\sin \mu t$$

$$h_2(t) = p_2(t)\cos \mu t - q_2(t)\sin \mu t$$

and

$$h_1(t) \not\equiv 0, \quad h_2(t) \not\equiv 0.$$

Then (31) implies

$$\frac{h_1(t)}{\hat{h}_1(t)} \equiv \frac{h_2(t)}{\hat{h}_2(t)} \equiv M(t) \tag{38}$$

and

$$p_1(t)q_2(t) \equiv q_1(t)p_2(t), \tag{39}$$

and, hence, if $q_1 q_2 \not\equiv 0$,

$$\frac{p_1(t)}{q_1(t)} \equiv \frac{p_2(t)}{q_2(t)} \equiv N(t). \tag{40}$$

We should note in connection with (40) that $q_1 \equiv 0$ implies $p_1 \not\equiv 0$ (since $h_1 \not\equiv 0$) and, hence, from (39) that $q_2 \equiv 0$. By symmetry, $q_1 q_2 \equiv 0$ implies

$$\begin{cases} h_1(t) = p_1(t)\cos \mu t, & h_2(t) = p_2(t)\cos \mu t \\ \hat{h}_1(t) = p_1(t)\sin \mu t, & \hat{h}_2(t) = p_2(t)\sin \mu t. \end{cases} \tag{41}$$

We cannot, according to the hypotheses, have $q_1 \equiv 0$ and $p_2 \equiv 0$ (or $p_1 \equiv 0$, $q_2 \equiv 0$), i.e.,

$$h_1(t) = p_1(t)\cos \mu t$$

$$h_2(t) = q_2(t)\sin \mu t$$

for then common zeros of h_1 and h_2 are necessarily common zeros of p_1 and q_2 so that (31) would imply, since p_1 and q_2 belong to $B_\infty(\lambda/2)$, that $h_1 \equiv 0$, $h_2 \equiv 0$, contrary to hypothesis. For a similar reason, $q_1 \equiv 0$ (or $q_2 \equiv 0$) implies that the set S in (31) includes a lot of the zeros of $\cos \mu t$ in (41).

Now the function $M(t)$ in (39) is a meromorphic function, the quotient of two functions in $B_\infty(\alpha,\beta) \subset B_\infty(\beta)$. The zeros of $M(t)$ are zeros of h_1 not common to $\hat{h}_1$. Hence, if h_1 and $\hat{h}_1$ have no common zeros, i.e., if h_1 has no free zeros, then the zeros of $M(t)$ are precisely the zeros of h_1. The zeros of a band-limited function determine the function within an exponential factor which in turn depends on the (actual) spectral end points. It follows from a theorem of Titchmarsh[7] (with an additional minor argument) that the zeros of a function f in $B_\infty(\beta)$ whose spectral end points are *centered* about the origin, i.e., a function whose spectrum

is confined to $[-\beta',\beta']$, $(\beta' \leqq \beta)$ but to no smaller interval, determine the function within a constant multiplier. For such functions,

$$f(t) = At^m \prod_{k=1}^{\infty} \left(1 - \frac{t}{\tau_k}\right), \tag{42}$$

where

$$|\tau_{k+1}| \geqq |\tau_k| > 0$$

with the product converging conditionally (owing to the ordering of the zeros). In particular, (42) holds for a band-limited function which is real-valued on the real axis. Hence, if h_1 in Theorem 5 is real-valued on the real axis and has no free zeros, the zeros of $M(t)$ determine h_1 within a multiplicative constant. $M(t)$ is, in principle at least, determined by any non-null function in $B_\infty(\alpha,\beta)$, say h_2, which vanishes on S. There is by hypothesis at least one such function. We may state this result as follows:

Theorem 6: *Let h_1 and h_2 belong to $B_\infty(\alpha,\beta)$ and be real-valued on the real axis and have no free zeros. Then,*

$$h_1(\tau_k) = h_2(\tau_k) \quad \textit{for all } \tau_k \textit{ in } S,$$

where S is a set of uniqueness for $B_\infty(\lambda)$, $\lambda = \beta - \alpha$, implies (if $h_2 \not\equiv 0$)

$$h_1(t) \equiv Ah_2(t).$$

Actually, for the problem at hand, we are interested in sets S which consist of points $\{t_k\}$, where h_1 and h_2 change sign. If this set has an upper density in excess of λ/π, then it is well known (see Levinson[8] for example) that S is a set of uniqueness for $B_\infty(\lambda)$. So in Theorem 6 we may take S to be any set $\{t_k\}$ where the number $\nu(T)$ of t_k in the interval $(0,T)$ satisfies

$$\limsup_{T\to\infty} \frac{\nu(T)}{T} > \lambda/\pi. \tag{43}$$

Roughly speaking, if h_1 and h_2 just vanish together (not necessarily change sign together) more often than $\cos \lambda t$, then the conclusion follows. We know, furthermore, that real-valued functions in $B_\infty(\alpha,\beta)$ must change sign (on either half line), again roughly speaking, at least as often as $\cos \alpha t$.

Theorem 7 (*from Ref. 9*): *Let h be a real-valued function in $B_\infty(\alpha,\beta)$, $h \not\equiv 0$, and denote by $\sigma(T)$ the number of sign changes of $h(t)$ in the interval $(0,T)$. Then,*

$$\liminf_{T\to\infty} \frac{\sigma(T)}{T} \geqq \alpha/\pi.$$

Hence, we have (since lim sup $\geqq$ lim inf):

Theorem 8: In Theorem 6 if $h_1 \not\equiv 0$, we may always take S to be the set $\{t_k\}$ where $h_1(t)$ changes sign provided $\alpha > \lambda$; i.e., provided $\alpha > \beta/2$.

The necessity of the strict inequality $\alpha > \beta/2$ in Theorem 8 is shown by the one-parameter family

$$h(t;a) = \text{Re}\,\{(1 + iae^{it})e^{it}\}, \quad -\tfrac{1}{2} < a < \tfrac{1}{2}$$
$$= \cos t - a \sin 2t = (1 - 2a \sin t)\cos t. \quad (44)$$

Here, $h(t;a)$ belongs to $B_\infty(1,2)$ and

$$\text{sgn}\,\{h(t;a)\} = \text{sgn}\,\{\cos t\}, \quad -\tfrac{1}{2} < a < \tfrac{1}{2}. \quad (45)$$

Also, $h(t;a)$ has no free zeros, which follows by identifying f in (13) as

$$f(t) = (1 + iae^{it})e^{-it/2},$$

which is clearly zero-free in the closed upper half-plane and, hence, by Corollary 3.2, $h(t;a)$ has no free zeros. Yet all members of the family have the same sign. There are similar examples for $B_\infty(m,n)$, m and n positive integers, $m < 2n$.

If $h_1(t)$ changes sign at $\{t_k\}$, then

$$\text{sgn}\,\{h_1(t)\} \equiv \text{sgn}\,\{h_2(t)\}$$

is a stronger statement than

$$h_1(t_k) = h_2(t_k) = 0.$$

By replacing the latter condition by the former, we can with a little more work obtain the conclusion of Theorem 6 by allowing h_1 and h_2 to have only real, simple, free zeros. (Note that $h(t)$ may have a high-order zero, say at $t = 0$, and yet have only a simple free zero there that would require only that $\hat{h}(t)$ have a simple zero at $t = 0$.) This is the most we could hope for in view of Theorem 4 and the example in eq. (44).

We denote by $Z(\alpha,\beta)$ the class of (real) bandpass functions that have no free zeros other than simple, real, free zeros. That is,

Definition: $Z(\alpha,\beta)$, $0 < \alpha < \beta < \infty$, consists of all real-valued functions $h(t)$ of the form

$$h(t) = \text{Re}\,\{f(t)e^{i\mu t}\},$$

where $\mu = (\alpha + \beta)/2$ and $f(t)$ belongs $B_\infty(\lambda/2)$, $\lambda = \beta - \alpha$, and has no pair of complex conjugate zeros and no real zeros that are not simple.

We should note that $Z(\alpha,\beta)$ includes all real-valued functions in $B_\infty(\alpha,\beta)$ that have only real simple zeros. For if h_0 is such a function, then f_0 in the above representation has no pair of complex conjugate zeros, since these are common zeros of h_0 and $\hat{h}_0$. Similarly, f_0 can have no

multiple real zeros, since these also belong to h_0 and $\hat{h}_0$, and h_0 has only simple real zeros.

Theorem 9: *Let h_1 and h_2 belong to $Z(\alpha,\beta)$. Then,*

$$\operatorname{sgn}\{h_1(t)\} = \operatorname{sgn}\{h_2(t)\}, \quad -\infty < t < \infty$$

implies

$$h_1(t) = Ah_2(t), \quad -\infty < t < \infty$$

provided $\sigma(T)$, the number of sign changes of $h_1(t)$ in $(0;T)$, satisfies

$$(i) \quad \limsup_{T\to\infty} \frac{\sigma(T)}{T} > \frac{\beta-\alpha}{\pi}.$$

Furthermore, (i) is always satisfied if $h_1 \not\equiv 0$ and (ii) $\alpha > \beta/2$.

Proof: We may assume that $h_1 \not\equiv 0$ and, hence, that $h_2 \not\equiv 0$. Otherwise the conclusion is trivially true. Then, since h_1 and h_2 vanish together on a set of uniqueness for $B_\infty(\lambda)$, we have from Theorem 5,

$$\frac{h_1(t)}{\hat{h}_1(t)} \equiv \frac{h_2(t)}{\hat{h}_2(t)} \equiv M(t).$$

The poles and zeros of $M(t)$ are determined by any non-null function in $B_\infty(\alpha,\beta)$ that vanishes at the points of sign change. The zeros of $M(t)$ identify the zeros of h_i that are not common to $\hat{h}_i (i = 1,2)$. That is, the free zeros of h_i are missing. All we have to show is that the locations of the free zeros of, say h_1, are uniquely determined by the zeros of M and the points of sign change of h_1 and h_2. It would then follow that h_1 and h_2 have the same set of zeros, and then the conclusion follows from (42).

Denote by $\{\xi_k\}$ the free (real, simple) zeros of h_1 and by $\{\tau_k\}$ the zeros of M and define

$$\Pi_1(t) = \prod_k \left(1 - \frac{t}{\xi_k}\right) e^{t/\xi_k} \tag{46}$$

$$\Pi_0(t) = \prod_k \left(1 - \frac{t}{\tau_k}\right) e^{t/\tau_k}, \tag{47}$$

where we have assumed, as a matter of convenience in writing, that $\xi_k \neq 0$, $\tau_k \neq 0$. [When the zeros of h_1 are thus separated into two sets, the exponential factors are generally required to make the infinite products in (46) and (47) converge. We could have, for example, $\xi_k = k$, $k = 1,2,\cdots$.] We have by the Hadamard factorization Theorem*

$$h_1(t) = h_1(0)e^{ct}\Pi_0(t)\Pi_1(t) \tag{48}$$

* See Ref. 2, page 22.

for some real c, which is irrelevant to the argument. The $\{\tau_k\}$ appear in conjugate pairs so both $\Pi_0(t)$ and $\Pi_1(t)$ are real-valued. We may assume that $h_1(0) > 0$. Then,

$$\text{sgn}\,\{h_1(t)\} = \text{sgn}\,\{\Pi_0(t)\}\,\text{sgn}\,\{\Pi_1(t)\} \quad -\infty < t < \infty. \tag{49}$$

Now $\Pi_0(t)$ and, hence, sgn $\{\Pi_0(t)\}$ are (in principle) given and sgn $\{h_1(t)\}$ is known [except at even-order zeros of $h_1(t)$]. We have

$$\text{sgn}\,\{\Pi_1(t)\} = \text{sgn}\,\{\Pi_0(t)\}\,\text{sgn}\,\{h_1(t)\} \quad \text{for almost all } t, \tag{50}$$

and since the zeros of $\Pi_1(t)$ are real and simple, (50) defines them uniquely; i.e., they are the points where the function on the right *changes* sign.

VI. THE ZEROS OF FULL-CARRIER LOWER-SIDEBAND SIGNALS

In connection with condition (28), which is a sufficient condition for a function to have no free zeros and, hence, to belong to $Z(\alpha,\beta)$, it is worth noting that the condition is also a sufficient condition for the function to have only real simple zeros. In particular, functions of the form

$$h(t) = \text{Re}\,[\{1 + x(t) - i\hat{x}(t)\}e^{i\beta t}], \tag{51}$$

where

$$|x(t)| < 1 \tag{52}$$

and

$$x,\hat{x}\ \text{(real) belong to } B_\infty(\lambda), \quad 0 \leqq \lambda < \beta, \tag{53}$$

which are called "full-carrier" lower-sideband signals, have only real simple zeros. We have the following more general result.

Theorem 10: Let f be a (non-null) bounded band-limited function whose spectrum is confined to the interval $[-\lambda,\lambda]$, $0 \leqq \lambda < \infty$, but to no smaller interval; i.e., $e^{i\mu t}f(t)$ does not belong to $B_\infty(\lambda)$ for any μ different from zero. Also, let $f(\tau)$, $\tau = t + iu$, be zero free in the closed lower half-plane $u \leqq 0$. Then the zeros of the function h defined by

$$h(t;\mu) = \text{Re}\,\{f(t)e^{i\mu t}\} \tag{54}$$

are real and simple provided $\mu > 0$, or provided $\mu \geqq 0$ if $f \neq$ constant.

Note that $h(t;\mu)$ in (54) need not be bandpass; i.e., we do not require $\mu > \lambda$. The function is just a special kind of band-limited function. The result is independent of λ so long as $\lambda < \infty$. The only significance of λ is to indicate that the spectral end points are centered about the origin.

Proof: The conclusion is trivial for f = constant, so we assume that $f \neq$ constant and, consequently, has an infinite number of zeros. From (42) we have

$$f(t) = f(0) \prod_{k=1}^{\infty} \left(1 - \frac{t}{\tau_k}\right), \tag{55}$$

where the product converges conditionally with the provision $|\tau_{k+1}| \geqq |\tau_k|$. We have by hypothesis

$$\tau_k = a_k + ib_k, \quad \text{where} \quad b_k > 0. \tag{56}$$

Then,

$$\begin{aligned} 2h(t) &= e^{i\mu t} f(t) + e^{-i\mu t} \bar{f}(t) \\ &= e^{i\mu t} f(t) \left\{1 + \frac{\bar{f}(t)}{f(t)} e^{-i2\mu t}\right\}. \end{aligned} \tag{57}$$

We have

$$B(t) \equiv \frac{\bar{f}(t)}{f(t)} = \frac{\bar{f}(0)}{f(0)} \frac{\prod_{k=1}^{\infty} \left(1 - \frac{t}{\bar{\tau}_k}\right)}{\prod_{k=1}^{\infty} \left(1 - \frac{t}{\tau_k}\right)} = \frac{\bar{f}(0)}{f(0)} \prod_{k=1}^{\infty} \frac{\left(1 - \frac{t}{\bar{\tau}_k}\right)}{\left(1 - \frac{t}{\tau_k}\right)}, \tag{58}$$

where the last product converges absolutely since

$$\sum \left(\frac{1}{\bar{\tau}_k} - \frac{1}{\tau_k}\right) = 2 \sum \frac{\text{Im}\{\tau_k\}}{|\tau_k|^2}$$

converges absolutely.* Since

$$\left| \frac{1 - \frac{t + iu}{a_k - ib_k}}{1 - \frac{t + iu}{a_k + ib_k}} \right|^2 = \frac{(a_k - t)^2 + (u + b_k)^2}{(a_k - t)^2 + (u - b_k)^2} < 1 \quad \text{for } u < 0, \tag{59}$$

we have

$$|B(t + iu)| < 1 \quad \text{for } u < 0. \tag{60}$$

Hence,

$$|B(t + iu)e^{-i2\mu(t+iu)}| = e^{2\mu u} |B(t + iu)| < 1 \quad \text{for } u < 0. \tag{61}$$

Therefore, since $f(\tau)$ does not vanish on the real axis nor in the lower half-plane, it follows from (57) and (61) that $h(\tau)$ has no complex zeros.

* See Ref. 2, Theorem 6.3.14, page 86.

Now $\log f(\tau)$ is analytic in the upper half-plane and we may write

$$f(t) = |f(t)|e^{i\varphi(t)}, \tag{62}$$

where

$$\varphi(t) = \text{Im}\,\{\log f(t)\} \quad \text{and (say)} \tag{63}$$

$$0 \leqq \varphi(0) < 2\pi. \tag{64}$$

Then,

$$h(t) = |f(t)| \cos\{\varphi(t) + \mu t\}. \tag{65}$$

Since $f(t)$ does not vanish on the real line, the zeros of $h(t)$ are the zeros of $\cos\{\varphi(t) + \mu t\}$. Now if $\cos\{\varphi(t_k) + \mu t_k\} = 0$, then

$$c'(t_k) \equiv \frac{d}{dt} \cos(\varphi(t) + \mu t)|_{t=t_k} = -\{\mu + \varphi'(t_k)\} \sin\{\varphi(t_k) + \mu t_k\} \tag{66}$$

or

$$|c'(t_k)| = |\mu + \varphi'(t_k)|. \tag{67}$$

We have

$$B(t) = e^{-2i\varphi(t)} \tag{68}$$

and

$$\frac{B'(t)}{B(t)} = -2i\varphi'(t) = \sum \left\{\frac{1}{t - \bar{\tau}_k} - \frac{1}{t - \tau_k}\right\} = -2i \sum \frac{b_k}{|t - \tau_k|^2} \tag{69}$$

or

$$\varphi'(t) = \sum \frac{b_k}{|t - \tau_k|^2} > 0. \tag{70}$$

Hence, if $\mu \geqq 0$,

$$|c'(t_k)| > 0 \tag{71}$$

and, therefore, all zeros of h are real and simple. Since $\text{Im}\,\{f(t)e^{i\mu t}\} = \text{Re}\,\{-if(t)e^{i\mu t}\}$ the conclusion of Theorem 10 also holds for $\text{Im}\,\{f(t)e^{i\mu t}\}$ and, since $\varphi'(t) > 0$ and since the zeros of $\cos\{\mu t + \varphi(t)\}$ and $\sin\{\mu t + \varphi(t)\}$ interlace ($\mu > 0$), we have proved

Theorem 11: If f satisfies the hypotheses of Theorem 10, then the functions

$$h_1(t;\mu) = \text{Re}\,\{f(t)e^{i\mu t}\}$$

$$h_2(t;\mu) = \text{Im}\,\{f(t)e^{i\mu t}\}$$

have all real, simple, interlacing zeros for $\mu > 0$.

In applying Theorems 10 and 11 to functions of the form (51), we set

$$g(t) = 1 + x(t) - i\hat{x}(t).$$

The spectrum of g is confined to an interval $[-\alpha,0]$ (where $\alpha \leqq \lambda$) but not to an interval $[-\alpha,\epsilon]$, where $\epsilon < 0$. Otherwise $\{1 + x(t)\} > 0$ would belong to $B_\infty(\epsilon,\alpha)$ and would, therefore, have an infinite number of sign changes, which is a contradiction. We suppose further that the spectrum is not confined to a smaller interval; i.e., that α is the left end point of the spectrum. We then set

$$f(t) = g(t) \exp\left\{\frac{i\alpha t}{2}\right\},$$

so that f meets the hypotheses of the theorem. Then writing

$$h = \text{Re}\,\{g(t) \exp(i\beta t)\}$$

$$= \text{Re}\left\{f(t) \exp\left(i\beta t - \frac{i\alpha t}{2}\right)\right\}$$

we may state the result as

Corollary 10.1: A function of the form (51) has only real simple zeros when the condition in (53) is replaced by $\beta > \lambda/2 \geqq 0$.

When $\beta > \lambda$, as in (53), h has a Hilbert transform $\hat{h}(t) = \text{Im}\,[\{1 + x(t) - i\hat{x}(t)\}e^{i\beta t}]$. So we have

Corollary 11.1: A function h of the form (51) and its Hilbert transform $\hat{h}$ have only real, simple, interlacing zeros.

We state one more result which follows from the proof of Theorem 10:

Theorem 12: Let f be a bounded (non-null) band-limited function whose spectrum is confined to the interval $[\alpha,\beta]$ but to no smaller interval, and let $f(\tau)$, $\tau = t + iu$, be zero-free in the upper half-plane $u \geqq 0$. Then the phase function $\varphi(t)$, defined uniquely by

(*i*) $\varphi(t)$ *is continuous*

(*ii*) $0 \leqq \varphi(0) < 2\pi$

(*iii*) $f(t) = |f(t)|e^{i\varphi(t)}$

satisfies

(*iv*) $\varphi'(t) \leqq \dfrac{\alpha + \beta}{2}$.

It follows, in particular, if $x(t)$ is a positive function in $B_\infty(\lambda)$ and has a Hilbert transform $\hat{x}(t)$ and

$$\varphi(t) = \tan^{-1} \frac{\hat{x}(t)}{x(t)},$$

$$-\frac{\pi}{2} < \varphi(t) < \frac{\pi}{2},$$

then

$$\varphi'(t) \leqq \lambda/2.$$

(Since x is positive, the smallest interval containing the spectrum of $x + i\hat{x}$ is $[0,\lambda']$ for some $\lambda' \leqq \lambda$.) We note that without further qualification, x must be bounded away from zero in order to obtain a (finite) lower bound for φ', as the example

$$x(t) = 1 + a \cos t, \quad (a = 1 - \epsilon)$$

shows.

VII. DISCUSSION AND CONCLUSIONS

The zeros of a bandpass function h that can be moved around without destroying the bandpass property of h; i.e., the free zeros of h play a key role in the problem here and it is safe to assume that they will be important in other problems. We have shown (Theorem 1) that the free zeros of h are simply the common zeros of h and its Hilbert transform $\hat{h}$ (whether or not $h(t)$ is real). It follows (Theorem 2) that moving a free zero of h simply alters its Hilbert transform in the same way; i.e., only the corresponding (common) zero of $\hat{h}$ is moved.

If we are given a large enough subset S of the zeros of h, then (Theorem 5) S determines $h/\hat{h}$. Without further qualification of S or h, this is all that S determines. If real-valued h has enough sign changes, slightly more (roughly speaking) than $\cos \lambda t$, where λ is the width of the passband (of the whole class), then the zero crossings $\{t_k\}$ constitute a set S which determines $h/\hat{h}$. This, without further qualification, is all the information the zero crossings may convey. If, in addition, it is known that h has no free zeros, then under the stipulated conditions $\{t_k\}$ determines h within a constant multiplier.

If h has free zeros, then we cannot determine (a multiple of) $h(t)$ from sgn $\{h(t)\}$, because (Theorem 4) there are other functions in the same class having the same signum function. In this connection, we note the following:

In the representation

$$h(t;\mu) = \text{Re}\,\{f(t)e^{i\mu t}\}, \quad \mu > \lambda/2,$$

where f is regarded as a fixed (complex-valued) function in $B_\infty(\lambda/2)$ and μ as a parameter, the zero crossings of $h(t;\mu)$ for arbitrarily large μ give no more information about f than for $\mu = 3\lambda/2 + \epsilon$, $\epsilon > 0$ (when the band spans less than an octave). The free zeros of $h(t;\mu)$, which are crucial to identifying h (or f) are invariant with μ.

These results may be generalized and specialized in various ways. We should note a specialization to functions of the form

$$h(t) = \cos \mu t - q(t) \sin \mu t,$$

where $q(t)$, real, belongs to $B_\infty(\lambda/2)$ and $\mu > \lambda/2$. This describes a common sort of phase modulation. Here $h(t)$ has no free zeros because the corresponding function in (13),

$$f(t) = 1 + iq(t),$$

clearly has no real zeros, and if ξ is a complex zero of f we have

$$q(\xi) = i$$

and, hence, since $q(t)$ is real,

$$q(\bar{\xi}) = -i$$

and so

$$f(\bar{\xi}) = 2.$$

Thus (Corollary 3.1) h has no free zeros. Then, if we consider two functions h_1 and h_2 of this form and return to the basic identity (37), we have $p_1 = p_2 = 1$ and

$$\hat{h}_1(t)h_2(t) - \hat{h}_2(t)h_1(t) = q_1(t) - q_2(t),$$

which belongs to $B_\infty(\lambda/2)$ rather than $B_\infty(\lambda)$. Now,

$$h_i(k\pi/\mu) = \cos k\pi = (-1)^k, \quad k = 0, \pm 1, \pm 2, \cdots, \quad (i = 1,2)$$

so $h_i(t)$ has at least as many sign changes as $\cos \mu t$. Thus, if $\mu > \lambda/2$ (just enough for high-pass), the zero crossings $\{t_k\}$ of h_i constitute a set of uniqueness for $B_\infty(\lambda/2)$, which is all we need to conclude that

$$\operatorname{sgn} h_1(t) \equiv \operatorname{sgn} h_2(t)$$

implies

$$q_1(t) - q_2(t) \equiv 0$$

i.e.,

$$h_1(t) \equiv h_2(t).$$

In this case the recovery problem is much simpler than in the general case. Here we are given

$$q(t_k) = \cot \mu t_k, \quad \text{all } t_k \text{ in } S, \quad q \text{ in } B_\infty(\lambda/2),$$

and seek $q(t)$; i.e., knowing $p(t) = 1$ vastly simplifies the problem.

As to generalizations, the results may be extended to bandpass functions that are not bounded (e.g., sample functions of gaussian processes). We can replace $B_\infty(\lambda)$ by $B(\lambda)$, which consists of restrictions to the real line of entire functions of exponential type λ whose growth (on the real line) is less than exponential (see Ref. 7). The zeros of these entire functions have ordinary densities, separately in the right and left half-planes, which are equal and do not exceed λ/π.[7] Hence, sets $\{t_k\}$ of upper density greater than λ/π constitute sets of uniqueness for $B(\lambda)$.

It is clear from the Hadamard factorization

$$f(t) = f(0)e^{ct} \prod_{k=1}^{\infty} \left(1 - \frac{t}{\tau_k}\right) e^{t/\tau_k}$$

that the zeros $\{\tau_k\}$ of real-valued f in $B(\lambda)$ determine f within a constant multiplier. Since the τ_k occur in conjugate pairs, the product is real-valued on the real axis and, hence, the exponent c must be real. Then c will be determined by the condition that the growth on the real axis be less than exponential.

Then we define $B(\alpha,\beta)$ analogous to $B_\infty(\alpha,\beta)$, and for h in $B(\alpha,\beta)$, we let h be defined by the right-hand side of (6) with f_1 and f_2 in $B(\lambda/2)$, and simply call it the generalized Hilbert transform of h. It is not important what we call it; the free zeros of h are still the common zeros of h and $\hat{h}$, or equivalently the common zeros of p and q. Then Theorem 7 must be generalized to $B(\alpha,\beta)$. It is clear that the proof in Ref. 9 extends easily, so all the uniqueness results may be extended to $B(\alpha,\beta)$.

In connection with this generalization, it might be interesting to study the free zeros of sample functions of bandpass gaussian processes $\{h\}$. The free zeros are going to be very rare (in the ergodic case) to say the least. It may be advisable to begin the study with the case of periodic sample functions.

There are still other questions that arise in connection with the problem considered here. For example, we have not shown that given an arbitrary real h in $B_\infty(\alpha,\beta)$ there is a corresponding function in $Z(\alpha,\beta)$ having the same signum function. The difficulty occurs when h has an infinite number of free zeros which, for example, may be complex and restricted to the right half-plane and have positive density there. (Such functions can be constructed on the basis of Corollary 3.1.) The Hadamard product composed of the free zeros will then not even belong to the broader class $B(\cdot)$ just discussed. The remaining zeros will not have equal densities in the right and left half-planes and, hence, the Hadamard product composed of these zeros will not belong to $B(\cdot)$. There seems to be no way to replace the free zeros with non-free zeros and obtain a function in $Z(\alpha,\beta)$ with the same signum function as h. It appears

that the argument can be completed to prove that the proposition is false.

Another problem is that of characterizing those h for which there is not another distinct function in the *whole* class $B_\infty(\alpha,\beta)$ having the same signum function. Of course, h must belong to $Z(\alpha,\beta)$ but now other arguments of a Fourier nature are required. The end points of the spectrum play an important role in this problem. For example, $\cos \alpha t$ and $\cos \beta t$ are special functions in $Z(\alpha,\beta)$ which for $\beta < 2\alpha$ meet the conditions of the problem, a result we state without proof. It appears that the "full-carrier" sideband signals, which have spectrum at one or the other end points, are also special functions of this type when $\beta < 2\alpha$. The decay of $h(t)$ also enters in the problem; i.e., $(1 + t^2)h(t)$ must not belong to $B_\infty(\alpha,\beta)$. The basic idea is that one should not be able to multiply $h(t)$ by a positive function and obtain a function in $B_\infty(\alpha,\beta)$. This obviously will be possible if the spectrum of h is confined to $[\alpha',\beta']$ (and $[-\beta',-\alpha']$), where $\alpha < \alpha' < \beta' < \beta$.

Given $h(t)$ in a form other than (13) with an explicit factorization of f, it is obviously difficult to determine whether or not h belongs to $Z(\alpha,\beta)$. However, it is easy to synthesize functions in $Z(\alpha,\beta)$, e.g., the full-carrier sideband signals.

The problem of actually recovering functions in $Z(\alpha,\beta)$ from their zero crossings appears to be difficult (to say the least) under the most general conditions for uniqueness. A general "method" suggested by the proof of Theorem 9 requires first finding any non-null test function in $B_\infty(\alpha,\beta)$ that merely vanishes at the points of sign change or some subset of the points that constitute a set of uniqueness for $B_\infty(\lambda)$. However, this in itself is a difficult, if not intractable, problem except in the simple periodic case. Assuming such a test function to be found, it will, in general, have complex free zeros and/or real free zeros which the sought after function h does not have. So, in effect, the test function and its Hilbert transform must be factored to discard common zeros (free zeros), which amounts to finding the zeros of $M(t)$ in (38), or the poles and zeros of $N(t)$ in (40). Then one constructs as in the proof of Theorem 9, a function $\Pi_0(t)$ with the zeros of $M(t)$, i.e., the non-free zeros of h. Then the missing (real simple) free zeros of h can be determined by comparing the sign changes of $\Pi_0(t)$ and the given sign changes of $h(t)$ as in (50).

The overall recovery procedure is obviously hopeless except in the case of periodic functions. There may be some simpler procedure under more restrictive hypotheses; e.g., condition (27), ensuring that f be zero-free in the closed upper half-plane. Condition (28), ensuring that f be zero-free in the closed lower half plane, was shown to imply that the corresponding h (e.g., a full-carrier lower-sideband signal) has all real simple zeros, in which case h can be recovered by forming an infinite product having simple zeros at the points of sign change. This fact, aside from

questions of practicality, might suggest a preference in full-carrier sideband transmission for the lower sideband.

The results here have theoretical interest in that they provide a satisfactory answer to the general question as to what information (in our sense) is conveyed by the zero crossings of bandpass functions. As far as practicality is concerned, the results cannot be extrapolated with abandon to "almost bandpass" functions. Although there is no argument with the assertion that practical signals can be closely approximated with bandpass signals, it does not follow that there even exists a bandpass signal (to which the results apply) with the same zero crossings as the practical signal, much less one which has the same zero crossings and is everywhere close to the practical signal. Clearly one must have a very severely constrained class of signals in order to assert that the zero crossings "closely" determine the signals.

APPENDIX

Here we sketch a proof of the fact

$$\lim_{T\to\infty} \int_{-T}^{T} \left(1 - \frac{|t|}{T}\right) f(t)e^{-i\omega t}dt = 0, \quad \text{for } |\omega| > \lambda, \quad f \text{ in } B_\infty(\lambda). \tag{72}$$

First we set

$$h(t) = h(t;\omega) = f(t)e^{-i\omega t}, \quad f \text{ in } B_\infty(\lambda) \tag{73}$$

and observe that for $|\omega| - \lambda = \alpha > 0$ (ω = real), h belongs to the class $H_\infty(\alpha)$ consisting of all bounded functions h (high-pass functions) satisfying

$$\int_{-\infty}^{\infty} g(t)h(t)dt = 0 \quad \text{all } g \text{ in } B_1(\alpha). \tag{74}$$

Indeed, $f(t)g(t)$ belongs to $B_1(\alpha + \lambda)$ and, hence, its Fourier transform is continuous and vanishes outside $(-\beta,\beta)$, $\beta = \alpha + \lambda$.

Let us then define

$$C(t;T) = 1 - \frac{|t|}{T}, \quad |t| \leqq T \tag{75}$$
$$= 0, \quad |t| > T.$$

Then we wish to prove

$$\lim_{T\to\infty} \int_{-\infty}^{\infty} C(t;T)h(t)dt = 0, \quad h \text{ in } H_\infty(\alpha) \quad (\alpha > 0). \tag{76}$$

There are several ways to prove this. In Ref. 4 we used the notion of the "unbiased" integral of $h(t)$, denoted by $h^{(-1)}(t)$, which is a particular integral of h also belonging to $H_\infty(\alpha)$. In general, we may define the nth unbiased integral of h by

$$h^{(-n)}(t) = \int_{-\infty}^{\infty} h(x)K_n(t-x)dx, \quad n = 1, 2, \cdots,$$

$$h \text{ in } H_\infty(\alpha) \tag{77}$$

where K_n is any kernel of L_1 whose Fourier transform satisfies

$$\int_{-\infty}^{\infty} K_n(t)e^{-i\omega t}dt = (i\omega)^{-n} \quad \text{for } |\omega| \geq \alpha. \tag{78}$$

Then we can show that $h^{(-n)}$ in fact does satisfy

$$\int_a^b h^{(-n)}(t)dt = h^{(-n-1)}(b) - h^{(-n-1)}(a)$$

$$(-\infty < a < b < \infty). \tag{79}$$

It suffices to show this for $n = 0$, $h^{(0)} \equiv h$, and then use induction.

Achieser[10] shows (in another context) that the minimal L_1-norm kernels have norm

$$\|K_n\|_1 = \alpha^{-n}\frac{4}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^{k(n+1)}}{(2k+1)^{n+1}} = \alpha^{-n}M_n. \tag{80}$$

Then we have, integrating twice by parts,

$$\int_{-\infty}^{\infty} C(t;T)h(t)dt = \frac{1}{T}\{-2h^{(-2)}(0) + h^{(-2)}(T) + h^{(-2)}(-T)\} \tag{81}$$

and, hence,

$$\left|\int_{-\infty}^{\infty} C(t;T)h(t)dt\right| \leq \frac{4M_2}{\alpha^2 T}\sup_t |h(t)|, \quad h \text{ in } H_\infty(\alpha). \tag{82}$$

Then (76) and (72) follow from (82).

Actually, for h of the form (73) we can replace M_2 in (82) by 1. For h having one-sided spectrum, say the half line $[\alpha,\infty)$, $\alpha > 0$, we only require

$$\int_{-\infty}^{\infty} K_n(t)e^{-i\omega t}dt = (i\omega)^{-n} \quad \text{for } \omega \geq \alpha > 0. \tag{83}$$

Here we may obtain the minimal-norm kernels simply by making their

Fourier transforms even about $\omega = \alpha$. It then follows from convexity that

$$K_n(t) = (i)^{-n} p_n(t) e^{i\alpha t}, \quad \text{where } p_n(t) > 0 \tag{84}$$

and

$$\int_{-\infty}^{\infty} |K_n(t)| dt = \int_{-\infty}^{\infty} p_n(t) dt = \alpha^{-n}. \tag{85}$$

In general, if one defines a class of bounded functions having a spectral gap (a,b) by an orthogonality condition similar to (74), then a simple modification of the proof gives the gratifying result that their Fourier integrals are actually summable $(C,1)$ to zero in the gap.

REFERENCES

1. H. Voelcker and A. Requicha, "Clipping and Signal Determinism: Two Algorithms Requiring Validation," IEEE Trans. Commun., June 1973, pp. 738–744.
2. R. P. Boas, Jr., *Entire Functions*, New York: Academic Press, 1954.
3. R. Paley and N. Wiener, *Fourier Transforms in the Complex Domain*, Amer. Math. Soc. Colloq. Publ., *XIX*, New York, 1934. Reprinted, Providence, R.I., 1960, p. 13.
4. B. F. Logan, "Integrals of High-Pass Functions," unpublished work.
5. P. L. Butzer and R. J. Nessel, *Fourier Analysis and Approximation*, Vol. 1, New York and London: Academic Press, 1971 (Hilbert Transforms, pp. 303–354).
6. K. Hoffman, *Banach Spaces of Analytic Functions*, Englewood Cliffs, N.J.: Prentice-Hall, 1962, p. 128.
7. E. C. Titchmarsh, "The Zeros of Certain Integral Functions," Proc. London Math. Soc., (2), *25*, 1926, pp. 283–302.
8. N. Levinson, *Gap and Density Theorems*, Amer. Math. Soc. Colloq. Publ., *XXVI*, New York, 1940, pp. 25–26.
9. B. F. Logan, "Properties of High-Pass Signals," Doctoral Dissertation, Elec. Eng. Dept., Columbia U., New York, 1965.
10. N. I. Achieser, *Theory of Approximation*, New York: Ungar, 1956, p. 189.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 56, No. 4, April 1977
Printed in U.S.A.

Pseudorandom Noise Loading for Noise Power Ratio and Law of Addition

By R. H. MOSELEY, S. B. PIRKAU, and R. B. SWERDLOW

(Manuscript received July 29, 1976)

A noise-loading measurement set is described, which uses coherent averaging to measure intermodulation noise that lies below thermal noise. This technique also lends itself to a new way of measuring the addition of intermodulation noise over a string of repeaters. Noise loading and law of addition data are presented.

I. INTRODUCTION

Noise loading has been used for many years to test wideband communications systems.[1] In this technique, the system load is simulated by thermal noise of the same bandwidth and total power as the actual load. This is particularly effective for equipment designed to carry hundreds of telephone circuits, because the multiplexing together of many independent circuits naturally leads to a total signal that is gaussian. Small frequency bands (notches), usually a few kilohertz wide, are eliminated from the thermal noise load at several places in the band. This noise signal is applied to the device under test, and the noise that appears at the notch frequencies after the device is detected and measured. The noise in these narrow notches is a fairly accurate measure of the circuit noise that a customer would overhear. This noise may be broken down into several components: thermal, intermodulation, interference, and so on. In modern, long-haul, repeatered telephone systems, intermodulation noise is usually a big contributor that must be held at or below the thermal noise, which limits the power at which each repeater may be operated. Until now, intermodulation noise was usually measured by raising the output power until intermodulation became far larger than the thermal noise. This is possible since intermodulation noise increases with an increase of input power and thermal noise does not. The difficulty with this approach is that the noise is then being measured outside of the equipment-design range. Another important factor in estimating the total intermodulation noise for many tandem

repeaters is that, since intermodulation depends on the input signal, there is the possibility of coherent addition of this noise between repeaters. The degree of coherency, or law of addition, has been estimated in the past by connecting several repeaters in tandem and measuring the intermodulation noise of the string. Since many repeaters (about 12 or more) are needed to estimate the addition accurately, this is difficult to do in the laboratory.

In 1971, work was started on a new noise-loading scheme to measure intermodulation noise that is below thermal noise and to estimate the law of addition. The scheme replaces the thermal-noise source by one that is periodic but appears to be noise-like: a pseudorandom noise source. The intermodulation noise is produced from the signal through the nonlinearity of the repeater. It has the same basic period as the input signal and so may be coherently detected and summed at the device output. As successive periods are summed, the intermodulation noise will add on a voltage basis, but the thermal noise will add on a power basis. Thus, a 3-dB enhancement is achieved with each doubling of the number of periods of the source.

A test set to perform this measurement was completed in 1973. The basic layout, special features, and test results of the instrument are described in Section II. Estimation of intermodulation noise for many repeaters in tandem, using the pseudorandom noise loading (PRNL) test set, is described in Section III, and test results are presented. The chief advantages of this test set are summarized in the conclusions, Section IV.

II. OVERVIEW OF OPERATION

The pseudorandom noise loading (PRNL) test set is comprised of three main units: a transmitter section to generate the broadband noise source, a receiver section that includes an analog-to-digital converter and memory for coherent addition of the measured signal, and a master timing unit to achieve repetitive and coherent operation of the transmitter and receiver sections. The noise signal may be applied to the test device at either baseband, IF or RF (4 GHz).

The power in the notch is preselected at the output of the device being tested by the PRNL receiver. The receiver has a noise figure of 10 dB or better at IF and 8 dB at baseband. Due to these low noise figures, the noise floor below which signal averaging is required is usually set by the noise figure of the test device. Up to 42-dB improvement in signal to noise is possible by the averaging process, corresponding to the summation of 16,384 successive periods of the pseudorandom test signal. Ninety seconds of averaging is required to achieve the 42-dB improvement.

The ultimate capability is determined by noise in the notch that is

Table I — Attainable notch depth

Baseband	Signal Power (dBm)	Notch Depth (dB)	Notch Frequencies (MHz)
380 kHz to 20 MHz IF	0	80	3.886, 11.7, 17.842
60 to 80 MHz RF	2	96	63.886, 71.7, 77.842
Any 20-MHz channel in the 3.7- to 4.2-GHz band	6	35	No notch filters at RF; notches at baseband and IF are translated to the RF channel

coherent with the signal. This is due to leakage of signal through the notch filter and intermodulation from the transmitter at higher output power. These limits are given in Table I.

The notch depth is enhanced when the test device has gain. The gain allows source levels to be low, which, in turn, reduces intermodulation from active devices in the source. A notch depth limit of 120 dB due to filter characteristics can be realized at IF and RF with sufficient gain.

2.1 PRNL transmitter

Shown in Fig. 1 is a block diagram of the PRNL transmitter. The noise source is a 25-stage shift register operating at 300 MHz. The output is a very long pseudorandom digital sequence at the bit rate of 300 Mb/s. The digital signal is first band limited from 380 kHz to 20 MHz by bandpass filter BP1 to form a noise source having statistics close enough to gaussian to approximate the telephone load. Reference 2 gives the

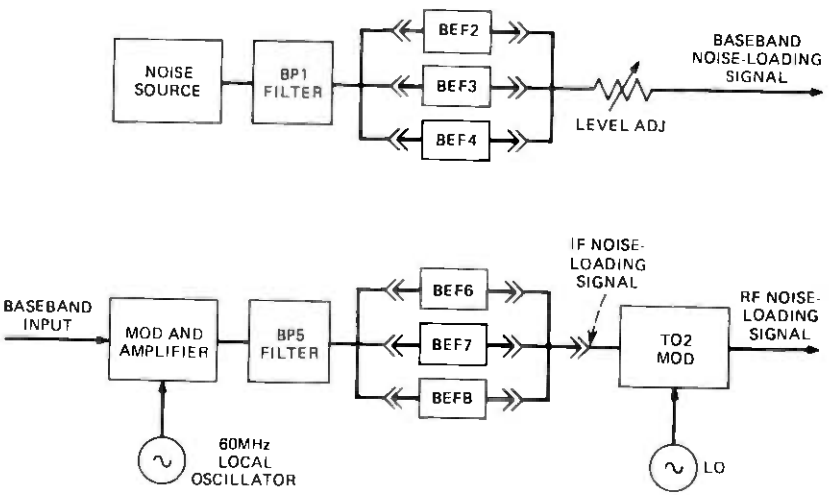


Fig. 1—PRNL transmitter.

statistics of the signal for various filter bandwidths. As a comparison, separate noise-loading curves were run on the same amplifier, using thermal and pseudorandom noise sources. Within the range where the intermodulation noise was sufficiently above thermal noise level, the two results agreed to 0.1 dB, which is within the measurement accuracy of the test.

Next, one of three band-elimination filters (BEF2, BEF3, and BEF4) is selected. These filters are narrow band-stop filters with an 80-dB rejection band, 6 kHz wide, centered at 3.886, 11.7, and 17.842 MHz, respectively. The signal is next level adjusted at baseband. Lowering the signal level before all the active devices in the signal path assures lowest intermodulation from the test set and, consequently, the deepest notches. This signal may be used to load at baseband. To load at IF, the signal is then up converted to an IF center frequency of 70 MHz by a modulator and 60-MHz local oscillator (LO).

The IF signal is first amplified and then passed through BPF5. This filter passes the upper sideband while attenuating the lower sideband by 60 dB and the LO signal by 90 dB. At this point in the transmitter, the notch depth (for a broadband power of 0 dBm) has been reduced to 50 dB due to intermodulation in the modulator.

To obtain a notch depth of 120 dB (at $f_c \pm 2$ kHz, where f_c is the notch-center frequency), the proper IF band-stop filter (BEF6, 7, or 8) is selected to align in frequency with the translated baseband notch. The IF notch frequencies are 63.886, 71.7, and 77.842 MHz. As shown in Fig. 1, this signal is applied to the device under test for IF noise loading. When an RF (4-GHz) noise-loading signal is required, an up converter and the associated tuners and filters are used to translate the noise load to RF. Notch depths at RF are limited to 46 dB by the up converter for broadband power levels of 0 dBm.

2.2 PRNL receiver

A block diagram of the receiver is shown in Fig. 2. For detection of a 4-GHz signal, RF preselection provides a 1-MHz signal centered on the notch frequency. This signal is then attenuated and down converted to IF. Attenuation is necessary to insure that the intermodulation products produced in the down converter are at least 20 dB lower than the signal to be measured. To insure coherency, the same LO source used to up convert to RF must be used to down convert to IF.

The first function performed at IF is that of preselection by narrowband (±1.5-kHz) crystal filters. Since the signal being preselected may be 80 dB lower than the loading signal, multistage preselection and amplification is used. The multistage process produces a gain of 75 dB and out-of-band rejection of 120 dB. The noise figure of the IF receiver

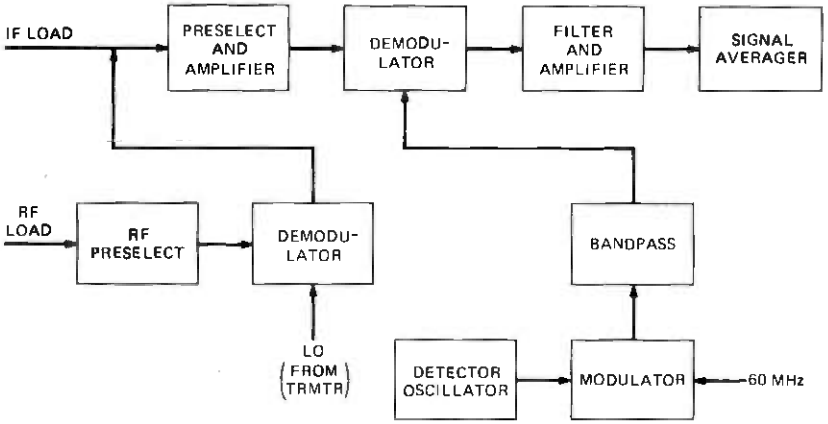


Fig. 2—PRNL receiver.

is determined by the preselect filter which varies between 5.4 and 8 dB of loss, depending on the frequency of preselection. The preselected signal is then down converted to baseband in a single stage of demodulation. The LO for this stage of demodulation is generated by mixing the IF transmitter LO (60 MHz) and the detector oscillator (set to 6 kHz below the baseband notch frequency) and selecting the sum frequency. The resulting notch signal is centered at 6 kHz, whereupon it is filtered and detected.

2.3 Final detection

The final detection process is performed by the signal-averaging unit. This is a commercial unit slightly modified to meet the specific requirements of PRNL. The functions performed by the averager are analog-to-digital conversion, storage, summation of the digitized signal, and the ability to play back the stored signal in real time for measurement with a power meter. The analog-to-digital converter takes a 9-bit sample of the input signal every 20 μs and stores it in an 18-bit register. Since there are 1024 word locations in the register, a repetitive signal with a period of 20 ms could be handled. In practice, a period of 5.28 ms is used and 256 samples taken during each period. The number, N, of successive sweeps by the averager is preselected for values of M ($2^M = N$) up to 14. The coherency of the detection process is preserved by starting successive averager sweeps at the start time of the pseudorandom source and at a fixed phase of the detection oscillator. This timing is performed by the master timing circuit described in Section 2.4.1. The digitized sample points are summed to corresponding locations in memory for each sweep.

Finally, the stored word in the averager can be played back in real time.

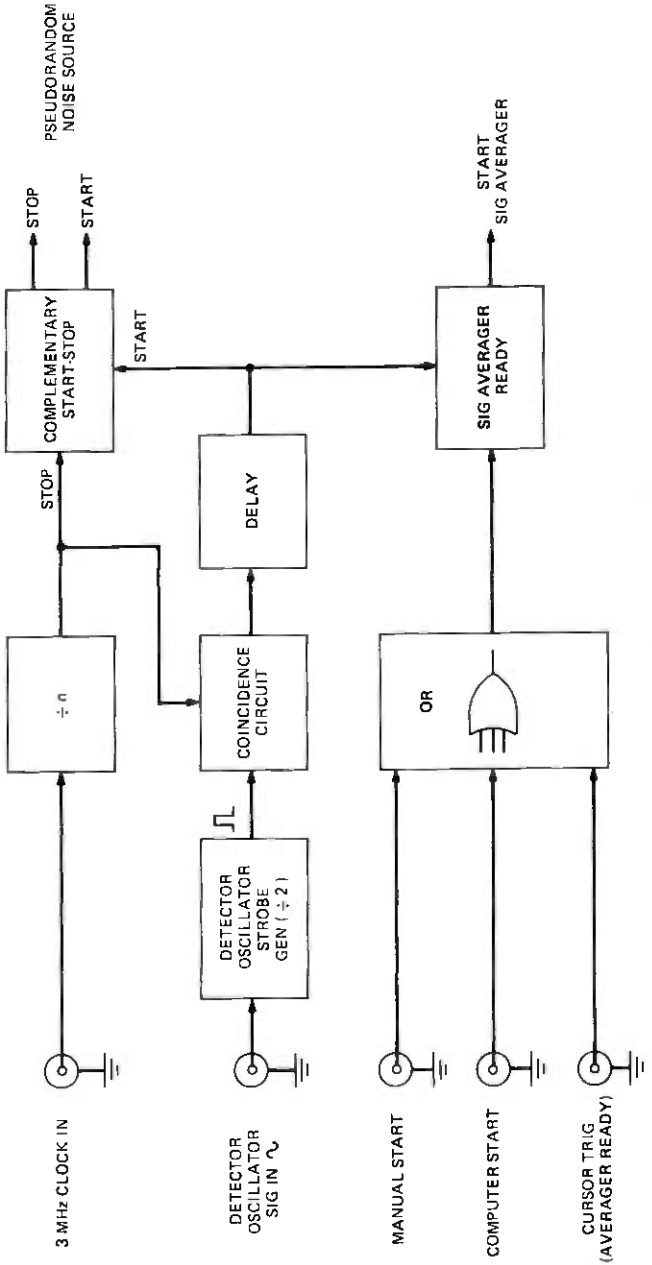


Fig. 3—Master timing circuit.

The real-time play-back signal can be displayed on an oscilloscope, applied to a spectrum analyzer, or to a power meter. The signal averager is also interfaced with a computer. The computer can be used to control the measurement and to further process the data. The computer interface is described in Section 2.4.2.

2.4 Special features

2.4.1 Master timing circuit

The heart of the noise loading set is the master timing circuit. The three major functions of the circuit are to

(*i*) Lock the start of the pseudorandom data transmitter and signal averager to a common oscillator (detector oscillator), which demodulates the signal in the notch to baseband.

(*ii*) Start the signal averager.

(*iii*) Start-stop the pseudorandom data transmitter.

Of the three functions listed above, synchronization (lock) is most important since it is essential for coherent signal addition. A simplified block diagram of the master timing circuit is shown in Fig. 3.

The critical section of this circuit is the coincidence function. The coincidence circuit phase-locks the relative high-frequency notch-detector strobe pulse (3.6 to 17.8 MHz) to the low-frequency divide-by-*n* counter output pulse (approximately 200 Hz). The resulting output pulse is used to initiate functions (*ii*) and (*iii*) above, thus insuring coherent operation of the PRNL set.

A second PRNL set, with 6-GHz capabilities, is under construction. This set is discussed in more detail in Appendix C. The major difference (besides the 6-GHz capability) is the absence of the above described coincidence circuit. In the new PRNL set, the LO as well as all notch-detect LOs are derived from one master oscillator, thereby assuring a constant phase relationship between all LO signals. The disadvantage of this design is, of course, less flexibility due to fixed notch-detect LO frequencies.

2.4.2 Computer interface

Another one of the features of the PRNL set is its operating capability through computer control, which permits unattended operation and data processing for law of addition measurements. The setup is illustrated in Fig. 4.

The program that operated the noise-loading set could operate either through a teletype request or automatically. In either mode, the computer sends a start command to the PRNL set. The PRNL set synchronizes this computer start command with its internal start commands and

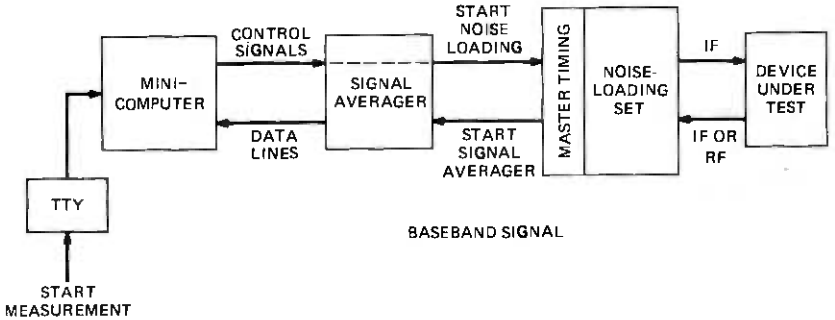


Fig. 4—Computer-controlled noise-loading data collection.

executes a measurement. Upon completion of the measurement, the PRNL set sends a finish command to the computer which, in turn, initiates a data transfer from the memory of the signal averager to the computer disk memory for storage. The automatic mode of operation was used to make repetitive measurements through the use of the real-time clock of the computer. In this mode, a measurement can be requested, either at a specified time, or at a specific interval.

2.5 Calibration

The purpose of calibrating the noise-loading test set is to obtain an absolute measure of the intermodulation signal in a notch at the output of the device under test and to establish the signal power density near the notch frequency. The difference in these measurements yields the noise-power ratio (NPR) which is a measure of the linearity of the device.

The precise calibration relates the power as measured at the signal-averager output for one cycle of the averager to the power in the notch at the output of the test device. To do this, the receiver was calibrated to determine the equivalent noise bandwidth of each set of preselect filters, and the gain of the receiver was measured for the center frequency of each set of preselection filters.

The intermodulation signal in the notch at the output of the device under test is determined by level adjusting the power as measured at the signal-averager output by the measured receiver gain and the amount of averaging performed. The signal-power density near the notch frequency is found by measuring the total output power at the output of the test device and subtracting the selectivity of the receiver.

2.6 NPR results

Two important features of the PRNL test set are the averaging feature and the multifrequency capability. Both are important for characterizing intermodulation noise in a repeater. The averaging feature permits NPR

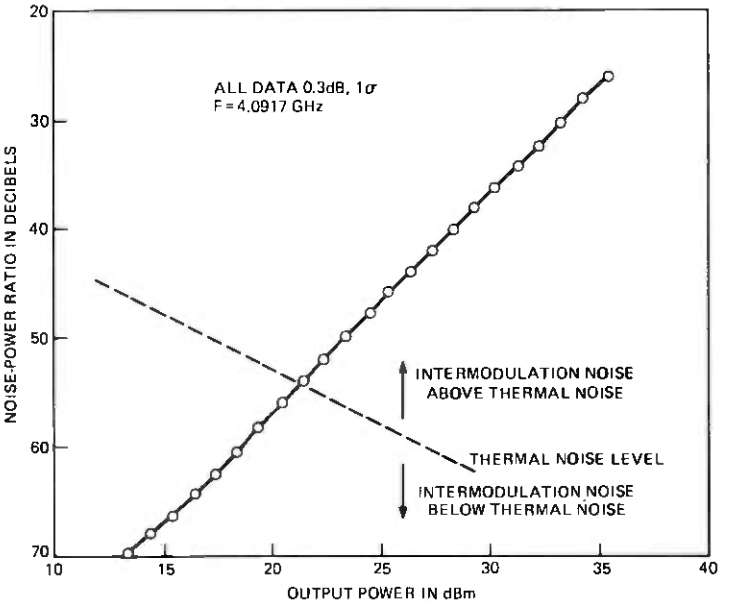


Fig. 5—Noise loading results for a 4-GHz TWT.

measurements below thermal noise. The multifrequency feature allows measurements on portions of the repeater so that the effect of these parts can be seen on the overall NPR performance. The data shown below illustrate both these points.

Examples of data showing the averaging and multifrequency capabilities of the set are shown in Figs. 5 and 6. In Fig. 5, the device under

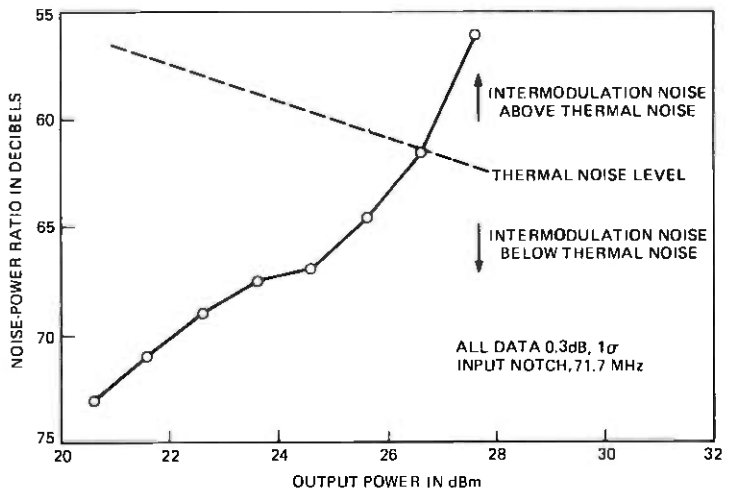


Fig. 6—Noise loading results for an SSB radio transmitter (70-MHz input, 4-GHz output).

test was a 4-GHz traveling-wave tube (TWT) measured RF-to-RF. The figure shows the measured NPR (ordinate) as a function of the TWT output power (abscissa). At an output power of 21.4 dBm, the thermal-noise power in the notch equalled the intermodulation (IM) power in the notch. The test set residual IM was 50 dB lower due to the high TWT gain. The 2:1 slope indicates dominance of third-order intermodulation products. Multifrequency capability of the PRNL set is illustrated in Fig. 6. Shown are the results of an IF-to-RF noise-loading measurement on an experimental single-sideband transmitter that includes a TWT. In this measurement, the thermal and intermodulation powers in the notch are equal for an output power of 28.5 dBm. The NPR performance of the additional units (besides the TWT) that make up a transmitter account for the noticeable change in the results shown in Figs. 5 and 6.

III. TANDEM PERFORMANCE AND THE LAW OF ADDITION

Telephone transmission over more than a few miles requires repeaters to maintain adequate speech volume. Since intermodulation is produced by the signal at each repeater, intermodulation from tandem repeaters may be coherent and add on a voltage basis. For instance, in a system of 150 repeaters, the difference in output intermodulation power at the end of the string between coherent and random addition is 22 dB. It would be helpful to have a way of quantifying and predicting this addition for a few repeaters in the laboratory, since direct measurement of many tandem repeaters in the laboratory or field can be expensive.

We show below that the intermodulation power at the end of a tandem string is related only to the crosscorrelation of the intermodulation signals from all pairs of repeaters. Therefore, the crosscorrelation can be used as a quantitative measure of the buildup of intermodulation noise. Since the PRNL set can measure the crosscorrelation, when linked to a computer, it is a tool that we can use to predict the addition of intermodulation products for many similar repeaters by measurements on a few repeaters in the field or laboratory. Data are shown to substantiate the relationship between total intermodulation noise power and the crosscorrelation. The theory and results here hold only for linear systems. Digital and band-limiting repeaters using frequency modulation (FM) cannot be handled by these techniques.

3.1 The model for intermodulation addition

Figure 7 is a schematic representation of a model of a repeater used to study intermodulation noise addition for linear repeaters. The assumption is made that the repeater can be split into two parts: a purely linear part (linear filter) followed by a purely nonlinear part (gain). This is not the most general case, but works well for microwave radio re-

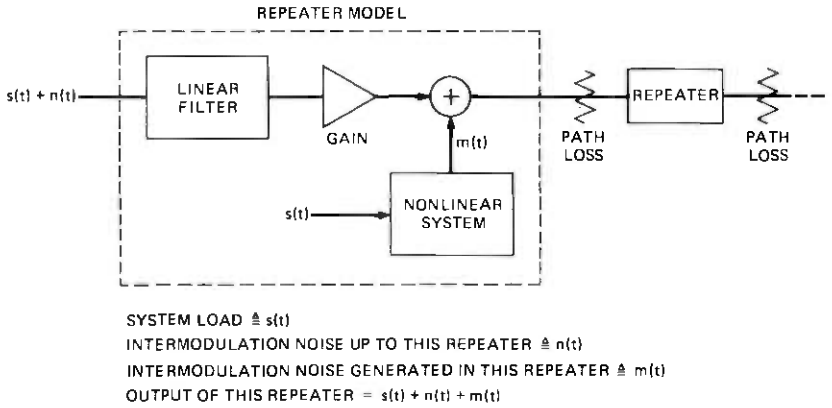


Fig. 7—System model for tandem response.

peaters, where the power stage that generates most of the intermodulation noise is broadband, and for certain cable repeaters.

The critical point of this model is that the intermodulation produced by the repeater, $m(t)$, depends only on the signal, $s(t)$, not the incoming intermodulation, $n(t)$. This is easily demonstrated. Suppose the nonlinear device is third order so that its intermodulation noise increases 3 dB for every dB increase in total power. Since the intermodulation must be small for a practical system, take it to be 40 dB down from the signal. In the next repeater, the signal will produce intermodulation 40 dB down from its level; but the incoming intermodulation will generate intermodulation that is 120 dB down from the signal-induced intermodulation. Such low-level signals will not affect system performance and may be ignored. In other words, the system is purely linear to small signals such as intermodulation. With the assumption that the noise generated at a repeater does not depend on the incoming noise, the total noise at the end of a tandem connection of repeaters is just the sum of these separately produced noise signals. The power of the total noise then depends on the average of products of the noise signals, which are correlations of these signals.

3.2 Experimental validation of model

The point we wish to test is that the total intermodulation from all repeaters in a tandem connection is the sum of the intermodulation from each repeater, where the intermodulation from each repeater is produced independently of the intermodulation from all other repeaters. To demonstrate this, the PRNL set was connected to the input of the first repeater and the output of the last repeater of four experimental single-sideband repeaters that were connected in tandem. The NPR of this arrangement was measured as a function of power, with all repeaters at

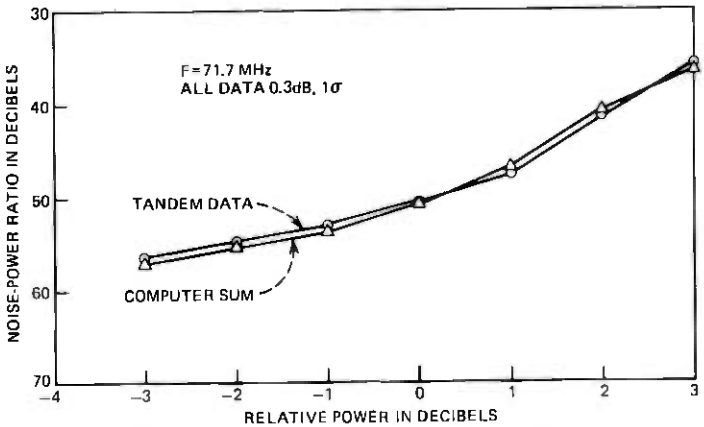


Fig. 8—Noise-loading tandem measurement.

nominal operating conditions. This result is plotted in Fig. 8 as the tandem data curve. Next, the first repeater was kept at nominal load, but the gain controls in the remaining three repeaters were adjusted to keep the transmitters, which generate most of the intermodulation, at low power. In this way, the intermodulation at the end had to be only that from the first repeater.* Notice that the intermodulation traversed all three remaining repeaters, just as when all repeaters were at nominal power. This is important when the repeaters are frequency selective, with even small amounts of delay distortion across the band, since the intermodulation from each repeater passes through different numbers of repeaters and so is distorted differently. For many repeaters in tandem, this effect can lead to considerable decorrelation of the products.

Measurement of intermodulation from each repeater proceeded in the same way with the repeater to be measured at nominal power and all others at low power. The PRNL set recorded the voltage waveform of the intermodulation from each measurement and transferred the resultant digital representation to a computer, as described in Section 2.4.2. These voltage waveforms were added together in the computer the power was computed for the composite waveform, and the results plotted in Fig. 8 labeled computer sum. The two curves are well within measurement accuracy, which substantiates the claim of addition of intermodulation.

The intermodulation is normally below the thermal noise, but in this experiment, it was at least 20 dB under the excess thermal noise created when repeaters were operated at low gain. The experiment could not have been performed without the noise-reduction capability of the PRNL set.

* This technique was proposed by H. Miedema of Bell Laboratories.

3.3 Tandem performance in terms of correlations

In Appendix A, formulas are derived for the intermodulation noise power over a string of repeaters in terms of the correlations between the noise from pairs of repeaters. The correlations available from the noise-loading set are calculated as

$$M_{ij}(0) = \int_{-T/2}^{T/2} v_i(t)v_j(t)dt, \tag{1}$$

where T is the period of the pseudorandom source, and $v_i(t)$, $v_j(t)$ are the intermodulation-noise voltages measured at the input of the PRNL set. This quantity was calculated using a computer interface with the set, as described in Section 2.4.2. As pointed out in Appendix A, eq. (10), $M_{ij}(0)$ is a sample of the broadband intermodulation-noise spectrum, $\mu_{ij}(\omega)$. The particular frequency sampled depends on $G(\omega)$, the preselect filter transfer function.

The intermodulation noise in terms of these correlations is given by eq. (7), repeated here:

$$p_n = \int_{-T/2}^{T/2} v^2(t)dt = \sum_{i=1}^{N} M_{ii}(0) + \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} M_{ij}(0), \tag{2}$$

where p_n is the intermodulation-noise power at the end of N repeaters in tandem and $v(t)$ the intermodulation-noise voltage. The correlation is defined as

$$\rho = \frac{M_{ij}(0)}{\sqrt{M_{ii}(0)M_{jj}(0)}} \tag{3}$$

and is a function of frequency, since $M_{ij}(0)$ is a function of frequency.

If ρ is a constant for a particular type of repeater, then eq. (2) may be used to extrapolate results on just a few repeaters, for which ρ has been measured, to N in tandem. Other formulae are given in Appendix A, where the effects of the linear portions of the repeaters may be studied. $M_{ij}(0)$ includes the effects of the linear filter part of the repeater, as well as the nonlinear part, as expressed in eq. (12) in Appendix A:

$$\mu_{ij}(\omega) = \psi_{ij}\hat{H}_i(\omega)\hat{H}_j^*(\omega). \tag{4}$$

$\hat{H}_i(\omega)$ is the overall linear transfer function from the output of the ith repeater to the Nth repeater, as illustrated in Fig. 9. According to our model, ψ_{ij}, the correlation between the intermodulation noise from the nonlinear element of repeater i to that of repeater j, is independent of frequency.

3.4 Experimental results

Correlations were calculated for the four experimental repeaters mentioned in Section 3.2 using the correlation technique described of

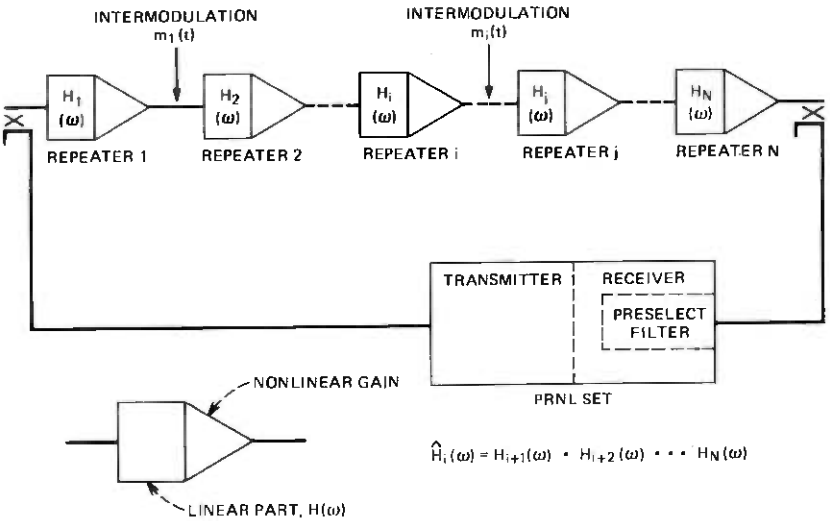


Fig. 9—Intermodulation measurement for tandem repeaters.

reducing the power on all repeaters but the one of interest in order to isolate the intermodulation noise from that repeater. The total noise of all four in tandem was also measured. Results of the correlations for various power settings at the 71.7-MHz notch are given in Fig. 10; the tandem results are shown in Fig. 11.

In Fig. 10, the circles give the mean correlation of the measurements, and the error bars indicate one standard deviation from the mean as

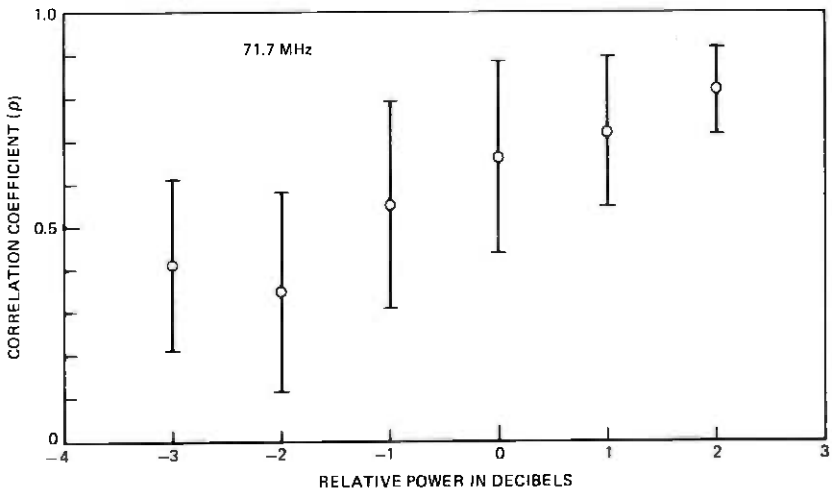


Fig. 10—Intermodulation correlation vs power at 71.7 MHz.

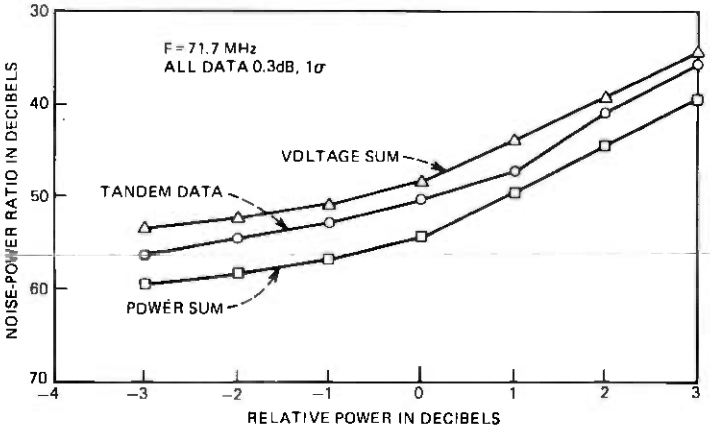


Fig. 11—Noise-loading tandem measurement.

computed from the data. The error in a correlation measurement is reviewed in Appendix B and shown to be about 10 percent. Thus, the large deviations observed must be attributed to the variability in intermodulation correlation between the repeaters. Although the data shown are all positive, negative correlations are possible and were observed.

The curve labeled "tandem data" in Fig. 11 is the actual total NPR measured for the four repeaters in tandem at power levels corresponding to those in Fig. 10. For comparison, results assuming power ($\rho = 0$) or voltage ($\rho = 1$) addition are included by appropriately summing the NPRs from each repeater at a given power level.

A rough comparison of the results of Figs. 10 and 11 can be made by referring to eq. (18) in Appendix A. This indicates, with some manipulation, that the tandem-data curve should have an NPR less than that for the power-sum curve by $10 \log[1 + (N - 1)\rho] = 10 \log(1 + 3\rho)$. For example, at a relative power of −2 dB, the spread in correlation implies an NPR worse than power addition by 1.1 to 4.5 dB; for +1 dB relative power, the spread is 4.2 to 5.7 dB. The maximum is 6 dB for $\rho = 1$, or voltage addition. The broad range in correlation of Fig. 10 therefore corresponds to a broad range in the possible tandem-data curves of Fig. 11, of which the curve shown is just one possibility. The filter characteristics of these repeaters were fairly flat at the midband of 70 MHz, so the range of values of Fig. 10 were probably due to variability in the intermodulation noise produced by the nonlinear elements.

We can see from Fig. 10 that power addition does not occur anywhere in the measured power range for this type of repeater, despite the clear dissimilarity between repeaters. This means that, for a large number of repeaters, intermodulation addition will approach a voltage law, as predicted by eq. (19) in Appendix A.

IV. CONCLUSIONS

The pseudorandom noise-loading set has made accurate, reliable, reproducible noise-loading measurements of intermodulation noise that is below thermal noise. This has permitted testing single-sideband repeaters in their nominal operating power range, tests that would not have been possible using conventional noise loading. The set has multifrequency capability that allows testing of pieces of a heterodyne system as well as the whole. Another unique capability of the set, when linked to a computer, is the computation of correlation of intermodulation noise between different repeaters. Estimates of the addition of intermodulation products on a tandem connection of repeaters can be made from these correlations.

APPENDIX A

Tandem Response Using Correlation

The tandem response of a single-sideband system can be calculated, using the repeater model shown in Fig. 9. The basic assumption of this model, as explained in the text, is that the intermodulation is produced by the signal load and not by other intermodulation signals. Three calculations are made in this appendix. First, the intermodulation power as measured by a noise loading set on a tandem connection of repeaters is shown to depend only on the crosscorrelation of intermodulation noise on pairs of repeaters. Crosscorrelation can be computed from the digital output of the PRNL set. Secondly, the crosscorrelations are broken down to show explicitly the relationship between the crosscorrelation of intermodulation noise and the linear filter part of each repeater. These two results on correlations permit prediction of the addition of intermodulation for many repeaters from measurements on a few. Lastly, the relationship between correlations and a quantity, known as the law of addition, are given. This quantity is often used as a figure of merit for a tandem connection of repeaters.

A.1 Intermodulation in terms of correlations

For this calculation, refer to Fig. 9. The PRNL set is being used to measure N repeaters in tandem, with each repeater modelled as in Fig. 7. The preselect filter on the noise-loading set has been brought out separately; its transfer function is $H(\omega)$, and its impulse response is $h(t)$. Now suppose that the intermodulation time function recorded at the PRNL set from the ith repeater is $v_i(t)$, and from the jth, $v_j(t)$. These functions might be measured using the technique explained in Section 3.2 of the text. The crosscorrelation of these two functions for zero shift can be found as[4]

$$M_{ij}(0) = \int_{-T/2}^{T/2} v_i(t)v_j(t)dt, \tag{5}$$

where T is the period of the PRNL source. This number may be calculated as a sum, using the digital representation of $v(t)$. Since the total intermodulation noise is

$$v(t) = \sum_{i=1}^{N} v_i(t), \tag{6}$$

the power is

$$p_n = \int_{-T/2}^{T/2} v^2(t)dt = \sum_{i=1}^{N} M_{ii}(0) + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} M_{ij}(0), \tag{7}$$

which is the desired result. In Sections A.1 through A.3, the units are volts, kilohms, and milliwatts. All voltages are normalized by dividing them by the square root of the effective input resistance of the PRNL set.

A.2 Effect of linear filters on the correlation

Consider the effect of the preselect filter of the PRNL set on an intermodulation measurement. Call the broadband intermodulation noise from each repeater, as seen at the final repeater, $q_i(t)$ for the ith repeater. If $g(t)$ is the impulse response of the preselect filter, $G(\omega)$ its Fourier transform, then

$$v_i(t) = q_i(t)*g(t), \tag{8}$$

where $v_i(t)$ is the intermodulation measured by the PRNL set, as given in the last section, and the asterisk is convolution. Then, by the techniques of Ref. 4,

$$M_{ij}(\tau) = \int_{-\infty}^{\infty} |G(\omega)|^2 \mu_{ij}(\omega) e^{j\omega\tau} d\omega \tag{9}$$

and

$$M_{ij}(0) = \int_{-\infty}^{\infty} |G(\omega)|^2 \mu_{ij}(\omega) d\omega, \tag{10}$$

where $M_{ij}(t)$ is the crosscorrelation of v_i with v_j, and $\mu_{ij}(\omega)$ is the spectrum of the correlation of the broadband noise, $q_i(t)$ with $q_j(t)$. This says simply that the preselect filter acts as a window on $\mu_{ij}(\omega)$. Usually only $M_{ij}(0)$ is computed, since this is what is needed for the power estimation from eq. (7). The addition of products is a function of ω, and $G(\omega)$ selects the ω of interest. For example, in Fig. 10, the ω is $2\pi \times 71 \times 10^6$ Mrad/s.

Now consider the effect of the linear filters associated with the repeaters. According to Fig. 9, $\hat{H}_i(\omega)$ is the net tandem linear response of the repeaters *after* the ith one until the end of the string. If $m_i(t)$ is the intermodulation noise generated in the ith repeater, then

$$q_i(t) = m_i(t)*\hat{h}_i(t), \tag{11}$$

where $q_i(t)$ is the broadband intermodulation voltage, as defined earlier, and $\hat{h}_i(t)$ is the impulse response, the Fourier transform of $\hat{H}_i(\omega)$. By the arguments of Ref. 4,

$$\mu_{ij}(\omega) = \psi_{ij}\hat{H}_i(\omega)\hat{H}_j^*(\omega), \tag{12}$$

where ψ_{ij} is the Fourier transform of the correlation between $m_i(t)$ and $m_j(t)$, and the asterisk indicates complex conjugate. Equation (12), in conjunction with (10) and (7), permits calculation of the noise on a string of repeaters from measurements on individual repeaters. Telephone repeaters are built the same for ease of manufacture. In the ideal case, $H_i(\omega) = H(\omega)$ for all values of i; furthermore, $\psi_{ij} = \psi_c$ for all $i \neq j$, and $\psi_{ii} = \psi_s$, for all i values of j. Equation (12) reduces to

$$\begin{aligned} \mu_{ij}(\omega) &= \psi_c(\omega)H^{N-i}(\omega)[H^{N-1}(\omega)]^* \qquad i \neq j \\ &= \psi_s(\omega)|H|^{2(N-i)} \qquad i = j. \end{aligned} \tag{13}$$

Equation (7) becomes

$$p_n = \int_{-\infty}^{\infty} |G(\omega)|^2 d\omega \left[\sum_{i=1}^{N} \psi_s |H|^{2(N-i)} + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \psi_c H^{N-i}(\omega)H^{N-j}(\omega)^* \right]. \tag{14}$$

A.3 Law of addition

In systems calculations, it is common to make the simplifying assumptions that the intermodulation power from all repeaters is the same and that the correlation of the intermodulation is the same between any two repeaters. Neither assumption is true, but since a large number of repeaters is involved, average quantities may be used if the distributions of noise and correlation are narrow. If, in this case, the total intermodulation power in a narrow band over a tandem connection of repeaters N is P_N in dBm, and the intermodulation power of any one repeater in the same narrow band is P in dBm, then the law of addition, Δ, is defined as

$$P_N = P + \Delta \log_{10} N. \tag{15}$$

The value of Δ is 10 for incoherent (power) addition and 20 for coherent (voltage) addition. The advantage of using the law of addition is that $\Delta \log_{10} N$ and the required NPR for a third-order dominated system both enter the system noise equations as the sum, $\Delta \log_{10} N$ + NPR. Therefore, for the same noise performance, the NPR must improve by the same amount that $\Delta \log_{10} N$ grows larger.

To relate Δ to the correlation of the intermodulation noise between any two repeaters, assume that $H(\omega) = 1$, so that

$$\rho = M_{ij}(0)/M_{ii}(0) \tag{16}$$

and

$$p = M_{ii}(0), \tag{17}$$

where $10 \log p = P$. Then eq. (7) becomes

$$p_N = Np[1 + (N - 1)\rho]. \tag{18}$$

The definition of eq. (15) gives

$$\Delta = 10 \left\{1 + \frac{\log_{10}[1 + (N - 1)\rho]}{\log_{10} N}\right\}. \tag{19}$$

Whereas ρ is a constant for any length system, Δ depends on N. Notice that, for $\rho \neq 0$, $\Delta \to 20$ for $N \to \infty$. In any case, $\rho = 0$ corresponds to $\Delta = 10$, and $\rho = 1$ to $\Delta = 20$, as would be expected.

APPENDIX B

Errors in Noise Loading Measurements

Errors in a measurement of intermodulation power were given in Section II and are reviewed here. In addition, errors in correlation measurements are given.

B.1 Intermodulation power and NPR

In Section II, the basic error in an intermodulation power measurement with this set was given as ± 0.25 dB, or 0.14 dB 1 σ. All errors are independent, so that error bars on Fig. 10 for assumed power or voltage addition are derived hy summing the variances of the errors for each repeater measurement.

B.2 Errors in estimates of correlation coefficient

The correlation coefficient is defined as

$$\rho_{ij} = \frac{M_{ij}(0)}{\sqrt{M_{ii}(0)M_{jj}(0)}}. \tag{20}$$

Amplitude errors occur in the measurement of intermodulation, but these do not affect the correlation. The significant sources of error are extraneous delays between the two signals being measured and sample size. Extraneous delays are delays introduced by the measurement equipment, rather than true signal delays in the repeatered string under test. In heterodyne single-sideband repeaters, such extraneous delay may

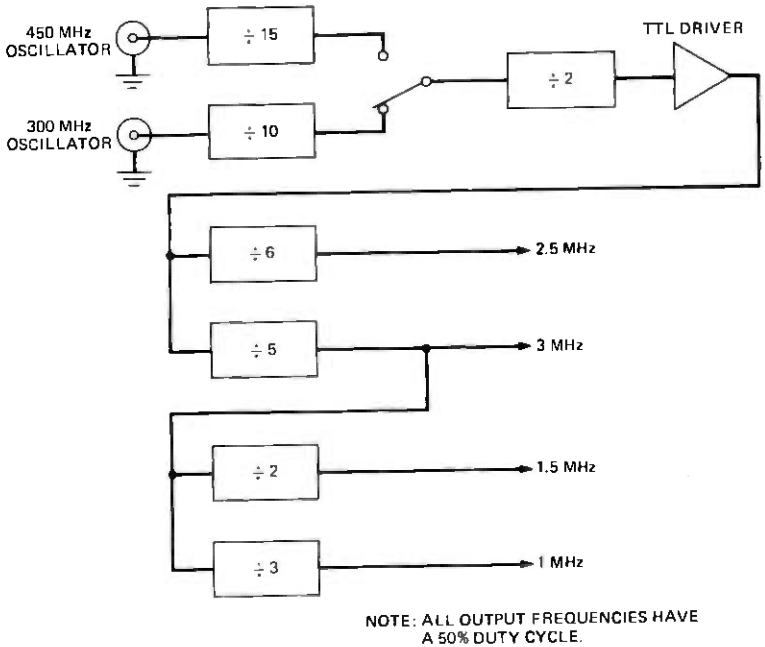


Fig. 12—Digital oscillator divider chain.

come from improperly functioning phase-locked loops. The error due to these delays in our data was estimated to have a standard deviation of 3 percent of the computed correlation.

The error in the correlation coefficient due to sample size may be estimated by assuming that the intermodulation noise in a notch is gaussian. For the noise-loading set used, the line spacing was about 200 Hz with a notch width of nearly 3 kHz, giving 15 lines for an intermodulation signal. This number of sine waves is well approximated by gaussian statistics.[5] The approximate mean and variance of an estimate of the correlation coefficient is $(\frac{1}{2})\log(1 + \rho/1 - \rho)$ and $1/(n - 3)$, respectively, where n is the number of points, and ρ is the estimate.[6] For the 255 data points available from the averager, this comes to a standard deviation of the error of about 6 percent. The standard deviation of the total error in an estimate of correlation is about 7 percent.

APPENDIX C

6-GHz Noise Loading Set

In the latter part of 1974 and early 1975, another PRNL set was designed for the 6-GHz band. Besides a much improved physical design, this set was designed with all LOs derived from a single source. This change reduced the cost of LO generation and also simplified the master

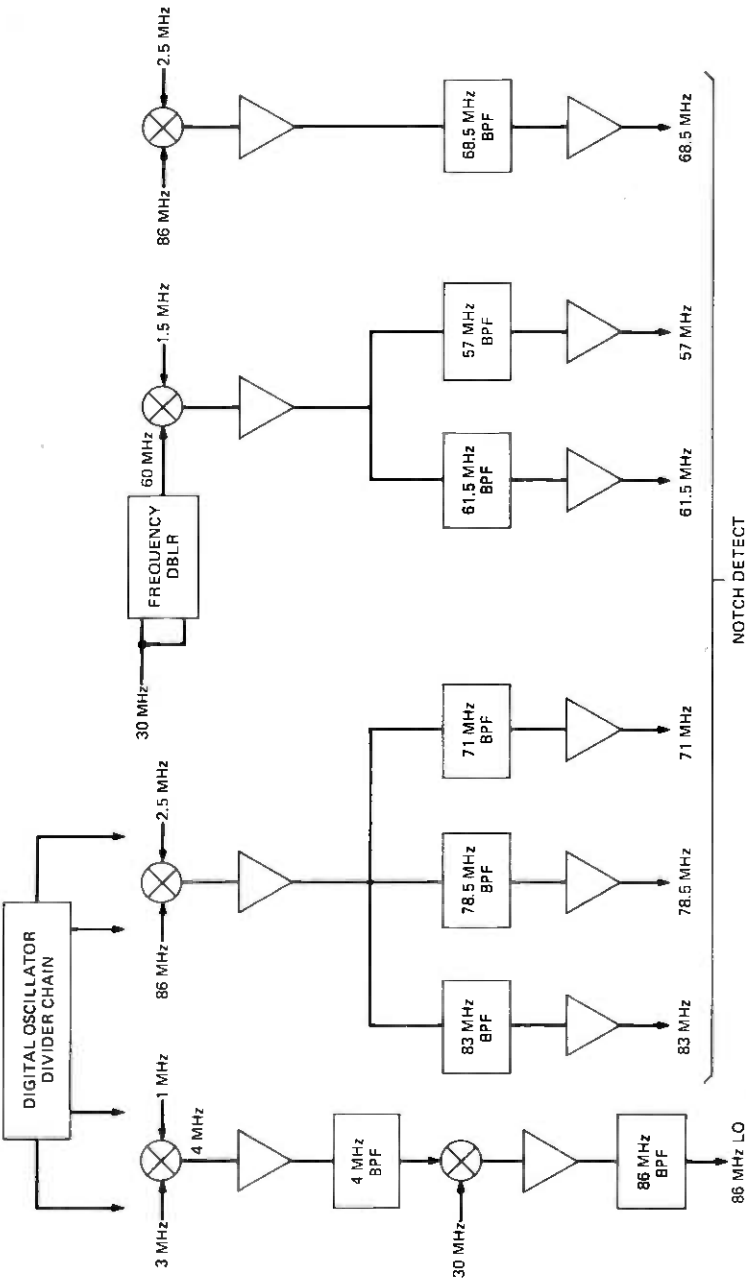


Fig. 13—Local oscillator frequency generation.

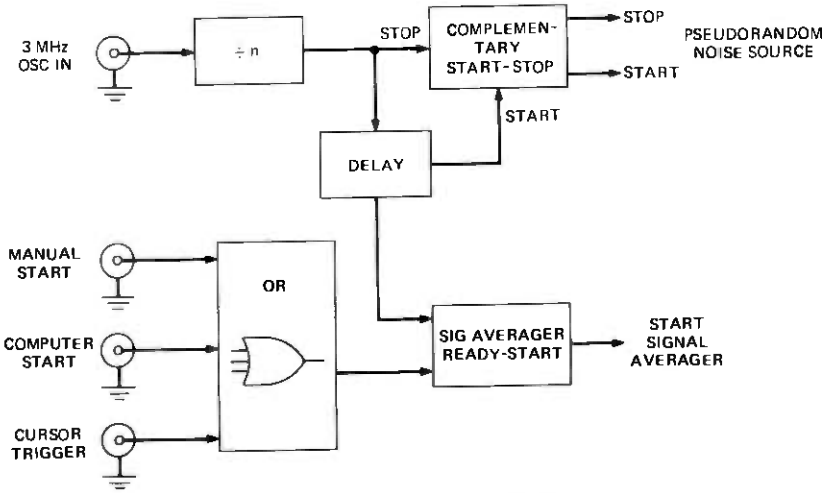


Fig. 14—Master timing circuit.

timing circuit. In addition, the set was designed to operate in the 4-GHz band with minimal changes. These design changes are described below.

In the 6-GHz set, only one master oscillator is used, either 300 MHz or 450 MHz (see Fig. 12), depending on whether a 20-MHz or 30-MHz IF is desired.

Using digital dividers, a number of basic frequencies are derived from this single oscillator. Double balanced mixers (see Fig. 13) are used to combine the appropriate basic frequencies as well as notch-detector frequencies. Crystal filters select the local oscillators so that spurs are more than 70 dB down. This keeps signal leakage into the notches at tolerable levels.

The block diagram for the new master timing circuit is shown in Fig. 14. The major difference between this and the previous circuit is the absence of the coincidence circuit. This circuit could be eliminated since all local oscillator frequencies, as well as all notch-select frequencies, are derived directly from one master oscillator. If the pseudorandom source repetition frequency is made a multiple of the lowest common denominator for all these LOs, coherence is assured. This lowest common denominator is 0.5 MHz, and the source frequency is usually chosen as 192.3 Hz, i.e., 2600 periods of the 0.5 MHz frequency.

Each single cycle of the test signal has to be slightly larger than the 5.12-ms single-sweep average period. This is achieved by setting the divide-by-n counter to the desired period. The complementary start-stop circuit, the delay circuit to assure a minimum reset period, the combining of Start commands (manual start, computer start, cursor trigger),

and the ready circuit are all similar to those shown in Fig. 3. The principal difference is simply the absence of the coincidence circuit. The strobe circuit is not needed since all notch-detect frequencies are coherently related to the 3-MHz oscillator frequency used to set the source period.

Output pulse jitter for this circuit is less than 0.25 ns, which makes it possible to achieve 42-dB signal-to-noise improvement routinely.

REFERENCES

1. *Transmission Systems for Communication*, Fourth Edition, Bell Laboratories, Inc., 1970.
2. P. O. Roberts, and R. H. Davis, "Statistical Properties of Smoothed Maximal-length Linear Binary Sequences," Proc. IEE, *113*, No. 1 (January 1966), pp. 190–196. See also additions by I. G. Cumming, "Autocorrelation Function and Spectrum of a Filtered, Pseudorandom Binary Sequence," Proc. IEE, *114*, No. 9 (September 1967), pp 1360–1364.
3. K. Y. Chang, "Intermodulation Noise and Products Due to Frequency-Dependent Nonlinearities in CATV Systems," IEEE Trans. Commun., *COM-23*, No. 1 (January 1975), pp. 142–155. Unpublished studies by C. D. Anderson, J. W. Smith, and L. L. Sheets of Bell Laboratories.
4. Y. W. Lee, *Statistical Theory of Communication*, New York: John Wiley, 1960, Chapter 13.
5. M. Slack, "The Probability Distributions of Sinusoidal Oscillations Combined in Random Phase," IEE Journal, *93*, Part III (March 1946), pp. 76–86.
6. K. A. Brownlee, *Statistical Theory and Methodology in Science and Engineering*, New York: John Wiley, 1961, p. 365.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 56, No. 4, April 1977
Printed in U.S.A.

On the Occurrence of Caustics in the Drawdown Zone of Silica Fibers

By P. G. SIMPKINS, T. D. DUDDERAR, J. McKENNA and J. B. SEERY

(Manuscript received September 16, 1976)

Caustics created by internal reflections and interface refraction have been observed in the molten drawdown zones of fibers during manufacture in a laser furnace. These regions of high luminescence are associated with the rapid change in cross-sectional area that occurs in the drawdown zone. Observations made with solidified samples of the molten drawdown zone reveal the nature of the caustic envelopes. Algorithms which incorporate ray tracing techniques have been developed and used with surface profile measurements to generate caustic loci. The results from these numerical experiments are in good agreement with observations. The caustics are found to be particularly sensitive to asymmetries in the drawdown profile. A simple technique for monitoring the maximum gradient of the drawdown zone is also described.

I. INTRODUCTION

The drawdown zone of a silica fiber manufactured in a laser furnace has a cross section that varies rapidly along its axis. When this zone is internally illuminated with a collimated beam parallel to the axis, a region of high-intensity radiation is observed (see Fig. 1). The radiation is the result of multiple internal reflections and a refraction of the beam caused by the rapidly varying boundary of the glass. The locus where the rays of geometrical optics form an envelope is known as a caustic. Along such a surface the luminous intensity is a maximum.[1] Normal to the caustic, the intensity is characterized mathematically by an Airy function, decaying exponentially on one side, the shadow region, and varying harmonically in the opposite direction. For simple geometries, the techniques used to calculate the caustics are straightforward; e.g., a collimated beam reflected from a semicircular concave mirror forms a caustic which is part of an epicycloid. When the reflecting boundaries

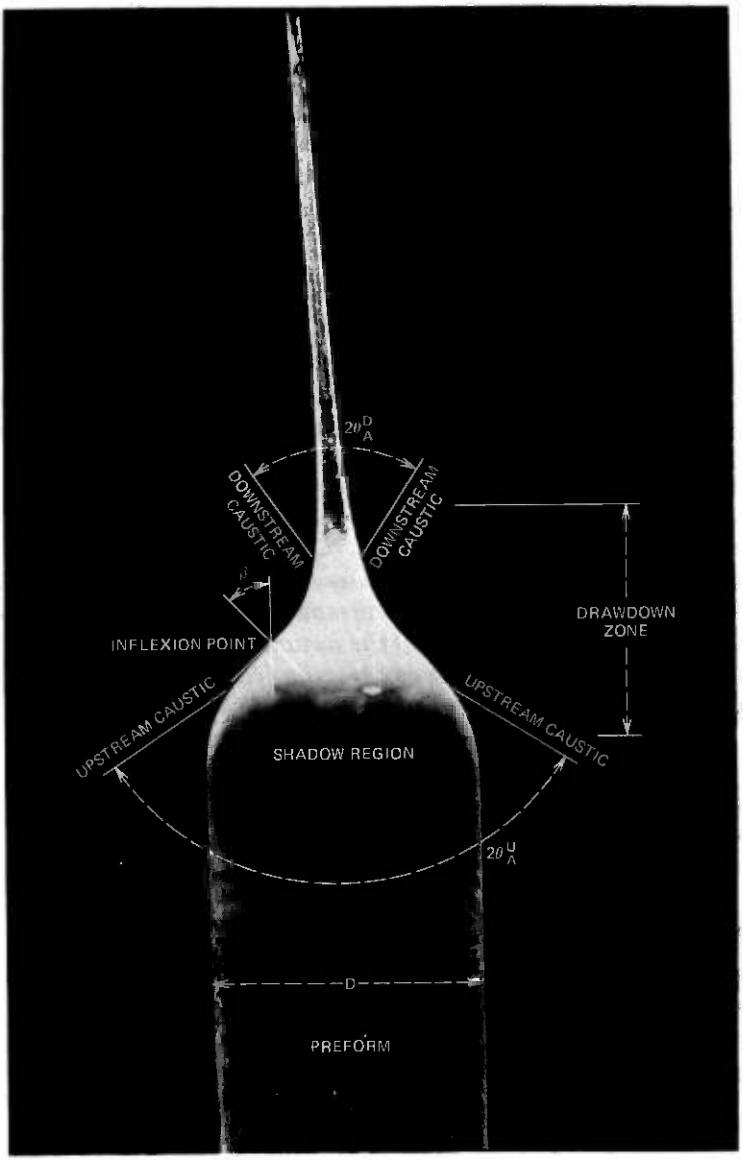


Fig. 1—High-intensity drawdown zone labeled with the relevant geometrical parameters.

are not given analytically, however, the calculation of the caustic geometry is more laborious.

It is our hypothesis that light reflected and refracted from the drawdown zone in the form of caustics can provide an extremely sensitive measure of the surface geometry and, possibly, such physical quantities

as the index of refraction and temperature. Observations made during an actual silica-fiber drawing operation confirm the existence of caustics in the molten zone. In that case, the emission is the result of the high-intensity radiation along numerous ray paths, many of which are skewed.

In this paper, we restrict our attention to an examination of the solidified drawdown zones of a number of fused-silica samples. For the most part, we concentrate on the radiation produced by an internal plane wave traveling along a sample; this is not, however, a limitation as experiments with a diffused source will show. Our observations are used to illustrate that the caustics originate from light rays traveling through different regions within the cross section. Furthermore, we show that the drawdown contour plays an important role in defining the caustic geometry and that small changes in the profile gradient can produce large variations in the caustic patterns. For a given index of refraction, ray theoretic analysis and geometrical optics confirm that the observed far-field caustic phenomenon can be predicted with good accuracy provided the boundary of the drawdown zone is well defined. Conversely, careful observation of the caustic patterns might be used to determine important aspects of the drawdown zone's geometry. We also show that a reflected caustic generated by external illumination of the drawdown zone may be used to provide a practical and accurate measure of the maximum profile gradient.

We consider two limiting cases and describe these as an "upstream" and a "downstream" caustic, in order to designate their nominal location relative to the silica flow during the drawing operation. These caustics do not appear in the drawdown zone of preforms being heated in an induction furnace since the profile gradient is too small. However, a geometrically relevant caustic can always be generated by external illumination of the drawdown zone regardless of how small the profile gradient becomes. This point will be discussed further below.

II. EXPERIMENTAL METHODS

Solidified samples of the transition region between the preform and the fiber were cut and polished square to the axis on the preform end some 50 mm ahead of the drawdown zone. Each sample began at a preform of a uniform diameter between 6.5 mm and 8 mm, and terminated at a fiber diameter of typically 200 μm.* All the samples were of fused silica with an index of refraction, n, of 1.457 at a wavelength of 643.8 nm.[2] Individual samples were mounted in a gimballed mirror mount with micrometer adjustment. The sample was usually illuminated internally

* Although the drawdown region continues much further—to a final fiber diameter of about 100 μm—this region is of no interest to the current study.

by a light beam entering the preform parallel to the fiber axis. This technique generated the upstream and downstream caustics. The sample could also be illuminated externally by a beam normal to the fiber axis and incident to the drawdown zone. This alternative generated a reflected caustic from which the maximum gradient of the drawdown profile could be determined. An Argon Ion or He-Ne laser, $\lambda = 514.5$ and 632.8 nm, respectively, provided the illumination. In either case, the beam was expanded through a spatial filter and collimated to a diameter greater than the preform. Particular regions of the preform cross section were selectively illuminated via adjustable slits and diaphragms. When required, the collimated beam was altered by introducing a diffuser at the polished end of the sample. The experimental arrangement is illustrated in Fig. 2.

Images of the caustic patterns emitted from a particular sample were recorded in planes parallel and perpendicular to its axis. These far-field caustic images (see Figs. 3 and 4 for typical examples) were obtained by mounting the film in a Polaroid* film holder and exposing it in a darkened room, or, by photographing the image formed on a suitable screen in the same plane. By translating the sample relative to the reference plane, the locus and orientation of the caustic was determined. The precise location of the caustic at the surface of the sample was obtained by adjusting a toolmaker's height gauge, in contact with the surface, until it interrupted the display of the far-field caustic image.

III. EXPERIMENTAL RESULTS

3.1 Origins of the caustics

A diaphragm aperture was used to vary the incident-beam diameter to determine which annular regions within a transverse cross section of the sample transmitted light that contributed to each caustic. With a ground glass viewing screen in position 1 or 3, Fig. 2, the incident beam diameter was reduced until the caustics disappeared. Figure 3 shows typical far-field photographs taken during these tests. When the beam and preform diameters were equal, both upstream and downstream caustics were visible. When the beam diameter was reduced approximately five percent, the downstream caustic vanished, while the upstream caustic remained unaffected. The upstream caustic remained intense until the beam diameter was reduced to between 60 and 55 percent of the sample diameter, when it began to lose intensity rapidly. These observations show that the downstream caustic originates from a region very near the shoulder of the drawdown. The upstream caustic has its source close to the inflexion point of the drawdown zone.

When the image plane was oriented parallel to the preform axis, po-

* Registered trademark of Polaroid Corporation.

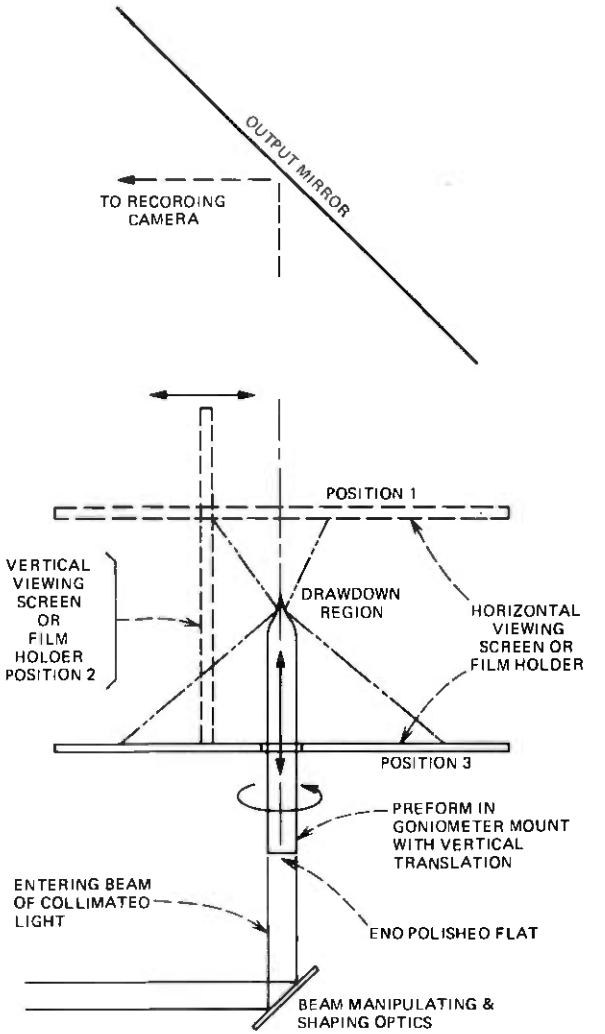


Fig. 2—Experimental arrangements for measuring various caustic loci.

sition 2 in Fig. 2, the caustic pattern for a symmetric sample became hyperbolas. The photographs in Fig. 4 show a comparison of the caustic images recorded using beams of different diameters with and without a diffuser. The principal observation is that both caustics can be seen in the full beam photographs. Although the detailed structure that occurs with the collimated beam has been eliminated by the diffuser, the skew rays introduced thereby do not eliminate the caustics. A reduction in the collimated beam diameter below that of the preform quickly eliminates the downstream caustics. However, they can still be seen whenever

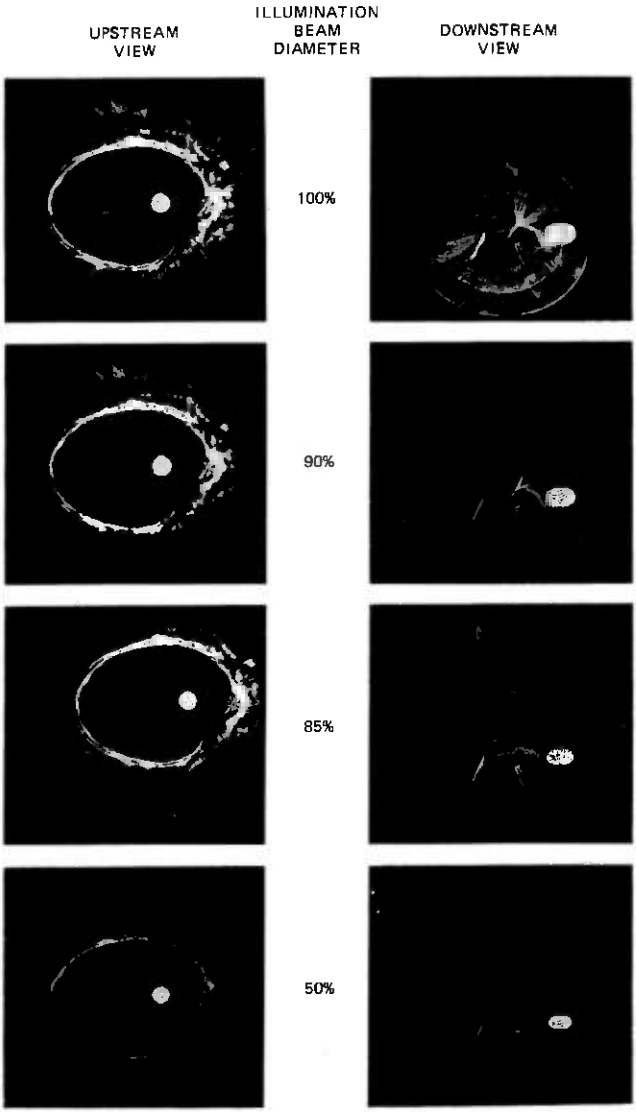


Fig. 3—Variations of the far-field caustic images (positions 1 and 3) with changes in illuminating-beam diameter; sample 1.

the diffuser is placed in the beam because the light scattered towards the surface can be reflected to form the downstream caustic. Significantly, as the beam is stopped down further, so that the upstream caustic also starts to disappear, the resolution of both caustics becomes relatively more apparent with the diffused beam. Considerable additional caustic structure, which is completely obscured in the patterns recorded with

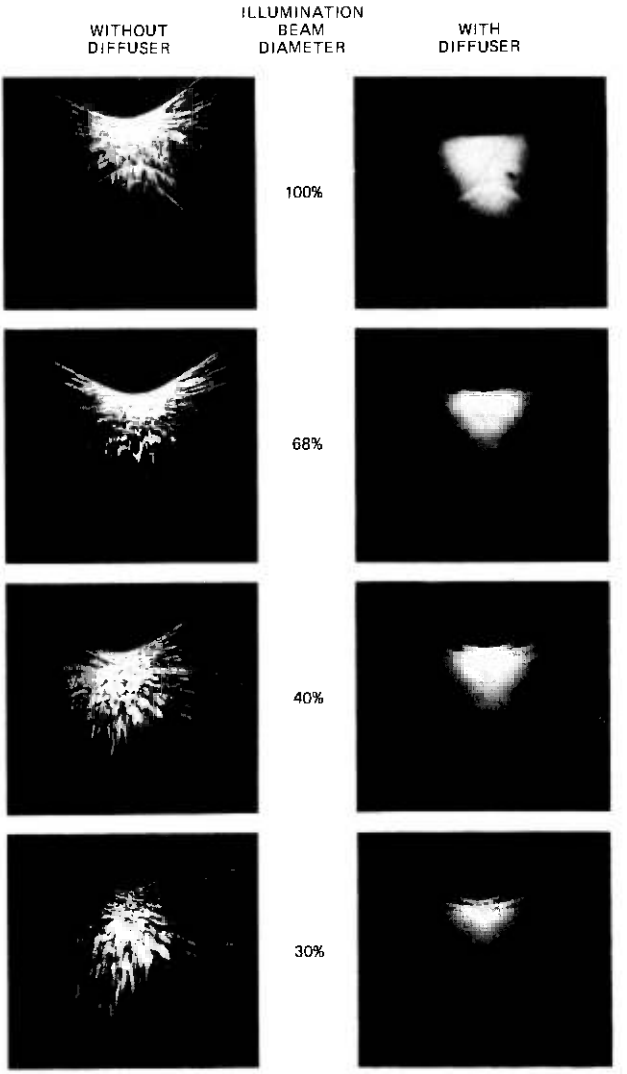


Fig. 4—Variations of the far-field caustic images (position 2) with collimated and diffused illuminating beams of changing diameter side views.

the fully collimated beam, becomes visible in patterns recorded over much longer times with diffused beams of substantially reduced diameter.

3.2 Propagation paths of the emergent caustics

Accurate* determinations of the propagation directions of the caustics

* See Appendix for estimates of accuracy.

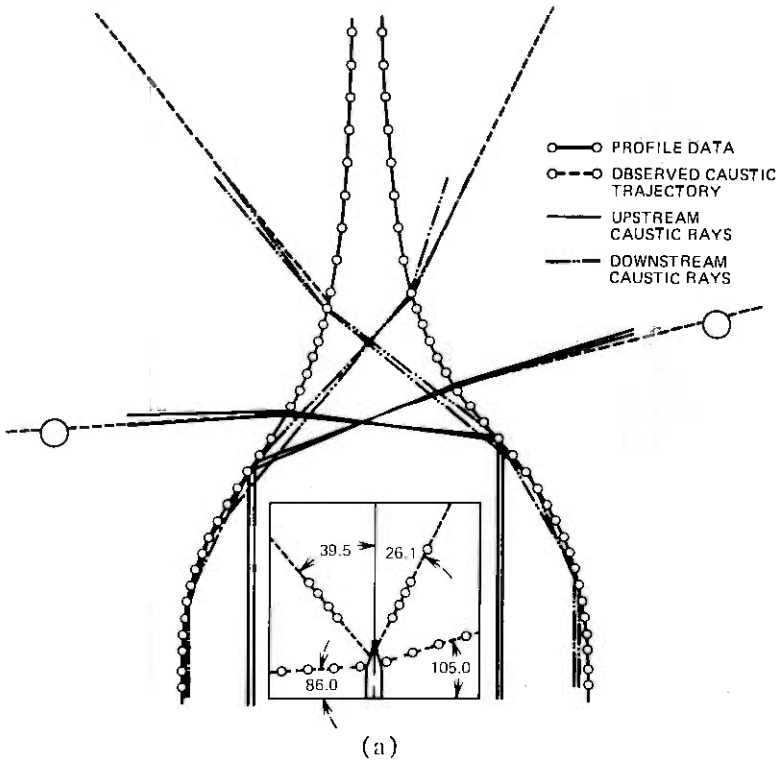


Fig. 5—(a) Comparison of the measured far-field caustic trajectories and the graphical ray traces for adjacent rays bounding the caustic ray.

and the precise loci of their emergence from the preform were made as described in Section II. This information was subsequently compared with emerging caustics identified by geometrical ray tracing through a known preform geometry. The ray tracing was performed graphically on a fiftyfold enlargement. Figure 5a shows, on a necessarily reduced scale, the excellent agreement between the experimental observation and the graphical ray tracing in a typical sample, such as sample 4.

From the above comparisons, we conclude that the observed caustics do indeed result from light following the paths determined by the graphical ray tracing. In Fig. 5a we see that the upstream caustic is generated by rays traveling down the sample at about the half radius which reflect internally once. These rays then cross the sample and strike the diametrically opposite side at such an angle that they emerge upstream of the downstream caustics. A caustic represents a limiting phenomenon at which the field of emerging refracted rays folds back upon itself. Thus, the upstream caustic represents the furthermost up-

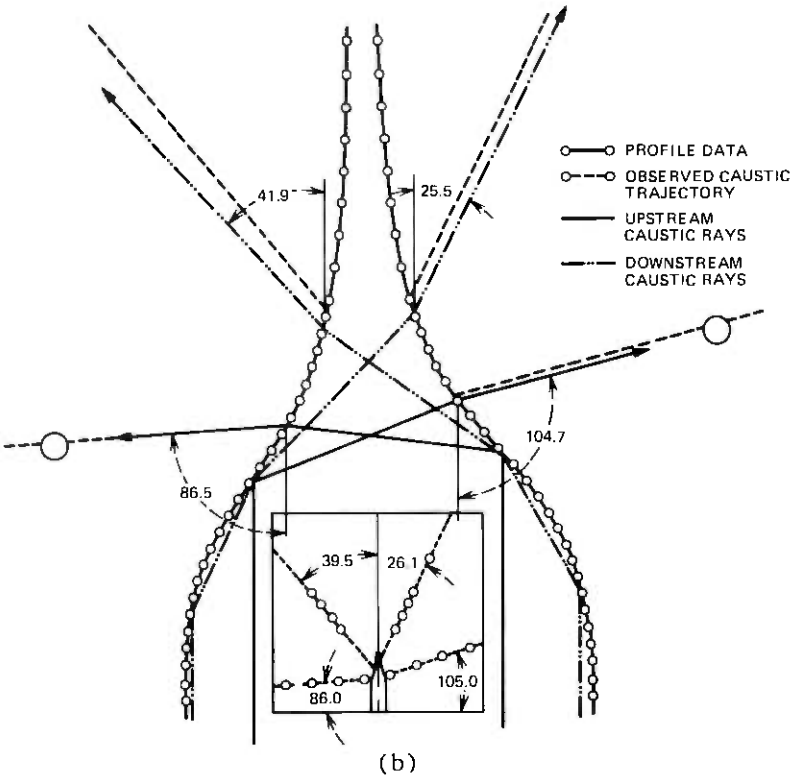


Fig. 5—(b) Comparison of the same observations with the numerical calculations.

stream limit that rays can emerge from the preform. It is possible, however, that the upstream caustic will not appear because the field of emerging rays is terminated at the critical angle. Then, before the field can fold back upon itself, it is internally reflected. The upstream image for such a condition is shown in Fig. 6b where no caustic is seen at the boundary between the light and dark regions. A similar internal reflection occurs locally with the original sample shown in Fig. 3. The asymmetry of the caustic observed in Fig. 3 and its local extinction are due to asymmetries in the drawdown profile, the effects of which will be discussed in Section 3.3.

Referring now to the downstream caustics, we have already established in Section 3.1 that these arise from light traveling down the sample very close to its surface. Their behavior is similar to the upstream caustic; however, the graphical ray tracing reveals an important distinction between the two families. The downstream caustic is formed by rays that have reflected twice before crossing the sample and emerging (see Fig. 5a). As noted earlier, there is excellent agreement between the constructed and observed caustic locations. This caustic is unique in that

Table I — Synopsis of the drawdown sample geometries

Sample	Station	Taper L/D*	$\gamma = \left(\frac{dy}{dx}\right)_{max}$	β = arc cot γ
1	3	0.73	0.87	48.9°
	6	0.73	0.92	47.3°
	9		1.01	44.6°
	12		0.98	45.6°
2	3	0.62	1.02	44.5°
	6	0.63	1.12	41.7°
	9		1.05	43.5°
	12		0.99	45.2°
3	3	0.44	1.95	27.1°
	6	0.46	1.89	27.9°
	9		1.89	27.9°
	12		2.14	25.0°
4	3	0.94	0.82	50.6°
	6	0.97	0.67	56.0°
	9		0.72	54.2°
	12		0.88	48.8°

* The taper information relates to pairs of profiles, so the information in row 3 refers to profiles 3 and 9 while information in row 6 refers to profiles 6 and 12.

no other ray paths give rise to a downstream caustic with a greater cone angle.

3.3 Studies of geometrical effects

3.3.1 Upstream caustics

Four samples with different geometries were studied to develop an understanding of the relationship between the shape of the drawdown region and the caustic loci. These samples, described in Table I, were selected to span the range of geometries to be expected from the laser drawing process. Every sample was examined in two orthogonal planes, parallel to the axis. These planes intersect the drawdown zone in four curves labelled 3, 6, 9 and 12 when viewed from the fiber end. The drawdown taper (L/D) ranges from $0.4 \leq L/D \leq 1.0$. Here D is the preform diameter, and L is the distance along the axis from a diameter of 0.98D to one of 0.15D. The maximum gradient, $\gamma = (dy/dx)_{max}$, ranged from 0.67 to 2.14. Since the caustic formation depends upon interactions between the ray trajectories and opposite boundaries of the drawdown zone, asymmetries on any particular cross section are of considerable significance.

The most extensive studies were made on sample 1 and have been discussed earlier. The other three samples were examined only with a full collimated beam (see Fig. 2) to establish the caustic geometries. Typical far-field caustic patterns from each sample are given in Figs. 6 and 7. The upstream images (see Fig. 6) illustrate the diversity of patterns that arise because of variations in the profile geometry. Figure 6a

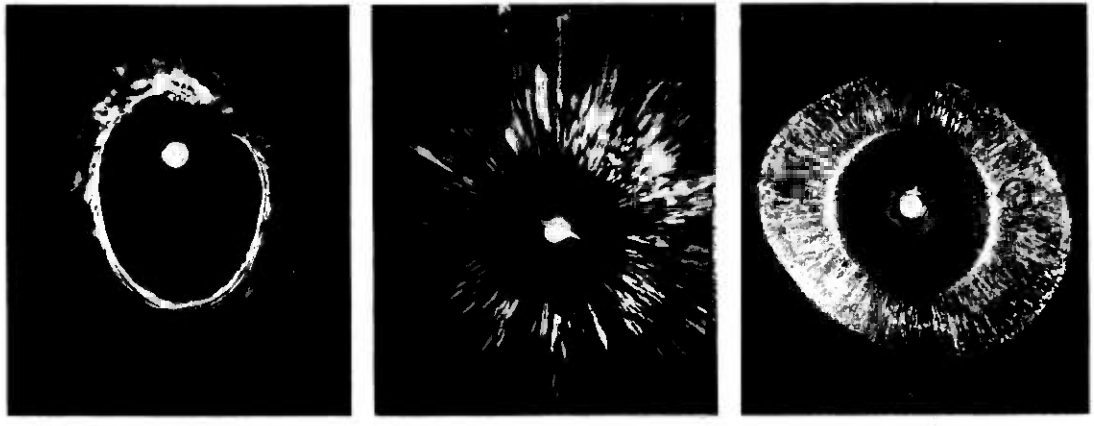


Fig. 6—Far-field upstream (position 3) caustic images observed from three different samples, illustrating the effects of geometrical variations.

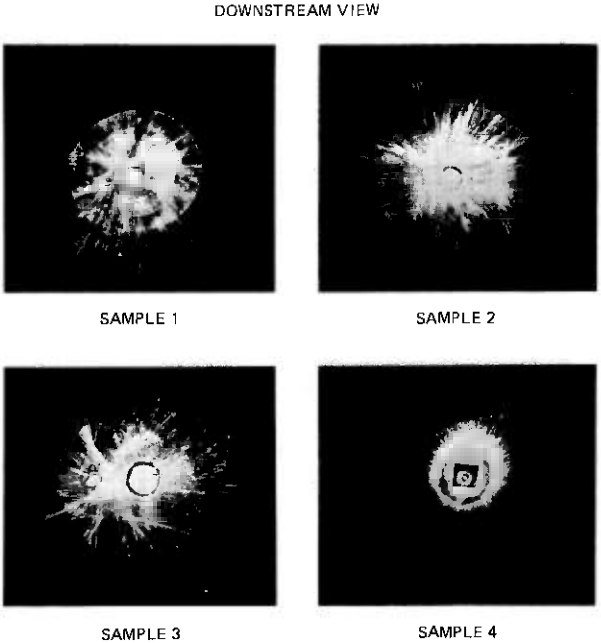


Fig. 7—Far-field downstream (position 1) caustic images observed from four different samples.

shows a single caustic with considerable asymmetry, one segment of which vanishes where internal reflection occurs. Figure 6b shows an occurrence in which there are no upstream caustics, which is again the result of internal reflection. Referring to Table I, we note that the angles of the normals to the profile, β, for samples 1 and 2 are in the neighborhood of 45°. Here $\beta = \text{arc cot } \gamma$. Since the upstream caustic ray paths intercept each boundary once, this assures that the significant incident rays approach the second surface at angles close to the critical value of 43.2° ($n = 1.46$). In these circumstances, the ray trajectories for sample 1 are such that the incident angle, ϕ, is usually less than, but close to, the critical angle. On the other hand, for sample 2, ϕ is everywhere greater than the critical angle. Snell's law predicts an increased sensitivity of the refracted ray to changes in the incident ray path in the vicinity of the critical angle; this is illustrated in Fig. 8. Clearly in regions close to the critical angle small changes in incidence are magnified considerably. Thus, sample 1 shows large caustic eccentricity associated with modest profile asymmetry, while sample 2 exhibits no upstream caustic at all. It does, of course, exhibit a transition between dark and light regions in the upstream field. However, the locus of this transition is associated with the limiting rays of an internal reflection.

Consider now Fig. 6c, in which one observes the appearance of *two*

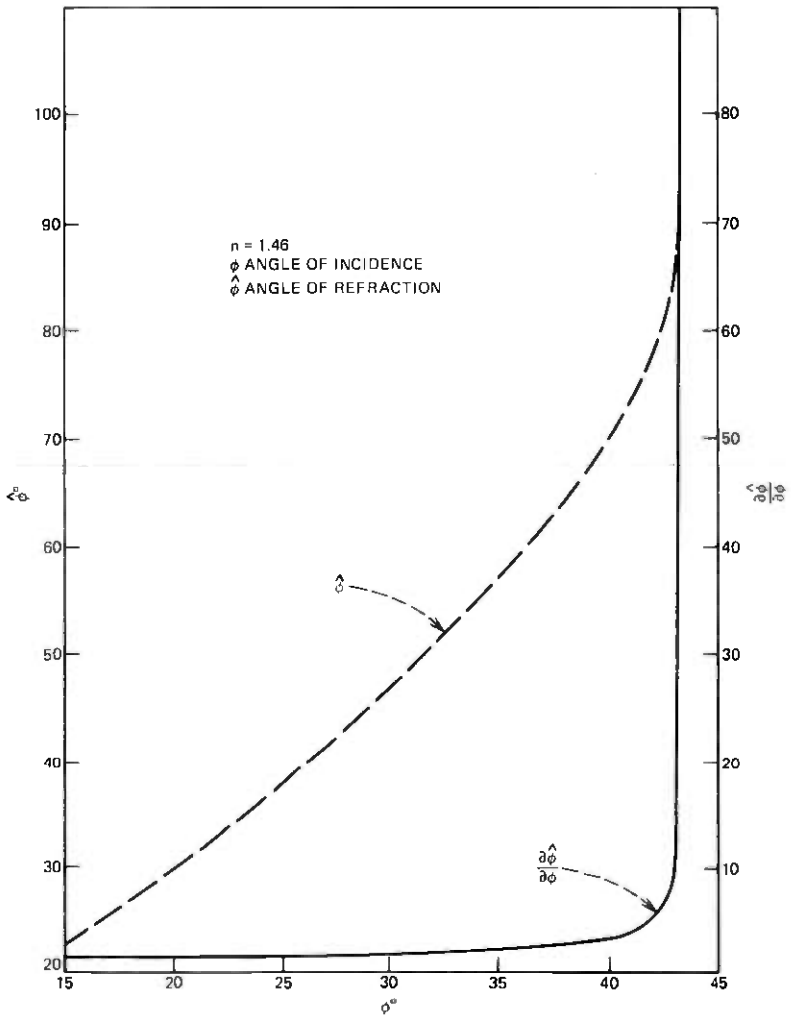


Fig. 8—Sensitivity of the angle of refraction, $\hat{\phi}$, to changes in the angle of incidence, ϕ, as ϕ approaches the critical angle.

upstream caustic boundaries. The condition arises because of the complicated reflections and refractions that occur in this sharply tapered drawdown region. The severity of the profile gradient allows some of the incident beam to refract from drawdown zone on the first interception with the boundary.* The small percentage of light that is internally reflected due to the low surface reflectance traverses the drawdown zone

* This emitted light forms a caustic at the inflexion point which is refracted downstream at a negative angle. This phenomenon is not classified as a 'downstream' caustic because it intercepts the drawdown surface only once before it is emitted.

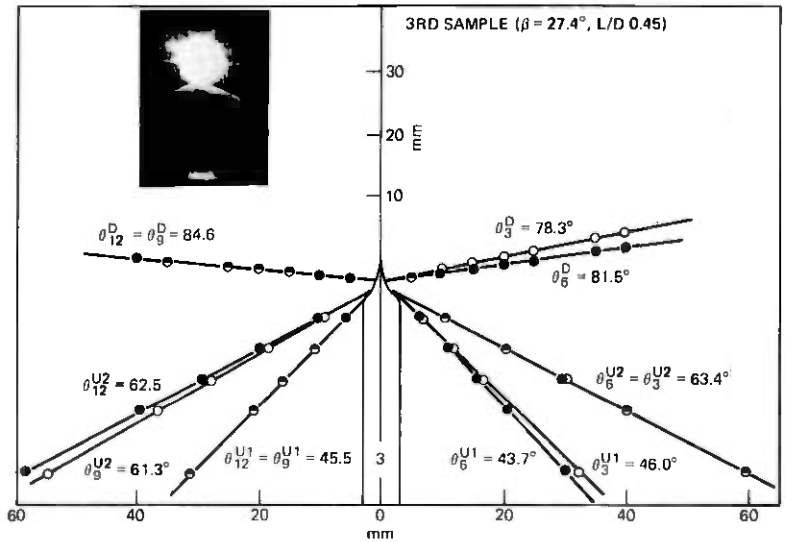


Fig. 9—Far-field caustic trajectories taken from sample 3, showing two upstream and a single downstream caustic. Each individual caustic can be identified on the accompanying photograph.

and emerges in the upstream direction bounded between two caustics (Fig. 6c).

Sample 4 is the shallowest of all those examined. Consequently, the ray trajectories are straightforward although the upstream caustic propagates downstream over an appreciable region of the polar angle. For that reason, no upstream photographs were made, and all data were recorded from position 2 (Fig. 2). Those measurements are given in the inserts of Fig. 5. The formation of the upstream caustic is the same as in sample 1, but since the local slope is much shallower, the incident ray angles are much less than the critical angle. Hence, the upstream caustic pattern for sample 4 is not as sensitive to deviations in the drawdown asymmetry.

3.3.2 Downstream caustics

Far-field images of the downstream caustics are shown in Fig. 7. The image for sample 3 is not apparent because it emerged almost normal to the axis. Data for this sample was therefore recorded from the side view, position 2, and is given in Fig. 9. Compared with the upstream caustics, the downstream ones are relatively well behaved and vary continuously with the surface geometry. The local caustic extinction found in the upstream cases does not now occur because the incident ray trajectories relative to the emergent surface are far removed from the critical angle. Careful analysis of these data shows that the downstream

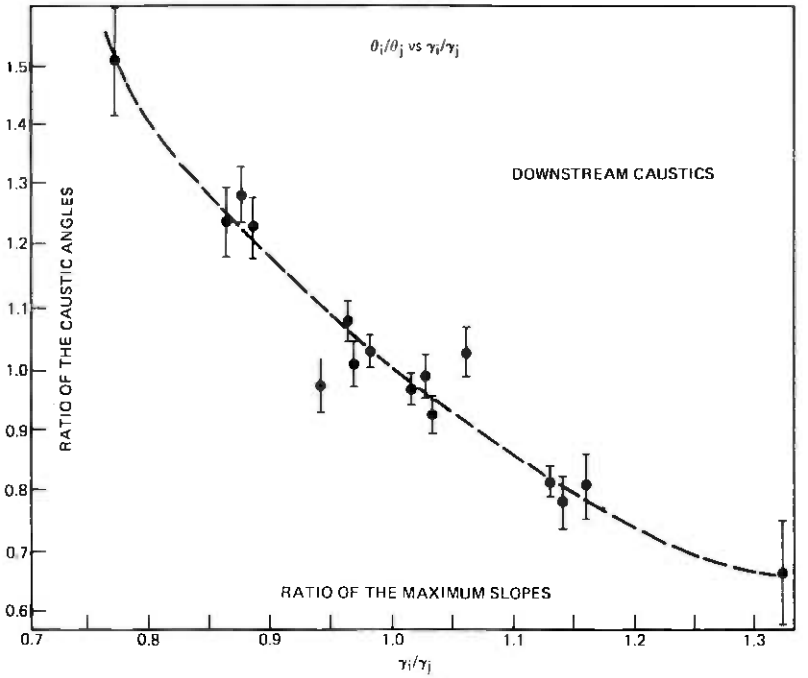


Fig. 10—Correlation of the downstream caustic angles with the maximum drawdown slope for all samples.

caustic asymmetry can be related to the asymmetry of the drawdown geometry. Figure 10 presents a correlation of the ratios of the caustic angles θ_i/θ_j, with the ratio of the maximum slopes γ_i/γ_j for all samples. The error bars represent the uncertainty in the measurement of the caustic angles, typically in absolute terms $\pm 1°$. With the exception of sample 3, the maximum gradients of the samples were obtained by spline fitting the original profile data measured with the Nikon comparator. The exterior reflected caustic technique was used to measure the gradient of sample 3 because its profile was too steep and irregular to be measured with the comparator and analyzed accurately. The significant aspect of the curve shown in Fig. 10 is the trend toward decreasing caustic angle ratio with increasing slope ratio. The scatter observed arises from the uncertainty in determining the slopes, and also from the probable dependence of the caustic asymmetry on factors other than the maximum gradient.

IV. MATHEMATICAL AND NUMERICAL ANALYSIS OF THE INTERNALLY ILLUMINATED CAUSTICS

In the preceding sections, we discussed the optical phenomena seen when the drawdown zone of a silica fiber is appropriately illuminated.

The experimental evidence presented shows that many important features of these phenomena could be explained by geometrical optics in terms of families of rays and their envelope curves or caustics. In this section, we derive equations for the rays in these families as well as equations for their caustic curves. These equations have been evaluated numerically, and we compare these numerical results with experimental results. Since the index of refraction of the glass is uniform, all the rays are straight lines between intercepts with the glass-air boundary. The continuation of a ray after its intercept with the boundary is determined by the laws of reflection and refraction of geometrical optics.[1] We assume that initially all rays are parallel to the axis of the preform, as is indicated in Fig. 11. A given initial ray in the preform and the axis of the perform determine a plane, called the initial plane. If the drawdown zone were axially symmetric, it is easily seen that at every reflection or refraction at the boundary, the ray would remain in that plane. In all the fibers we have considered, the drawdown zone is not axially symmetric. We have continued to make the approximation in our calculations, however, that any ray always remains in its initial plane. The good agreement between theory and experiment leads us to believe the error introduced by this approximation is small. Nevertheless, because the two curves in which the initial plane intersects the drawdown zone are different, any asymmetry of the drawdown is taken into account.

The derivation of the equations for the rays and the caustic curves is essentially the same for the double- and triple-intercept rays. As shown in Fig. 11, we express the boundary curves in the (ξ,η) coordinate system and the equations of the rays and caustics in the (x,y) coordinate system. A given ray is parameterized by the abscissa ξ of the point at which it first intercepts the boundary curve $g(\xi)$ of the drawdown zone. We then trace the ray until it emerges from the preform and write the equation of the emergent ray as

$$a(\xi)x + b(\xi)y = c(\xi). \tag{1}$$

Equation (1) represents a one-parameter family of straight lines with the parameter ξ. The envelope of this family of lines is a caustic curve.[1] To obtain the envelope, we differentiate eq. (1) with respect to ξ, $' = d/d\xi$,

$$a'(\xi)x + b'(\xi)y = c'(\xi). \tag{2}$$

The solution $x = x(\xi)$, $y = y(\xi)$ of eqs. (1) and (2) is the caustic curve in parametric form.[3]

Consider first the family of double-intercept rays, and refer to Fig. 11a. Here normal vectors to the curves are denoted by **n**, tangent vectors by t. The ray is first reflected from the lower curve at the point $[\xi, -g(\xi)]$.

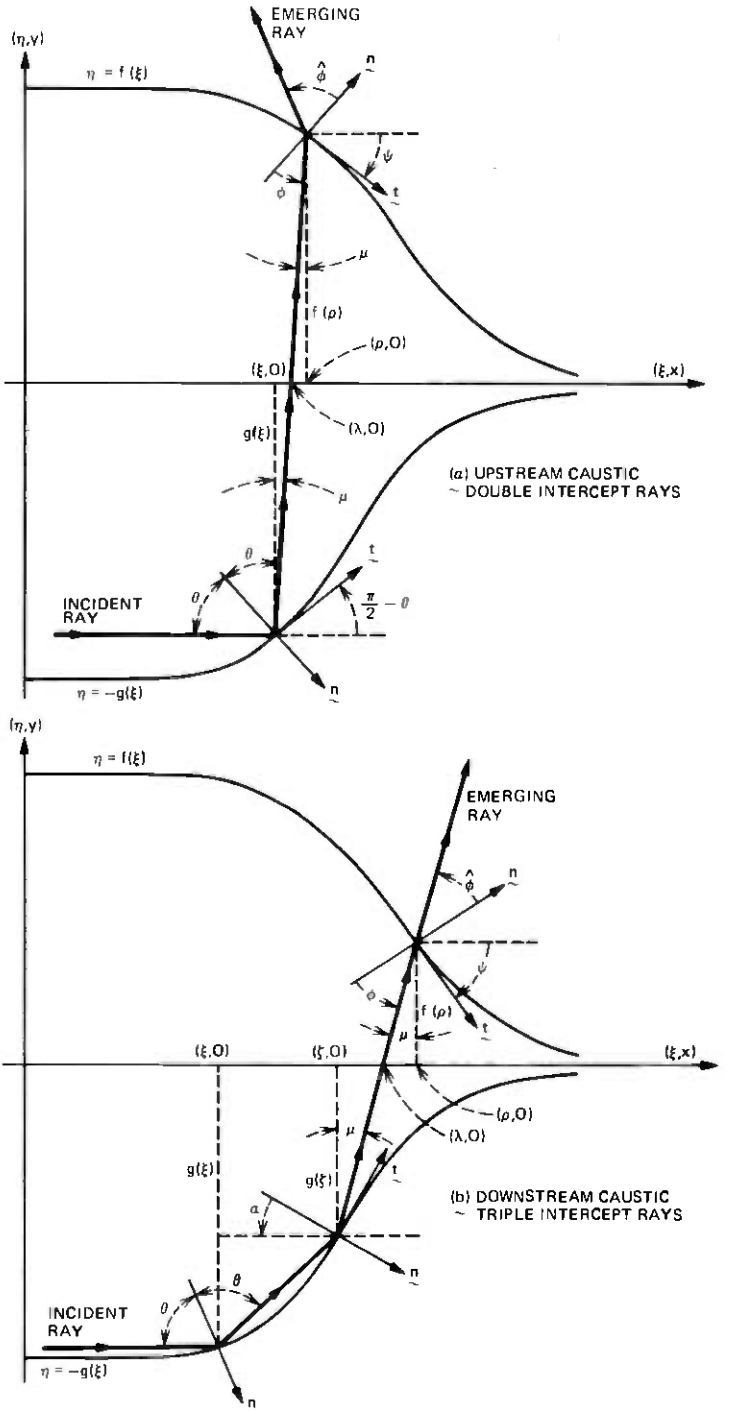


Fig. 11—Nomenclature used for the analysis of the caustic loci and for the ray-tracing calculations. All angles are positive measured in the counterclockwise sense.

Since the angle of incidence equals the angle of reflection, elementary geometry yields

$$\cot\theta = \tan\left[\frac{\pi}{2} - \theta\right] = -g'(\xi), \tag{3}$$

$$\theta' = g''(\xi)\sin^2\theta, \tag{4}$$

$$\mu = 2\theta - \frac{\pi}{2}. \tag{5}$$

Equation (4) is obtained by differentiating eq. (3), and eqs. (3) and (4) together determine $\theta(\xi)$ and $\theta'(\xi)$. Note that $g'(\xi) < 0$, so (3) determines $\theta(\xi)$ as an angle in the first quadrant. The use of eq. (5) for μ and simple trigonometry show that

$$\lambda = \xi - g(\xi)\cot 2\theta. \tag{6}$$

The ray next intersects the upper curve at the point$[\rho,f(\rho)]$. Observing the triangle with vertices $(\lambda,0)$, $(\rho,0)$, and $[\rho,f(\rho)]$, and making use of (5), we find

$$\rho - \lambda = -f(\rho)\cot 2\theta. \tag{7}$$

Since θ and λ are known functions of ξ from (3) and (6), we can determine $\rho(\xi)$ as the solution of (7). If eq. (7) is differentiated with respect to ξ and eqs. (3), (4), and (6) are used, it can be shown that

$$\rho'(\xi) = \left[-g'(\xi) - \frac{g''(\xi)}{g'(\xi)}\{g(\xi) + f(\rho)\}\right] \times \left[\sin 2\theta + \frac{df}{d\rho}(\rho)\cos 2\theta\right]^{-1}. \tag{8}$$

Next, by definition,

$$\tan\psi = \frac{df}{d\rho}(\rho), \tag{9}$$

and by differentiating (9) we get

$$\psi' = \frac{d^2f}{d\rho^2}(\rho)\rho'\cos^2\psi. \tag{10}$$

Simple geometry and Snell's law require that

$$\phi = \frac{\pi}{2} - \psi - 2\theta, \qquad \phi' = -\psi' - 2\theta', \tag{11}$$

$$\sin\hat{\phi} = n\sin\phi, \qquad \hat{\phi}' = (n\cos\phi/\cos\hat{\phi})\phi', \tag{12}$$

where n is the index of refraction of the glass. The equation of the emerging ray is

$$(x - \rho)\cos(\hat{\phi} + \psi) + (y - f(\rho))\sin(\phi + \psi) = 0. \tag{13}$$

If eq. (13) is differentiated with respect to ξ and eq. (9) is used, we get

$$-(x - \rho) \sin(\hat{\phi} + \psi) + [y - f(\rho)] \cos(\hat{\phi} + \psi) = F(\xi), \qquad (14)$$

where

$$F(\xi) = (\rho' \cos \hat{\phi})/\{(\hat{\phi}' + \psi') \cos \psi\}. \qquad (15)$$

Equations (13) and (14) can be solved for x and y to give the parametric equations of the caustic curve

$$x(\xi) = \rho - F(\xi) \sin(\hat{\phi} + \psi), \qquad y(\xi) = f(\rho) + F(\xi) \cos(\hat{\phi} + \psi). \qquad (16)$$

We evaluated $x(\xi)$ and $y(\xi)$ for the double-intercept rays by choosing a value of ξ and then successively determining θ, θ', ρ, ρ', ψ, ψ', ϕ, ϕ', $\hat{\phi}$, and $\hat{\phi}'$ from eqs. (3) through (12). These values were then used to determine $x(\xi)$ and $y(\xi)$ from (15) and (16). Values of $x(\xi)$ and $y(\xi)$ were evaluated at a number of closely spaced values of ξ and then plotted. The treatment of the triple-intercept rays, shown in Fig. 11b, uses the same type of argument. The calculations were carried out on a Honeywell 6000 digital computer. Most of the calculations involved only the use of standard trigonometric subroutines and either the Newton-Raphson method[4] or Brent's algorithm.[5]

The actual profile functions were approximated by sixth-order B-splines[6] possessing four continuous derivatives (a sixth order B-Spline is a piecewise polynomial of degree five). The approximations were obtained by making least squares fits of the B-Splines to the original profile data measured with the Nikon comparator. The fitting was done using an algorithm and subroutine developed by N. L. Schryer of Bell Laboratories.

In forming the caustic of double-intercept rays, only those rays were employed that intercepted the curve $\eta = -g(\xi)$ once and made an angle less than the critical angle with the normal to the surface at its intersection with the curve $\eta = f(\xi)$. Similarly, in forming the caustic of triple-intercept rays, only those rays were used that intercepted the curve $\eta = -g(\xi)$ twice in succession and made an angle less than the critical angle with the normal to the curve $\eta = f(\xi)$ on intersecting it.

In Figs. 12a and 12b, we show the caustics for sample 1 which lie in the 6–12 initial plane. This is the same sample shown in Fig. 1, although the initial plane of Fig. 1 is not the 6–12 plane. In both Figs. 12a and 12b, $f(\xi)$ is profile 6 and $g(\xi)$ is profile 12. Only those caustics corresponding to rays initially incident on profile 12 are shown. These caustics are representative of all those we have calculated, although of course there are important differences of detail. In both figures the heavy curves are the caustics, while the lighter straight lines are typical rays emerging from their point of tangency with the caustic.

For sample 1, all the incident rays were totally internally reflected at

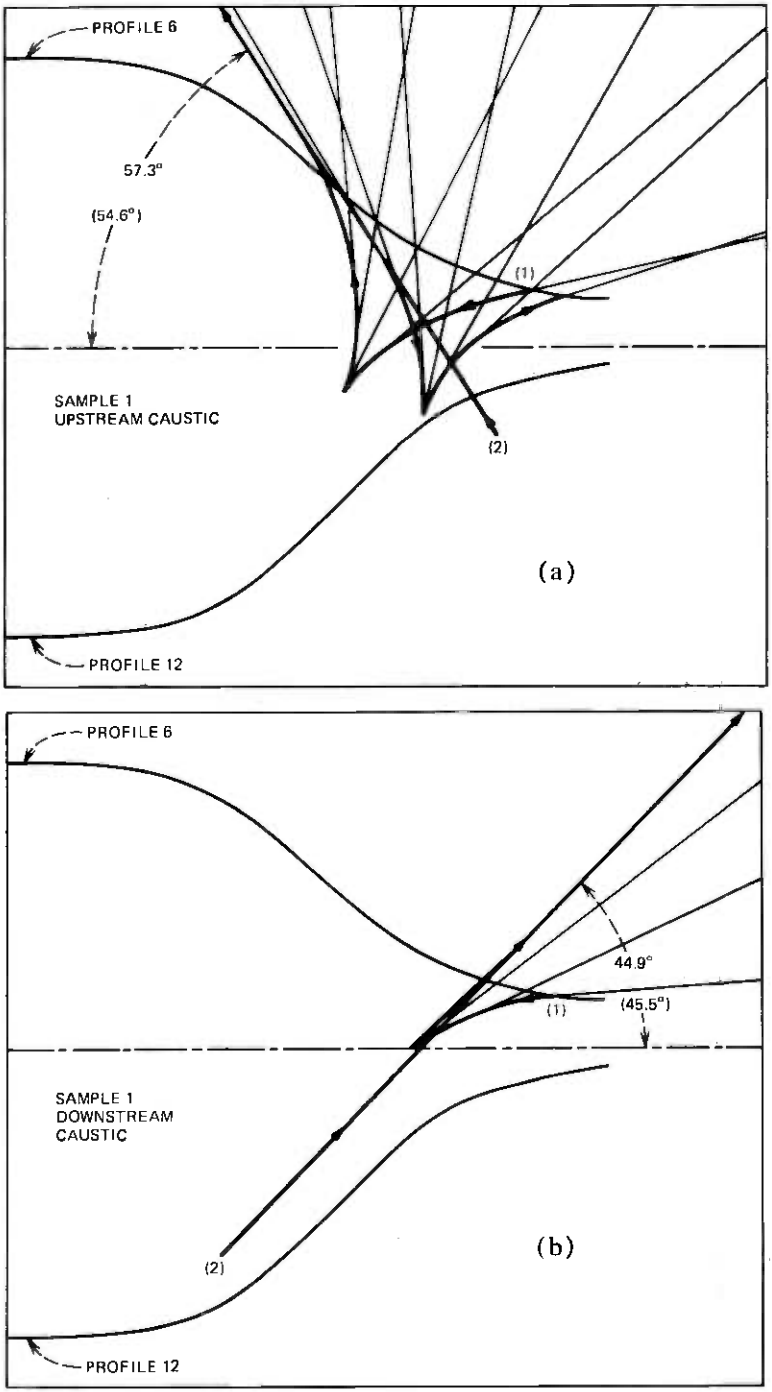


Fig. 12—Calculated loci of the upstream and downstream caustics in sample 1. The observed far-field caustic angles are given in parentheses.

their first intercept with the surface for the double-intercept rays, and for the first two consecutive intercepts for the triple-intercept rays. Sample 1 had a preform radius $r_0 = 3.242$ mm, and all the incident rays of the double-intercept family of Fig. 12a lay in the strip $0.160\, r_0 \leqq g(\xi) \leqq 0.929\, r_0$, while all the incident rays of the triple-intercept family of Fig. 12b lay in the strip $0.942\, r_0 \leqq g(\xi) \leqq 0.973\, r_0$.

The caustic curves themselves for both the double- and triple-intercept families consist of two branches. The point of the caustic $[x(\xi),y(\xi)]$ in Fig. 12a, which corresponds to the smallest permissible value of ξ, is labeled (1). As ξ increases, the caustic is traced out in the direction shown by the arrows. It has a cusp below the axis of the preform and then emerges from side 6 and goes to infinity in the direction shown. The second branch returns from infinity at the bottom of the figure [shown schematically by (2)], crosses the lower boundary, has two cusps, and ends near the upper boundary. Note that all the finite cusps are virtual, that is inside the silica. The important point to note is that the tangents to both branches of the curve at infinity are parallel and are superimposed. This part of the caustic, which goes to infinity, we referred to earlier as the emergent or far-field upstream caustic. The emergent rays thus form a fan parameterized by ξ. As ξ increases, the rays, initially pointing downstream, rotate counterclockwise until they reach the limiting position tangent to the caustic at infinity, and then rotate back in a clockwise direction. It is the accumulation of the rays in the direction of the tangent to the caustic at infinity which produces the observed far-field image. The angle obtained from our calculations between the axis of the preform and the tangent to the caustic at infinity was 57.3°. The measured value was 54.6°. The caustic curve of the triple-intercept family of rays (Fig. 12b), although simpler, has the same general structure. The calculated value of the angle between the axis of the preform and the tangent to the caustic at infinity was 44.9° and the measured value was 45.5°.

Although we have consistently talked about caustic curves, there is also a caustic surface, which is the envelope of the caustic curves. Although this caustic surface is not symmetric about the preform axis, it can nevertheless be roughly visualized by rotating the caustic curves in Figs. 12a and 12b about the preform axis. The upstream caustic surface describes reasonably well the upstream separation between the light and dark regions shown in Fig. 1. However, the downstream separation between the light and dark regions in the drawdown zone has not been completely explained on the basis of the caustics we have been studying.

In Table II we have summarized the calculated and observed angles between the preform axis and the tangents to the caustics at infinity for the four samples studied. The station number labels the profile curve

Table II — A comparison of the observed and calculated caustic angles for four drawdown samples

		Caustic Angles in Degrees, θ			
		Upstream		Downstream	
Sample	Station	Calculation	Observation	Calculation	Observation
1	3	54.5	46.1	46.2	49.4
	6	57.3	54.6	44.9	45.5
	9	75.3	72.5	37.3	40.0
	12	65.3	54.2	42.9	46.8
$2\theta_A = \theta_3 + \theta_9$		129.8	118.6	83.5	89.4
$2\theta_A = \theta_6 + \theta_{12}$		122.6	108.8	87.8	92.3
2	3	No emerging caustic,	(46.2)	NC*	57.2
	6	extinction boundary	(47.6)	49.2	53.0
	9	at 47	(46.9)	NC	56.8
	12		(48.8)	64.6	65.2
$2\theta_A = \theta_3 + \theta_9$		No caustic	(93.1)	NC	114.0
$2\theta_A = \theta_6 + \theta_{12}$		No caustic	(96.4)	113.8	118.2
3	3	48.0[64.1]	46.0[63.4]	83.6	78.3
	6	NC	43.7[63.4]	NC	81.5
	9	49.0[65.1]	45.5[61.3]	80.7	84.6
	12	NC	45.5[62.5]	NC	84.6
$2\theta_A = \theta_3 + \theta_9$		97.0[129.2]	91.5[124.7]	164.3	162.9
$2\theta_A = \theta_6 + \theta_{12}$		NC	89.2	NC	166.1
4	3	NC	101.0	NC	29.0
	6	86.5	86.0	41.9	39.5
	9	NC	86.6	NC	37.2
	12	104.7	105.0	25.5	26.1
$2\theta_A = \theta_3 + \theta_9$		NC	187.6	NC	66.2
$2\theta_A = \theta_6 + \theta_{12}$		191.5	191.0	66.6	65.5

* NC means not calculated.

through which the observed rays emerge. In Fig. 5b we show the observed and calculated caustics in the 6–12 plane for sample 4. The agreement between theory and experiment is very good in most cases. Several cases are of special interest. No upstream caustic was observed in sample 2, and none was exhibited by the model. Two upstream caustics were observed in sample 3, and the model also exhibited these two caustics and the agreement between the model and experiment is good. In short, we feel that our geometrical optics model has explained a variety of complicated effects very well.

V. DISCUSSION

Table II shows the agreement between the experimental observations of the emitted caustic angles and the predicted values. Most of the comparisons agree within 6 percent. The only exceptions are the upstream caustics at the 3 and 12 o'clock orientations on sample 1. As explained in Section 3.3, these upstream caustics are associated with an approach to the critical angle where the geometrical dependence becomes

exaggerated beyond the precision of the profile data. In other words, the accuracy of the profile measurement is insufficient to compute properly the emerging angles, $\hat{\phi}$, when the differential quantity, $d\hat{\phi}/d\phi$, is so large and changing so rapidly, see Fig. 8. The above exceptions are related to a singularity in the dependence of the upstream caustic on the maximum profile slope at which the caustic disappears. The caustic vanishes when the maximum slope approaches unity, causing it to intercept the opposite side of the drawdown at an angle greater than the critical and thus be reflected internally. At a slightly greater slope, depending on the index of refraction, the incidence of the internal rays at the first side is less than the critical angle. Consequently, most of the incident light will refract from the glass instead of reflecting across the drawdown. The low-intensity light which is reflected internally also forms a caustic originating at the inflexion point. That caustic crosses the drawdown zone at even steeper angles as the maximum slope increases until it intercepts the opposite side at an angle less than the critical angle and emerges. For sufficiently steep gradients the fan of light behind the caustic can also arrive at angles incident to the surface within the critical angle and emerge to form yet another caustic by refraction from the curved profile of the drawdown, e.g., sample 3.

The angles given for the upstream caustic directions of sample 2 are shown parenthetically in Table II because they represent limiting rays governed by internal reflection rather than caustics. As such they can be expected to equal angles very nearly tangent to the maximum slopes, Table I, and/or angles complimentary to the critical angle, 43.2°. The observed values satisfy these criteria quite well. The bracketed values given in Table II represent calculated and observed caustic angles for the second upstream caustic generated by sample 3 and discussed earlier.

A comparison of the measured included caustic angles with the average tapers computed for each orthogonal profile pair is shown in Fig. 13. There it can be seen that as the taper, L/D and β decrease, e.g. with slower drawing rates or higher temperatures, the included angles of the upstream caustics likewise decrease and those of the downstream caustics increase. In addition we have found that, aside from the anomalous extinction when β is near the critical angle, the number of caustics increases as L/D decreases, and the interaction between reflected and refracted light wave fronts becomes more complex. Conversely, as L/D increases, the included angle of the downstream caustic decreases continuously until that caustic disappears entirely by internal reflection. In extreme cases, such as long tapered drawdowns from furnaces, the upstream caustic rotates so far downstream that it also reflects internally, and ultimately no light is emitted. In that case ($L/D \gg 1$), the drawdown acts as a waveguide for the light entering the preform and

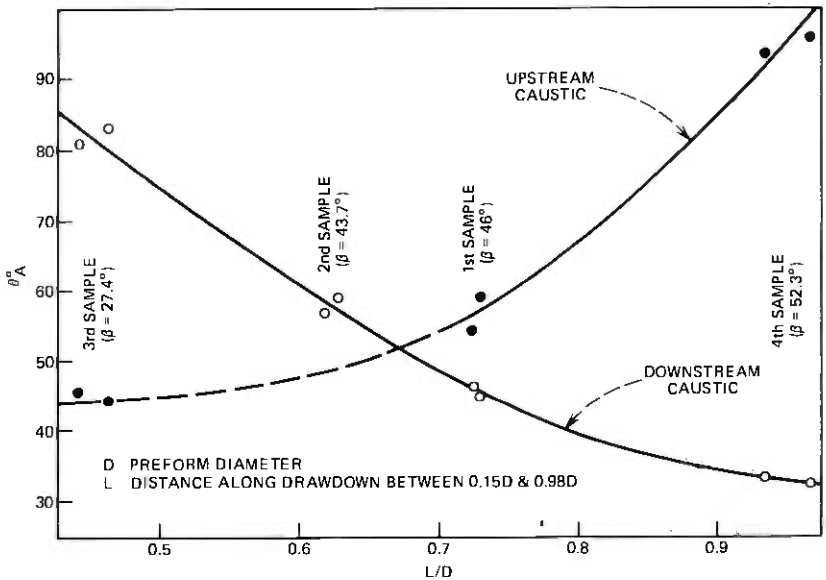


Fig. 13—Correlation of the included angles of the far-field caustics with the drawdown taper.

transmits all of it into the drawn fiber. This behavior is actually observed in furnace-drawn fiber drawdowns.

VI. SUMMARY

Our studies have shown that two unique families of caustics can be formed by light emanating from within a drawdown. We have chosen to designate these two families as "upstream" and "downstream" caustics; their character is rigorously defined in terms of the number of intercepts the trajectory makes with the drawdown profile. A very gradually tapered drawdown will emit no light and no caustics are formed. As the slope increases, light will appear in a downstream direction where the two intercept rays strike the opposite side at less than a critical angle, eventually forming an "upstream" caustic headed *downstream*. As the slope increases still further, the propagation direction of the caustic rotates upstream. Soon the downstream caustic appears and progresses in a similar manner with increasing slope. A singularity in the dependence of the upstream caustic on the maximum profile slope occurs near unity, causing the caustic to disappear by internal reflection.

External illumination of the drawdown zone by a collimated beam produces a far-field caustic pattern. This reflection caustic provides a direct measure of the maximum slope of the drawdown. Obviously the difference between the behavior of the externally illuminated caustic

and the internally illuminated downstream caustic, which is also slope related, will evolve from the optical properties of the glass and the secondary geometrical properties of the drawdown profile.

We have also found that the far-field images of all the caustics exhibit asymmetries related to geometrical asymmetries of the drawdown zone. The downstream caustic asymmetry is a fairly direct measure of the asymmetry of the slopes of the opposite drawdown profiles, while the upstream caustic responds to a more complicated interplay of geometrical factors. The enhanced geometrical sensitivity at incident angles near the critical produces an exaggerated asymmetry sensitivity in far-field upstream caustic images.

VII. ACKNOWLEDGMENTS

We wish to thank the following individuals for their assistance and helpful suggestions: S. I. Feldman, R. E. Jaeger, N. L. Schryer, E. A. Sigety, and J. R. Simpson.

APPENDIX

Estimations of accuracy

Profile measurements of the solidified samples were made on a Nikon optical comparator. Data were recorded every 0.13 mm (0.005 in.) along the profile to an accuracy of 0.005 mm (0.0002 in.). These data were used in the geometrical ray tracing procedures and the numerical computations. During the experiments in which the caustic images were recorded and/or measured, translatory motions were determined to an accuracy of 0.025 mm. All detailed analysis of the caustics were made by direct measurement of the images displayed on suitable screens. The caustic location was established with a metric scale and a vernier rule. We estimate that, with the screen in a horizontal position (e.g., position 1 or 3, Fig. 2), the caustics can be located, with respect to the preform, to within ± 0.1 mm vertically and ± 0.2 mm horizontally. When the beam size was reduced by slits or iris diaphragms, the aperture was measured to within a 0.001 mm with a traveling microscope.

Measurements of the maximum gradient of the drawdown profile using the external reflected caustic (see Section II), were found to be repeatable to within 1 percent. These results were more easily and more accurately obtained by the caustic technique than by analysis of the data obtained by the Nikon comparator.

REFERENCES

1. M. Born and E. Wolfe, *Principles of Optics*, Chap. III, London: Pergamon Press, 1959.

2. *Handbook of Chemistry and Physics*, 55th Edition, CRC Press, 1974.
3. D. J. Struik, *Lectures on Classical Differential Geomtry*, Cambridge, Mass.: Addison-Wesley, 1950, page 73.
4. R. W. Hamming, *Numerical Methods for Scientists and Engineers*, New York: McGraw-Hill, 1962, page 81.
5. R. P. Brent, *Algorithms for Minimization Without Derivatives*, New Jersey: Englewood Cliffs, Prentice-Hall, 1973, page 48.
6. H. B. Curry and I. J. Schoenberg, "On Polya Frequency Functions IV," J. Anal. Math., *17*, 1969, pp. 77–107.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 56, No. 4, April 1977
Printed in U.S.A.

The Effects of Traffic Load Variation on Measurement Accuracy

By D. W. HILL

(Manuscript received June 10, 1976)

A method is developed to quantify the effect of day-to-day variation in offered load on the accuracy of functions of traffic measurements. The method is applicable to any smooth function of the standard trunk-measurements—i.e., peg count, overflow, and usage. As an example, the accuracy of the trunks-required estimator for probability-engineered, full-access trunk groups is approximated. A sensitivity analysis shows that the major contributor to the variance of the estimator is day-to-day variation.

I. INTRODUCTION

Most of the traffic-engineering procedures in use in the Bell System require data collected over periods of up to several days. It has long been recognized that the daily offered loads estimated from trunk-group data show considerable variability even for data taken during the same hour of successive days.[1,2] This variability can cause significant differences between the observed blocking and the objective grade of service. It also can induce large fluctuations in the estimation of network requirements and must be accounted for in the traffic-engineering procedures. (For details on a model for this day-to-day load variation, see the Appendix.)

Two important applications that can be affected by day-to-day load variation are trunk servicing and trunk forecasting. The former is the use of traffic measurements to determine when trunk groups are significantly overloaded or underloaded. In this case, not allowing for day-to-day load variation can cause either repeated rearrangements (churning) or, more typically because of the emphasis on providing good service, overprovision of the traffic network. Trunk forecasting is the prediction of future network requirements. The accuracy of the forecast is strongly influenced by the day-to-day load variation. That accuracy, in turn, affects the procedures used to implement the forecast.

Most earlier work has attempted to quantify the effect of day-to-day load variation on averages of functions of traffic data.[2–6] Several trunk-engineering practices are now based on that work.[7] However, most of the previous studies that have examined statistical accuracy (as measured by variances) of functions of traffic measurements other than offered load have assumed that the true offered load (also called the source load) was a constant (e.g., Ref. 8). The purpose of this study is to provide an extension of the earlier work on traffic-measurement accuracy (which usually considered the stochastic nature of traffic and the effects of finite sampling) to include the effect of day-to-day load variation.

The most general work on the accuracy of single-hour measurements is that of Neal and Kuczura.[8] That work is used as a starting point for a more general model developed in Section II. The new model can be used to estimate the accuracy of any sufficiently differentiable function of the standard traffic measurements—i.e., peg count (number of arrivals), overflow, and usage. The model is used in Section III to approximate the standard deviation of estimates of the number of trunks required for probability-engineered groups. Section IV illustrates the application of the results of Section III to trunk servicing.

This paper uses concurrent work on mathematical models for day-to-day variation.[9] The reader should be familiar with that work, or for a short description of the main results, see the Appendix.

II. THEORETICAL RESULTS

For completeness, a brief review of notations and definitions is included here.[8] On each day, the measurements are taken over a time period denoted as $(0,t]$, with t usually taken to be one hour. The standard trunk-group measurements are:

(*i*) $A(t)$ is the measured number of arrivals (peg count) in $(0,t]$.

(*ii*) $0(t)$ is the measured number of overflows in $(0,t]$.

(*iii*) $L_d(t)$ is the measured usage based on a discrete scan [typically by a 100-second-scan traffic usage recorder (TUR)] of the number of busy trunks in $(0,t]$.*

It was found during this study that when day-to-day variation is included, the additional effect of the sampling errors in $L_d(t)$ is negligible for data from the message trunk network (see Section 3.3.5). Consequently, $L(t)$ will be used throughout the paper with the results being equally valid for the discrete, 100-second scan measurement, $L_d(t)$. The triple of measurements $(A(t),O(t),L(t))$ is taken for an interval of

* A continuous scan of the number of busy trunks (i.e., the total usage) is considered and is denoted by $L(t)$.

length t on each of n days. The total collection of data is described as $(A_i(t),O_i(t),L_i(t))$, $i = 1,\cdots,n$.

For the model assumed in this study, the trunk group contains c trunks, whose holding times are assumed to be from a negative exponential distribution with mean h. During the ith measurement interval, the arrival epochs form a renewal process with mean interarrival time λ_i^{-1}. The offered load during the ith interval is assumed to be constant and given by $a_i = \lambda_i h$, and the peakedness of the traffic is z.[1] The offered source loads $a_1,\cdots,a_n$ are assumed to be independent and identically distributed (iid) according to a specified probability distribution, Γ, which will be assumed to be a gamma distribution.[4] Note that because a_i and a_j, $i \neq j$, are independent, the processes associated with them are also independent.*

Any customer who arrives when there is an idle server will enter service immediately. Because this is a study for trunk groups with typically low blocking, a customer arriving to find all servers busy is assumed to depart and has no further effect on the system; i.e., customer retrials will be ignored.

2.1 The approximation

Let ξ_{ij}, $j = 1,2,3$, $i = 1,\cdots,n$ be the $3n$ random variables representing the data; i.e., $\xi_{i1} = A_i(t)/t$, $\xi_{i2} = O_i(t)/t$, and $\xi_{i3} = L_i(t)/t$. For a fixed i, each ξ_{ij}, $j = 1,2,3$ is a random variable whose parameters are functions of another random variable a_i. Denote the mean and conditional mean of ξ_{ij} by

$$\bar{\theta}_j = E[\xi_{ij}] = E[E(\xi_{ij}|a_i)],$$

and

$$\theta_j(a_i) = E[\xi_{ij}|a_i].$$

Then, setting $\underline{\xi} = (\xi_{11},\xi_{12},\xi_{13},\xi_{21},\cdots,\xi_{n3})$ implies that the mean of $\underline{\xi}$ is

$$\underline{\bar{\theta}} = E(\underline{\xi}) = (\bar{\theta}_1,\bar{\theta}_2,\bar{\theta}_3,\bar{\theta}_1,\cdots,\bar{\theta}_3).$$

Now consider any differentiable function $g(x_{11},\cdots,x_{n3}) = g(\underline{x})$. Expand g in a Taylor series about $\underline{\bar{\theta}}$ neglecting terms of order greater than one to get

$$g(\underline{\xi}) \approx g(\underline{\bar{\theta}}) + \sum_{i=1}^{n}\sum_{j=1}^{3} \left.\frac{\partial g(\underline{x})}{\partial x_{ij}}\right|_{\underline{x}=\underline{\bar{\theta}}} (\xi_{ij} - \bar{\theta}_j). \tag{1}$$

* Studies have shown that the day-to-day variation in peakedness is small and is negligible for most network engineering applications. Recent studies have also shown that the effect of a systematic variation in load (e.g., as a function of the day of the week) is also negligible.[10] Finally, simulation data from this study have indicated that including the effect of calls with different exponential holding times on one group (e.g., effective and ineffective attempts) have little effect on the results.

The mean of $g(\xi)$ is approximated by

$$E[g(\xi)] \approx g(\theta) + \sum_{i=1}^{n} \sum_{j=1}^{3} \left. \frac{\partial g(\underline{x})}{\partial x_{ij}} \right|_{\underline{x}=\underline{\theta}} E(\xi_{ij} - \bar{\theta}_j)$$
$$= g(\underline{\theta}).$$

For the variance of $g(\xi)$,

$$E[g(\xi) - g(\underline{\theta})]^2 \approx E \left\{ \left[\sum_{i=1}^{n} \sum_{j=1}^{3} \left. \frac{\partial g(\underline{x})}{\partial x_{ij}} \right|_{\underline{x}=\underline{\theta}} (\xi_{ij} - \bar{\theta}_j) \right] \right.$$
$$\left. \times \left[\sum_{k=1}^{n} \sum_{l=1}^{3} \left. \frac{\partial g(\underline{x})}{\partial x_{kl}} \right|_{\underline{x}=\underline{\theta}} (\xi_{kl} - \bar{\theta}_l) \right] \right\}$$
$$= \sum_{i=1}^{n} \sum_{j=1}^{3} \left. \frac{\partial g(\underline{x})}{\partial x_{ij}} \right|_{\underline{x}=\underline{\theta}} \sum_{k=1}^{n} \sum_{l=1}^{3} \left. \frac{\partial g(\underline{x})}{\partial x_{kl}} \right|_{\underline{x}=\underline{\theta}}$$
$$\times E[(\xi_{ij} - \bar{\theta}_j)(\xi_{kl} - \bar{\theta}_l)]. \quad (2)$$

Because a_i and a_k are independent for $k \neq i$, the expectation in (2) vanishes unless $k = i$, in which case it is[11]

$$E[(\xi_{ij} - \bar{\theta}_j)(\xi_{il} - \bar{\theta}_l)] = E \operatorname{Cov}(\xi_{ij}, \xi_{il} | a_i) + \operatorname{Cov}(\theta_j(a_i), \theta_l(a_i)). \quad (3)$$

Because (3) does not depend on the subscript i (the a_i are iid), we may drop it and replace ξ_{ij} by $\xi_{\cdot j}$ and ξ_{il} by $\xi_{\cdot l}$ in the sequel. Substituting (3) into (2) provides the approximation for the variance*

$$E[g(\xi) - g(\underline{\theta})]^2 \approx \sum_{i=1}^{n} \sum_{j=1}^{3} \sum_{l=1}^{3} \left. \frac{\partial g(\underline{x})}{\partial x_{ij}} \right|_{\underline{x}=\underline{\theta}} \left. \frac{\partial g(\underline{x})}{\partial x_{il}} \right|_{\underline{x}=\underline{\theta}}$$
$$\times [E \operatorname{Cov}(\xi_{\cdot j}, \xi_{\cdot l} | a) + \operatorname{Cov}(\theta_j(a), \theta_l(a))]. \quad (4)$$

2.2 Computational considerations

The term in brackets in (4) is given by

$$\int [\operatorname{Cov}(\xi_{\cdot j}, \xi_{\cdot l} | a) + (\theta_j(a) - \bar{\theta}_j)(\theta_l(a) - \bar{\theta}_l)] d\Gamma(a). \quad (5)$$

The functions in the integrand given in Ref. 8 are too complicated for the integral to be computed exactly, hence numerical quadrature is required.

In previous work several different quadrature schemes have been used on integrals similar to that in eq. (5). Because the functions are usually smooth, these schemes are generally successful. For this study, a compound 7th-order Newton-Cotes form was chosen.[12] The tails of the gamma distribution tend to zero sufficiently quickly that the infinite region of integration can be truncated to a finite region with no problem.

* When $g(\xi)$ includes the averaging of n days of data, each term $\partial g(\underline{x})/\partial x_{ij}$ contains the factor $1/n$, so that the variance of $g(\xi)$ is of the order of $(1/n)$.

III. EXAMPLE: VARIABILITY OF THE TRUNKS-REQUIRED ESTIMATE

The methods described in Section II were applied to the specific problem of computing the standard deviation, $\sigma(\hat{c})$, of the trunks-required estimate, $\hat{c}$, for probability engineered full-access trunk groups. The function g in this case is defined by a set of algorithms described in the Appendix. The partial derivatives of g are approximated by divided differences. Approximations were computed for several typical cases to cover a reasonable range of engineering interest. To test the accuracy of the approximations, they were compared with corresponding sample standard deviations from a simulation. These results are described in the following two sections. An application using these results to compute probability intervals for estimates of trunk required is illustrated in Section IV.

3.1 Basic calculations

For the main results, the existing trunk group size is fixed and the mean offered load $\bar{a}$, peakedness z, and levels of day-to-day variation are varied over the range of interest. Following Bell System practice, it is assumed that measurements are taken over a 20-day period (i.e., 20 independent one-hour measurement intervals) and that the trunk group is to be designed for an average-blocking objective of 0.01 (denoted $\overline{B}0.01$). It is also assumed that the calls have a mean holding time of 180 seconds. The sensitivity of the results to these assumptions are described in Section 3.3.

To validate the theoretical computations, sample variances from a simulation program were computed for each of several sets of input conditions (each variance was computed from a sample of size 50, which was large enough to give stable results and still be computationally feasible). For some input sets, the simulation runs were repeated to provide an indication of the variability of the estimated standard deviation. (An analytic approach would require computations of the 4th-order moments and cross-moments of the measurements and was not practical.)

3.2 Results

Results were computed for trunk groups ranging in size from 10 to 68 circuits. Illustrations of typical results are presented in Fig. 1 showing plots of $\sigma(\hat{c})$ as a function of input offered load and peakedness on trunk groups with 68 circuits. It has been observed in actual data that peakedness and level of day-to-day variation are correlated. Hence, combinations of peakedness and levels of day-to-day variation were selected to cover most values encountered in practice, with $z = 1$ and low variation selected to illustrate groups which first-routed traffic; $z = 4$ and medium,

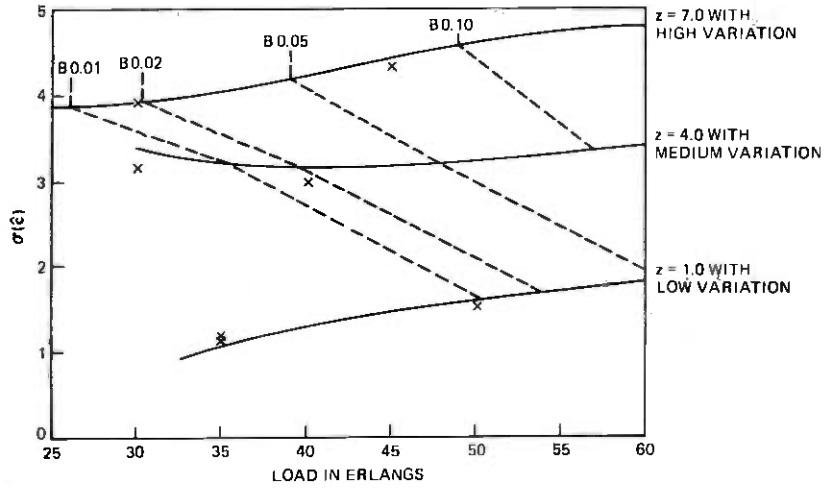


Fig. 1—Standard deviation of trunk estimate, $\sigma(\hat{c})$, vs offered load on a 68-trunk group.

and $z = 7$ and high were related to illustrate groups serving overflow traffic. In all cases, it was assumed that measurements are taken on a trunk group of the specified size, and the number of trunks needed to achieve an average blocking of 0.01 was estimated.

The input loads in general correspond to a range of blocking values from less than 0.01 to greater than 0.15. Except for $z = 1$, $\sigma(\hat{c})$ has a minimum at a load that corresponds to an observed blocking in the 0.02 to 0.03 range.* For lower blocking values and $z > 1$, $\sigma(\hat{c})$ increases as the load decreases because the coefficients of variation (standard-deviation-to-mean ratio) of the measurements, especially overflow, increase. As the load increases, $\sigma(\hat{c})$ also increases; however, the coefficient of variation of trunks required decreases slowly, probably because the coefficient of variation of the offered load decreases with increasing $\bar{a}$, causing the coefficients of variation of the measurements to decrease.[9]

3.3 Parameter sensitivity

The sensitivity of $\sigma(\hat{c})$ to the various parameters is illustrated in Figs. 2 and 3 and is described below.

3.3.1 Blocking objective

As illustrated in Fig. 2a, the first parameter tested was the design blocking. For the design range of 0.01 to 0.03,† there is very little change

* As discussed in the Appendix, when $z = 1$, the peakedness is not estimated, which causes the $\sigma(\hat{c})$ curve to have a different shape.

† The Bell System design objective is 0.01, but in some private networks and other administrations, higher values are used.

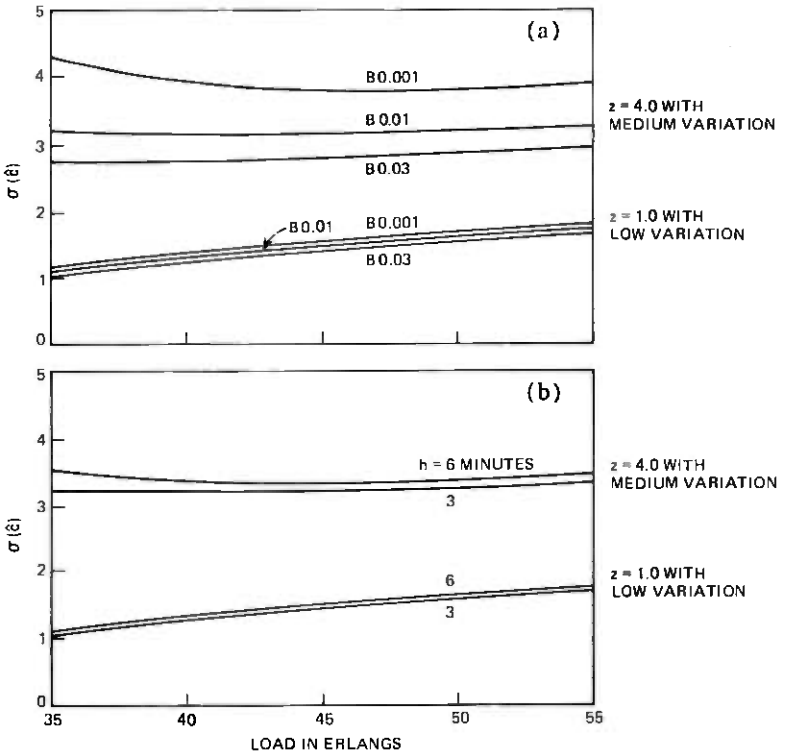


Fig. 2—Parameter sensitivity for a 68-trunk group. (a) Blocking objective. (b) Call-holding time.

in $\sigma(\hat{c})$ when $z = 1$ and low variation. For blocking in this range, the results for 0.01 can be used as an upper bound. Changes in the design blocking have more of an effect for more variable data, as illustrated by the data for $z = 4$ and medium variation.

3.3.2 Call-holding time

As the call-holding time h increases, the relative length of the one-hour measurement interval decreases. The result is a relative decrease in the amount of data available and a resultant increase in the standard deviation of the measurements.[8] However, for a fixed observed-load variance and for holding times in the range of 3 to 6 minutes, the effect is mostly offset by a decrease in the true day-to-day variation of the source load.[9] This is illustrated in Fig. 2b.

3.3.3 Level of day-to-day variation

The assumption to which the results are most sensitive is the level of day-to-day variation. The day-to-day variation of offered source-load is characterized by four levels, called *no*, *low*, *medium*, and *high*.

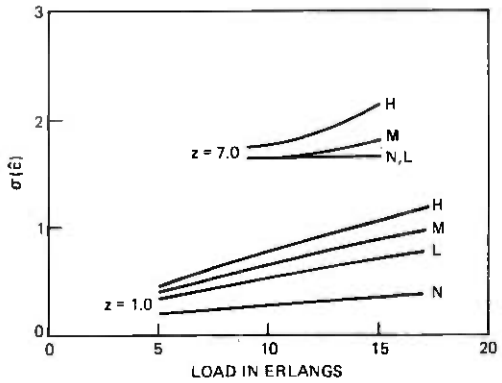


Fig. 3—Effect of level of day-to-day variation on a 10-trunk group.

Figure 3 provides a comparison of results for different levels of variation for a 10-trunk group. The labels, *N*, *L*, *M*, and *H* on the figure correspond to *no*, *low*, *medium*, and *high* variation, respectively.* The largest relative impact is for the case $z = 1$ when changing from *no* to *low* variation. In fact, the standard deviation of the trunk estimate doubles when such a change is made. Thus, inclusion of day-to-day variation in the model has a significant effect on the standard deviation of trunk estimates.

For large peakedness, shown by $z = 7$, as the load decreases, the curves coalesce. This rather unexpected behavior can be explained by the model for day-to-day variation (see Ref. 9). For groups with small loads, low day-to-day variation, and large z, the variance component due to sampling in a finite measurement-interval may be most or all of the total variance of the observed load. In this situation, the day-to-day component in the model of observed-load variance decreases to zero. For most practical applications, the large values of z are associated with medium or high levels of day-to-day variation, and this phenomenon does not occur.

3.3.4 Number of samples

As stated earlier, all of the plots are based on an average of 20 hours of data. If the number of hours, n, of available data is different from 20, the effect can be determined analytically from (4). The standard deviation for n hours is computed by multiplying the given results by the factor $(20/n)^{1/2}$.

3.3.5 Effect of usage sampling errors

The effects introduced by a discrete measurement of usage (TUR)

* These levels are defined in the Appendix.

Table I — Comparison of continuous and discrete usage measurement effects

				$\sigma(\hat{c})$ using	
c	a	z	*Level*	L_d	L
10	7.0	1.0	Low	0.4007	0.4372
10	7.0	4.0	Medium	1.012	1.043
40	30.0	4.0	Medium	2.138	2.174
40	30.0	7.0	High	2.949	2.999
68	40.0	7.0	High	4.172	4.211

were found to be negligible for the message-trunk network. Typical results of calculations using $L_d(t)$ and $L(t)$ are presented in Table I. The first four columns of Table I are the trunk group size, load, peakedness, and level of day-to-day variation. The next two columns are the approximations of the standard deviation of the trunk estimates with $L_d(t)$ and $L(t)$, respectively. The difference between the last two columns is negligible for traffic-engineering applications. Note that the relationship of the two columns is the opposite of what might be expected. This results from a bias in the asymptotic approximation for $\mathrm{Var}[L(t)]$ for the small loads included in the region of integration.[8]

IV. APPLICATION: PROBABILITY INTERVALS FOR TRUNK ESTIMATES

One of the first applications for the methods described in this paper was the development of probability intervals for trunk estimates. The intervals are used to determine if the estimated number of trunks required for a given circuit is (statistically) significantly different from the number presently in service. If the difference between the estimate and the current number is within the interval, then that difference is considered to be the result of the statistical nature of the data. Such a difference should not be the cause for action.

These intervals have application in two different areas of the trunk-engineering process. First, they provide an upper limit for the accuracy that can be attained by the trunk-forecasting process. Sources of error that have not been included here, such as wiring errors and load-projection errors, must increase the variability of the data. Second, they enable a trunk-servicer to evaluate the output of a mechanized trunk-servicing system and to determine if (and where) network rearrangements are necessary.

Data from the simulation described earlier indicate that the estimates of trunks required appear to have a normal probability distribution, with mean and variance computed as described in Section II. Using this in-

* This is not true when day-to-day variation is ignored.[8]

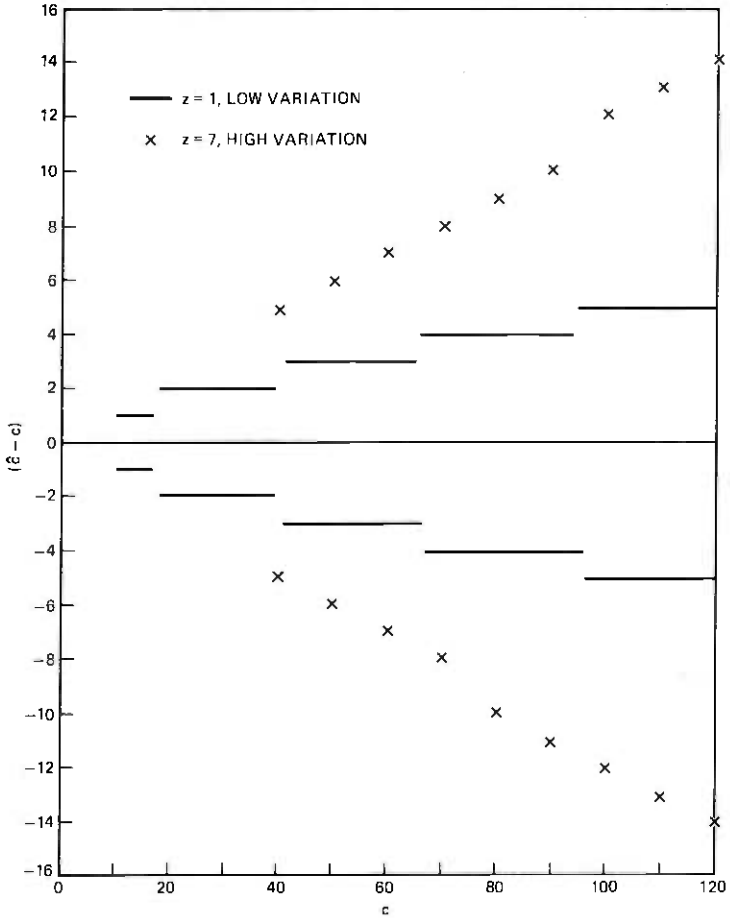


Fig. 4—Ninety-five-percent probability intervals for trunk estimates.

formation, the probability intervals for the trunk estimates have been constructed. Examples of such probability intervals are shown in Fig. 4. The solid line on Fig. 4 shows 95-percent intervals for trunk groups with $z = 1$ and low day-to-day variation. For example, for a correctly engineered 50-trunk group, 95 percent of the estimates of trunks required, based on traffic measurements, will lie between 47 and 53 trunks. The crosses on Fig. 4 show similar data for traffic with $z = 7$ and high variation.

The method described in Section II has also been successfully applied to approximate the variance of observed blocking for probability-engineered trunk groups. That result has been used to develop intervals of acceptable measured blocking for use in network servicing.

V. SUMMARY

A method for computing the effect of day-to-day variation in offered load on functions of traffic measurements has been presented. The method, which is applicable to any smooth function of the common traffic measurements, was used to compute the standard deviations of estimates of trunks required for probability-engineered trunk groups. In the associated sensitivity analysis, the daily variation in source load was identified as the significant contributor to the total variation of the estimate. In fact, day-to-day variation was so large that it was possible to neglect errors introduced by a discrete 100-second-scan measurement of the usage. (This extends a result derived analytically by Hayward for traffic with Poisson arrivals.[13])

The variability in trunk estimates depends very strongly on levels of peakedness and day-to-day variation. For first-route traffic and low day-to-day variation, probability intervals for trunk estimates grow slowly with c, while for higher levels of variation they expand rapidly. These results are presently being used to develop methods to assist trunk engineers in the forecasting and servicing of the traffic network.

APPENDIX

Details of the Computational Models

This appendix contains some of the engineering details necessary for the computations in Sections III and IV. The first section gives a brief description of the model for day-to-day load variation used in Section 3.1.[9] The second section discusses the conversion of traffic measurement data into estimates of the trunks required to meet an objective grade-of-service.

A.1 Model for day-to-day variation

Four levels of day-to-day load variation described as *no*, *low*, *medium*, and *high* are used for trunk engineering. For the latter three classes, the variance v of the measured (observed) offered loads is related to the mean offered load $\bar{a}$ by the formula

$$v = 0.13\,(\bar{a})^{\phi},$$

where $\phi = 1.5$, 1.7, or 1.84 for low, medium or high, respectively. The mean $\bar{a}$ is assumed to be constant during the measurement period (any variation of $\bar{a}$ during the measurement period will cause the estimated peakedness to be larger).

The variance of the observed loads is composed of two parts: the true source-load variance and the variance contributed by estimating the traffic parameters from data collected over a finite measurement interval.[9] The latter component is given by $2\bar{a}z/(t/h)$ where $\bar{a}$ is the mean

of the daily loads, z is the peakedness of the offered traffic, t is the length of the measurement interval, and h is the mean call holding time. Thus, by subtraction, the source-load variance is assumed to be

$$\mathrm{Var}(a) = \max\left\{0.13\bar{a}^{\phi} - \frac{2\bar{a}z}{(t/h)}, 0\right\}.$$

A more detailed discussion is given in Ref. 9.

A.2 Trunk-engineering process

The trunk-engineering process starts with an estimation of the traffic parameters obtained from trunk-group measurements. Time-consistent busy-hour measurements of the number of arrivals (peg count), the number of overflows, and usage are gathered for a period of several days (up to 20 business days when all data are available). They are then used to estimate the mean of the busy-hour loads and the peakedness of the offered traffic. The mean load is computed by averaging the hourly loads

$$\hat{a}_i = \frac{L_{di}(t)}{1 - \dfrac{O_i(t)}{A_i(t)}},$$

where $L_{di}(t)$, $O_i(t)$, and $A_i(t)$ are defined in Section II. The sample mean, $\hat{a} = 1/n\ \Sigma\ \hat{a}_i$ and sample variance $\hat{v} = 1/(n-1)\ \Sigma\ (\hat{a}_i - \hat{a})^2$ are computed next. In practice, the level of day-to-day load variation is selected by picking the value of ϕ as the one that provides the closest agreement between v and $\hat{v}$. For the computations here, ϕ was assigned by the program input. The next step in the traffic-engineering process is to apply a correction for the effect of retrials on $\hat{a}$ and $\overline{B} = 1/n\ \Sigma\ [O_i(t)/A_i(t)]$. (Because the region for the usual application of these results was for small blocking values, the retrial correction was not included in the analysis.)

The traffic peakedness is estimated by an iterative procedure. Given c, $\hat{a}$, and $\overline{B}$, a preliminary estimate of z is determined so that the theoretical blocking predicted by the equivalent random method matches the observed blocking.[1] The preliminary estimate of z is adjusted to correct for day-to-day variation using the procedure described in Ref. 7 to give the corrected estimate, $\hat{z}$, of the peakedness. In the case of trunk groups known to serve only first-offered traffic (i.e., none of the traffic has overflowed from some other group), the theoretical value of $z = 1$ is assumed and no estimation of peakedness is performed.

Once $\hat{a}$ and $\hat{z}$ have been determined, the number of trunks required to satisfy the engineering objective (usually $\overline{B}0.01$) can be determined from established trunk-capacity tables or appropriate computer algo-

rithms. These algorithms specify that any fractional trunk-requirement will be rounded up unless it is less than 0.3. This rounding rule induces the slight nonsymmetry seen in Fig. 4.

REFERENCES

1. R. I. Wilkinson, "Theories for Toll Traffic Engineering in the U.S.A.," B.S.T.J., *35*, No. 2 (March 1956), pp. 421–454.
2. C. Palm, "Waiting Times when Traffic has Variable Mean Intensity," Ericsson Review, *24* (1947), pp. 102–108.
3. L. von Sydow, "Some Aspects on the Variations in Traffic Intensity," Teleteknik, *1* (1957), pp. 58–64.
4. R. I. Wilkinson, "A Study of Load and Service Variations in Toll Alternate Route Systems," Proc. Second Int'l. Teletraffic Congress, The Hague (1958), Paper No. 29.
5. W. S. Hayward and R. I. Wilkinson, "Human Factors in Telephone Systems and Their Influence on Traffic Theory, Specially with Regard to Future Facilities," Proc. Sixth Int'l. Teletraffic Congress, Munich (1970), Paper No. 431.
6. R. I. Wilkinson, "Some Comparisons of Load and Loss Data with Current Teletraffic Theory," B.S.T.J., *50*, No. 8 (October 1971), pp. 2807–2834.
7. R. I. Wilkinson, "Nonrandom Traffic Curves and Tables," unpublished work.
8. S. R. Neal and A. Kuczura, "A Theory of Traffic Measurement Errors for Loss Systems with Renewal Input," B.S.T.J., *52*, No. 10 (December 1973), pp. 967–990.
9. D. W. Hill and S. R. Neal, "The Capacity of a Probability-Engineered Trunk Group," B.S.T.J., *55*, No. 7 (September 1976), pp. 831–842.
10. P. J. Burke, "Sensitivity of Average Blocking and Overflow to Finite Population and High-Day, Low-Day Effects," unpublished work.
11. H. H. Panjer, "On the Decomposition of Moments by Conditional Moments," Amer. Statist., *27*, No. 4 (1973), pp. 170–171.
12. E. Isaacson and H. B. Keller, *Analysis of Numerical Methods*, New York: John Wiley, 1966, pp. 308ff.
13. W. S. Hayward, "The Reliability of Telephone Traffic Load Measurements by Switch Counts," B.S.T.J., *31*, No. 2 (March 1952), pp. 357–377.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 56, No. 4, April 1977
Printed in U.S.A.

Reliability Considerations for Multiple-Spot-Beam Communication Satellites

By A. S. ACAMPORA

(Manuscript received September 29, 1976)

Reliability considerations associated with multiple-spot-beam satellite systems are explored. If each coverage area is serviced by a single transponder, then loss of transponders due to failure eliminates all service to the areas covered by those transponders. Thus, failures are quite costly compared to a system employing global coverage with multiple transponders, where a limited number of transponder failures results in a slight increase in the traffic demand upon the survivors. Since the total orbital weight of a satellite is fixed, any redundant hardware deployed to improve reliability reduces the number of active transponders that can be supported, and a highly efficient redundancy strategy must be employed. Cold standby redundancy with complete spare interconnectivity is studied and appropriate reliability formulas are derived. A specific satellite concept dominated by final power amplifier failures is studied in detail, and it is found, for typical failure rates, that a 27-percent reduction in capacity must be accepted to provide for a single satellite lifetime reliability of 99 percent. Various techniques for employing the in-orbit redundancy of a spare satellite are investigated to increase reliability while minimizing capacity reduction. Bounds are derived and a reliability of at least 99 percent is shown possible for a system containing three active satellites plus one spare, at a cost of 9 percent in potential capacity of the active satellites.

I. INTRODUCTION

Multiple-spot-beam antennas for communication satellites offer the potential for greatly increasing the traffic capacity of the satellite since the allocated frequency band can be reused in the various spot beams.[1–4] In such a configuration, we might reduce the number of required satellite transponders by allocating a single wideband transponder, consisting of a receiver and a transmitter, to each antenna port as shown in Fig. 1.

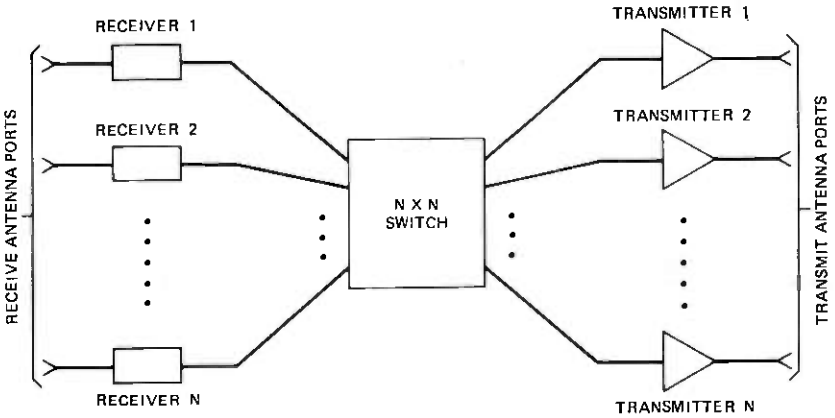


Fig. 1—Satellite transponder.

An on-board switch is provided to route messages originating within the coverage area of one antenna beam to their destinations in the coverage areas of other beams. Access to the satellite is by time division multiple access (TDMA), and the up and down link frequency bands are the same for all beams. This arrangement is often referred to as satellite-switched TDMA.

If each coverage area is serviced by a single transponder, failure of that transponder eliminates all communications to that area. Thus, reliability requirements are generally much higher than in earlier satellite systems employing frequency-division multiple access global beam coverage wherein failure of a single transponder results in partially reduced traffic handling capability to all users.[5,6]

The higher reliability requirement of multiple-spot-beam satellite systems implies that either (*i*) additional redundant hardware must be provided and/or (*ii*) a more efficient strategy be adopted for deploying existing redundant units. Since the total orbital weight of the satellite is fixed by launch vehicle capability, it is desirable to maintain the weight of redundancy-related equipment at a minimum in order that as much weight as possible be available for active communication transponders. In this paper, several configurations employing cold standby redundancy[7–10] are proposed and explored, and additional weight vs reliability trade-offs are established. Consideration is limited to failures of the final high power traveling wave tube amplifiers (TWTAs) since the relatively high failure rate and large weight of these devices would dominate in a more detailed study. Exact results are provided for single satellite systems; from these, reliability bounds for multiple satellite systems, including an in-orbit spare, are derived for a variety of philosophies for spare satellite utilization.

In Section II, the mathematical formulas pertaining to single satellite reliability employing cold standby M-on-N redundancy (M standby spare units for N active units with full interconnectivity) are derived assuming that the cold failure rate is zero. From these, we can find, at each point in time, the probability that j transponders have failed, $0 \leqq j \leqq N$.

These formulas are applied in Section III to a specific satellite concept, appropriate for a Thor-Delta launch, to determine transponder reliability vs capacity trade-offs under a weight constraint. Also investigated in Section III is the interconnectivity trade-off; that is, reliability vs total capacity for conditions other than full interconnectivity.

In Section IV, upper and lower reliability bounds are derived for satellite systems containing an arbitrary number of active satellites plus an identical in-orbit spare. Various approaches to spare satellite utilization, differing in degree of complexity, are explored. Results indicate that significant reliability improvements can be achieved through appropriate utilization of the spare satellite. These results, however, are to be interpreted as indicative of a trend rather than as concrete design tools since the probability of catastrophic failure of an active satellite, requiring complete utilization of the spare satellite, was assumed to be zero in deriving the results.

II. RELIABILITY FORMULAS

In this section, a model for predicting the reliability of a multibeam communication satellite employing M-on-N standby redundancy is developed and analyzed. Complete interconnectivity is assumed, that is, any of the M cold standby redundant units may be substituted for any of the N active units upon failure of one of the latter. The cold standby failure rate is assumed to be zero, as is the failure rate of the interconnecting redundancy switches, and the failure rate for each active unit, λ, is assumed to be constant. Failures are assumed to be characterized by a Poisson point process. An expression for the probability that N units are operational at any point in time is easily established;[11] this result is extended to find the probability of having j active channels as a function of time, $0 \leqq j \leqq N$.

Let the system state j be the number of operational units at time t, $0 \leqq j \leqq N + M$, and let $P_j(t)$ be the probability of finding the system in state j at time t. Then, since there are N active units for states $j = N, N + 1, \ldots, N + M$, the following relationships hold for those states:

$$P_{N+M}(t + \Delta t) \cong (1 - N\lambda\Delta t)P_{N+M}(t), \tag{1}$$

$$P_j(t + \Delta t) \cong (1 - N\lambda\Delta t)P_j(t) + N\lambda\Delta t P_{j+1}(t), \; N \leqq j < N + M, \tag{2}$$

where the approximations become exact as $\Delta t \to 0$. Similarly, since there are j active units for states $j = 0, 1, \ldots, N$,

$$P_j(t + \Delta t) \cong (1 - j\lambda\Delta t)P_j(t) + (j + 1)\lambda\Delta t P_{j+1}(t),\ 0 \leqq j < N. \quad (3)$$

Letting $\Delta t \to 0$, the following set of differential equations results for the system state probabilities:

$$\frac{dP_{N+M}}{dt} = -N\lambda P_{N+M}. \quad (4)$$

$$\frac{dP_j}{dt} = -N\lambda P_j + N\lambda P_{j+1},\ N \leqq j < N + M. \quad (5)$$

$$\frac{dP_j}{dt} = -j\lambda P_j + (j + 1)\lambda P_{j+1},\ 0 \leqq j < N. \quad (6)$$

Initial conditions are:

$$P_{N+M}(0) = 1. \quad (7)$$

$$P_j(0) = 0,\ 0 \leqq j < N + M. \quad (8)$$

Solutions of these equations are readily obtained for $N \leqq j \leqq N + M$. The solutions are:

$$P_j(t) = \frac{(N\lambda t)^{N+M-j}e^{-N\lambda t}}{(N + M - j)!},\ N \leqq j \leqq N + M. \quad (9)$$

Solutions for $0 \leqq j < N$ are considerably more difficult and are derived in the Appendix.

The exact solutions are rather cumbersome, provide little physical insight into the benefits of standby redundancy, and require high precision arithmetic to obtain numerical results. Approximate formulas are therefore derived in the Appendix, valid for the region of interest $(N - j)\lambda t/(M + 3) \ll 1$. Letting Q_j be the probability of exactly j surviving operational transponders, we obtain:

$$Q_N = \sum_{j=N}^{N+M} P_j \cong 1 - \left[\frac{M + 2}{M + 2 - N\lambda t}\right]\frac{(N\lambda t)^{M+1}}{(M + 1)!}e^{-N\lambda t}. \quad (10)$$

$$Q_{N-1} = P_{N-1}(t) \cong \left[\frac{M^2 + 5M + 6 + \lambda t}{(M + 2)(M + 3 - \lambda t)}\right]\left[\frac{(N\lambda t)^{M+1}e^{-N\lambda t}}{(M + 1)!}\right]. \quad (11)$$

$$Q_j = P_j(t) = \frac{N!N^M(M + 3)^2(\lambda t)^{M+N-j}e^{-N\lambda t}}{j!(N - j - 1)!(M + 2)!(M + 3 - \lambda t)^{N-j}},\ 0 \leqq j \leqq N - 2. \quad (12)$$

III. SATELLITE TRANSPONDER WEIGHT AND RELIABILITY

We will now study the reliability vs weight trade-off for a particular satellite concept, appropriate for a Thor-Delta launch weight class. The following assumptions are made:

(*i*) The satellite can support the weight of 11 transponders without any traveling wave tube redundancy. (Normalized weight = 11).

(*ii*) The weight of a traveling wave tube amplifier (including power supply) is equal to one-third the weight of a single transponder (including the weight of prime power sources for that transponder).

(*iii*) Each switch point needed for redundancy has a weight equal to 1/15th the weight of a traveling wave tube amplifier.

These assumptions are consistent with weight and power budget requirements of a system described elsewhere.[12]

From these requirements, we see that the following redundancy arrangements are possible:

(A) 10 active transponders
2 standby TWTAs
Complete redundant interconnectivity
(Normalized weight = 10.9)

(B) 9 active transponders
4 standby TWTAs
Complete redundant interconnectivity
(Normalized weight = 10.7)

(C) 8 active transponders
6 standby TWTAs
Complete redundant interconnectivity
(Normalized weight = 10.5)

(D) 9 active transponders
6 standby TWTAs
Redundancy provided in 2 on 3 arrangement
(Normalized weight = 11.2)

Numerical results have been obtained for TWTA failure rates between 1500 and 6000 fits (one fit is a failure rate of 1 per 10^9 hours). These extremes are based upon (*i*) actual space experience and life test data for space qualified TWTAs, (*ii*) changes in TWT electrical design objectives such as power rating, operating frequency, and multi-mode capability, and (*iii*) projections for future technology improvements.

Shown in Fig. 2 are the reliability predictions vs time for an 11-transponder satellite with no redundancy. Results are plotted for a failure

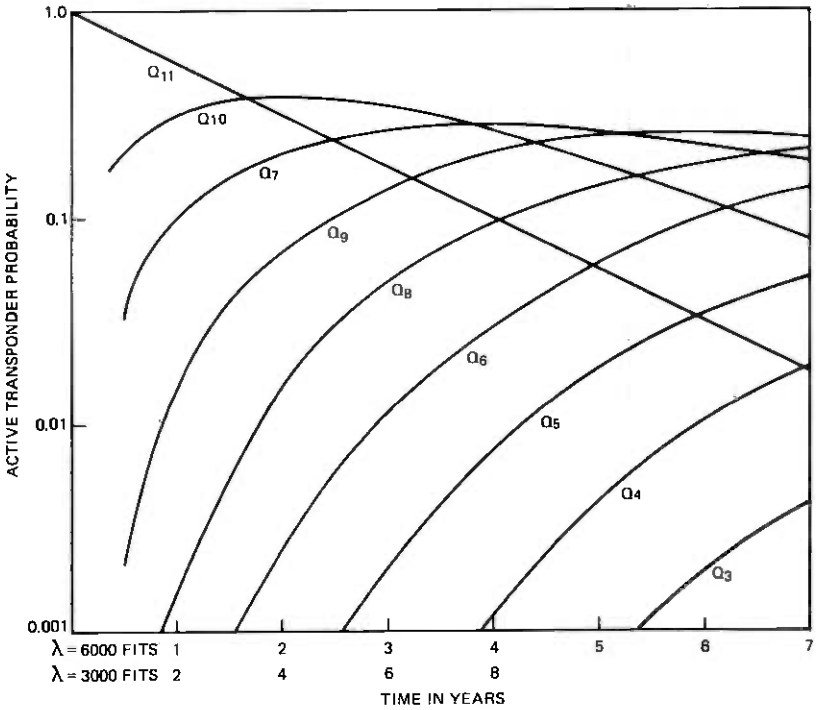


Fig. 2—Active transponder probabilities for a satellite with 11 active transponders and no spare TWTAs. Q_J is the probability that exactly J transponders are operational at time t, $0 \leqq J \leqq 11$.

rate of 6000 fits. Since the reliability predictions are dependent upon the product, λt, predictions can easily be inferred for other failure rates by scaling the time axis by the ratio of $6000/\lambda$, λ being the failure rate under consideration. Such a scaling has been performed in Fig. 2 for $\lambda = 3000$ fits.

Results for Cases A–D are shown, respectively, in Figs. 3 through 6. Since Case D employs less than full redundant interconnectivity, we solve equations (4) through (8) for $N = 3$, $M = 2$, and use these to calculate the desired probabilities:

$$Q_9 = P\{3 \text{ active transponders in each of 3 groups}\} = (Q_3')^3. \tag{13}$$

$$Q_8 = P\{3 \text{ active transponders in each of 2 groups and 2 active transponders in the third group}\} = 3(Q_3')^2 Q_2'. \tag{14}$$

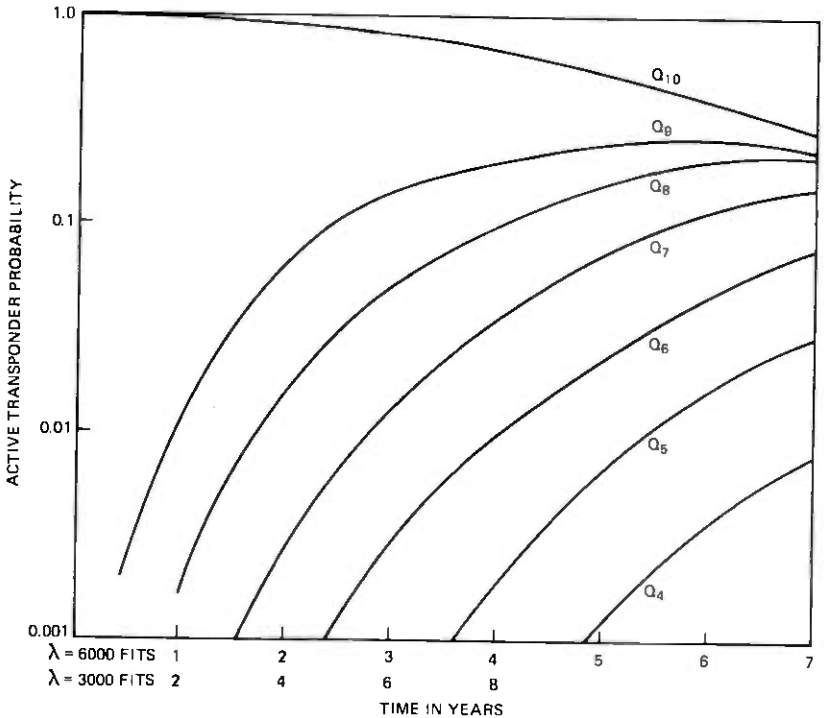


Fig. 3—Active transponder probabilities for a satellite with 10 active transponders and 2 spare TWTAs. Each transponder can access any spare amplifier. Q_J is the probability that exactly J transponders are operational at time t, $0 \leq J \leq 10$.

$$Q_7 = P\{3 \text{ active transponders in each of 2 groups and 1 active transponder in the third, or 2 active transponders in each of 2 groups and 3 active transponders in the third}\}$$
$$= 3(Q_3')^2 Q_1' + 3(Q_2')^2 Q_3', \text{ etc.} \quad (15)$$

Figures (2) through (6) represent a mesh of the exact solutions of the Appendix and the approximate solutions (10) through (12); that is, for large t, the exact expressions have been plotted, while, for small t, the approximate solutions are used since they provide better accuracy than possible via numerical evaluation of the exact solutions. In all cases, agreement was better than 5 percent at the cross-over point between the exact and approximate solutions.

We now interpret the results appearing in Fig. (2) through (6). A seven-year satellite life is assumed. We see from Fig. 2 that service is quite unreliable in the absence of redundancy. For a TWTA failure rate of 6000 fits, the probability of providing service to all coverage areas is less than 2 percent at the end of satellite life; this figure rises to about

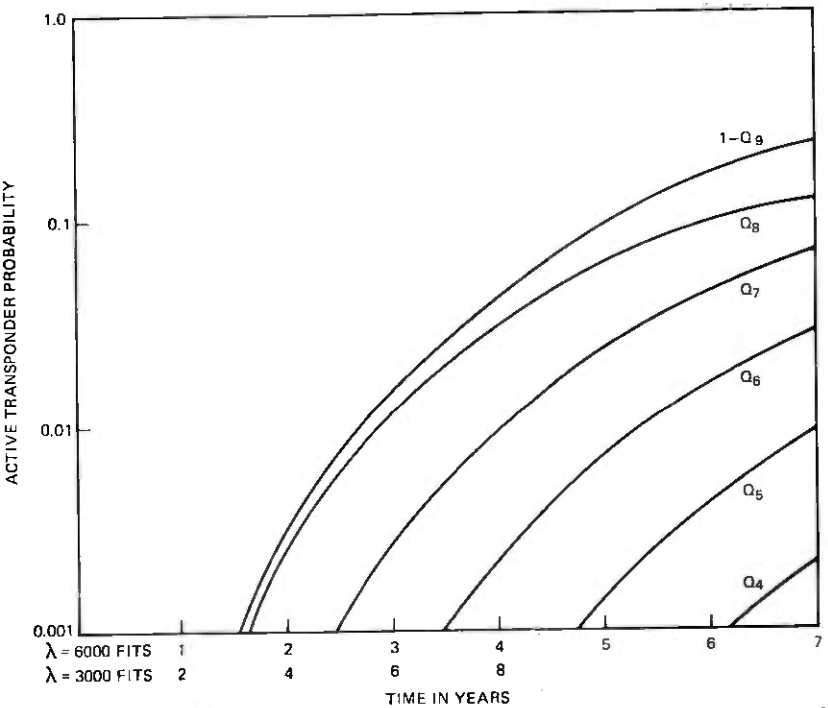


Fig. 4—Active transponder probabilities for a satellite with 9 active transponders and 4 spare TWTAs. Each transponder can access any spare amplifier. Q_J is the probability that exactly J transponders are operational at time t, $0 \leqq J \leqq 9$.

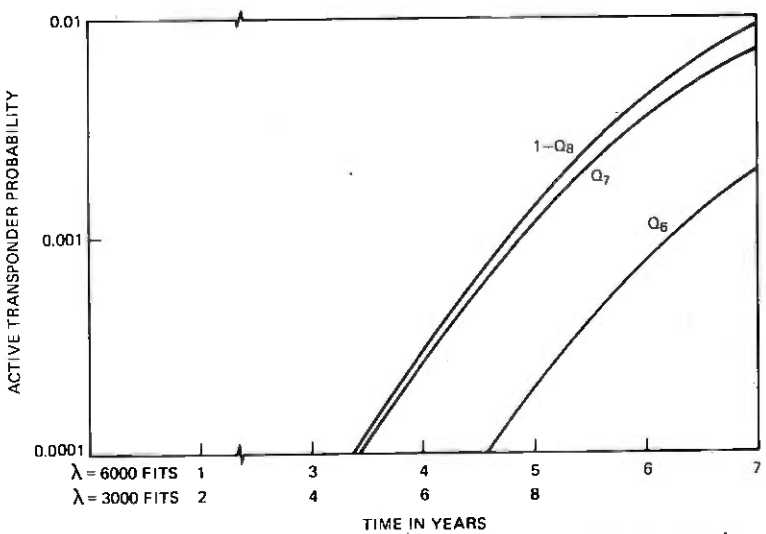


Fig. 5—Active transponder probabilities for a satellite with 8 active transponders and 6 spare TWTAs. Each transponder can access any spare amplifier. Q_J is the probability that exactly J transponders are operational at time t, $0 \leqq J \leqq 8$.

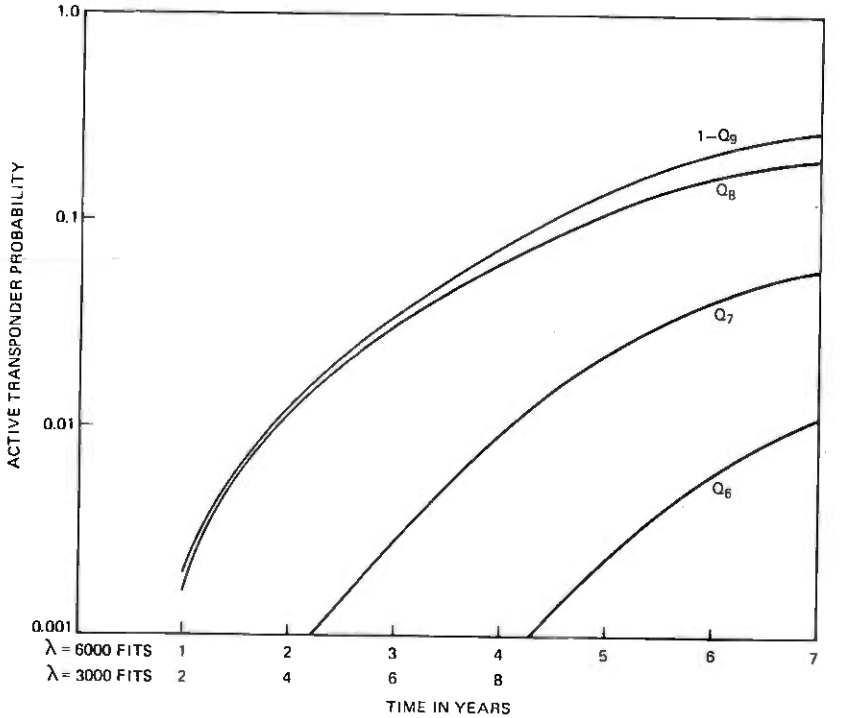


Fig. 6—Active transponder probabilities for a satellite with 9 active transponders and 6 spare TWTAs. Each transponder within a group of 3 can access only the two spare amplifiers assigned to that group. Q_J is the probability that exactly J transponders are operational at time t, $0 \leq J \leq 9$.

13 percent for a failure rate of 3000 fits and to 36 percent for $\lambda = 1500$ fits. In fact, the probability of providing service to seven or fewer of the original 11 overage areas is 41 percent for $\lambda = 3000$ fits. Performance is unacceptable for a satellite system employing multiple spot beams, and this case will not receive any further attention.

The expected number of surviving transponders for Cases A through D are plotted in Fig. 7 as functions of time. From these curves we conclude that Case A (10 active transponders, 2 spares, full interconnectivity) provides the maximum number of expected transponder years. From Fig. 3, we see that for this case, the probability of losing at least one transponder by satellite end-of-life is 72 percent for $\lambda = 6000$ fits, 26 percent for $\lambda = 3000$ fits, and 6 percent for $\lambda = 1500$ fits. However, the probability of having seven or more active transponders at satellite end-of-life is 88 percent for $\lambda = 6000$ fits, 99 percent for $\lambda = 3000$ fits, and 99.99 percent for $\lambda = 1500$ fits. Thus, such an arrangement is probably unacceptable for multiple spot-beam satellites, but is a prime candidate for a global coverage satellite since it provides maximum ex-

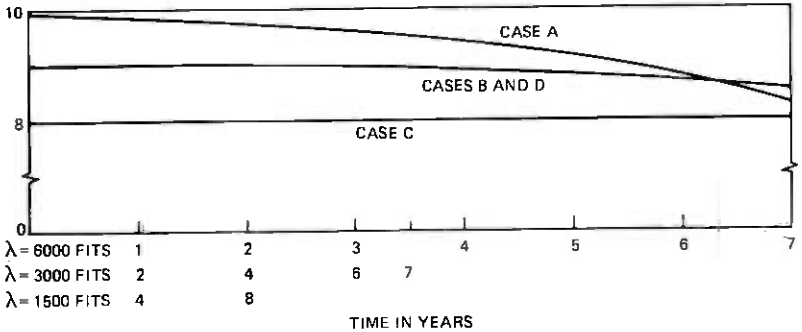


Fig. 7—Expected number of active transponders vs time. Cases A, B, C, and D are as defined in the text.

pected transponder years with high probability of providing at least 70-percent throughput at satellite end-of-life.

For a multiple spot-beam satellite, we see that Case C provides the most dependable service, but that the expected number of transponder years is lower than that for Cases B or D. For Case B, we see that for $\lambda = 6000$ fits, the probability of at least one transponder failing has risen to 24 percent by end of satellite life, and that the probability that no more than one transponder fails is only 88 percent. However, for $\lambda = 3000$ fits, Case B provides for complete service with a 97-percent probability at end of satellite life; the probability that no more than one transponder fails is 99 percent. For $\lambda = 1500$ fits, the corresponding probabilities are 99.9 percent and 99.99 percent. Thus, for a TWTA failure rate of 1500 fits, Case B appears superior to Case C in that an additional transponder is provided with high probability over the life of the satellite. Conversely, for a high TWTA failure rate (6000 fits), Case C is to be preferred since it provides for more dependable service toward the satellite end-of-life.

It is noted further that for a multiple spot-beam satellite, Case B is marginally preferable to Case D if the probability of providing complete service over the life of the satellite is to be maximized (97 percent vs 95 percent for $\lambda = 3000$ fits). Case D, however, provides for a smaller probability that more than one transponder fails (0.5 percent vs 1 percent). Since Case B requires twice the number of switch points to provide for complete interconnectivity, its marginal advantage might disappear if the switch reliability were to be included in the analysis. Conversely, Case D requires slightly more weight than Case B. Thus, for the particular parameters selected for this study, Case B and D are virtually indistinguishable; both provide for highly reliable service over the life of the satellite if the TWTA failure rate is sufficiently small.

We conclude this section with a plot of satellite capacity vs end-of-life reliability for each case discussed (satellite weight constraint). Fig. 8a

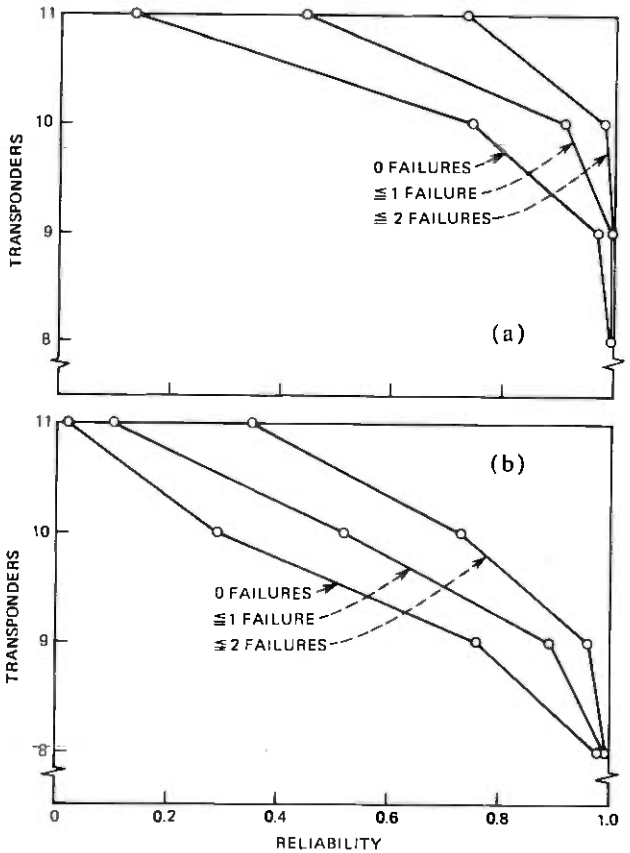


Fig. 8—Probability of success after 7 years vs the original number of active transponders. Success is defined in one of three ways: no failures, no more than one failure, or no more than two failures. (a) $\lambda = 3000$ fits. (b) $\lambda = 6000$ fits.

is applicable for $\lambda = 3000$ fits; Fig. 8b for $\lambda = 6000$ fits. For these plots, reliability is defined as having no more than 0, 1, or 2 failed transponders.

IV. APPLICATION TO MULTI-SATELLITE SYSTEMS

We now apply the results of Sections II and III to determine bounds on the success probability, P_s, of a system containing S identical active satellites. Success is defined as having all design transponders ($S \times N$ total) available at end-of-life. All satellites are assumed to be launched simultaneously.

Obviously, at any point in time, $P_s(t) = Q_N^S(t)$, where $Q_N(t)$ is the probability of all transponders being active on any given satellite at time t. Results for the cases studied in Section III appear in Table I.

Table I — Satellite system success probability; unutilized in-orbit spare

S	Case A	Case B	Case C	Case D
		(a) λ = 6000 fits		
2	0.08	0.57	0.98	0.52
3	0.02	0.44	0.97	0.37
4	0.01	0.33	0.96	0.27
		(b) λ = 3000 fits		
2	0.54	0.94	0.999	0.90
3	0.41	0.91	0.999	0.86
4	0.30	0.89	0.998	0.81
		(c) λ = 1500 fits		
2	0.82	0.996	0.9999	0.984
3	0.74	0.994	0.9999	0.976
4	0.67	0.992	0.9999	0.968

From Table I, we conclude that, with the exception of Case C, the success probability is unacceptably low at the higher fit rates. For λ = 1500 fits, Case B is relatively attractive.

Since, by any measure, the cost associated with total failure of one satellite of a satellite system is prohibitive, satellite systems generally contain an in-orbit spare. In the following, we explore the possibility of using this spare to improve reliability of systems containing S active satellites plus an identical spare.

Suppose we use the spare satellite by assigning to it all the traffic of the first active satellite to experience a transponder failure. Upon such a failure, each ground station permanently reroutes all communications with the failed satellite to the spare. From eq. (9), the probability $r(t)dt$ that any one of the active satellites loses its first transponder in $[t, t + dt]$, given that the first loss occurs somewhere within $[0, T]$, is:

$$r(t)dt = \frac{N\lambda \dfrac{(N\lambda t)^M e^{-N\lambda t}}{M!} dt}{\displaystyle\int_0^T N\lambda \frac{(N\lambda t)^M e^{-N\lambda t}}{M!} dt} = \frac{N\lambda P_N(t)dt}{N\lambda \displaystyle\int_0^T P_N(t)dt}. \tag{16}$$

The probability u that the spare satellite then successfully completes the mission of the failed active satellite is given by:

$$u = \int_0^T r(t)Q_N(T - t)dt, \tag{17}$$

where Q_N is given by eq. (10). Since, for the mission to be successful, the remaining active satellites must be operational at end-of-life, and since the spare, if needed, can be assigned to any of S satellites, we obtain

Table II — Satellite system success probability. Spare satellite replaces first active to fail

No. Active Satellites	Case A λ = 1500 fits	Case A λ = 3000 fits	Case B λ = 3000 fits	Case B λ = 6000 fits	Case C λ = 6000 fits
1	0.975	0.961	0.997	0.970	0.999
2	0.938	0.872	0.993	0.883	0.996
3	0.903	0.762	0.988	0.773	0.993
4	0.866	0.621	0.983	0.659	0.989

$$P_S = Q_N^S + SN\lambda u Q_N^{S-1} \int_0^T P_N(t)dt. \tag{18}$$

The evaluation of eq. (17) is straightforward and is omitted. Evaluation of (18) for Cases A, B, and C of Section II yields the results appearing in Table II.

By comparing these results against those of Table I, we see that the availability of the spare satellite has greatly enhanced the probability of mission success. However, for Case A, we again conclude that the success probabilities are too low for systems employing two or more active satellites. In what follows, we explore the advantages of deploying the spare satellite in a different manner.

Suppose the spare satellite is used by assigning individual transponders on the spare satellite, as needed, to replace failed transponders on the active satellites. Each ground station, then, must be capable of dynamically routing that portion of its traffic destined for the coverage areas of the failed transponders to the spare satellite. Since, for this study, the spare satellite is identical to the active satellites, we conclude that the mission is successful if (*i*) the total number of failed transponders on all active satellites is $\leq N$, (*ii*) no two transponders servicing the same footprint fail on the active satellites, and (*iii*) the assigned transponders on the spare satellite survive after activation for the remainder of the mission. From this, we derive upper and lower bounds on the probability of success as follows.

Let p_i be the probability that i transponders have failed on an active satellite; i.e., $p_i = Q_{N-i}$. Then, the number of ways, W, that i_1 out of N transponders fail on active satellite 1, i_2 out of N transponders (covering footprints different from the i_1 failed transponders of satellite 1) fail on satellite 2, . . . , and i_S out of N transponders (covering footprints different from the i_1 failed transponders of satellite 1 and the i_2 failed transponders of satellite 2, etc.) fail on satellite S is given by:

$$W_{i_1,i_2,\ldots,i_S} = \binom{N}{i_1} \times \binom{N-i_1}{i_2} \times \ldots \times \binom{N-i_1-i_2-\ldots-i_{S-1}}{i_S}$$

$$= \frac{N!}{i_1!i_2!\ldots i_S!(N-i_1-i_2-\ldots-i_S)!}, \tag{19}$$

where $\binom{k}{j}$ represents the binomial coefficient $k!/j!(k-j)!$ The total number of outcomes T in which i_j transponders fail on the jth satellite, $1 \leqq j \leqq S$, is given by

$$T_{i_1,i_2,\ldots i_S} = \binom{N}{i_1} \times \binom{N}{i_2} \times \ldots \times \binom{N}{i_S}. \tag{20}$$

The probability of mission success is given by

$$P_s = \sum_{i_1,\ldots,i_S} P(s|i_1,\ldots,i_S) p_{i_1} p_{i_2} \ldots p_{i_S} \tag{21}$$

Obviously, $P(s|i_1, i_2, \ldots, i_S) = 0, i_1 + i_2 + \ldots + i_S > N,$
for then the spare satellite has too few transponders to provide complete coverage.

Let us define B_k as the probability that none of k specific transponders on the spare satellite have failed at mission end-of-life, given that all spare TWTAs on the spare satellite are available to these k transponders, and that the remaining $N-k$ transponders on the spare satellite are unactivated. The k active transponders are activated at the beginning of the mission; the probabilities B_k are obtained by solving eqs. (4) through (8) for $N = k$.

Now, for a given set $\{i_j | 1 \leqq j \leqq S\}$, all outcomes of failures are equally likely. Thus, assuming that the transponders on the spare satellite needed to take over for failed transponders are available and do not fail, we see that

$$P(S|i_1,\ldots,i_S) \leqq \frac{W_{i_1,\ldots,i_S}}{T_{i_1,\ldots,i_S}}. \tag{22}$$

Similarly, by assuming that $i_1 + i_2 + \ldots i_S$ transponders are initially activated on the spare satellite and that spare TWTAs are available only to these transponders, we see that

$$P(S|i_1,\ldots,i_S) \geqq \frac{W_{i_1,\ldots,i_S}}{T_{i_1,\ldots,i_S}} B_{i_1+i_2+\ldots+i_S}. \tag{23}$$

Thus,

$$\sum_{\substack{i_1,\ldots,i_S \\ i_1+i_2+\ldots+i_S \leqq N}} p_{i_1} p_{i_2} \ldots p_{i_S} \frac{(N-i_1)!(N-i_2)!\ldots(N-i_S)!}{(N!)^{S-1}(N-i_1-i_2-\ldots-i_S)!} B_{i_1+i_2+\ldots+i_S}$$

$$\leqq P_s \leqq \sum_{\substack{i_1,\ldots,i_S \\ i_1+i_2+\ldots+i_S \leqq N}} p_{i_1} p_{i_2} \ldots p_{i_S} \frac{(N-i_1)!(N-i_2)!\ldots(N-i_S)!}{(N!)^{S-1}(N-i_1-i_2\ldots-i_S)!}. \tag{24}$$

We apply these bounds to Case A for $\lambda = 3000$ fits and to Case B for $\lambda = 6000$ fits. Results appear in Table III. We note that dynamic tran-

Table III — Satellite system success probability. Spare satellite transponders assigned dynamically without priority

No. Active Satellites	Case A	Case B
1	$0.998 \leqq P_s \leqq 1.0$	$0.999 \leqq P_s \leqq 1.0$
3	$0.946 \leqq P_s \leqq 0.951$	$0.951 \leqq P_s \leqq 0.953$

sponder assignments on the spare satellite have raised the success probability for Cases A and B appreciably.

Another redundancy strategy, applicable to improving the reliability of systems employing two or more active satellites, also evolves about dynamic assignment of the transponders of the spare satellite as required. For this scheme, however, all available spare TWTAs on the active satellite are allocated on a priority basis such as to maximize the probability of success. For example, if there are two active satellites and one has a single failed transponder whose traffic load has been assumed by the spare satellite, and the transponder covering the same footprint should fail on the second active satellite, then any available spare TWTAs on the second active satellite would be assigned to that failed transponder; if such a spare had previously been brought on line to replace a different failed transponder, the spare satellite would then be employed to pick up the traffic of this second transponder. Again, each ground station must dynamically route that portion of its traffic destined for the coverage areas of the failed transponders to the spare satellite.

The analysis of Section II is not directly applicable to this scheme. This analysis, however, can be applied to a slightly modified scheme whereby not only the spare TWTAs, but also those originally assigned to active transponders, are assigned, as needed, on a priority basis to maximize the probability of mission success. Then, the mission is successful, provided that (*i*) no more than N transponders have failed on the active satellites, and (*ii*) no needed transponders on the spare have failed. The probability of success, therefore, is upper-bounded by:

$$P_s \leqq \sum_{\substack{i_1, i_2, \ldots, i_S \\ i_1 + \ldots + i_S \leqq N}} p_{i_1} p_{i_2} \cdots p_{i_S} \tag{25}$$

Assuming, as before, that $i_1 + i_2 + \ldots + i_S$ transponders needed on the spare satellite begin to fail at time $t = 0$ and that all TWTAs on the spare satellite are available to these transponders, we obtain a lower bound as:

$$P_s \geqq \sum_{\substack{i_1, \ldots, i_S \\ i_1 + i_2 + \ldots + i_S \leqq N}} p_{i_1} p_{i_2} \cdots p_{i_S} B_{i_1 + i_2 + \ldots + i_S} \tag{26}$$

We apply these results to a system containing three active satellites plus a spare. For Case A, with a failure rate of 3000 fits, the success probability exceeds 0.992, while for Case B, with a failure rate of 6000 fits, the success probability exceeds 0.995. For both these examples, the upper bound given by eq. 25 is within 0.02 percent of unity.

V. CONCLUSION

The concept of satellite-switched TDMA was briefly reviewed, and the need for highly reliable transponders was discussed. Since, in a satellite-switched TDMA system, each coverage area is serviced by a single transponder, failure of that transponder eliminates service to that coverage area.

Reliability formulas were derived for the probability of having j out of N active transponders operational, $0 \leqq j \leqq N$, as a function of time for a system employing M-on-N cold standby redundancy. From these, approximate formulas, valid for highly reliable systems, were found. These results were then applied to study the reliability vs the number of transponder trade-off under a total transponder weight constraint. A specific concept, allowing 11 transponders with no redundancy, was studied. It was assumed that each transponder was equivalent in weight to two TWTAs plus interconnecting redundancy switches. Failure rate for the TWTA was varied between 1500 and 6000 fits.

Results indicated that for $\lambda = 6000$ fits, 6-on-8 redundancy with complete interconnectivity is required to achieve a success probability greater than 99 percent after seven years. Thus, for this case, reliability requirements result in a capacity reduction of 27 percent. For a lower failure rate of 3000 fits, 4-on-9 redundancy with complete interconnectivity would provide for 97.3-percent reliability; the capacity reduction is then 18 percent. For $\lambda = 1500$ fits, 4-on-9 redundancy with complete interconnectivity would provide 99.8-percent reliability.

Various schemes employing a spare satellite to increase the reliability of a system employing S active satellites were proposed and studied. Such a spare satellite is generally provided to protect against catastrophic failure; the probability of catastrophic failure was assumed to be zero. We saw that for a failure rate of 1500 fits, a reliability of 97.5 percent could be achieved for a single active satellite system employing 2-on-10 complete interconnectivity redundancy; the spare was utilized by assuming the entire burden of the active satellite when the first transponder failure occurred. Similarly, a 99.7-percent reliability could be achieved for a 4-on-9 redundancy and a failure rate of 3000 fits. For the former case, reliability falls to under 95 percent when two or more active satellites are contained in the system. By contrast, a 6-on-8 strategy employing complete interconnectivity provides for reliability exceeding 99 percent for $\lambda = 6000$ fits and 3 active satellites.

To improve reliability of the 2-on-10 and 4-on-9 concepts at higher TWT failure rates, a different utilization of the spare satellite was explored. This utilization consisted of dynamically assigning to the spare satellite only the traffic of failed transponders, rather than the entire traffic load of the first active satellite to lose a transponder. Such a scheme would require that each ground terminal be capable of communicating, simultaneously, with all active satellites plus the spare. Then, for a TWTA failure rate of 3000 fits, a system employing a single active satellite with 2-on-10 redundancy plus an identical spare satellite would have a success probability exceeding 99.8 percent. Thus, high reliability is achieved at the expense of 9-percent traffic-handling reduction of the active satellite potential. Under the same conditions, a system containing three active satellites plus a spare would have a success probability between 94.6 percent and 95.1 percent.

To increase the reliability of multiple satellite systems still further, a concept again utilizing dynamic transponder allocation to the spare satellite was proposed. For this concept, however, the TWTAs of the active satellites were assigned among the active transponders on a priority basis to prevent failure when possible. It was then found that a system employing three active satellites plus a spare, each of 2-on-10 redundancy, would have a success probability exceeding 99.2 percent for a TWTA failure rate of 3000 fits; a 3-active plus spare satellite system, each of 4-on-9 redundancy, would have a success probability exceeding 99.5 percent for a TWTA failure rate of 6000 fits.

Thus, we conclude that for the type of satellite-switched TDMA system studied, we must accept a 27-percent reduction of capacity to achieve a 99-percent probability of losing no transponder if we do not utilize the presence of a spare satellite. If the spare satellite is properly employed, and if the probability of catastrophic failure is vanishingly small, we achieve even higher success probabilities at the expense of 9 percent in traffic-handling capability. Since a greater volume of traffic can then be handled, and since the total traffic demand upon the satellite system grows with the number of service areas which can be interconnected at a rate greater than the traffic demand of individual additional areas, we conclude that utilization of the spare satellite, as proposed above, is highly desirable.

APPENDIX

In this appendix, we obtain exact solutions to eqs. (4) through (6), subject to initial conditions (7) and (8). Approximate solutions are also derived. For $N \leqq j \leqq N + M$, the exact solutions are:

$$P_j(t) = \frac{(N\lambda t)^{N+M-j} e^{-N\lambda t}}{(N + M - j)!}, \quad N \leqq j \leqq N + M. \tag{9}$$

For $0 \leqq j < N$, we solve via Laplace transform techniques. From eq. (6),

$$P_j(s) = \frac{(j+1)\lambda P_{j+1}(s)}{s + j\lambda}. \tag{27}$$

Thus,

$$P_{N-1}(s) = \frac{N\lambda P_N(s)}{s + (N-1)\lambda} \tag{28}$$

$$P_{N-2}(s) = \frac{N(N-1)\lambda^2 P_N(s)}{[s + (N-1)\lambda][s + (N-2)\lambda]} \tag{29}$$

$$\vdots$$

$$P_j(s) = \frac{N!\lambda^{N-j}}{j!} \frac{P_N(s)}{\prod_{k=1}^{N-j} [s + (N-k)\lambda]}. \tag{30}$$

From eq. (9),

$$P_N(s) = \frac{(N\lambda)^M}{(S + N\lambda)^{M+1}}. \tag{31}$$

Thus,

$$P_j(s) = \frac{N!(N\lambda)^M\lambda^{N-j}}{j!} \left[\frac{1}{(s + N\lambda)^{M+1}}\right] \times \left[\frac{1}{\prod_{k=1}^{N-j} \{s + (N-k)\lambda\}}\right] \qquad 0 \leqq j < N. \tag{32}$$

The various P_j's are seen to possess a multiple pole of order $M + 1$ at $s = -N\lambda$ and simple poles at $s = -(N-k)\lambda$, $k = 1, \ldots, N-j$. All poles are in the left-half complex s-plane. We find $P_j(t)$ by inverting a partial fraction expansion. The result is:

$$P_j(t) = \frac{N!N^M}{j!} e^{-N\lambda t} \sum_{i=1}^{N-j} \frac{1}{i^{M+1} \prod_{\substack{k=1 \\ k \neq i}}^{N-j} (i-k)} \times \left[e^{i\lambda t} - \sum_{l=0}^{M} \frac{(i\lambda t)^{M-l}}{(M-l)!}\right]. \tag{33}$$

We now obtain approximate formulas for $P_j(t)$, $0 \leqq j \leqq N$. Let

$$I = e^{i\lambda t} - \sum_{l=0}^{M} \frac{(i\lambda t)^{M-l}}{(M-l)!} \tag{34}$$

$$= \sum_{l=M+1}^{\infty} \frac{(i\lambda t)^l}{l!} \tag{35}$$

$$= \frac{(i\lambda t)^{M+1}}{(M+1)!}\left[1 + \frac{i\lambda t}{M+2} + \frac{(i\lambda t)^2}{(M+3)(M+2)} + \dots\right]. \tag{36}$$

For $i\lambda t/(M+3) \ll 1$,

$$I \cong \frac{(i\lambda t)^{M+1}}{(M+1)!}\left[1 + \frac{i\lambda t}{M+2} + \frac{(i\lambda t)^2}{(M+2)(M+3)} + \frac{(i\lambda t)^3}{(M+2)(M+3)^2} + \dots\right] \tag{37}$$

$$\cong \frac{(i\lambda t)^{M+1}}{(M+1)!}\left[\frac{M+3}{M+2}\right]\left[\frac{M+2}{M+3} + \frac{i\lambda t}{M+3} + \frac{(i\lambda t)^2}{(M+3)^2} + \frac{(i\lambda t)^3}{(M+3)^3} + \dots\right] \tag{38}$$

$$\cong \frac{(i\lambda t)^{M+1}}{(M+1)!}\left[\frac{M+3}{M+2}\right]\left[-\frac{1}{M+3} + 1 + \frac{i\lambda t}{M+3} + \frac{(i\lambda t)^2}{(M+3)^2} + \dots\right] \tag{39}$$

$$\cong \frac{(i\lambda t)^{M+1}}{(M+1)!}\left[\frac{M+3}{M+2}\right]\left[\frac{1}{1 - \dfrac{i\lambda t}{M+3}} - \frac{1}{M+3}\right]. \tag{40}$$

Now, it is readily shown that

$$\frac{1}{\prod_{\substack{k=1 \\ k \neq i}}^{N-j} (i-k)} = \frac{(-1)^{N-j-i}}{(N-j-i)!(i-1)!}. \tag{41}$$

Thus,

$$P_j(t) \cong \frac{N!N^M e^{-N\lambda t}}{j!} \sum_{i=1}^{N-j} \frac{(-1)^{N-j-i}}{(N-j-i)!(i-1)!i^{M+1}} \times \left[\frac{(i\lambda t)^{M+1}}{(M+1)!}\right]\left[\frac{M+3}{M+2}\right]\left[\frac{1}{1 - \dfrac{i\lambda t}{M+3}} - \frac{1}{M+3}\right]. \tag{42}$$

For $\lambda t/(M+3) \ll 1$,

$$\frac{1}{1-\dfrac{i\lambda t}{M+3}} \cong \left[\frac{1}{1-\dfrac{\lambda t}{M+3}}\right]^i. \tag{43}$$

Thus,

$$P_j(t) \cong \frac{N!N^M e^{-N\lambda t}(-1)^{N-j}}{j!} \sum_{i=1}^{N-j} (-1)^{i} i \binom{N-j}{i} \times \left[\left(\frac{1}{1-\dfrac{\lambda t}{M+3}}\right)^i - \frac{1}{M+3}\right]. \tag{44}$$

Let

$$S = \sum_{i=1}^{N-j} (-1)^{i} i \binom{N-j}{i} \left(\frac{1}{1-\dfrac{\lambda t}{M+3}}\right)^i. \tag{45}$$

Now, for any β,

$$(1-\beta)^{N-j} = \sum_{i=0}^{N-j} \binom{N-j}{i} (-\beta)^i \tag{46}$$

$$\frac{d}{d\beta}[(1-\beta)^{N-j}] = -(N-j)(1-\beta)^{N-j-1} \tag{47}$$

$$= \sum_{i=0}^{N-j} -i \binom{N-j}{i} (-\beta)^{i-1} \tag{48}$$

$$= \frac{1}{\beta} \sum_{i=0}^{N-j} \binom{N-j}{i} i(-\beta)^i \tag{49}$$

$$= \frac{1}{\beta}\left[S\Big|_{\beta = \dfrac{1}{1-\dfrac{\lambda t}{M+3}}}\right]. \tag{50}$$

Thus,

$$S = -\left[\frac{1}{1-\dfrac{\lambda t}{M+3}}\right](N-j)\left[1-\frac{1}{1-\dfrac{\lambda t}{M+3}}\right]^{N-j-1}. \tag{51}$$

Now,

$$\sum_{i=1}^{N-j} (-1)^i i \binom{N-j}{i} = S\Big|_{\beta=1} \tag{52}$$

$$= \begin{cases} 0, j \leqq N-2 \\ -1, j = N-1 \end{cases} \tag{53}$$

Thus,

$$P_{N-1}(t) \cong \left[\frac{M^2+5M+6+\lambda t}{(M+2)(M+3-\lambda t)}\right]\left[\frac{(N\lambda t)^{M+1}e^{-N\lambda t}}{(M+1)!}\right] \tag{54}$$

$$P_j(t) \cong \frac{N!N^M(M+3)^2(\lambda t)^{M+N-j}e^{-N\lambda t}}{j!(N-j-1)!(M+2)!(M+3-\lambda t)^{N-j}} \, 0 \leqq j \leqq N-2. \tag{55}$$

Finally, we obtain an approximate formula for the probability Q_N that all N transponders are active at time t. We note that

$$Q_N(t) = \sum_{j=N}^{N+M} P_j(t) \tag{56}$$

$$= e^{-N\lambda t} \sum_{j=N}^{N+M} \frac{(N\lambda t)^{N+M-j}}{(N+M-j)!} \tag{57}$$

$$= e^{-N\lambda t}\left[e^{N\lambda t} - \sum_{k=M+1}^{\infty} \frac{(N\lambda t)^k}{k!}\right]. \tag{58}$$

For $N\lambda t/(M+2) \ll 1$,

$$Q_N(t) \cong e^{-N\lambda t}\left\{e^{N\lambda t} - \frac{(N\lambda t)^{M+1}}{(M+1)!} \times \left[1 + \frac{N\lambda t}{M+2} + \frac{(N\lambda t)^2}{(M+2)^2} + \ldots\right]\right\} \tag{59}$$

$$\cong 1 - \frac{(N\lambda t)^{M+1}e^{-N\lambda t}}{(M+1)!}\left[\frac{1}{1-\frac{N\lambda t}{M+2}}\right]. \tag{60}$$

Thus,

$$Q_N(t) \cong 1 - \left[\frac{M+2}{M+2-N\lambda t}\right]\frac{(N\lambda t)^{M+1}}{(M+1)!}e^{-N\lambda t}. \tag{61}$$

REFERENCES

1. L. C. Tillotson, "A Model of a Domestic Satellite Communication System," B.S.T.J., *47*, No. 10, (December 1968) pp. 2111–2137.

2. R. Cooperman and W. G. Schmidt, "Satellite-Switched SDMA and TDMA System for Wideband Multi-Beam Satellite," ICC Conference Record, 1973.
3. R. A. Rapuano and N. Shimasaki, "Synchronization of Earth Stations to Satellite Switched Sequences," A1AA 4th Communications Satellite Systems Conference, Washington, D.C. (1972).
4. C. R. Carter and S. S. Haykin, "Precision Synchronization to a Switching Satellite Using PSK Signals," ICC Conference Record, 1974.
5. J. G. Puente, W. G. Schmidt, and A. M. Werth, "Multiple-Access Techniques for Commercial Satellites," Proc. IEEE, *59*, No. 2 (February 1971), pp 218–229.
6. S. B. Bennett and I. Dostis, "Design of the Intelsat IV Transponder," A1AA 4th Communications Satellite Systems Conference, Washington, D.C. (1972), Paper No. 72-535.
7. W. G. Bouricius, et. al., "Reliability Modeling for Fault-Tolerant Computers," IEEE Trans. Comput., *C-20*, (November, 1971) pp. 1306–1311.
8. D. S. Taylor, "A Reliability and Comparative Analysis of Two Standby System Configurations," IEEE Trans. Rel., *R-22*, (April 1973) pp 13–19.
9. D. S. Taylor, "Unpowered to Powered Failure Rate Ratio: A Key Reliability Parameter," IEEE Trans. Rel., *R-23*, No. 1, (April 1974), pp.
10. J. J. Stiffler, "On the Efficiency of R-on-M Redundancy," IEEE Trans. Rel., *R-23*, No. 1, (April 1974), pp 37–43.
11. B. V. Gnedenko Yu. K. Belyayev, and A. D. Solovyev, *Mathematical Methods of Reliability Theory*, New York: Academic Press, 1969.
12. D. O. Reudink, unpublished work.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 56, No. 4, April 1977
Printed in U.S.A.

The Theory of Uniform Cables—Part I: Calculation of Propagation Parameters

By T. A. LENAHAN

(Manuscript received October 5, 1976)

An algorithm for computing the electromagnetic propagation modes and their associated propagation constants (i.e., the loss and phase shift per length) is rigorously developed for uniform cables. The conductors (including the shield, if present) are assumed to have circular cross sections and to be covered by two layers of dielectric insulation. By means of the algorithm, it is now possible to compute the propagation parameters of uniform cables for all frequencies below the microwave range. The algorithm consists of calculating the eigenvectors and eigenvalues of a certain matrix, which are the set of conductor voltages of the modes and the associated propagation constants, respectively. The matrix in question is computed by methods developed in a companion paper on charge densities.

I. INTRODUCTION

Multipair cables consist of a collection of insulated wires surrounded by other dielectric materials, all of which are usually enclosed in a jacketed, metallic shield for mechanical protection and electrical isolation. The wires in a cable are generally helically twisted in pairs and, in practice, other, often undesired, nonuniformities occur along the cable. Nevertheless, the uniform cable, whose wires are straight and parallel, defined by the property that all its longitudinal cross sections are identical, has successfully modeled some aspects of electrical propagation (such as loss per length) over a pair in a cable.[1] Also, some cables are directly modeled as uniform cables; so there is a need for studying uniform cables. Moreover, further experience using the uniform-cable model may lead to a more exact model for nonuniform multipair cables.

A rigorous analysis of electromagnetic propagation over uniform cables, starting from Maxwell's equations, was performed by Carson[2] for cables having homogeneous dielectric material separating the wires.

Subsequently, Mead[3] applied his method to the shielded balanced pair to get a set of formulas for the primary constants (R, L, C, G).

Later, Kuznetsov[4] developed another procedure for calculating the propagation parameters of uniform cables. He too assumed the dielectric separating the wires was homogeneous; also, he did not consider the presence of a shield. His procedure was abstractly stated, but he did derive a concrete formula for the loss and phase-shift per length in the special case of a single pair in free space (see Ref. 4, Chap. 2). Both Kuznetsov and Carson assumed that the frequency of excitation was high enough that the associated skin depth was less than the radius of the wires.

In this paper, the electrical behavior of uniform cables is analyzed from first principles along the lines developed by Kuznetsov. The analysis pertains to all frequencies of excitation below the microwave range, it pertains to cables with or without a shield, and it provides for two layers of dielectric on the conductors* (including the shield). The final result, making use of results from the accompanying paper,[5] is an algorithm, suitable for use on a digital computer, that computes the various propagation parameters of the cable.

The analysis starts with the concept of a modal solution. For a sinusoidal excitation of frequency $\omega/2\pi$, a modal solution for the electric field has the form,

$$\tilde{E} = \exp(i\omega t - \gamma z)\, E,$$

where the Cartesian coordinate z runs along the cable, γ denotes a propagation constant of the cable (giving the loss and phase-shift per length of the mode), and the vector field E is independent of time (t) and z. Thus, the electric field distribution for a modal solution is the same for all transverse cross sections of the cable, apart from a multiplying factor that gives the loss and phase shift per length for the mode, and similarly for the magnetic field. The general solution is a linear combination of modal solutions.

For propagating solutions, in particular, Kuznetsov has shown that there is a potential function V(i.e., a solution of Laplace's equation) defined in the dielectric-region, independent of z, and constant on the surface of the conductors, such that external to the conductors

$$E = -\nabla V + i_z E_z,$$

where i_z denotes the unit vector in the z-direction and the function E_z denotes the z-component of E. (His argument for this electrostatics approximation is enlarged in the next section to account for the shield

* This permits the modeling of a new type of wire insulation recently introduced into cable manufacture: a layer of expanded insulation surrounded by a "skin" of solid insulation. The insulation on the shield is included for formal completeness.

and inhomogeneities in the dielectric). The various propagation modes of the cable, called here the excitation modes, are specified by the set of values for V on the conductors, which are simply the conductor voltages.

The set of wire voltages at $z = 0$ (the shield is always assumed to be grounded) is represented by the vector

$$\mathbf{v} = (v_1, \cdots, v_M),$$

where M is the number of wires in the cable. Certain particular voltage-vectors correspond to excitation modes of the cable. Usually, there are M such normalized voltage vectors, each corresponding to a distinct propagation constant. But it can happen that two or more linearly independent voltage vectors correspond to the same propagation constant; in this case, the propagation constant corresponds to the subspace spanned by these vectors, and the dimension of the subspace is called the multiplicity of the propagation constant. For either situation the collection of excitation modes and their associated propagation constants are the specific propagation parameters that are sought. The usual primary constants of a transmission line or mode, if desired, can be determined by standard formulas[6] from the propagation constant in conjunction with the capacitance matrix.

Under the idealized conditions that the conductors are perfectly conducting and the dielectric is homogeneous, there is a single propagation constant,

$$\gamma = ik_e = i\omega(\epsilon\mu)^{1/2}, \tag{1}$$

where ϵ and μ denote the permittivity and permeability of the dielectric, respectively. In such a case, the multiplicity of ik_e is M. This is the so-called TEM mode,[6] where the longitudinal components of the electric and magnetic fields are zero. Other modes are evanescent, since the frequency range considered here lies below their cutoff frequencies. Since the dielectric in the cable may be inhomogeneous and the conductors are good but not perfect, the propagation solutions or modes are perturbations of the TEM mode, the result being in general a set of excitation modes having distinct propagation constants.

In the next section, the algorithm for computing the excitation modes and associated propagation constants is derived. The modes correspond to the eigenvectors of an $M \times M$ matrix, the propagation constants are simple functions of the associated eigenvalues, and the matrix itself is related to the capacitance and admittance matrices of the cable. In the third section, the algorithm is applied to the shielded balanced pair and a high-frequency approximation is indicated. Experimental testing of these results is reported in Ref. 7.

II. PROPAGATION MODES FOR UNIFORM CABLES

In this section, an algorithm for computing the propagation modes and their associated propagation constants is derived for a uniform cable. The analysis pertains to cables having M straight and parallel wires of circular cross section, enclosed by a circular metallic shield.

The geometry, materials, and excitation of the structure to be considered are specified as follows:

(*i*) The radius of the mth conductor is denoted R_m for $m = 0, 1, \cdots, M$, where R_0 refers to the inside radius of the shield. The shield has a uniform thickness th.

(*ii*) The conductors have two layers of insulation with thickness $(R_{m1} - R_m)$ and $(R_{m2} - R_{m1})$ for the wires $(m = 1, \cdots, M)$ and $(R_0 - R_{01})$ and $(R_{01} - R_{02})$ for the shield. The permittivities (possibly complex) are ϵ_m and ϵ_{m1} for the first and second layers, respectively, for $m = 0, 1, \cdots, M$. In general, the subscript "0" will refer to the shield.

(*iii*) The materials are nonmagnetic, and the permeability is μ throughout.

(*iv*) The material in the interstitial space outside the conductors and their insulation has permittivity ϵ (possibly complex).

(*v*) The excitation frequency $\omega/2\pi$ satisfies $(\omega/2\pi)R_0(\epsilon\mu)^{1/2} \ll 1$, that is, the wavelength is much greater than the radius of the cable. Since R_0 generally exceeds 0.1 inch, this means that the frequency is below the microwave range.

(*vi*) The conductivity of a conductor is denoted σ_m and $\omega\epsilon \ll \sigma_m$ for $m = 0, 1, \cdots, M$.

(*vii*) When the cross-sectional plane of the cable is viewed as the complex plane, the centers of the conductors are associated with the complex numbers $b_0 = 0, b_1, \cdots, b_M$. A typical cross section is indicated in Fig. 1.

The electrical behavior of a cable is described by the electric and magnetic fields, $\tilde{E}$ and $\tilde{H}$, which satisfy Maxwell's equations. A modal solution, as discussed in Section I, has the form,

$$\tilde{E} = \exp(i\omega t - \gamma z)E$$
$$\tilde{H} = \exp(i\omega t - \gamma z)H, \tag{2}$$

where the vector fields E and H are independent of z and t. In terms of their longitudinal and transverse parts,

$$E = E^{tr} + i_z E_z$$
$$H = H^{tr} + i_z H_z, \tag{3}$$

where the vector fields E^{tr} and H^{tr} are the projections of E and H, respectively, onto the transverse plane. Maxwell's equations imply that

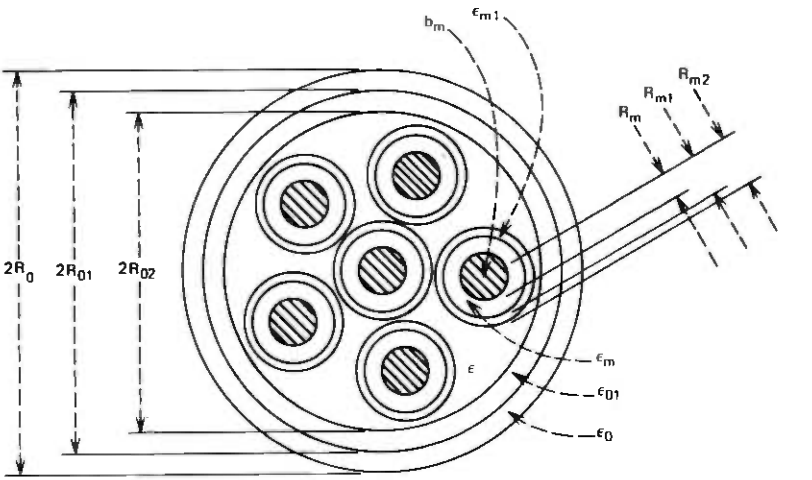


Fig. 1—Typical cable cross section.

(see Ref. 4, page 3)

$$-X^2E^{tr} = \gamma\nabla E_z + i\omega\mu\ \nabla H_z \times i_z$$
$$-X^2H^{tr} = \gamma\nabla H_z + (k^2/i\omega\mu)\ \nabla E_z \times i_z, \tag{4}$$

where taking $\tilde{\epsilon}$ as the permittivity as a function of position,

$$k^2 = \begin{cases} & \text{in } m\text{th conductor} \\ k_m^2 \equiv -i\omega\mu\ \sigma_m & m = 0, 1, \cdots, M \\ k_e^2 \equiv \omega^2\mu\tilde{\epsilon} & \text{outside the conductors,} \end{cases} \tag{5}$$

where the subscript "e" refers to the region exterior to the conductors, and

$$X^2 = k^2 + \gamma^2. \tag{6}$$

The form of the solution in the space separating the conductors is determined by the electrostatics approximation. Kuznetsov[4] has validated this approximation in three steps:

(i) Both E_{ze} and H_{ze} and their derivatives are determined to better than one part in $2.4/k_eR_0$ as solutions of Laplace's equation.

(ii) There is a function V such that $E^{tr} = -\nabla V$ and $\nabla^2 V = 0$.

(iii) On each conductor the tangential part of E^{tr} is zero; so V is constant on the conductors.

For point (i), the original proof must be enlarged to deal with the presence of a shield and inhomogeneities in the dielectric; this is done in Appendix A. Point (ii) is a direct consequence of point (i) because in eq.

(4) both the divergence and curl of $\nabla H_z \times i_z$ are zero, which implies it is the gradient of a solution to Laplace's equation, and point (*iii*) comes from the estimate, $(k_e^2/k_m^2 \ll 1$ for all $m)$ [see Ref. 4, eq. (2.3) and Appendix 1].

The potential which is one on the kth conductor and zero on the others is denoted V_k $k = 1, \cdots, M$. These form a basis set since for any potential V (which is zero on the shield) there are coefficients v_k $k = 1, \cdots, M$ such that

$$V = \sum_{k=1}^{M} v_k V_k. \tag{7}$$

Likewise, when E_{ze} (k) denotes the longitudinal component that corresponds to the potential V_k $k = 1, \cdots, M$, then

$$E_{ze} = \sum_{k=1}^{M} v_k E_{ze}(k) \tag{8}$$

corresponds to the potential V. The coefficients v_k $k = 1, \cdots, M$ form the excitation vector

$$\mathbf{v} = (v_1, \cdots, v_M). \tag{9}$$

To find those excitation vectors that correspond to propagation modes and their associated propagation constants, a set of constraints is developed, one for each conductor, by relating V and E_{ze} at the conductor surfaces. At the outside surface of the mth wire (W_m), eq. (4) implies

$$-X_e^2 E_{\rho me} = \gamma \frac{\partial E_{ze}}{\partial_{\rho m}} + \frac{i\omega\mu}{R_m} \frac{\partial H_{ze}}{\partial \phi_m}, \quad m = 1, \ldots M, \tag{10}$$

where (ρ_m, ϕ_m) denote polar coordinates based at the center of the mth conductor. But $E_{\rho me} = -\partial V/\partial \rho_m$, so

$$X_e^2 \int_0^{2\pi} \frac{\partial V}{\partial \rho_m} d\phi_m = \gamma \int_0^{2\pi} \frac{\partial E_{ze}}{\partial \rho_m} d\phi_m \quad m = 1, \cdots, M, \tag{11}$$

where the derivatives are evaluated on W_m. By eqs. (7) and (8), this is expressed as

$$X_e^2 \sum_{k=1}^{M} v_k \int_0^{2\pi} \frac{\partial V_k}{\partial \rho_m} d\phi_m = \gamma \sum_{k=1}^{M} v_k \int_0^{2\pi} \frac{\partial E_{ze}(k)}{\partial \rho_m} d\phi_m \quad m = 1, \cdots, M. \tag{12}$$

The mk-element of the $M \times M$ capacitance matrix C is

$$C_{mk} = \epsilon_m R_m \int_0^{2\pi} \frac{\partial V_k}{\partial \rho_m} d\phi_m, \tag{13}$$

and by eq. (4) for the transverse magnetic field, the mk-element of the $M \times M$ admittance matrix Y is

$$Y_{mk} = \oint_{W_m} H_{\phi_m}(k)\, ds = \frac{i\omega\epsilon_m R_m}{X_e^2} \int_0^{2\pi} \frac{\partial E_{ze}(k)}{\partial \rho_m}\, d\phi_m. \tag{14}$$

Thus, the set of constraints in eq. (12) is jointly represented by the matrix equation,

$$i\omega\, C\mathbf{v} = \gamma Y\mathbf{v}, \tag{15}$$

where the excitation vector $\mathbf{v}$ is viewed here as a column vector.

But, as will be shown, the matrix Y has the form,

$$Y = \frac{i\omega\gamma}{X_e^2}\, \tilde{Y}, \tag{16}$$

where the matrix $\tilde{Y}$ is independent of γ. Consequently, eq. (15) can be expressed

$$C\mathbf{v} = (\gamma^2/X_e^2)\, \tilde{Y}\mathbf{v} \text{ or } C^{-1}\, \tilde{Y}\mathbf{v} = (X_e^2/\gamma^2)\, \mathbf{v}. \tag{17}$$

This means that to correspond to an excitation mode, the excitation vector $\mathbf{v}$ must be an eigenvector of the matrix $C^{-1}\tilde{Y}$, and conversely. If λ is the associated eigenvalue, then the corresponding propagation constant γ is

$$\gamma = ik_e(1 - \lambda)^{-1/2}. \tag{18}$$

The remaining problem is to calculate the matrices C and $\tilde{Y}$.

The elements of the capacitance matrix C can be computed by the method developed in Ref. 5 (see in particular Section III). This consists of inverting a matrix and is a standard operation on a digital computer. The elements of the $\tilde{Y}$-matrix are evaluated by deriving a boundary-value problem for the function E_{ze} (in the region separating the conductors), then again applying methods from Ref. 5 to solve it.

As the first step in determining $\tilde{Y}$, in Appendix B it is shown that E_{ze} satisfies the following conditions:

(*i*) $\nabla^2 E_{ze} = 0$, where ∇^2 denotes the Laplacian, $(\partial^2/\partial x^2) + (\partial^2/\partial V^2)$.

(*ii*) Across dielectric interfaces E_{ze} is continuous and $\delta(\partial E_{ze}/\partial\rho) = \gamma\delta(\partial V/\partial\rho)$, where $\partial/\partial\rho$ denotes the normal derivative and δ denotes the difference or jump.

(*iii*) $(\partial E_{ze}/\partial\rho_m) - D_\omega E_{ze} = \gamma(\partial V/\partial\rho_m)$ at $\rho_m = R_m$ for $m = 0, 1, \cdots, M$, where D_ω is a linear operator such that for $m = 0, 1, \cdots, M$

$$D_\omega: e^{in\phi_m} \to (X_{cm}/\alpha_{nm})\, e^{in\phi_m} \quad n = 0, \pm 1, \pm 2, \ldots, \tag{19}$$

and the coefficients α_{nm} (which depend on ω) are defined in eq. (58) for $m = 0$ and in eq. (52) for $m \neq 0$, both in Appendix B. In other words, in

condition (*iii*), D_ω alters the Fourier series of E_{ze} on each conductor by multiplying the respective Fourier components.

This boundary problem is simplified by putting

$$E_{ze} = \gamma V + E'_{ze}. \tag{20}$$

Then the corresponding conditions on the function E'_{ze} are

(*i*) $\nabla^2 E'_{ze} = 0$.

(*ii*) Both E'_{ze} and its derivatives are continuous across dielectric interfaces.

(*iii*) $(\partial E'_{ze}/\partial \rho_m) - D_\omega E'_{ze} = \gamma D_\omega V$ at $\rho_m = R_m$ for $m = 0, 1, \cdots, M$.

The function E'_{ze} restricted to a conductor surface is represented by a Fourier series expansion and this by an infinite vector whose components are the Fourier coefficients. If, consistent with the notation in Ref. 5, $\mathbf{u}$ denotes the concatenation of these vectors for each conductor and $\mathbf{p}$ the concatenation of the vectors corresponding to $\epsilon_m(\partial E'_{ze}/\partial \rho_m)$ $m = 0, 1, \cdots, M$, then condition (*iii*) can be expressed by the matrix equation,

$$\mathbf{p} - \tilde{D}_\omega \mathbf{u} = \tilde{D}_\omega \mathbf{V}, \tag{21}$$

where $\tilde{D}_\omega$ denotes the diagonal matrix with elements $\epsilon_m X_m/\alpha_{nm}$ for $0 \leqq m \leqq M$ and $n = 0, 1, 2, \cdots$. Also, there is a vector β and matrices T_0, H_0, and G_0 (defined in Ref. 5) such that

$$\mathbf{u} = T_0\beta \text{ and } \mathbf{p} = G_0\beta - H_0\mathbf{u}, \tag{22}$$

where the subscript "0" emphasizes the context of homogeneous dielectric, in accordance with condition (*ii*) for E'_{ze}. It follows that

$$(G_0 - H_0T_0 - \tilde{D}_\omega T_0)\,\beta = \gamma \tilde{D}_\omega \mathbf{V} \tag{23}$$

or

$$\beta = -\gamma[(H_0 + \tilde{D}_\omega)T_0 - G_0]^{-1}\,\tilde{D}_\omega \mathbf{V}. \tag{24}$$

But it follows from eq. (14) and eq. (16) that for $1 \leqq m, k \leqq M$,

$$\tilde{Y}_{mk} - C_{mk} = \frac{\epsilon_m R_m}{\gamma} \int_0^{2\pi} \frac{\partial E'_{ze}(k)}{\partial \rho_m}\, d\phi_m = 2\pi R_m P_{0k}/\gamma = 2\pi\epsilon\beta_{0k}/\gamma, \tag{25}$$

where the latter equality is shown in Ref. 5, eq. (9). Therefore, the mk element of the $\tilde{Y}$-matrix is computed by carrying out the matrix inversion in eq. (24) for $V = V_m$ to give β_{0k}. Of course, in practice, this inversion is performed on a truncated approximation of the matrix $[(H_0 + \tilde{D}_\omega)T_0 - G_0]$.

III. APPLICATION AND EXTENSIONS

In this section, the general results of the last section are applied to the so-called shielded balanced pair. Also, extensions of the results to cables

without a shield and cables with circular holes in the interstitial dielectric are discussed.

The shielded balanced pair (SBP) consists of a pair of identical wires placed symmetrically inside a circular shield (the center of the shield lies on the line joining the centers of the wires and midway between them). Since the structure is symmetric about the perpendicular bisector of the line joining the wire centers, there will be a balanced mode: the potential and fields will be antisymmetric about this line and, in particular, the wire-voltages will be equal and opposite in sign. The problem is to determine the propagation constant (γ) of the SBP. For convenience, the dielectric will be assumed to be homogeneous with permittivity ϵ.

The excitation vector $\mathbf{v} = (1, -1)$ is an eigenvector for the matrix $C^{-1}\tilde{Y}$ and the corresponding eigenvalue is

$$\lambda = (\tilde{Y}_{11} - \tilde{Y}_{12})/(C_{11} - C_{12}). \tag{26}$$

Equation (18) gives γ in terms of λ. When $V_{12} \equiv V_1 - V_2$,

$$C_{11} - C_{12} = \epsilon \oint_{W_1} \frac{\partial V_{12}}{\partial \rho_1} ds \equiv Q_1, \tag{27}$$

which is the total charge on wire no. 1 (W_1), and

$$\tilde{Y}_{11} - \tilde{Y}_{12} = (\epsilon/\gamma) \oint_{W_1} \frac{\partial E_{ze}}{\partial \rho_1} ds. \tag{28}$$

These two quantities are evaluated as in section II (relative to the potential function V_{12}): Q_1 by methods from Ref. 5, and $\tilde{Y}_{11} - \tilde{Y}_{12}$ by eqs. (24) and (25).

A suggestive formula for γ is obtained by applying Green's identity to eq. (28). Since both V_{12} and E_{ze} are antisymmetric functions, and since $V_{12} = 0$ on the surface of the shield (Sh),

$$\oint_{W_1} \frac{\partial E_{ze}}{\partial \rho_1} ds = \frac{1}{2} \oint_{\Gamma} \frac{\partial E_{ze}}{\partial \eta} V_{12}\, ds, \tag{29}$$

where Γ denotes the entire boundary of the dielectric (W_1, W_2, Sh) and $\partial/\partial\eta$ denotes the normal derivative on Γ (into the dielectric). Since both E_{ze} and V_{12} satisfy Laplace's equation, Green's identity (see Ref. 8, Vol. 2, page 252) implies that the derivative in the integrand can be switched to give

$$\epsilon \oint_{W_1} \frac{\partial E_{ze}}{\partial \rho_1} ds = \frac{\epsilon}{2} \oint_{\Gamma} \frac{\partial V_{12}}{\partial \eta} E_{ze}\, ds = \oint_{W_1} q_1 E_{ze}\, ds + \frac{1}{2} \oint_{Sh} q_0 E_{ze}\, ds, \tag{30}$$

where q_1 and q_0 denote the charge densities on W_1 and Sh, respectively. Thus,

$$\gamma = ik_e \left[1 - (Q_1)^{-1} \oint_{W_1} q_1(E_{ze}/\gamma)ds - (2Q_1)^{-1} \oint_{Sh} q_0(E_{ze}/\gamma)ds \right]^{-1/2}; \quad (31)$$

the Fourier components of $(E_{ze}/\gamma) - V_{12}$ on W, and Sh are $T_0\beta$ with β given by eq. (24). This formula for γ indicates the contribution from each conductor. A comparison between measurement and theory for the SBP based on eq. (31) is presented in Ref. 7.

For high frequencies (i.e., ω for which $\omega\mu\sigma R_1^2 \gg 1$), the formula for E_{ze} can be simplified by neglecting the first term of condition (*iii*) (see Ref. 4, eq 5.8, where X_e should be X_i). This gives

$$E_{ze} = -(\gamma/\epsilon)D_\omega^{-1} q_m \quad (m = 0, 1, 2). \quad (32)$$

This approximation together with the approximation

$$(1 - \delta)^{-1/2} \cong 1 + \tfrac{1}{2}\delta$$

gives

$$\gamma = ik_e \left\{ 1 + (2Q_1)^{-1} \left[\oint_{W_1} q_1(D_\omega^{-1} q_1)ds + \tfrac{1}{2} \oint_{Sh} q_0(D_\omega^{-1} q_0)ds \right] \right\}. \quad (33)$$

The integrals in (31) and (33) can be expressed in terms of the Fourier coefficients of the charge density for convenience in calculating these expressions.

The development in Section II can be extended to cables without a shield and to cables with circular holes in the interstitial dielectric. The calculation of the charge components, and hence, the capacitance and admittance matrices, in both of these situations was treated in Ref. 5, Section III. Again, eq. (17) must be solved for eigenvectors and eigenvalues; but in the absence of a shield, the capacitance matrix is not invertible and the problem is what Kato[9] calls an eigen-problem in the generalized sense. In principle, these can be solved on a digital computer as before.

The results presented here in conjunction with the computation techniques of the companion paper on charge densities allow computation of the excitation modes and their associated propagation constants for uniform cables under a broad range of conditions. These include capacitive and resistive unbalance, caged shields that consist of a collection of unconnected wires, and many decades of excitation frequen-

cies. As the simplest, nontrivial example, the formulation was applied to the shielded balanced pair. As indicated in Ref. 7, the calculations agree reasonably well with measurements.

APPENDIX A

The Electrostatics Approximation

The proof of the electrostatics approximation is broken into three parts in Section III. In this appendix the first part is verified; for the other two parts the reader is referred to Kuznetsov's work.[4] It is shown in particular that within an error for E_{ze} of one part in $(2.4/k_eR_0)$ (typically 250 for 10 MHz),

$$\nabla^2 E_{ze} = 0, \tag{34}$$

and the same argument will hold for H_{ze}.

As indicated in eq. (1.5) of Ref. 5,

$$(\nabla^2 + X_e^2)E_{ze} = 0. \tag{35}$$

If E_{ze}^0 is a function such that $E_{ze}^0 = E_{ze}$ on Γ, E_{ze}^0 is continuous across dielectric interfaces and $\delta\,(\partial E_{ze}^0/\partial\rho) = \delta(\partial E_{ze}/\partial\rho)$ there, and

$$\nabla^2 E_{ze}^0 = 0, \tag{36}$$

then $E_{ze} = E_{ze}^0 + E_{ze}^1$, where $E_{ze}^1 = 0$ on Γ, E_{ze}^1 and its derivatives are continuous across dielectric interfaces, and

$$(\nabla^2 + X_e^2)E_{ze}^1 = -X_e^2 E_{ze}^0. \tag{37}$$

When ψ_1 denotes the first eigenfunction of $-\nabla^2$ that is zero on Γ, then the L^2-norm of E_{ze}^1 is bounded by

$$\|E_{ze}^1\| \leqq \left|\frac{X_e^2}{X_e^2 - \lambda_1^2}\right| \|E_{ze}^0\|, \tag{38}$$

where λ_1 is the eigenvalue associated with ψ_1. The smallest eigenvalue of $-\nabla^2$ in the circle of radius R_0 relative to a boundary condition of zero on the circumference is $2.4048/R_0$ [2.4048 is the first zero of $J_0(x)$] and (see Ref. 8, Vol. 1, page 409, Theorem 3)

$$\lambda_1^2 \geqq (2.4048/R_0)^2. \tag{39}$$

Hence,

$$\|E_{ze}^1\| \leqq (X_eR_0/2.4048)^2 \|E_{ze}^0\|, \tag{40}$$

and since $X_e \sim k_e$,

$$E_{ze} \cong E_{ze}^0, \tag{41}$$

which was to be proved. Clearly, the same argument holds for the x and y derivatives of E_{ze}.

APPENDIX B

Boundary Problem for E_{ze}

From the various conditions on E_z, a boundary problem for E_{ze} can be deduced. The result is stated as three conditions in Section II. In this appendix these conditions are verified.

(*i*) From the electrostatics approximation, $\nabla^2 E_{ze} = 0$.

(*ii*) Since E_z is a tangential component of E, it is continuous across interfaces. The difference in normal-derivative across the dielectric interfaces is deduced from eq. (4). Since H_{ze} is continuous across interfaces,

$$-\delta(X_e^2 E_{\rho me}) = \gamma\delta\left(\frac{\partial E_{ze}}{\partial \rho_m}\right) \quad m = 0, 1, \cdots, M \tag{42}$$

at the interfaces. But by the usual boundary condition for the normal E field, $\delta(k_e^2 E_{\rho e}) = 0$; so for $m = 0, 1, \cdots, M$,

$$\delta\left(\frac{\partial E_{ze}}{\partial \rho_m}\right) = -\gamma\,\delta(E_{\rho e}) = \gamma\delta\left(\frac{\partial V}{\partial \rho_m}\right) \tag{43}$$

at the dielectric interfaces.

(*iii*) By the same reasoning as in (*ii*)

$$\frac{\partial E_{ze}}{\partial \rho_m} - \frac{\partial E_{zm}}{\partial \rho_m} = -\gamma(E_{\rho e} - E_{\rho m}) \text{ at } \rho_m = R_m, \tag{44}$$

where E_{zm} and $E_{\rho m}$ refer to components of E in the mth conductor for $m = 0, 1, \cdots, M$. Since

$$E_{\rho m} = (k_e^2/k_m^2)E_{\rho e} \ll E_{\rho e} \tag{45}$$

for all m, eq. (44) to a good approximation becomes

$$\frac{\partial E_{ze}}{\partial \rho_m} - \frac{\partial E_{zm}}{\partial \rho_m} = \gamma\frac{\partial V}{\partial \rho_m} \quad \text{for } m = 0, 1, \cdots, m. \tag{46}$$

Finally, it is proved that

$$\frac{\partial E_{zm}}{\partial \rho_m} = D_{\omega m} E_{zm} \quad \text{for } m = 0, 1, \cdots, M, \tag{47}$$

and since $E_{zm} = E_{ze}$ at $\rho_m = R_m$, this will verify the third condition. This is proved first for the wires, then for the shield.

Inside a wire,

$$\nabla^2 E_{zm} + X_m^2 E_{zm} = 0 \quad m = 1, \cdots, M, \tag{48}$$

and since $\gamma^2 \ll X_m^2$, the quantity X_m^2 can be replaced by k_m^2. The solution for E_{zm} has the form,

$$E_{zm} = \sum_n a_n \frac{J_n(k_m \rho_m)}{J_n(k_m R_m)} \exp(in\phi), \tag{49}$$

where the a_n are the Fourier coefficients of E_z at $\rho_m = R_m$ and J_n denotes the nth order Bessel function of the first kind. The radial derivative at $\rho_m = R_m$ is

$$\frac{\partial E_{zm}}{\partial \rho_m} = k_m \sum_n a_n \frac{J_n'(k_m R_m)}{J_n(k_m R_m)} \exp(in\phi_m) = k_m D_{\omega m} E_{zm}, \tag{50}$$

where $D_{\omega m}$ is a linear operator, which for $m = 1, \cdots, M$,

$$D_{\omega m}: \exp(in\phi_m) \rightarrow (k_m/\alpha_{nm}) \exp(in\phi_m) \quad n = 0, \pm 1, \pm 2, \cdots, \tag{51}$$

and

$$\alpha_{nm} = \frac{J_n(k_m R_m)}{J_n'(k_m R_m)}. \tag{52}$$

Inside the shield,

$$\nabla^2 E_{z0} + k_0^2 E_{z0} = 0, \tag{53}$$

where X_0^2 has been replaced by k_0^2. The solution has the form,

$$E_{z0} = \sum_n a_n \frac{b_n J_n(k_0 \rho_0) + c_n Y_n(k_0 \rho_0)}{b_n J_n(k_0 R_0) + c_n Y_n(k_0 R_0)} \exp(in\phi_0), \tag{54}$$

where the a_n are the Fourier coefficients of E_z at $\rho_0 = R_0$, the b_n and c_n are determined by reference to the boundary condition on the outside surface of the shield ($\rho_0 = R_0 + th \equiv R_0'$), and Y_n denotes the nth order Bessel function of the second kind. Assuming the cable is not in the vicinity of sources or other conductors, the potential outside will be zero; hence, there is no charge on the outer surface of the shield, and by eq. (4) $\partial E_z/\partial \rho_0$ is continuous at $\rho_0 = R_0'$. Since outside the cable, E_z has the form

$$E_z = \sum_{n \neq 0} f_n \left(\frac{R_0'}{\rho_0}\right)^{|n|} \exp(in\phi_0), \tag{55}$$

for some f_n, it follows from the continuity of $\partial E_z/\partial \rho_0$ that in eq. (54),

$$b_n = Y_{n-1}(k_0 R_0') \quad \text{and} \quad c_n = -J_{n-1}(k_0 R_0') \quad n = 0, \pm 1, \pm 2, \cdots. \tag{56}$$

The radial derivative of E_z at $\rho_0 = R_0$ is

$$\frac{\partial E_{z0}}{\partial \rho_0} = k_0 \sum_n (a_n/\alpha_{n0}) \exp(in\phi_0), \tag{57}$$

where

$$\alpha_{n0} = \frac{Y_{n-1}(k_0R_0')J_n(k_0R_0) - J_{n-1}(k_0R_0')Y_n(k_0R_0)}{Y_{n-1}(k_0R_0')J_n'(k_0R_0) - J_{n-1}(k_0R_0')Y_n'(k_0R_0)} \tag{58}$$

$$n = 0, \pm 1, \pm 2, \cdots.$$

Hence, at $\rho_0 = R_0$

$$\frac{\partial E_{z0}}{\partial \rho_0} = k_0 D_{\omega 0} E_{z0}, \tag{59}$$

where $D_{\omega 0}$ is a linear operator as in eq. (51) with $m = 0$. The direct sum[10] of the $D_{\omega m}$ for $m = 0, 1, \cdots, M$ is D_ω, as given in eq. (19).

REFERENCES

1. R. C. Sacks, unpublished work
2. J. R. Carson, "The Rigorous and Approximate Theories of Electrical Transmission Along Wires," B.S.T.J., 7, No. 1 (January 1928), pp. 11–25.
3. S. P. Mead, "Shielded Pairs of Wires," Patent 2,034,032, March 1936.
4. P. I. Kuznetsov and R. L. Stratonovich, *The Propagation of Electromagnetic Waves in Multiconductor Transmission Lines*, New York: Pergamon Press, 1964, Chap. 1, 2.
5. T. A. Lenahan, "The Theory of Uniform Cables—Part II: Calculation of Charge Components," B.S.T.J., this issue, pp. 611–625.
6. R. B. Adler, L. J. Chu, and R. M. Fano, *Electromagnetic Energy Transmission and Radiation*, New York: John Wiley, 1960, p. 180 and Chap. 9.
7. T. A. Lenahan, "Experimental Test of Propagation-Parameter Calculations for Shielded Balanced Pair Cables," B.S.T.J., this issue, pp. 627–636.
8. R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. 1, 2, New York: Interscience Publishers, 1966.
9. T. Kato, *Perturbation Theory for Linear Operators*, New York: Springer-Verlag, 1966, pp. 416–417.
10. K. Hoffman and R. Kunze, *Linear Algebra*, Englewood Cliffs, N.J.: Prentice-Hall, 1961, p. 159.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 56, No. 4, April 1977
Printed in U.S.A.

The Theory of Uniform Cables—Part II: Calculation of Charge Components

By T. A. LENAHAN

(Manuscript received October 5, 1976)

The analysis of electromagnetic propagation over uniform cables depends on the calculation of the charge densities on the conductors, relative to a potential function that is not necessarily constant on the conductors. By considering such a potential function as the real part of an analytic function, two Laurent series are derived, one of which involves the Fourier components of the potential function and its associated charge densities on the conductors. The second series accounts for the relative location of the conductors. The two series are equated to give a system of linear equations that can be solved for the charge components. The results obtained, which apply to uniform cables whose conductors (including the shield, if present) have circular cross sections and are covered with two layers of dielectric insulation, can be used to calculate the propagation modes and propagation constants of the cable.

I. INTRODUCTION

Multiconductor cables for telecommunications have distributions of charge on each conductor. The surface charge density on the conductors is proportional to the normal derivative of a potential function (i.e., a solution to Laplace's equation), which is defined in the region separating the conductors and is constant on the conductors. The constant values of the potential are the voltages of the conductors, and the proportionality constant is the permittivity of the material next to the conductor. Also, generalized charge densities are defined for potential functions, such as the longitudinal component of the electric field, which are not constant on the conductors.

In the present work an algorithm is developed for computing the charge densities in this generalized sense for uniform cables. Thus, the conductors of the cable are assumed to be straight and parallel, so that each transverse cross section is identical. The wires are assumed to have

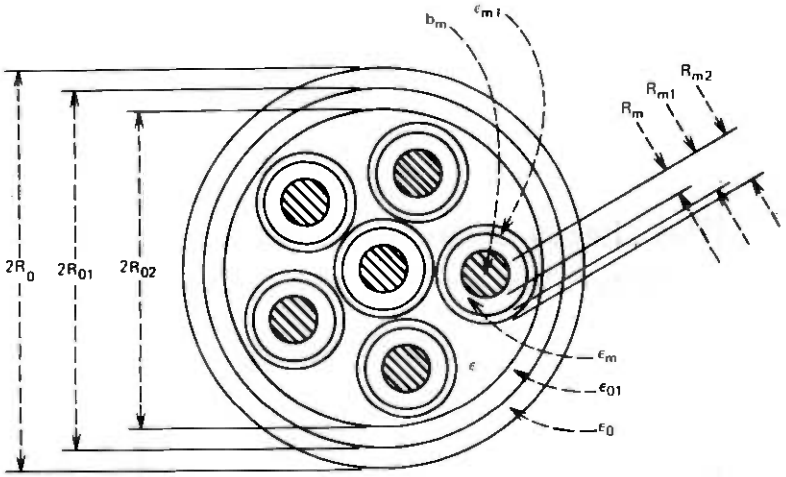


Fig. 1—Typical cable cross section

circular cross sections and to be covered by two circularly symmetric layers of homogeneous dielectric material. Surrounding the collection of wires is a circular, metallic shield, and it too is assumed to have two uniform layers of dielectric on its inside surface. A typical cross section is shown in Fig. 1.

Interest in the charge densities was spurred by the recent finding that the modes of a cable and their associated propagation constants can be expressed directly in terms of the charge densities when, as in Kuznetsov's work,[1] low frequencies are excluded. Subsequently, it was found that the analysis could be extended to low frequencies by using the notion of generalized charge densities associated with the longitudinal component of the electric field. These matters are developed in detail in Ref. 2.

Previous work on calculating charge on the conductors of a cable usually involved the simplifying assumption of homogeneous dielectric in the region separating the conductors. For two identical wires in free space, an explicit formula is available for all the Fourier components of the charge density on a wire (see Ref. 1, page 41). Goluzin,[3] using the theory of complex variables, has developed a continued fraction expansion for a potential function in a region bounded by circles; the expansion converges under certain conditions on the size and location of the circles. Nordgard[4] has developed an algorithm involving a matrix inversion for computing Fourier components of the charge densities for a pair in a shield. Also, the capacitance elements (the zero-order component of a charge density) have been calculated in a variety of circumstances.[5,6]

For inhomogeneous dielectrics, the calculation of charge density has

been approached by a finite-difference technique[7] coupled with a Fourier analysis[8] of the normal derivative to give the Fourier components. This technique, by calculating more information than is needed, has proved to involve more computer expense than the technique to be described here.

In this paper, techniques from the theory of complex variables are used to develop a set of linear equations for the charge components. Following Goluzin (see Ref. 3), the potential function is viewed as the real part of an analytic function; then two Laurent series are calculated for it about each conductor. The first is expressed in terms of the Fourier components of the charge density and the potential function on each conductor; the second is expressed in terms of singularities located at the centers of the conductors (at infinity for the shield). Since the two Laurent series must be identical, their corresponding coefficients can be equated to give a system of linear equations from which the charge components can be calculated. For convenience, the system of equations is expressed in matrix form.

The matrix equations are developed in the next section, with details of the calculations and proofs relegated to appendices. In Section III, the development is summarized and then extended to apply to cables without shields, to cables with holes in the dielectric, and to more general boundary problems that arise in determining the admittance matrix and other propagation parameters of the cable. Numerical examples and experimental testing of this work is presented in Ref. 9.

II. CALCULATION OF THE CHARGE COMPONENTS

The cable to be considered consists of M straight and parallel wires with circular cross sections surrounded by a circular shield of inside radius R_0. The wires and the inside surface of the shield are covered by two different layers of dielectric insulation. For the mth wire ($1 \leq m \leq M$) the radius is R_m; the permittivity of its first layer of dielectric is ϵ_m with thickness $R_{m1} - R_m$; and the permittivity of its second layer is ϵ_{m1} with thickness $R_{m2} - R_{m1}$. On the shield (referred to as the $m = 0$ conductor), the permittivities of the first and second layers are ϵ_0 and ϵ_{01} with thickness $R_0 - R_{01}$ and $R_{01} - R_{02}$, respectively. The permittivity of the material separating the insulated conductors is ϵ.

When the cross-sectional plane of the cable is viewed as the complex plane, the centers of the conductors can be specified as complex numbers. These are denoted b_m ($m = 0, 1, \cdots, M$), where $b_0 = 0$ refers to the center of the shield. A typical cross section is shown in Fig. 1.

With (ρ_m, ϕ_m) as polar coordinates based at the center of the mth conductor ($m = 0, 1, \cdots, M$), a potential function U is assumed to satisfy the following conditions:

(i) In the region separating the conductors, $\nabla^2 U = 0$, where $\nabla^2 =$

$(\partial^2/\partial x^2) + (\partial^2/\partial y^2)$ and (x, y) denote Cartesian coordinates in the plane (i.e., U satisfies Laplace's equation).

(ii) At the surface of the mth conductor ($\rho_m = R_m$), U has the Fourier series:

$$U = \sum_n u_{nm} \exp(in\phi_m) \qquad m = 0, 1, \cdots, M. \tag{1}$$

(iii) When $\tilde{\epsilon}$ denotes the permittivity as a function of position, then both U and $\tilde{\epsilon}(\partial U/\partial \rho_m)$ are continuous across dielectric interfaces about the mth conductor for $m = 0, 1, \cdots, M$. The charge density associated with U at the mth conductor is

$$p_m = \epsilon_m \left.\frac{\partial U}{\partial \rho_m}\right|_{\rho_m = R_m} = \sum_n p_{nm} \exp(in\phi_m) \qquad m = 0, 1, \cdots, M. \tag{2}$$

Thus, the problem is to determine the Fourier components, p_{nm}.

The boundary problem for U is put into the context of complex variables by viewing the cross-sectional plane of the cable as the complex plane. Then, in the region separating the conductors, there is an analytic function $f(z)$ (unique within addition of an imaginary constant) such that U is the real part of $f(z)$. And $f(z)$ can be represented as a Laurent series.[10]

In Appendix A, a Laurent series is calculated for $f(z)$ in a neighborhood of b_m in terms of the Fourier series of U and ϵ_m $(\partial U/\partial \rho_m)$ at the surface of the mth conductor [eq. (1) and eq. (2), respectively] for $m = 0, 1, \cdots, M$. The result, eq. (26) of Appendix A, shows the coefficients of the Laurent series about $z = b_m$ depending only on p_{nm} and u_{nm} ($n = 0, \pm 1, \pm 2, \cdots$) with no explicit indication of interconductor coupling.

A second representation for $f(z)$, which is based on Cauchy's integral formula is

$$f(z) = f_0(z) + \sum_{m=1}^{M} f_m(z), \tag{3}$$

where

$$f_0(z) = \beta_{00} + \sum_{n=1}^{\infty} \beta_{n0} \left(\frac{z}{R_0}\right)^n \tag{4}$$

is analytic everywhere inside the shield and

$$f_m(z) = \beta_{0m} \ell n \left(\frac{z - b_m}{R_m}\right) + \sum_{n=1}^{\infty} \beta_{nm} \left(\frac{z - b_m}{R_m}\right)^{-n} \tag{5}$$

is analytic everywhere outside the mth wire for $m = 1, \cdots, M$ (see Ref. 3). The coefficients $\beta_{00}, \cdots, \beta_{0M}$ are real, and in general the β_{nm} (for $1 \leqq m \leqq M$ and $n > 0$) are complex. In Appendix B, a second Laurent series is calculated for $f(z)$ in a neighborhood of b_m by combining these

functions. The result is eq. (38) for $m = 0$ and eq. (39) for $m = 1, \cdots, M$.

The two Laurent series must be identical since they represent the same function $f(z)$. Therefore, their coefficients can be equated, and, as shown in Appendix C, this leads to systems of linear equations in u_{nm}, p_{nm}, and β_{nm} for $m = 0, 1, \cdots, M$ and $n = 0, 1, \cdots$. It suffices to deal with non-negative values of n because U and p_m are real quantities for all m (hence, $u_{-nm} = u^*_{nm}$ and $p_{-nm} = p^*_{nm}$, where * denotes complex conjugate).

The systems of equations are conveniently expressed in matrix form with the various coefficients combined into vectors. Accordingly, the following infinite vectors are defined:

$$\begin{aligned} \mathbf{u_m} &= (u_{0m}, u_{1m}, \cdots) \\ \mathbf{p_m} &= (p_{0m}, p_{1m}, \cdots) \\ \beta_{\mathbf{m}} &= (\beta_{0m}, \beta_{1m}, \cdots), \end{aligned} \tag{6}$$

all for $m = 0, 1, \cdots, M$ and the joint vectors,

$$\begin{aligned} \mathbf{u} &= (\mathbf{u_0}, \cdots, \mathbf{u_m}) \\ \mathbf{p} &= (\mathbf{p_0}, \cdots, \mathbf{p_m}) \\ \beta &= (\beta_0, \cdots, \beta_{\mathbf{m}}). \end{aligned} \tag{7}$$

As indicated in Appendix C, when these vectors are taken as column vectors, there are matrices T, G, and H such that

$$T\beta = \mathbf{u} \text{ and } \mathbf{p} = G\beta - H\mathbf{u}. \tag{8}$$

In particular, G and H are such that

$$p_{0m} = (\epsilon/R_m)\beta_{0m} \quad m = 1, \cdots, M. \tag{9}$$

If only the zero-order components of the charge-densities are of interest, then it suffices to invert the T-matrix to give

$$\beta = T^{-1}\mathbf{u}, \tag{10}$$

whereupon eq. (9) is used. When all the components of the charge densities are of interest, then combining the equations in (8) gives

$$\mathbf{p} = (GT^{-1} - H)\mathbf{u}. \tag{11}$$

(That the T-matrix is invertible follows from the well-known uniqueness of solutions to Laplace's equation with prescribed values on the boundary.)

In practice, the infinite vectors $\mathbf{u_m}$, $\mathbf{p_m}$, and $\beta_{\mathbf{m}}$ $m = 0, 1, \cdots, M$ are truncated to give N-vectors, and the matrices T, G, and H are truncated to $(M + 1)N \times (M + 1)N$ matrices. The matrix operations indicated in

eq. (11) are then carried out on the truncated matrices to give an approximation to the first N components of the charge density on each conductor.

A detailed study of the matrix truncation has not been carried out, though the cables studied in Ref. 9 provide some experience in this matter. For the cables whose wires and shield were mutually separated by more than one wire diameter (the 754E and Focal), it more than sufficed to consider five harmonics on the conductors, including the zero order (i.e., $N = 5$). When N was set to 6, there was a difference in the coefficient of the dominant harmonic (zero order) of less than 1 in 10,000; furthermore, the coefficient of the extra harmonic was four orders of magnitude less than that of the zero-order harmonic. For the cable whose wires and shield were separated by a small fraction of one wire diameter (the Proximity cable), eight harmonics were required: the coefficient of the seventh-order harmonic was about $1/15$ that of the zero-order harmonic. When N was set to 10, again there was a difference in the dominant coefficients of less than 1 in 10,000; and the extra coefficients were two orders of magnitude less than that of the zero-order harmonic.

A second practical matter is the presence of conjugation operators in the T-matrix. When it is known (e.g., by symmetry) that the coefficients $\beta_{\ell m}$ are real for all ℓ and m, then the conjugation operator has no effect and can be ignored. The case where the conductor centers are collinear is handled in this way.

In general, the coefficients have an imaginary part. Then $\beta_{\ell m}$ is expressed as the two-vector

$$\begin{pmatrix} \text{Re } \beta_{\ell m} \\ \text{Im } \beta_{\ell m} \end{pmatrix}$$

and the element of the T-matrix, $T_{km}(n, \ell)$, which multiplies $\beta_{\ell m}$, is expressed as

$$\begin{pmatrix} \text{Re } T_{km}(n, \ell) & -\text{Im } T_{km}(n, \ell) \\ \text{Im } T_{km}(n, \ell) & \text{Re } T_{km}(n, \ell) \end{pmatrix}$$

The first component of the matrix product is the proper real part and the second is the proper imaginary part of the product, $T_{km}(n, \ell)\beta_{\ell m}$. When a conjugation operator appears, the 2×2 matrix above must be multiplied on the right by

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

(this corresponds to changing the sign of the imaginary part of $\beta_{\ell m}$). The result is the matrix

$$\begin{pmatrix} \text{Re } T_{km}(n, \ell) & \text{Im } T_{km}(n, \ell) \\ \text{Im } T_{km}(n, \ell) & -\text{Re } T_{km}(n, \ell) \end{pmatrix}.$$

When these matrices are substituted for the elements of the T-matrix, then the resulting matrix can be inverted in the usual way.

III. SUMMARY AND EXTENSIONS

Matrix equations have been derived which relate $\mathbf{u}$, $\mathbf{p}$, and the auxiliary vector β. These are

$$T\beta = \mathbf{u} \text{ and } \mathbf{p} = G\beta - H\mathbf{u}.$$

From these equations, the vector $\mathbf{p}$ can be determined when $\mathbf{u}$ is given. When the potential function U is the constant 1 on the mth conductor and zero on the others, the quantity

$$C_{km} = 2\pi R_k p_{0k} = 2\pi\epsilon\beta_{0k} \quad 1 \leqq k, m \leqq M \tag{12}$$

is the km-element of the $M \times M$ capacitance matrix C. Thus, the capacitance matrix is calculated as a special case of the analysis in Section II.

These results can be readily extended to cables without a shield and with some modification to cables with circular holes in the interstitial dielectric. When no shield is present, two changes must be made:

(i) Components of $\mathbf{u}$, $\mathbf{p}$, and β associated with $m = 0$ must be eliminated, and the corresponding submatrices in the matrices T, G, and H must be eliminated.

(ii) The T-matrix must be bordered by one row and one column to give

$$\bar{T} = \begin{pmatrix} 0 & \mathbf{c} \\ \mathbf{c}^t & T \end{pmatrix} \text{ with } \mathbf{c} = (\mathbf{e}_1, \cdots, \mathbf{e}_1), \tag{13}$$

where $\mathbf{e}_1 = (1, 0, 0, \cdots)$ is repeated M-times. Then the first equation in (8) becomes

$$\bar{T}\begin{pmatrix} \xi \\ \beta \end{pmatrix} = \begin{pmatrix} Q \\ \mathbf{u} \end{pmatrix}, \tag{14}$$

where Q is the total charge of the cable and ξ is a constant to be determined. The second condition comes from the requirement that the total charge be specified, but that the potential on the boundary (the surface of the conductors) be specified only within an additive constant (ξ). This holds for any exterior problem for Laplace's equation in two dimensions.[11]

Circular holes in the interstitial dielectric are treated like extra conductors with unspecified potential values at their surface ($\rho_h = R_h$). But if u_{nh} denote the Fourier components of the potential function at the surface of the hole, then inside the hole,

$$U = \sum_n u_{nh} \left(\frac{\rho_h}{R_h}\right)^n \exp(in\phi_h); \tag{15}$$

and if ϵ_h is the permittivity of the material in the hole, then

$$p_h = \epsilon \left.\frac{\partial U}{\partial \rho_h}\right|_{\text{out}} = \epsilon_h \left.\frac{\partial U}{\partial \rho_h}\right|_{\text{in}} = \sum_n (n\epsilon_h u_{nh}/R_h) \exp(in\phi_h), \quad (16)$$

where the normal-derivatives are evaluated on the outside and inside surfaces of the hole as indicated. Therefore, in terms of their Fourier components

$$\mathbf{p}_h = D_h \mathbf{u}_h, \quad (17)$$

where D_h is the diagonal matrix with main diagonal $\{n\epsilon_h/R_h\}$ for $n = 0, 1, 2, \cdots$. Also, by eq. (46)

$$\mathbf{p}_h = G_h \boldsymbol{\beta}_h - H_h \mathbf{u}_h. \quad (18)$$

From eq. (17) and eq. (18), $\boldsymbol{\beta}_h$ can be calculated in terms of $\mathbf{u}_h$. This in conjunction with the matrix-equation $T\boldsymbol{\beta} = \mathbf{u}$ is sufficient to determine first $\boldsymbol{\beta}_h$, then $\boldsymbol{\beta}$, and then by the second equation in (8), $\mathbf{p}$. Details are not supplied here.

The problem can also be generalized by specifying more complicated boundary conditions on the conductors. For example, the longitudinal component of the electric field satisfies a boundary condition of the form,

$$\mathbf{p} - S\mathbf{u} = \mathbf{g}, \quad (19)$$

(see Ref. 2) where S is some matrix and $\mathbf{g}$ some vector. It follows immediately from (8) that

$$(G - HT - ST)\boldsymbol{\beta} = \mathbf{g}. \quad (20)$$

Thus, under certain conditions on matrix S, $\boldsymbol{\beta}$ is obtained by inverting $(G - HT - ST)$, $\mathbf{u}$ is obtained as $T\boldsymbol{\beta}$, and $\mathbf{p}$ is then obtained from eq. (19).

The latter problem is involved in determining the elements of the admittance matrix for the cable (see Ref. 2, eq. 21). As shown in Ref. 2, this is an intermediate step for calculating the propagation modes of the cable and their associated propagation constants. The finite-difference technique, which had been used to calculate charge densities,[7,8] was not capable of dealing with this type of boundary condition. But even for cases such as the calculation of the capacitance matrix where the finite difference technique could be applied, the technique described here has a cost savings of more than one order of magnitude for 0.1 percent accuracy. Thus, the technique has proved to be flexible in solving potential problems associated with uniform cables and relatively inexpensive.

APPENDIX A

The First Laurent Series

When the cross-sectional plane of the cable is viewed as the complex plane, the potential function U is the real part of an analytic function

$f(z)$. In this appendix a Laurent series is deduced for $f(z)$ in a neighborhood of each conductor, starting from the Fourier series for the charge density and U at the surface of the conductor, as given in eq. (1) and eq. (2).

In the first layer of insulation for the mth conductor (for $R_m \leqq \rho_m \leqq R_{m1}$ $m = 1, \cdots, M$ or $R_0 \geqq \rho_0 \geqq R_{01}$), the potential function is

$$U(\rho_m, \phi_m) = u_{0m} + p_{0m}(R_m/\epsilon_m)\ell n\ \xi_m + \tfrac{1}{2} \sum_{n \neq 0} \times \exp(in\phi_m)[u_{nm}(\xi_m^n + \xi_m^{-n}) + p_{nm}(R_m/n\epsilon_m)(\xi_m^n - \xi_m^{-n})], \quad (21)$$

where $\xi_m = (\rho_m/R_m)$. This is validated by noting that it satisfies Laplace's equation and it satisfies the boundary conditions of eq. (1) and eq. (2) at $\rho_m = R_m$ (i.e., when $\xi_m = 1$).

In the second layer of insulation (for $R_{m1} \leqq \rho_m \leqq R_{m2}$ $m = 1, \cdots, M$ or $R_{01} \geqq \rho_0 \geqq R_{02}$), the potential function is

$$U(\rho_m, \phi_m) = u_{0m} + p_{0m}(R_m/\epsilon_m)\ \ell n r_m + p_{0m}(R_m/\epsilon_{m1})\ell n \xi_{m1} + \tfrac{1}{4} \sum_{n \neq 0} \exp(in\phi_m)[u_{nm}G_{nm}^{+}(\rho_m) + p_{nm}(R_m/n\epsilon_m)G_{nm}^{-}(\rho_m)], \quad (22)$$

where $\xi_{m1} = (\rho_m/R_{m1})$,

$$r_m = (R_{m1}/R_m), \quad \delta_m = (\epsilon_{m1}/\epsilon_m) \quad (23)$$

and

$$G_{nm}^{\pm}(\rho_m) = (r_m^n \pm r_m^{-n})(\xi_{m1}^n + \xi_{m1}^{-n}) + (\delta_m)^{-1}(r_m^n \mp r_m^{-n})(\xi_{m1}^n - \xi_{m1}^{-n}). \quad (24)$$

This satisfies Laplace's equation and the continuity conditions on U and $\tilde{\epsilon}(\partial U/\partial \rho_m)$ at the interface $\rho_m = R_{m1}$ (i.e., when $\xi_{m1} = 1$ and $\xi_m = r_m$). For U this is obvious, but to check the continuity of $\tilde{\epsilon}(\partial U/\partial \rho_m)$, it is useful to refer to the calculation,

$$(n/\rho_m)H_{nm}^{\pm}(\rho_m) \equiv \frac{\partial}{\partial \rho_m} G_{nm}^{\pm}(\rho_m) = (n/\rho_m)[(r_m^n \pm r_m^{-n})(\xi_{m1}^n - \xi_{m1}^{-n}) + (\delta_m)^{-1}(r_m^n \mp r_m^{-n})(\xi_{m1}^n + \xi_{m1}^{-n})]. \quad (25)$$

Outside the insulation (for $\rho_m \geqq R_{m2} +$, $m = 1, \cdots, M$ or $\rho_0 < R_{02}$), the analytic function whose real part matches U is

$$f(z) = u_{0m} + p_{0m}(R_m/\epsilon)\kappa_m + p_{0m}(R_m/\epsilon)\ell n \left(\frac{z - b_m}{R_m}\right) + \tfrac{1}{4} \sum_{n=1}^{\infty} \left(\frac{z - b_m}{R_m}\right)^n [u_{nm}A_{nm} + p_{nm}(R_m/n\epsilon_{ni})B_{nm}] + \tfrac{1}{4} \sum_{n=1}^{\infty} \left(\frac{z - b_m}{R_m}\right)^{-n} [u_{nm}^{*}E_{nm} + p_{nm}^{*}(R_m/n\epsilon_m)F_{nm}], \quad (26)$$

where * denotes complex conjugation,

$$\kappa_m = \delta_{m1}\delta_m \ell n r_m + \delta_{m1}\ell n r_{m1} - \ell n r_m r_{m1}, \tag{27}$$

$$r_{m1} = (R_{m2}/R_{m1}), \qquad \delta_{m1} = (\epsilon/\epsilon_{m1}), \tag{28}$$

and with $G^{\pm}_{nm} = G^{\pm}_{nm}(R_{m2})$ and $H^{\pm}_{nm} = H^{\pm}_{nm}(R_{m2})$,

$$\begin{aligned} A_{nm} &= (r_m r_{m1})^{-n}(G^{+}_{nm} + (\delta_{m1})^{-1}H^{+}_{nm}) \\ B_{nm} &= (r_m r_{m1})^{-n}(G^{-}_{nm} + (\delta_{m1})^{-1}H^{-}_{nm}) \\ E_{nm} &= (r_m r_{m1})^{n}(G^{+}_{nm} - (\delta_{m1})^{-1}H^{+}_{nm}) \\ F_{nm} &= (r_m r_{m1})^{n}(G^{-}_{nm} - (\delta_{m1})^{-1}H^{-}_{nm}). \end{aligned} \tag{29}$$

Substituting $z = b_m + R_{m2}\exp(\mathrm{i}n\phi_m)$ into $f(z)$ and taking the real part gives

$$\begin{aligned} &u_{0m} + p_{0m}(R_m/\epsilon)(\kappa_m + \ell n r_m r_{m1}) \\ &\quad + \tfrac{1}{8}\sum_{n\neq 0}\exp(\mathrm{i}n\phi_m)[u_{nm}(A_{nm}(r_m r_{m1})^n + E_{nm}(r_m r_{m1})^{-n}) \\ &\qquad + p_{nm}(R_m/n\epsilon_m)(B_{nm}(r_m r_{m1})^n + F_{nm}(r_m r_{m1})^{-n})], \end{aligned}$$

since $u_{-nm} = u^{*}_{nm}$ and $p_{-nm} = p^{*}_{nm}$. As is easily checked, this matches U in eq. (22) for $\rho_m = R_{m2}$ (i.e., for $\xi_{m1} = r_{m1}$). Likewise, it is straightforward to check that the ρ_m-derivative of the real part of $f(z)$ matches $(\epsilon_{m1}/\epsilon_m)\partial U/\partial\rho_m$ evaluated at $\rho_m = R_{m2}$-. This validates eq. (26) for $f(z)$.

When the dielectric is homogeneous, then $r_m = r_{m1} = \delta_m = \delta_{m1} = 1$. It follows that in this case $\kappa_m = 0$ and $A_m = B_{nm} = E_{nm} = -F_{nm} = 4$.

APPENDIX B

The Second Laurent Series

An analytic function $f(z)$ in the region separating the conductors has the form given in eq. (3) through eq. (5). In this appendix, a single Laurent series is derived for $f(z)$ in a neighborhood of each conductor (including the shield) by combining these equations.

The functions $f_m(z)$, as given in eq. (4) and eq. (5), are analytic in a neighborhood of b_k for $m \neq k$ and $k \neq 0$; hence, they are represented by a power series about $z = b_k$,

$$f_m(z) = \sum_{n=0}^{\infty} c_{mk}(n)\left(\frac{z - b_k}{R_k}\right)^n \qquad m = 0, 1, \cdots, M. \tag{30}$$

The coefficients $c_{mk}(n)$ are related to the derivatives of $f_m(z)$ at $z = b_k$ by the formula

$$c_{mk}(n) = \frac{(R_k)^n}{n!} \frac{\partial^n f_m}{\partial z^n}\bigg|_{z=b_k} \qquad n = 0, 1, 2, \cdots; \tag{31}$$

in particular, $c_{mk}(0) = f_m(b_k)$.

The results of the calculation in eq. (31) are indicated at the end of this appendix. When the formula is evaluated for $f_0(z)$, as given in eq. (4), the result is eq. (40), and for $f_m(z)$, as given in eq. (5), the result is the combination of eq. (41) and eq. (42). The combined Laurent series for $f(z)$ about $z = b_k$ is indicated in eq. (39).

In a neighborhood of the surrounding shield, the Laurent series of eq. (5) holds for $f_m(z)$ $(m = 1, \cdots, M)$. This is rewritten as

$$f_m(z) = \beta_{0m}\, \ell n \left(\frac{z}{R_0}\right) + \beta_{0m}\, \ell n \left(\frac{R_0}{R_m}\right) + g_m(z), \tag{32}$$

where

$$g_m(z) = \beta_{0m}\, \ell n \left(\frac{z - b_m}{z}\right) + \sum_{\ell=1}^{\infty} \beta_{\ell m} \left(\frac{z - b_m}{R_m}\right)^{-\ell}. \tag{33}$$

Since $g_m(\infty) = 0$, it follows that $g_m(R_0^2/z)$ is analytic in a neighborhood of $z = 0$; so it is represented there by a power series,

$$g_m\left(\frac{R_0^2}{z}\right) = \sum_{n=1}^{\infty} c_{m0}(n) \left(\frac{z}{R_0}\right)^n; \tag{34}$$

and accordingly,

$$g_m(z) = \sum_{n=1}^{\infty} c_{m0}(n) \left(\frac{z}{R_0}\right)^{-n}. \tag{35}$$

The coefficients $c_{m0}(n)$ for $m = 1, \cdots, M$ are obtained from the formula,

$$c_{m0}(n) = \frac{R_0^n}{n!} \frac{d^n}{dz^n} \left[g_m\left(\frac{R_0^2}{z}\right)\right]\bigg|_{z=0} \qquad n = 1, 2, \cdots. \tag{36}$$

The result of applying this formula to eq. (33) is

$$c_{m0}(n) = \left(\frac{b_m}{R_0}\right)^n \left(-\frac{\beta_{0m}}{n} + \sum_{\ell=1}^{n} \beta_{\ell m} \binom{n-1}{\ell-1} \left(\frac{b_m}{R_m}\right)^{-\ell}\right). \tag{37}$$

$$n = 1, 2, \cdots$$

The combined Laurent series for $f(z)$ in a neighborhood of the shield is

$$f(z) = \sum_{n=0}^{\infty} \beta_{n0} \left(\frac{z}{R_0}\right)^n + \sum_{m=1}^{M} \left[\beta_{0m}\, \ell n \left(\frac{z}{R_0}\right) + \beta_{0m}\, \ell n \left(\frac{R_0}{R_m}\right) + \sum_{n=1}^{\infty} c_{m0}(n) \left(\frac{z}{R_0}\right)^{-n}\right]. \tag{38}$$

The combined Laurent series for $f(z)$ about $z = b_k$ $(k \neq 0)$ is

$$f(z) = \sum_{m \neq k} c_{mk}(0) + \beta_{0k} \ell n \left(\frac{z - b_k}{R_k}\right) + \sum_{n=1}^{\infty} \beta_{nk} \left(\frac{z - b_k}{R_k}\right)^{-n} + \sum_{n=1}^{\infty} \left(\frac{z - b_k}{R_k}\right)^{n} \left(\sum_{m \neq k} c_{mk}(n)\right). \quad (39)$$

The coefficients obtained by carrying out the calculations in eq. (31) are

$$c_{0k}(n) = \left(\frac{b_k}{R_k}\right)^{-n} \sum_{\ell=n}^{\infty} \beta_{\ell 0} \binom{\ell}{n} \left(\frac{b_k}{R_0}\right)^{\ell} \quad k \neq 0, n = 0, 1, 2, \cdots \quad (40)$$

and for $m, k \neq 0$, $m \neq k$

$$c_{mk}(0) = \beta_{0m} \ell n \left|\frac{b_k - b_m}{R_m}\right| + \sum_{\ell=1}^{\infty} \beta_{\ell m} \left(\frac{b_k - b_m}{R_m}\right)^{-\ell} \quad (41)$$

$$c_{mk}(n) = \left(\frac{b_m - b_k}{R_k}\right)^{-n} \times \left[-\frac{\beta_{0m}}{n} + \sum_{\ell=1}^{\infty} \beta_{\ell m} \binom{\ell + n - 1}{n} \left(\frac{b_k - b_m}{R_m}\right)^{-\ell}\right] \quad (42)$$

for $n = 1, 2, \cdots$.

APPENDIX C

Equate the Two Laurent Series

Equating the constant, logarithmic, nth power, and $-n$th power terms $(n \geq 1)$, respectively, in the Laurent series about the shield [i.e., in eq. (26) and eq. (38)] gives the linear equations,

$$(i)\ \beta_{00} + \sum_{m=1}^{M} \beta_{m0} \ell n \frac{R_0}{R_m} = u_{00} + p_{00}(R_0/\epsilon)\kappa_0$$

$$(ii)\ \sum_{m=1}^{M} \beta_{0m} = p_{00}(R_0/\epsilon)$$

$$(iii)\ \beta_{n0} = \tfrac{1}{4}[u_{n0}A_{n0} + p_{n0}(R_0/n\epsilon_0)B_{n0}]$$

$$(iv)\ \sum_{m=1}^{M} c_{m0}(n) = \tfrac{1}{4}[u^{*}_{n0}E_{n0} + p^{*}_{n0}(R_0/n\epsilon_0)F_{n0}]. \quad (43)$$

For the kth wire $(k = 1, \cdots, M)$, the corresponding equations are

$$(i)\ \sum_{m \neq k} c_{mk}(0) = u_{0k} + p_{0k}(R_k/\epsilon)\kappa_k$$

$$(ii)\ \beta_{0k} = p_{0k}(R_k/\epsilon)$$

$$(iii)\ \sum_{m \neq k} c_{mk}(n) = \tfrac{1}{4}[u_{nk}A_{nk} + p_{nk}(R_k/n\epsilon_k)B_{nk}]$$

$$(iv)\ \beta_{nk} = \tfrac{1}{4}[u^{*}_{nk}E_{nk} + p^{*}_{nk}(R_k/n\epsilon_k)F_{nk}]. \quad (44)$$

The charge coefficients can be solved from (*ii*) and (*iii*) in the first set and from (*ii*) and (*iv*) in the second set to give

$$
\begin{aligned}
p_{00} &= (\epsilon/R_0) \sum_{m=1}^{M} \beta_{0m} \\
p_{n0} &= (4\epsilon_0 n/R_0 B_{n0})\beta_{n0} - (\epsilon_0 n A_{n0}/R_0 B_{n0})u_{n0} \\
p_{0k} &= (\epsilon/R_k)\beta_{0k} \\
p_{nk} &= (4\epsilon_k n/R_k F_{nk})\beta^*_{nk} - (\epsilon_k n E_{nk}/R_k F_{nk})u_{nk}.
\end{aligned} \tag{45}
$$

In terms of the vectors defined in (6) (viewed as column vectors), these equations are expressed

$$
\mathbf{p}_k = G_k \beta_k - H_k \mathbf{u}_k \qquad k = 0, 1, \cdots, M, \tag{46}
$$

where the $n\ell$-elements of these matrices for $n, \ell = 0, 1, \cdots$ are

$$
G_0(n, \ell) = \begin{cases} (0; e_1; \cdots; e_1) & n = 0 \\ (4\epsilon_0 n/R_0 B_{n0}) & n = \ell \neq 0 \\ 0 & \text{otherwise} \end{cases}
$$

$$
G_k(n, \ell) = \quad (k = 1, \cdots, M) \begin{cases} (\epsilon/R_k) & n = \ell = 0 \\ (4\epsilon_k n/R_k F_{nk})(*) & n = \ell \neq 0 \\ 0 & n \neq \ell \end{cases}
$$

$$
H_0(n, \ell) = \begin{cases} (\epsilon_0 n A_{n0}/R_0 B_{n0}) & n = \ell \\ 0 & n \neq \ell \end{cases}
$$

$$
H_k(n, \ell) = \quad (k = 1, \cdots, M) \begin{cases} (\epsilon_k n E_{nk}/R_k F_{nk}) & n = \ell \\ 0 & n \neq \ell \end{cases}
$$

In terms of the joint vectors defined in (7),

$$
\mathbf{p} = G\beta - H\mathbf{u}, \tag{47}
$$

where G and H are the direct sum of the matrices, G_k and H_k, respectively, for $k = 0, 1, \cdots, M$ (see Ref. 12, p. 159).

The charge components can be eliminated in the original set of equations by means of the identity,

$$
\begin{aligned}
B_{nk}E_{nk} - A_{nk}F_{nk} &= (32/\delta_k \delta_{k1}) \\
n = 0, 1, \cdots \quad &\text{and} \quad k = 0, \cdots, M.
\end{aligned} \tag{48}
$$

The result is the set of equations,

$$(i)\ \beta_{00} + \sum_{m=1}^{M} \beta_{0m}\left(\ell n \frac{R_0}{R_m} - \kappa_0\right) = u_{00}$$

$$(ii)\ -(\delta_0\delta_{01}F_{n0}/8)\beta_{n0} + (\delta_0\delta_{01}B_{n0}/8)\sum_{m=1}^{M} c^*_{m0}(n) = u_{n0}$$

$$(iii)\ \sum_{m\neq k} c_{mk}(0) - \beta_{0k}\kappa_k = u_{0k}$$

$$(iv)\ -(\delta_k\delta_{k1}F_{nk}/8)\sum_{m\neq k} c_{mk}(n) + (\delta_k\delta_{k1}B_{nk}/8)\beta^*_{nk} = u_{nk}$$

$$\text{for } k = 1, \cdots, M \quad \text{and} \quad n = 1, 2, \cdots. \tag{49}$$

These equations can be expressed in the matrix form

$$T\beta \equiv \begin{bmatrix} T_{00}T_{01}\cdots T_{0M} \\ T_{10}T_{11}\cdots T_{1M} \\ \vdots \\ T_{M0}T_{M1}\cdots T_{MM} \end{bmatrix}\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_M \end{bmatrix} = \begin{bmatrix} \mathbf{u}_0 \\ \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_M \end{bmatrix} \equiv \mathbf{u}, \tag{50}$$

where β and $\mathbf{u}$ are viewed as column vectors. The symbols T_{km} for $k, m = 0, \cdots, M$ denote submatrices of T defined as follows:

$$T_{00}(n, \ell) = \begin{cases} 1 & n = \ell = 0 \\ (\delta_0\delta_{01}F_{n0}/8) & n = \ell \neq 0 \\ 0 & n \neq \ell \end{cases}$$

$$T_{0m}(n, \ell) = \quad (m = 1, \cdots, M)$$

$$\begin{cases} \ell n \dfrac{R_0}{R_m} - \kappa_0 & n = \ell = 0 \\ -(\delta_0\delta_{01}B_{n0}/8n)\left(\dfrac{b_m}{R_0}\right)^n \ (*) & \ell = 0;\ n = 1, 2, \cdots \\ (\delta_0\delta_{01}B_{n0}/8)\left(\dfrac{b_m}{R_0}\right)^n\left(\dfrac{b_m}{R_m}\right)^{-\ell}\dbinom{n-1}{\ell-1} & \ell = 1, \cdots, n;\ n = 1, 2, \cdots \\ 0 & \text{otherwise} \end{cases}$$

$$T_{k0}(n, \ell) = \quad (k = 1, \cdots, M)$$

$$\begin{cases} \left(\dfrac{b_k}{R_0}\right)^\ell & n = 0;\ \ell = 0, 1, \cdots \\ (\delta_k\delta_{k1}F_{nk}/8)\dbinom{\ell}{n}\left(\dfrac{b_k}{R_k}\right)^{-n}\left(\dfrac{b_k}{R_0}\right)^\ell & 1 \leqq n \leqq \ell \\ & \ell = 1, 2, \cdots \\ 0 & \text{otherwise} \end{cases}$$

$$T_{kk}(n, \ell) = \begin{cases} -\kappa_k & n = \ell = 0 \\ (\delta_{k1}\delta_k B_{nk}/8)(*) & n = \ell \neq 0 \\ 0 & n \neq \ell \end{cases}$$
$$(k = 1, \cdots, M)$$

$$T_{km}(n, \ell) = \begin{cases} \ell n \left|\frac{b_k - b_m}{R_m}\right| & n = \ell = 0 \\ \left(\frac{b_k - b_m}{R_m}\right)^{-\ell} & n = 0;\ \ell = 1, 2, \cdots \\ (\delta_k\delta_{k1}F_{nk}/8n)\left(\frac{b_m - b_k}{R_k}\right)^{-n} & \ell = 0; \\ & n = 1, 2, \cdots \\ -(\delta_k\delta_{k1}F_{nk}/8)\binom{\ell + n - 1}{n}\left(\frac{b_m - b_k}{R_k}\right)^{-n}\left(\frac{b_k - b_m}{R_m}\right)^{-\ell} & \end{cases}$$
$$(k, m \neq 0 \quad k \neq m)$$

$$n, \ell \neq 0.$$

The symbol (*) throughout means that complex conjugation is to be performed. These elements of the various submatrices are read off the appropriate coefficients in the system of equations in (49) in conjunction with the equations for the $c_{mk}(\cdot)$ in eqs. (37), (40), (41), and (42).

REFERENCES

1. P. I. Kuznetsov and R. L. Stratonovich, *The Propagation of Electromagnetic Waves in Multiconductor Transmission Lines*, New York: Pergamon Press, 1964, Chap. 1, 2.
2. T. A. Lenahan, "The Theory of Uniform Cables—Part I: Calculation of Propagation Parameters," B.S.T.J., this issue, pp. 597–610.
3. G. M. Goluzin, "Solution of the Principal Plane Problems of Mathematical Physics for the Laplace Equation and Multiply Connected Domains Limited by Circles (Method of Functional Equations)," Mat Sbornik, *41*, No. 2 (1934), pp. 246–276.
4. J. D. Nordgard, "The Capacitances and Surface-Charge Distributions of a Shielded Balanced Pair," IEEE Trans. Microw. Theory Tech., *24*, No. 2 (Feb. 1976), pp. 94–100.
5. J. W. Craggs, "The Determination of Capacity for Two-Dimensional Systems of Cylindrical Conductors," Quart. J. Math. Oxford, Series I, *17* (1946), pp. 131–137.
6. C. M. Miller, "Capacitances of a Shielded Balanced-Pair Transmission Line," B.S.T.J., *51*, No. 3 (March 1972), pp. 759–776.
7. H. W. Friesen, unpublished work.
8. W. T. Anderson, unpublished work.
9. T. A. Lenahan, "Experimental Test of Propagation-Parameter Calculations for Shielded Balanced Pair Cables," B.S.T.J., this issue, pp. 627–636.
10. L. V. Ahlfors, *Complex Analysis*, New York: McGraw Hill, 1953, p. 147.
11. G. Hsiao and R. C. MacCamy, "Solution of Boundary Value Problems by Integral Equations of the First Kind," SIAM Rev., *15*, No. 4 (October 1973), pp. 687–705.
12. K. Hoffman and R. Kunze, *Linear Algebra*, Englewood Cliffs, N.J.: Prentice-Hall, 1961.

THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 56, No. 4, April 1977
Printed in U.S.A.

Experimental Test of Propagation-Parameter Calculations for Shielded Balanced Pair Cables

By T. A. LENAHAN

(Manuscript received October 5, 1976)

Experimental verification is presented for a recently developed method for calculating propagation parameters, such as loss and phase-shift per length, of uniform cables. For the shielded balanced pair (SBP), a computer program has been written to carry out the calculations. In this paper, these calculations are compared with measurements of three different cables that can be modeled as SBPs. Agreement between calculation and measurement was within 3.3 percent over the frequency range 50 Hz to 10 MHz. The greatest deviation occurred in a cable with substantial air space in the interstitial dielectric; for the other two cables agreement was within 2 percent.

I. INTRODUCTION

A method for calculating propagation parameters of uniform cables is presented in Ref. 1. The method depends on algorithms for determining the charge densities on the wires and shield, and this is presented in Ref. 2. For the shielded balanced pair (SBP) in particular, a computer program has been written to carry out the required calculations. To test the method, three different SBP cables were fabricated and, with the cables laid straight, their attenuation and phase-shift per length were measured over a broad range of frequencies and then compared with calculated values. The results of this experimental test are presented in this paper.

The three cables were constructed to conform to the SBP model. In this model, the two wires are identical, they are straight and parallel, and they are symmetrically located inside a cylindrical, metallic shield of uniform thickness. Also, the dielectric material separating the wires and shield is homogeneous and isotropic. A cross section of an ideal SBP is shown in Fig. 1.

The first cable is an untwisted version of a standard twisted-pair cable built by Western Electric, called the 754E. This cable deviates somewhat

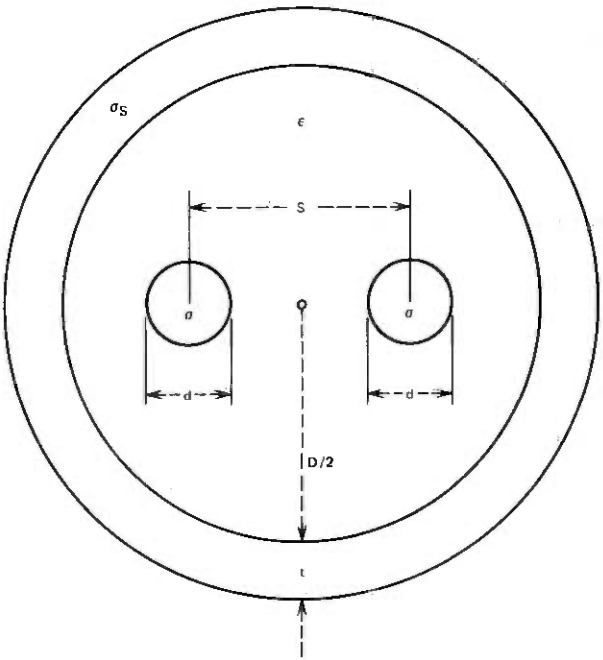
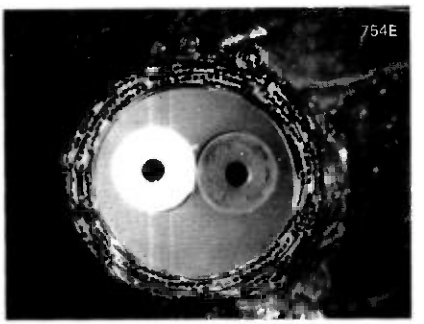


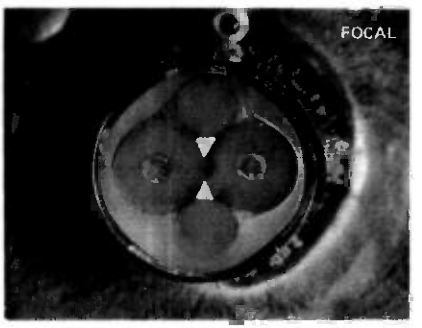
Fig. 1—Cross section of a typical shielded balanced pair.

from the SBP model since its shield instead of being solid is braided from fine Cu wires coated with Ag. The second cable was constructed by Bell Laboratories and is called the FOCAL cable. For this cable, a polyethylene belt was extruded over two insulated 19-gage wires (as in the 754E cable) and bordered by two polyethylene rods for stability, resulting in substantial air space. This was then covered by a solid Al shield. The third cable was also constructed by Bell Laboratories; it was designed to accentuate the proximity effect by having the wires and shield relatively close to each other. This is called the proximity cable. Photographs of the cross sections of each cable are shown in Fig. 2.

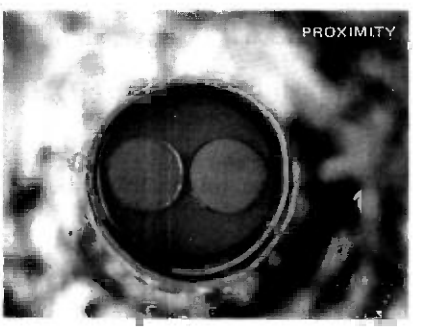
The geometric and material parameters used in the SBP model for these three cables correspond to certain electrical measurements of the cable. Though the parameters are specified, in practice the nominal values are neither accurate nor constant enough along the cable to permit testing the model. Also, the SBP model, though essentially descriptive of the cable, does not account for any unbalance, air spaces in the dielectric, eccentricity of the wires and shield, axial variations of the wires and other nonuniformities along the cable, and other practical matters. Accordingly, with the exception of cable length, the input parameters are computed (in a way now to be described) as average values from electrical measurements.



(MAGNIFIED 8x)



(MAGNIFIED 9x)



(MAGNIFIED 20x)

Fig. 2—Photographs of cable cross sections.

For each cable, the wires were made of Cu; and when the temperature and purity are known, the conductivity, σ, is determined to within 0.5 percent. The temperature is assumed to be 73°F, which corresponds to a σ of 5.73749 mho/m. The diameter of the wires d is computed from the measurement of their DC resistance by the formula,

$$R_{DC} = \frac{\ell}{\sigma(\text{Area})} = \frac{4\ell}{\pi\sigma d^2} \text{ or } d = 2\left(\frac{\ell}{\pi\sigma R_{DC}}\right)^{1/2}, \quad (1)$$

where ℓ denotes the length of the cable. The distance S separating the two wire centers is computed from the measurement of low-frequency inductance L by the formula,

$$L = \frac{\mu_0}{\pi} \ell n \left(\frac{2S}{d}\right) + \frac{\mu_0}{4\pi} \text{ or } S = \frac{d}{2} \exp\left[\frac{\pi L}{\mu_0} - 1/4\right], \tag{2}$$

where μ_0 denotes the free-space permeability (see Ref. 3, page 155). This formula assumes that the materials comprising the cable are nonmagnetic. It also assumes that the wires are straight; so the cable must be unreeled to get the proper measurement of L.

The "effective" dielectric constant ϵ of the model is chosen so that the measured and calculated values of the phase shift per mile β agree at a frequency of 10 MHz. In the FOCAL and proximity cables there were air spaces in the cable, and the effective dielectric constant for these represents a kind of average of the dielectric constants for air and polyethylene. The value of ϵ is estimated from the formula,

$$\beta(10 \text{ MHz}) = \omega\sqrt{\epsilon\mu_0} + \delta = 337.286 \sqrt{\frac{\epsilon}{\epsilon_0}} + \delta, \tag{3}$$

in units of radians per mile, where ω denotes the excitation frequency in radians per second, ϵ_0 denotes the free-space permittivity, and δ is a positive number that corrects for proximity effect (e.g., calculation has shown this is approximately 4 rad/mi for the 754E cable and 36 rad/mi for the proximity cable). Since δ is positive, an upper bound for ϵ is

$$\epsilon_u = [\beta(10 \text{ MHz})/337.286]^2 \epsilon_0.$$

Experience has shown that, apart from ϵ, δ is relatively insensitive to changes in the input parameters; hence, it can be estimated as the calculated β at 10 MHz (with nominal dimensions and with $\epsilon = \epsilon_u$) minus the measured β at 10 MHz. When this estimate of δ is inserted into eq. (3), an estimate of ϵ can be obtained that usually gives a calculated β at 10 MHz within 1 percent of measurement.

The inside diameter of the shield D is chosen so that the calculated mutual capacitance matches the measured value (though a variation of this is used for the FOCAL cable). This calculation presumes that the parameters d, S, and ϵ have already been obtained. The thickness t and the conductivity σ_s of the shield are selected to match the DC resistance of the shield according to the formula,

$$R_{DC,S} = \frac{\ell}{\sigma_s(\text{Area})} = \frac{\ell}{\sigma_s \pi t(D + t)}. \tag{4}$$

There is one degree of freedom left in selecting the pair (σ_s, t); this is used to minimize the deviation between measured and calculated values of

Table I — Initial measurements on 754E cable

Length	R_{DC} Cond. 1	R_{DC} Cond. 2	R_{DC} Shield	C_m	C_{1g}	C_{2g}	$L_{50\ Hz}$
1043 ft	9.009Ω	8.700Ω	1.0044Ω	13.072 nf	18.24 nf	18.02 nf	0.2706 mh

loss per mile (α) over the range of frequencies for which there are measurements.

By means of this procedure a set of input parameters was chosen for each cable. Parameters for which there was latitude in choosing were selected by trial and error, and the result is not optimal in any respect.

In general, it was found that calculations of α agreed with measurements to within 3.3 percent for frequency samples over the range 50 Hz to 10 MHz. The best agreement was for the proximity cable, which had a deviation within 1.6 percent over the range 5 kHz to 10 MHz. For β, agreement was always within 1 percent.

In the next section, the detailed results of the experiment are presented. In the final section, conclusions are drawn and application to multipair cables is discussed.

II. RESULTS

In this section, the results of the experimental test are presented. First, initial measurements were made on a Wheatstone bridge of mutual capacitance, capacitance to ground of both wires, DC resistances, and low-frequency inductance for the three cables. These, together with the length of the samples, are shown in Tables I, IV, and VII for the 754E, FOCAL, and proximity cables, respectively.

Next, measurements of loss α in dB/mi and phase-shift β in rad/mi were made on the computer-operated transmission measuring set[4] (COTMS) corrected for varying terminal impedance at low frequency[5] for frequencies from 4 kHz to 10 MHz. For lower frequencies, 50 Hz to 10 kHz, values of α and β were deduced from bridge measurements of AC resistance, AC inductance, and capacitance. It was found in this

Table II — Parameter estimates for 754E cable

	Description	Estimates
σ	Cu wire at 73°	5.73749×10^7 mho/m
d	19 gage (35.9 mils)	34.84 mils
S	117 to 127 mils	116 mils
ϵ	Polyethylene	$2.288\ \epsilon_0$
D	290 mils	280 mils
t	Shield has braided wire construction	38 mils
σ_s	about 32 mils thick	1.292×10^7 mho/m

Table III — Calculated vs measured α and β for 754E cable

	Calc. α (dB/mi)	Meas. α (dB/mi)	Diff. (%)	Calc. β (rad/mi)	Meas. β (rad/mi)	Diff. (%)
50 Hz	0.267	0.265	<1	0.0309	0.0305	1
100 Hz	0.376	0.376	<1	0.0438	0.0437	<1
500 Hz	0.826	0.822	<1	0.0998	0.0992	<1
1 kHz	1.143	1.135	<1	0.1446	0.1438	<1
5 kHz	2.209	2.169	1.8	0.387	0.386	<1
10 kHz	2.700	2.682	<1	0.649	0.656	<1
20 kHz	3.047	3.074	<1	1.176	1.185	<1
50 kHz	3.507	3.580	2.0	2.817	2.809	<1
80 kHz	4.013	4.045	<1	4.457	4.437	<1
100 kHz	4.372	4.383	<1	5.53	5.52	<1
500 kHz	9.391	9.263	1.4	26.54	26.52	<1
1 MHz	13.109	13.085	~1	52.48	52.45	<1
5 MHz	28.85, $F_p = 0$ 29.97, $F_p = 1000$ μrad	29.830	—	258.38	258.06	<1
10 MHz	40.66, $F_p = 0$ 42.88, $F_p = 1000$ μrad	42.796	—	514.84	514.09	<1

process that the COTMS measurements of α at low frequencies had to be increased (~1 percent for 10 kHz and ~2 percent for 5 kHz) for the 754E and FOCAL cables, presumably because the low-loss values went beyond the sensitivity of the machine, which is $\frac{1}{100}$ dB. The losses in the proximity cable were large enough at all frequencies to be within the range of sensitivity for COTMS.

The parameters of the SBP model for the three cables were calculated as indicated in the introduction when feasible. But each cable deviated to some extent from the model, so there were some exceptions and extra considerations to accommodate the deviations. This is discussed in the appendix. The parameters selected for the SBP model are given in Tables

Table IV — Initial measurements on FOCAL cable

Length	R_{DC} Cond. 1	R_{DC} Cond. 2	R_{DC} Shield	C_m	C_{1g}	C_{2g}	L_{50Hz}
249.2 ft	2.0086Ω	2.0083Ω	0.3968Ω	2.956nf	4.098nf	4.069nf	0.065mh

Table V — Parameter estimates for FOCAL cable

	Description	Estimates
σ	Cu wire at 73°	5.73749×10^7 mho/m
d	19 gage (35.9 mils)	36.068 mils
S	125 mils	119.27 mils
ϵ	Polyethylene & air spaces	2.06 ϵ_0
D	280 mils	278 mils
t	Solid Al sheet 10-mil thick with overlap	9.22 mils
σ_s		3.54×10^7 mho/m

Table VI — Calc. vs measured α and β for FOCAL cable

	Calc. α (dB/mi)	Meas. α (dB/mi)	Diff. (%)	Calc. β (rad/mi)	Meas. β (rad/mi)	Diff. (%)
50 Hz	0.247	0.251	1.5	0.029	0.029	—
100 Hz	0.349	0.354	1.5	0.041	0.041	—
500 Hz	0.764	0.775	1.5	0.093	0.094	1
1 kHz	1.054	1.069	1.5	0.134	0.136	1
5 kHz	2.012	2.034	1	0.364	0.370	1
10 kHz	2.491	2.516	<1	0.617	0.628	1
20 kHz	2.890	2.799	3.3	1.117	1.124	<1
50 kHz	3.339	3.268	2.1	2.651	2.677	<1
80 kHz	3.764	3.735	<1	4.192	4.223	<1
100 kHz	4.069	4.076	<1	5.213	5.223	<1
500 kHz	8.397	8.229	2.1	25.122	25.071	<1
1 MHz	11.751	11.421	2.9	49.714	49.752	<1
5 MHz	25.757, $F_p = 0$ 26.289, $F_p =$ 500 μrad	25.884	—	244.973	245.048	<1
10 MHz	35.260, $F_p = 0$ 37.320, $F_p =$ 500 μrad	37.194	—	488.238	488.179	<1

II, V, and VIII for the 754E, FOCAL, and proximity cables, respectively.

Calculations of α and β for the model over the range of measured frequencies are given in Tables III, VI, and IX for the 754E, FOCAL, and proximity cables, respectively. These tables also give the measured values and the percent deviation in the comparison of calculated and measured values.

The calculations for the models agreed with measurement for all frequencies considered and all cables to within 3.3 percent for α and within 1 percent for β. For the 754E cable, calculation of α agreed with measurement to within 2 percent, and for the proximity cable, agreement

Table VII — Initial measurements on proximity cable

Length	R_{DC} Cond. 1	R_{DC} Cond. 2	R_{DC} Shield	C_m	C_{1g}	C_{2g}	$L_{50\ Hz}$
634.4 ft	3.273 Ω	3.267 Ω	4.63 Ω	43.396 nf	30.59 nf	28.80 nf	0.0773 mh

Table VIII — Parameter estimates for proximity cable

	Description	Estimates
σ	Cu wire at 73°	5.73749×10^7 mho/m
d	17 gage (45.3 mils)	45.06 mils
S	50.3 mils	47.74 mils
ϵ	Polyethylene with air spaces	$2.132\epsilon_0$
D	105.6 mils	109.9 mils
t	4-mil-thick Al	5.3 mils
σ_s	Shield with overlap	3.365×10^7 mho/m

Table IX — Calculated vs measured α and β for proximity cable

	Calc. α (dB/mi)	Meas. α (dB/mi)	Diff. (%)	Calc. β (rad/mi)	Meas. β (rad/mi)	Diff. (%)
5 kHz	4.109	4.096	<1	0.671	0.671	<1
10 kHz	5.284	5.310	<1	1.118	1.114	<1
20 kHz	7.150	7.122	<1	1.970	1.976	<1
50 kHz	11.998	11.895	<1	4.211	4.229	<1
80 kHz	15.817	15.670	<1	6.241	6.263	<1
100 kHz	18.016	18.037	<1	7.538	7.573	<1
500 kHz	44.219	44.953	1.6	31.338	31.248	<1
1 MHz	67.295	67.498	<1	59.445	59.231	<1
5 MHz	181.01, $F_p = 0$ 181.60, $F_p = 500$ μrad	179.42	~1	271.03	270.957	<1
10 MHz	270.02, $F_p = 0$ 271.34, $F_p = 500$ μrad	267.18	~1	528.09	528.12	<1

was within 1.6 percent. Agreement between calculation and measurement for the FOCAL cable was the poorest of the three (with a maximum deviation of 3.3 percent). This may be attributed to the presence of substantial air space in the dielectric region and to the short length of the sample (249.2 ft). Since the sensitivity of COTMS is $\frac{1}{100}$ dB, for this cable the measurements are uncertain to about (249.2/5280) (1/200) $\simeq$ 0.1 dB/mi, an amount which exceeds the difference of calculation and measurement for frequencies below 500 kHz.

The power factor F_p of the dielectric is generally important to the loss of a cable when the frequency exceeds 1 MHz. Since an independent measurement of F_p was not available for this experiment, α is calculated for power factors equal to 0 and to some upper bound. This gives an interval of possible values for α. As shown in Table III, the measured loss for the 754E cable at 5 MHz and 10 MHz lies in such an interval when $0 \leqq F_p \leqq 1000$ μrad. This range of F_p is within the usual range of values estimated from measurements of pairs in a multipair cable.[6] The measured loss for the FOCAL cable at 5 MHz and 10 MHz lies within the calculated range when $0 \leqq F_p \leqq 500$ μrad, as shown in Table VI. The calculated loss for the proximity cable at 5 MHz and 10 MHz (assuming $F_p = 0$) exceeds the measurements, so no estimate of F_p is possible for this cable. If, however, F_p is assumed to be 500 μrad, for example, then the error is 1.5 percent at 10 MHz.

III. SUMMARY AND CONCLUSIONS

The SBP model has been used to represent three different single-pair cables. Though each of these departs from the model to some extent, the cables are close enough to SBPs so that the model represents them with sufficient accuracy for engineering purposes. The calculations agreed

with measurement in all cases and all frequencies considered to within 3.3 percent, and usually it was closer. Though more controlled tests may be devised to test the SBP theory, the tests presented here indicate enough agreement between theory and practice that it could be used to predict the electrical behavior of a nominally balanced pair in a shield.

By a similar controlled experiment, the theory of the uniform cable[1,2] might be checked for unbalanced pairs in a shield and for other uniform cables. Given a test cable, the DC resistances would give the wire diameters of the model, the low-frequency inductances would give the wire separations, the capacitance measurements would give the size and location of the shield, and the range of phase velocities would give an effective dielectric constant. The shield thickness and conductivity could be estimated as before. In general, there will be several excitation modes having distinct propagation constants; their measurements could then be compared with calculations on the deduced model, as was done for the SBP.

The major application of the theory of the SBP, or more generally, of uniform cables is to a pair in a multipair cable. This extension is complicated by these factors:

(*i*) Nonuniformities along the cable arise from twisting, stranding, cabling, and various perturbations.

(*ii*) The shield of a pair effectively consists of the other pairs around it; thus, the effective shield is caged, not solid.

(*iii*) Small unbalance of 1 to 2 percent is almost always present.

(*iv*) The dielectric materials around the wires often are inhomogeneous.

Detailed consideration of points (*ii*), (*iii*), and (*iv*) in the context of uniform cables is in progress, but consideration of twisting and stranding goes beyond the scope of the uniform cable model.

IV. ACKNOWLEDGMENTS

The author wishes to acknowledge R. C. Sacks for many discussions on this work and experimental support.

APPENDIX A

Details of Parameter Selection for the Model

Each of the three cables deviates from the SBP model to some extent. In this appendix, details are provided of the parameter selection process used to accommodate these deviations.

Initial measurements of the 754E cable indicated a resistive unbalance of 3 percent. This means the wire diameter d and, hence, the wire separation S could not be exactly determined. Instead, a range of possible

values for d and S was indicated, and selection was made within these ranges to give acceptable agreement between calculated and measured values of α for the frequencies considered.

The FOCAL and proximity cables have insignificant resistive unbalance; hence, unambiguous values of d and S can be calculated for them as indicated in the introduction. Though all the cables have some capacitive unbalance (with a maximum of 5 percent for the proximity cable), this has no direct effect on the selection of parameters for the model.

A significant departure from the model is the presence of air spaces in the dielectric region for the FOCAL and proximity cables. This is indicated by the photographs in Fig. 2 and by the effective dielectric constants, $2.06\epsilon_0$ for the FOCAL cable and $2.132\epsilon_0$ for the proximity cable, which are lower than the dielectric constant of polyethylene ($2.29\epsilon_0$). The air spaces cause no direct ambiguity in selecting parameters for the model, as did the resistive unbalance, but they do seem to affect the appropriateness of the model itself. For the FOCAL cable in particular, a set of parameters for the model could not be found that satisfies all the conditions of the introduction and that gives a loss (α) within 5 percent of measurement for all frequencies considered. Instead, a compromise model is used whose mutual capacitance is 2.8-percent lower than measured.

For the proximity cable, no compromise model is required despite the presence of some air space, and the mutual capacitance of the model agrees with measurement to within 0.1 percent. The effective dielectric computed for the 754E cable indicates an absence of air space, and again the mutual capacitance of the model agrees with measurement to within 0.1 percent.

REFERENCES

1. T. A. Lenahan, "The Theory of Uniform Cables—Part I: Calculation of Propagation Parameters," B.S.T.J., this issue, pp. 597–610.
2. T. A. Lenahan, "The Theory of Uniform Cables—Part II: Calculation of Charge Components," B.S.T.J., this issue, pp. 611–625.
3. E. Weber, *Electromagnetic Theory, Static Fields and Their Mapping*, New York: Dover Publications, 1965.
4. W. J. Geldart, G. D. Haynie, and R. G. Schleich, "A 50 Hz–250 MHz Computer Operated Transmission Measuring Set," B.S.T.J., *45*, No. 5 (May–June 1969), pp. 1339–1381.
5. T. D. Nantz and R. C. Sacks, unpublished work, June 1974.
6. J. Kreutzberg, private communication, May 1976.

Contributors to This Issue

Anthony S. Acampora, B.S.E.E., M.S.E.E., and Ph.D. from the Polytechnic Institute of Brooklyn in 1968, 1970, and 1973, respectively. Bell Laboratories, 1968–. At Bell Laboratories, Mr. Acampora has worked in the fields of high power microwave transmitters and radar signal processing. He is currently engaged in satellite systems research and channel coding theory. Eta Kappa Nu, Sigma Xi, and IEEE.

Thomas D. Dudderar, B.S.M.E., 1957, Lehigh University; Sc.M., 1961, New York University; Ph.D., 1966, Brown University; Bell Laboratories, 1966–. Mr. Dudderar works in the Materials Science and Engineering Division. He has published research papers on mechanical properties of materials, photoelasticity, and holographic interferometry and has been awarded patents on mechanical testing techniques and stress analysis using holographic interferometry.

Douglas W. Hill, B.S., 1964, California Institute of Technology; M.S. (E.E.), 1965, Stanford University; Ph.D. (Mathematics), 1973, University of New Mexico; U.S. Air Force 1965–1969; Bell Laboratories 1973–1976; Motorola 1976–. While at Bell Laboratories, Mr. Hill studied the accuracy of traffic engineering procedures and developed new methods for trunk network administration. Member, SIAM, IMS, IEEE.

Terrence A. Lenahan, S.B., S.M., 1964, Massachusetts Institute of Technology; Ph.D., 1970, University of Pennsylvania; Bell Laboratories 1970—. Mr. Lenahan is presently concerned with the mathematical analysis of multipair communication cables. Member, SIAM, American Mathematical Society, American Association for the Advancement of Science.

Benjamin F. Logan, Jr., B.S. (Electrical Engineering), 1946, Texas Technological College; M.S., 1951, Massachusetts Institute of Technology; Eng.D.Sc. (Electrical Engineering), 1965, Columbia University; Bell Laboratories, 1956—. While at MIT, Mr. Logan was a research as-

sistant in the Research Laboratory of Electronics, investigating characteristics of high-power electrical discharge lamps. Also at MIT he engaged in analogue computer development at the Dynamic Analysis and Control Laboratory. From 1955 to 1956 he worked for Hycon-Eastern, Inc., where he was concerned with the design of airborne power supplies. He joined Bell Laboratories as a member of the Visual and Acoustics Research Department, where he was concerned with the processing of speech signals. Currently, he is a member of the Mathematical Research Department. Member, Sigma Xi, Tau Beta Pi.

James McKenna, B.Sc. (Mathematics), 1951, Massachusetts Institute of Technology; Ph.D. (Mathematics), 1961, Princeton University; Bell Laboratories, 1960–. Mr. McKenna has done research in quantum mechanics, electromagnetic theory, and statistical mechanics. He has recently been engaged in the study of nonlinear partial differential equations that arise in solid-state device work, in the theory of stochastic differential equations, and the theory of elastic wave propagation.

Richard H. Moseley, Associate Degree (EE), 1953, Connecticut State Technical Institute; BS(EE), 1965, MS(EE), 1969, Newark College of Engineering; Bell Laboratories, 1953—. Mr. Moseley worked on various defense systems until joining the Single Sideband Microwave Radio Department in 1970. He is currently engaged in single sideband system design.

Siegfried B. Pirkau, A.A.S., New Hampshire Vocational Technical College (NHVTC), 1969; Bell Laboratories, 1969—. Mr. Pirkau's first assignment was work involved with computer-simulated Picturephone® coding. In 1971 he began work on SSB radio in the Repeater Evaluation Group, where he became involved with the design and construction of special test equipment.

Judith B. Seery, B.A., 1968, College of St. Elizabeth; M.S., 1972, New York University; Bell Laboratories, 1968–. Ms. Seery does computing and analysis in the Mathematics and Statistics Research Center. She has recently participated in problems in fiber optics, minimal spanning networks, and multidimensional scaling. Member, Mathematical Association of America, Association for Women in Mathematics.

Peter G. Simpkins, Diploma in Technology, 1957, University of London; M.S., 1960, California Institute of Technology; Ph.D., 1965, Imperial College, London; AVCO Corporation 1965–1968; Bell Laboratories, 1968–. Mr. Simpkins is currently working in the Materials Research Laboratory. He has published articles on gas dynamics, fluid mechanics, underwater acoustics, and fracture mechanics. During a leave of absence in 1974, he was Senior Research Fellow at the University of Southampton.

Richard B. Swerdlow, B.S.E.E., 1960, University of Pennsylvania (Moore School); M.S.E.E., 1962, Massachusettes Institute of Technology; Bell Laboratories, 1963—. Mr. Swerdlow has supervised groups in the analysis of antiballistic missile systems and a single-sideband (SSB) radio transmission system. He is currently working on special problems in carrier transmission. Mr. Swerdlow has been awarded two patents.

Abstracts of Papers by Bell System Authors Published in Other Journals

CHEMISTRY

Assignment of a Ligand in Stellacyanin by a Pulsed Electron Paramagnetic Resonance Method. W. B. Mims and J. Peisach, Biochemistry, *15*, No. 17 (1976), pp. 3863–3869. The electron spin echo decay envelope for the blue copper protein, stellacyanin, and for a number of other Cu(II) complexes has been studied. Particular attention was given to the form of the "nuclear modulation" patterns, which show the effects of coupling between the electron spin and the neighboring nuclei. The envelopes for the hydrated cupric complex and for copper(II) glycylglycine were essentially the same and indicative of the coupling to protons. The peptide complex contains nitrogen nuclei coupled directly to Cu(II), but the coupling constant is so large for these nuclei that a modulation pattern ascribable to ^{14}N is not seen. For copper(II) bovine serum albumin, on the other hand, a contribution due to the coupling of the remote nitrogen belonging to a histidyl imidazole ligand was observed. The modulation pattern for this complex and for stellacyanin closely resembled one another, strongly suggesting that an imidazole is ligated to the copper in this blue protein.

Auger Studies of Au Diffusion Through Pt Films: Dependence on Pt Thickness and Annealing Ambient. C. C. Chang and G. Quintana, Appl. Phys. Lett., *29*, (Nov. 1976), pp. 453–454. The outdiffusion kinetics of Au through Pt/Au couples with ~2000 Å to ~10,000 Å Pt films, heat treated between 250 and 350°C in 1 atm. N_2 and in vacuum, were determined using Auger electron spectroscopy. The Au outdiffusion time, in 1 atm. N_2 ambient, was approximately proportional to $d^{1.3}$ for Pt thickness $2000 \leq d \leq 10{,}000$ Å. Au was found to outdiffuse through ~10,000 Å Pt at a rate *fifteen* times faster in 1 atm. N_2 than in vacuum, at 350°C.

Electrodeposition of Gold. Depolarization Effects Induced by Heavy Metal Ions. J. D. E. McIntyre and W. F. Peck, Jr., J. Electrochem. Soc., *123* (December 1976), pp. 1800–1813. Addition of trace quantities (ppm) of heavy metal ions such as Pb(II) or Tl(I) to soft gold electroplating baths induces a marked cathode depolarization effect that extends the current density range in which bright, uniform, fine-grained deposits can be obtained. It is shown that this phenomenon arises from a catalytic electrochemical displacement reaction made possible by the underpotential deposition of depolarizer adatoms on gold and specific adsorption of depolarizer ions in the double-layer.

Photovoltage Studies of Clean and Oxygen Covered Gallium Arsenide. S. C. Dahlberg, Surface Sci, *59* (1976), pp. 83–96. The photovoltage of GaAs was measured by the retarding potential electron beam technique. The photovoltage light-intensity dependence deviates from steady state models. The photovoltage shows additional structure near the bandgap and in the below bandgap region; it is lower for uncleaned surfaces. Work function shows that oxygen adsorption is second order.

Photovoltage Studies of n-Type InP (100). S. C. Dahlberg, Surf. Sci., *60* November (1976), pp. 231–238. The photovoltage of InP (100) decreases as the intensity and/or duration of the light exposure decreases. The photovoltage spectra has been studied as a function of temperature and it shows considerable structure above the bandgap energy. Both the photovoltage and work function decrease sharply after Ar bombardment, probably due to preferential sputtering of P.

Polycrystalline Thin Film InP/CdS Solar Cells. K. J. Bachmann, J. L. Shay, S. Wagner, and E. Buehler, Appl. Phys. Lett., *29*, No. 2 (July 15, 1976), pp. 121–123. Polycrystalline films of p-type InP were prepared via chemical vapor deposition on glass, molybdenum, and carbon substrates. Cd or Zn were used as dopants yielding net acceptor concentrations $N_A - N_D \leq 2 \times 10^{18}\,\mathrm{cm}^{-3}$. Polycrystalline InP/CdS solar cells fabricated by evaporating CdS onto such InP films on carbon have a solar power conversion efficiency of 3 percent.

ELECTRICAL AND ELECTRONIC ENGINEERING

Spin Flip Raman Laser at Wavelengths Up to 16.8 μm. C. K. N. Patel, T. Y. Chang, and V. T. Nguyen, Appl. Phys. Lett., *28*, No. 10 (May 15, 1976), pp. 603–605. An InSb spin flip raman (SFR) laser is pumped with an optically pumped NH_3 laser line at 780.515 cm^{-1} to obtain tunable first Stokes laser radiation at wavelengths up to 16.8 μm. We report results on the power output, tunability, and preliminary spectroscopy of UF_6.

MATERIALS SCIENCE

Low-Field Depoling Characteristics of $Pb(Zr,Ti)O_3$ Ceramics. J. B. Koeneman, *59*, No. 9–10 (September–October 1976), pp. 59–61. Low-electric-field depoling measurements were made on $Pb(Zr,Ti)O_3$ ceramics. The apparent activation energy for the depoling process was calculated to be 33 kcal/mol. The data indicate that the materials tested should be very stable when subjected to depoling fields on the order of one-fourth of the coercive field.

Oxidation Induced Stacking Faults in n- and p-Type (100) Silicon. S. P. Murarka and G. Quintana, J. Appl. Phys., *48* (January 1977), pp. 46–51. The formation of stacking faults during thermal oxidation of silicon has been investigated. The length and the density of stacking faults, in both n- and p-type 2-inch-diameter (100) silicon wafers obtained from various manufacturers, were determined as a function of time and temperature of oxidation in dry and steam ambients. Two categories of stacking faults were established.

A Simple Titanium and Nickel Sublimation Pump (TNSP). C. A. Haque, *13*, No. 5 (September/October 1976), pp. 1088–1090. One filament of a commercially available Titanium Sublimation Pump (TSP) was replaced with a nickel filament on a tungsten support. This Titanium and Nickel Sublimation Pump (TNSP) besides having the attributes of a TSP, effectively minimizes the hydrogen contamination problem omnipresent in most stainless steel ultrahigh vacuum chambers, with nickel acting as a getter for hydrogen.

Sources of Oxidation Induced Stacking Faults in Czochralski Silicon Wafers: A One-to-One Correlation With Native Defects. G. A. Rozgonyi, S. Mahajan, M. H. Read, and D. Brasen, Appl. Phys. Lett., *29* (November 1, 1976), pp. 531–533. Using optical microscopy/etch pit techniques for the delineation of defects in [100] Czochralski silicon wafers, we have made a one-to-one correlation between bulk stacking faults in oxidized wafers and etch hillocks identified at the same sites before oxidation. Transmission electron microscopy of the hillock defects shows them to be clusters of precipitates ranging in size from 0.01 μm to 0.3 μm.

MATHEMATICS AND STATISTICS

On the Set of Distances Determined by the Union of Arithmetic Progressions. F. R. K. Chung and R. L. Graham, Ars Combinatoria., *1* (1976), pp. 57–76. Suppose a set of real numbers $A = \{a_1 < a_2 < \cdots < a_t\}$ is a union of $n \geq 2$ arithmetic progressions, each with common difference 1. Let $\Delta(A)$ denote $\{a_{k+1} - a_k : 1 \leq k < t\}$. It is shown that $|\Delta(A)|$, the number of elements of $\Delta(A)$, satisfies $|\Delta(A)| \leq 3n - 3$ and that this inequality is best possible. A similar result with $3n - 3$ replaced by $3n$ holds when A lies on a circle.

Simulation and Extension of a Minimum Mean Squared Error Estimator in Comparison With Stein's. H. D. Vinod, Technometrics *18* (November 1976), pp. 491–496. We discuss a fixed point solution of the iterative process underlying Farebrother's minimum mean squared error (MSE) estimator. A simulation study favors Stein's

shrunken least squares estimator and our fixed point solution over ordinary least squares. Farebrother's estimator almost never has the minimum MSE.

MECHANICAL AND CIVIL ENGINEERING

Optimal Aseismic Design of Building and Equipment. S. C. Liu, M. R. Dougherty, F. Neghabat, J. Eng. Mech. Div., *EM3*, (June 1976), pp. 395–414. A method for earthquake-resistant design was developed and applied to the problem of protecting industrial equipment in multistory buildings. In contrast with deterministic methods of analysis, attention is focused on the random properties of the environment and structural response. This approach has the advantage of including earthquake damage cost, protection cost, and reliability with conventional design factors. The results disclose cost effective and optimal designs.

PHYSICS

The A. C. Stark Shift for High Light Intensities. P. F. Liao and J. E. Bjorkholm, Bull. Amer. Phys. Soc., *36*, No. 26 (June 28, 1976), pp. 1543–1545. We report measurements of optically induced energy-level shifts produced by nonresonant light in sodium vapor. Intensities are sufficiently high such that the shifts depart substantially from the linear behavior predicted by second-order perturbation theory.

Determination of Nonradiative Decay Rate in Electron-Hole Drops in Ge at 1.6°K. R. F. Leheny, J. Shah, and M. Voos,* Solid State Commun. *20* (November 1976), pp. 819–821. We show that microwave photoconductivity measurements of optically excited carriers in Ge at 1.6°K can be used to determine the importance of nonradiative recombination within electron-hole liquid drops. Our results show that the nonradiative lifetime is 80 μs from which we calculate a radiative efficiency of 0.5 ± 0.1 for the condensed phase. *Groupe de Physique des Solides, De L'Ecole Normale Superieure, Paris, France.

Grain-Boundary Electromigration in Thin Films. I. Low-Temprature Theory. K. L. Tai and M. Ohring,* J. Appl. Phys., *48*, No. 1 (January 1977), pp. 28–35. A macroscopic mathematical theory is presented accounting for grain-boundary diffusion and electromigration in the presence of a simultaneous flux of atoms into the surrounding bulk lattice. The model employs a semi-infinite bicrystal geometry with a constant source at the origin, and both integral and numerical solutions to the subsequent non-steady-state transport equations are given. A comparison between the present theory and a previous treatment based on an extension of the Fisher analysis will be made. Application to recent results in thin films will be discussed. *Department of Metallurgy, Stevens Institute of Technology.

Grain-Boundary Electromigration in Thin Films. II. Tracer Measurements in Pure Au. K. L. Tai and M. Ohring,* J. Appl. Phys., *48*, No. 1 (January 1977), pp. 36–45. The first *direct* measurement of the grain-boundary ion drift velocity in thin Au films over the temperature range 120–250°C is reported. Central regions on the narrow stripe conductors were selectively embedded with ^{195}Au tracer atoms and the extent of the subsequent transport were evaluated by a high-resolution autoradiography technique employing the scanning electron microscope. *Department of Metallurgy, Stevens Institute of Technology.

The Nucleation and Growth of an Epitaxial Monolayer of Cd on Ge(111): A Simultaneous Rheed-MB Study. K. J. Matysik, J. Appl. Phys., *47* (October 1976), pp. 4359–4363. The correlation between structures and mass spectrometric molecular beam spectra is presented. The spectra are discussed in terms of the atomistic rate theory of nucleation. The results suggest that a convention of nucleation rate theory, that growing clusters act as monomer capture centers which increase in capture efficiency with time, does not hold at submonolayer coverages.

A Photometric Ellipsometer for Measuring Flux in a General State of Polarization. D. E. Aspnes, Surface Sci. *56* (June 1976), pp. 161–9. The theory and operation of a rotating analyzer/compensator ellipsometer capable of measuring all four Stokes parameters of generally polarized flux is described. Applications presented include an investigation of the effects of internally and externally stray and scattered light and surface roughness on measured values of the dielectric function.

Photovoltaic Properties and Junction Formation in $CuInSe_2$. B. Tell and P. M. Bridenbaugh, J. Appl. Phys. *47*, No. 2 (February 1976), pp. 619–620. Studies of diffusion and photovoltaic effects in $CuInSe_2$ pn junctions are reported. Junctions were formed by annealing Zn, Cd, and Cu plated p-type samples at temperatures from 200–450°C. The most efficient junctions are formed by 5–10 minute anneals at 200°C with a calculated interdiffusion coefficient $\sim 5 \times 10^{-10}$ cm^2/s.

SYSTEMS ENGINEERING AND OPERATIONS RESEARCH

Sea Plow IV—An Underseas Vehicle for Burying Ocean Cable. G. S. Cobb, Oceans, September 1976, pp. 19B1–19B6. A newly developed SG cable system, schedule for TAT-6, will have 1½ times the channel capacity of all present transatlantic cables combined, meaning significantly higher cost of providing interim service in the event of a trawler break. To provide the cable protection required for TAT-6 both in terms of burial performance and operating-depth capability, SEA PLOW IV was designed and tested over a 18-month period beginning in January, 1974 and was successfully used to bury both shore ends of TAT-6 during August and October, 1975.