THE BELL SYSTEM TECHNICAL JOURNAL

Statistical Properties of a Sum of Sinusoids and Gaussian Noise and its Generalization to Higher Dimensions

By JOEL GOLDMAN

This paper investigates the statistical properties of the sum, S, of an n-dimensional Gaussian random vector, N, plus the sum of M vectors, $X_1, \cdots, X_M$, having random amplitudes and independent arbitrary orientations in n-dimensional space. We derive expressions for the probability density function (p.d.f.) and distribution function (d.f.) of S and of its length, $|S|$, as series expansions involving only the moments of $|X_i|$, $i = 1, \cdots, M$. In addition, we find the p.d.f. and d.f. of the projection of S onto 1-dimensional space. Our results are generalizations of the n = 2-dimensional problem of finding the statistical properties of a sum of constant-amplitude sinusoids having independent uniformly distributed phase angles plus Gaussian noise. The latter problem has been treated by Rice[1] and Esposito and Wilson,[2] but our results can also deal with sinusoids having random amplitudes. When n = 3, our findings treat, in the presence of a Gaussian vector, the classical problem of "random flights" dating back to Rayleigh. Some calculations for the 2- and 3-dimensional problem are presented, and an application to coherent phase-shift-keying communications systems is discussed.

I. INTRODUCTION

In a number of problems arising in communications systems, in multipath phenomena, and in other areas, the determination of the

statistical properties of a sum of sinusoids and Gaussian noise is important for evaluating system performance. For this reason there has been interest in this problem for a number of years. Rice[1] first investigated the statistical properties of the sum of a single constant-amplitude sinusoid and Gaussian noise. Later, Esposito and Wilson[2] considered this same problem but with two constant-amplitude sinusoids having independent uniformly distributed phase angles. More recently, Rice[3] studied the properties of a sum of M sinusoids and Gaussian noise. In this paper, we look at the natural generalization of this problem to n-dimensional space; namely, we determine the statistical properties of the sum of an n-dimensional Gaussian random vector plus the sum of M vectors having random amplitudes and independent arbitrary orientations in n-dimensional space. In the special case when $n = 2$, our results are applicable to the type of problems considered by Rice and Esposito and Wilson, but they can also deal with any number of sinusoids with *random* amplitudes. When $n = 3$, our findings treat, in the presence of a Gaussian vector, the classical problem of "random flights" dating back to Rayleigh.

In Section II we give a definition of spherically symmetric random n-vectors and state a theorem which characterizes the form of such vectors in an n-dimensional spherical coordinate system. We consider M independent spherically symmetric vectors, $\mathbf{X}_1, \cdots, \mathbf{X}_M$, and define $\mathbf{S} = \sum_{i=1}^{M} \mathbf{X}_i$. Using our characterization theorem, we show that the even moments, $E[|\mathbf{S}|^{2k}]$, $k = 1, 2, \cdots$, can be easily expressed in terms of only the moments of $|\mathbf{X}_i|$, $i = 1, \cdots, M$. Then with the normal vector $\mathbf{N} \sim \mathfrak{N} (0, \sigma^2 \mathbf{I})$ independent of the $\mathbf{X}_i$'s, we derive in Section III the probability density functions (p.d.f.'s) and distribution functions (d.f.'s) of $\mathbf{S} + \mathbf{N}$ and of $|\mathbf{S} + \mathbf{N}|$ as series expansions involving the moments of $|\mathbf{S}|$. In addition, we derive the p.d.f. and d.f. of the projection of $\mathbf{S} + \mathbf{N}$ onto 1-dimensional space in terms of a similar series expansion. When $n = 2$ and $M = 2$, we check that our results agree with those of Esposito and Wilson for two constant-amplitude sinusoids.

Last, in Section IV we present some calculations for the 2- and 3-dimensional problems, and discuss some aspects of the computational procedure that we use. Certain of these calculations provide results for the probability of error of a binary coherent phase-shift-keying communications system operating in the presence of M co-channel interferers and Gaussian noise. These results extend previously published computations.[4,5] Additionally, our findings can be used to find

the probability of error of this system operating in the presence of more general types of interference.

II. SPHERICAL SYMMETRY

The generalization of sinusoids with uniformly distributed phase angles are "spherically symmetric" vectors defined in the following way (see Refs. 6 and 7):

Definition: A random n-vector $\mathbf{X} = (X_1, \cdots, X_n)$, $n \geqq 1$, is *spherically symmetric with matrix* ϱ if and only if the covariance matrix of $\mathbf{X}$ exists,[*] $E(\mathbf{X}) = \mathbf{0}$, and the joint characteristic function of $\mathbf{X}$ is of the form:[†]

$$\Phi_{\mathbf{X}}(\mathbf{u}) = E[e^{j\mathbf{u}\mathbf{X}'}] = h[(\mathbf{u}\varrho\mathbf{u}')^{\frac{1}{2}}] \tag{1}$$

for some function h on $[0, \infty)$ and where ϱ is some $n \times n$ (symmetric) positive definite matrix.[‡] Actually, h and ϱ are defined only up to positive multiplicative factors. However, in this paper we are only concerned with spherically symmetric vectors with $\varrho = \mathbf{I} =$ identity matrix. Then h is uniquely determined and $\Phi_{\mathbf{X}}(\mathbf{u}) = h(|\mathbf{u}|)$. We denote such a spherically symmetric vector by the notation "$\mathbf{X}$ is s.s."

Note that if $\mathbf{X}_1$ and $\mathbf{X}_2$ are two independent s.s. vectors, then clearly $\mathbf{X}_1 + \mathbf{X}_2$ is also s.s.

The probability density function of an s.s. vector $\mathbf{X}$ can be found by Bochner's theorem.[8] If $h(|\mathbf{u}|)$ is absolutely integrable, then the p.d.f. of $\mathbf{X}$ is:

$$p_{\mathbf{X}}(\mathbf{x}) = g_n(|\mathbf{x}|), \tag{2}$$

where

$$g_n(r) = \frac{1}{(2\pi)^{n/2}} \frac{1}{r^{(n-2)/2}} \int_0^\infty h(\lambda)\lambda^{n/2} J_{(n-2)/2}(\lambda r) d\lambda \quad r > 0, \quad n \geqq 1.$$

Thus, if $\mathbf{X}$ is s.s., its p.d.f. is constant over every n-dimensional sphere centered about the origin. This vector is precisely what is meant by a "random flight" in a higher dimensional space.

For our purposes, a more suitable characterization of an s.s. vector is given by the following theorem proved in Ref. 9.

---

[*] Expected value will be denoted by $E(\cdot)$.
[†] We denote vectors by boldface characters: $\mathbf{u} = (u_1, \cdots, u_n)$. The character $\mathbf{u}'$ is the transpose of $\mathbf{u}$. The norm of $\mathbf{u}$ is denoted $|\mathbf{u}| = (\mathbf{u}\mathbf{u}')^{\frac{1}{2}}$.
[‡] For $n = 1$, a spherically symmetric random variable has an even characteristic function, $\Phi_{\mathbf{X}}(u) = h[\rho^{\frac{1}{2}}|u|]$.

Theorem 1: Suppose $\mathbf{X} = (X_1, \cdots, X_n)$, $n \geq 2$, *is s.s. Then there exists a unique set of random variables* $R \geq 0$, $\Phi_k \in [0, \pi]$, $k = 1, \cdots, n-2$, $\theta \in [0, 2\pi]$ *for which**

$$X_j = R \left(\prod_{k=1}^{j-1} \sin \Phi_k \right) \cos \Phi_j \qquad 1 \leq j \leq n - 2$$

$$X_{n-1} = R \left(\prod_{k=1}^{n-2} \sin \Phi_k \right) \cos \theta \tag{3}$$

$$X_n = R \left(\prod_{k=1}^{n-2} \sin \Phi_k \right) \sin \theta,$$

and furthermore $(R, \Phi_1, \cdots, \Phi_{n-2}, \theta)$ *are independent and have respective p.d.f.'s:*

$$p_R(r) = 2\pi^{n/2} \left[\Gamma \left(\frac{n}{2} \right) \right]^{-1} r^{n-1} g_n(r) \qquad r \geq 0$$

$$p_{\Phi_k}(\phi_k) = \Gamma \left(\frac{n - k + 1}{2} \right) \pi^{-\frac{1}{2}} \left[\Gamma \left(\frac{n - k}{2} \right) \right]^{-1} \sin^{n-1-k}\phi_k$$

$$0 \leq \phi_k \leq \pi \tag{4}$$

$$k = 1, \cdots, n - 2$$

$$p_\theta(\theta) = \frac{1}{2\pi} \qquad 0 \leq \theta < 2\pi$$

for the $g_n(\cdot)$ *of* (2).

Conversely, *if* $(R, \Phi_1, \cdots, \Phi_{n-2}, \theta)$ *are independent and have the p.d.f.'s given by* (4), *and* $\mathbf{X}$ *is defined as in* (3), *then* $\mathbf{X}$ *is s.s.*

The utility of this theorem lies in the fact that the random variables $(R, \Phi_1, \cdots, \Phi_{n-2}, \theta)$ are independent with specified p.d.f.'s. As an immediate corollary, we see from (2), (3), and (4) that:

Corollary 1: Suppose $\mathbf{X} = (X_1, \cdots, X_n)$, $n \geq 1$, *is s.s. Then its p.d.f. is given by:*

$$p_{\mathbf{X}}(\mathbf{x}) = (2\pi^{n/2})^{-1} \Gamma(n/2) \frac{p_{|\mathbf{x}|}(|\mathbf{x}|)}{(|\mathbf{x}|)^{n-1}}.$$

Moreover, for $j = 1, \cdots, n$ *and for all* i,

$$\frac{\Gamma(n/2) E[|\mathbf{X}|^{2i}]}{\Gamma[(n/2) + i]} = \frac{\Gamma(\frac{1}{2}) E[|X_j|^{2i}]}{\Gamma(\frac{1}{2} + i)}.$$

Using Theorem 1 we can prove:

Theorem 2: Suppose $\mathbf{X}_1, \cdots, \mathbf{X}_M$ *are independent s.s. n-vectors,* $n \geq 1$. *Let* $\mathbf{S}_j = \sum_{i=1}^{j} \mathbf{X}_i$, $j = 1, \cdots, M$, *and define*

---

* We define $\prod_{k=1}^{0} a_k = 1$.

$$\mu_j^{(2k)} = E[|\mathbf{S}_j|^{2k}], \qquad k = 0, 1, \cdots, \qquad j = 1, \cdots, M$$
$$\nu_\ell^{(2m)} = E[|\mathbf{X}_\ell|^{2m}], \qquad m = 0, 1, \cdots, \qquad \ell = 1, \cdots, M.$$

Put

$$c_{n,2m} \triangleq \frac{B[(2m+1)/2, (n-1)/2]}{B[\frac{1}{2}, (n-1)/2]},$$

where $B(\cdot, \cdot)$ *is the beta function. Denote*

$$\mathbf{u}_j^{(2m)} = (\mu_j^{(0)}, \mu_j^{(2)}, \cdots, \mu_j^{(2m)})$$
$$\mathbf{v}_j^{(2m)} = (\nu_j^{(0)}, \nu_j^{(2)}, \cdots, \nu_j^{(2m)})$$

and define $D_{n,j}$ *to be an* $(m+1) \times (m+1)$ *matrix whose* (k, ℓ)*th element equals*

$$\binom{2\ell - 2}{2k - 2} \frac{c_{n,2k-2}\, c_{n,2\ell-2k}}{c_{n,2\ell-2}} \nu_j^{(2\ell-2k)}, \qquad if \quad \ell \geqq k$$

and is 0 if $\ell < k$.

Then, for $j = 2, \cdots, M$, *and* $m = 0, 1, \cdots$,

$$\mu_j^{(2m)} = \sum_{i=0}^{m} \binom{2m}{2i} \frac{c_{n,2i}\, c_{n,2m-2i}}{c_{n,2m}} \mu_{j-1}^{(2i)} \nu_j^{(2m-2i)}. \tag{5}$$

In matrix form this is

$$\mathbf{u}_j^{(2m)} = \mathbf{u}_{j-1}^{(2m)} D_{n,j},$$

so that

$$\mathbf{u}_j^{(2m)} = \mathbf{v}_1^{(2m)} D_{n,2} \cdots D_{n,j-1} D_{n,j} \tag{6}$$

for $j = 2, \cdots, M$.

Proof: By Theorem 1 we have for each $\mathbf{X}_i$ a corresponding vector in spherical coordinate space:

$$\mathbf{X}_i \leftrightarrow (R_i, \Phi_{1,i}, \cdots, \Phi_{n-2,i}, \theta_i).$$

Since $\mathbf{S}_j$ is a sum of independent s.s. vectors, it is also s.s., so there are vectors corresponding to it:

$$\mathbf{S}_j \leftrightarrow (P_j, \xi_{1,j}, \cdots, \xi_{n-2,j}, \Psi_j).$$

Note that $\mu_j^{(2k)} = E[P_j^{2k}]$ and $\nu_\ell^{(2m)} = E[R_\ell^{2m}]$ and that

$$E[\cos^i \Phi_{1,j}] = E[\cos^i \xi_{1,j}]$$

$$= \frac{\Gamma(n/2)}{\Gamma(\frac{1}{2})\Gamma[\frac{1}{2}(n-1)]} \int_0^\pi \cos^i \alpha \, \sin^{n-2} \alpha \, d\alpha$$

$$= \begin{cases} c_{n,2i} & \text{if } i \text{ is even} \\ 0 & \text{if } i \text{ is odd.} \end{cases} \tag{7}$$

Denote the components of $\mathbf{S}_j$ and $\mathbf{X}_j$ as follows:

$$\mathbf{S}_j = (S_{1,j}, \cdots, S_{n,j})$$
$$\mathbf{X}_j = (X_{1,j}, \cdots, X_{n,j}),$$

so that

$$S_{1,j} = P_j \cos \xi_{1,j}$$
$$X_{1,j} = R_j \cos \Phi_{1,j}.$$

Also,

$$E[S_{1,j}^i] = E[P_j^i \cos^i \xi_{1,j}]$$
$$= E[P_j^i]E[\cos^i \xi_{1,j}] = 0 \tag{8}$$

if i is odd, and

$$E[S_{1,j}^{2m}] = E[P_j^{2m}]E[\cos^{2m} \xi_{1,j}]$$
$$= \mu_j^{(2m)} c_{n,2m} \tag{9}$$

and

$$E[X_{1,j}^{2k}] = E[R_j^{2k}]E[\cos^{2k}\Phi_{1,j}] = \nu_j^{(2k)} c_{n,2k}. \tag{10}$$

Noting that $S_{1,j-1}$ is independent of $X_{1,j}$ since $\mathbf{S}_{j-1}$ depends only on $\mathbf{X}_1, \cdots, \mathbf{X}_{j-1}$ which are independent of $\mathbf{X}_j$, the following equalities follow from (8) to (10):

$$\mu_j^{(2m)} c_{n,2m} = E[S_{1,j}^{2m}]$$
$$= E\{[S_{1,j-1} + X_{1,j}]^{2m}\}$$
$$= \sum_{i=0}^{2m} \binom{2m}{i} E\{[S_{1,j-1}]^i [X_{1,j}]^{2m-i}\}$$
$$= \sum_{i=0}^{2m} \binom{2m}{i} E[S_{1,j-1}^i]E[X_{1,j}^{2m-i}]$$
$$= \sum_{i=0}^{m} \binom{2m}{2i} E[S_{1,j-1}^{2i}]E[X_{1,j}^{2m-2i}]$$
$$= \sum_{i=0}^{m} \binom{2m}{2i} \mu_{j-1}^{(2i)} c_{n,2i} \nu_j^{(2m-2i)} c_{n,2m-2i}.$$

Hence,

$$\mu_j^{(2m)} = \sum_{i=0}^{m} \binom{2m}{2i} \frac{c_{n,2i} c_{n,2m-2i}}{c_{n,2m}} \mu_{j-1}^{(2i)} \nu_j^{(2m-2i)}.$$

The vector equation

$$\mathbf{u}_j^{(2m)} = \mathbf{u}_{j-1}^{(2m)} D_{n,j} \qquad j = 2, \cdots, M \tag{11}$$

follows immediately from this expression. Since $\mathbf{u}_1^{(2m)} = \mathbf{v}_1^{(2m)}$, eq. (11) implies that $\mathbf{u}_j^{(2m)} = \mathbf{v}_1^{(2m)} D_{n,2} \cdots D_{n,j-1} D_{n,j}$ for $j = 2, \cdots, M$. For $n \geqq 1$, we prove (5) directly.

In Reference 9 we proved the result:

Theorem 3: If $\mathbf{X}$ *is s.s. and independent of* $\mathbf{N} \sim \mathfrak{N}(\mathbf{0}, \sigma^2 \mathbf{I})$*, then the p.d.f. of* $\mathbf{X} + \mathbf{N}$ *is:*

$$p_{\mathbf{X}+\mathbf{N}}(\mathbf{z}) = \frac{1}{2\pi^{n/2}\sigma^2} \Gamma\left(\frac{n}{2}\right) \int_0^\infty dv \, p_{|\mathbf{X}|}(v) \, \frac{1}{[v|\mathbf{z}|]^{(n-2)/2}}$$

$$\times \exp\left[-\frac{1}{2\sigma^2}(v^2 + |\mathbf{z}|^2)\right] I_{(n-2)/2}\left(\frac{v|\mathbf{z}|}{\sigma^2}\right), n \geq 1, \quad (12)$$

where $I_\nu(\cdot)$ *is the* ν*th order modified Bessel function of the first kind.*

From this theorem we obtain the following:

Corollary 2: Suppose $\mathbf{X}$ *is s.s, independent of* $\mathbf{N} \sim \mathfrak{N}(\mathbf{0}, \sigma^2\mathbf{I})$*, and* $\sum_{i=0}^\infty [(1/2\sigma^2)^i/i!]E[|\mathbf{X}|^{2i}] < \infty$ *(for example, if* $|\mathbf{X}|$ *is a bounded random variable). Let* Z *be the projection of* $\mathbf{X} + \mathbf{N}$ *onto 1-dimensional space, i.e.,* Z *is (say) the first component of the vector* $\mathbf{X} + \mathbf{N}$*. Then for* $n \geq 1$ *the p.d.f.'s of* $\mathbf{X} + \mathbf{N}$*, of* $|\mathbf{X} + \mathbf{N}|$*, and of* Z *are given, respectively, by:*

$$p_{\mathbf{X}+\mathbf{N}}(\mathbf{z}) = \frac{\Gamma(n/2)}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}|\mathbf{z}|^2\right)$$

$$\times \sum_{i=0}^\infty \frac{L_i^{[(n-2)/2]}(|\mathbf{z}|^2/2\sigma^2)(-1/2\sigma^2)^i E[|\mathbf{X}|^{2i}]}{\Gamma[(n/2)+i]}, \quad (13)$$

$$p_{|\mathbf{X}+\mathbf{N}|}(v) = \frac{2}{(2\sigma^2)^{n/2}} v^{n-1} \exp\left(-\frac{1}{2\sigma^2}v^2\right)$$

$$\times \sum_{i=0}^\infty \frac{L_i^{[(n-2)/2]}(v^2/2\sigma^2)(-1/2\sigma^2)^i E[|\mathbf{X}|^{2i}]}{\Gamma[(n/2)+i]}, \quad (14)$$

and

$$p_Z(z) = \frac{\Gamma(n/2)}{(2\pi\sigma^2)^{\frac{1}{2}}} \exp\left(-\frac{z^2}{2\sigma^2}\right)$$

$$\times \sum_{i=0}^\infty \frac{L_i^{(-\frac{1}{2})}(z^2/2\sigma^2)(-1/2\sigma^2)^i E[|\mathbf{X}|^{2i}]}{\Gamma[(n/2)+i]}, \quad (15)$$

where $L_i^{(\alpha)}(\cdot)$ *are the generalized Laguerre polynomials. In addition, for* $n \geq 1$ *the "distribution functions" of* $|\mathbf{X} + \mathbf{N}|$ *and of* Z *are given, respectively, by:*

$$\Pr\{|\mathbf{X} + \mathbf{N}| > a\} = \frac{1}{\Gamma(n/2)} \Gamma\left(\frac{n}{2}, \frac{a^2}{2\sigma^2}\right) - \left(\frac{a^2}{2\sigma^2}\right)^{n/2} \exp\left(-\frac{a^2}{2\sigma^2}\right)$$

$$\times \sum_{i=1}^\infty \frac{L_{i-1}^{(n/2)}(a^2/2\sigma^2)(-1/2\sigma^2)^i E[|\mathbf{X}|^{2i}]}{i\Gamma[(n/2)+i]} \quad (16)$$

and

$$\Pr\{Z > a\} = \tfrac{1}{2}\,\mathrm{erfc}\left(\frac{a}{(2\sigma^2)^{\frac{1}{2}}}\right) - \frac{a}{2}\frac{\Gamma(n/2)}{(2\pi\sigma^2)^{\frac{1}{2}}}\exp\left(-\frac{a^2}{2\sigma^2}\right)$$

$$\times \sum_{i=1}^{\infty}\frac{L_{i-1}^{(\frac{1}{2})}(a^2/2\sigma^2)(-1/2\sigma^2)^{i}E[|\mathbf{X}|^{2i}]}{i\Gamma[(n/2)+i]}, \quad (17)$$

where $\Gamma(\cdot,\cdot)$ *is the incomplete gamma function (Ref. 10, p. 337) and* *erfc* $(\cdot)$ *is the complementary error function.*

Proof: From Ref. 10, p. 242, we have the generating function

$$e^{-t^2}\frac{1}{(tx)^{\alpha}}I_{\alpha}(2tx) = \sum_{i=0}^{\infty}\frac{L_i^{(\alpha)}(x^2)(-1)^{i}t^{2i}}{\Gamma(\alpha+i+1)}. \quad (18)$$

With $t = v/\sqrt{2\sigma^2}$, $x = |\mathbf{z}|/\sqrt{2\sigma^2}$, and $\alpha = (n-2)/2$, we substitute (18) into (12) to get:

$$p_{\mathbf{X+N}}(\mathbf{z}) = \frac{1}{2\pi^{n/2}\sigma^2}\,\Gamma\left(\frac{n}{2}\right)\exp\left[-\frac{1}{2\sigma^2}|\mathbf{z}|^2\right]\int_0^{\infty}dv\,p_{|\mathbf{X}|}(v)$$

$$\times \frac{1}{(2\sigma^2)^{(n-2)/2}}\sum_{i=0}^{\infty}\frac{L_i^{[(n-2)/2]}(|\mathbf{z}|^2/2\sigma^2)(-1/2\sigma^2)^{i}v^{2i}}{\Gamma[(n/2)+i]}$$

$$= \frac{\Gamma(n/2)}{(2\pi\sigma^2)^{n/2}}\exp\left(-\frac{1}{2\sigma^2}|\mathbf{z}|^2\right)$$

$$\times \sum_{i=0}^{\infty}\frac{L_i^{[(n-2)/2]}(|\mathbf{z}|^2/2\sigma^2)(-1/2\sigma^2)^{i}E[|\mathbf{X}|^{2i}]}{\Gamma[(n/2)+i]}, \quad (19)$$

assuming that the interchange of integration and expectation is valid. The second assertion of the corollary follows from Corollary 1 since

$$p_{|\mathbf{X+N}|}(|\mathbf{z}|) = \frac{2\pi^{n/2}}{\Gamma(n/2)}|\mathbf{z}|^{n-1}p_{\mathbf{X+N}}(\mathbf{z}). \quad (20)$$

To prove (15) we note that $Z = X_1 + N_1$ is a 1-dimensional s.s. random vector and we apply eq. (13) (with $n = 1$) and the second part of Corollary 1 to obtain the desired result.

Next, to show (16) we integrate (14) over the interval (a, ∞) and utilize the relationships:

$$\int_a^{\infty}v^{n-1}\exp\left(-\frac{v^2}{2\sigma^2}\right)dv = \frac{1}{2}(2\sigma^2)^{n/2}\Gamma\left(\frac{n}{2},\frac{a^2}{2\sigma^2}\right) \quad (21)$$

and for $i \geqq 1$,

$$\int_a^\infty v^{n-1} \exp\left(-\frac{v^2}{2\sigma^2}\right) L_i^{[(n-2)/2]}\left(\frac{v^2}{2\sigma^2}\right) dv$$
$$= -\frac{1}{2} a^n \exp\left(-\frac{a^2}{2\sigma^2}\right) \frac{L_{i-1}^{(n/2)}(a^2/2\sigma^2)}{i}. \quad (22)$$

Equation (21) is given in Ref. 10, p. 337, and (22) is proved in the appendix.

Finally, to obtain (17) we integrate (15) and use eq. (22) with $n = 1$ and the definition:

$$\text{erfc } (x) = \frac{2}{\pi^{\frac{1}{2}}} \int_x^\infty \exp (-t^2) dt.$$

It remains to justify the various interchanges of integration and expectation or summation. For example, to validate the interchange in (19) it suffices (Ref. 11, pp. 28–29) to show that

$$\sum_{i=0}^\infty \frac{(1/2\sigma^2)^i}{\Gamma[(n/2)+i]} E\left[|\mathbf{X}|^{2i}\right] \left| L_i^{[(n-2)/2]}\left(\frac{|\mathbf{z}|^2}{2\sigma^2}\right)\right| < \infty. \quad (23)$$

Since (Ref. 12, p. 207) $|L_i^{(\alpha)}(y)| \leqq e^{y/2}\Gamma(\alpha + i + 1)/i!\Gamma(\alpha + 1)$, the expression in (23) is less than or equal to:

$$\sum_{i=0}^\infty \frac{(1/2\sigma^2)^i}{\Gamma[(n/2)+i]} E\left[|\mathbf{X}|^{2i}\right] \exp\left(\frac{|\mathbf{z}|^2}{4\sigma^2}\right) \frac{1}{i!} \frac{\Gamma[(n/2)+i]}{\Gamma(n/2)}$$

which is finite by hypothesis.

The utility of this corollary lies in the fact that we can evaluate the various p.d.f.'s and d.f.'s knowing only the moments of $|\mathbf{X}|$ and not the entire distribution of $\mathbf{X}$.

III. STATISTICAL PROPERTIES OF THE SUM OF INDEPENDENT SPHERICALLY SYMMETRIC VECTORS AND GAUSSIAN NOISE

For simplicity, we combine the results of Corollary 2 and Theorem 2 into:

Theorem 4: Suppose $\mathbf{X}_1, \cdots, \mathbf{X}_M$ *are independent s.s. n-vectors,* $n \geqq 1$, *with moments* $\nu_\ell^{(2m)} = E[|\mathbf{X}_\ell|^{2m}]$, $m = 0, 1, \cdots,$ *and* $\ell = 1, \cdots, M$, *which are also independent of* $\mathbf{N} \sim \mathfrak{N}(\mathbf{0}, \sigma^2 \mathbf{I})$. *Let* $\mathbf{S} = \sum_{i=1}^M \mathbf{X}_i$ *and assume that* $\sum_{i=0}^\infty (1/2\sigma^2)^i/E[|\mathbf{S}|^{2i}]i! < \infty$ *(for example, if the* $|\mathbf{X}_i|$*'s are bounded random variables). Let* Z *be the projection of* $\mathbf{S} + \mathbf{N}$ *onto 1-dimensional*

space. Then the following relations for p.d.f.'s and d.f.'s are valid:

$$p_{S+N}(z) = \frac{\Gamma(n/2)}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}|z|^2\right)$$
$$\times \sum_{i=0}^{\infty} \frac{L_i^{[(n-2)/2]}(|z|^2/2\sigma^2)(-1/2\sigma^2)^i\mu_M^{(2i)}}{\Gamma[(n/2)+i]}, \quad (24)$$

$$p_{|S+N|}(v) = \frac{2}{(2\sigma^2)^{n/2}} v^{n-1} \exp\left(-\frac{1}{2\sigma^2}v^2\right)$$
$$\times \sum_{i=0}^{\infty} \frac{L_i^{[(n-2)/2]}(v^2/2\sigma^2)(-1/2\sigma^2)^i\mu_M^{(2i)}}{\Gamma[(n/2)+i]}, \quad (25)$$

$$p_Z(z) = \frac{\Gamma(n/2)}{(2\pi\sigma^2)^{\frac{1}{2}}} \exp\left(-\frac{z^2}{2\sigma^2}\right)$$
$$\times \sum_{i=0}^{\infty} \frac{L_i^{(-\frac{1}{2})}(z^2/2\sigma^2)(-1/2\sigma^2)^i\mu_M^{(2i)}}{\Gamma[(n/2)+i]}, \quad (26)$$

$$\Pr\{|S+N|>a\} = \frac{1}{\Gamma(n/2)} \Gamma\left(\frac{n}{2}, \frac{a^2}{2\sigma^2}\right) - \left(\frac{a^2}{2\sigma^2}\right)^{n/2}$$
$$\times \exp\left(-\frac{a^2}{2\sigma^2}\right) \sum_{i=1}^{\infty} \frac{L_{i-1}^{(n/2)}(a^2/2\sigma^2)(-1/2\sigma^2)^i\mu_M^{(2i)}}{i\Gamma[(n/2)+i]}, \quad (27)$$

and

$$\Pr\{Z>a\} = \frac{1}{2}\operatorname{erfc}\left(\frac{a}{(2\sigma^2)^{\frac{1}{2}}}\right) - \frac{a}{2}\frac{\Gamma(n/2)}{(2\pi\sigma^2)^{\frac{1}{2}}}$$
$$\times \exp\left(-\frac{a^2}{2\sigma^2}\right) \sum_{i=1}^{\infty} \frac{L_{i-1}^{(\frac{1}{2})}(a^2/2\sigma^2)(-1/2\sigma^2)^i\mu_M^{(2i)}}{i\Gamma[(n/2)+i]}. \quad (28)$$

The moments $\mu_M^{(2i)} \triangleq E[|S|^{2i}]$ are determined by the recurrence relations

$$\mu_j^{(2i)} = \sum_{k=0}^{i} \binom{2i}{2k} \frac{c_{n,2k}\, c_{n,2i-2k}}{c_{n,2i}} \mu_{j-1}^{(2k)}\nu_j^{(2i-2k)} \quad (29)$$

for $j = 2, \cdots, M$ with

$$c_{n,2m} = \frac{B[(2m+1)/2,\ (n-1)/2]}{B[\frac{1}{2},\ (n-1)/2]},$$

or by the matrix equation (6).

We next look at some special cases.

A. *n = 2-dimensional space*

When $n = 2$, eqs. (24) through (29) reduce in an obvious manner.

The incomplete gamma function in (27) equals (Ref. 10, p. 339):

$$\Gamma\left(1, \frac{a^2}{2\sigma^2}\right) = \exp\left(-\frac{a^2}{2\sigma^2}\right). \tag{30}$$

The Laguerre polynomials $L_i^{(-\frac{1}{2})}$ and $L_i^{(\frac{1}{2})}$ in (26) and (28) can be expressed in terms of Hermite polynomials $H_i(\cdot)$ using the relations (Ref. 10, p. 240):

$$L_i^{(\frac{1}{2})}(x) = \frac{(-1)^i x^{-\frac{1}{2}} H_{2i+1}(x^{\frac{1}{2}})}{i!\, 2^{2i+1}} \tag{31a}$$

and

$$L_i^{(-\frac{1}{2})}(x) = \frac{(-1)^i H_{2i}(x^{\frac{1}{2}})}{i!\, 2^{2i}}. \tag{31b}$$

For example, eq. (28) can be rewritten as:

$$\Pr\{Z > a\} = \frac{1}{2}\,\mathrm{erfc}\left(\frac{a}{\sqrt{2\sigma^2}}\right) + \frac{1}{\pi^{\frac{1}{2}}}\exp\left(-\frac{a^2}{2\sigma^2}\right)$$
$$\times \sum_{i=1}^{\infty} \frac{H_{2i-1}(a/\sqrt{2\sigma^2})(1/2\sigma^2)^i \mu_M^{(2i)}}{(i!)^2 2^{2i}}. \tag{32}$$

We also check that

$$\binom{2i}{2k}\frac{c_{n,2k}\, c_{n,2i-2k}}{c_{n,2i}} = \binom{i}{k}^2$$

so that

$$\mu_2^{(2i)} = \sum_{k=0}^{i}\binom{i}{k}^2 \mu_1^{(2k)}\nu_2^{(2i-2k)}$$

$$= \sum_{k=0}^{i}\binom{i}{k}^2 \nu_1^{(2k)}\nu_2^{(2i-2k)},$$

$$\mu_3^{(2i)} = \sum_{\ell=0}^{i}\sum_{k=0}^{\ell}\binom{i}{\ell}^2\binom{\ell}{k}^2 \nu_1^{(2k)}\nu_2^{(2\ell-2k)}\nu_3^{(2i-2\ell)},$$

and so forth.

Consider the type of problem investigated by Rice[1,3] and Esposito and Wilson[2] in determining the p.d.f.'s of the envelope and instantaneous value of

$$z(t) = \sum_{k=1}^{M} A_k \cos(w_k t + \hat{\theta}_k) + n(t),$$

where each $A_k \geqq 0$ is independent of $\hat{\theta}_k$ and $\hat{\theta}_k$ is uniformly distributed on $[0, 2\pi)$. Assume that the pairs $\{(A_k, \hat{\theta}_k)\}$ are independent of each other and of $n(\cdot)$. Suppose $n(t)$ is the result of the passage of zero-mean white stationary Gaussian noise through a bandpass symmetrical

filter. Then $n(t)$ can be written as (Ref. 13, pp. 142–148):

$$n(t) = n_1(t) \cos w_o t - n_2(t) \sin w_o t,$$

where $n_1(t)$ and $n_2(t)$ are zero-mean independent stationary low-pass Gaussian processes with

$$\sigma^2 = E[n(t)]^2 = E[n_1(t)]^2 = E[n_2(t)]^2.$$

Let $\theta_k(t) = (w_k - w_o)t + \hat{\theta}_k$ and thus:

$$
\begin{aligned}
z(t) &= \sum_{k=1}^{M} A_k \cos [w_o t + \theta_k(t)] + n(t) \\
&= \sum_{k=1}^{M} A_k \cos [w_o t + \theta_k(t)] + n_1(t) \cos w_o t - n_2(t) \sin w_o t \\
&= \left[\sum_{k=1}^{M} A_k \cos \theta_k(t) + n_1(t) \right] \cos w_o t \\
&\qquad\qquad - \left[\sum_{k=1}^{M} A_k \sin \theta_k(t) + n_2(t) \right] \sin w_o t \\
&= A(t) \cos [w_o t + \gamma(t)].
\end{aligned}
$$

At any time t_o, let $\theta_k = \theta_k(t_o)$, $n_1 = n_1(t_o)$, and $n_2 = n_2(t_o)$. Put

$$\mathbf{X}_k = (A_k \cos \theta_k, A_k \sin \theta_k), \qquad k = 1, \cdots, M,$$

and

$$\mathbf{N} = (n_1, n_2).$$

Then $\sum_{k=1}^{M} \mathbf{X}_k + \mathbf{N}$ is s.s., so by Theorem 1 it has the form $(\Gamma \cos \Psi, \Gamma \sin \Psi)$, where $\Gamma \geq 0$ is independent of Ψ and Ψ is uniformly distributed on $[0, 2\pi)$. It follows that

$$
\begin{aligned}
z(t_o) &= \Gamma \cos \Psi \cos w_o t_o - \Gamma \sin \Psi \sin w_o t_o \\
&= \Gamma \cos (w_o t + \Psi);
\end{aligned}
$$

that is, $\Gamma = A(t_o)$ and $\Psi = \gamma(t_o)$. Hence, at any time t_o, $A(t_o)$ and $\gamma(t_o)$ are independent and $\gamma(t_o)$ is uniformly distributed on $[0, 2\pi)$. Moreover, the p.d.f. of the "envelope" $A(t_o)$ is the p.d.f. of $\Gamma = |\sum_{k=1}^{M} \mathbf{X}_k + \mathbf{N}|$ which can be determined from (25). Thus we can find the p.d.f. of the envelope of the sum of Gaussian noise plus *any number* of sinusoids with *random* amplitudes and independent uniformly distributed phase angles. The case considered by Esposito and Wilson[2] was that of $M = 2$, $A_1 = a = \text{constant}$ and $A_2 = b = \text{constant}$, in which case

$$
\begin{aligned}
\nu_1^{(2m)} &= E[|\mathbf{X}_1|^{2m}] = a^{2m}, \\
\nu_2^{(2m)} &= E[|\mathbf{X}_2|^{2m}] = b^{2m},
\end{aligned}
$$

and

$$\mu_2^{(2i)} = \sum_{k=0}^{i} \binom{i}{k}^2 a^{2k} b^{2i-2k}.$$

The envelope p.d.f. is then:

$$p_\Gamma(v) = \frac{v}{\sigma^2} \exp[-v^2/2\sigma^2]$$

$$\times \sum_{i=0}^{\infty} \frac{L_i(v^2/2\sigma^2)(-1/2\sigma^2)^i}{i!} \left[\sum_{k=0}^{i} \binom{i}{k}^2 a^{2k} b^{2i-2k} \right],$$

which agrees with the result in Esposito and Wilson [Ref. 2, eq. (12)]. This expression was obtained earlier by Goldman.[14]

To find the p.d.f. of $z(t)$ at some time instant t_o, note that:

$$z(t_o) = A(t_o) \cos[w_o t_o + \gamma(t_o)]$$
$$= \Gamma \cos(w_o t_o + \Psi).$$

Since Ψ is uniformly distributed on $[0, 2\pi)$ and the cosine function has period 2π, the p.d.f. of $\Gamma \cos(w_o t_o + \Psi)$ is the same as that of $\Gamma \cos \Psi$. Recall that $(\Gamma \cos \Psi, \Gamma \sin \Psi) = \sum_{k=1}^{M} \mathbf{X}_k + \mathbf{N}$. Thus, $\Gamma \cos \Psi$ is the first component of the 2-dimensional vector $\sum_{k=1}^{M} \mathbf{X}_k + \mathbf{N}$ and, from eq. (26), its p.d.f. is:

$$p_{z(t_o)}(z_1) = p_{\Gamma \cos \Psi}(z_1)$$

$$= \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} \exp\left(-\frac{1}{2\sigma^2} z_1^2\right) \sum_{i=0}^{\infty} \frac{(-1/2\sigma^2)^i \mu_M^{(2i)}}{i!}$$
$$\times L_i^{(-\frac{1}{2})} \left(\frac{z_1^2}{2\sigma^2}\right). \quad (33)$$

In Esposito and Wilson's example, this becomes

$$p_{z(t_o)}(z_1) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} \exp\left(-\frac{1}{2\sigma^2} z_1^2\right) \sum_{i=0}^{\infty} \frac{(-1/2\sigma^2)^i}{i!}$$
$$\times L_i^{(-\frac{1}{2})} \left(\frac{z_1^2}{2\sigma^2}\right) \left[\sum_{k=0}^{i} \binom{i}{k}^2 a^{2k} b^{2i-2k} \right],$$

which agrees with their eq. (29).

We also check that the d.f. in (27) is the same as that obtained in eq. (18) of Ref. 2, when we use the fact that (Ref. 10, p. 241):

$$x L_{i-1}^{(1)}(x) = i[L_{i-1}^{(0)}(x) - L_i^{(0)}(x)].$$

Finally, consider a binary coherent phase-shift-keying communications system operating in the presence of Gaussian noise and M co-channel interferers modeled by a sum of constant amplitude sinu-

soids with independent, uniformly distributed phase angles, $\theta_1, \cdots, \theta_M$. (Details of this model and system may be found in Refs. 4 and 5.) The probability of error in such a system is:[4,5]

$$\mathrm{Pe} = \mathrm{Pr} \left\{ \sum_{i=1}^{M} b_i \cos \theta_i + N_1 > a \right\},$$

where $N_1 \sim \mathfrak{N}(0, \sigma^2)$ and a is the amplitude of the transmitted (desired) signal. This probability of error is given by the expression in (32) and agrees with the result found in Refs. 4 and 5. However, eq. (32) can also be used to find the probability of error in this system for a more general class of co-channel interference consisting of a sum of uniformly phased sinusoids having also independent *random* amplitudes.

B. $n = 3$-dimensional space

When $n = 3$, eqs. (24) through (29) reduce in a straightforward way. Equations (24) to (26) and (28) can also be expressed in terms of Hermite polynomials by employing eq. (31). The incomplete gamma function in (27) can be written in terms of tabulated functions by use of the relations (Ref. 10, pp. 339–340):

$$\Gamma(c + 1, x) = c\Gamma(c, x) + x^c e^x$$

and

$$\Gamma(\tfrac{1}{2}, x) = \pi^{\frac{1}{2}} \operatorname{erf}(x^{\frac{1}{2}}),$$

where $\operatorname{erf}(\cdot)$ is the error function.

The recurrence relation for the moments becomes, for $M = 2$,

$$\mu_2^{(2i)} = \sum_{k=0}^{i} \binom{2i}{2k} \left(\frac{2k + 2}{2k + 1} \right) \left(\frac{2i - 2k + 2}{2i - 2k + 1} \right) \left(\frac{2i + 1}{2i + 2} \right) \nu_1^{(2k)} \nu_2^{(2i-2k)},$$

and so on for higher values of M.

IV. SOME COMPUTATIONS

The form of the expressions in (24) through (28) is quite similar, and so the computer programs used for their evaluation were only slight modifications of one basic (Fortran IV) program. Different values of n could also be treated easily. The basic program required computation of a sum of the form:

$$\sum_{i=0}^{\infty} \frac{L_i^{(\alpha)}(x)(-1/2\sigma^2)^i \mu_M^{(2i)}}{\Gamma[(n/2) + i]}, \tag{34}$$

where x is a variable.
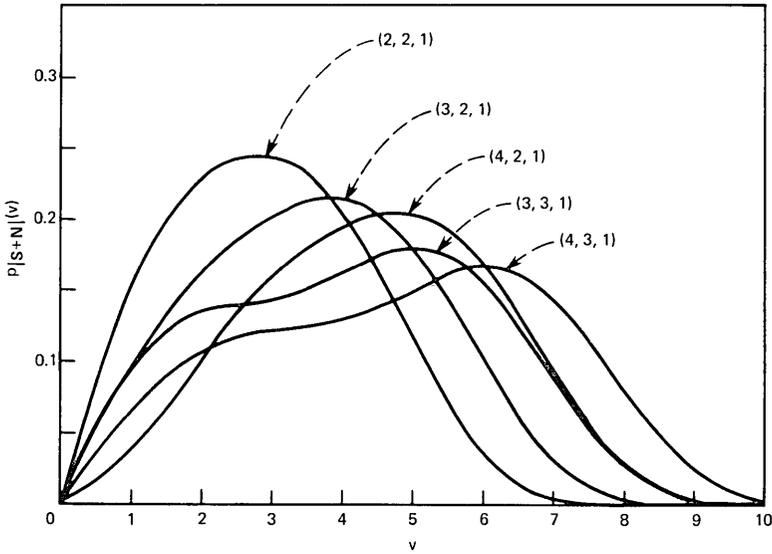
Fig. 1—Plots of p.d.f. of $|S + N|$ for $\sigma^2 = 1$, $M = 3$, $n = 2$, and different sets of vector lengths (b_1, b_2, b_3).

In one part of our program, the moments $\mu_M^{(2i)}$ were determined from eq. (29):

$$\mu_j^{(2i)} = \sum_{k=0}^{i} C_n(i, k)\mu_{j-1}^{(2k)}\nu_j^{(2i-2k)}, \qquad j = 2, \cdots, M, \tag{35}$$

where

$$C_n(i, k) \triangleq \binom{2i}{2k} \frac{c_{n,2k}\, c_{n,2i-2k}}{c_{n,2i}}. \tag{36}$$

Using the definition of $c_{n,2m}$ and properties of the beta function, we can show that the coefficients $C_n(i, k)$ are equal to:

$$C_n(i, k) = \frac{\Gamma(i + 1)\Gamma(n/2)\Gamma[(n/2) + i]}{\Gamma(k + 1)\Gamma(i - k + 1)\Gamma[(n/2) + k]\Gamma[(n/2) + i - k]}. \tag{37}$$

To efficiently compute these coefficients and to eliminate "overflow" problems, we utilized the simple recurrence relation

$$C_n(i, k) = \frac{(i - k + 1)[(n/2) + i - k]}{k[(n/2) + k - 1]} C_n(i, k - 1), \quad k \geqq 1, \tag{38}$$
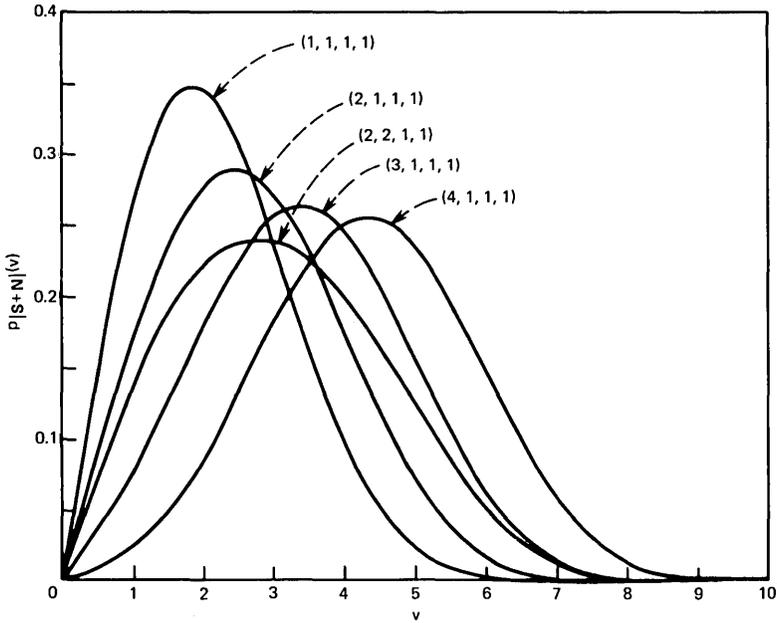
Fig. 2—Plots of p.d.f. of $|S + N|$ for $\sigma^2 = 1$, $M = 4$, $n = 2$, and different sets of vector lengths (b_1, b_2, b_3, b_4).
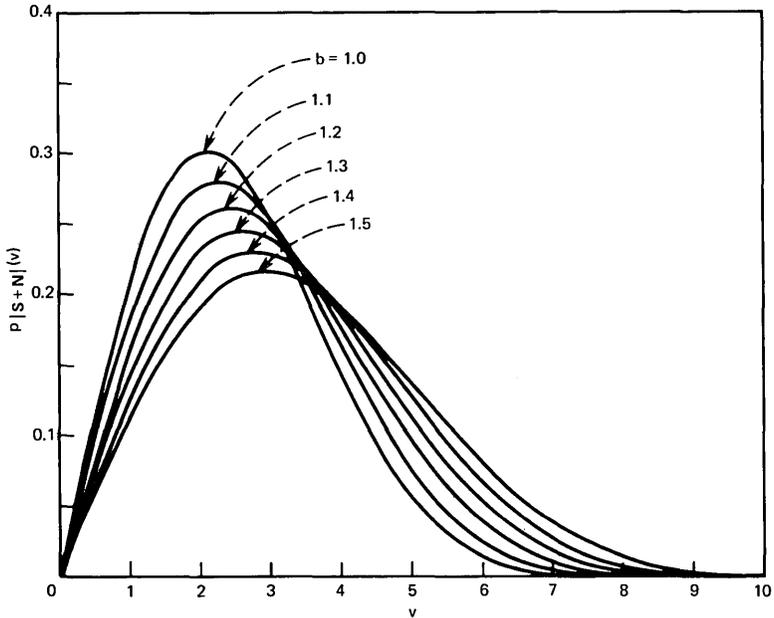
Fig. 3—Plots of p.d.f. of $|S + N|$ for $\sigma^2 = 1$, $M = 6$, $n = 2$, and vector lengths all equal to b.
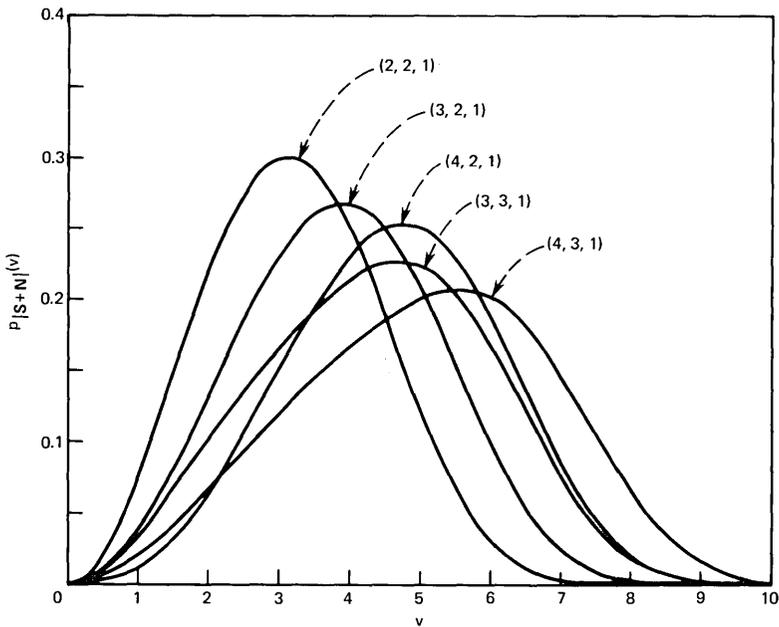
Fig. 4—Plots of p.d.f. of $|S + N|$ for $\sigma^2 = 1$, $M = 3$, $n = 3$, and different sets of vector lengths (b_1, b_2, b_3).
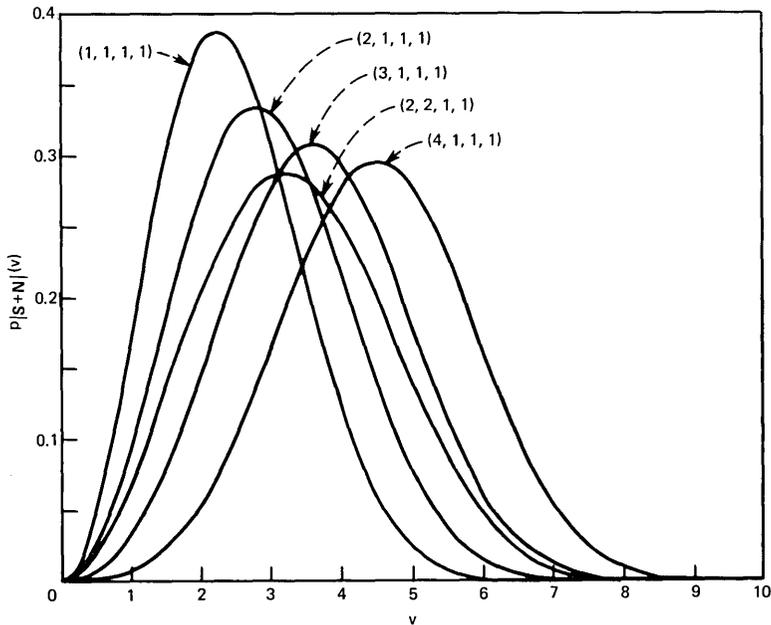
Fig. 5—Plots of p.d.f. of $|S + N|$ for $\sigma^2 = 1$, $M = 4$, $n = 3$, and different sets of vector lengths (b_1, b_2, b_3, b_4).
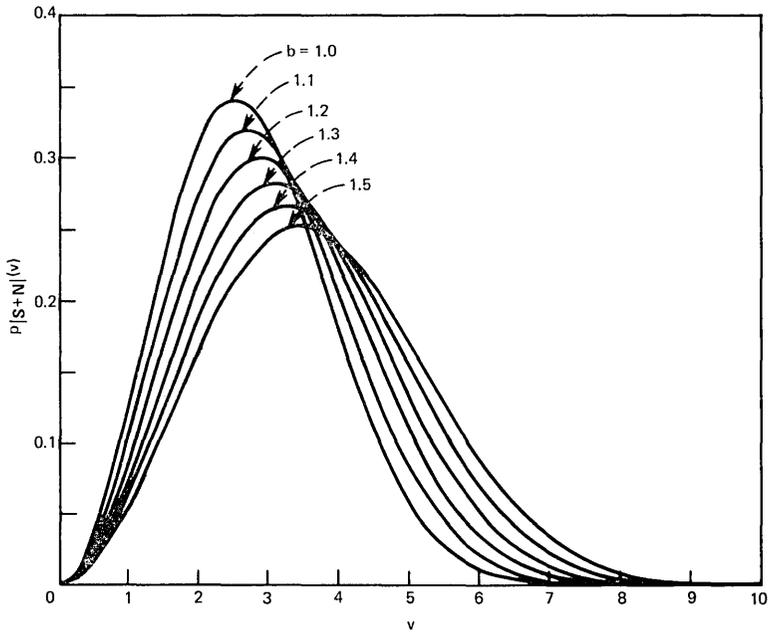
Fig. 6—Plots of p.d.f. of $|S + N|$ for $\sigma^2 = 1$, $M = 6$, $n = 3$, and vector lengths all equal to b.
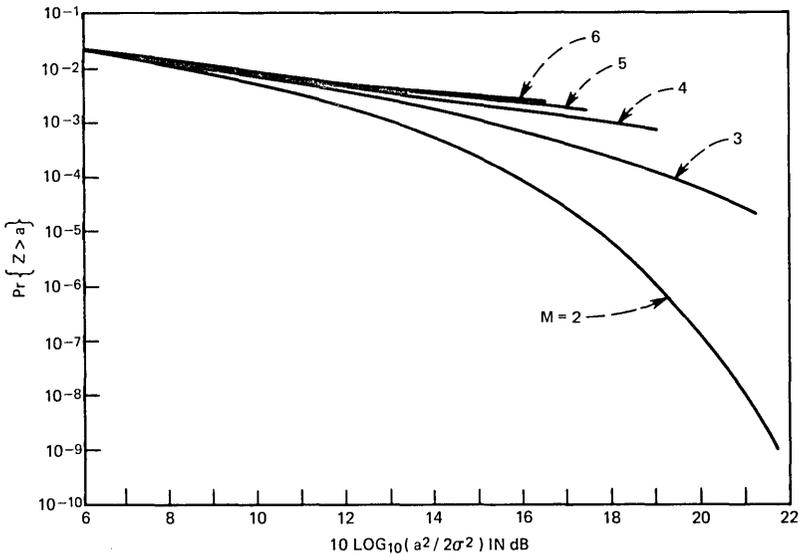
Fig. 7—Plots of $\Pr \{Z > a\}$ for $n = 2$, $10 \log_{10} (a^2/\sum_{i=1}^{M} b_i^2) = 6$ dB, $b_1 = \cdots = b_M$, and for various values of M.

Fig. 8—Plots of Pr $\{Z > a\}$ for $n = 2$, $10 \log_{10} (a^2 / \sum_{i=1}^{M} b_i^2) = 8$ dB, $b_1 = \cdots = b_M$, and for various values of M.

together with $C_n(i, 0) \equiv 1$. To evaluate $\mu_M^{(2i)}$ from eq. (35), particular sets of moments $\{\nu_j^{(2i)}\}$ could be read into the program. However, for simplicity we chose spherically symmetric vectors having constant lengths $b_1, \cdots, b_M$.

The second part of the program was concerned with computation of

$$L_i^{(\alpha)}(x) \left(- \frac{1}{2\sigma^2} \right)^i \Big/ \Gamma \left(\frac{n}{2} + i \right).$$

In order to avoid "overflow" difficulties, we actually computed

$$L_i^{(\alpha)}(x) \lambda^i \Big/ \Gamma \left(\frac{n}{2} + i \right) \qquad \text{with} \qquad \lambda = - \frac{1}{2\sigma^2} \left(\sum_{i=1}^{M} b_i \right)^2.$$

Fig. 9—Plots of Pr $\{Z > a\}$ for $n = 2$, $10 \log_{10} (a^2/\sum_{i=1}^{M} b_i^2) = 10$ dB, $b_1 = \cdots = b_M$, and for various values of M.

To do this efficiently we used the iterative identity:

$$\frac{L_i^{(\alpha)}(x)\lambda^i}{\Gamma[(n/2) + i]} = \frac{(2i + \alpha - 1 - x)}{i\Gamma[(n/2) + i]} \lambda^i L_{i-1}^{(\alpha)}(x)$$

$$- \frac{(i + \alpha - 1)}{i\Gamma[(n/2) + i]} \lambda^i L_{i-2}^{(\alpha)}(x), \qquad i \geq 2 \quad (39)$$

[which follows from the Laguerre polynomial recurrence relation (Ref. 10, p. 241)], together with the fact that $L_0^{(\alpha)}(x) = 1$ and $L_1^{(\alpha)}(x) = \alpha + 1 - x$.

The final part of the program computed the sum:

$$\sum_{i=0}^{\infty} \frac{L_i^{(\alpha)}(x)\lambda^i}{\Gamma[(n/2) + i]} \hat{\mu}_M^{(2i)}, \qquad (40)$$

where $\hat{\mu}_M^{(2i)} = \mu_M^{(2i)}/(\sum_{i=1}^{M} b_i)^{2i}$. [The factor $1/(\sum_{i=1}^{M} b_i)^{2i}$ was built into

the computation of eq. (35) in order to find $\hat{\mu}_M^{(2i)}$.] A convergence check was provided to end the summation after additional terms did not change any significant digits. Although the program was written to handle up to 200 terms in the sum, many computations required less than 50 terms. As Esposito and Wilson[2] also noted, for certain values of x, σ^2, and $\{b_i\}$, the terms in (40) alternate in sign and have magnitudes of the order 10^{15}. For these cases, precision and convergence could not be guaranteed. The typical CPU time required to compute eq. (40) for 200 values of x was about 10 to 20 seconds in double precision arithmetic on the IBM 370/165 system.

Some representative results of these computations are shown in Figs. 1 to 12. Figures 1 to 6 are plots of $p_{|S+N|}(v)$ as a function of v for $\sigma^2 = 1$, for various values of n and M, and for s.s. vectors having constant lengths b_1, $\cdots$, b_M. Curves for $n = M = 2$ were given in Ref. 2. Figures 7 to 12 are plots of Pr $\{Z > a\}$ versus the quantity $10 \log_{10} (a^2/2\sigma^2)$ for fixed values of the quantity $10 \log_{10} (a^2/\sum_{i=1}^{M} b_i^2)$ and for various values of n and M. In these curves, for simplicity, we took $b_1 = b_2 = \cdots = b_M$. As we discussed in the last section, the plots in Figs. 7 to 12 represent the probability of error of a binary coherent
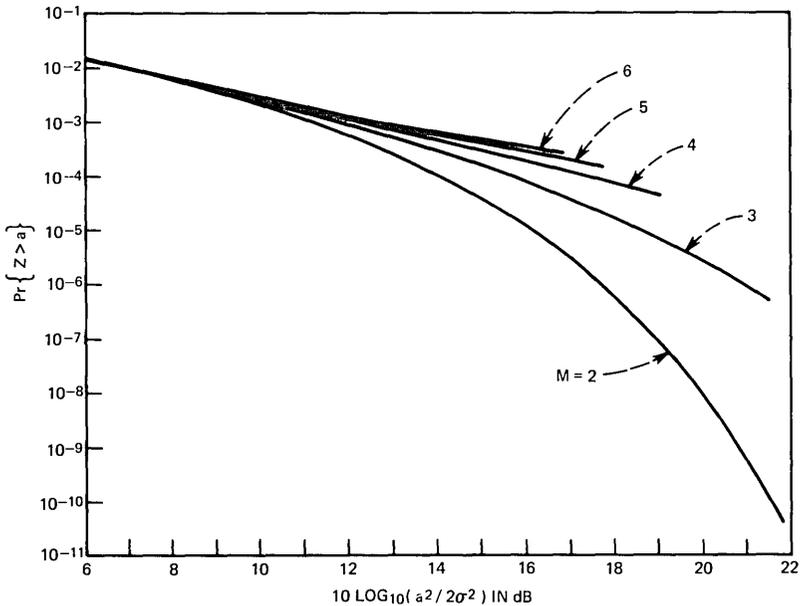
Fig. 10—Plots of Pr $\{Z > a\}$ for $n = 3$, $10 \log_{10} (a^2/\sum_{i=1}^{M} b_i^2) = 6$ dB, $b_1 = \cdots = b_M$, and for various values of M.
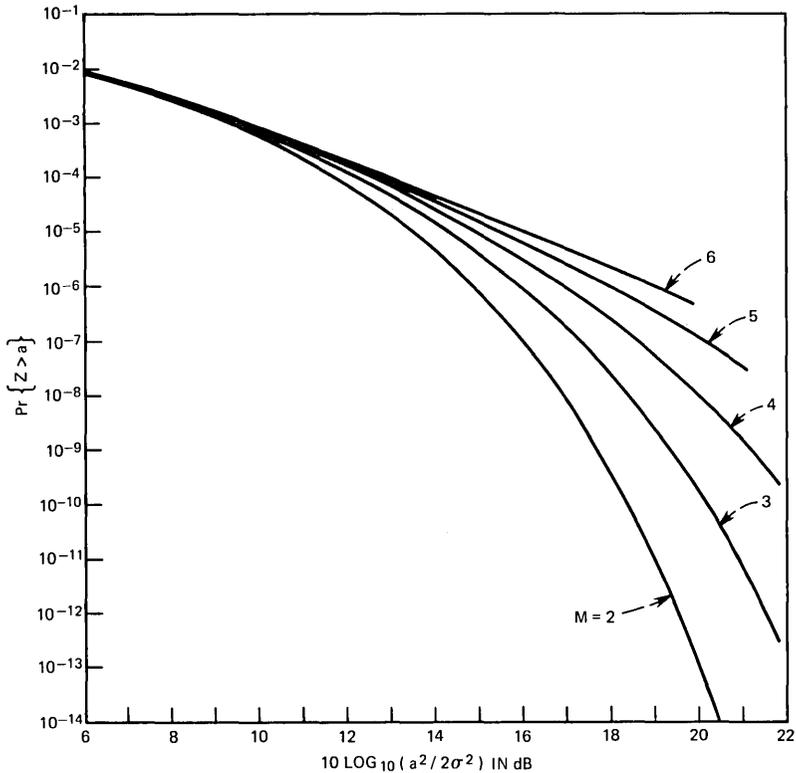
Fig. 11—Plots of Pr $\{Z > a\}$ for $n = 3$, $10 \log_{10} (a^2/\sum_{i=1}^{M} b_i^2) = 8$ dB, $b_1 = \cdots = b_M$, and for various values of M.

phase-shift-keying system versus signal-to-noise ratio,

$$\text{SNR} = 10 \log_{10} (a^2/2\sigma^2) \qquad \text{(dB)},$$

for fixed values of signal-to-interference ratio,

$$\text{SIR} = 10 \log_{10} \left(a^2 \Big/ \sum_{i=1}^{M} b_i^2 \right) \qquad \text{(dB)}.$$

These results extend those previously found in Refs. 4 and 5 to larger values of SNR and smaller values of SIR.

V. CONCLUSION

In this paper we presented expressions for the p.d.f. of a sum of spherically symmetric random vectors plus a Gaussian vector in n-dimensional space. We also found expressions for the p.d.f. and d.f. of the length of this sum and of the projection of this sum onto 1-dimen-
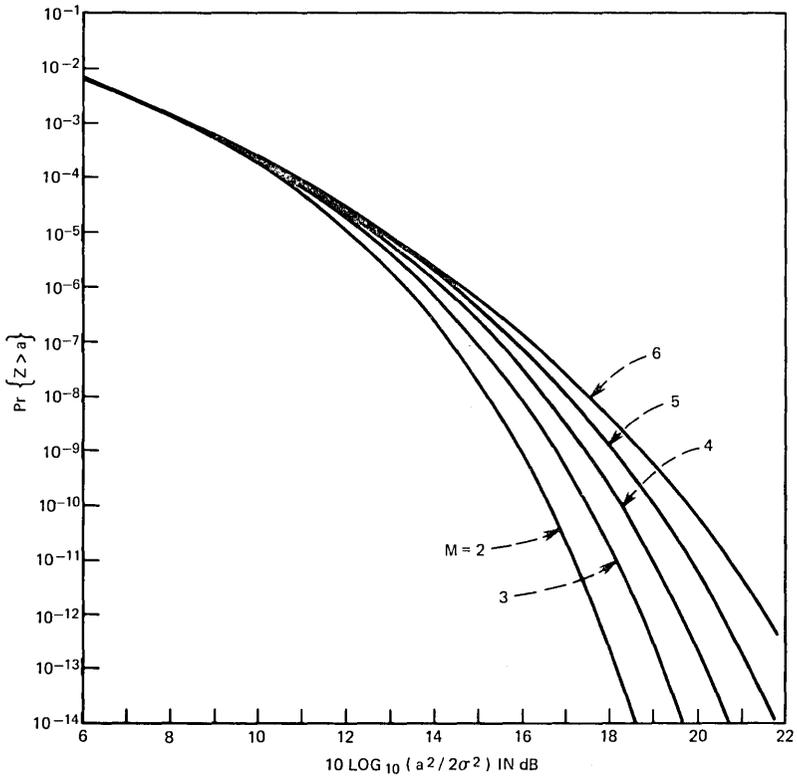
Fig. 12—Plots of Pr $\{Z > a\}$ for $n = 3$, $10 \log_{10} (a^2/\sum_{i=1}^{M} b_i^2) = 10$ dB, $b_1 = \cdots = b_M$, and for various values of M.

sional space. All of these expressions were series expansions involving only the moments of the length of the sum of the s.s. vectors. These moments could be found from recurrence relations also derived in the paper. Some computations of the p.d.f.'s and d.f.'s were presented for the 2- and 3-dimensional cases, and an application to a communications system was discussed. However, as pointed out earlier in Refs. 2 and 3, there are sometimes difficulties in evaluating these p.d.f.'s and d.f.'s for certain parameter values, even for the case of s.s. vectors having constant lengths.

APPENDIX

To prove eq. (22) we use the fact (Ref. 10, p. 241) that, for $i \geqq 1$,

$$\frac{d}{dt} \left[e^{-t} t^{\alpha+1} L_{i-1}^{(\alpha+1)}(t) \right] = i e^{-t} t^{\alpha} L_i^{(\alpha)}(t).$$

Integrating this expression over the interval (a, ∞) yields

$$- e^{-a}a^{\alpha+1}L_{i-1}^{(\alpha+1)}(a) = i \int_a^\infty e^{-t}t^\alpha L_i^{(\alpha)}(t)dt. \tag{41}$$

Equation (22) follows from (41) after a simple change of variables.

REFERENCES

1a. S. O. Rice, "Mathematical Analysis of Random Noise (concluded)," B.S.T.J., *24*, No. 1 (January 1945), pp. 46–156.
1b. S. O. Rice, "Statistical Properties of a Sine Wave Plus Random Noise," B.S.T.J., *27*, No. 1 (January 1948), pp. 109–157.
2. R. Esposito and L. R. Wilson, "Statistical Properties of Two Sinewaves in Gaussian Noise," IEEE Trans. Info. Theory, *IT-19*, March 1973, pp. 176–183.
3. S. O. Rice, "Probability Distributions for Noise Plus Several Sine Waves—The Problem of Computation," unpublished work.
4. V. K. Prabhu, "Error Rate Considerations for Coherent Phase-Shift Keyed Systems with Cochannel Interference," B.S.T.J., *48*, No. 3 (March 1969), pp. 743–767.
5. A. S. Rosenbaum, "Binary PSK Error Probabilities with Multiple Cochannel Interferences," IEEE Trans. Comm. Tech., *COM-18*, June 1970, pp. 241–253.
6. K. Yao, "A Representation Theorem and Its Applications to Spherically-Invariant Random Processes," IEEE Trans. Info. Theory, *IT-19*, September 1973, pp. 600–608.
7. I. F. Blake and J. B. Thomas, "On a Class of Processes Arising in Linear Estimation Theory," IEEE Trans. Info. Theory, *IT-14*, January 1968, pp. 12–16.
8. S. Bochner, *Lectures on Fourier Integrals*, Annals of Mathematical Studies, Study 42, Princeton, N.J.: Princeton University Press, 1959.
9. J. Goldman, "Detection in the Presence of Spherically Symmetric Random Vectors," unpublished work.
10. W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics*, New York: Springer-Verlag, 1966.
11. W. Rudin, *Real and Complex Analysis*, New York: McGraw-Hill, 1966.
12. A. Erdelyi, *et al.*, *Higher Transcendental Functions*, Vol. II, New York: McGraw-Hill, 1953.
13. D. Sakrison, *Communication Theory: Transmission of Waveforms and Digital Information*, New York: John Wiley and Sons, 1968.
14. J. Goldman, "Multiple Error Performance of PSK Systems with Cochannel Interference and Noise," IEEE Trans. Comm. Tech., *COM-19*, August 1971, pp. 420–430.

Experimental Investigation of a Linear 500-Element 3-Phase Charge-Coupled Device

By C. H. SÉQUIN

A linear, 500-element, 3-phase charge-coupled device, originally built as a high-resolution linear image sensor, has been chosen as a representative structure for single-level, 3-phase charge-coupled devices to exemplify the performance of such devices in detail.

Charge-handling ability and transfer efficiency have been studied as a function of various design parameters and operating conditions. Most of the observed functional dependences are well understood and agree with the expectations based on model calculations. However, various problems are encountered in these structures. An unusually wide spread of the performance of different devices and slow instabilities are observed. They are attributed to a lack of control of the interface potential in the gaps between the transfer electrodes.

Some emphasis is placed on a more detailed description of the various measurement techniques used. These techniques are of a general interest since they are applicable to other charge-transfer devices.

I. INTRODUCTION

The principle of charge coupling was first conceived using a 3-phase technique with simple, nondirectional transfer electrodes.[1] This approach using all identical electrodes in a single level of metallization has been successfully demonstrated on several successive designs of linear devices; each structure has a greater number of elements with smaller electrode length than the previous design.[2-4] Correspondingly, improved efficiency at higher frequencies has been observed.

The charge-coupled device (CCD) discussed in this paper was designed as a high-resolution linear image sensor with 500 three-phase elements at a spatial period of 18 μm, capable of reading half a line of

a normal page of printed material.[4] The device has also been demonstrated as an analog shift register able to delay a line of the *Picturephone®* video signal.[5] In these demonstrations, a few of the best devices have been used, and their performance has been described in the context of these particular demonstrations only. The following discussion enlarges the description of the performance of 3-phase single-level metal CCD's by reporting more representative results obtained in a study of many devices. The goal of this study was to obtain an understanding of the functional dependences of charge-handling ability, transfer efficiency, and dark current on operating potentials, frequency, and temperature. In addition, devices with different electrode lengths and different fabrication technologies have been compared. However, similar devices on different slices show a wide performance spread and slow instabilities and show that some results are even irreproducible. All this is attributed to a lack of control of the interface potential in the gaps between the transfer electrodes.

The techniques used to determine signal-handling ability, transfer efficiency, or dark current profiles are described in some detail, since they are applicable to the investigation of other charge-transfer devices.

II. THE 500-ELEMENT DEVICE

The device discussed in this paper was designed as a linear image sensor with high-resolution density. At the same time, the device was supposed to serve as a test structure for the high-frequency performance of 3-phase CCD's. To obtain good transfer efficiency, a "smooth" interface potential profile with no barriers or pockets is required, which transfers the minority carriers under the influence of electric fringe fields from electrode to electrode. Extensive computer modeling studies[6] have shown that this condition can be achieved in a 5×10^{14} cm^{-3} silicon substrate with 3000 Å of SiO$_2$ as an insulator and electrodes shorter than 5 μm. If, in addition, the gaps between the electrodes are made 3 μm or smaller, the device can be operated at pulse amplitudes of 15 V and no barriers will form underneath the transfer gaps for charge densities at the Si-SiO$_2$ interface ranging from about 5×10^{10} to 2×10^{11} cm^{-2}. Thus, the unit cell was made as small as possible using available technology. Individual electrodes, nominally 3 μm long, were arranged at a spatial period of 6 μm, leading to a cell length of 18 μm. The calculated transfer time constant for electrons lies in the subnanosecond range. Ideally, the transfer efficiency in the frequency range of this study should then be limited by the effects of
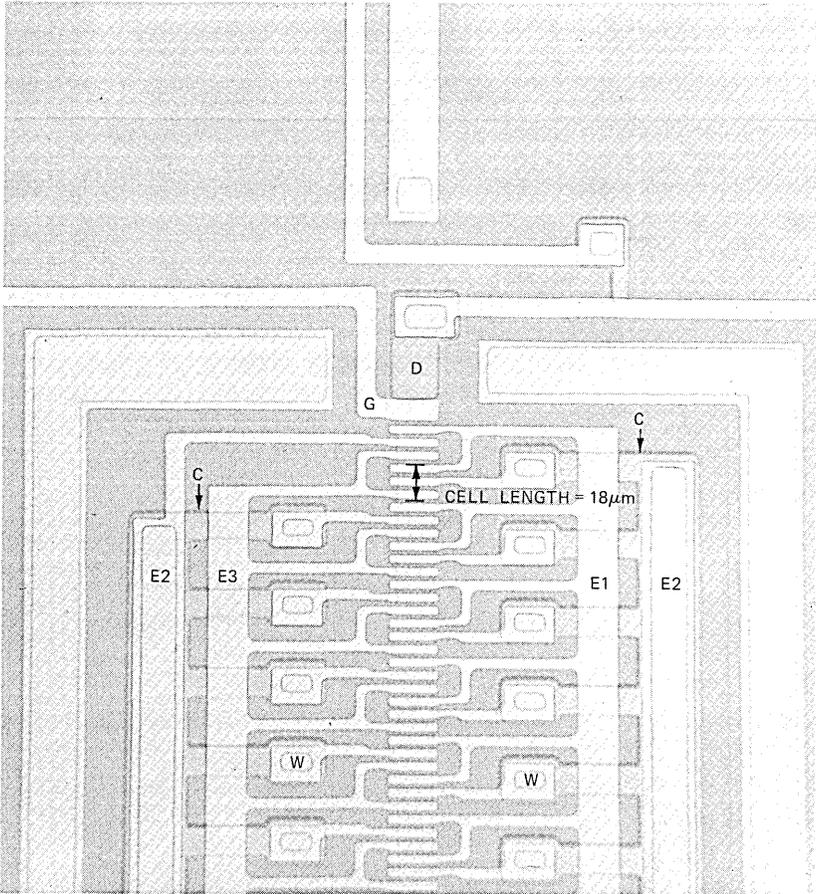
Fig. 1—One end of the linear 500-element 3-phase CCD. Shown are the three sets of transfer electrodes (E1, E2, and E3), the input diode (D), and the input gate (G). Electrodes E2 are alternately contacted to diffused crossunders (C) through contact windows (W) lying on either side of the transfer channel.

interface states. This was indeed found true for the best devices, which showed no problems associated with the bare transfer gaps.

A 3-phase CCD needs one crossing of electrodes per element, which in this device are realized with diffused crossunders. They are introduced by the same phosphorous diffusion (10^{18} cm^{-3}) that forms the input output diodes. The crossunder bus line that addresses electrode system $E2$ (Fig. 1) is repeated on either side of the transfer channel, and the transfer electrodes $E2$ in subsequent elements are contacted

alternately to either one of them. In the resulting structure (Fig. 1), which has a periodicity corresponding to two CCD elements, considerably larger and thus more reliable contact windows can be formed than in a conventional approach with only one bus line.

The channel width, defined by a channel-stopping boron diffusion (10^{17} cm^{-3}), was designed to be 15 μm, taking into account a lateral diffusion of 1.5 μm. This width was chosen as a compromise between the preference for higher signal levels and the reduction in the vertical resolution that would be caused by too wide a channel if the device is used as a line scanner with no separate means to restrict the light-sensitive area.

The transfer channel is terminated at both ends by input output diodes. They are electrostatically shielded from the pulsed transfer electrodes by a gate electrode, which in normal operation is kept at a dc potential. The whole device is surrounded by large substrate contacts which serve as points of reference for the driving pulses. A special small substrate contact near the output diode can serve as a reference ground for the video signal. Including these features, the device dimensions are 230 μm $\times$ 9150 μm.

Devices were fabricated originally on 10-ohm-cm p-type silicon substrates, with 3000 Å of SiO$_2$. Devices with 1500 Å of SiO$_2$, or with a double insulator structure consisting of 1200 Å of SiO$_2$ and 500 Å of Al$_2$O$_3$, were also built for comparison.

In most devices, the transfer electrodes were chemically etched out of 1500 Å of RF-diode sputtered tungsten. Some batches, however, used Al or backsputter delineated[7] Ti-Pd-Ni metallization. With two sets of masks, various exposure times of the photoresist, and various etching procedures, the electrode length could be varied in different batches from about 1.5 μm to 4.5 μm. After metallization, the devices were subjected to annealing treatments to reduce the interface state density. Most commonly, the devices with refractory electrodes were heated in a hydrogen atmosphere at 700 °C for 1 hour; the Al devices were annealed at 380 °C. A considerable improvement in transfer efficiency was normally observed.

In operation, many devices showed a strong sensitivity to changes of the ambient, which could be demonstrated by breathing onto the surface. To reduce these effects, some devices were protected with a second dielectric level, such as 1 μm of a phosphorous glass or 1000 Å silicon nitride.

Most of the results presented in the following sections were obtained on devices with 3000 Å of SiO$_2$, with electrodes 3.5 μm long of RF-

diode sputtered tungsten, and with no protective dielectric layer on top of the electrodes. Deviations in behavior due to different technologies will be discussed in a special section.

III. EXPERIMENTAL PROCEDURE

After screening tests on the uncut wafer, working devices were mounted on ceramic substrates for the investigation. A few selected devices have been demonstrated as line image sensors[4] or analog delay lines.[5] The following sections present a detailed report of studies carried out on several different devices with reasonably good performance.

The mounted devices were investigated in a test setup, illustrated in Fig. 2, built around an optical microscope. A set of TTL logic, three sets of pulse drivers with different rise times, and several preamplifiers with various bandwidths were used to investigate the devices in the frequency range from 1 kHz to 17 MHz. The devices could be operated
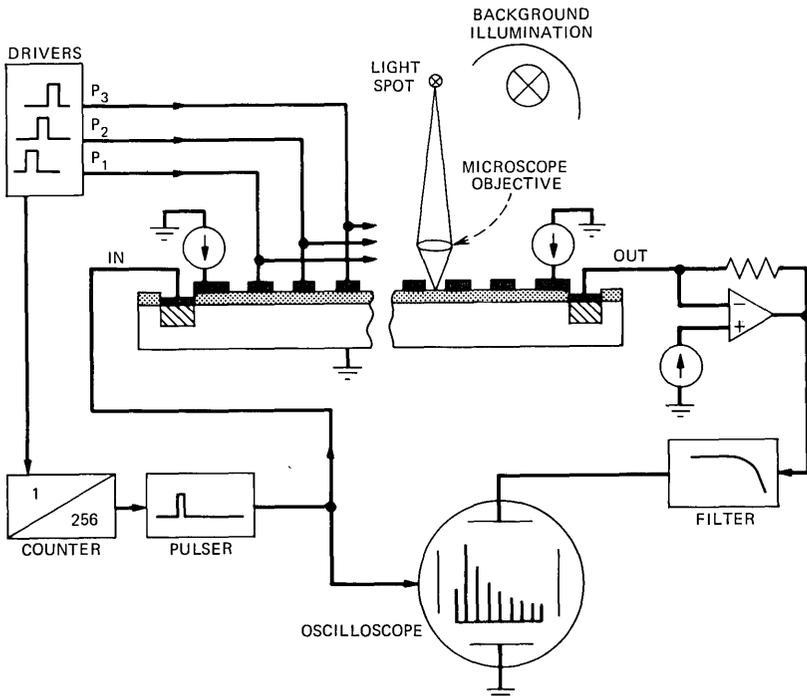

Fig. 2—Schematic layout of test setup.

as analog shift registers with the driving pulses applied continuously and charge packets injected periodically through the input diode. Alternatively, the devices could be held in the storage mode, integrating the dark current or carriers generated by light projected onto it. A sharp spot of light, about 2 μm in diameter, projected through the microscope, was especially useful in these investigations. By moving the spot along the device, singularities in transfer behavior could be localized and often associated with visually observable defects. The built-in object illumination of the microscope was used in addition to provide a fairly uniform background.

The accumulated charge could be read out at various clock rates after various integration times. The driving pulse shapes, pulse amplitudes, and bias potentials, and the mutual overlap between subsequent pulses could also be varied to study their effects on the performance of each device.

IV. CHARGE HANDLING

The basic task of any charge-transfer device is to carry charge along the transfer channel. An important characterization of a device is, therefore, the maximum amount of charge that can be handled at a given amplitude of the driving pulses. In the following section, the functional dependence of the charge handling ability on various operating parameters will be investigated.

The maximum amount of charge that can be carried in the potential wells underneath the transfer electrodes for a given set of operating conditions is measured by observing the output signal as the device is driven into saturation. In the shift-register mode this can be achieved electrically by injecting more and more charge from the input diode. In the storage mode a light spot is used to fill a single well until it starts to spill into neighboring elements. In both cases, beyond saturation the output signal pulse starts to widen rapidly and its amplitude increases only slowly and often in an irregular manner.

The amount of charge that a transfer pad can hold is approximately determined by the product of the oxide capacitance underneath the pad C_{ox} and the applied voltage $V_R + V_P$. Fringe effects increase that capacitance somewhat. On the other hand, not all the applied voltage will appear across the oxide. The interface potential for a full bucket will still be larger than zero. In fact, it cannot be lower than the barriers produced by the isolating electrodes, which are kept at a voltage V_R, without spilling charge. Neglecting fringe effects and the influence of the depletion capacitance, one expects a linear relationship
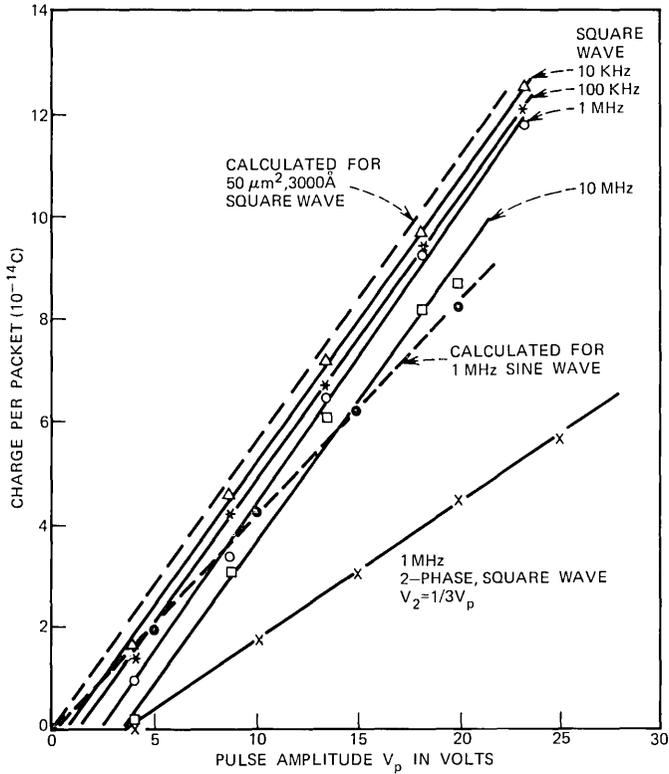
Fig. 3—Charge-handling ability as a function of applied pulse amplitudes for various waveforms.

between the driving pulse amplitude V_P and the signal-handling ability: $Q = C_{ox} \cdot V_P$.

Experimental results derived from the saturation value of the output signal show good agreement with this approximation (Fig. 3). For ordinary 3-phase operation with square pulses, the slope of the curves corresponds to an electrode capacitance of 5.6×10^{-15} F, which in turn corresponds to an area of 50 μm^2 on 3000 Å of SiO$_2$. With a channel width of 15 μm, this yields a calculated effective electrode length of 3.3 μm, which agrees with the observed length within the measurement accuracy of 0.5 μm.

Curves taken at various pulse rates are parallel but do not fall on top of each other. Somewhat lower charge handling is observed as the frequency increases, and the corresponding extrapolated curves cross the abscissa at higher values of V_P. Comparison of measurements

taken at a fixed clock rate but with pulses that have a different fall time show reduced signal amplitude for the slower pulses. This indicates that the frequency dependence of the charge-handling ability may originate from the degradation of the driving pulses. Stray capacitance and the resistance of the diffused crossunders slow down the response time and reduce the pulse amplitude at higher frequencies.

The device was also operated with sine waves and the maximum signal charge plotted as a function of peak-to-peak amplitude. The expected values in this mode are 75 percent of the square-wave operation, since, when one phase is at its maximum value V_P, the two neighbors are at a voltage $V_P(1 + \cos 120°)/2 = 0.25\ V_P$. Experimentally observed values fall well upon the calculated straight line through the origin. These measurements show a much smaller frequency dependence.

A 3-phase device can also be operated in an asymmetric 2-phase mode by leaving one set of electrodes at an intermediate dc potential V_2. Using a simplistic model that neglects fringe effects, one would
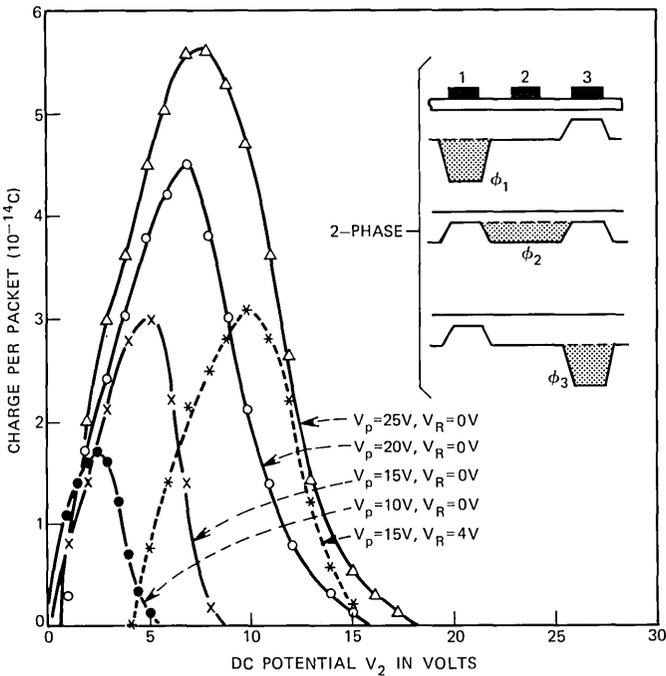
Fig. 4—Charge-handling ability in asymmetric 2-phase mode of operation (see inset) as a function of the dc potential of the static phase for various pulse amplitudes.

expect that the largest signal current could be carried when this intermediate potential V_2 is at half the pulse voltage V_P, and that for this case the signal handling should be half that of the normal 3-phase operation.

Experimentally, it is seen that for all measured values of V_P, the maximum signal current is reached for a dc potential V_2 equal to about one-third of V_P (Fig. 4). For this case, the charge-handling ability is about 40 percent of that observed in normal 3-phase operation. This is due in part to the fact that for higher values of V_2 the transfer efficiency decreases, owing to the formation of insuppressible barriers between the dc phase and an adjacent phase that is fully turned on. Also, in this structure that has gap widths comparable to the electrode lengths, the gaps themselves may play a significant role in the charge-storage process. The inset of Fig. 4 illustrates that fact. In this model the dc phase V_2 turned on to $V_P/3$, together with the two adjacent gaps, can store the same amount of charge as the well underneath phase 1 or 3 in the asymmetric 2-phase mode.

The important role of the potential in the gaps is also expressed in a strong dependence of the performance on the resting potential V_R. The inset of Fig. 5 shows the function of V_R serving as a bias on top
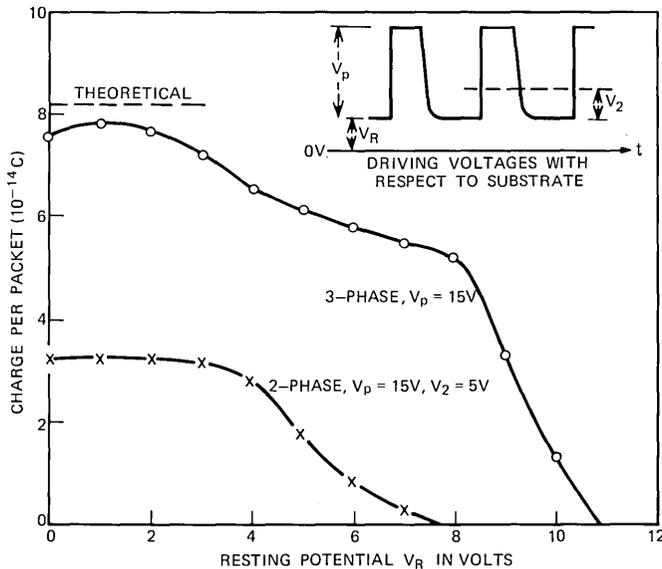
Fig. 5—Charge-handling ability as a function of resting potential V_R (see inset for definition) for normal 3-phase operation and for the asymmetric 2-phase mode.

of which the pulse potentials and, if applicable, V_2 are floating. Figure 5 then shows the dependence of the charge-handling ability on the rest potential V_R. In the normal 3-phase mode, the operating range of V_R is about 8 V. This is a rather high value for a 3-phase CCD. Typically, these devices are much more sensitive to V_R and show ranges of only 1 to 3 V. The range of V_R increases for higher pulse voltages V_P, because the stronger fringe fields can suppress larger barriers in the gap. In the asymmetrical 2-phase mode, where the voltage difference on adjacent pads is smaller, a smaller operating range of V_R is found.

Switching all electrodes to the same dc potential allows one to operate the device as a long IGFET. Figure 6 shows two sets of curves. The dashed lines are the drain current versus gate voltage curves taken in a point-by-point measurement starting at low values of V_G. The electrodes were then held at a dc potential of 30 V for 90 minutes, and the measurements were repeated working from high toward low values of V_G. The strong hysteresis observed is produced by the slow time constants involved in charging the gaps to the potential of the electrodes.

Values for the carrier mobility deduced from steady-state curves obtained after the device had been sitting at a certain gate potential for a sufficiently long time are in fair agreement with measurements taken on ordinary test IGFET's with continuous gate electrodes. In both cases, the values range around 700 cm^2 V^{-1} s^{-1}.

V. TRANSFER INEFFICIENCY

5.1 Introduction

For practically all applications of a charge-coupled device, the most crucial parameter of the device is its transfer inefficiency. Figure 7 shows some calculated output pulse trains, produced in response to the injection of single charge packets into the input of devices with different overall transfer performance. This computation was done using a linear small-signal approximation,[8] which assumes that in each transfer every charge packet leaves a fixed fraction of its charge ϵ behind, regardless of signal amplitude or the charge contained in previous stations. The overall performance of the device is suitably characterized with a "transfer inefficiency product" $n\epsilon$ multiplying the number of transfers n with the fraction of charge ϵ left behind in each transfer. Experimentally, this $n\epsilon$ product is determined by comparing the observed output pulse train with calculated model plots. For large values of $n\epsilon$ the delay of the maximum amplitude of the output pulse train is measured with respect to the calculated exit time
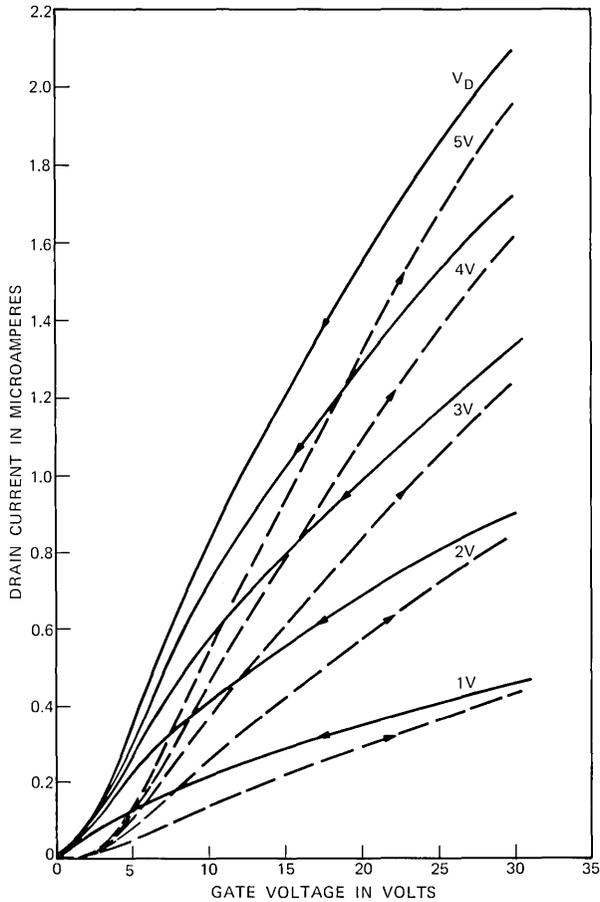
Fig. 6—Operation of the CCD as an IGFET by tying all electrodes together. Hysteresis is observed due to leakage of charge from the electrodes into the transfer gaps.

from an ideal device. This delay, expressed in a number of time slots given by the inverse of the clock frequency, is numerically equal to the transfer inefficiency product $n\epsilon$.[8]

5.2 Results

A linear model describes the actual behavior of a real device imperfectly. Small charge packets following a string of empty buckets show the biggest degradation. Figure 8 illustrates the dependence of the transfer efficiency on signal amplitude and on background charge. A
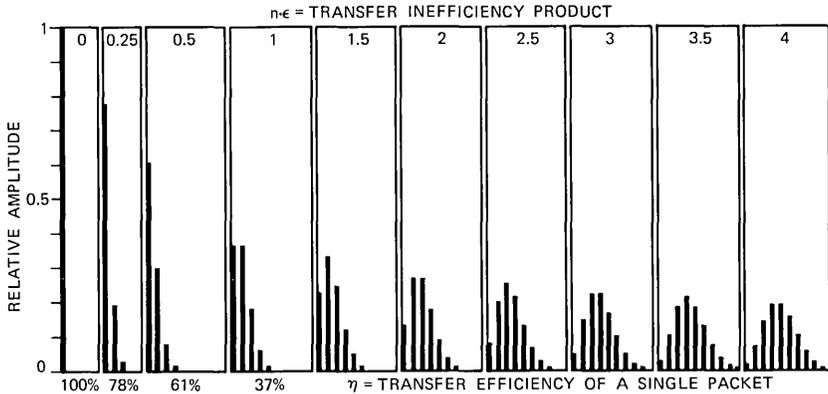
Fig. 7—Appearance of a single-charge packet at the output after transfer through a CCD with a total inefficiency product $n\epsilon$ (linear model).

sharp spot of light with a visible diameter of about 2 μm was focused on the gap between two electrodes 150 transfers from the output end. By varying the integration time, the linear amplitude range of the device was determined. The output pulse train was then monitored as a function of signal level [Fig. 8(a)]. The inefficiency product decreases from 0.8 to 0.5 as the signal is increased from a fraction of 0.5 to 1.0 of the linear amplitude range, which was defined earlier as the saturation point. If more charge is injected, it can no longer be held by a single potential well. Some charge then overflows into the neighboring stations, forward as well as backward. In the output pulse train, some charge is observed to come out earlier than the proper time slot.

The influence of background charge was studied by illuminating the device uniformly at various intensities. A signal charge packet corresponding to half a full well was injected with the sharp light spot and the output monitored as a function of the amount of background charge [Fig. 8(b)]. The inefficiency product decreases from 0.8 to 0.4 as the background charge is increased from 0 to a fraction of 0.5 of the linear range. The first 20 percent of background charge yields the most significant improvement in performance. The measurements of Fig. 8 have been taken at the clock rate of 1 MHz. The behavior is typical for frequencies below 2 MHz.

In most of the following experiments the transfer inefficiency was measured by using the device as an analog shift register. Input diode and input gate were kept at dc potentials such that the diode would just trickle a small amount of current into the device and thus provide some background charge. Every 256th clock pulse, the input diode was
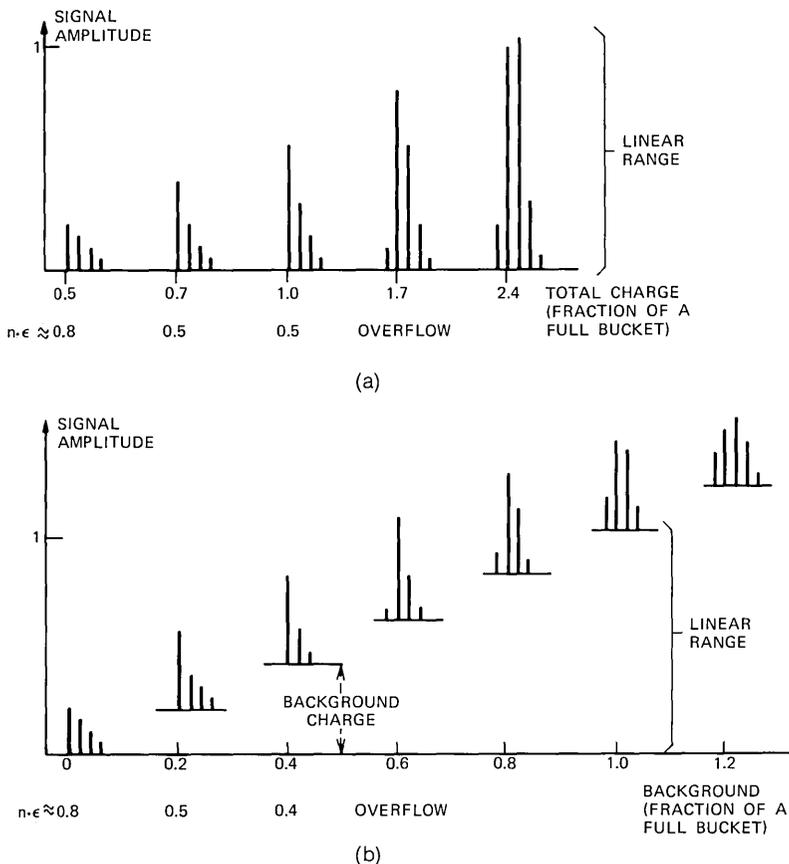
Fig. 8—Experimentally observed signal amplitudes (a) as a function of the total charge placed into one packet, (b) as a function of background charge with constant input signal charge (1 MHz).

pulsed to a more negative value for a time interval of about one-third of the clocking period and, thus, a well-defined packet of charge was injected. The charge packet was due to leave the device 500 clock pulses later. At that point, a time mark was generated that served as a reference point.

The signal from the output diode was led into a linear preamplifier consisting of a cascode stage with an active load, an emitter follower, and a current-feedback branch to the gate of the input J-FET. A set of three different amplifiers was used to cover the full range from 1 kHz to 17 MHz. The potential of the output diode was normally kept at about 10 V and the output gate at a dc potential of 2 to 5 V.

Figure 9 shows the output from two 500-element devices operated in that mode at 1 MHz. Comparison with the calculated plots in Fig. 7 shows that the two pulse trains correspond to inefficiency products of about 0.25 [Fig. 9(a)] and 3.0 [Fig. 9(b)]. However, in both cases, the tail of the pulse train extends much further than in the calculated examples. In Fig. 9(a) at least five stations carry an observable amount of charge and in Fig. 9(b) the tail extends well beyond the range of the picture. This is due to the nonlinear dependence of transfer efficiency on signal amplitude. Owing to the shape of the potential well, small packets lose a larger fraction of their charge to the following packets by trapping effects in interface states.[9] The charge packets forming the tail of the pulse train are thus transferred less and less efficiently and become more and more delayed. This nonlinear behavior is also the reason that the addition of a small amount of background charge can drastically improve the performance.
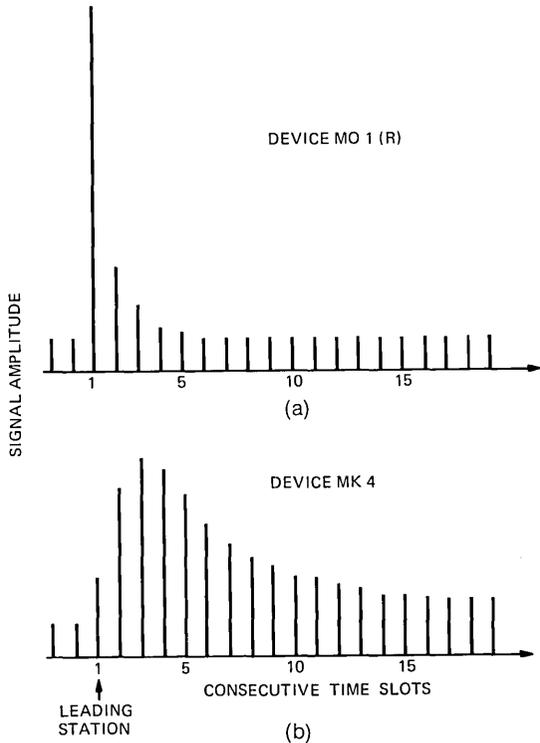
Fig. 9—Two experimentally observed signal outputs that illustrate disagreement with linear model.

The transfer efficiency of the devices studied is a sensitive function of most of such operating parameters as waveform, amplitude, and bias of the driving pulses. The dependence on signal amplitude and background charge has already been mentioned. In the following experiments, the latter two parameters were adjusted by trial and error to give the best results. Background charge ranged typically from 20 to 50 percent.

The devices are also sensitive to any changes occurring near the surface of the silicon. Some devices have shown strong sensitivity to the ambient atmosphere. Breathing slightly onto the device can change its performance considerably. A majority of the mounted devices were thus covered with a phosphorous glass to protect the bare gaps between the electrodes from the influence of the ambient. Though the sensitivity to the atmosphere could be strongly reduced by this means, the devices still showed a dependence on the history of the investigation. Prolonged operation (over one hour) at high potentials often considerably impaired the performance at lower potentials. The devices did, however, recover and resume good performance at lower potentials after they had been turned off for a few hours. These kinds of instabilities can make the experiments very tedious. To reduce their effects on the results as much as possible, the measurements have been performed first at the lower voltages and then extended to larger pulse amplitudes.

The quoted inefficiency products refer to the linear model except where otherwise stated. This seems to be justified since in the following experiments the emphasis is on the functional dependence rather than on absolute values. This approach simplifies interpretation for the reader since he can visualize that an inefficiency product $n\epsilon$ which ranges between i and $i + 1$ means that the ith station after the proper output time slot carries the maximum amount of charge.

Figure 10 shows the results of an experiment designed to demonstrate the time dependence of the performance of a device. After the device was turned off for several days, an inefficiency product of 1.8 was measured. The device was then completely flooded with current for 15 minutes by grounding the input diode while the pulses were left applied to the electrodes. The input diode was then returned to the normal condition and the performance measured at pulse amplitudes of 20 V. Figure 10 shows the strong degradation and subsequent recovery. A few seconds after the return to measurement conditions the $n\epsilon$ value was about 20 and then recovered to 2.7 within one hour and to 2.4 by the next day.

It is believed that the change in performance is due to a migration
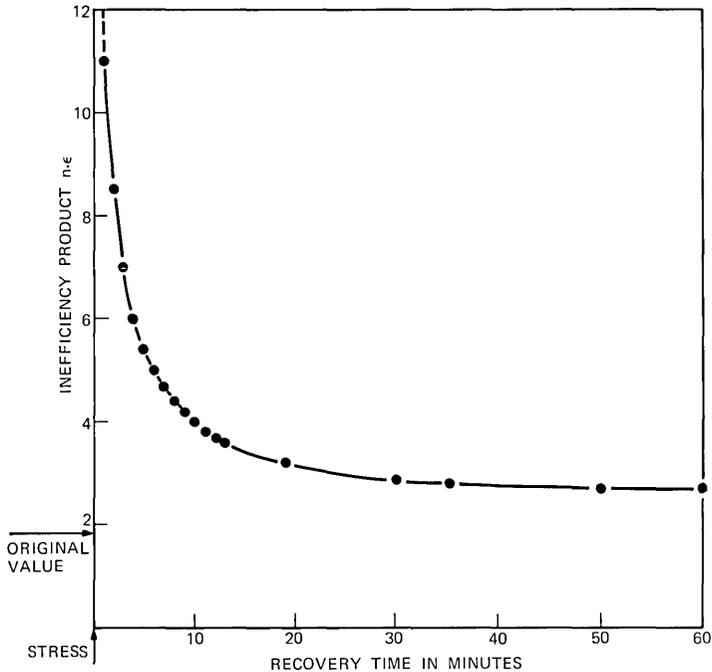
Fig. 10—Strong degradation of performance and subsequent slow recovery intro-duced by strong saturation of the transfer channel.

of charge at the interface between the gate dielectric and the deposited protective P glass. The excessive charge in the flooded transfer channel represents a ground plane that terminates the field lines from the transfer electrodes. Since this charge sheet extends through the whole device, it is also present under the gaps between the electrodes. Field lines from the edges of the transfer electrodes have strong lateral components that can move charge along the outer surface of the gate dielectric and, thus, charge the surface above the gaps more positively. This generates potential pockets in the silicon which can trap part of the signal charge. In normal operation of the device, the forces that charge up the gaps are absent since the charge resides under the gaps only for a fraction of a nanosecond.

The response to saturation is not equally strong in all devices. The state of the surface of the gate dielectric before the deposition of the protective P glass probably plays an important role. In the following measurements, prolonged saturation of the devices was carefully avoided.
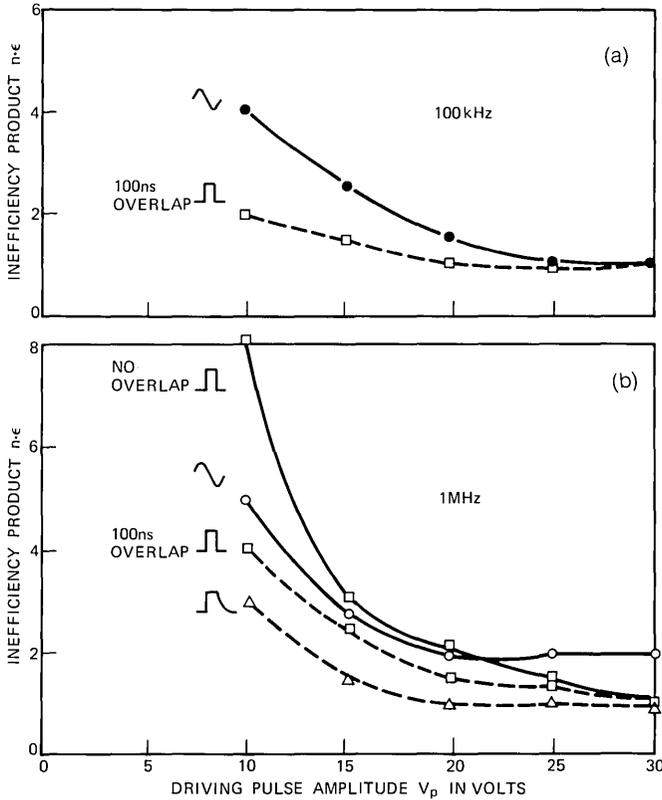
Fig. 11—Transfer inefficiency product as a function of driving pulse amplitude for various waveforms at (a) 100 kHz and (b) 1 MHz.

A set of experiments was performed to determine the best operating conditions at a given frequency. Values of $n\epsilon$ have been determined as a function of driving pulse waveforms, amplitudes, and dc bias. Figure 11 shows the dependence on waveform. At lower amplitudes, shaped square pulses with slower trailing edges give best results. For higher amplitudes, the performance is fairly independent of the driving pulse form, including sine waves.

The strong differences in the performance for the cases of square pulses of 10 V amplitude with mutual overlaps of 0 ns and 100 ns are further explored in Fig. 12. The pulses are about 300 ns long with rise and fall times of about 10 ns. The $n\epsilon$ values are measured as a function of the overlap of the driving pulses. The results are plotted in two different ways. In Fig. 12(a) the overlap in time cor-
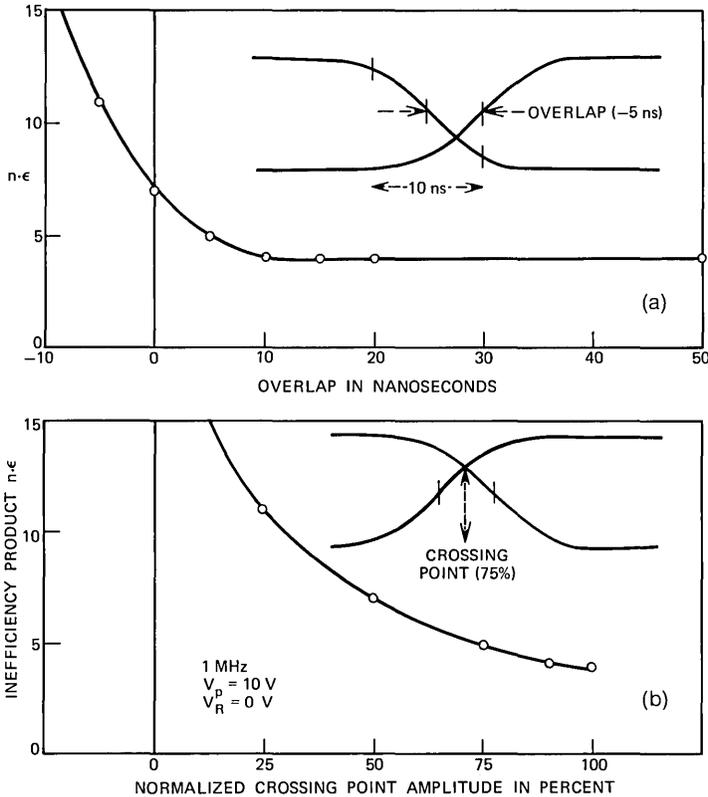
Fig. 12—Transfer inefficiency product as a function of mutual overlap for square driving pulses plotted in two different ways (see inset for definitions).

responding to the 50-percent point of the waveforms is measured in nanoseconds. A negative overlap thus means that the waves cross at less than 50 percent of their peak values. The best performance is reached for overlaps of more than 10 ns. Figure 12(b) shows the same results plotted as a function of the amplitude at the crossing point. The best performance is reached for a crossing point higher than 90 percent of peak amplitude, which corresponds to overlaps of more than 10 ns.

Square pulses with 10-ns overlap have been used to evaluate the influence of the resting potential on the performance. In Fig. 13(a), the dependence on pulse amplitude is first established. For pulse amplitudes higher than about 10 V, which seem to be necessary to override some barrier in the bare gaps, performance improves slowly but

monotonically toward higher amplitudes. For a pulse amplitude of 15 V, the influence of the resting potential V_R was studied [Fig 13(b)]. The device is operable in a range of V_R up to about 10 V with the best performance near 6 V. Again, it has to be pointed out that such a wide operating range has not been observed too often.

In these conditions in which there are square pulses with 10-ns overlap, the transfer efficiency was measured for clock rates between 1 kHz and 17 MHz. At each frequency, the amount of background charge and the signal level were adjusted to give the best results. Background charge ranged from 15 to 30 percent and signal amplitudes from 30 to 50 percent of a full bucket. The devices were operated as shift registers in the continuous wave mode. Single charge packets were injected every 256 clock pulses. Over more than three orders of magnitude from 1
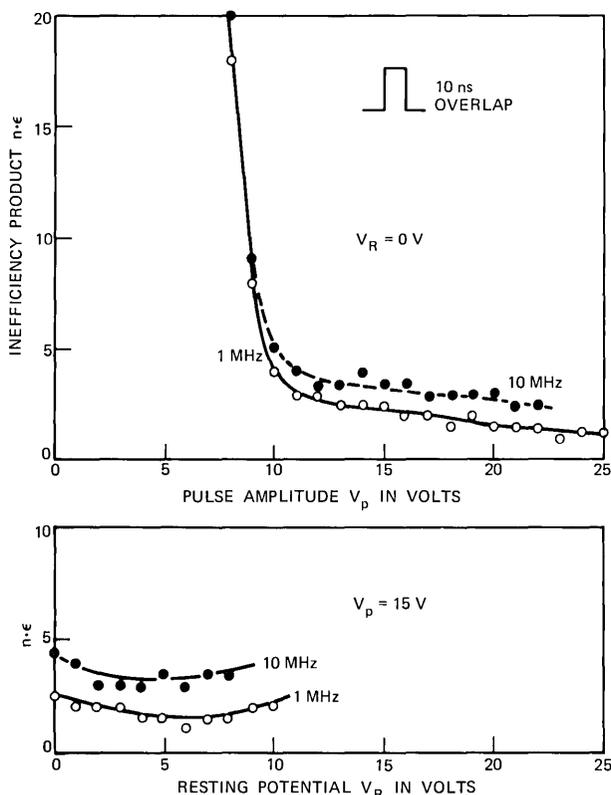
Fig. 13—Transfer inefficiency product as a function of (a) pulse amplitude and (b) resting potential for clock frequencies of 1 MHz and 10 MHz.

CLOCK FREQUENCY IN MEGAHERTZ
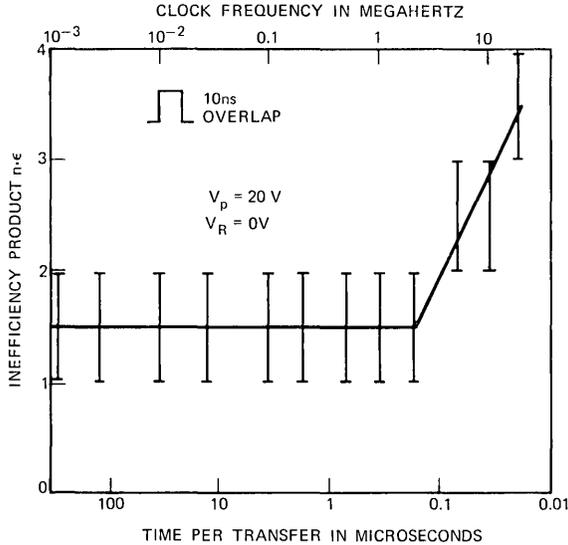
Fig. 14—Transfer inefficiency product as a function of clock frequency or available transfer time.

kHz to 2 MHz no significant differences in performance were measurable (Fig. 14). Then, from 2 MHz to 17 MHz the $n\epsilon$ value rose from about 1.5 to 3.5. This decrease in performance does not stem from the free-charge-transfer mechanism itself. The time available per transfer, even at the highest frequency where it is 20 ns, is still long compared to the calculated transfer time constants, which are well below 1 ns.[6] An increase in the interface state density toward the conduction band edge could account for the decrease in performance at the highest frequencies.[9] Furthermore, above 5 MHz, the driving pulses as observed on the connector to the device were far from perfect, and additional degradation might have been produced by stray capacitance and by the diffused crossunders on the device itself.

5.3 Discussion

To study the transfer performance of a CCD, an optimum number of transfers exist, depending upon the performance of the device. If the device is too short, the degradation is too small to be measured accurately. If the device is too long, the degradation is large and, thus, again $n\epsilon$ is difficult to measure accurately. Values of $n\epsilon$ on the order of one are most easily measured.

For long devices another problem arises. The measured functional dependences might be smeared due to nonuniformities of some physical parameters along the transfer channel, such as flat-band voltage or electrode width. Figure 15 illustrates the integrated $n\epsilon$ values measured with a spot of light injected at various distances from the output diode. It can be seen that $n\epsilon$ is not linearly increasing along the device. In some other devices, sharp steps in the $n\epsilon$ curves have been observed which often could be correlated with an obvious physical defect, such as a partly missing transfer electrode.

The observed inefficiency products ranged from 0.2 to several hundred. They are distributed in a log-normal manner around a value of 20 with a standard deviation of about a factor of five. This spread is too high to be explained by variations in interface state densities. These devices are, however, very sensitive to variations in the fixed oxide charge in the transfer gaps. Too little or too much charge can lead to barriers or pockets in the interface potential, both of which strongly impair the transfer efficiency of the device. These effects associated with the transfer gaps are strong enough to override other parameters
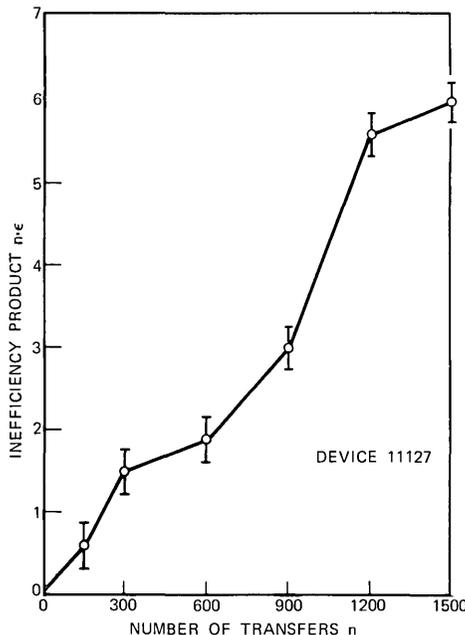
Fig. 15—Transfer inefficiency product from various points of the device to the output. The charge was injected with a light spot.

that could influence the measured inefficiency products. A survey of the storage times of MOS capacitors, of the average dark currents of working devices, and of their n_ϵ values was performed on 35 slices in dynamic probe tests; although the observed values for the storage times and for the dark currents correlated reasonably well from slice to slice, no correlation between n_ϵ and either of the other two measurements could be found.

The best results observed are $n_\epsilon = 0.2$, which corresponds to $\epsilon = 1.3 \times 10^{-4}$, a value comparable to results reported on 2-phase devices.[10,11] Values between 10^{-4} and 10^{-3} can be expected with interface states densities in the low 10^{10} cm^{-2} eV^{-1}.[9] The majority of devices, however, have ϵ values between 10^{-3} and 10^{-2}. These values lie in the range of what has been obtained with ordinary bucket-brigade devices.[12,13] This similarity leads to the suggestion that gapped devices often work in a similar mode. Fixed barriers in the gaps keep a reservoir of carriers underneath each electrode. The signal charge modulates the barrier height, and transfer efficiencies comparable to bucket-brigade structures can be expected.

VI. DARK CURRENT

When the linear CCD is used as a line imaging device, one set of electrodes is switched to an integration potential V_I during the time that charge is being integrated. The minority carriers generated by the incoming light are then collected by the potential wells underneath that set of electrodes. After a sufficient charge pattern has been accumulated, the stored information is read out in serial form.

In the absence of any illumination, minority carriers are still generated by thermal effects and are collected in the nearest potential wells. The generation rate is not necessarily uniform over the whole device and, thus, this dark current can generate a pattern of its own. Figure 16 shows the readout signal of a fairly nonuniform device after integration times of 250 ms and 500 ms where in the latter case the highest peaks of the signal have already reached saturation. During readout each charge packet picks up a little bit of dark current from all the locations it passes on its way to the output diode. From all the other locations on the input end of the device, it had already received a dark current contribution during the readout of the previous line when the "empty" packet was moved from the input diode to the integration site. Therefore, the same integral contribution is added to every charge packet. This uniform component can be subtracted or, if the
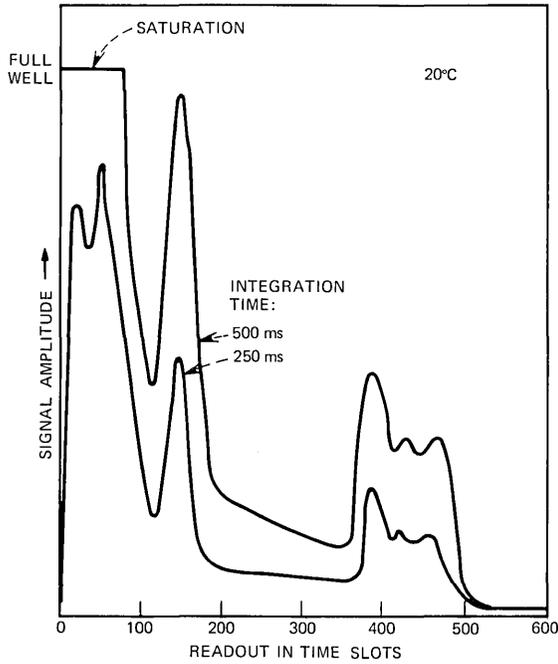
Fig. 16—Integrated dark current profile for two different integration times.

readout time is kept short compared to the integration time, it can be neglected.

For low-light-level-imaging applications, the dark current should be as low as possible. Cooling the device is a possible means to reduce the thermal carrier generation. For a simple demonstration the device was mounted on an open cooling block. Figure 17 shows the results for a temperature interval from $-20\,°C$ to $+60\,°C$. As expected, the inefficiency product showed no significant changes except below $-15\,°C$ where the formation of ice degraded the operation of the device. The dark current measurements displayed in Fig. 17 were taken near element 150 in the signal shown in Fig. 16. Within the measurement accuracy they follow the calculated dependence given by the intrinsic carrier density n_i.

There are mainly two mechanisms that produce a dark current component which is proportional to n_i.[14,15] The generation current arising from bulk states in a 5-μm-wide depletion region is on the order of 6 nA/cm² for a minority carrier lifetime of 100 μs, which typically
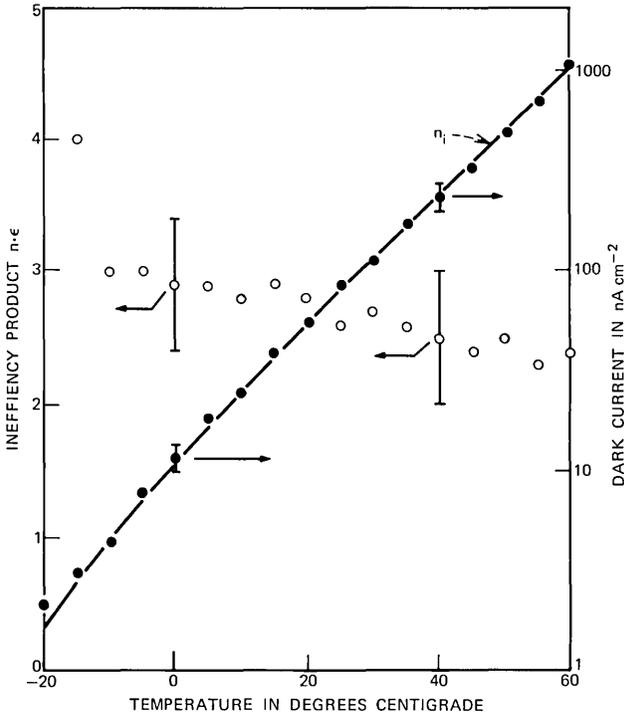
Fig. 17—Measured inefficiency products (left ordinate) and local dark current generation (right ordinate) compared to temperature dependence of intrinsic carrier density n_i (matched at 20°C).

can be expected in a good sample. The generation current arising from interface states with an estimated density of $2 \times 10^{10}/cm^2/eV$ at midband is on the order of 4.6 nA/cm². Thus, surface and bulk dark current contributions are of the same order of magnitude. The shape of the thermal relaxation curve of the MOS capacitor can be analyzed to determine which is the dominant component in a particular sample.[16] In a CCD the situation is complicated by the fact that the two components normally do not have the same active generation area. While the strong depletion region is mainly localized underneath one electrode per element only, a small depletion region extends underneath all electrodes and thus the surface contribution stems from an area that is several times larger.

Several devices showed a fairly uniform dark current background on the order of 10 nA/cm^2, but superposed by highly localized point defects. The temperature dependence of some of these localized

generation centers does not show the simple proportionality with n_i, but has a range of weaker temperature dependences. At a given temperature, the number of observable current sources as well as their strength depends strongly on the electric field, i.e., on the potential V_I applied to the integrating electrode set. Most defects seem to fill the integrating potential well and then stop their activity as would be expected from states that are only active in the depleted bulk of the silicon substrate. Other defects, however, possibly located at the interface, continue to generate minority carriers and fill adjacent wells and eventually a whole transfer channel. These observations indicate that there is a variety of localized defects.

VII. LIGHT SENSITIVITY

The described device can be used in slow-scan-imaging applications as a simple line scanner that integrates the information incident upon the transfer region itself. The transfer electrodes are opaque and, thus, about 50 percent of the light incident on the transfer region is lost. The resolution of such a system in the direction of the electronic scan has been discussed elsewhere[17] and experimental results presented.[4] The resolution in the direction of the mechanical scan depends on the effective light-collecting line width of the device. This width has been measured by probing the device with a very narrow spot of light (approximately 2 μm wide) produced by an incandescent bulb from which the IR radiation has been filtered out. This spot was moved along two different lines across the transfer channel of the device (see Fig. 18). Line A lies beside one of the integrating electrodes, and the generated charge, thus, will spill mainly into the potential well underneath. Carriers generated deep down in the bulk can reach adjacent wells by diffusion. To measure the total amount of charge, the signals of the two adjacent stations were added to the main station. To eliminate effects of transfer inefficiency, the experiment was performed close to the output diode.

In a second experiment, the spot of light was moved along line B lying midway between two integrating electrodes. The charge was then distributed more or less equally into the two potential wells and the sum of the two signals was used in plotting Fig. 18.

Both experiments yield the same sensitivity across the channel. The 50-percent point is about 10 μm outside the edge of the transfer channel, indicating that the channel stopping diffusion does not provide adequate definition of the optical integration region. The equivalent line width of this image sensor is, thus, 35 μm or about twice as large as
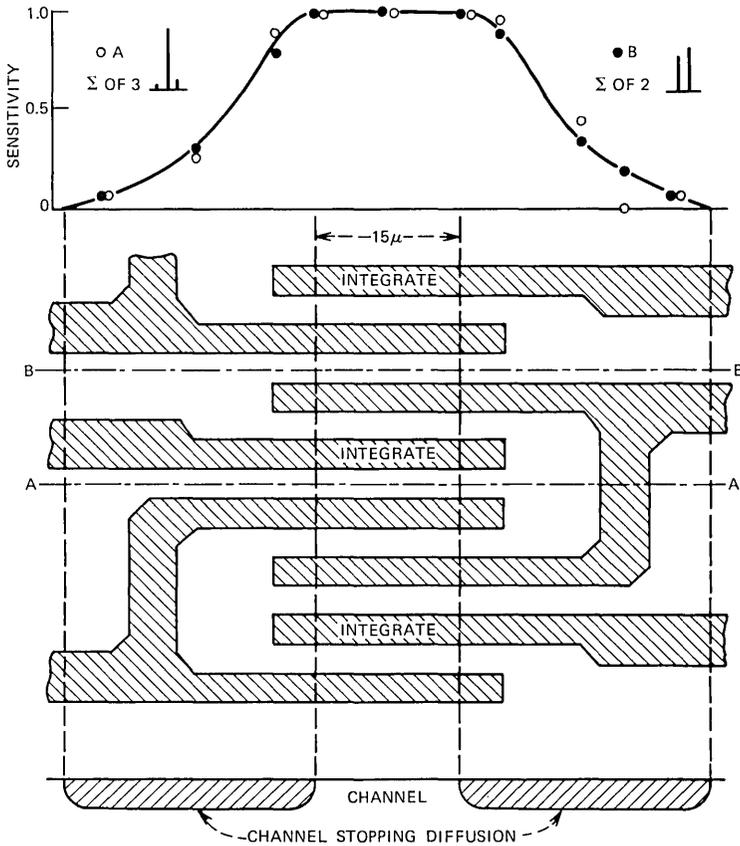
Fig. 18—Normalized light sensitivity measured with a narrow light spot in two different traces across the transfer channel.

the length of a CCD element. Such a device should, therefore, give a much better resolution in the direction of the electronic scan than in the direction of the mechanical scan. This agrees with experimental observations. To improve the vertical resolution of this imaging system to match the horizontal resolution, the light-sensitive line width would have to be confined to about half the element length.

VIII. VARIATIONS IN TECHNOLOGY

Devices with different insulator thickness, electrode length, and metallization have been built to study the influence of these parameters on device behavior. While the effect on signal-handling ability followed a predictable pattern, the influence on transfer inefficiency was often

concealed by the large spread of the observed values. Still some trends became evident.

If operated at the same potentials, a thinner insulator leads generally to a higher signal-handling ability, as expected from the higher capacitance. When the operating voltages are lowered proportionally to the reduced insulator thickness, the devices show poor performance. High potential differences between adjacent transfer electrodes are still necessary to produce fringe fields that can overcome the fixed potential barriers in the gap. With a thinner insulator, the dependence on V_R becomes more critical, and the devices show a narrower operating range. This trend can be compensated for if, at the same time, the gaps are narrowed. With an oxide thickness of 1500 Å, 2 μm seemed to be an appropriate gap width to achieve an average operating range for V_R on the order of 1 V. No significant change in behavior could be attributed to the replacement of 1500 Å of SiO_2 with a double insulator structure consisting of 1200 Å SiO_2 and 500 Å Al_2O_3.

Three different metallizations have been compared: W, Al, Ti-Pd-Ni. No specific difference in performance was observed in devices with chemically etched W or Al electrodes. On devices with a Ti-Pd-Ni metallization on a SiO_2-Al_2O_3 insulator a backsputtering process[7] was used to obtain the required accuracy in the delineation of the transfer electrodes. Figure 19 compares the results of the dynamic probe tests on all operating backsputter delineated devices with a control batch with chemically etched electrodes on the same double insulator. The backsputtered devices showed $n\epsilon$ values that were on the average about a factor of six higher than the values obtained on devices with W or Al electrodes. Among possible causes, differences in the interface state density underneath the electrodes and variations in gap width due to a possible undercutting of the Ti have been ruled out experimentally. Thus, it is conjectured that the backsputter process degrades the integrity of the Si-SiO_2 interface in the region of the gaps where the metal is thinned to within 1000 Å of the insulator surface during backsputtering.[18] In spite of the wide spread of values, this particular trend in $n\epsilon$ was clearly visible, because its origin itself is associated with the transfer gaps, which are the single most significant cause for high transfer inefficiency.

To reduce the sensitivity to the ambient, some slices were protected with a dielectric level of, for example, 1 μm of phosphorous glass or 1000 Å of silicon nitride. While the reaction to such simple tests as "breathing onto the device" was strongly reduced, transfer efficiency did not improve on the average, nor did the slow instabilities disappear.
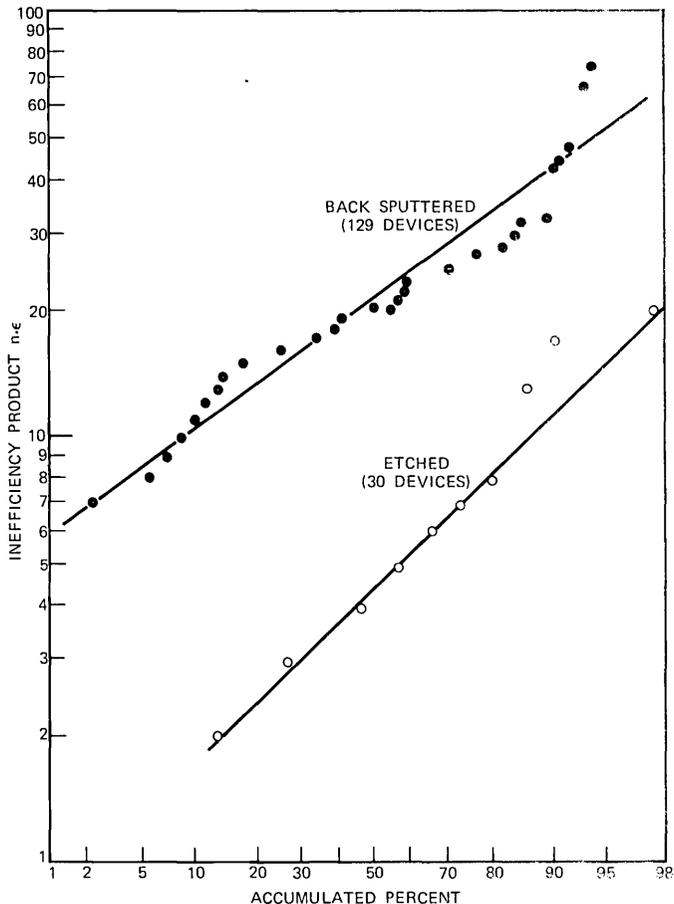
Fig. 19—Plot of $n\epsilon$ distribution for all operating backsputter delineated devices and for a control batch with etched electrodes.

The results generally leave the impression that the condition of the surface of the gate insulator in the gaps during the deposition of the protective dielectric is crucial to the final performance of the device.

In another attempt to control the potential in the gaps, six finished devices with 1500 Å of SiO_2 and 3 μm gaps, which orginally operated with $n\epsilon$ products of 5 to 10, were overcoated with a strip of poly-crystalline silicon, with sheet resistances ranging from 10^8 to 10^{10} ohms/square. In two cases the $n\epsilon$ product improved to values around two and in one case even below one. The performance of the other

three devices worsened and showed some erratic behavior. Still, these results suggest that a properly developed resistive sea might possibly yield a solution to the gap problem.

IX. SUMMARY

Signal-handling ability, transfer efficiency, and dark-current profiles of over a hundred individual linear CCD's have been studied. The total spread of the result is surprisingly wide. Many erratic effects have been observed but not investigated in detail and, therefore, are not fully understood. On several devices which showed reasonably good performance and no erratic behavior, a thorough investigation of the functional dependence of the performance on various parameters has been carried out. These dependences are well understood and can be explained with simple models. Deviations from these simple models as well as the limitations in performance seem to be mainly associated with the gaps between the transfer electrodes. Poorly controlled surface potential can lead to the formation of barriers or pockets that produce poor transfer efficiency, slow instabilities, and nonuniformities in large devices.

Three-phase CCD's with gaps between the electrodes have been an invaluable tool for the investigation of charge coupling and for an early demonstration of charge-coupled image sensors. But the technologies presently used to make these structures do not produce reliable devices with consistent performance. With some technological effort a solution to the gap problem can probably be found, for instance in the application of a resistive sea on the surface of the device. It is questionable, however, if it is worthwhile to put such a development effort into this structure. In big devices the large number of narrow gaps cause a serious reduction in yield. Furthermore, the peripheral structures in more complex devices might require more than one level of metallization. It seems more advantageous to build the actual CCD with overlapping gates and to provide a completely sealed channel.

X. ACKNOWLEDGMENTS

REFERENCES

1. W. S. Boyle and G. E. Smith, "Charge-Coupled Semiconductor Devices," B.S.T.J., *49*, No. 4 (April 1970), pp. 587–593.
2. M. F. Tompsett, G. F. Amelio, and G. E. Smith, "Charge-Coupled 8-bit Shift Register," Appl. Phys. Lett., *17*, 1970, pp. 111–115.
3. M. F. Tompsett, G. F. Amelio, W. J. Bertram, R. R. Buckley, W. J. McNamara, J. C. Mikkelsen, and D. A. Sealer, "Charge-Coupled Imaging Devices: Experimental Results," IEEE Trans. Electron. Devices, *ED-18*, 1971, pp. 992–996.
4. W. J. Bertram, D. A. Sealer, C. H. Séquin, and M. F. Tompsett, "Recent Advances in Charge-Coupled Imaging Devices," IEEE INTERCON Digest of Papers, 1972, pp. 292–293.
5. M. F. Tompsett and E. J. Zimany, "Use of Charge-Coupled Devices for Analog Delay," IEEE Jour. Solid-State Circuit, *SC-8*, 1973, pp. 151–157.
6. G. F. Amelio, "Computer Modeling of Charge-Coupled Device Characteristics," B.S.T.J., *51*, No. 3 (March 1972), pp. 705–730.
7. E. F. Labuda, to be published.
8. W. B. Joyce and W. J. Bertram, "Linearized Dispersion Relation and Green's Function for Discrete-Charge-Transfer Devices with Incomplete Transfer," B.S.T.J., *50*, No. 6 (July-August 1971), pp. 1741–1759.
9. M. F. Tompsett, "The Quantitative Effects of Interface States on the Performance of CCDs," IEEE Transactions on Electron Devices, *ED-20*, 1973, pp. 45–55.
10. W. F. Kosonocky and J. E. Carnes, reported at the ISSCC in Philadelphia, February 17, 1972.
11. N. A. Patrin, A. Bhattacharyya, M. L. Joshi, and J. J. Chang, "Memory Potential of CCD," Nerem 82 Record, 1972, pp. 157–160.
12. L. Boonstra and F. L. J. Sangster, "Progress on Bucket-Brigade Charge Transfer Devices," reported at the ISSCC in Philadelphia, February 17, 1972.
13. C. N. Berglund and H. J. Boll, "Performance Limitations of the IGFET Bucket-Brigade Shift Register," IEEE Trans. Electron Devices, *ED-19*, 1972, pp. 852–860.
14. See, for example, A. J. Grove, *Physics and Technology of Semiconductor Devices*, New York: Wiley, 1967, pp. 298–302.
15. G. F. Amelio, W. J. Bertram, and M. F. Tompsett, "Charge Coupled Imaging Devices: Design Considerations," IEEE Trans. Electron Devices, *ED-18*, 1971, pp. 986–992.
16. F. P. Heiman, "On the Determination of Minority Carrier Lifetime from the Transient Response of an MOS Capacitor," IEEE Trans. Electron Devices, *ED-14*, 1967, pp. 781–784.
17. C. H. Séquin, "Interlacing in Charge-Coupled Imaging Devices," IEEE Trans. Electron Devices, *ED-20*, 1973.
18. B. B. Kosicki, private communication.

Sampled-Data System Approach to Model Time-Division Switches

By C. K. LIU and T. L. WANG

The sampling switch in a time-division switching system is, in general, different from the sampler of sampled-data system theory. A general approach is developed for characterizing such a switch as an ideal sampler plus some modified transfer functions. With this approach, a time-division switching circuit containing a sampling switch can be converted easily to a typical sampled-data system, and the well-established mathematical tools for sampled-data systems, such as the Z-transform, can be applied. In addition, a simplified approach is described that will lead to a very good approximation of the "exact" solution.

I. INTRODUCTION

The transfer function approach developed for sampled-data systems has proved to be a very powerful tool for analyzing time-division systems.[1-3] It yields information useful for both analysis and synthesis of the system. However, its application is often limited due to the fact that the sampling switch in a time-division circuit is different from the sampler of sampled-data-system theory. This difference can be seen from the fact that the voltage at the output side of a sampler in a sampled-data system is always zero between sampling instants, while the voltage at the output side of a sampling switch in a time-division circuit is not necessarily zero between sampling instants, if, for example, the switch is connected to a capacitor. As a consequence, one cannot treat a time-division circuit as a sampled-data system unless the sampling switch can be modeled by a sampler plus a modified system-transfer function.

Of the few who have worked on time-division-system analysis,[1-9] only Desoer[1] has come close to using functional blocks to model a sampling switch, but no general approach has been developed. It is

the purpose of this paper to present a general approach for solving this problem. With this general approach, any time-division circuit containing a sampling switch can be converted to a typical sampled-data system, and the well-established mathematical tools for sampled-data systems, such as the Z-transform, can be applied.

II. FORMULATION

In a sampled-data system, the sampled signal is related to the original signal by a sampling device such as is shown in Fig. 1. The output of the samples is a train of amplitude-modulated pulses. The interval T between the consecutive pulses is called the sampling period, and the pulse width p is referred to as the sampling duration. In the ideal case, we assume that the sampler operates in zero time so that the pulse width p is equal to zero. Then the output of an ideal sampler is a train of amplitude-modulated impulses and is related to the input by

$$v^*(t) = \sum_{n=0}^{\infty} v(nT)\delta(t - nT),\qquad(1)$$

where δ is the Dirac Delta function. We note that whether the operating time of the sampler is zero or not, the sampled voltage is always zero between samplings.

A switch operating periodically in a time-division system is not equivalent to a sampler in a sampled-data system, because the signal at the output side of a switch is not necessarily zero between samplings. However, if an ideal amplifier with zero output and infinite input impedances is added to the switch, as shown in Fig. 2, then the output signal of the amplifier is equal to zero between samplings. In fact, if the sampling duration is much shorter than the sampling period, then an input-output signal relation identical to (1) can be obtained. Thus, if the switch in a time-division circuit is followed by an amplifier, then
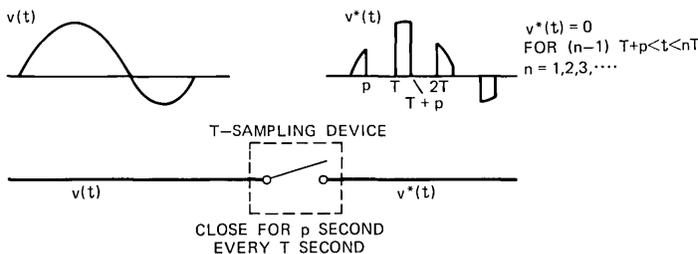
Fig. 1—Sampling device.
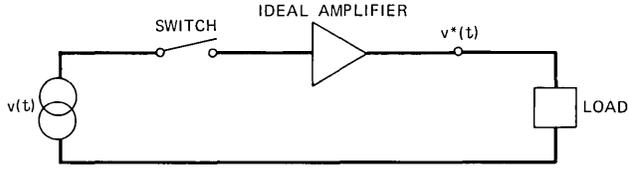
Fig 2—Time-division switch followed by an ideal amplifier.

the sampled-data system techniques are directly applicable. In practice, this is not always the case. Sampling in a time-division system is frequently performed by a switch connected directly to a time-division bus. Our objective is to characterize the switch by a sampler plus some modified system-transfer function so that any time-division system employing periodic sampling can be treated as a standard sampled-data system.

In general, the time-division system we are interested in has the form shown in Fig. 3. It consists of two networks connected by a switch in series with some finite impedance Z_o. The switch is closed periodically for a brief interval of p seconds every T seconds. The smallest time constant of the input signal and the sampling period T are both much greater than p. Referring to Fig. 3, we define $v_{12}(t)$ as the difference between $v_1(t)$ and $v_2(t)$, the voltages at terminals 1 and 2, respectively; $v_{oc}(t)$ is defined as the open circuit Thevenin equivalent voltage at terminal 1 and $i(t)$ as the current in the switch. The current $i(t)$ can be found from the equivalent circuit (Fig. 4) obtained by connecting the driving source $e(t) = v_{oc}(t)$ in series with the time-division switch and impedances Z_0, Z_1, and Z_2, where Z_1 and Z_2 are the output impedance of network 1 and input impedance of network 2, respectively. As the switch is closed only for a time interval from $t = nT$ to
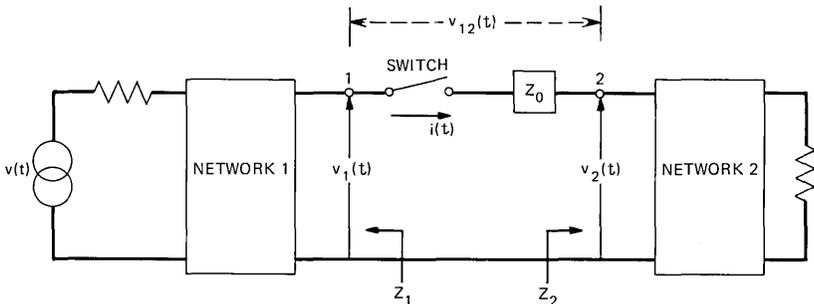
Fig. 3—General time-division system.
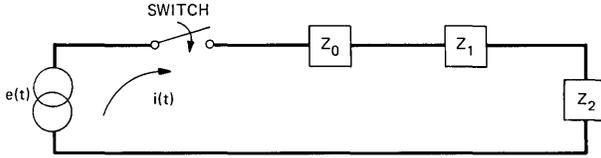
Fig. 4—Equivalent circuit for solving for $g(t)$.

$t = nT + p$, $n = 0, 1, 2, \cdots$, we may express $i(t)$ as

$$i(t) = \begin{cases} i_n(t) & nT \leqq t \leqq nT + p \\ 0 & \text{otherwise} \end{cases} \qquad (2)$$

or

$$i(t) = \sum_{n=0}^{\infty} i_n(t) \cdot [u(t - nT) - u(t - nT - p)], \qquad (3)$$

where $u(t)$ is the unit step-function, T is the sampling period, and p is the sampling duration.

To solve for $i_n(t)$, we let the switch in the equivalent circuit (Fig. 4) close only for a time interval from $t = nT$ to $t = nT + p$. The driving source $e(t)$ should be modified to $v_{oc}(t) - v_{10}(nT^-) - v_{20}(nT^-)$, where

$$\begin{aligned} v_{10}(t) &= v_{oc}(t) - v_1(t) = i(t) \circ z_1(t) \\ v_{20}(t) &= v_2(t) = i(t) \circ z_2(t) \end{aligned} \qquad (4)$$

and $\circ$ denotes the convolution product. Note that $v_{10}(nT^-)$ and $v_{20}(nT^-)$ are the voltages across Z_1 and Z_2 just before the switch is closed. For the small time interval $nT < t < nT + p$, $v_{oc}(t) \simeq v_{oc}(nT)$.[10] Therefore, $e(t) \simeq v_{oc}(nT) - v_{10}(nT^-) - v_{20}(nT^-)$. Defining*

$$\begin{aligned} v_d(t) &= v_{oc}(t) - v_{10}(t^-) - v_{20}(t^-) \\ &= v_{oc}(t) - v_{10}(t - \epsilon) - v_{20}(t - \epsilon), \end{aligned} \qquad (5)$$

where $\epsilon > 0$ is an arbitrarily small number, the current $_ni(t)$ can be expressed as

$$i_n(t) = \mathcal{L}^{-1} \left[\frac{v_d(nT)}{S} \cdot Y(S) \cdot e^{-nTS} \right], \qquad (6)$$

where $Y(S) = 1/[Z_0(S) + Z_1(S) + Z_2(S)]$ is the admittance function of the equivalent circuit and $\mathcal{L}^{-1}$ denotes inverse Laplace trans-

---

* From (5) we note that at the sample instant $t = (nT)$, $v_d(nT) = v_1(nT^-) - v_2(nT^-)$ for $n \geqq 1$, and $v_d(0) = v_{oc}(0) = 0$ for a physically realizable system. Hence, $v_d(nT) = v_{12}(nT^-)$, the difference between $v_1(t)$ and $v_2(t)$ just before the switch is closed.

formation. Substituting (6) into (3), we have

$$i(t) \;=\; \sum_{n=0}^{\infty} v_d(nT)\,\mathcal{L}^{-1}\left[\frac{1}{S}\cdot Y(S)\cdot e^{-nTS}\right]$$

$$\cdot\,[u(t-nT)-u(t-nT-p)]. \quad (7)$$

Equation (7) suggests that we may characterize the switch by a sampler plus a transfer function $G(S)$, as shown in Fig. 5, to relate v_d and i by

$$I(S) \;=\; V_d^*(S)G(S), \quad\quad (8)$$

where $I(S)$, $V_d^*(S)$, and $G(S)$ are the Laplace transforms of $i(t)$, $v_d^*(t)$, and $g(t)$, respectively (similar notations will be used hereafter without explanation). By the definition of impulse-modulated signal,

$$v_d^*(t) \;=\; \sum_{n=0}^{\infty} v_d(nT)\delta(t-nT). \quad\quad (9)$$

In the time domain, eq. (8) corresponds to the convolution integral

$$i(t) \;=\; \int_0^t v_d^*(\tau)g(t-\tau)d\tau. \quad\quad (10)$$

Substituting (9) into (10) and integrating we have

$$i(t) \;=\; \sum_{n=0}^{\infty} v_d(nT)g(t-nT)u(t-nT). \quad\quad (11)$$

Comparing (7) and (11), we have

$$g(t)u(t) \;=\; \mathcal{L}^{-1}\left[\frac{1}{S}\cdot Y(S)\right]\cdot[u(t)-u(t-p)] \quad\quad (12)$$

and

$$G(S) \;=\; \mathcal{L}\left\{\mathcal{L}^{-1}\left[\frac{1}{S}\cdot Y(S)\right]\cdot[u(t)-u(t-p)]\right\}, \quad\quad (13)$$

where $\mathcal{L}$ is the Laplace transformation operator. Equation (13) yields the transfer function we need to characterize the switch. Note that the function $\mathcal{L}^{-1}[(1/S)\,Y(s)]$ is the current $i(t)$ in Fig. 4 with a unit dc driving source. Once $G(S)$ is found, a functional block diagram
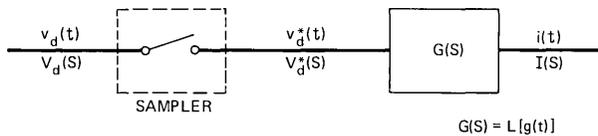
Fig. 5—Characterization of the switch.
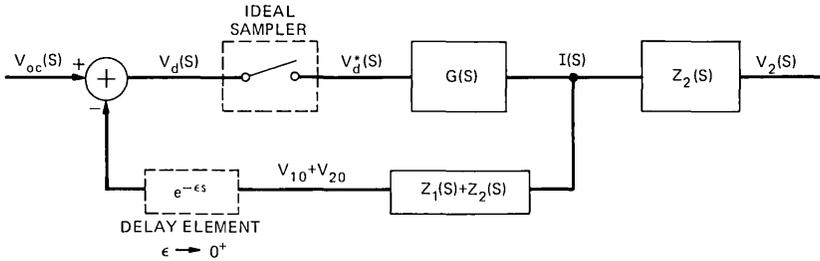
Fig. 6—Transfer function block diagram between V_{oc} and V_2.

describing the signal flow from network 1 to network 2 can be constructed easily, as shown in Fig. 6. Now the system in Fig. 3 is converted to a standard sampled-data system.

The transfer function from v_{oc} to v_2 can be obtained from Fig. 6:

$$V_d(S) = V_{oc}(S) - V_{10}(S)e^{-\epsilon S} - V_{20}(S)e^{-\epsilon S}$$
$$= V_{oc}(S) - I(S)[Z_1(S) + Z_2(S)]e^{-\epsilon S}$$
$$= V_{oc}(S) - V_d^*(S)G(S)[Z_1(S) + Z_2(S)]e^{-\epsilon S}$$
$$V_d^*(S) = V_{oc}^*(S) - V_d^*(S)[GZ_1(S)e^{-\epsilon S} + GZ_2(S)e^{-\epsilon S}]^*$$

or

$$V_d^*(S) = \frac{V_{oc}^*(S)}{1 + [GZ_1(S)e^{-\epsilon S} + GZ_2 e^{-\epsilon S}]^*}, \tag{14}$$

where $GZ_i(S) = G(S)Z_i(S)$ and $i = 1, 2$, and $\epsilon > 0$ is arbitrarily small. From (14),

$$V_2(S) = I(S)Z_2(S) = V_d^*(S)G(S)Z_2(S)$$

$$= \frac{GZ_2(S)}{1 + [GZ_1(S)e^{-\epsilon S} + GZ_2(S)e^{-\epsilon S}]^*} V_{oc}^*(S). \tag{15}$$

To find $[GZ_i(S)e^{-\epsilon S}]^*$, we need to know the relationship between $gz_i^*(t)$ and $[gz_i(t - \epsilon)]^*$. The function $g(t)$ is defined in (12):

$$g(t) = \mathcal{L}^{-1} \left\{ \frac{1}{S[Z_0(S) + Z_1(S) + Z_2(S)]} \right\} [u(t) - u(t - p)]$$
$$= h(t)u(t) - h(t)u(t - p),$$

where

$$h(t) = \mathcal{L}^{-1} \left\{ \frac{1}{S[Z_0(S) + Z_1(S) + Z_2(S)]} \right\}.$$

Note that $h(t)$ is the step-response of a linear passive network and, thus,

is continuous for all $t > 0$. Now

$$gz_i(t) = \int_{-\infty}^{t} g(\tau)z_i(t - \tau)d\tau$$

$$= \int_{-\infty}^{t} h(\tau)z_i(t - \tau)u(\tau)d\tau - \int_{-\infty}^{t} h(\tau)z_i(t - \tau)u(\tau - p)d\tau.$$

From the convolution of two continuous time functions, $gz_1(t)$ will be continuous for all $0 < t < p$ and $t > p$. Since $p < T$, $gz_i(t)$ will be continuous at $t = nT$ for all $n \geqq 1$, i.e.,

$$gz_i(nT - \epsilon) = gz_i(nT) \tag{16}$$

as ϵ approaches zero. At $t = 0$, since $gz_i(t) = 0$ for all $t < 0$, we have

$$gz_i(0 - \epsilon) \triangleq 0. \tag{17}$$

From (16) and (17),

$$\sum_{n=0}^{\infty} gz_i(nT - \epsilon)\delta(t - nT) = \sum_{n=1}^{\infty} gz_i(nT)\delta(t - nT)$$

$$= \sum_{n=0}^{\infty} gz_i(nT)\delta(t - nT) - gz_i(0)\delta(t)$$

for arbitrarily small ϵ. Therefore,

$$[gz_i(t - \epsilon)]^* = gz_i^*(t) - gz_i(0)\delta(t) \tag{18}$$

and

$$[GZ_i(S)e^{-\epsilon S}]^* = GZ_i^*(S) - gz_i(0). \tag{19}$$

From (15) and (19), we have:

$$\frac{V_2(S)}{V_{oc}^*(S)} = \frac{GZ_2(S)}{1 - gz_1(0) - gz_2(0) + GZ_1^*(S) + GZ_2^*(S)} \tag{20}$$

and

$$\frac{V_2^*(S)}{V_{oc}^*(S)} = \frac{GZ_2^*(S)}{1 - gz_1(0) - gz_2(0) + GZ_1^*(S) + GZ_2^*(S)} . \tag{21}$$

If we are interested in the transfer function $V_2(S)/V_1^*(S)$, then since $V_1(S) = V_{oc}(S) - I(S)Z_1(S)$, we have:

$$V_1^*(S) = V_{oc}^*(S) - V_d^*(S)GZ_1^*(S). \tag{22}$$

Substituting (14) into (22),

$$V_1^*(S) = V_{oc}^*(S) \frac{1 + GZ_2^*(S) - gz_1(0) - gz_2(0)}{1 + GZ_1^*(S) + GZ_2^*(S) - gz_1(0) - gz_2(0)} . \tag{23}$$

From (20), (21), and (23), we have:

$$\frac{V_2(S)}{V_1^*(S)} = \frac{GZ_2(S)}{1 - gz_1(0) - gz_2(0) + GZ_2^*(S)} \tag{24}$$

and

$$\frac{V_2^*(S)}{V_1^*(S)} = \frac{GZ_2^*(S)}{1 - gz_1(0) - gz_2(0) + GZ_2^*(S)}. \tag{25}$$

Since $gz_i(0) = \lim_{S \to \infty} S \cdot GZ_i(S)$, it will be equal to zero when the function $GZ_i(S)$ has at least two more poles then zeros. In such cases, (20), (21), (24), and (25) become

$$\frac{V_2(S)}{V_{oc}^*(S)} = \frac{GZ_2(S)}{1 + GZ_1^*(S) + GZ_2^*(S)} \tag{20A}$$

$$\frac{V_2^*(S)}{V_{oc}^*(S)} = \frac{GZ_2^*(S)}{1 + GZ_1^*(S) + GZ_2^*(S)} \tag{21A}$$

$$\frac{V_2(S)}{V_1^*(S)} = \frac{GZ_2(S)}{1 + GZ_2^*(S)} \tag{24A}$$

$$\frac{V_2^*(S)}{V_1^*(S)} = \frac{GZ_2^*(S)}{1 + GZ_2^*(S)}. \tag{25A}$$

As a simple example, let us refer to Fig. 7, which shows an ideal sample-and-hold circuit with a capacitor C. For this circuit $Z_0 = Z_1 = 0$, $Z_2 = 1/CS$, and it can easily be found that $G(S) = C$. Therefore,

$$G(S)Z_2(S) = \frac{1}{S}. \tag{26}$$

Since it has only one more pole than it has zeros, eq. (24) should be used to find $V_2(S)/V_1^*(S)$:

$$\frac{V_2(S)}{V_1^*(S)} = \frac{1/S}{[1 - u(0)] + [1/S]^*} = \frac{1 - e^{-TS}}{S}. \tag{27}$$

The general approach used to characterize the time-division switch can be summarized as follows:

(i) Form an equivalent circuit (Fig. 4) by using a unit step voltage source to drive the impedances Z_0, Z_1, and Z_2 connected in series, where Z_1 is the output impedance of $N1$ and Z_2 is the input impedance of $N2$.

(ii) Solve for the current $i(t)$ in the equivalent circuit.

(iii) Let $g(t) = i(t)[u(t) - u(t - p)]$. Calculate $G(S) = \mathcal{L}\{g(t)\}$ and $GZ_i(S) = G(S)Z_i(S)$, $i = 1, 2$.
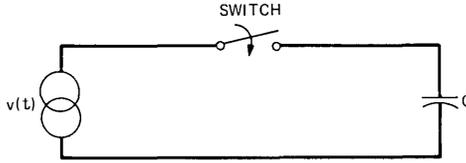
Fig. 7—Ideal sample-and-hold circuit.

(*iv*) Now the energy transfer between $N1$ and $N2$ can be described by a sampler plus some transfer functions. Either one of the following formulas may be used:

$$\frac{V_2(S)}{V_{oc}^*(S)} = \frac{GZ_2(S)}{1 - gz_1(0) - gz_2(0) + GZ_1^*(S) + GZ_2^*(S)} \qquad (28)$$

$$\frac{V_2(S)}{V_1^*(S)} = \frac{GZ_2(S)}{1 - gz_1(0) - gz_2(0) + GZ_2^*(S)}, \qquad (29)$$

where v_{oc} is the open circuit voltage at terminal 1 in Fig. 3. When the function $GZ_i(S)$ has at least two more poles than zeros, we know immediately that $gz_i(0) = 0$ and can be removed from the above formulas.

III. AN APPLICATION

The switch we model here is a practically realizable sample-and-hold switch for a time-division switching system. It is shown in Fig. 8. The series resistor R represents the gate resistance during sampling. The series inductor represents the lead inductance whose value depends on the bus structure.

For this circuit, $Z_1 = 0$, $Z_2 = 1/CS$, and $Z_0 = SL + R$. Therefore, $g(t)$ can be found by solving a simple series RLC circuit with unit dc input and a switch closed at $t = 0$ and open at $t = p$ for the rest of the time. The result is

$$g(t) = \frac{1}{2L\sqrt{\alpha^2 - \omega_0^2}} \, (e^{-\beta_1 t} - e^{-\beta_2 t})[u(t) - u(t - p)], \qquad (30)$$

GATE RESISTANCE R=30 Ω
LEAD INDUCTANCE L
SAMPLE-AND-HOLD CAPACITANCE C = 1000pF
SAMPLING PERIOD T = 83.3 μs
SAMPLING INTERVAL p = 300ns
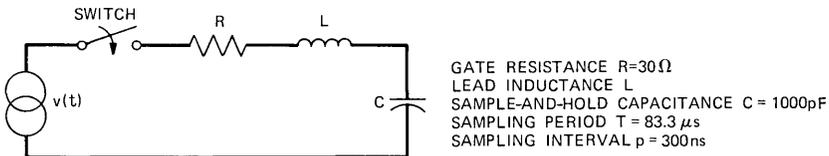
Fig. 8—Practical sample-and-hold switch.

where $\alpha = R/2L$, $\omega_0 = 1/\sqrt{LC}$, $\beta_1 = \alpha - \sqrt{\alpha^2 - \omega_0^2}$, and $\beta_2 = \alpha + \sqrt{\alpha^2 - \omega_0^2}$. From (30) we have:

$$G(S) = \frac{1}{2L\sqrt{\alpha^2 - \omega_0^2}} \left\{ \frac{1 - e^{-p(S+\beta_1)}}{S + \beta_1} - \frac{1 - e^{-p(S+\beta_2)}}{S + \beta_2} \right\}. \qquad (31)$$

The transfer function $H(S) = V_2(S)/V_1^*(S)$ can now be found from

$$H(S) = \frac{V_2(S)}{V_1^*(S)} = \frac{G(S)Z_2(S)}{1 + [G(S)Z_2(S)]^*}. \qquad (32)$$

The calculation of $G(S)Z_2(S)$ and $[G(S)Z_2(S)]^*$ is given in the appendix. With $G(S)Z_2(S)$ and $[G(S)Z_2(S)]^*$ known, we have

$$H(S) = \frac{1 - e^{-TS}}{S} \cdot \frac{\beta_1\beta_2}{\beta_1 - \beta_2} \left[\frac{1 - e^{-p(S+\beta_1)}}{S + \beta_1} - \frac{1 - e^{-p(S+\beta_2)}}{S + \beta_2} \right]$$
$$\cdot \frac{1}{1 + ke^{-TS}}, \qquad (33)$$

where

$$k = \frac{\beta_1 e^{-\beta_2 p} - \beta_2 e^{-\beta_1 p}}{\beta_2 - \beta_1}. \qquad (34)$$

From (33),

$$H^*(S) = \frac{V_2^*(S)}{V_1^*(S)}$$

$$= \frac{1}{1 + \dfrac{1 - e^{-TS}}{e^{-TS}} \cdot \dfrac{1}{1 + k}}$$

$$= \frac{1 + k}{e^{TS} + k}. \qquad (35)$$

When the driving function is sinusoidal, we have:

$$H^*(j\omega) = \frac{1 + k}{e^{j\omega T} + k} \qquad (36)$$

$$|H^*(j\omega)| = \frac{1 + k}{\sqrt{1 + k^2 + 2k \cos \omega T}}. \qquad (37)$$

This shows that, for $k \neq 0$, the magnitude of the voltage gain at the sampling instant will be a function of frequency. The maximum (or minimum, depending upon the sign of k^*) occurs at $\omega T = \pi$, i.e., the

---

* It can be easily verified from (34) that k will be positive only if the equivalent RLC circuit is under damped.

f = fs/16

f = fs/8
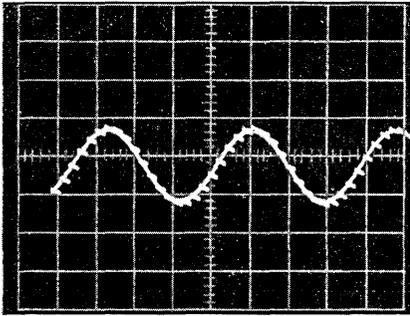
OUTPUT SIGNAL
INPUT SIGNAL
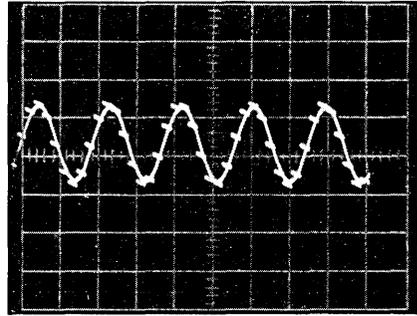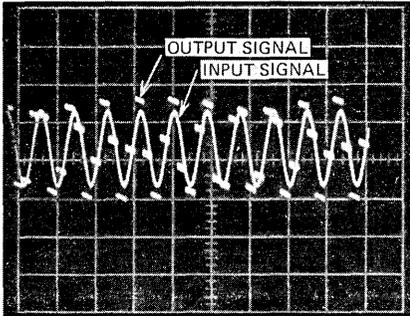
f = fs/4
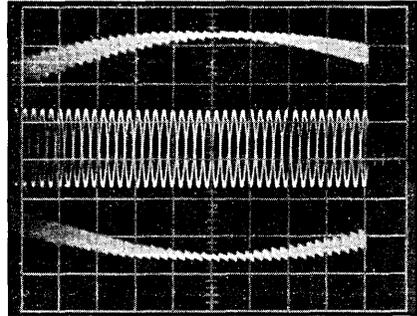
f = fs/2

Fig. 9—Variation of magnitude of output signal at sampling gate with respect to input frequency f.

half sampling frequency:

$$|H^*(j\omega)|_{\text{extreme}} = \frac{1 + k}{|1 - k|}. \tag{38}$$

Figure 9 shows the laboratory observation of such an effect for $k \simeq \frac{1}{3}$.

When the lead inductance is negligibly small, i.e., $L \to 0$, then $\beta_2 \to \infty$ and $\beta_1 \to 1/RC$. Equations (33), (35), and (37) now become:

$$H(S) = \frac{1 - e^{-TS}}{S} \frac{a}{S + a} \frac{1 - e^{-ap}e^{-pS}}{1 - e^{-ap}e^{-TS}} \tag{39}$$

$$H^*(S) = \frac{1 - e^{-ap}}{e^{TS} - e^{-ap}} \tag{40}$$

$$|H^*(j\omega)| = \frac{1 - e^{-ap}}{\sqrt{1 + e^{-2ap} - 2e^{-ap} \cos \omega T}}, \tag{41}$$

where $a = 1/RC$.

In this case, the RLC-series circuit reduces to an RC circuit. In a practical time-division-switching system, the typical values might be: sampling period $T = 83.3$ μs (sampling at 12 kHz), sampling duration $p = 300$ ns, hold-capacitor $C = 1000$ pF, and gate resistance $R = 30$ Ω. A simple calculation will show that

$$ap = \frac{p}{RC} = \frac{300 \times 10^{-9}}{30 \times 10^{-9}} = 10. \tag{42}$$

Hence, $e^{-ap} = e^{-10} \simeq 0$ and the transfer function $H(S)$ in (39) becomes

$$H(S) \simeq \frac{1 - e^{-TS}}{S} \cdot \frac{\dfrac{1}{RC}}{S + \dfrac{1}{RC}}, \tag{43}$$

which indicates that the switch-and-hold circuit can be approximately considered as an ideal sample-and-hold device in series with an RC circuit, as shown in Fig. 10.

It is also interesting to note that if both the gate resistance R and the lead inductance L approach 0, we will have an ideal sample-and-hold switch. From (43), we can see that $H(S)$ will approach the ideal sample-and-hold transfer function $1 - e^{-TS}/S$ as expected.

IV. AN APPROXIMATION

In this section, we shall present a simplified approach which in general leads to a very good approximation of the results found by the general approach described in Section III. The basic idea here is to approximate the current $i(t)$ in the switch by an impulse-modulated signal,

$$i(t) \simeq \tilde{\imath}(t), \tag{44}$$

and characterize the switch by the energy transfer during the sampling duration:

$$v_2(nT^+) = v_2(nT^-) + \gamma[v_1(nT^-) - v_2(nT^-)], \tag{45}$$
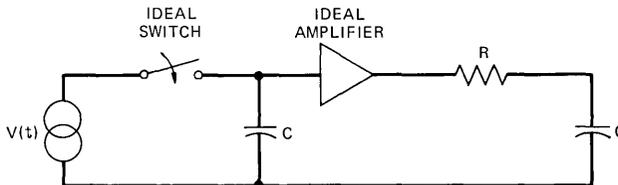
Fig. 10—Equivalent circuit for Fig. 8 when L approaches 0.

where nT^- represents the instant just before the switch is closed, and nT^+ the instant just after the switch is reopened. The determination of γ, which is related to the transfer loss, will be discussed later. From (45) we have:

$$v_2^*(t) = \gamma[v_1(t - \epsilon)]^* + (1 - \gamma)[v_2(t - \epsilon)]^* \tag{46}$$

and†

$$V_2^*(S) = \gamma[V_1(S)e^{-\epsilon S}]^* + (1 - \gamma)[V_2(S)e^{-\epsilon S}]^*, \tag{47}$$

where ϵ is arbitrarily small.

Substituting

$$V_1(S) = V_{oc}(S) - \tilde{I}^*(S)Z_1(S)$$

and

$$V_2(S) = \tilde{I}^*(S)Z_2(S)$$

into (47), we have

$$\tilde{I}^*(S)Z_2^*(S) = \gamma[V_{oc}(S)e^{-\epsilon S}]^* + \tilde{I}^*(S)\{-\gamma[Z_1(S)e^{-\epsilon S}]^* \\ + (1 - \gamma)[Z_2(S)e^{-\epsilon S}]^*\}. \tag{48}$$

As $V_{oc}(t)$ is continuous for all $t \geq 0$, and $z_i(t)$ is continuous for all $t > 0$,

$$\tilde{I}^*(S)Z_2^*(S) = \gamma V_{oc}^*(S) + \tilde{I}^*(S)\{-\gamma[Z_1^*(S) - z_1(0)] \\ + (1 - \gamma)[Z_2^*(S) - z_2(0)]\}$$

or

$$\tilde{I}^*(S) = \frac{\gamma V_{oc}^*(S)}{\gamma[Z_1^*(S) + Z_2^*(S)] - \gamma z_1(0) - (\gamma - 1)z_2(0)}. \tag{49}$$

Therefore,

$$\frac{V_2(S)}{V_{oc}^*(S)} = \frac{\tilde{I}^*(S)Z_2(S)}{V_{oc}^*(S)}$$

$$= \frac{\gamma Z_2(S)}{\gamma[Z_1^*(S) + Z_2^*(S)] - \gamma z_1(0) - (\gamma - 1)z_2(0)}, \tag{50}$$

and

$$\frac{V_2(S)}{V_1^*(S)} = \frac{\tilde{I}^*(S)Z_2(S)}{V_{oc}^*(S) - \tilde{I}^*(S)Z_1^*(S)}$$

$$= \frac{\gamma Z_2(S)}{\gamma Z_2^*(S) - \gamma z_1(0) - (\gamma - 1)z_2(0)}. \tag{51}$$

---

† Note that as $v_i(t)$ may not be continuous at $t = nT$, $n = 0, 1, \cdots$,

$$[V_i(S)e^{-\epsilon S}]^* \neq V_i^*(S) - v_i(0).$$

To determine the constant γ, we note that the change in v_2 during the sampling duration is a function of the current in the sampling switch:

$$\Delta v_2(nT) = v_2(nT^+) - v_2(nT^-)$$
$$= i_n(t) \cdot z_2(t) |_{t=nT+p}. \tag{52}$$

As mentioned earlier, $i_n(t)$ can be solved from Fig. 4 with the driving source $e(t) = v_d(nT)$ and the switch closed at $t = nT$. Since $v_{oc}(t)$ is continuous for all $t \geqq 0$, $v_d(nT) = v_{12}(nT^-) = v_1(nT^-) - v_2(nT^-)$. Hence, in the equivalent circuit of Fig. 4, we let $e(t) = v_{12}(nT^-)$ and close the switch at $t = nT$. Solving this circuit, the current will be $i_n(t)$ and the voltage across Z_2 at $t = nT + p$ will be $\Delta v_2(nT)$; i.e.,

$$\Delta v_2(nT) = \mathcal{L}^{-1}\left\{ \mathcal{L}\left[\frac{v_{12}(nT^-)}{S} \cdot Y(S) \cdot e^{-nTS} \cdot Z_2(S) \right] \right\}_{t=nT+p}, \tag{53}$$

where $Y(S) = 1/[Z_0(S) + Z_1(S) + Z_2(S)]$. From (53),

$$\gamma = \frac{\Delta v_2(nT)}{v_{12}(nT^-)}$$

$$= \mathcal{L}^{-1}\left\{ \mathcal{L}\left[\frac{1}{S} \cdot Y(S) \cdot e^{-nTS} \cdot Z_2(S) \right] \right\}_{t=nT+p}$$

$$= \mathcal{L}^{-1}\left\{ \mathcal{L}\left[\frac{1}{S} \cdot Y(S) \cdot Z_2(S) \right] \right\}_{t=p}. \tag{54}$$

Hence, in Fig. 4, if we let $e(t) = u(t)$ and close the switch at $t = 0$, then the voltage across Z_2 at $t = p$ will be the value of γ. After γ is found, either (50) or (51) may be used to enable us to replace the switch by an ideal sampler plus a transfer function, as shown in Fig. 11.

To illustrate how this approach works, let us return to the practical sample-and-hold switch in Fig. 8. As stated in the last section, $Z_1 = 0$, $Z_0 = R + SL$, and $Z_2 = 1/CS$. Solving the series RLC circuit with the

$$H(S) = \frac{\gamma Z_2(S)}{\gamma[Z_1^*(S)+Z_2^*(S)] - \gamma z_1(0) - (\gamma-1) z_2(0)}$$
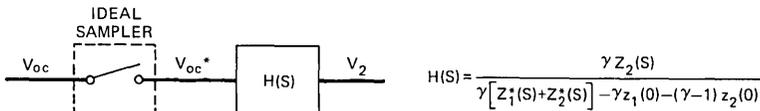
Fig. 11—Transfer function diagram for the approximate approach.

driving source of $u(t)$, the current $i_0(t)$ is found as shown in (30):

$$i_0(t) = \frac{1}{2L\sqrt{\alpha^2 - \omega_0^2}} (e^{-\beta_1 t} - e^{-\beta_2 t}). \tag{55}$$

Now γ, the voltage across Z_2 at $t = p$, can be found:

$$\gamma = \frac{1}{C} \int_0^p i_0(t)dt$$

$$= \frac{\beta_1\beta_2}{\beta_2 - \beta_1} \left[\frac{1}{\beta_1} (1 - e^{-\beta_1 p}) - \frac{1}{\beta_2} (1 - e^{-\beta_2 p}) \right]$$

$$= 1 + k, \tag{56}$$

where

$$k = \frac{\beta_1 e^{-\beta_2 p} - \beta_2 e^{-\beta_1 p}}{\beta_2 - \beta_1}$$

is the same k given by (34) in the last section. Therefore,

$$\tilde{H}(S) = \frac{V_2(S)}{V_1^*(S)} = \frac{(1 + k) \dfrac{1}{CS}}{(1 + k) \left[\dfrac{1}{CS} \right]^* - \dfrac{k}{C}}$$

$$= \frac{1 - e^{-TS}}{S} \frac{1 + k}{1 + ke^{-TS}}. \tag{57}$$

We now want to show that $\tilde{H}(S)$ of (57) is a good approximation of $H(S)$ of (33). From (33), we have:

$$H(S) = \frac{1 - e^{-TS}}{S} \frac{F(S)}{1 + ke^{-TS}}, \tag{58}$$

where

$$F(S) = \frac{\beta_1\beta_2}{\beta_2 - \beta_1} \left[\frac{1 - e^{-p(S+\beta_1)}}{S + \beta_1} - \frac{1 - e^{-p(S+\beta_2)}}{S + \beta_2} \right]. \tag{59}$$

To show that $H(S) \simeq \tilde{H}(S)$, we want to show that $F(S) \simeq 1 + k$. Since p is small, we have:

$$F(S) \simeq \frac{\beta_1\beta_2}{\beta_2 - \beta_1} \left[p - \frac{p^2}{2} (S + \beta_1) - p + \frac{p^2}{2} (S + \beta_2) \right]$$

$$= \frac{\beta_1\beta_2}{2} p^2,$$

and

$$1 + k = 1 + \frac{\beta_1 e^{-\beta_2 p} - \beta_2 e^{-\beta_1 p}}{\beta_2 - \beta_1} = \frac{\beta_2(1 - e^{-\beta_1 p}) - \beta_1(1 - e^{-\beta_2 p})}{\beta_2 - \beta_1}$$

$$\simeq \frac{1}{\beta_2 - \beta_1} \left[\beta_2 \left[\beta_1 p - \frac{\beta_1^2 p^2}{2} \right] - \beta_1 \left[\beta_2 p - \frac{\beta_2^2 p^2}{2} \right] \right]$$

$$= \frac{\beta_1 \beta_2}{2} p^2.$$

Hence, $F(S) \simeq 1 + k$ and $H(S) \simeq \tilde{H}(S)$.

Finally, we note that as Z_0 approaches zero, the current $i(t)$ does approach an impulse-modulated function. Thus, the approach described in this section will always lead to the true answer when $Z_0 = 0$. For example, when R, $L \to 0$, the switch we modeled above becomes the ideal sample-and-hold switch. In this case $k \to 0$ and $\tilde{H}(S)$ of (57) approaches $1 - e^{-TS}/S$, the familiar ideal sample-and-hold transfer function. It can also be easily seen that if we start with this ideal sample-and-hold switch, i.e., $Z_0 = Z_1 = 0$ and $Z_2 = 1/CS$, then $\gamma = 1$ and $Z_2^*(S) = 1/C(1 - e^{-TS})$. From (51), we shall again have $V_2(S)/V_1^*(S) = 1 - e^{-TS}/S$ as expected.

V. ACKNOWLEDGMENT

APPENDIX

Calculation of $G(S)Z_2(S)$ and $[G(S)Z_2(S)]$ *

Since $Z_2(S) = 1/CS$, $\beta_1 = \alpha - \sqrt{\alpha^2 - \omega_0^2}$, $\beta_2 = \alpha + \sqrt{\alpha^2 - \omega_0^2}$, and $\omega_0^2 = 1/LC$, we have:

$$GZ(S) = G(S)Z_2(S) = \frac{1}{LC} \frac{1}{2\sqrt{\alpha^2 - \omega_0^2}} \cdot \frac{1}{S}$$

$$\cdot \left\{ \frac{1 - e^{-p(S+\beta_1)}}{S + \beta_1} - \frac{1 - e^{-p(S+\beta_2)}}{S + \beta_2} \right\}$$

$$= \frac{\beta_1 \beta_2}{\beta_2 - \beta_1} \cdot \frac{1}{S} \left\{ \frac{1 - e^{-p(S+\beta_1)}}{S + \beta_1} - \frac{1 - e^{-p(S+\beta_2)}}{S + \beta_2} \right\} \qquad (60)$$

and

$$GZ^*(S) = \frac{\beta_1 \beta_2}{\beta_2 - \beta_1} \left\{ \sum_{i=1}^{2} \frac{(-1)^{i+1}}{\beta_i} \right.$$

$$\left. \times \left[\frac{1}{S} - \frac{e^{-p(S+\beta_i)}}{S} - \frac{1}{S + \beta_i} + \frac{e^{-p(S+\beta_i)}}{S + \beta_i} \right]^* \right\}.$$

Hence,

$$GZ^*(S) = \frac{\beta_1\beta_2}{\beta_2 - \beta_1} \left\{ \sum_{i=1}^{2} \frac{(-1)^{i+1}}{\beta_i} \left[\frac{1}{1 - e^{-TS}} - \frac{e^{-\beta_i p}e^{-TS}}{1 - e^{-TS}} \right. \right.$$

$$\left. \left. - \frac{1}{1 - e^{-TS}e^{-\beta_i T}} + \frac{e^{-\beta_i T}e^{-TS}}{1 - e^{-TS}e^{-\beta_i T}} \right] \right\}$$

$$= \frac{\beta_1\beta_2}{\beta_2 - \beta_1} \sum_{i=1}^{2} \frac{(-1)^{i+1}}{\beta_i} \left[\frac{1 - e^{-\beta_i p}e^{-TS}}{1 - e^{-TS}} - 1 \right]$$

$$= \frac{\beta_1\beta_2}{\beta_2 - \beta_1} \cdot \frac{e^{-TS}}{1 - e^{-TS}} \left[\frac{1 - e^{-\beta_1 p}}{\beta_1} - \frac{1 - e^{-\beta_2 p}}{\beta_2} \right]$$

$$= \frac{1}{\beta_2 - \beta_1} \cdot \frac{e^{-TS}}{1 - e^{-TS}} \left[\beta_2 - \beta_1 + \beta_1 e^{-\beta_2 p} - \beta_2 e^{-\beta_1 p} \right]$$

$$= (1 + k) \frac{e^{-TS}}{1 - e^{-TS}}, \tag{61}$$

where

$$k = \frac{\beta_1 e^{-\beta_2 p} - \beta_2 e^{-\beta_1 p}}{\beta_2 - \beta_1}. \tag{62}$$

REFERENCES

1. C. A. Desoer, "A Network Containing a Periodically Operated Switch Solved by Successive Approximations," B.S.T.J., *36*, No. 6 (November 1957), pp. 1403–1428.
2. K. W. Cattermole, "Efficiency and Reciprocity in Pulse Amplitude Modulation, Part I—Principles," Proc. IEE (London), *105*, Part B (September 1958), pp. 449–462.
3. D. B. James and T. L. Wang, unpublished work, 1971.
4. W. R. Bennett, "Steady-State Transmission Through Networks Containing Periodically Operated Switches," IRE Trans. on Circuit Theory, *2*, No. 1 (March 1955), pp. 17–22.
5. T. H. Crowley, unpublished work, 1956.
6. C. A. Desoer, "Transmission Through a Linear Network Containing a Periodically Operating Switch," IRE WESCON Convention Record, *2*, Part 2 (1958), pp. 34–41.
7. G. B. Thomas, "Synthesis of Input and Output Networks for a Resonant Transfer Gate," IRE Int. Conv. Record, *9*, Part 9 (1961), pp. 236–243.
8. A. Fettweis, "Steady-State Analysis of Circuits Containing a Periodically Operated Switch," IRE Trans. Circuit Theory, *CT-6* (September 1959), pp. 252–260.
9. M. L. Liou and F. R. Mastromonaco, "Exact Analysis of Linear Circuits Containing a Periodically Operated Switch Using the State Space Approach," *Symp. on Circuit Theory*, 1968.
10. B. C. Kuo, *Discrete-Data Control Systems*, Englewood Cliffs, N.J.: Prentice-Hall, 1970, p. 16.

An Optical Repeater With High-Impedance Input Amplifier

By J. E. GOELL

A 6.3-Mb/s repeater for fiber optic communication systems is described which incorporates a high-impedance input amplifier. It is shown that by utilizing an input circuit with a time constant which is long compared to the bit interval and equalizing after the signal has been sufficiently amplified to set the signal-to-noise ratio, thermal noise can be decreased. As a result, a reduction can be realized in the required signal and, with an avalanche detector, in the optimum gain.

The repeater, which was realized in a compact form employing standard integrated circuits, utilizes a GaAs light-emitting diode as its optical source. Other features include automatic gain and threshold controls and recovered timing.

I. INTRODUCTION

Digital communication systems utilizing low-loss optical fibers are presently being investigated. The realization of fibers with losses as low as 4 dB/km[1] has opened the way for numerous applications. System configurations will depend on such factors as fiber dispersion, fiber cost, desired information capacity, and terminal costs. Transmission rates near 6.3 Mb/s are attractive because fiber group delay dispersion is not expected to be a problem even with an incoherent source, and a wide variety of low-cost integrated circuits are applicable. If, as now appears likely, the fiber cost is low, then space multiplex could be an attractive alternative to time multiplex in achieving high capacity.

The repeater described here incorporates a high-impedance input amplifier which is similar in approach to ones that have been used for other applications incorporating a capacitive detector such as nuclear particle counters[2] and television cameras.[3] As a result of the high-input

impedance, the power required to achieve a specified error rate is reduced, as is the optimum gain if an avalanche detector is used. The latter advantage is important since it eases the fabrication of the detector diode and increases its thermal stability.

The repeater employs return-to-zero pulses with 50-percent duty cycle. The only word pattern restriction is that an occasional "one" be included so timing can be recovered and signal level can be determined. The optical signal is generated by a gallium arsenide light-emitting diode (LED) operating at 0.9 μ wavelength.[4] Diodes of the type employed have been built with output powers of up to about 5 mW. Optical powers to 1 mW have been coupled from these diodes into a fiber with a 0.63 numerical aperture. The repeater was fabricated in a compact form using standard integrated circuits. Automatic gain and threshold controls were provided so the optical input power could vary over a wide range. Clamping was employed to prevent baseline wander with an unbalanced data content of the signal, and timing was extracted by a phase-locked loop.

II. THEORY

A typical circuit for a photodiode driving an input amplifier is shown in Fig. 1(a) and its equivalent circuit in Fig. 1(b). The current generator i_d is the photo current. R_r is the dc return resistor for the detector, and i_r is the noise generator associated with it. The capacitor C_d is the
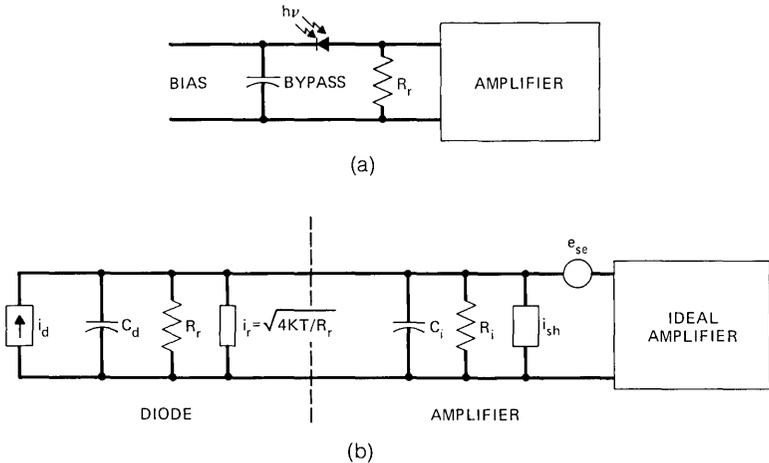
Fig. 1—(a) Input circuit. (b) Equivalent circuit.

output capacitance of the diode, C_i is the input capacitance to the amplifier (excluding feedback effects), and R_i is the input resistance of the amplifier. The quantities e_{se}^2 and i_{sh}^2 are the spectral noise densities of the series voltage and shunt current generators which characterize the noise properties of the amplifier.

A common approach to circuit design has been to set

$$\tau = RC$$
$$< 1/\text{baud rate},$$

where

$$R = \frac{R_r R_i}{R_r + R_i}$$

and

$$C = C_d + C_i$$

to minimize noise while not introducing significant intersymbol interference.

To achieve this, i_r and i_{sh} are often major sources of noise. Therefore, from the standpoint of noise, it is preferable to make R_r very large and employ an amplifier with low i_{sh}, even if $\tau \gg$ baud interval, to amplify the signal sufficiently to set the signal-to-noise ratio, and then to equalize the resulting distortion to eliminate intersymbol interference.

Two possible limitations to the high-impedance approach exist, both related to the low-frequency component of the signal developed across the detector. The difference in voltage between a long string of "ones" and a long string of "zeros" is proportional to the dc load resistance on the diode. Thus, the required dynamic range of the amplifiers preceding the equalizer increases with increasing R. Furthermore, with an avalanche detector this voltage could change the avalanche gain since it is in series with the diode bias. These two factors, which increase in importance as baud rate decreases, ultimately limit the magnitude of the detector load resistance.

In the remainder of this section the relationships between error rate, signal power, and the circuit parameters are discussed for a binary signal with both states equally likely. It is assumed that Gaussian noise statistics apply, that dark current is negligible, and that the amplifiers preceding the equalizer are linear.

Personick[5] has shown that for a pulse of average power p_o with avalanche gain of mean square $\langle g^2 \rangle$, if we assume the optical pulses are distinct, the ratio of the pulse peak to the root mean square thermal noise in the baseband circuit is equal to the ratio of the average cur-

rent of the received pulse to the square root of the quantity

$$I^2(p_o) = \frac{\eta e^2 p_0 f_b \langle g^2 \rangle}{h\nu} + n_t, \tag{1}$$

where n_t, the mean square thermal noise current* weighted to correct for input and output pulse shape, is given by

$$n_t = \left[\left(i_{sh}^2 + \frac{2kT}{R_r} + \frac{e_{se}^2}{R^2} \right) I_1 + (2\pi f_b C)^2 e_{se}^2 I_2 \right] f_b. \tag{2}$$

The constants are

$$T = \text{temperature}$$
$$h = \text{Plank's constant}$$
$$\nu = \text{optical frequency}$$
$$\eta = \text{detector quantum efficiency}$$
$$e = \text{electron charge}$$
$$f_b = \text{bit rate.}$$

The weighting functions I_1 and I_2, which take account of pulse shape, are given by

$$I_1 = \frac{f_b}{2\pi} \int_{-\infty}^{\infty} \left| \frac{H_{out}(\omega)}{H_p(\omega)} \right|^2 d\omega \tag{3}$$

$$I_2 = \frac{1}{(2\pi)^3 f_b} \int_{-\infty}^{\infty} \left| \frac{H_{out}(\omega)}{H_p(\omega)} \right|^2 \omega^2 d\omega, \tag{4}$$

where $H_{out}(\omega)$ is the Fourier transform of the output voltage pulse shape, $H_p(\omega)$ is the Fourier transform of the optical power pulse shape, and the pulses have been normalized so that the area of the optical pulse is unity, as is the magnitude of the output pulse at the center of the time slot. The functions I_1 and I_2 can be shown to depend only on pulse shape relative to the time slot length, not on baud rate.

The probability of error can be readily derived from eq. (1), assuming Gaussian noise statistics. For Gaussian-distributed noise, the probability that the noise current will exceed a value D is given by

$$P(D) = \frac{1}{2} \text{erfc} \left(\frac{D}{I(p_o)\sqrt{2}} \right),$$

where erfc is the error function complement and I the root mean square

* Note that amplifier shot noise has been included in n_t.

noise current.[6] We assume an ideal regenerator where a "one" is produced if the input exceeds a threshold level D and a "zero" otherwise. Then if a "zero" is transmitted, the condition that the probability will be P_e that the noise will exceed the decision threshold is given by

$$D > QI(0), \tag{5}$$

where

$$Q = \sqrt{2} \ \text{erfc}^{-1} (2P_e). \tag{6}$$

Similarly, so that the probability that the noise will not exceed the signal when a "one" is transmitted, the expected value of the signal, given by $p_{\max}\eta e\langle g\rangle/h\nu$, where $\langle g\rangle$ is the mean avalanche gain and $p_{\max}$ the average power for all "ones," must exceed the threshold by Q times the noise current; that is,

$$\frac{p_{\max}\eta e\langle g\rangle}{h\nu} - D > QI_{\max}. \tag{7}$$

From eqs. (1), (5), and (7), the average power required to achieve a specified error probability with avalanche gain is given by

$$p = \frac{h\nu Q}{2\eta} \left[\frac{Q\langle g^2\rangle f_b}{\langle g^2\rangle} + \frac{2}{\langle g\rangle e} n_t^{\frac{1}{2}} \right]; \tag{8}$$

where $\langle g^2\rangle$ is the mean square avalanche gain. (In the case without avalanche gain when thermal noise predominates, the first term in the bracket of eq. (8) can be neglected.) For avalanche photodiodes, it has been found that

$$\langle g^2\rangle = \langle g\rangle^{2+x} \tag{9}$$

and for silicon units $x = 0.5$.

A value of $\langle g\rangle$ exists which optimizes performance.[7] The value which minimizes the power required to achieve a given error rate, found by minimizing eq. (8), is given by

$$g_{\text{opt}} = \frac{2^{4/3}}{(ef_bQ)^{\frac{2}{3}}} n_t^{\frac{1}{3}}. \tag{10}$$

At optimal gain the power required to achieve a specified error rate is given by

$$p = \frac{3h\nu Q}{\eta e g_{\text{opt}}} n_t^{\frac{1}{2}} \tag{11}$$

$$p = \frac{3h\nu Q^{5/3} f_b^{\frac{2}{3}} n_t^{\frac{1}{6}}}{2^{4/3}\eta e^{\frac{1}{3}}}. \tag{12}$$

It is interesting to note that eq. (11) can also be put in the form

$$p = \frac{3}{g_{\text{opt}}} p',$$ (13)

where p' is the required power without avalanche gain. From eqs. (8), (10), and (12), the required power without avalanche gain, the optimum avalanche gain, and the required power with optimum avalanche gain are proportional to the second, third, and sixth roots of the thermal noise, respectively.

The series noise source of a junction field effect transistor (FET) is virtually independent of the circuit parameters in the normal range of operation, and the shunt noise source is negligible at 6.3 Mb/s without input tuning. For an FET[8]

$$e_{se}^2 \approx 2kT \left(\frac{0.7}{g_m} \right),$$ (14)

where g_m is the transconductance of the device, so the thermal noise referred to the detector is

$$n_t \approx 2kT f_b \left[\left(\frac{1}{R_r} + \frac{0.7}{g_m R^2} \right) I_1 + \frac{0.7 (2\pi f_b C)^2}{g_m} I_2 \right],$$ (15)

assuming all the noise is due to the first stage of the input amplifier. For a good FET the input resistance is virtually infinite, so

$$R = R_r.$$

In a common-source configuration, C is the sum of the drain-gate, gate-source, diode, and gate wiring capacitances.

III. CIRCUITRY

Figure 2 is a block diagram of the repeater, which was constructed in a $5\frac{1}{4} \times 4 \times 1\frac{1}{2}$ inch enclosure. The main signal path is represented by heavy lines and boxes. The signal, which is detected by either a PIN or a silicon avalanche photodiode, is first amplified by a high-impedance input amplifier. Following this, additional gain is provided by an SN52733 integrated video amplifier. Next, the signal is equalized, then further amplified by another SN52733 video amplifier and filtered by a single-section maximally flat LC filter with a 7-MHz bandwidth which, in combination with the other amplifiers, gives a 3-dB point of about 6.3 MHz. From this point the signal is fed to the timing circuits, the automatic gain and threshold circuits, and the regenerator. Finally, the regenerated signal is amplified and applied to the LED.
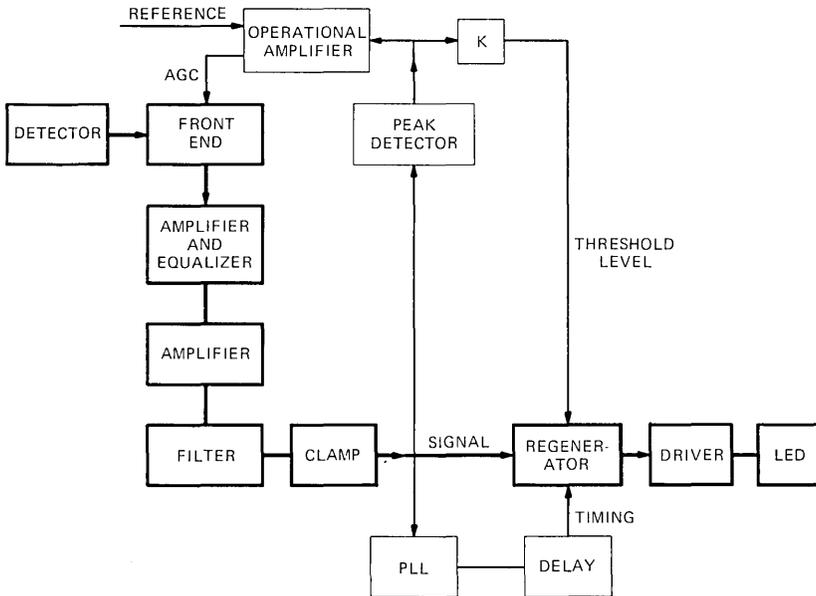
Fig. 2—6.3 Mb/s repeater.

3.1 *Input amplifier*

A 3-stage input amplifier was employed, as shown in Fig. 3. The amplifier consists of a 2N4416 junction field effect transistor followed by a 3N159 tetrode amplifier, and finally a 2N4416 in a source follower configuration. It was found experimentally that input amplifiers with a 2N4416 input stage had an input noise equivalent power about 1 dB less than with the 3N159. However, the tetrode can provide more gain for a single stage because of its low drain to first gate capacitance. Furthermore, the tetrode is well suited to automatic gain control since the g_m of the device is highly dependent on the voltage applied to the second gate. Thus, the configuration of Fig. 3 was chosen. It was found that the input noise dropped by 6 dB when the first gate of the tetrode was shorted to ground, so 75 percent of the thermal noise originates in the first stage. Thus, this configuration is very close to optimum. The source follower was provided to decouple the input amplifier from the subsequent circuits.

3.2 *Equalization*

In order to compensate for the distortion introduced by the long input time constant, the circuit of Fig. 4 was employed. With a source
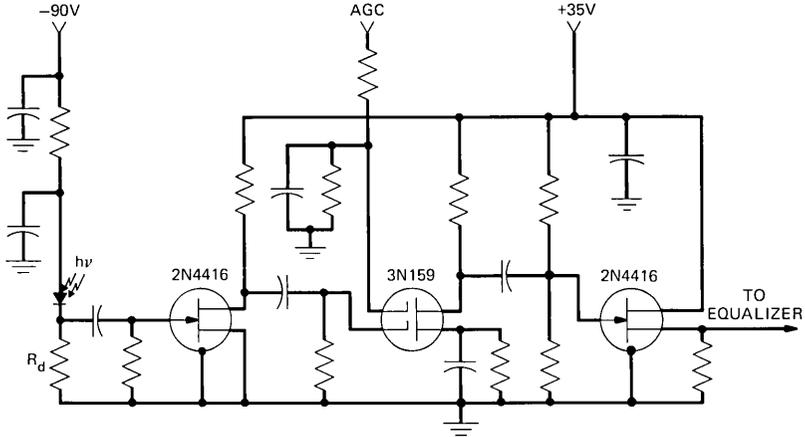
Fig. 3—Input amplifier.

of resistance R_s and a load whose resistance is included in R_2, it can be shown that the transfer function has a pole at

$$s = -\frac{1}{C_1 R_1}\left(1 + \frac{R_1}{R_s + R_2}\right)$$

and a zero at

$$s = -\frac{1}{C_1 R_1}.$$

Thus, the position of the zero can be adjusted by varying either C_1 or R_1 and, as long as

$$R_1 \gg R_s + R_2,$$

the pole will be above the band of interest and have a negligible effect.

3.3 Clamping and Peak Detection

Since the amplifiers of the repeater are ac-coupled, the dc level would be a function of word pattern unless suitable provisions are made.
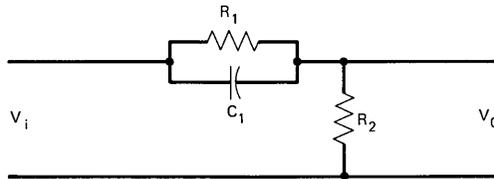
Fig. 4—RC equalizer.
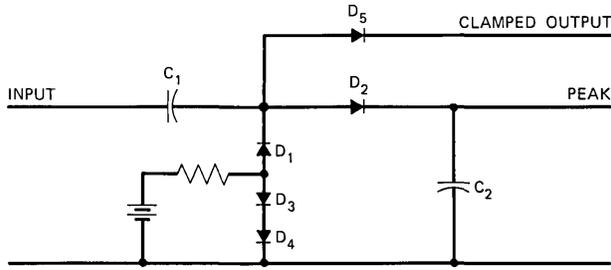
Fig. 5—Clamp and peak detector.

In addition, provision to measure signal level must be incorporated to set the threshold and to control gain. Clamping and peak detection were employed to solve these problems.

By incorporating both automatic threshold and gain controls, the AGC gain need only be high enough to assure that the phase-locked loop will function properly and to prevent compression. This reduces the tendency toward instability because of a high-gain feedback loop.

Figure 5 shows the circuit employed. Diode D_1 and capacitor C_1 serve as the clamp, and diode D_2 and capacitor C_2 serve as the peak detector. The diodes D_3, D_4, and D_5 are included to cancel the diode drops of D_1 and D_2.

3.4 Digital Circuits

The timing coincidence and regenerator circuits, Figs. 6 and 7, share an SN72810 dual comparator and an SN7474 dual D flip-flop. The comparator has the property

$$V_0 = 1, \qquad V_1 > V_2$$
$$V_0 = 0, \qquad V_1 < V_2,$$

where 1 and 0 represent states. For the D flip-flop, the output Q takes
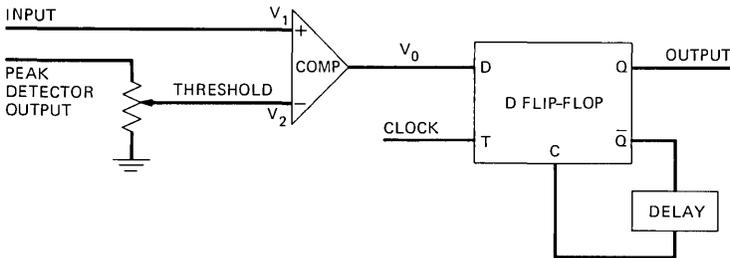
Fig. 6—Regenerator.
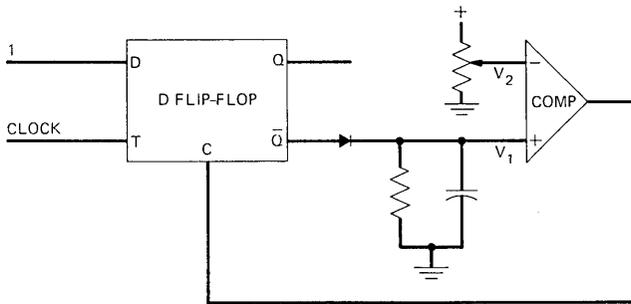
Fig. 7—Timing coincidence.

on the value applied to the D input when the clock input changes from low to high. This value is held until either the clock again shifts from low to high or the clear c is returned to ground. The output $\bar{Q}$ is the complement of Q.

In the regenerator, Fig. 6, the comparator serves as a quantizer and the D flip-flop retimes the signal. On the flip-flop, feedback from $\bar{Q}$ to c is provided so the D flip-flop output will be a pulse of proper duration whenever the clock goes positive and the D input is high.

The circuit shown in Fig. 7 is used to adjust timing coincidence between the phase-locked loop output and the regenerator. The D input to the flip-flop is kept high at all times. Adjusting the input to V_2 of the comparator adjusts the delay between the time when a positive clock pulse is applied and the voltage of the clear (c) drops to a sufficiently low value to clear the flip-flop. The rising edge of the $\bar{Q}$ output is used to trigger the regenerator.

3.5 LED and driver

The LED driver consists of a cascade of two emitter followers. The driver is capable of generating 1.5-A pulses into a diode load. For the tests to be described, the diode was driven with 300-mA peak current pulses and generated optical pulses of about 0.3-mW peak power.

IV. RESULTS

Both signal-to-noise ratio and error-rate measurements were made to evaluate the input amplifier and repeater performance. The 2N4416 JFET employed for all the tests had an input capacitance of 5 pF and g_m of 0.006 mho. An additional picofarad of capacitance was added by the diode load circuit. The measurements without gain were per-

formed with an SGD-040A PIN photodetector which had a capacitance of 2 pF and a quantum efficiency of 83 percent. For the measurements with gain, a TIXL56 silicon avalanche photodetector was employed. This diode had a capacitance of 1 pF and a quantum efficiency of 55 percent. For both diodes, the noise calculated from the measured dark current was negligible.

The constants I_1 and I_2 have been evaluated by Personick[5] for rectangular optical pulses and pulses with a raised cosine spectrum and maximum eye opening at the regenerator input. He found $I_1 = 0.6$ and $I_2 = 0.26$. The theoretical curves presented here were obtained using these values.

Measurements of noise equivalent power, that is, optical power to achieve a unity signal-to-noise ratio at the regenerator, were first made to evaluate the performance of the input amplifier. As shown in Fig. 8, the results for the input amplifier of Fig. 3 with equalization closely approximates those predicted from eq. (8) by setting $Q = 1$. The difference between the theoretical and the experimental curve is mainly due to the noise of the second stage, which was about 1 dB with the 1-MΩ diode load resistor.
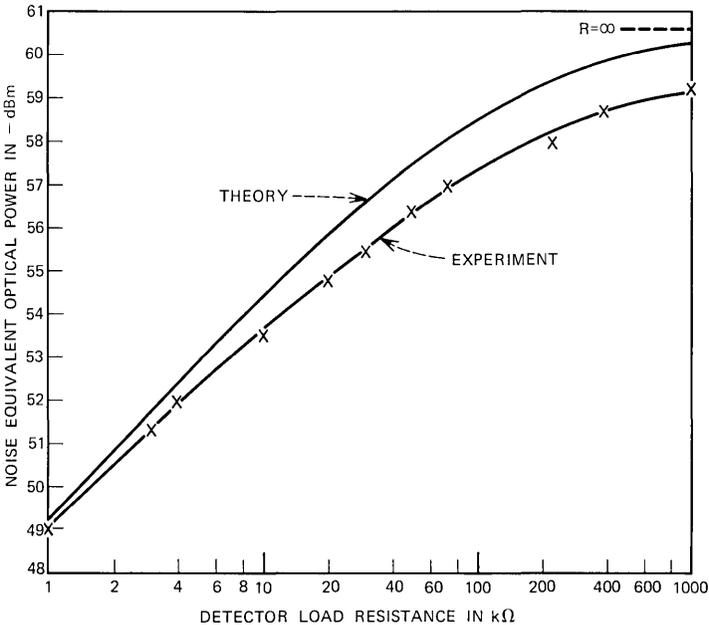
Fig. 8—Noise equivalent power circuit vs. diode load resistance.

Other input amplifiers were also tried. It was found that with a common-drain first stage and a common-source second stage, with a 1-MΩ diode load resistor, the noise equivalent power was about 1 dB higher. With smaller values of diode load resistance, where the resistor noise predominates, the common-drain input amplifier's performance was identical to that of the common-source input amplifier. Similar results were obtained with a cascade configuration using a bipolar transistor for the second stage.

Figure 9 shows "eye" diagrams taken at key points in the receiver with a 1-MΩ diode load resistance. Before equalization, the "eye" is fully closed, as is expected. After equalization, the "eye" is almost fully open. The regenerated pulse was photographed with recovered timing.

The tracking of the threshold level with peak signal is shown in Fig. 10. The tracking in conjunction with the AGC is adequate, as will be apparent from the error performance. With germanium or Schottky barrier diodes in the clamping and peak detection circuit, the tracking could have been held even closer to the ideal, had this been necessary. The AGC also functioned properly. The signal level could be held within about a 20-percent range with the power up to 10 dB above the signal required for 10^{-8} error probability. Greater range could be achieved by cascading tetrode FET stages, if required.
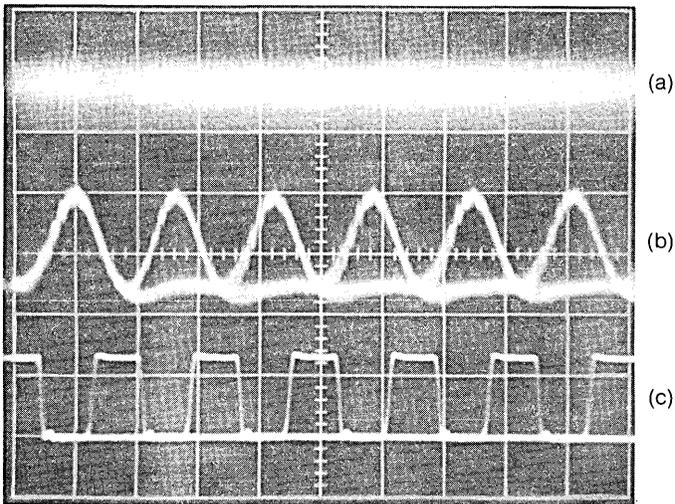
Fig. 9—"Eye" diagrams: (a) Before equalization. (b) After equalization. (c) Regenerated pulse.
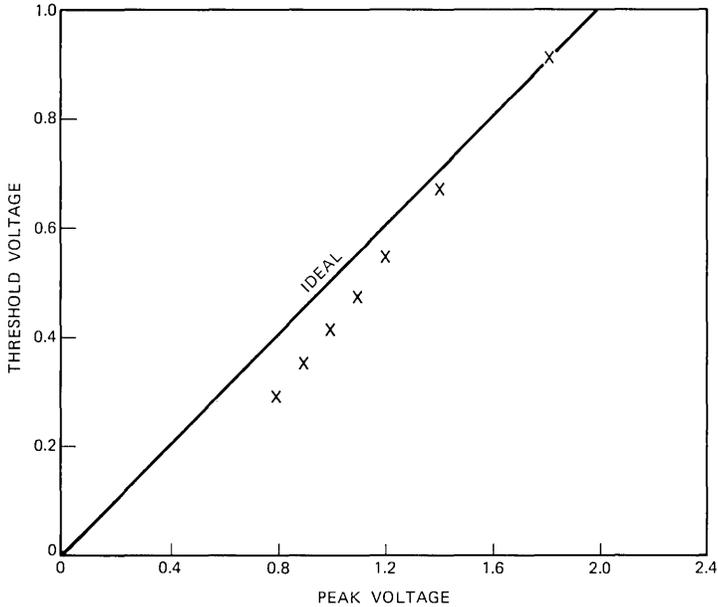
Fig. 10—Threshold voltage vs. peak voltage.

Error probability measurements were made under a variety of conditions. The results are shown in Fig. 11. The signal source was a 2^{15}-1 bit pseudo-random word generator. The signal consisted of 15 bit blocks each separated by a zero. The measurements were made with external timing and performance was optimized at each point, except the points indicated with x's. For these points, which were taken to check the performance of the automatic gain and threshold controls, as well as the timing recovery, the system was optimized with recovered timing and AGC at an error probability of about 10^{-8}. Then all the points were taken without further adjustment to the repeater.

The theoretical curve for a 4-kΩ diode load resistor is shown along with the measured points. These points were taken with the source-follower input amplifier because this amplifier introduced less intersymbol interference with the 4-kΩ diode load resistor. The improvement with a 1-MΩ diode load resistor over a 4-kΩ one was about 8 dB, which is in agreement with the theory.

Measurements of repeater performance were made with a TIXL56 silicon avalanche photodiode. This diode exhibits a significant diffusion tail. A second stage of RC equalization was employed to remove the resulting intersymbol interference.
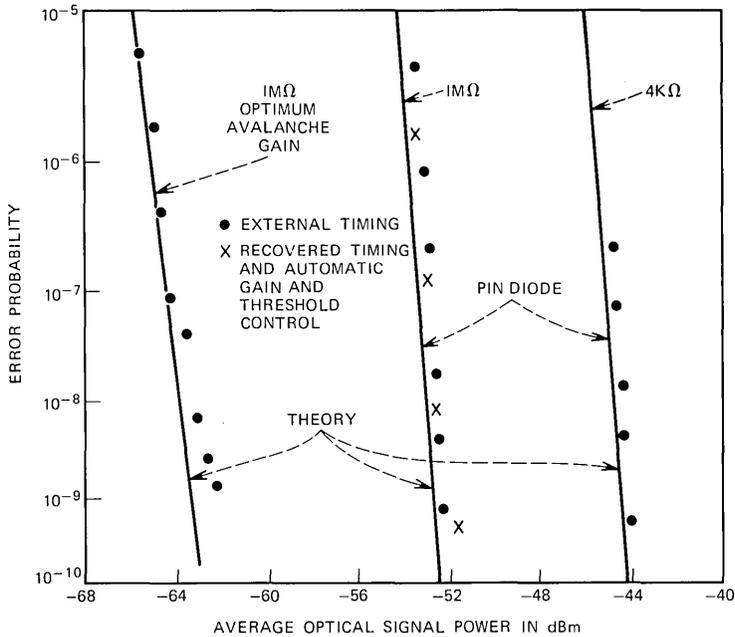
Fig. 11—Error probability vs. average optical signal power.

Equation (10) indicates that the optimum avalanche gain is a function of error rate. However, since it is not practical to optimize the gain at extremely low error probability, the gain was optimized at an error probability of 5×10^{-7} and then held constant for all the points.

Table I shows a comparison of the measured and predicted values of g_{opt} for 4-kΩ and 1-MΩ diode load resistance. In view of the assumption of noise statistics and diode characteristic, the agreement is satisfactory.

V. CONCLUSIONS

It has been demonstrated that a simple compact low-cost* repeater suitable for fiber optic applications can be built which functions close to theory at 6.3 Mb/s. The high-impedance input amplifier and its associated equalizer were realized in a straightforward manner, and compression did not turn out to be a serious problem. A significant reduction in required signal power to achieve a specified error rate with-

---

* The cost of the active components was about $30, exclusive of the detector and LED. An SGD-040 PIN detector costs about $15 and a TIXL56 avalanche detector $65.

Table I

Detector Load Ohms	Theoretical Gain	Measured Gain
4 K	171	188
1 M	44	62.2

out avalanche gain was achieved. With an avalanche detector, the optimum gain was greatly reduced with high impedance input, as predicted. Thus, the temperature stability will be greatly increased and the diode fabrication requirements eased with an avalanche detector.

VI. ACKNOWLEDGMENT

The author wishes to acknowledge the advice and assistance of W. M. Muska in the fabrication of the repeater.

REFERENCES

1. D. B. Keck, P. C. Schultz, and F. Zimar, "Attenuation of Multimode Glass Optical Waveguide," Appl. Phys. Lett., *21*, No. 5 (September 1, 1972).
2. A. B. Gillespie, *Signals, Noise, and Resolution in Nuclear Counter Amplifiers,* New York: McGraw-Hill, 1953.
3. O. H. Schade, Sr., "A Solid-State Low-Noise Preamplifier and Picture-Tube Drive Amplifier for a 60 mHz Video System," RCA Review, *29*, No. 1 (March 1968), p. 3.
4. C. A. Burrus and R. W. Dawson, "Small-Area High-Current GaAs Electroluminescent Diodes and a Method of Operation for Improved Degradation Characteristics," Appl. Phys. Lett., *17*, No. 3 (August 1970), pp. 97–99.
5. S. D. Personick, "Receiver Design for Digital Fiber Optic Communication Systems, Pt. I," B.S.T.J., *52* (July 1973), pp. 843–886.
6. W. R. Bennett and J. R. Davey, *Data Transmission,* New York: McGraw-Hill, 1965, pp. 100, 101.
7. H. Melchior and W. T. Lynch, "Signal and Noise Response of High Speed Germanium Avalanche Photodiodes," IEEE Trans. on Elec. Devices, *ED 13*, No. 12 (December 1966), pp. 829–838.
8. A. van der Ziel, *Noise, Sources, Characterization, Measurement,* Englewood Cliffs, N.J.: Prentice-Hall, 1970, pp. 74–76.

Slab-Coupled Waveguides

By E. A. J. MARCATILI

The slab-coupled waveguide, consisting of a dielectric rod lying on a slab that in turn covers a substrate, is a multidielectric waveguide that includes such special cases as the single-material fiber, the rib waveguide, and the strip-loaded film guide. These guides have recently become known as potentially useful either for long-distance optical transmission or for integrated optics.

Simple, closed-form, approximate solutions have been found to describe the following properties of the guide: number of modes, their field configurations and propagation constants, numerical aperture, requirements for single-mode operation, field penetration in the slab, tolerance to curvature of the guide axis, dispersion, and impulse response.

I. INTRODUCTION

Descriptions of three novel dielectric waveguides of wide potential use in long-distance optical transmission and in integrated optics have appeared in the literature recently. These guides are the single-material fiber[1,2] (Fig. 1) made of low-loss undoped fused silica; the rib waveguide[3] (Fig. 2) made of two materials, and the strip-loaded film guide[4,5] (Fig. 3) made of three materials. In all these fibers n_3 is air or an inert atmosphere, while in a more general guide it could be another dielectric.

Although these guides have different shapes and different distributions of refractive indices, they have essential elements in common that make them close relatives of the same family. A more generic member of this family of waveguides (Fig. 4), from which all the others can be deduced, is a fiber of arbitrary cross section at a distance l from a slab mounted on a substrate. The way in which this guide operates is simpler to understand than the others, and is described below.

The modal spectrum of the fiber ($l = \infty$) is shown in Fig. 5(a). In this example, five modes are guided and their axial propagation constants k_z lie between kn_3 and kn_2 where k is the free-space propagation constant $2\pi/\lambda$. Smaller propagation constants than kn_3 belong to a
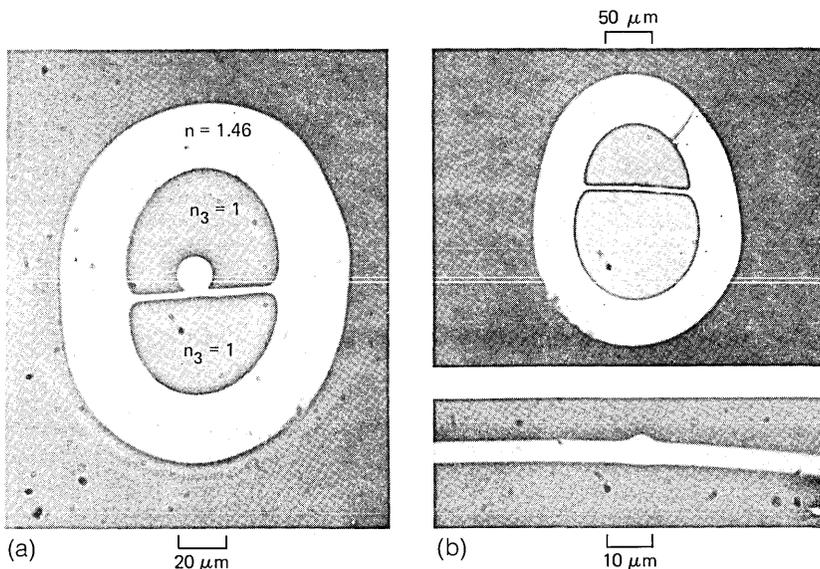
Fig. 1—Photographs of an experimental (a) multimode single-material fiber and (b) single-mode single-material fiber (top), with magnified core region (bottom) $(n > n_3)$.

continuum of radiating modes that are unimportant for this discussion. On the other hand, the isolated slab $(l = \infty)$ supports modes with q extrema in the y direction. For simplicity we will assume that the field is well confined within the slab. The field components vary sinusoidally along x, y, and z and the respective propagation constants k_x, $k_y = \pi q/t$ and k_z are related by the characteristic equation

$$k_z = \sqrt{k^2 n^2 - k_x^2 - \left(\frac{\pi q}{t}\right)^2}.$$

Since k_x can take any value between zero and infinity, the propagating
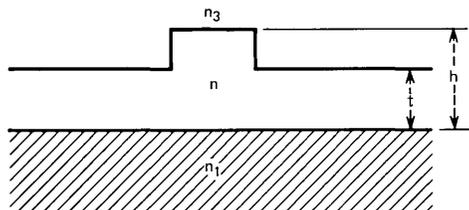
Fig. 2—Rib waveguide $(n > n_1$ and $n_3)$.
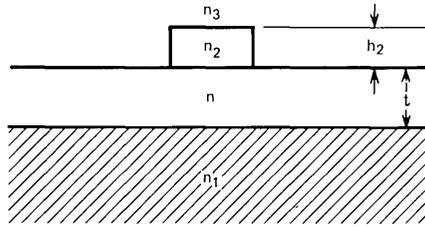
Fig. 3—Optical strip line ($n > n_1$, n_2 and n_3).

modes of the slab with one maximum along y, ($q = 1$) constitute a continuum with axial propagation constants ranging from zero to

$$\sqrt{k^2 n^2 - \left(\frac{\pi}{t}\right)^2},$$

as shown in Fig. 5(b). Similarly, propagating slab modes with two extrema along y, ($q = 2$) constitute another continuum with propagation constants k_z ranging from zero to

$$\sqrt{k^2 n^2 - \left(\frac{2\pi}{t}\right)^2},$$

as shown in Fig. 5(c), and so on.

Now let us imagine that, as in Fig. 4, the fiber and the slab are separated by a finite distance l which is far enough that their respective spectra are only slightly perturbed by the coupling. Modes with the same propagation constant k_z will couple to each other. Therefore,
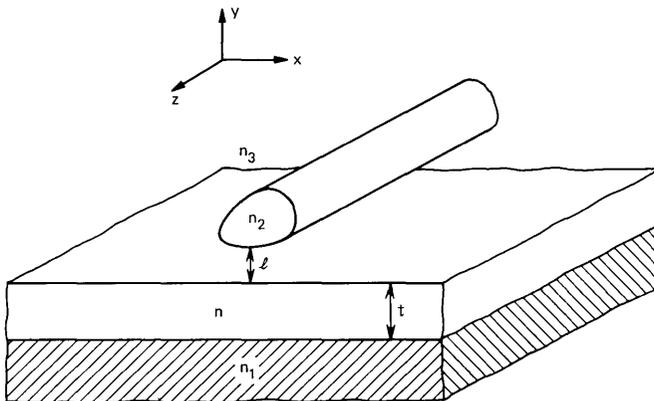
Fig. 4—Slab-loaded waveguide.
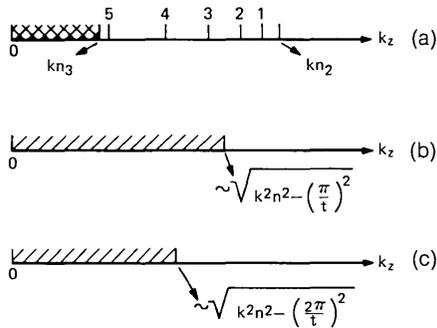
Fig. 5—Modal spectrum in: (a) isolated fiber of Fig. 4; (b) isolated slab of Fig. 4 (only for modes with one half period across t; (c) isolated slab of Fig. 4 (only for modes with two half periods across t).

modes 1 and 2 of the fiber will remain guided without attenuation, though there is indeed an electromagnetic field in the slab that decays exponentially in the x direction (Fig. 4) away from the fiber. Mode 3 will couple to the slab mode with the same k_z of the spectrum in Fig. 5(b); modes 4 and 5 will couple to modes of spectra in Figs. 5(b) and 5(c), etc. The net result is that modes 3, 4, and 5 of the fiber will be attenuated by coupling to slab modes. The smaller the distance l between fiber and slab, the tighter the coupling and consequently the higher the attenuation of these leaky modes. From Fig. 5 it becomes obvious that by adjusting the thickness t of the slab, the number of lossless modes can be selected.

The most important point from this discussion is that only fiber modes with axial propagation constants k_z larger than the propagation constant

$$\sqrt{k^2 n^2 - \left(\frac{\pi}{t}\right)^2}$$

[Fig. 5(b)] of the slab's fundamental mode are lossless. Therefore, throughout the paper we will be concerned with the coupling between the field in the core's guide and the fundamental mode of the slab.

Having identified the three basic elements of this waveguiding structure, a slab, a guide (also referred to as a fiber or strip), and the coupling between them, we will use the generic name of slab-coupled guides, fibers, or strips for all the members of this prolific family. Of course, we will reserve the names given by the original authors to identify individual guides.

The general solution of the slab-coupled guide in Fig. 4 would encompass as particular cases those in Figs. 1, 2, and 3. However, only
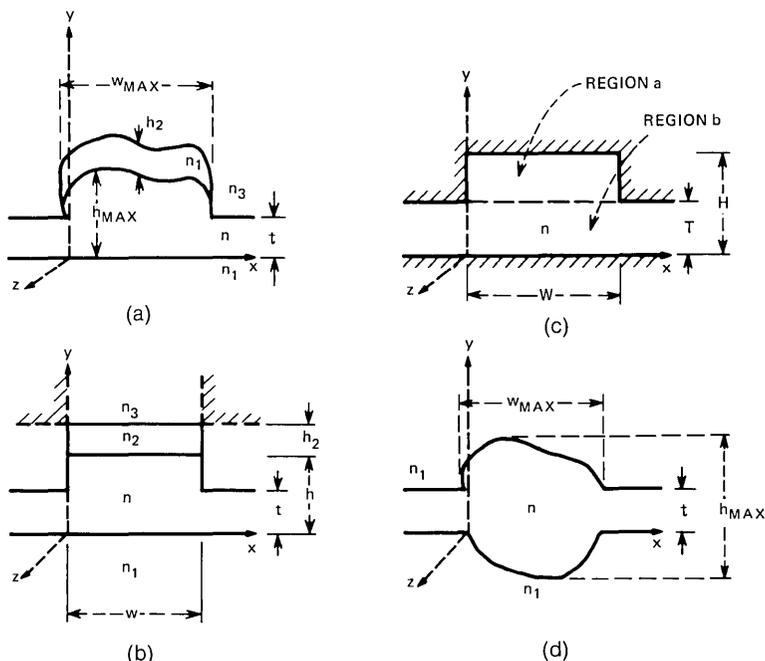
Fig. 6—Slab-coupled waveguide. (a) Original guide. (b) and (c) Equivalent simpler guides. (d) Single-material guide.

two extreme cases seem amenable to closed-form calculations: the case of feeble coupling described in Ref. 6 and the case of strong coupling occurring when the separation between fiber and slab vanishes, and which is the subject of this paper.

In Section II we consider the properties and characteristics of a somewhat generalized slab-coupled guide [Fig. 6(a)]. These are the following:

(i) the equivalence to a much simpler guide shown in Fig. 6(c),
(ii) the number of guided modes,
(iii) their propagation constants and field configurations,
(iv) the numerical aperture,
(v) the design for single-mode operation,
(vi) the field penetration in the slabs,
(vii) the tolerance of the guide to the curvature of its axis.

These general results are applied to the multimode and single-mode single-material fibers, rib guides, and strip-loaded guides in Sections III, IV, V, and VI, respectively. Furthermore, dispersion and impulse

response in single-material fibers are considered in Sections III and IV. The simple but burdensome mathematics involved are placed in the appendices to this paper.

II. SOLUTION OF THE SLAB-COUPLED GUIDE

Consider the somewhat generalized slab-coupled guide of Fig. 6(a). It is shown in Appendix B that if

(*i*) most of the electromagnetic energy travels within the region of refractive index n,

(*ii*) the height and the width of the core are almost constants,

(*iii*) there are no turning points within the core $(0 < x < w_{\max})$ or, in other words, exponential decay of the field components in the region of refractive index n occurs only in the slabs, and

$$(iv) \qquad\qquad n > \begin{cases} n_1, \\ n_2, \\ n_3 \end{cases} \qquad\qquad (1)$$

then the four-dielectric guide with somewhat arbitrarily shaped core in Fig. 6(a) is equivalent to the single-dielectric guide with rectangular core in Fig. 6(c). Surrounding the dielectric of refractive index n, there is a material into which there is no field penetration and the x and y field components of the guided modes vanish at the interface. The dimensions of this equivalent guide are, according to eqs. (83), (84), and (85),

$$T = t(1 + c_t), \qquad\qquad (2)$$

$$W = w(1 + c_w), \qquad\qquad (3)$$

$$H = h(1 + c_h), \qquad\qquad (4)$$

in which the quantities c_t, c_w, and c_h, which are small compared to unity, are from (73), (74), (79), (86) to (89), (100), and (101).

$$c_t = \begin{cases} \dfrac{h}{t}\left(\dfrac{1}{v_1} + \dfrac{1}{v_3}\right) & \text{for modes polarized along } x \\[3mm] \dfrac{h}{tn^2}\left(\dfrac{n_1^2}{v_1} + \dfrac{n_3^2}{v_3}\right) & \text{for modes polarized along } y \end{cases} \qquad (5)$$

$$c_w = \begin{cases} \dfrac{2h}{wv_3}\dfrac{n_3^2}{n^2} & \text{for modes polarized along } x \\[3mm] \dfrac{2h}{wv_3} & \text{for modes polarized along } y \end{cases} \qquad (6)$$

$$c_h = \begin{cases} \dfrac{1}{v_1} + \dfrac{1}{v_2} \tanh\left(\dfrac{h_2}{h} v_2 + \tanh^{-1} \dfrac{v_2}{v_3} \right) \\[4pt] \qquad\qquad\qquad \text{for modes polarized along } x \\[10pt] \dfrac{n_1^2}{n^2 v_1} + \dfrac{n_2^2}{n^2 v_2} \tanh\left(\dfrac{h_2}{h} v_2 + \tanh^{-1} \dfrac{n_3^2 v_2}{n_2^2 v_3} \right) \\[4pt] \qquad\qquad\qquad \text{for modes polarized along } y, \end{cases} \tag{7}$$

where

$$v_{1,2,3} = kh\sqrt{n^2 - n_{1,2,3}^2}\,,^\dagger \tag{8}$$

$$h = \sqrt{\frac{h_{\max}}{w_{\max}}}\, s, \tag{9}$$

$$w = \sqrt{\frac{w_{\max}}{h_{\max}}}\, s, \tag{10}$$

and $h_{\max}$ and $w_{\max}$ are the maximum height and width of the portion of the core with refractive index n in Fig. 6(a) and s is its cross-sectional area. All these expressions are valid for

$$v > \begin{cases} \dfrac{\pi}{2} - \tan^{-1} \sqrt{\dfrac{n^2 - n_1^2}{n_1^2 - n_3^2}} \qquad \text{for modes polarized along } x \\[10pt] \dfrac{\pi}{2} - \tan^{-1} \dfrac{n_3^2}{n^2} \sqrt{\dfrac{n^2 - n_1^2}{n_1^2 - n_3^2}} \qquad \text{for modes polarized along } y, \end{cases} \tag{11}$$

where

$$v = kt\sqrt{n^2 - n_1^2}. \tag{12}$$

Four parameters, then, n, T, W, and H, determine the guide in Fig. 6(c) and we proceed to characterize its transmission properties.

The guided modes are hybrid; however, the longitudinal field components (along z) are small compared to the transverse ones; therefore, the modes are almost transverse electromagnetic (TEM). Within the core, these transverse field components vary sinusoidally along x and y. Within the slab, the field components also vary sinusoidally along y but decay exponentially away from the core. All these components vanish at the edge of the guide.

There are two families of modes, E_{pq}^y and E_{pq}^x. The first family, Fig. 7(a), is mostly polarized along y, and the main transverse field components are E_y and H_x. Within the core, a mode has p field extrema (approximately p half periods) along x and q field extrema along y.

---

† Throughout this paper numbers or letters separated by commas must be considered one at a time.
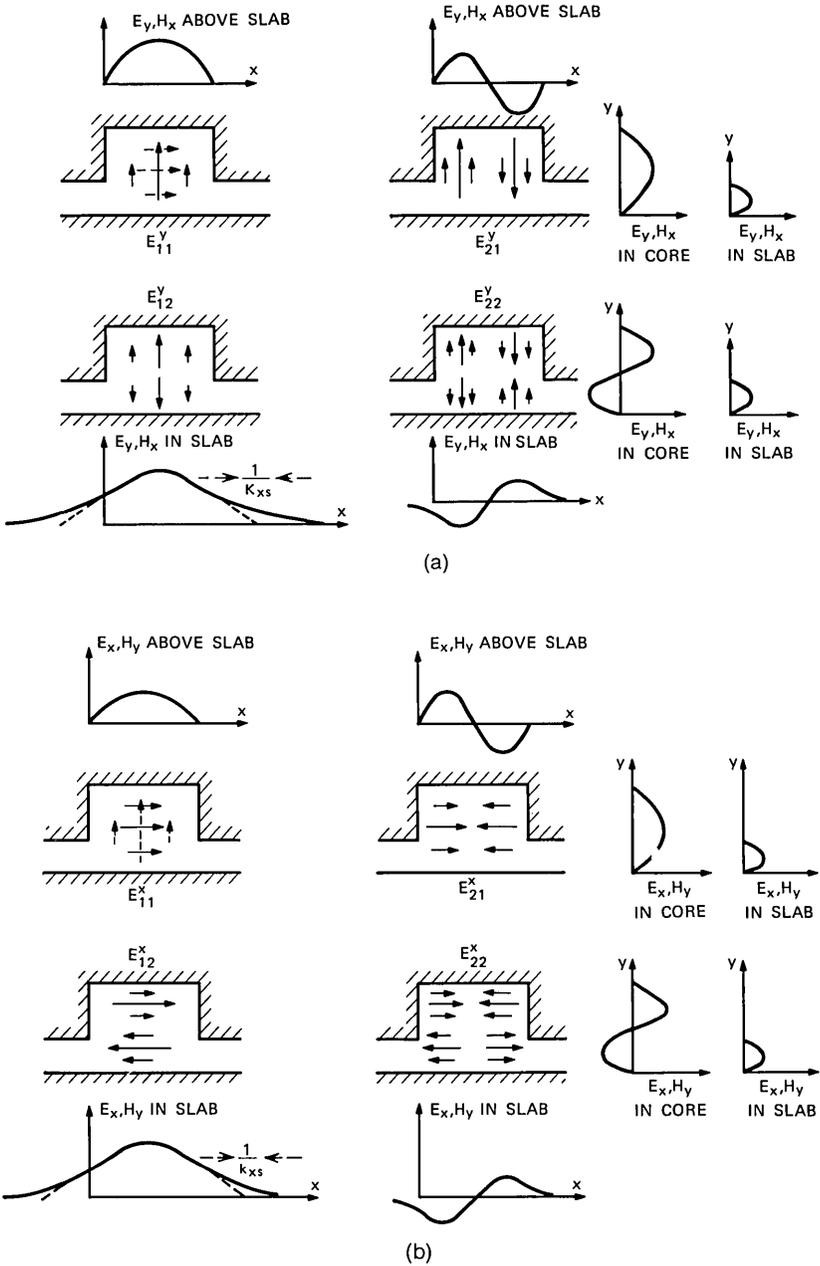
Fig. 7—Two families of modes. (a) E_{pq}^y modes and (b) E_{pq}^x modes.

The second family of modes, Fig. 7(b), is mostly polarized along x and the main transverse components are E_x and H_y.

For both families, the axial propagation constant k_z and the field penetration in the slabs d_{pq}, that is the distance over which the field components decay in the slabs by $1/e$, are according to (98) and (99),

$$k_z = \sqrt{k^2 n^2 - \left[\frac{\pi p}{W(1 + c_q)} \right]^2 - \left(\frac{\pi q}{H} \right)^2} \tag{13}$$

$$k_{xs} = \frac{1}{d_{pq}} = \frac{\pi}{T} \sqrt{1 - \left[\frac{pT}{W(1 + c_q)} \right]^2 - \left(\frac{qT}{H} \right)^2}, \tag{14}$$

where c_q taken from (97) is

$$c_q = \frac{2}{\pi} \frac{T^2}{WH} \frac{1}{\sqrt{1 - \left(\frac{qT}{H} \right)^2}}. \tag{15}$$

The highest-order modes, which we will designate with indices $p = P$ and $q = Q$, are those for which the penetration depth d_{PQ} is infinite. For them, eqs. (13) and (14) are reduced to

$$\sqrt{n^2 - \frac{k_{z\min}^2}{k^2}} = \frac{\lambda}{2T} = \text{N.A.} \tag{16}$$

and

$$\frac{PT}{W} - \sqrt{1 - \left(\frac{QT}{H} \right)^2} = \frac{2}{\pi} \frac{T^2}{WH}. \tag{17}$$

While for ordinary fibers, the numerical aperture (N.A.) $\sqrt{n^2 - n_e^2}$ is an exclusive function of the refractive indices of the core n, and the cladding n_e, the N.A. of a slab-coupled guide defined in (16) is mostly a function of the wavelength and the slab thickness. The longer the wavelength and the thinner the slab, the larger the numerical aperture. This statement is true provided that the inequality (11) is satisfied.

Naturally, an equivalent cladding refractive index

$$n_e = \sqrt{n^2 - \left(\frac{\lambda}{2T} \right)^2} \tag{18}$$

is derived by equating $k_{z\min}/k$ to n_e.

The easiest property to observe in slab-coupled guides is, probably, the number of spots of the highest order mode guided. That number is the product of P and Q which are related to each other by eq. (17).
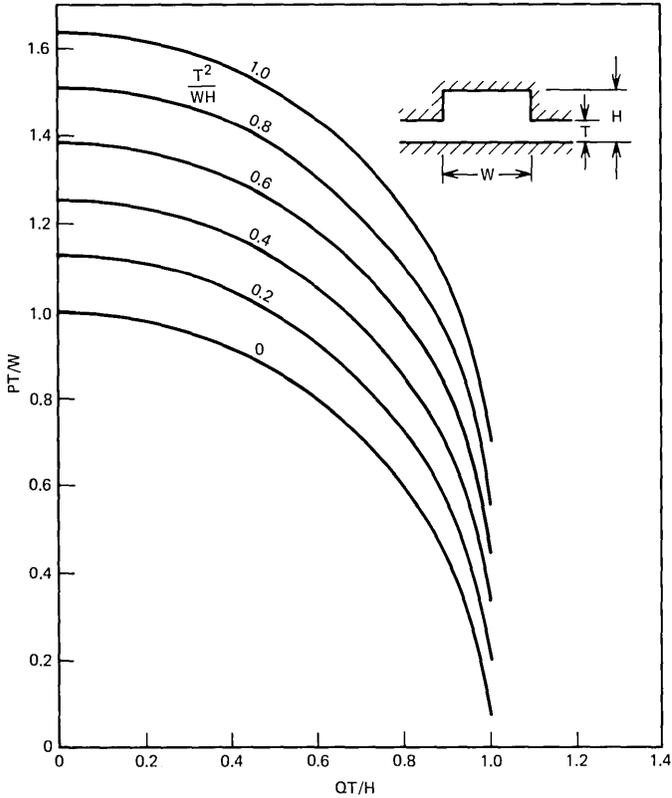
Fig. 8—Waveguide dimensions for $E_{PQ}^{x,y}$ mode at cutoff.

A plot using PT/W and QT/H as coordinates and T^2/WH as the parameter is shown in Fig. 8. Given T/W and T/H, the parameter T^2/WH, which selects one of these curves in Fig. 8, is also known. For the ordinate T/W corresponding to $P = 1$, the abscissa $Q_{max} T/H$ is determined and from it the maximum number of half periods Q_{max} of the modes $E_{1Q_{max}}^{x,y}$ in the y direction. Similarly for the abscissa T/H corresponding to $Q = 1$, the ordinate $P_{max} T/W$ yields the maximum number of half periods P_{max} of the modes $E_{P_{max}1}^{x,y}$ in the x direction.

  Example: For $T/H = 0.5$ and $T/W = 0.1$, the values $P_{max} = 8$ and $Q_{max} = 1$ (rounded off to the immediate lower integer) are obtained.

  The explicit values of P_{max} and Q_{max} derived from (17) are

$$P_{max} = \frac{W}{T}\left[\sqrt{1 - \left(\frac{T}{H}\right)^2} + \frac{2}{\pi}\frac{T^2}{WH}\right] \tag{19}$$

and

$$Q_{\max} = \frac{H}{T}\left[1 - \left(\frac{T}{W}\right)^2\left(1 - \frac{2T}{\pi H}\right)^2\right]^{\frac{1}{2}}. \tag{20}$$

From this last equation,

$$T < \frac{H}{Q_{\max}}. \tag{21}$$

Consequently, for any guided mode the slab thickness T is always smaller than the half period of the mode in the core along y. This justifies one of the assumptions in Appendix B.

The guide is largely overmoded if

$$\frac{T}{W} \ll 1 \tag{22}$$

and

$$\frac{T}{H} \ll 1; \tag{23}$$

then, the number of modes for each polarization deduced from (19) and (20) results in

$$N = \frac{\pi}{4}\frac{WH}{T^2}. \tag{24}$$

Unlike ordinary fibers, the number of modes of a slab-coupled guide is mostly determined by its geometry.

To dimension the slab-coupled guide for single-mode operation, eq. (17) has been plotted in Fig. 9 using T/H and T/W as variables plus two sets of parameters $P = 1$, $Q = 2$ and $P = 2$, $Q = 1$. The coordinates of the first line give the dimensions of a guide with the E_{12}^x and E_{12}^y modes at cutoff while those in the second line yield the guide dimension with the E_{21}^x and E_{21}^y modes at cutoff. The solid portions of both curves determine the smallest possible ratios T/W and T/H compatible with single-mode guidance. The discrete numbers on the curves indicate the ratio H/W. For a square core $(H = W)$ we deduce from the figure that $T/W \simeq T/H \simeq 0.5$.

For the fundamental mode, the field penetration in the slabs, d_{11}, is obtained from (14), making $p = q = 1$. With rearranged terms, eq. (14) reads

$$\frac{W}{T} = \frac{1}{\sqrt{1 - \left(\frac{T}{H}\right)^2 - \left(\frac{T}{\pi d_{11}}\right)^2}} - \frac{2T}{\pi H}\frac{1}{\sqrt{1 - \left(\frac{T}{H}\right)^2}} \tag{25}$$
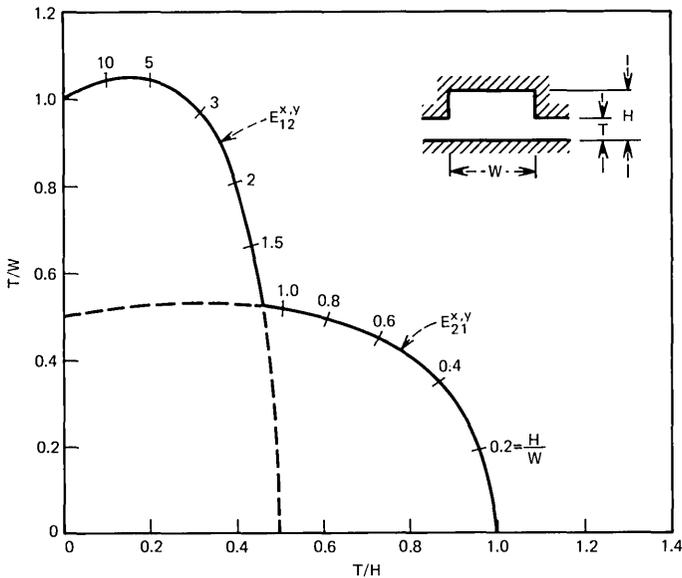
Fig. 9—Waveguide dimensions for $E_{12}^{x,y}$ and $E_{21}^{x,y}$ modes at cutoff.

and it is plotted as solid lines in Fig. 10 using T/H and T/W as co-ordinates and T/d_{11} as the parameter.

In the same figure, the dotted line is a reproduction of the curve of Fig. 9 corresponding to the cutoff condition of the $E_{12}^{x,y}$ and $E_{21}^{y,x}$ modes. The intersection of this curve with the others yields, then, the field penetration of the fundamental modes in guides designed to be at cutoff for the next higher order modes.

The region between the solid curve with parameter $T/d_{11} = 0$ (infinite field penetration in the slab) and the dotted curve delimits the possible choices of T/H and T/W for single-mode waveguides. The region within the dotted curve corresponds to multimode wave-guides.

Let us consider now the attenuation of the fundamental mode due to radiation induced by the curvature of the guide's axis. If the guide axis is bent in the plane of the slab along a constant radius of curvature R, the attenuation of the fundamental mode in a 90° bend[7,8] is pro-portional to

$$R \exp\left[-\frac{1}{6\pi^2 n^2} \frac{\lambda^2 R}{d_{11}^3} \right] \tag{26}$$

and it is negligibly small if[7]

$$R \geqq 24 \left(\frac{\pi n}{\lambda} \right)^2 d_{11}^3. \tag{27}$$

The tolerable radius of curvature decreases rapidly with d_{11}.

Shorter radii of curvature can be negotiated in the plane perpendicular to the slabs if, as it happens in general, the field penetrations from the slabs into the media of indices n_1 and n_3 are smaller than d_{11}.

We can reuse Fig. 10 by substituting the parameter T/d_{11} with its equivalent

$$6.2 \left(\frac{n^2 T^3}{R\lambda^2} \right)^{\frac{1}{3}}$$

deduced from the equality in (27). For single-mode waveguides, the shortest d_{11} and consequently the shortest tolerable radius of curvature is achieved for $T/H \simeq T/W \simeq 0.5$.

The pertinent calculations for curvature-induced losses in multimode slab-coupled guides are carried on in Section III, where multimode single-material fibers are considered.
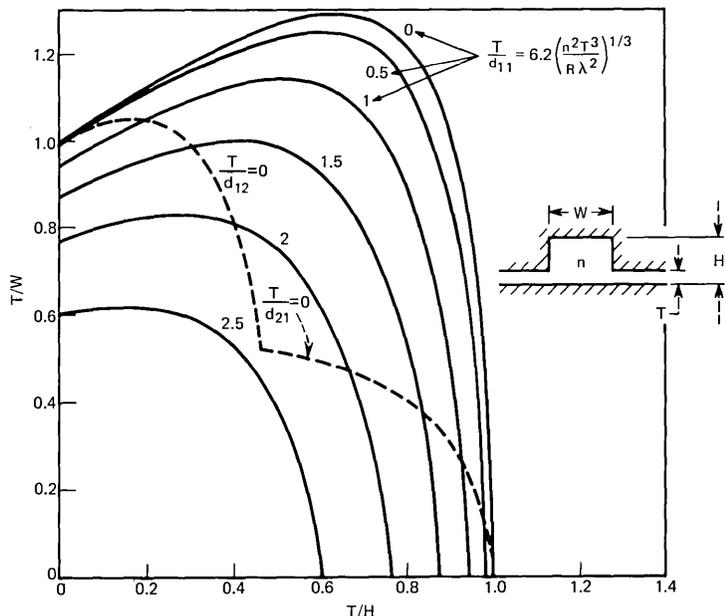
Fig. 10—Field penetration in slab d_{11} and tolerable radius of curvature R for fundamental mode.

III. MULTIMODE SINGLE-MATERIAL FIBERS

Single-material fibers supporting any number of modes, Fig. 1(a), are characterized by

$$n_1 = n_2 = n_3 = 1. \tag{28}$$

Under these circumstances, the location of the slab with respect to the core in Fig. 6(a) is not taken into account by the theory presented in this paper and, consequently, a more general cross section of single-material fibers is shown in Fig. 6(d). Figure 6(c) is still its equivalent (Appendix B).

Multimode single-material fibers satisfy not only (28) but also the following inequalities

$$\left.\begin{array}{c} \dfrac{h}{t} \\[2mm] \dfrac{w}{t} \\[2mm] \dfrac{2t}{\lambda} \end{array}\right\} \gg 1. \tag{29}$$

Therefore, parameters (2), (3), and (4) defining the guide are substantially simplified

$$T = t \tag{30}$$

$$W = w \tag{31}$$

$$H = h \tag{32}$$

and are valid for all polarizations.

According to (30), (31), and (32), the electromagnetic field is well confined within the guide. Using these values, the numerical aperture (16), the equivalent external refractive index (18), the number of modes for both polarizations (24), and the propagation constant for each mode (13) are

$$\text{N.A.} = \frac{\lambda}{2t} \ll 1, \tag{33}$$

$$n_e = \sqrt{n^2 - \left(\frac{\lambda}{2t}\right)^2} \cong n\left[1 - \frac{1}{8}\left(\frac{\lambda}{nt}\right)^2\right], \tag{34}$$

$$N = \frac{\pi}{2}\frac{S}{t^2}, \tag{35}$$

and

$$k_z = \sqrt{k^2 n^2 - \left(\frac{\pi p}{w}\right)^2 - \left(\frac{\pi q}{h}\right)^2}, \tag{36}$$

where S is the core cross-sectional area.

Unlike ordinary guides the number of modes N is independent of the free-space wavelength λ. In other words, by keeping λ fixed, the scale of the guide's cross section could be changed without varying the number of guided modes! How is it possible? The following is a plausible argument. For a given wavelength, if S is increased the number of guided modes in the core should increase, but simultaneously the number of modes that can escape through the enlarged slabs is also increased. The fact that both increases compensate for each other can only by justified with the mathematics.

Let us turn now to modal dispersion. Calling L the length of the guide and c the speed of light in free space, the group delay spread between any mode with propagation constant k_z and the fundamental one[9] (which has a propagation constant very close to that of a plane wave in a medium of refractive index n) is

$$\tau = \frac{L}{c}\frac{d}{dk}(k_z - kn). \tag{37}$$

With the help of (36)

$$\tau = \frac{Ln}{c}\left(\frac{kn}{k_z} - 1\right). \tag{38}$$

The maximum time spread occurs for the highest order mode which has the smallest k_z value, kn_e. Then, using (34) for the value of n_e,

$$\tau_{\max} = \frac{L}{c}n\left(\frac{n}{n_e} - 1\right) = \frac{L}{8nc}\left(\frac{\lambda}{t}\right)^2. \tag{39}$$

The impulse response is similar to that of clad fibers. A short impulse feeding equally all of the guided modes arrives at the other end as many impulses unequally displaced in time.[10] However, the power density of the arriving pulse is uniform over the time interval $\tau_{\max}$ given in (39) and zero elsewhere. This impulse response width being inversely proportional to the square of the slab thickness can be shortened by increasing t.

Since there is more familiarity with clad fibers than with single-material fibers, it is of interest to make a comparison between them,

assuming both guide the same number of modes N, and have the same modal dispersion spread τ_{max}. For a clad fiber of radius a and core and cladding refractive indices n and $n(1 - \Delta)$, those values are[11,9]

$$N = \left(\frac{2\pi a n}{\lambda} \right)^2 \Delta \tag{40}$$

and

$$\tau_{max} = \frac{L}{C} n\Delta. \tag{41}$$

These two equations together with (35) and (39) are plotted in Fig. 11. The group of curves on the lower part corresponds to single-material guides, and those on the upper part correspond to the equivalent clad fiber. Dotted lines are for the modal dispersion spread and solid lines for the number of modes. The parameters are either the core diameter of the clad fiber or the square root of the core cross section of the single-material fiber normalized in both cases to the free-space wavelength.

 Example: For a dispersion spread τ of 26 ns/km and $N/n = 100$, the single-material fiber dimensions are $t\sqrt{n}/\lambda = 4$ and $\sqrt{S}/\lambda = 31.6$, while those of the equivalent clad fiber are $n\Delta = 0.008$ and $2a/\lambda = 36$.

 If the multimode single-material fiber is bent, all the modes become somewhat lossy; however, as in ordinary clad fibers the radiation loss is significant only for those modes whose plane wave components exceed the critical angle. Unlike clad guides, though, a bend in the plane of the slabs produces higher losses than a similar bend in the perpendicular plane. This is due to the fact that, for a given mode, the field penetration in the slabs is far larger than in the material of index n_1.

 The single-material fiber bent in the plane of the slabs on a radius of curvature R has a numerical aperture N.A.$'$ and guides a number of modes[12] N', both of them smaller than the N.A. (33) and the number of modes (35) of the straight guide. As a matter of fact,

$$\text{N.A.}' = \frac{\lambda}{2t} \left[1 - \frac{w}{R} \left(\frac{2nt}{\lambda} \right)^2 \right] \tag{42}$$

and

$$N' = \frac{\pi}{2} \frac{S}{t^2} \left[1 - 2 \frac{w}{R} \left(\frac{2nt}{\lambda} \right)^2 \right]. \tag{43}$$

 Only half of the modes remain guided if

$$R_{\frac{1}{2}} = w \left(\frac{4nt}{\lambda} \right)^2. \tag{44}$$
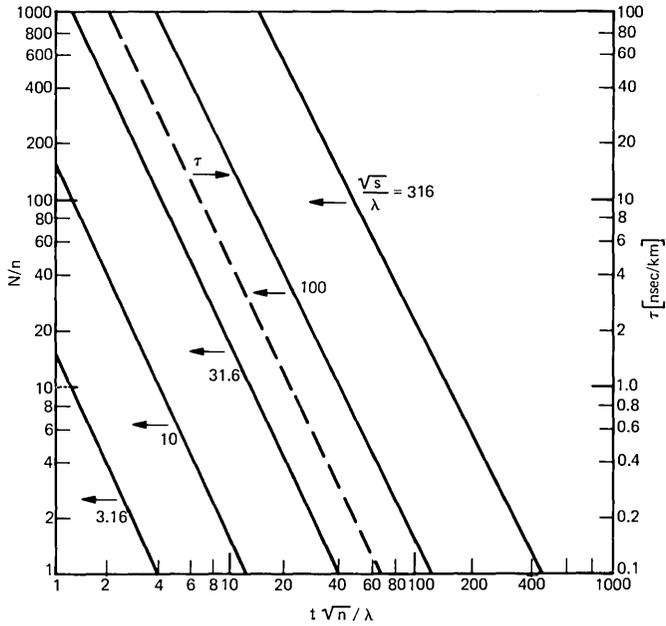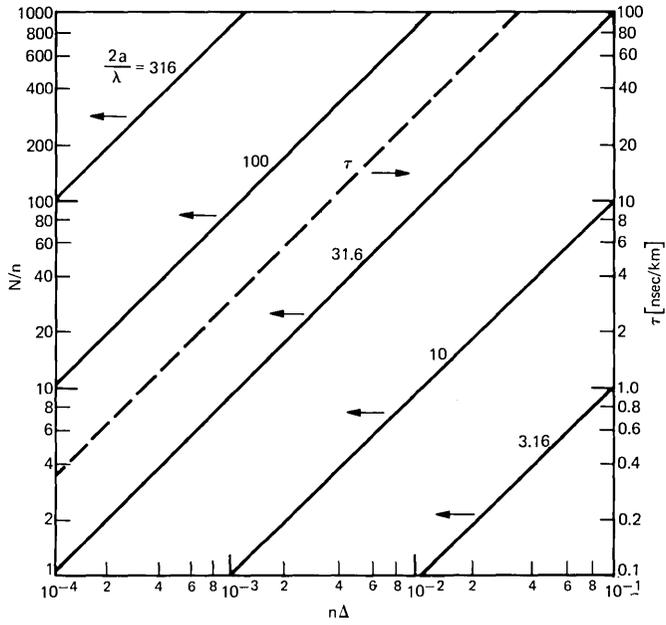
Fig. 11—Multimode single-material fiber and its equivalent clad fiber.

Example: For $n = 1.5, t/\lambda = 2$ and $w = 50\,\mu$, from the previous formula follows

$$R_{\frac{1}{2}} = 7.2 \text{ mm},$$

a small radius indeed.

IV. SINGLE-MODE SINGLE-MATERIAL FIBERS

In single-mode single-material fibers, Fig. 1(b), of interest for optical communication, all of the dimensions are large compared to the wavelength of operation. Therefore, eqs. (28) through (32) are valid and all formulas and figures of Section II related to the fundamental mode in slab-coupled guides are applicable just by changing T, H, and W into t, h, and w, respectively.

The dimensional requirements for single-mode operation with the next higher order at cutoff are determined by the solid line in Fig. 9. For example, it may be that for splicing purposes it is desirable to have a core of square cross section; then, from that figure,

$$h = w = 2t. \tag{45}$$

As in the multimode case, these dimensional requirements are independent of the wavelength and, consequently, the cross section of the guide can be scaled to satisfy other demands, such as relief of splicing tolerances, simplicity of fabrication, etc.

The propagation constant of the fundamental mode for both polarizations (13), the field penetration in the slabs (14), and the tolerable radius of curvature (27) are

$$k_z = \sqrt{k^2 n^2 - \pi^2 \left[\frac{1}{h^2} + \frac{1}{w^2(1 + c_1)^2} \right]}, \tag{46}$$

$$\frac{1}{d_{11}} = \pi \sqrt{\frac{1}{t^2} - \frac{1}{h^2} - \frac{1}{w^2(1 + c_1)^2}}, \tag{47}$$

and

$$R = \frac{24}{\pi} \left(\frac{n}{\lambda} \right)^2 \frac{t^3}{\left[1 - \frac{t^2}{h^2} - \frac{t^2}{w^2(1 + c_1)^2} \right]^{\frac{3}{2}}}, \tag{48}$$

where

$$c_1 = \frac{2}{\pi} \frac{t^2}{wh} \frac{1}{\sqrt{1 - \left(\frac{t}{h} \right)^2}}. \tag{49}$$

Normalized values of d_{11} and R can be found as parameters in Fig. 10

Table I

t/h	0.33	0.5	0.9	0.33
$\frac{t}{(\mu m)}$ \ t/w	0.53	0.51	0.33	0.1
2	0.23 mm	0.32 mm	2.6 mm	74 mm
5	2.1	4.9	40.0	115
10	17.0	40.0	330.0	925
Mode at Cutoff	$E_{21}^{x,y}$	$E_{21}^{x,y}$	$E_{21}^{x,y}$	$E_{12}^{x,y}$

for different values of t/h and t/w. Points on the dotted curve belong to single-mode guides with either the $E_{12}^{x,y}$ or $E_{21}^{x,y}$ modes at cutoff.

Continuing with the practical example above in which $h = w = 2t$, we obtain either from (47) or from Fig. 10

$$d_{11} = 0.42t. \tag{50}$$

With the help of (48) or Fig. 10, one can calculate a table of tolerable radii of curvature (Table I) for the fundamental mode in a guide with either the $E_{12}^{x,y}$ or $E_{21}^{x,y}$ modes at cutoff and assuming $n = 1.5$ and $\lambda = 1\ \mu m$.

Shorter radii of curvature are achieved for smaller t and t/h ratios if the guide is designed for the $E_{21}^{x,y}$ modes at cutoff (see Table I, first three columns). As seen from the last column, guides designed for $E_{12}^{x,y}$ modes at cutoff have longer tolerable radii of curvature. The guidance is not as tight and consequently less desirable.

Let us turn to dispersion. Knowing k_z (46) it is possible to calculate the dispersion $(L/c)(\partial k_z/\partial k)$ in a guide of length L. However, a more interesting result is the guide response to a Gaussian input pulse of $1/e$ width T_o. Following standard techniques to calculate responses through linear devices, one finds the output to be close to another Gaussian whose $1/e$ width is

$$T = T_o \sqrt{1 + \left[\frac{2L}{c^2 T_0^2} \frac{\partial^2 k_z}{\partial k^2} \right]^2_{k=k_o}}. \tag{51}$$

The second derivative is to be calculated at the wave number k_o of the carrier and c is the free-space speed of light.

For a given length of fiber L', the input pulse width T_0 that minimizes the output pulse width $(T' = \sqrt{2}T_o)$ is related to L' by

$$L' = \frac{c^2 T_0^2}{2} \frac{1}{\left(\dfrac{\partial^2 k_z}{\partial k^2} \right)_{k=k_o}}. \tag{52}$$

Assuming (46) to be applicable and $c_1 \ll 1$,

$$L' = 4\pi n \frac{c^2 T_0^2}{\lambda^3} \frac{h^2 w^2}{h^2 + w^2}. \tag{53}$$

Example: For

$$n = 1.5,$$
$$T_o = 10 \text{ ps},$$
$$h = w = 20 \ \mu m,$$

and

$$\lambda = 1 \ \mu m,$$

then,

$$L' = 34 \text{ km}.$$

As expected, the waveguide dispersion of the fundamental mode in a single-material fiber is very small and material dispersion may be more significant.[10]

V. RIB WAVEGUIDES

These slab-coupled guides, Figs. 2 and 6(b), are characterized by

$$n_2 = n_3 = 1, \tag{54}$$

$$\frac{n - n_1}{n} \ll 1, \tag{55}$$

and h slightly larger than t. Substituting (54) and (55) in (2), (3), and (4), the dimensions of the equivalent guides in Fig. 6(c) are

$$T = \begin{cases} t\left(1 + \dfrac{1}{v}\right) & \text{for } E_{pq}^x \text{ modes} \\[2mm] t\left(1 + \dfrac{n_1^2}{n^2 v}\right) & \text{for } E_{pq}^y \text{ modes}, \end{cases} \tag{56}$$

$$W = w \qquad \text{for } E_{pq}^{xy} \text{ modes}, \tag{57}$$

and

$$H = \begin{cases} h\left(1 + \dfrac{t}{hv}\right) & \text{for } E_{pq}^x \text{ modes} \\[2mm] h\left(1 + \dfrac{n_1^2 t}{n^2 hv}\right) & \text{for } E_{pq}^y \text{ modes}, \end{cases} \tag{58}$$

where

$$v = kt\sqrt{n^2 - n_1^2}. \tag{59}$$

Simplified by (54) and (55), eq. (11) says that expressions (56) and (58) are valid provided

$$v > \begin{cases} \dfrac{\pi}{2} - \sqrt{\dfrac{n^2 - n_1^2}{n_1^2 - 1}} & \text{for } E_{pq}^x \text{ modes} \\[4mm] \dfrac{\pi}{2} - \dfrac{1}{n^2}\sqrt{\dfrac{n^2 - n_1^2}{n_1^2 - 1}} & \text{for } E_{pq}^y \text{ modes.} \end{cases} \tag{60}$$

Using the values T, W, and H given by (56), (57) and (58) in previous equations and figures, the following results can be ascertained:

(*i*) Propagation constants of different modes and polarizations (13),

(*ii*) Maximum number of half periods P_{max} and Q_{max}, in the highest order modes $E_{1Q_{max}}^{x,y}$ and $E_{P_{max}1}^{x,y}$ [(19) and (20) or Fig. 8],

(*iii*) Dimensions of the guide for single-mode operation with the next higher order mode at cutoff [(17) or Fig. 9],

(*iv*) Field penetration in slabs d_{11} for the fundamental modes [(25) or Fig. 10],

(*v*) Tolerable radii of curvature in the plane of the slabs for the fundamental mode [(27) or Fig. 10].

VI. STRIP-COUPLED GUIDE

This guide, Figs. 3 and 6(b), is characterized by

$$n_3 = 1, \tag{61}$$

$$\frac{n - n_1}{n} \ll 1, \tag{62}$$

$$\frac{n - n_2}{n} \ll 1, \tag{63}$$

and

$$h = t. \tag{64}$$

Substituting (61) through (64) in (2), (3), and (4), the dimensions of the equivalent guide in Fig. 6(c) are

$$T = \begin{cases} t\left(1 + \dfrac{1}{v_1}\right) & \text{for } E_{pq}^x \text{ modes} \\[4mm] t\left(1 + \dfrac{n_1^2}{n^2 v_1}\right) & \text{for } E_{pq}^y \text{ modes,} \end{cases} \tag{65}$$

$$W = w \qquad\qquad\qquad \text{for } E_{pq}^{x,y} \text{ modes,} \tag{66}$$

$$H = \begin{cases} t\left(1 + \dfrac{1}{v_1} + \dfrac{1}{v_2}\tanh\dfrac{h_2 v_2}{t}\right) & \text{for } E_{pq}^x \text{ modes} \\[4mm] t\left(1 + \dfrac{n_1^2}{n^2 v_1} + \dfrac{n_2^2}{n^2 v_2}\tanh\dfrac{h_2 v_2}{t}\right) & \text{for } E_{pq}^y \text{ modes,} \end{cases} \tag{67}$$

where

$$v_{1,2} = kt\sqrt{n^2 - n_{1,2}^2}. \tag{68}$$

These expressions are valid only if, according to (11) simplified by (61) through (64),

$$v_1 > \begin{cases} \dfrac{\pi}{2} - \sqrt{\dfrac{n^2 - n_1^2}{n_1^2 - 1}} & \text{for } E_{pq}^x \text{ modes} \\[4mm] \dfrac{\pi}{2} - \dfrac{1}{n^2}\sqrt{\dfrac{n^2 - n_1^2}{n_1^2 - 1}} & \text{for } E_{pq}^x \text{ modes.} \end{cases} \tag{69}$$

It should be noticed that H cannot be made much larger than T just by increasing h_2 because the field decays almost exponentially along y in the material of index n_2, Fig. 6(b). As a matter of fact, the maximum value H is

$$h_{\max} = \begin{cases} t\left(1 + \dfrac{1}{v_1} + \dfrac{1}{v_2}\right) & \text{for } E_{pq}^x \text{ modes} \\[4mm] t\left(1 + \dfrac{n_1^2}{n^2 v_1} + \dfrac{n_2^2}{n^2 v^2}\right) & \text{for } E_{pq}^y \text{ modes.} \end{cases} \tag{70}$$

As in the rib waveguide the values of T, W, and H in (65), (66), and (67) can be entered in previous formulas and figures to find propagation constants of modes polarized along either x or y (13); number of guided modes [(19) and (20) or Fig. 8]; dimensions of guide for fundamental mode operation with the next higher order at cutoff [(17) or Fig. 9]; field penetration in slabs for the fundamental modes [(25) or Fig. 10]; tolerable radius of curvature for the fundamental mode [(27) or Fig. 10].

ACKNOWLEDGMENT

APPENDIX A

Approximate solution of the slab

Consider the two-layer slab in Fig. 12(a). Propagating along z, the field components are independent of x and the characteristic equa-
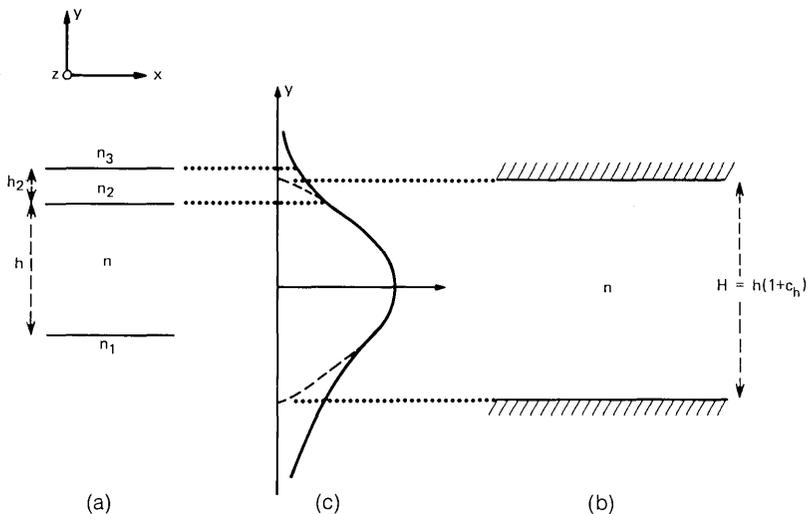
Fig. 12—(a) Two-layered slab. (b) Equivalent slab. (c) Field distributions (E_x and H_y or E_y and H_x) in original slab (solid line) and in equivalent slab (dotted line).

tion[13] is

$$\pi q = \psi + \tan^{-1} \frac{K_1 \psi}{\sqrt{v_1^2 - \psi^2}}$$

$$+ tan^{-1} \left\{ \frac{K_2 \psi}{\sqrt{v_2^2 - \psi^2}} \tanh \left[\frac{h_2}{h} \sqrt{v_2^2 - \psi^2} + \tanh^{-1} \frac{K_3}{K_2} \sqrt{\frac{v_2^2 - \psi^2}{v_3^2 - \psi^2}} \right] \right\}, \quad (71)$$

where

$$\psi = k_y h \tag{72}$$

is the electrical height of the slab of index n; q is the number of field extrema within the slab;

$$K_{1,2,3} = \begin{cases} \dfrac{n_{1,2,3}^2}{n^2} & \text{for polarization along } y \\ 1 & \text{for polarization along } z, \end{cases} \tag{73}$$

$$v_{1,2,3} = kh\sqrt{n^2 - n_{1,2,3}^2}. \tag{74}$$

Assuming

$$n > \begin{cases} n_1 \\ n_2, \\ n_3 \end{cases} \tag{75}$$

the field components vary sinusoidally along y in the medium of index n and exponentially in the others.

Real solutions of (71), which correspond to guided modes, exist for values of v_1 larger than

$$v_{min} = \pi \left(q - \frac{1}{2} \right) - \tan^{-1} \left\{ K_2 \sqrt{\frac{n^2 - n_1^2}{n_1^2 - n_2^2}} \right.$$
$$\left. \cdot \tanh \left[\frac{h_2}{h} \sqrt{v_2^2 - v_1^2} + \tanh^{-1} \frac{K_3}{K_2} \sqrt{\frac{n_1^2 - n_2^2}{n_1^2 - n_3^2}} \right] \right\}. \quad (76)$$

The simple asymptotic solution of (71) for

$$v_1 \gg v_{min} \quad (77)$$

is

$$\psi = \frac{\pi q}{1 + c_h}, \quad (78)$$

where

$$c_h = \frac{K_1}{v_1} + \frac{K_2}{v_2} \tanh \left(\frac{h_2}{h} v_2 + \tanh^{-1} \frac{K_3}{K_2} \frac{v_2}{v_3} \right). \quad (79)$$

The asymptotic solution (78) and the exact solution have been plotted in Figs. 13(a), (b), and (c) for several cases of interest.

Since the percentile errors are small, even close to the cutoff values of v_1, expression (78) will be used throughout the paper.

Now we proceed to find a much simpler slab [Fig. 12(b)] that is equivalent to that of Fig. 12(a) in the sense that both have the same propagation constants

$$k_y = \frac{\pi q}{h(1 + c_h)} \quad (80)$$

and

$$k_z = \sqrt{k^2 n^2 - k_y^2}. \quad (81)$$

That is indeed the case if the slab in Fig. 12(b) has refractive index n, height

$$H = h(1 + c_h), \quad (82)$$

and is surrounded by a hypothetical dielectric* such that all the transverse field components (components along x and y) vanish at the interfaces. This dielectric plays only the role of confining the electromagnetic field within the slab.

---

* This hypothetical dielectric has infinite conductivity if $E_x \neq 0$ or infinite permeability if $H_x \neq 0$.
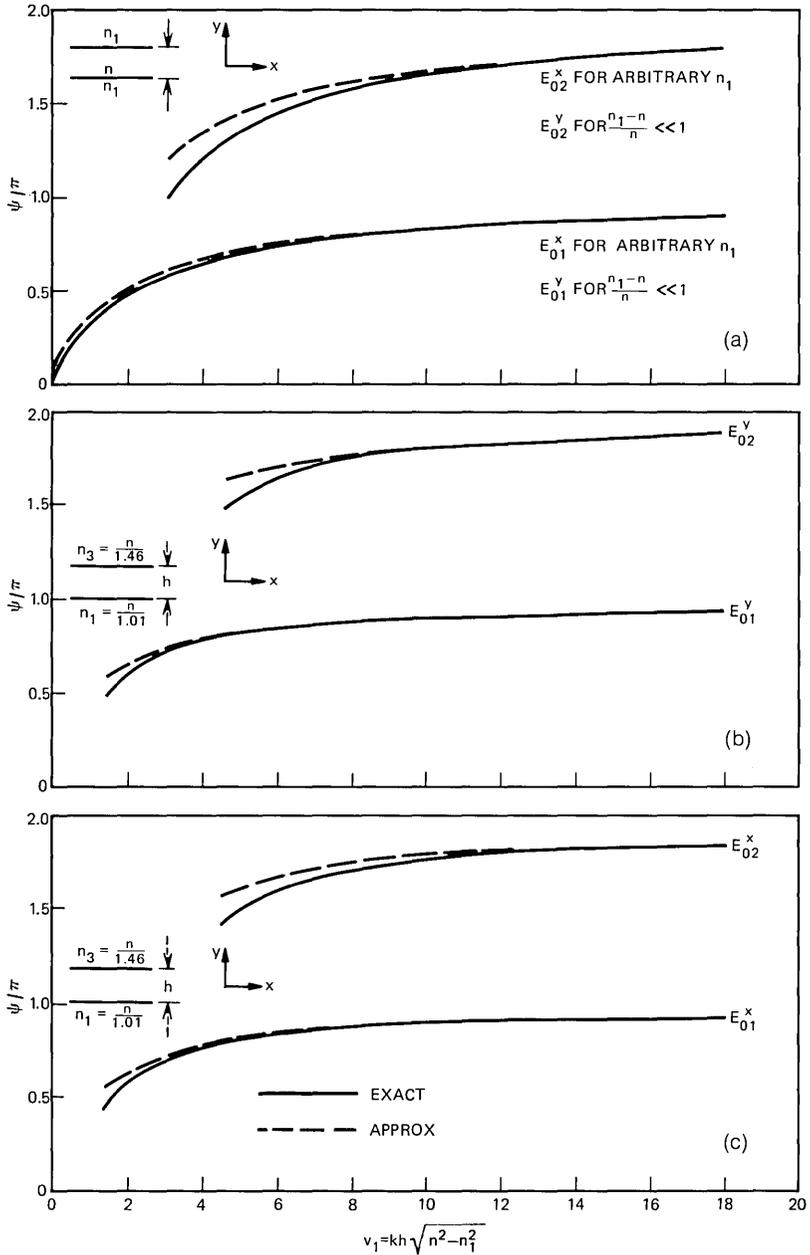
Fig. 13—Exact and approximate width of dielectric slabs. (a) Symmetric slab.
(b) Asymmetric slab (polarization along y). (c) Asymmetric slab (polarization
along x).

A geometrical interpretation of the equivalence of the slabs can be gained from the field intensity distributions shown in Fig. 12(c).

APPENDIX B

Solution of the slab-coupled guide

The exact solution of Maxwell's equations for the dielectric guide whose cross section is shown in Fig. 6(b) is very difficult because the boundaries are not analytical. However, a good quantitative insight can be gained if, as in Ref. 14, good guidance is assumed, that is, if most of the electromagnetic energy is contained within the guide. Then the field in the shaded areas, Fig. 6(b), can be ignored and the slab solution of Appendix A can be applied independently to each of the finite slabs of widths h, w, and t that make the guide. Thus, another dielectric guide, Fig. 6(c), is derived which is equivalent to that in Fig. 6(b) in the sense of having the same axial propagation constant k_z and the same field penetration in the slabs of thicknesses t and T. Unlike the original guide that has four dielectrics, the equivalent one has a single dielectric of index n and is surrounded by a hypothetical material that forces the transverse field components (along x and y) to be negligibly small on the boundaries and confines the electromagnetic energy within the guide. The dimensions of this equivalent guide are related to the original one by the following expressions derived from (82),

$$H = h(1 + c_h), \tag{83}$$

$$W = w(1 + c_w), \tag{84}$$

and

$$T = t(1 + c_t), \tag{85}$$

where

$$c_w = 2 \frac{K_w}{v_w}, \tag{86}$$

$$c_t = \frac{h}{t} \left(\frac{K_1}{v_1} + \frac{K_3}{v_3} \right), \tag{87}$$

$$K_w = \begin{cases} \dfrac{n_3^2}{n^2} & \text{for polarization along } x, \\ 1 & \text{for polarization along } y, \end{cases} \tag{88}$$

and

$$v_w = kw\sqrt{n^2 - n_3^2}, \tag{89}$$

and, because of the choice of symbols, c_h coincides with (79).

To solve the boundary value problem of Fig. 6(c) we make further assumptions. One is that the slabs do not perturb the sinusoidal distribution of field in the core and vice versa. Another assumption, justified qualitatively in the text, is that only the fundamental mode of the slabs of thickness T contribute significantly to determine the propagation constants of the modes of the guide. Then, the characteristic equations of the core and slabs are

$$k^2 n^2 = \left(\frac{\pi p}{W_e} \right)^2 + \left(\frac{\pi q}{H} \right)^2 + k_z^2 \tag{90}$$

$$k^2 n^2 = -k_{xs}^2 + \left(\frac{\pi}{T} \right)^2 + k_z^2, \tag{91}$$

where p and q are the number of half periods along x and y, k_{xs} is the propagation constant in the x direction within the slabs, and W_e is the equivalent width of the core. W_e is somewhat different from W because of the field penetration from the core into the slabs.

To calculate W_e, we imagine the core divided by the dotted line in two regions, a and b. In region a the width is W. In region b the electrical width $\phi = k_x W$ is given by the slab equation[13]

$$\pi p = \phi + 2 \tan^{-1} \frac{\phi}{k_{xs} W}. \tag{92}$$

With the value of k_{xs} deduced from (90) and (91) and considering that $W_e \cong W$, eq. (92) becomes

$$\pi p = \phi + 2 \tan^{-1} \frac{\phi}{\sqrt{ \left(\frac{\pi W}{T} \right)^2 - \left(\frac{\pi q W}{H} \right)^2 - (\pi p)^2 }} \tag{93}$$

and its asymptotic solution is

$$\phi = \frac{\pi p}{1 + \dfrac{2T}{\pi W} \dfrac{1}{\sqrt{1 - \left(\dfrac{qT}{H} \right)^2}}}. \tag{94}$$

Following the procedure described in Appendix A, the equivalent width of the region b of the core results in

$$W_b = W \left[1 + \frac{2T}{\pi W} \frac{1}{\sqrt{1 - \left(\dfrac{qT}{H} \right)^2}} \right]. \tag{95}$$

We will assume the equivalent width of the core W_e to be a linearly weighted average of W and W_b. Therefore,

$$W_e = \frac{W(H - T) + W_bT}{H}$$

or

$$W_e = W(1 + c_q), \tag{96}$$

in which

$$c_q = \frac{2}{\pi} \frac{T^2}{WH} \frac{1}{\sqrt{1 - \left(\dfrac{qT}{H}\right)^2}}. \tag{97}$$

From (90), (91), and (96), we derive the explicit values of the propagation constant of a mode and the penetration depth d_{pq} in the slab over which the field decays by $1/e$:

$$k_z = \sqrt{k^2n^2 - \left[\frac{\pi p}{W(1 + c_q)}\right]^2 - \left(\frac{\pi q}{H}\right)^2} \tag{98}$$

and

$$k_{xs} = \frac{1}{d_{pq}} = \frac{\pi}{T}\sqrt{1 - \left[\frac{pT}{W(1 + c_q)}\right]^2 - \left(\frac{qT}{H}\right)^2}. \tag{99}$$

These results apply not only to the guides in Figs. 6(b) and 6(c) but also to the somewhat more general guide in Fig. 6(a) provided that (i) the curved edges depart only slightly from those in Fig. 6(b), (ii) no exponential decay of the field or, equivalently, no turning point is introduced by the wall deformations, and (iii) h and w are chosen to be

$$h = \sqrt{\frac{h_{max}}{w_{max}}} s \tag{100}$$

and

$$w = \sqrt{\frac{w_{max}}{h_{max}}} s, \tag{101}$$

where h_{max} and w_{max} are the maximum height and width of the core portion of index n in Fig. 6(a) and s, its cross-sectional area.

If exponential decay of the field were introduced by the wall deformation, the simple expression (98) developed for k_z would not be applicable.

The choice of h and w in (100) and (101) are derived from the WKBJ method[15] or from the almost obvious demands:

$$hw = s$$

and

$$\frac{h}{w} = \frac{h_{\max}}{w_{\max}},$$

which mean that both portions of core with index n in Figs. 6(a) and 6(b) have equal surface and equal aspect ratio.

After so many approximations one wonders about the percentile errors in the final results (98) and (99). We can have some impression of the precision achieved by checking (98) against the more exact results developed elsewhere[11] in order to dimension an optical fiber of circular cross section at cutoff for the second mode, assuming small difference of refractive indexes between core and cladding.

Calling a the radius of the fiber core, the pertinent values for (99) are

$$n_1 = n_2 = n_3,$$

$$t = 0,$$

$$h = w = \sqrt{\pi}a \qquad \text{from (100) and (101),}$$

$$c_h = c_w = \frac{2}{\sqrt{\pi}V} \qquad \text{from (79) and (86),}$$

$$W = H = \sqrt{\pi}a\left(1 + \frac{2}{\sqrt{\pi}V}\right) \qquad \text{from (83) and (84),}$$

$$c_q = 0 \qquad \text{from (97),}$$

and

$$d_{pq} = 0,$$

where

$$V = ka\sqrt{n^2 - n_1^2}. \tag{102}$$

Substituting the values of d_{pq}, c_q, H, and W in (99), one obtains for $p = 1$ and $q = 2$, or for $p = 2$ and $q = 1$,

$$V = 2.83,$$

while the exact result is 2.4. The error of 18 percent is small indeed considering that at cutoff the assumption of negligible field outside of the guide is crudely violated.

REFERENCES

1. P. Kaiser, E. A. J. Marcatili, and S. E. Miller, "A New Optical Fiber," B.S.T.J., *52*, No. 2 (February 1973), pp. 265–269.
2. P. Kaiser and H. W. Astle, "Low-Loss Single-Material Fibers Made from Pure Fused Silica," to be published in B.S.T.J., *53*, No. 7 (September 1974).
3. J. E. Goell, "Rib Waveguide for Integrated Optical Circuits," J. Appl. Optics, *12*, No. 12 (December 1973), pp. 2797–2798.
4. H. Noda, H. Furuta, and A. Ihaya, "A Novel Optical Waveguide for Integrated Optics," IEEE/OSA (1973), Conference on Laser Engineering and Applications, Washington, May 30–June 1, 1973.
5. V. Ramaswamy, "Strip-Loaded Film Waveguide," B.S.T.J., this issue, pp. 697–704.
6. J. A. Arnaud, "Transverse Coupling in Fiber Optics—Part II: Coupling to Mode Sinks," B.S.T.J., this issue, pp. 675–696.
7. E. A. J. Marcatili, "Bends in Optical Dielectric Guides," B.S.T.J., *48*, No. 7 (September 1969), pp. 2103–2132.
8. E. A. J. Marcatili and S. E. Miller, "Improved Relations Describing Directional Control in Electromagnetic Wave Guidance," B.S.T.J., *48*, No. 7 (September 1969), pp. 2161–2188.
9. D. Gloge, "Dispersion in Weakly Guiding Fibers," J. Appl. Optics, *10*, No. 11 (November 1971), pp. 2442–2445.
10. S. E. Miller, E. A. J. Marcatili, and Tingye Li, "Research Toward Optical-Fiber Transmission Systems," Proc. IEEE, *61*, No. 12 (December 1973), pp. 1703–1751.
11. D. Gloge, "Weakly Guiding Fibers," J. Appl. Optics, *10*, No. 10 (October 1971), pp. 2252–2258.
12. D. Gloge, "Bending Loss in Multimode Fibers with Graded and Ungraded Core Index," J. Appl. Optics, *11*, No. 11 (November 1972), pp. 2506–2512.
13. P. K. Tien, R. J. Martin, and G. Smolinsky, "Formation of Light-Guiding Interconnections in an Integrated Optical Circuit by Composite Tapered Film Coupling," J. Appl. Optics, *12*, No. 8 (August 1973), pp. 1909–1917.
14. E. A. J. Marcatili, "Dielectric Rectangular Waveguide and Directional Coupler for Integrated Optics," B.S.T.J., *48*, No. 7 (September 1969), pp. 2071–2102.
15. P. M. Morse and H. Feshback, *Methods of Theoretical Physics*, New York: McGraw-Hill, 1953, pp. 1092–1099.

Transverse Coupling in Fiber Optics
Part II: Coupling to Mode Sinks

By J. A. ARNAUD

The number of modes that can propagate without radiation loss in oversized waveguides is sharply reduced if the waveguide is coupled to a structure supporting radiation modes, the loss mechanism being analogous to Cerenkov radiation. The coupling formula derived in Part I[1] is used to evaluate the loss for a specific configuration: a reactive surface (e.g., a thin dielectric slab) acting as a waveguide, coupled to a semi-infinite dielectric acting as a mode sink. The method consists in first assuming that the substrate is finite in size and lossy and adding the losses associated with each substrate mode. The substrate dimensions are subsequently made infinite and the dissipation loss is made to vanish. The expression obtained for the radiation loss coincides with an expression obtained by solving the boundary value problem. The method is then applied to the problem of mode selection for dielectric rods coupled to dielectric slabs, which is of particular importance for optical communications and integrated optics. A 2-dB/m radiation loss is calculated for the first higher order mode when the rod radius is 10 μm, λ = 1 μm, n = 1.41, and the rod-to-slab spacing is 0.15 μm.

I. INTRODUCTION

An expression for the coupling between lossy single-mode open waveguides was derived in Part I.[1] We now investigate the coupling of a waveguide with finite cross section with a waveguide with infinite cross section (called a substrate), the latter supporting radiation modes. Radiation losses are suffered whenever the propagation constant h of the guided mode is smaller than the highest propagation constant h_s of the radiation modes carried by the substrate. Radiation then takes place at the Cerenkov angle $\theta = \cos^{-1}(h/h_s)$. By properly choosing the dimensions and permittivities of the waveguide and those of the

substrate, it is possible to reduce the number of modes that can propagate without attenuation (in the absence of dissipation and scattering losses). This arrangement is of great practical importance because optical fibers are usually highly overmoded to facilitate fabrication and splicing.[2] (For a coherent source, it is important to reduce the number of modes because different modes usually have different group velocities. If a short optical pulse is sent through the fiber, mode conversion takes place because of the imperfections of the fiber; this causes the pulse to spread in time.) The mode selection mechanism just described is also of practical importance in the microwave range for oversized waveguides such as oversized microstrips on dielectric substrates and oversized dielectric strips.* Multimoding in traveling wave tubes can also be avoided with the help of mode sinks.

We investigate the loss mechanism for two specific configurations. First, a reactive surface acting as a waveguide coupled to a semi-infinite dielectric acting as a mode sink. We show that, by adding the losses associated with each substrate mode, an expression for the total loss is obtained that coincides with an expression obtained by solving the boundary value problem. Then the method is applied to the problem of a dielectric rod coupled to a dielectric slab.[2] The case of dielectric rods coupled to dielectric cylinders supporting whispering gallery modes and acting as mode sinks[3] will be discussed in another paper.

II. RADIATION LOSSES IN SUBSTRATES—GENERAL FORMULA

To evaluate the radiation losses, let us first assume that the transverse dimensions of the substrate are finite, and let $h_{sz} = h_{sr} + ih_{si}$ be the propagation constant of a trapped mode in the substrate, with h_{sr} real and h_{si} real positive (the subscript s stands for "substrate").[†] If h_o denotes the propagation constant of a trapped mode of the waveguide in the absence of the substrate, the propagation constant h of the coupled wave is, from eq. (6a) in Part I,

$$h = h_o + \tfrac{1}{2}(h_{sz} - h_o) - [\tfrac{1}{4}(h_{sz} - h_o)^2 + C^2]^{\frac{1}{2}}, \qquad (1)$$

where $C^2 \equiv c_a c_b / P_a P_b$ denotes the coupling coefficient defined in Part I. The minus sign before the square root has been selected because it

---

* In the microwave range, there are no compelling reasons for using dielectric waveguides that are large compared with the wavelength in all dimensions, but we may want to use strips (either metallic or dielectric) whose widths exceed one wavelength for improved accuracy.

† The dependence of the field on time (t) and on the axial coordinate (z) is denoted $\exp[i(hz - \omega t)]$. This term is henceforth omitted.

corresponds to the mode whose field is concentrated in the waveguide cross section rather than in the substrate (that is, we require $h = h_o$ when $C^2 = 0$).

Let us now assume that h_o is real (lossless waveguide) and that

$$h_{si} \gg C. \tag{2}$$

Using this condition, eq. (2), we can expand the r.h.s. of eq. (1) in power series of C^2 and keep only the first two terms in the expansion. The loss is given by the imaginary part h_i of h. Because the imaginary part of C^2 can be neglected in the case that we consider, we have

$$h_i \approx C^2 h_{si} [(h_{sr} - h_o)^2 + h_{si}^2]^{-1}. \tag{3}$$

The total loss $\mathcal{L}$ experienced by the waveguide is now obtained by summing over the various modes of the substrate:

$$\mathcal{L} = \sum_\alpha C_\alpha^2 h_{si} [(h_{sr\alpha} - h_o)^2 + h_{si}^2]^{-1}, \tag{4}$$

where the subscript α refers to the substrate modes. We have assumed, for simplicity, that h_{si} does not depend on α. It is shown in the next section for a simple configuration that in the limit of dense substrate modes eq. (4) is in agreement with an exact result, obtained from a boundary value method.

If we let the cross-section area S of the substrate tend to infinity, the substrate modes become denser and denser, and the summation in eq. (4) can be replaced by an integral

$$\mathcal{L} = \lim_{S \to \infty} \sum_\alpha C_\alpha^2 h_{si} [(h_{sr\alpha} - h_o)^2 + h_{si}^2]^{-1}$$

$$= \int \mathcal{C}(h_{sr}) h_{si} [(h_{sr} - h_o)^2 + h_{si}^2]^{-1} dh_{sr}, \tag{5}$$

where we have defined a coupling density $\mathcal{C}$ by

$$\mathcal{C}(h_{sr}) dh_{sr} = \lim_{S \to \infty} \sum_\alpha C_\alpha^2,$$

the range of α being defined by the condition

$$h_{sr} < h_{sr\alpha} < h_{sr} + dh_{sr}. \tag{6}$$

This density exists because, as $S \to \infty$, the coupling coefficient C^2 decreases at least as fast as S^{-1}, the power in the substrate being proportional to S if the power density is kept a constant.

We can now let h_{si} tend to zero, the condition eq. (2) being preserved. The second factor in the integrand of eq. (5) is sharply peaked

at $h_{sr} = h_o$ and behaves as a symbolic δ-function. Thus, in the limit $h_{si} \to 0$ we have

$$\mathcal{L} = \pi \mathcal{C}(h_o). \tag{7}$$

It should be noted that the subscript α in eqs. (4) to (6) stands for three subscripts m, n, and s, where m refers to modes in the x direction, n refers to modes in the y direction (we assume for simplicity that the substrate modes are separable in Cartesian coordinates), and s refers to the state of polarization (e.g., H or E modes).

III. COUPLING TO A SEMI-INFINITE SUBSTRATE

Consider a reactive surface coupled to a semi-infinite dielectric (Fig. 1). We consider only H modes and assume that the field is
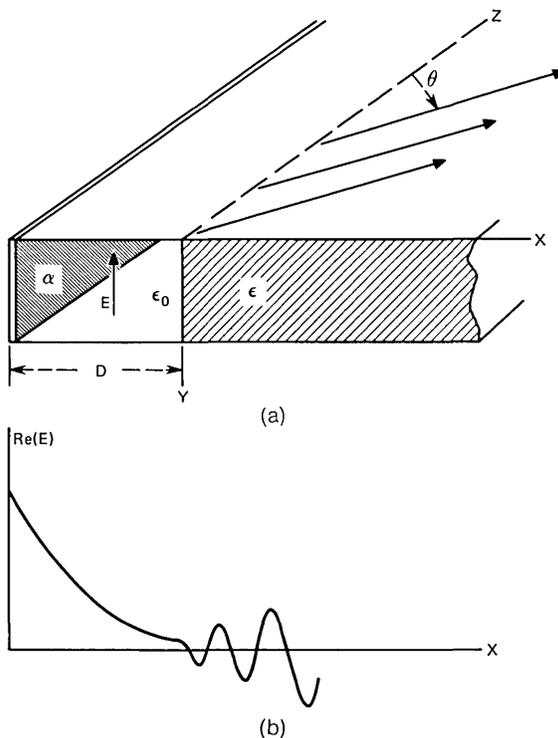
Fig. 1—(a) Reactive surface, with normalized susceptance α, coupled to a semi-infinite dielectric with permittivity $\epsilon \equiv n^2\epsilon_o$. For H modes, the structure is assumed terminated in the y direction by electric walls. Radiation takes place at the Cerenkov angle $\theta = \cos^{-1}\left[(k^2 + \alpha^2)^{\frac{1}{2}}/kn\right]$, $k \equiv 2\pi/\lambda$. (b) Variation of the field as a function of x.

independent of the y coordinate. Except for the changes $x \rightleftarrows y$ and $i \rightarrow -i$, we use Shevchenko's notation.[4]

For waves propagating along the z axis, the electric field has only a y component that we denote E. In a region with constant ϵ, E obeys the wave equation

$$d^2E/dx^2 + (\omega^2\epsilon\mu_o - h^2)E = 0,$$
$$H_x = -h(\omega\mu_o)^{-1}E,$$
$$H_z = (i\omega\mu_o)^{-1}dE/dx. \tag{8}$$

If ϵ has a finite discontinuity, E and dE/dx remain continuous. The general solution of eq. (8) for $\epsilon = \epsilon_o$ and ϵ are, respectively,

$$E = A^+e^{ix x} + A^-e^{-ix x}, \qquad (\epsilon_o) \tag{9a}$$

$$E_s = A_s^+e^{igx} + A_s^-e^{-igx}, \qquad (\epsilon) \tag{9b}$$

$$\chi^2 \equiv \omega^2\epsilon_o\mu_o - h^2, \tag{10a}$$

$$g^2 \equiv \omega^2\epsilon\mu_o - h^2 = u^2 + \chi^2,$$
$$u^2 \equiv \omega^2(\epsilon - \epsilon_o)\mu_o. \tag{10b}$$

The loss can be evaluated by solving the boundary value problem. At the reactive surface ($x = -D$), we have the condition (see Ref. 4)

$$dE/dx + \alpha E = 0, \qquad x = -D, \tag{11}$$

where α is a positive real number proportional to the susceptance of the surface.* We assume that, in the dielectric, the wave propagates away from the structure, that is,

$$E_s = A_s^+e^{igx}. \tag{12}$$

Note that h is expected to have a small positive imaginary part expressing the radiation loss in the dielectric. Assuming that ϵ is real, that is, that the dielectric is free of dissipation losses, eq. (10b) shows that g has a small negative imaginary part. Thus, the wave amplitude grows exponentially as the distance to the structure increases. This solution of Maxwell's equations is called a "leaky wave."[4] It is not difficult to show that the curves of constant irradiance in the dielectric are straight lines making with the z axis an angle $\theta = \cos^{-1}(h_o/kn)$ (Cerenkov angle).

---

*A thin dielectric slab with permittivity ϵ and thickness d, supported by a magnetic wall, is equivalent to a reactive surface with normalized susceptance $\alpha = \omega^2(\epsilon - \epsilon_o)\mu_o d$. An equivalent configuration, obtained by symmetry with respect to the magnetic wall, is a thin slab of width $2d$ with dielectrics symmetrically located on both sides. Note that α has the dimension of a propagation constant.

From eq. (12), the boundary condition at $x = 0$ is

$$dE/dx - igE = 0, \qquad x = 0. \tag{13}$$

From eqs. (9a), (11), and (13), we obtain the equation defining χ, or h,

$$(\chi - i\alpha)(\chi + g) = (\chi + i\alpha)(\chi - g) \exp(2i\chi D). \tag{14}$$

If we let αD tend to infinity, the reactive surface is uncoupled to the dielectric and eq. (14) reduces to $\chi \equiv \chi_o = i\alpha$; that is,

$$\chi^2 \equiv \chi_o^2 \equiv \omega^2 \epsilon_o \mu_o - h_o^2 = -\alpha^2, \tag{15a}$$

$$g^2 \equiv g_o^2 = \omega^2(\epsilon - \epsilon_o)\mu_o + \chi_o^2. \tag{15b}$$

Equation (15a) defines the propagation constant h_o of the uncoupled reactive surface.

Let us now consider

$$\exp(2i\chi_o D) \equiv \delta \tag{16}$$

as a small parameter and set

$$\chi = \chi_o + \chi_1 \delta + \cdots, \tag{17}$$
$$g = g_o + g_1 \delta + \cdots,$$

in eqs. (14) and (10b). Collecting terms of first order in δ we get

$$\chi_1 = 2i\alpha(i\alpha - g_o)/(i\alpha + g_o). \tag{18}$$

From eqs. (10a) and (17) we have, to first order,

$$\mathrm{Im}(h) = -(\alpha\delta/h_o)\,\mathrm{Re}(\chi_1). \tag{19}$$

Thus the loss $\mathfrak{L} \equiv \mathrm{Im}(h)$ is

$$\mathfrak{L} = 4\alpha^3 u^{-2} h_o^{-1} g_o \exp(-2\alpha D), \tag{20a}$$

or, explicitly, in terms of k, n, D, and α,

$$\mathfrak{L} = 4\alpha^3 [k^2(n^2 - 1)]^{-1}(k^2 + \alpha^2)^{-\frac{1}{2}}$$
$$\times [k^2(n^2 - 1) - \alpha^2]^{\frac{1}{2}} \exp(-2\alpha D). \tag{20b}$$

If the micron is used as the unit of length, the loss in dB/km is obtained by multiplying the r.h.s. of eq. (20b) by 8.7×10^9.

This expression for the loss, applicable to small couplings, can be obtained alternatively from the equality

$$h - h_o = \omega \int (\epsilon - \epsilon_o)\mathbf{E}^+ \cdot \mathbf{E}_p dS \Big/ \int (\mathbf{E}^+ \times \mathbf{H}_p - \mathbf{E}_p \times \mathbf{H}^+) \cdot d S, \tag{21}$$

where $(\mathbf{E}, \mathbf{H})$ and h_o denote the field and propagation constant of the

wave guided by the reactive surface in the absence of the dielectric and ($\mathbf{E}^+$, $\mathbf{H}^+$) denotes the field adjoint to ($\mathbf{E}$, $\mathbf{H}$) (see Part I). ($\mathbf{E}_p$, $\mathbf{H}_p$) and h denote the field and propagation constant in the presence of the dielectric. The integral in the numerator extends to the dielectric cross section, and the integral in the denominator extends to the whole cross section. Equation (21) is exact and is readily obtained from Maxwell's equations.[*] The field ($\mathbf{E}_p$, $\mathbf{H}_p$), unfortunately, is not known. It may differ considerably from the unperturbed field ($\mathbf{E}$, $\mathbf{H}$) when the dielectric supports modes almost synchronous with the waveguide mode. This is why this expression, eq. (21), is, in general, not practical to evaluate the coupling between waveguides, or waveguides and mode sinks. The configuration presently considered, however, is sufficiently simple to be handled on the basis of eq. (21).

For our case, eq. (21) becomes, with the approximation $h \approx h_o$,

$$h - h_o \approx - (\omega^2 \mu_o / 2h_o) \int_0^\infty (\epsilon - \epsilon_o) E E_p dx \Big/ \int_{-D}^\infty E^2 dx. \qquad (22)$$

The unperturbed field, normalized to unity at $x = -D$, is

$$E = \exp (i\chi x) \exp (i\chi D). \qquad (23)$$

The perturbed field is obtained by assuming as before an $\exp (igx)$ dependence in the dielectric, matching E and dE/dx at the vacuum-dielectric interface ($x = 0$), and stating that $E_p \approx 1$ at $x = -D$. We obtain

$$E_p = 2(1 + g/\chi)^{-1} \exp (i\chi D) \exp (igx), \qquad x \geqq 0. \qquad (24)$$

Substituting in eq. (22) and integrating, a result identical to eq. (20) is obtained.

Let us now apply to the same problem the method explained in Section II of this paper, which consists in adding the losses associated with each mode of the substrate. The coupling coefficient between two H modes, with fields E and E_s, was given in Part I. With our present notation we have

$$C^2 = \alpha^2 h_o^{-2} \Big(E^2 \Big/ \int E^2 dx \Big) \Big(E_s^2 \Big/ \int E_s^2 dx \Big), \qquad (25)$$

where the integrals are over the whole cross section, and E, E_s are defined at some point located between the two waveguides.

---

[*] The contribution at infinity is assumed to vanish. Thus, it is implicitly assumed that the rate of decay of the unperturbed field exceeds the rate of growth of the perturbed field. This condition is always satisfied for small couplings.

The field E of the reactive surface alone is, as we have seen,

$$E = \exp(-\alpha x). \tag{26}$$

Thus, at $x = 0$,

$$E^2 \Big/ \int_{-D}^{\infty} E^2 dx = 2\alpha \exp(-2\alpha D). \tag{27}$$

Let us consider next the dielectric alone and first assume that its thickness L_x is finite. By matching E and dE/dx at $x = 0$ and $x = L_x$, we obtain the field at the vacuum dielectric interface, and

$$E_s^2 \Big/ \int_{-\infty}^{+\infty} E_s^2 dx \approx E_s^2 \Big/ \int_0^{L_x} E_s^2 dx = 2g_o^2 u^{-2} L_x^{-1}. \tag{28}$$

Substituting eqs. (27) and (28) in eq. (25), we obtain the coupling coefficient

$$C^2 = 4\alpha^3 g_o^2 u^{-2} h_o^{-2} \exp(-2\alpha D) L_x^{-1}. \tag{29}$$

Let us now evaluate the number of modes $(N \; dh)$ in the dielectric whose propagation constants lie between h and $h + dh$. Because we are far from cut-off, the boundary condition is almost the same as for a metallic waveguide, $E = 0$. Thus, the condition on g is

$$g_m = m\pi/L_x, \qquad m = 1, 2, \cdots. \tag{30}$$

Using the relation

$$g_m^2 = \omega^2 \epsilon \mu_o - h^2, \tag{31}$$

the mode number density is, from eq. (30),

$$N = hg^{-1} L_x/\pi. \tag{32}$$

The radiation loss is obtained from eqs. (29), (32), and (7), and $h = h_o$, $g = g_o$,

$$\mathcal{L} = \pi C^2 N = 4\alpha^3 u^{-2} h_o^{-1} g_o \exp(-2\alpha D). \tag{33}$$

This result coincides with the result eq. (20) obtained by taking the limit of large D in the exact solution. The variation of the loss expressed in dB/km is given in Fig. 2 as a function of the normalized susceptance α of the surface, for $\lambda = 1 \; \mu m$, $\epsilon/\epsilon_o = 2$, and $D = 1.5$, 1.75, and $2 \; \mu m$.

For comparison, when the dielectric permittivity has the form $\epsilon = \epsilon_o + i\epsilon_i$ (the dielectric is perhaps a lossy foam) and the spacing D is chosen as large as consistent with a loss of 10 dB/km at $\alpha = 6.28$, the loss experienced is shown on the same figure as a dotted line. The comparison clearly shows the advantage of mode sinking over dissipation for mode selection.
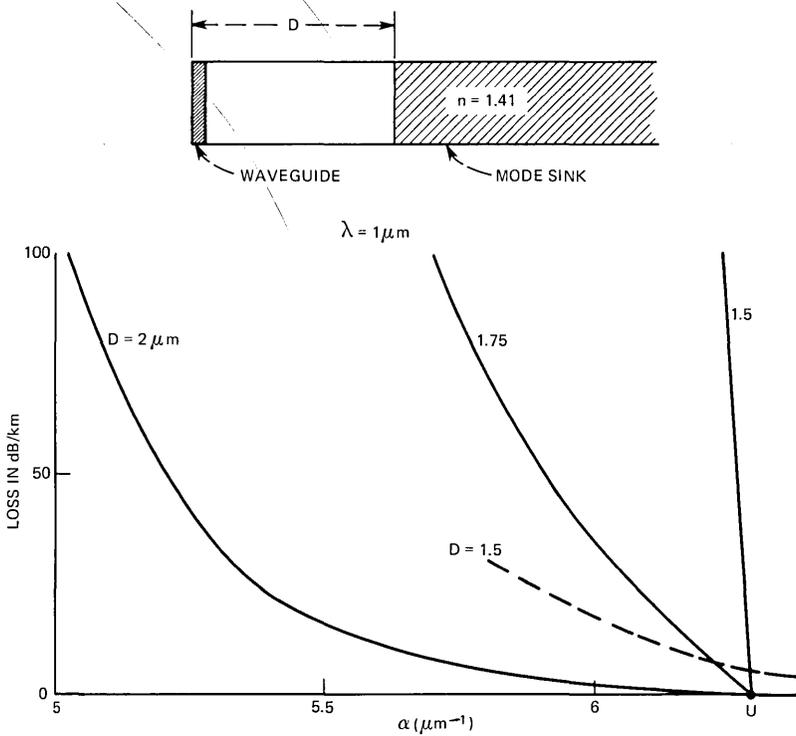
Fig. 2—Radiation loss in dB/km as a function of the normalized surface suscep-tance α of the waveguide for a wavelength $\lambda = 1\ \mu m$, $n^2 = 2$, and $D = 1.5, 1.75,$ and $2\mu m$. The dotted line is applicable to a dissipative dielectric.

IV. COUPLING TO PLANAR SUBSTRATES

Let us now consider a waveguide with propagation constant h_1 coupled to a substrate that extends to infinity in the y direction, but has a finite thickness in the x direction. This substrate is perhaps a reactive plane (e.g., a corrugated conductor) or a dielectric slab, as illustrated in Fig. 3. In any case, homogeneity of the substrate in the y, z plane is assumed.

Because of the assumed homogeneity of the substrate, plane wave solutions

$$E_s(x, y, z) = E_s(x) \exp\ (ih_{sy}y + ih_{sz}z), \tag{34}$$

where

$$h_{sz} = f(h_{sy}, \omega), \tag{35}$$

exist at some angular frequency ω (ω is now considered a fixed param-eter and is omitted).
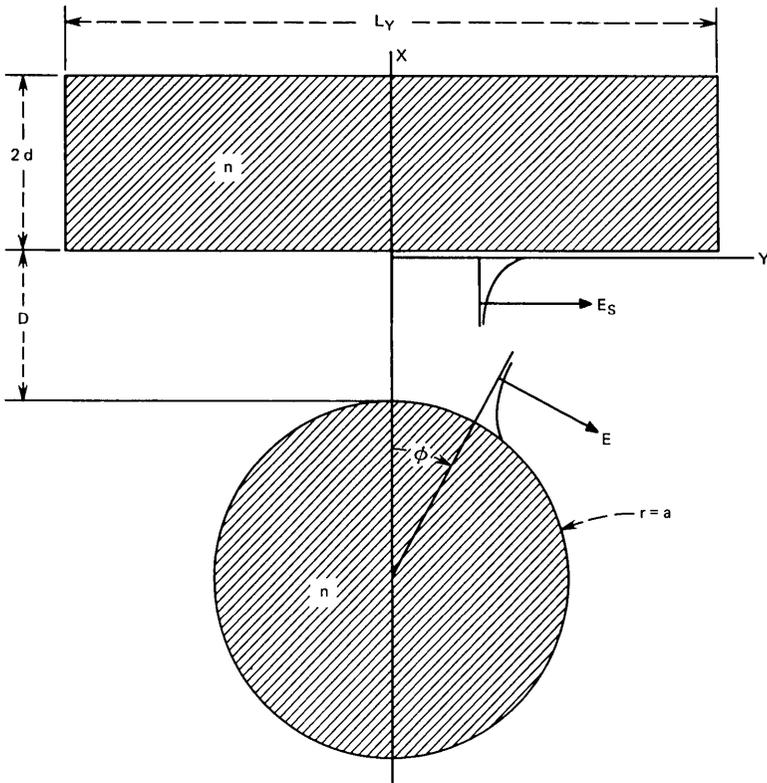
Fig. 3—Dielectric rod coupled to a dielectric slab. The rod field E is shown for the spurious H_{01} mode, and the slab mode is $H_{1\nu}$ (ν is a continuous index in the limit $L_y \to \infty$). Coupling takes place at $\phi \approx 0$.

In the discussion that follows, we consider only waveguide and substrate modes that are even in y. Assuming that f is even in h_{sy} and that the slab is terminated by electric walls, even modes satisfy the relation

$$h_{sy}L_y = 2n\pi, \qquad n = 0, 1, 2 \cdots, \tag{36}$$

where L_y denotes the width of the substrate. L_y will be later assumed to tend to infinity. The density N of even modes is from eqs. (35) and (36)

$$N = (df/dh_{sy})^{-1}(L_y/2\pi). \tag{37}$$

If the substrate is isotropic, with wave vector h_s, eq. (35) is

$$h_{sz} \equiv f(h_{sy}) = (h_s^2 - h_{sy}^2)^{\frac{1}{2}}, \tag{38}$$

and the mode density is, from eq. (37),

$$N = (h_{sz}/h_{sy})L_y/2\pi. \tag{39}$$

The loss is then obtained from eqs. (7) and (39).

$$\mathcal{L} = \tfrac{1}{2}[h_1/f^{-1}(h_1)]C^2L_y, \tag{40}$$

the coupling coefficient C^2 being evaluated from eq. (6) in Part I.

It should be noted that, when the propagation constant of the wave-guide mode (h_1) is just equal to the propagation constant (h_s) of the 2-dimensional substrate, h_{sy} is equal to zero and the loss, according to eq. (40), is infinite if C^2L_y remains finite. (This was not the case for the 3-dimensional mode sinks considered in Section III because, as $L_x \to \infty$, the field at the surface of the dielectric tends sufficiently rapidly to zero to make C^2L_x vanish in the limit.) This infinity at $h_1 = h_s$ would be removed if some finite dissipation loss in the substrate were present. Even in the absence of dissipation losses, the radiation loss remains finite at $h_1 = h_s$, because the perturbation method on which eq. (40) is based is no longer applicable. The peak in the loss curve predicted by eq. (40) (analogous to a sound barrier) is pronounced only for small couplings.

Our general result, eq. (40), is now applied to a dielectric rod coupled to a dielectric slab. The thickness and permittivity of the slab can always be chosen in such a way that only the fundamental mode of the rod propagates without radiation loss. The calculation of the loss of higher-order modes is carried out for the case where the rod diameter and the slab thickness are very large compared with the wavelength; that is, when the rod is highly multimoded in the absence of coupling.

Approximate expressions for the modes and propagation constant in the slab and the rod are given in the next subsections.

4.1 Modes of the slab

Let us consider first the modes in the dielectric slab. If the thickness $2d$ of the slab is very large (more precisely, if $\omega^2(\epsilon - \epsilon_o)\mu_o d^2 \gg 1$), the propagation constant of the fundamental H_1 mode is approximately given by the condition that the field E vanishes at the boundary

$$E(x, z') \approx E_{so} \cos(g_s x) \exp(ih_s z').$$

Thus, we have*

$$g_s^2 \equiv \omega^2 \epsilon \mu_o - h_s^2 = (\pi/2d)^2. \tag{41}$$

---

* A more accurate and general expression is (see Ref. 5) $g_s d = m(\pi/2)(1 - V^{-1})$ for H modes and $g_s d = m(\pi/2)(1 - n^{-2}V^{-1})$ for E modes, where $m = 1, 2 \cdots$ is the mode number and $V \equiv ud$. These expressions show that the H_1 mode that we are considering in this section is the fundamental mode; that is, the mode that has the largest propagation constant. The difference Δh in propagation constants is, for $m = 1$, equal to $d^{-1}(\pi/2knd)^2(1 - 1/n^2)^{\frac{1}{2}}$.

(The axial coordinate is denoted z' instead of z to avoid changing our notation when waves propagating at some angle to the z axis are considered. The origin of the x axis is, in this subsection, at the center of the slab.) The axial (z') and transverse (x) components of the magnetic field are, within the slab, as we have seen before

$$H_x = -h_s(\omega\mu_o)^{-1}E, \tag{42}$$

$$H_{z'} = (i\omega\mu_o)^{-1}dE/dx, \tag{43}$$

and the power per unit width is approximately

$$P \approx -\int_{-d}^{+d} EH_x dx = dh_s(\omega\mu_o)^{-1}E_{so}^2. \tag{44}$$

The field at the boundary is in fact not exactly equal to zero. To obtain its value, we use the fact that the dependence of E on x in vacuum is $\exp(-p_s x)$, where $p_s^2 = h_s^2 - \omega^2\epsilon_o\mu_o$, and the continuity of dE/dx. We obtain

$$E(d) = (\pi/2d)p_s^{-1}E_{so}. \tag{45}$$

Now let the slab have a finite width L_y with electric walls at $y = \pm L_y/2$. The modes even in y can be described as a superposition of two infinite slab waves whose propagation constants are such that

$$h_{sy} = \pm 2\pi n/L_y, \qquad n = 0, 1, 2, \cdots. \tag{46}$$

We have, by definition,

$$h_{sy}^2 + h_{sz}^2 = h_s^2, \tag{47}$$

h_s being given in eq. (41).

The field has all its components different from zero with the exception of E_x, which vanishes. The components E_y and H_z are obtained by adding the field of the two waves. We obtain

$$E_{sy} = 2h_{sz}h_s^{-1} \cos(h_{sy}y) \cos(\pi x/2d)E_{so}, \tag{48}$$

$$H_{sz} = -2(i\omega\mu_o)^{-1}h_{sz}h_s^{-1}(\pi/2d) \cos(h_{sy}y) \sin(\pi x/2d)E_{so}. \tag{49}$$

The energy flowing through the slab is obtained by multiplying P, given in eq. (44), by $2h_{sz}h_s^{-1}L_y$

$$P_s = 2h_{sz}(\omega\mu_o)^{-1}dL_yE_{so}^2. \tag{50}$$

The y component of the field at the boundary ($x = d$) is obtained from eq. (45) or directly from $H_{sz} = (i\omega\mu_o)^{-1}\partial E/\partial x$:

$$E_{sy}(d) = -p_s^{-1}i\omega\mu_oH_{sz}. \tag{51}$$

4.2 Modes of the rod

Let us now turn our attention to the modes of the dielectric rod. We assume that the radius a of the rod is much larger than the wavelength $(a \gg \lambda)$.

In the limit of large radii, the propagation constant of the fundamental HE_{11} mode is given (see the appendix) by the first root of $J_0(g_o a)$, namely,

$$g_o a \equiv (\omega^2 \epsilon \mu_o - h_o^2)^{\frac{1}{2}} a = 2.4 \cdots, \qquad a \to \infty. \tag{52}$$

The next higher order mode of the dielectric rod is the H_{01} mode.* In the limit of large radii, the boundary condition at $r = a$ is $E_\phi = 0$, as for a round metallic pipe. The propagation constant h_1 is therefore given by

$$J_1(g_1 a) = 0, \tag{53}$$

whose first root is

$$g_1 a \equiv (\omega^2 \epsilon \mu_o - h_1^2)^{\frac{1}{2}} a = 3.8 \cdots, \qquad a \to \infty. \tag{54}$$

Within our approximation, the field of the H_{01} mode in the rod $(r < a)$ has components

$$
\begin{aligned}
E_\phi &= J_1(g_1 r), \\
H_r &= -h_1(\omega\mu_o)^{-1} J_1(g_1 r), \\
H_z &= (i\omega\mu_o)^{-1} g_1 J_0(g_1 r),
\end{aligned}
\tag{55}
$$

and the energy flow is

$$P = -\int_0^a E_\phi H_r 2\pi r \, dr = \pi h_1(\omega\mu_o)^{-1} a^2 J_0^2(g_1 a). \tag{56}$$

To obtain the field E_ϕ at the boundary $(r = a)$, we use the fact that dE/dr is continuous and that the r dependence of E_ϕ in vacuum is approximately† $\exp(-p_1 r)$ where $p_1^2 \equiv h_1^2 - \omega^2 \epsilon_o \mu_o$. We obtain

$$E_\phi(a) = p_1^{-1} i\omega\mu_o H_z. \tag{57}$$

4.3 Synchronization conditions

For simplicity and because this is a case of practical significance, we assume that the rod and the slab have the same permittivity ϵ.

---

* The E_{01} and HE_{21} modes have almost the same propagation constant as the H_{01} mode for large rod radii. For small radiation losses, they can be considered independently of the H_{01} mode (see appendix).

† The exact dependence of E_ϕ on r is $K_0(p_1 r)$, where K_0 denotes the modified Bessel function of the second kind. For large arguments, $K_0(x) \approx (2/\pi x)^{\frac{1}{2}} \exp(-x)$ and $K_0'(x) \approx -(2/\pi x)^{\frac{1}{2}} \exp(-x) \approx -K_0(x)$.

The fundamental HE$_{11}$ mode of the rod is free of radiation loss if its propagation constant h_o given in eq. (52) is slightly larger than the propagation constant h_s in the slab given in eq. (41). For simplicity, we set $h_s = h_o$ or, equivalently, $g_s = g_o$; that is,

$$\pi/2d = 2.4/a, \tag{58a}$$

or

$$d = 0.65a. \tag{58b}$$

Thus the ratio of the slab thickness to rod diameter is 0.65. (In practice, the slab has finite dissipation losses and a finite width. Furthermore, it is difficult to control accurately the thickness of the slab. For these reasons, it might be preferable to choose the value of h_s midway between the propagation constants of the HE$_{11}$ and H$_{01}$ modes rather than equal to the propagation constant of the HE$_{11}$ mode. If the former condition were to hold, we would find that the slab thickness should be equal to half the rod diameter.) Figure 4 gives the propagation constants of the rod and the slab for $n = 1.41$ and a rod radius of 10 μm ($\lambda = 1$ μm).

Let us now consider one of the next higher order modes of the rod, the H$_{01}$ mode. This mode radiates into the substrate modes that have the same propagation constant along the z axis ($h_{sz} = h_1$). Using eq. (54), we obtain

$$\omega^2 \epsilon \mu_o - h_{sz}^2 = (3.8/a)^2. \tag{59}$$

Since

$$h_{sz}^2 + h_{sy}^2 = h_s^2, \tag{60}$$

and h_s has the value h_o given in eq. (52), we have

$$h_{sy}^2 = (3.8/a)^2 - (2.4/a)^2, \tag{61}$$

or

$$h_{sy} = 3.0/a. \tag{62}$$

In the next subsection we evaluate the coupling coefficient between the H$_{01}$ mode of the rod and the substrate mode defined by eq. (62).

4.4 Coupling coefficient

The contour of integration for the evaluation of the coupling coefficient being arbitrary, it is convenient to choose this contour as the rod boundary, $r = a$. Along that contour, the H_{01} mode field is a constant.
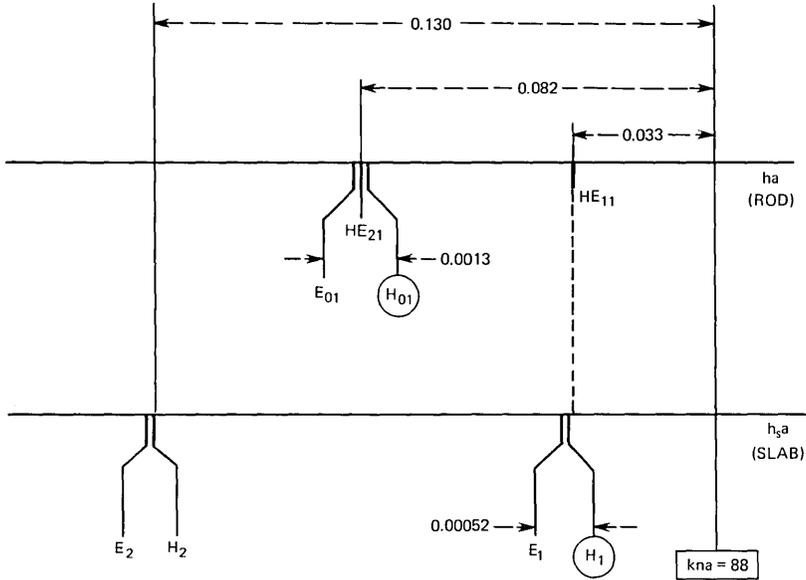
Fig. 4—Propagation constants (h) of the trapped modes of the rod and maximum value (h_s) of the propagation constants of the radiation modes in the slab. It is assumed that $n = 1.41$, $\lambda = 1\ \mu m$, and $a = 10\ \mu m$. The modes circled are those whose coupling is discussed in this paper.

Let ϕ denote the angle from the x axis shown in Fig. 3 and D the spacing between the rod and the slab. We have

$$x = -D - a(1 - \cos \phi),$$
$$y = a \sin \phi.$$
(63)

Because $a \gg \lambda$, the coupling takes place near the point of closest approach of the rod to the slab; that is, $\phi \approx 0$. We can therefore write

$$x \approx -D - a\phi^2/2,$$
$$y \approx a\phi.$$
(64)

The y dependence of the field slab is $\cos (h_{sy}y) = \cos (h_{sy}a\phi)$. However, since, according to eq. (62), h_{sy} is of the order of a^{-1}, the argument of the cosine function is small compared with unity in the range where the coupling is significant. Thus, we can neglect the dependence of the field of the slab on y. This approximation could be relaxed with little additional complication.

Using the above approximation, we obtain for the field of the H_{01} mode (rod) at $r = a$, from eqs. (55), (56), and (57),

$$H_z = (i\omega\mu_o)^{-1}g_1 J_0(g_1 a), \tag{65a}$$

$$E_\phi = p_1^{-1}i\omega\mu_o H_z, \tag{65b}$$

$$P = \pi h_1(\omega\mu_o)^{-1}a^2 J_0^2(g_1 a), \tag{65c}$$

where

$$g_1^2 \equiv \omega^2\epsilon\mu_o - h_1^2 = (3.8/a)^2,$$
$$p_1^2 \equiv h_1^2 - \omega^2\epsilon_o\mu_o \approx \omega^2(\epsilon - \epsilon_o)\mu_o \equiv u^2. \tag{66}$$

For the slab we have, at $r = a$, from eqs. (49), (50), and (51), setting the arbitrary constant E_{so} equal to unity, $h_{sz} \approx h_s$ and taking into account the exp $(p_s x)$ dependence of the field below the slab

$$H_{sz} = -2(i\omega\mu_o)^{-1}(\pi/2d) \exp\left[-p_s(D + a\phi^2/2)\right], \tag{67a}$$

$$E_{sy} = -p_s^{-1}i\omega\mu_o H_{sz}, \tag{67b}$$

$$P_s = 2h_s(\omega\mu_o)^{-1}dL_y, \tag{67c}$$

with

$$h_s \approx h_1 \approx kn,$$
$$p_s \approx p_1 \approx u = k(n^2 - 1)^{\frac{1}{2}}, \tag{68}$$
$$\pi/2d = 2.4/a.$$

The coupling coefficient C^2 is c^2/PP_s, where

$$c = a\int_{-\pi}^{+\pi} \left[E_\phi H_{sz} - E_{sy} \cos(\phi)H_z\right]d\phi. \tag{69}$$

From eqs. (65) and (67), it is apparent that the two terms in the integrand in eq. (69) are equal and add up if we make the approximation $\cos\phi \approx 1$. Thus,

$$c \approx 2aE_\phi \int_{-\infty}^{+\infty} H_{sz}d\phi. \tag{70}$$

Using eq. (67a) for H_{sz}, we have

$$\int_{-\infty}^{+\infty} H_{sz}d\phi = -2(i\omega\mu_o)^{-1}(\pi/2d) \exp(-p_s D)(2\pi/p_s a)^{\frac{1}{2}}, \tag{71}$$

if we make use of the identity

$$\int_{-\infty}^{+\infty} e^{-bx^2}dx = (\pi/b)^{\frac{1}{2}}. \tag{72}$$

Thus,

$$c = 4ap_1^{-1}g_1(i\omega\mu_o)^{-1}(\pi/2d)J_0(g_1 a) \exp(-p_s D)(2\pi/p_s a)^{\frac{1}{2}}, \tag{73}$$

and

$$C^2 = (32/\pi)g_1^2(\pi/2d)^3(u^3h^2aL_y)^{-1}\exp{(-2uD)}. \tag{74}$$

Since the mode number density is given by eq. (40), the loss

$$\mathcal{L} = \tfrac{1}{2}(h_{sz}/h_{sy})C^2L_y \tag{75}$$

is finally obtained from eqs. (74), (62), (66), and (68),

$$\mathcal{L} = 340n^{-1}(n^2 - 1)^{-\frac{3}{2}}(k^4a^5)^{-1}\exp{[-2(n^2 - 1)^{\frac{1}{2}}kD]}. \tag{76}$$

The loss in dB/km is obtained by multiplying the r.h.s. of eq. (76) by 8.7×10^9, the μm being used as the unit of length. Thus, for $n = 1.41$ and $n = 1.01$ we have, respectively,

$$\mathcal{L}_{\mathrm{d\,B/km}} = 1.35 \times 10^9\lambda_{\mu m}^{-1}(a/\lambda)^{-5}\exp{(-12.5D/\lambda)}, \qquad n = 1.41, \tag{77}$$

$$\mathcal{L}_{\mathrm{d\,B/km}} = 675 \times 10^9\lambda_{\mu m}^{-1}(a/\lambda)^{-5}\exp{(-1.76D/\lambda)}, \qquad n = 1.01. \tag{78}$$

For example, if $D = 0.15$ μm, $n = 1.41$, $\lambda = 1$ μm and $a = 40$ μm, we find that the radiation loss of the H_{01} mode is $\mathcal{L} = 2$ dB/km. If $D = 1$ μm, $n = 1.01$, $\lambda = 1$ μm, and $a = 40$ μm, the loss is as high as 1140 dB/km. The radiation loss is shown as a function of a/λ and D/λ in Figs. 5 and 6 for a wavelength of 1 μm, and for $n = 1.41$ and 1.01, respectively. The amount of loss required to prevent the power
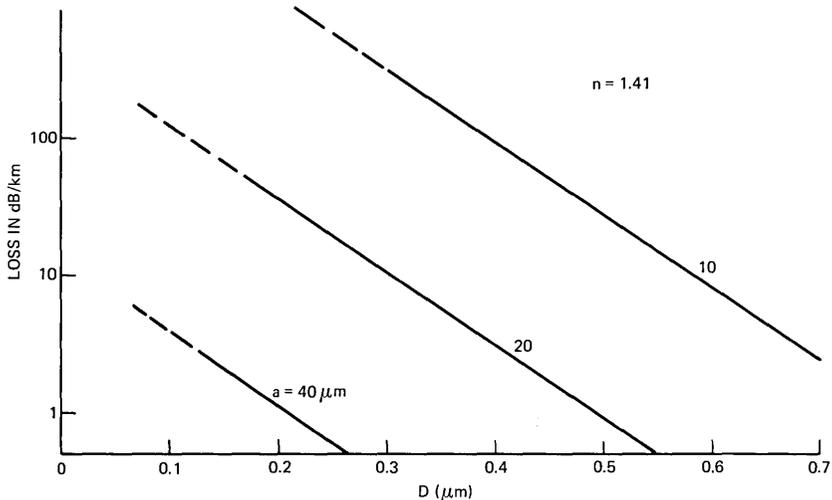
Fig. 5—Radiation loss in dB/km of the rod H_{01} mode in the slab as a function of spacing D with the rod radius a as a parameter, for $n_{\mathrm{rod}} = n_{\mathrm{slab}} = 1.41$. These curves are valid for large values of D.
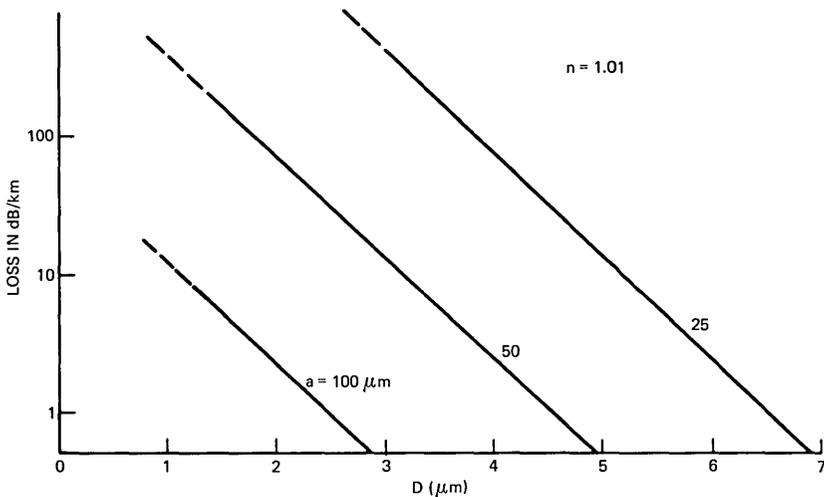
Fig. 6—Continuation of Fig. 5 for $n = 1.01$.

transferred to the H_{01} mode to be transferred back to the HE_{11} mode and to cause pulse spreading depends on the fiber irregularities and is not accurately known.

The above results are approximate and, to some extent, incomplete. In particular, the perturbation method that we used is not accurate when D is small. Also it would be useful to ascertain that the radiation losses of the other higher-order modes are at least equal to the loss calculated for the H_{01} mode. For some of these higher-order modes of the rod, it is necessary to take into account the higher-order modes of the substrate, both E and H, and this involves some complication.[9] In spite of these limitations, our result, eq. (76), should provide preliminary information concerning the mode-selection mechanism afforded by 2-dimensional mode sinks. In particular, the very fast dependence of the loss on the rod radius (a^{-5}) indicates that very large rods cannot be used if single-mode operation is to be achieved in air. However, if the gap between the rod and the slab is filled up with a material whose permittivity is only slightly smaller than the rod and slab permittivities, the rod radius a and the spacing D can be large, as Fig. 6 suggests.

V. ACKNOWLEDGEMENT

APPENDIX

Limit forms of the propagation constants in optical fibers

Two approximations can be made, applicable to low-order modes in highly multimoded fibers and to fibers with small transverse variation of permittivity. A simplified presentation is given in this appendix.

Low-order modes propagating in highly multimoded fibers correspond to waves propagating almost along the axial direction, z. The propagation constant h is therefore close to kn if n denotes the refractive index on axis. If the fiber refractive index is a constant within some contour and assumes a lower value outside that contour, the wave near a section of the contour can be assumed plane. Because it is incident at grazing angles, the electric and magnetic fields tend to zero compared with their values in the bulk. Thus, the electric and magnetic fields at the boundary of a dielectric rod vanish, compared to their values on axis, as the transverse dimensions of the rod tend to infinity for a given mode number.

For a round fiber with refractive index n and radius a, the exact equation defining h is, using the notation of the main text,[6]

$$\left[\frac{n^2 J_\nu'(u_1)}{u_1 J_\nu(u_1)} + \frac{K_\nu'(u_2)}{u_2 K_\nu(u_2)} \right] \left[\frac{J_\nu'(u_1)}{u_1 J_\nu(u_1)} + \frac{K_\nu'(u_2)}{u_2 K_\nu(u_2)} \right]$$
$$= \left[\frac{\nu h}{k} \frac{(u_1^2 + u_2^2)}{u_1^2 u_2^2} \right]^2, \quad (79)$$

the axial and azimuthal variations of the field being denoted $\exp(ihz + i\nu\phi)$ and

$$u_1 \equiv ga \equiv (k^2 n^2 - h^2)^{\frac{1}{2}} a,$$
$$u_2 \equiv pa \equiv (h^2 - k^2)^{\frac{1}{2}} a, \quad (80)$$
$$k \equiv \omega (\epsilon_o \mu_o)^{\frac{1}{2}}.$$

In the limit $a \to \infty$, u_2 tends to infinity and the second terms in the brackets on the l.h.s. of eq. (79) vanish $[K_\nu'(x)/K_\nu(x) \to -1$ if $x \to \infty]$. On the r.h.s. of eq. (79), h can be replaced by kn. Thus, it is apparent that eq. (79) becomes

$$J_\nu'(u_1)/J_\nu(u_1) = \pm \nu/u_1, \quad (81)$$

or, equivalently, using well-known formulas involving Bessel's functions and their derivatives:*

$$J_{\nu \pm 1}(u_1) = 0. \quad (82)$$

---

* We have $\nu J_\nu \mp x J_\nu' = x J_{\nu \pm 1}$ and (for later use) $\nu K_\nu \pm x K_\nu' = \mp x K_{\nu \pm 1}$.

Table I

ν	−2	−1	0	1	2	
EH:	$J_{-3} = 0$	$J_{-2} = 0$	$J_{-1} = 0$	$J_0 = 0$	$J_1 = 0$:	HE
ν < 0	$(\mu = -2)$	$(\mu = -1)$	$(\mu = 0)$	$(\mu = 1)$	$(\mu = 2)$	ν > 0
HE:	$J_{-1} = 0$	$J_0 = 0$	$J_1 = 0$	$J_2 = 0$	$J_3 = 0$:	EH

[Note: $J_{-\nu} = 0 \Longleftrightarrow J_\nu = 0$]

For symmetry reasons, modes with opposite values of ν have the same propagation constants. For $\nu = 2$, for instance, the propagation constants of the two sets of modes are given by the roots J_1 and J_3. For $\nu = -2$, they are given by the roots J_{-3} and J_{-1}. However, these are the same because $J_{-\nu} = (-)^\nu J_\nu$. Equation (82) was given by Snitzer.[7] For the HE₁₁ ($\nu = 1$) and H₀₁ ($\nu = 0$) modes, the relevant solutions of eq. (82) are the first roots of

$$J_0(u_1) = 0, \qquad u_{10} = 2.4 \cdots \tag{83a}$$

and

$$J_1(u_1) = 0, \qquad u_{10} = 3.8 \cdots. \tag{83b}$$

These are the results used in the main text.

Because the modes H₀₁, E₀₁, and HE₂₁ have almost the same propagation constants (see Table I), the validity of the calculations given in the main text can be questioned where the mode H₀₁ was considered independently of the two other modes. It is therefore important to evaluate the actual splitting between these three modes. For simplicity, we consider only the H₀₁ and E₀₁ modes. The expressions giving the exact propagation constants of the H₀₁ and E₀₁ modes are, setting $\nu = 0$ in eq. (79),

$$J_1(u_1)/u_1 J_0(u_1) = -K_1(u_2)/u_2 K_0(u_2), \qquad \text{(H}_0\text{)}, \tag{84}$$

and

$$J_1(u_1)/u_1 J_0(u_1) = -n^2 K_1(u_2)/u_2 K_0(u_2), \qquad \text{(E}_0\text{)}. \tag{85}$$

Setting

$$u_1 = u_o + \delta, \qquad \delta \ll 1, \tag{86}$$

where

$$J_1(u_o) = 0, \qquad u_o = 3.8 \cdots \tag{87}$$

on the l.h.s. of eqs. (84) and (85) and

$$u_2 = k(n^2 - 1)^{\frac{1}{2}} a \tag{88}$$

on the r.h.s., we obtain for the difference Δh in propagation constant between the H_{01} and E_{01} modes

$$\Delta h a = (3.8/kna)^2(1 - 1/n^2)^{\frac{1}{2}}. \tag{89}$$

Except for a numerical factor, this result is the same as for a slab (see Section 4.1). If $a = 10$ μm, $n = 1.41$, and $\lambda = 1$ μm, the beat wavelength $2\pi/\Delta h$ is, from eq. (89), equal to 5 cm.

The individuality of the H_{01} mode is preserved and the calculations given in the main text are valid if the loss $\mathcal{L}$ is small over that length (e.g., $\mathcal{L} \ll 1$ dB/cm for $a = 10$ μm). In fact, this restriction on $\mathcal{L}$ may be even less stringent than that calculated above because the degeneracy between the three modes may be lifted further by the presence of the slab when the coupling is increased.

The second approximation referred to at the beginning of this appendix is the scalar approximation widely used in optics. If the transverse variations of the medium permittivity are small, the x and y components of the field satisfy approximately the scalar Helmholtz equation

$$(\partial^2/\partial x^2 + \partial^2/\partial y^2)E_x + [k^2n^2(x, y) - h^2]E_x = 0. \tag{90}$$

A similar equation holds for E_y, which need not be written down.

Because all quantities are bounded in eq. (90), E_x and its first derivatives are continuous functions of x and y.

For the rod considered earlier, eq. (90) becomes, assuming an $\exp(i\mu\phi)$ dependence of E_x on ϕ,

$$
\begin{aligned}
d^2E_x/dr^2 + r^{-1}dE_x/dr + (k^2n^2 - h^2 - \mu^2/r^2)E_x = 0, \qquad r < a, \\
d^2E_x/dr^2 + r^{-1}dE_x/dr + (k^2 - h^2 - \mu^2/r^2)E_x = 0, \qquad r > a.
\end{aligned} \tag{91}
$$

These are differential equations for Bessel functions. The bounded solutions of eq. (91) are

$$
\begin{aligned}
E_x &= J_\mu(gr), & g^2 &\equiv k^2n^2 - h^2, & r < a \\
E_x &= AK_\mu(pr), & p^2 &\equiv h^2 - k^2, & r > a.
\end{aligned} \tag{92}
$$

Continuity of E_x and dE_x/dr imposes

$$J_\mu(u_1)/K_\mu(u_2) = (u_1/u_2)J_\mu'(u_1)/K_\mu'(u_2), \tag{93}$$

or, using the transformation formulas given before,

$$u_1 J_{\mu+1}(u_1)/J_\mu(u_1) = u_2 K_{\mu+1}(u_2)/K_\mu(u_2), \tag{94}$$

a result previously derived by Snyder[5] from the exact equation, eq.

(79). In the limit $a \to \infty$, eq. (94) reduces to

$$J_\mu(u_1) = 0, \tag{95}$$

in agreement with eq. (82). To each value of μ we must associate modes corresponding to the two states of polarization of the electromagnetic field. This is illustrated in Table I.

The physical significance of the scalar approximation is that if, for instance, a linearly polarized field, solution of eq. (90), is launched into a fiber, this field configuration is approximately maintained over a certain length. Eventually, however, the polarization is transformed because the two electromagnetic modes have slightly different real propagation constants as we have seen (for a report of experimental observations, see Ref. 8 in which the mode $\mu = \pm 1$ is illustrated in Figs. 3 and 4d) and/or different losses. The scalar approximation is useful to obtain approximate expressions for the propagation constants. This approximation is not applicable to the evaluation of radiation losses if these losses are polarization dependent. This is the case, for instance, if the propagation constant of the rod mode lies between the propagation constants of the slab E and H modes. Because the split between these two modes is very small, this is unlikely to happen unless the optical waveguide has been specially designed for that purpose. In that sense, the scalar approximation may be applied to problems of radiation losses.

REFERENCES

1. J. A. Arnaud, "Transverse Coupling in Fiber Optics—Part I: Coupling Between Trapped Modes," B.S.T.J., *53*, No. 2 (February 1974), pp. 217–224.
2. E. A. J. Marcatili, "Slab-Coupled Waveguides," B.S.T.J., this issue, pp. 645–674. Experimental results for the case $D = 0$ were reported by P. Kaiser, E. A. J. Marcatili, and S. E. Miller, "A New Optical Fiber," B.S.T.J., *52*, No. 2 (Feburary 1973), pp. 265–469.
3. J. A. Arnaud, "Note on the Use of Whispering Gallery Modes in Communication," September 1971, unpublished work.
4. V. V. Shevchenko, *Continuous Transitions in Open Waveguides*, Boulder, Colorado: The Golem Press, 1971, Chapter 2.
5. A. W. Snyder, "Asymptotic Expressions for Eigenfunctions and Eigenvalues of a Dielectric or Optical Waveguide," IEEE Trans. on Microwave Theory and Techniques, *MTT-17*, No. 12 (December 1969), pp. 1130–1138.
6. R. E. Collin, *Field Theory of Guided Waves*, New York: McGraw-Hill, 1960, p. 482.
7. E. Snitzer, "Cylindrical Dielectric Waveguide Modes," J. Opt. Soc. Amer., *51*, No. 5 (May 1961), pp. 491–498.
8. E. Snitzer and H. Osterberg, "Observed Dielectric Waveguide Modes in the Visible Spectrum," J. Opt. Soc. Amer., *51*, No. 5 (May 1961), pp. 499–505.
9. J. A. Arnaud, "Transverse Coupling in Fiber Optics—Part III: Bending Losses," to appear in B.S.T.J., *53*, No. 7 (September 1974).

Strip-Loaded Film Waveguide

By V. RAMASWAMY

Low-loss strip-loaded guides, consisting of 7059 glass film on fused quartz substrate with sputtered SiO_2 as the loading strip, have been investigated. The number of modes supported by the strip-loaded structures were determined experimentally and compared with the values predicted by the application of an equivalent index analysis. Agreement between theory and experiment is good in the case of the smaller number of modes which result from small loading, with the 7059 film thickness far away from cutoff.

Current interest in optical fibers and thin film devices for use in optical communication systems has prompted the development of several new guided wave structures. One of these is a single material (SM) fiber[1] representing an unclad fiber with a structural support. Basically, this is a planar slab waveguide structure with an increase in slab thickness in the central region where the guided light is concentrated. The region with the increased thickness can be considered a strip which loads a planar slab waveguide.[2]

Another type of a strip-loaded structure is shown in Fig. 1(a) where the planar waveguide has the higher index material and a strip of slightly lower index material acts as the loading. Since, within the region of the strip, most of the energy is confined in the *film*, requirements on the edge roughness of the strip are no longer as severe as in rectangular film waveguides[3] and therefore strip-loaded structures seem easier to fabricate.

Recently, Noda et al.[4] have demonstrated guiding in curved strip-loaded guides using a glass film waveguide with photoresist strip loading. They also analyzed the modal distribution of the structure by using a variational technique which, however, requires extensive computer calculations.
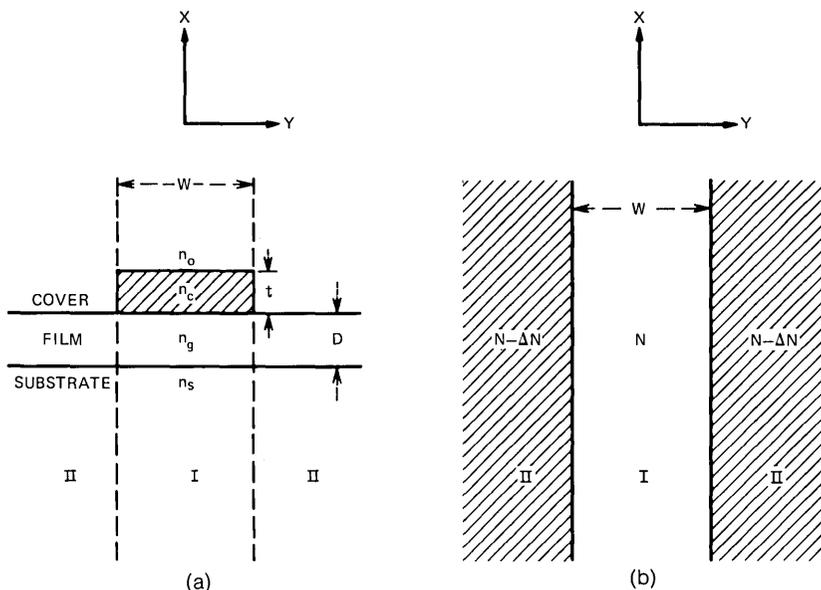
Fig. 1—Geometry of (a) strip-loaded film waveguide and (b) equivalent symmetrical waveguide.

In this paper we report studies on a strip-loaded guide consisting of 7059 glass film on quartz substrate with SiO_2 as the loading strip. We use this structure to explore the characteristics of strip-loaded guides in greater detail. We confirm the observations of Noda et al. on the guiding properties of strip-loaded structures. In addition, we use our low-loss structures to determine the number of modes supported by the strip-loaded guides and compare these results with the values predicted by the application of an equivalent index approach. This equivalent index concept is helpful in understanding the guiding characteristics of the relatively complex structure in a simple manner.

Figure 1(a) shows the geometry of a strip-loaded film waveguide. Wave propagation effects can be studied conveniently by dividing the structure into two regions. Region II represents an asymmetrical planar waveguide with n_o, n_g, and n_s as the refractive indices of the cover, film, and substrate, respectively. D is the thickness of the guiding film. Region I, on the other hand, includes a strip of superstrate of width W, thickness t, and index n_c.

The energy is confined in the x direction because of the high index film. When viewed from above, we see that the structure is symmetrical

about the x-axis and can be considered an equivalent symmetrical guide [Figure 1(b)] with the loaded section having a higher index, thereby providing the confinement of energy in the y direction. We assume, however, that this equivalent symmetrical guide is unbounded in the x direction. Each mode in regions I and II can be characterized by its own phase velocity and the corresponding effective refractive index,[5] $N_I = \beta_I/k$ and $N_{II} = \beta_{II}/k$. The difference in effective index between the two regions is responsible for the confinement of the energy within the loaded section in the plane of the film and is given by

$$\Delta N = \Delta\beta/k = N_I - N_{II}. \tag{1}$$

In order to determine the effective index N_I in Figure 1(b), we assume $W \gg t$ so that region I can be considered a planar 4-layered structure. We assume that $n_g > n_s, n_c > n_o$. Using Maxwell's equations and matching the tangential field components at the interfaces, we can obtain the transcendental equation describing the propagation characteristics of the TE modes

$$\kappa D = \phi_{co} + \phi_s + m\pi \qquad m = 0, 1, 2, 3 \cdots, \tag{2}$$

where ϕ_s and ϕ_{co} are the phase changes on total internal reflection at the film boundaries given by

$$\phi_s = \tan^{-1}\frac{\gamma_s}{\kappa}, \tag{3}$$

$$\phi_{co} = \tan^{-1}\frac{\gamma_c}{\kappa}\frac{1 - \eta e^{-2\gamma_c t}}{1 + \eta e^{-2\gamma_c t}}, \tag{4}$$

and the parameter

$$\eta = \frac{\gamma_c - \gamma_o}{\gamma_c + \gamma_o}. \tag{5}$$

The transverse propagation constant in each layer is given by

$$\gamma_o^2 = \beta^2 - (kn_o)^2, \tag{6}$$

$$\gamma_c^2 = \beta^2 - (kn_c)^2, \tag{7}$$

$$\kappa^2 = (kn_g)^2 - \beta^2, \tag{8}$$

$$\gamma_s^2 = \beta^2 - (kn_s)^2, \tag{9}$$

where $k = \omega/c = 2\pi/\lambda$ is the free-space propagation constant and β is the propagation constant in the planar structure of region I.
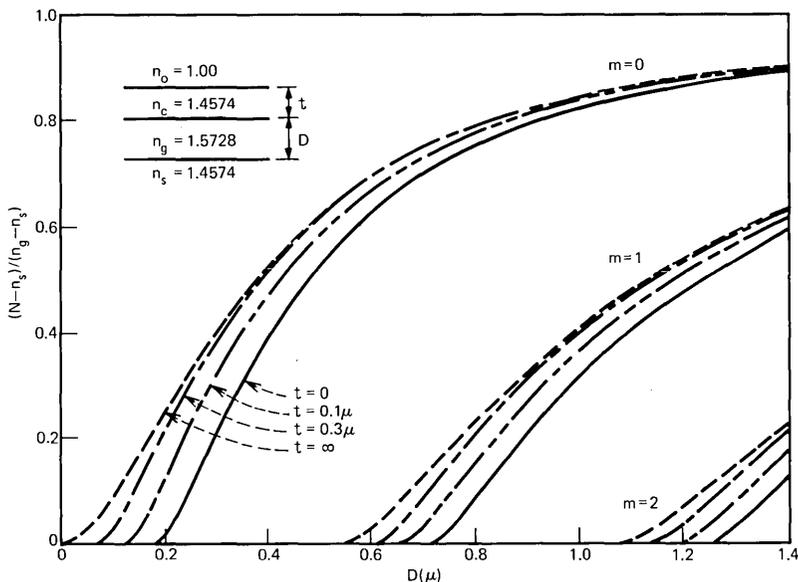
Fig. 2—Dependence of effective refractive index N on film thickness D and loading height t.

The effective index N can be obtained by solving eq. (2) by means of a computer. As an illustration, Fig. 2 shows the behavior of the effective index N as a function of film thickness D with the loading height t as the parameter for a specific set of values of the layer indices. By letting $t = 0$ in (4), we can obtain the propagation constant β of the planar guide in region II. We note the effective index $N(t)$ for a finite t is always larger than the case when $t = 0$. Therefore, the loaded region of the equivalent symmetrical guide [Fig. 1(b)] has an index N which is higher than that of the unloaded region by an amount ΔN. Figure 3 shows the plot of ΔN for the same parameters of the waveguide structure illustrated in Fig. 2. It is clear from Fig. 3 that, while ΔN increases with the height t of the loading, it does not have to be very large to achieve a reasonable value (of the order of 10^{-2}) of ΔN. In fact, the difference in ΔN between the cases $t = 0.3 \, \mu$ and $t = \infty$ is indeed very small.

The total number of modes p in the equivalent symmetrical guide is obtained from the cutoff condition[6] as

$$p = 1 + \frac{2W}{\lambda} \{N^2(t) - N^2(0)\}^{\frac{1}{2}} \approx 1 + \frac{2W}{\lambda} (2N\Delta N)^{\frac{1}{2}}, \quad \Delta N \ll 1. \quad (10)$$

By using the computed values for $N(t)$ and $N(0)$ in (10), the number of modes in the strip-loaded structure can be determined.

We made several strip-loaded guides using RF-sputtered 7059 film waveguides on glass on fused quartz substrates. The film thickness was above cutoff, allowing at least one propagating mode in the film without the loading. To avoid using photoresist in the sputtering system, a thin layer of SiO_2 was sputtered on the 7059 film first and then, using photolithographic techniques, the SiO_2 film was etched everywhere except over a strip region using buffered HF as the etchant. Etching was carried out in small steps to control the depth carefully. The loading strip width W was varied from 5 to 12.5 μ. The height of the loading strip was 0.1 to 0.4 μ, and the values for the thickness of the film were chosen to be between 0.2 to 0.5 μ.

A Gaussian He–Ne laser beam was apertured and used to excite the strip-loaded guide by means of a prism coupler placed directly on the strip (Fig. 4). Coupling to the strip-loaded guide was easier when the height t was small. In addition to the use of a rotating table in the
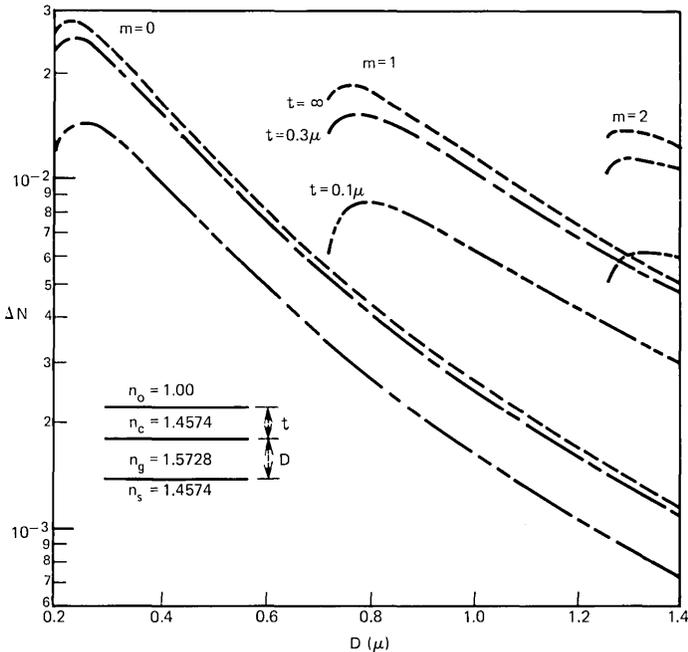
Fig. 3—Difference in effective refractive index between regions I and II vs film thickness D.
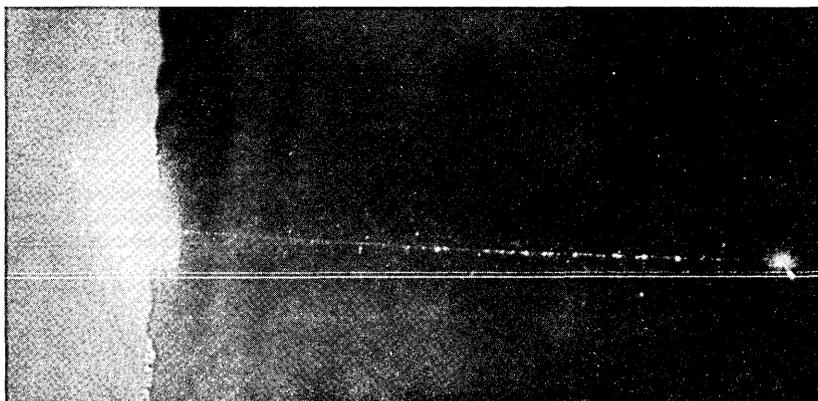
Fig. 4—Propagation of light beam guided in a strip-loaded guide. At the point where the strip is scratched, the light is radiated.

prism coupler arrangement, provision was made to tilt the entire assembly in the plane perpendicular to the plane of the table. Each mode was excited by varying the position of the beam as well as by changing the angle of excitation. Where there are only few modes and very little mode conversion, the modes can be identified by viewing the far field pattern, and the number of modes can be counted by varying the exciting conditions.

In order to compare the measured number of modes with the predictions of eq. (10), we determine N and ΔN from Figs. 2 and 3. For this purpose, the refractive index n_g and thickness D of the 7059 film was obtained by measuring the synchronous angles using a prism coupler.[7] The thickness t was measured using a Tally-Surf thickness measuring machine and the width W by viewing the structure in a microscope. The index of SiO_2 film was assumed to be the same as that of the quartz disc used in sputtering. By computing $N(t)$ and $N(0)$, the number of modes p was calculated and is shown in Fig. 5.

Each measured point ($\blacktriangle$) in Fig. 5 represents one guide structure. We find the agreement between theory and experiment is good in the case of smaller number of modes which result from small loading with the film thickness D far away from cutoff. Since it was rather difficult to excite the structure with a large strip thickness using the present techniques, investigation of structures with large strip thickness was not possible. Moreover, as the number of modes increased, the higher-order modes could not be resolved because of mode conversions resulting from imperfections. In the case of higher-order modes, the measured
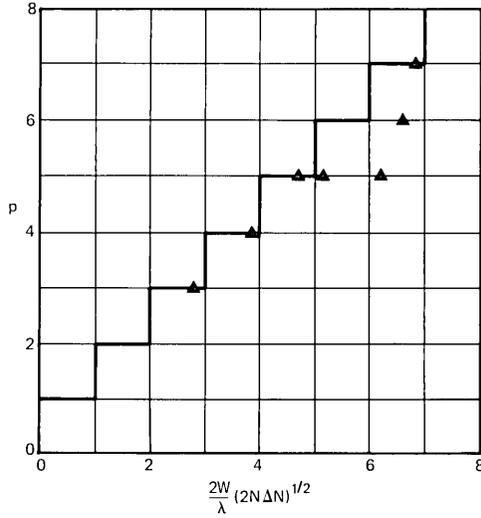
Fig. 5—Number of modes in the strip-loaded guide. The ▲ points represent the experimental results. Note that N and ΔN are functions of λ to be determined from Figs. 2 and 3.

losses in the structure were as high as 1.5 to 2 dB/cm when the strip edge roughness was ≈ 6000 Å and the lowest measured loss was 0.5 dB/cm for the fundamental mode. It is also interesting to note that the edge roughness problem once again becomes important with the increased loading resulting from the increase in the energy content in the strip.

A more generalized approach to these new guided wave structures has been developed by Marcatili[8] and our results are in agreement with his analysis.

REFERENCES

1. P. Kaiser, E. A. J. Marcatili, and S. E. Miller, "A New Optical Fiber," B.S.T.J., *82*, No. 2 (February 1973), p. 265–269.
2. Strip-loaded single material asymmetrical planar guides have also been referred to as rib waveguides. See J. E. Goell, "Rib Waveguides for Integrated Optical Circuits," to be published in Applied Optics.
3. J. E. Goell, "Electron-Resist Fabrication of Bends and Couplers for Integrated Optical Circuits," Appl. Opt., *12*, 729 (1973).

4. H. Noda, H. Faruka, and A. Ihaya, "A Novel Optical Waveguide for Integrated Optics," Conference on Laser Engineering and Applications, Washington,D. C., 1973.
5. R. Ulrich and R. J. Martin, "Geometrical Optics in Thin Film Light Guides," Appl. Opt., *10*, 2077 (1971).
6. D. Marcuse, "Light Transmission Optics," New York: Van Nostrand Reinhold, Co., p. 327, Equations (8.6-6).
7. P. K. Tien, R. Ulrich, and R. J. Martin, "Modes of Propagating Light Waves in Thin Deposited Semiconductor Films," Appl. Phys. Letters, *14*, 291 (1969).
8. E. A. J. Marcatili, "Slab-Coupled Waveguides," B.S.T.J., this issue, pp. 645–674.

Rayleigh Scattering and the Impulse Response of Optical Fibers

By D. MARCUSE

The impulse response of multimode optical fibers is distorted because each mode carries the signal at a different group velocity. Mode coupling tends to reduce the width of the impulse response. Rayleigh scattering, being the most fundamental scattering process in optical fibers, serves as a mode-coupling mechanism. However, it also causes radiation loss. The penalty of a seemingly apparent improvement of the impulse response through Rayleigh scattering is calculated in this paper. We conclude that, because of the high loss penalty, Rayleigh scattering is not a suitable technique for pulse-width improvement.

I. INTRODUCTION

The term "Rayleigh scattering" describes light scattering from re-fractive index inhomogeneities whose linear dimensions are much shorter than the wavelength of light. Most of the scattered light escapes from the core region of the fiber and enters the cladding or the space outside of the fiber. Some of the scattered power goes into other guided modes. Rayleigh scattering thus contributes to the losses in the fiber and also influences the impulse response through mode coupling.

Since mode coupling tends to improve the impulse response of optical fibers,[1,2] the question may be asked: How beneficial is Rayleigh scattering for light transmission in multimode fibers because of its mode-coupling ability? To answer this question we investigate the loss penalty that is incurred if Rayleigh scattering is assumed as the only mode-coupling mechanism.

For simplicity, our study is limited to a slab waveguide model (see Fig. 1) assuming that there is no variation of the refractive index or the light field in the y direction. Ignoring coupling between guided
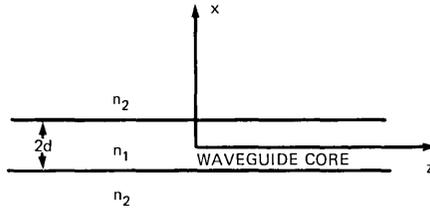
Fig. 1—Schematic of slab waveguide. The scattering centers are distributed randomly throughout the core and the outside medium. They are infinitely thin threads of slightly different refractive index extending in the y direction.

modes traveling in opposite directions, we calculate the width of the impulse response and the amount of scattering losses. These calculations allow us to establish the loss penalty. We find that the loss penalty for any significant pulse-width reduction caused by Rayleigh scattering is intolerably high. Thus, it is not feasible to improve the pulse dispersion of multimode fibers by intentionally implanting Rayleigh scatterers into the dielectric material of the fiber. However, improved pulse transmission is obtainable by using other carefully engineered mode-coupling mechanisms.[2]

II. THE COUPLING COEFFICIENT

The even guided TE modes of a slab waveguide consisting of a perfect dielectric are determined by the y component of its electric field.[3]

$$E_y = A \cos \kappa x \qquad |x| < d. \tag{1}$$

$$E_y = A \cos \kappa d \, e^{-\gamma(|x|-d)} \qquad |x| > d. \tag{2}$$

The odd guided modes are given by

$$E_y = A \sin \kappa d \qquad |x| < d. \tag{3}$$

$$E_y = \frac{x}{|x|} A \sin \kappa d \, e^{-\gamma(|x|-d)} \qquad |x| > d. \tag{4}$$

The magnetic field components are obtained by differentiation:

$$H_x = \frac{-i}{\omega \mu_0} \frac{\partial E_y}{\partial z} \tag{5}$$

and

$$H_y = \frac{i}{\omega \mu_0} \frac{\partial E_y}{\partial x}. \tag{6}$$

The factor $\exp[i(\omega t - \beta z)]$ is omitted from these and all subsequent

field equations. The width of the core of the slab is $2d$. The parameters κ and γ are defined as follows:

$$\kappa = (n_1^2 k^2 - \beta^2)^{\frac{1}{2}} \tag{7}$$

and

$$\gamma = (\beta^2 - n_2^2 k^2)^{\frac{1}{2}}, \tag{8}$$

with $k = \omega\sqrt{\epsilon_0\mu_0}$, $n_1 = $ core index, and $n_2 = $ cladding index. The propagation constant β is obtained as a solution of the eigenvalue equations:

$$\tan \kappa d = \frac{\gamma}{\kappa} \qquad \text{for even modes} \tag{9}$$

and

$$\tan \kappa d = -\frac{\kappa}{\gamma} \qquad \text{for odd modes.} \tag{10}$$

The amplitude coefficient is related to the power P carried by the mode

$$A = \left(\frac{2\gamma\omega\mu_0 P}{(1 + \gamma d)\beta} \right)^{\frac{1}{2}}. \tag{11}$$

In addition to guided modes, the slab with infinite cladding has radiation modes. The magnetic fields of the radiation modes follow from E_y by means of (5) and (6). The E_y component of the even radiation modes is[4]

$$E_y = B \cos \sigma x \qquad |x| < d \tag{12}$$

and

$$E_y = \left(\frac{2\omega\mu_0 P}{\pi\beta} \right)^{\frac{1}{2}} \cos [\rho(|x| - d) + \psi] \qquad |x| > d, \tag{13}$$

with ψ defined by

$$\tan \psi = \frac{\sigma \sin \sigma d}{\rho \cos \sigma d}. \tag{14}$$

The amplitude coefficient B is given by

$$B = \left(\frac{2\rho^2\omega\mu_0 P}{\pi\beta(\rho^2 \cos^2 \sigma d + \sigma^2 \sin^2 \sigma d)} \right)^{\frac{1}{2}}. \tag{15}$$

The parameters σ and ρ are defined by

$$\sigma = (n_1^2 k^2 - \beta^2)^{\frac{1}{2}} \tag{16}$$

and

$$\rho = (n_2^2 k^2 - \beta^2)^{\frac{1}{2}}. \tag{17}$$

Similarly, for the odd radiation modes we have

$$E_y = C \sin \sigma x \qquad |x| < d \tag{18}$$

and

$$E_y = \frac{x}{|x|} \left(\frac{2\omega\mu_0 P}{\pi\beta} \right)^{\frac{1}{2}} \sin \left[\rho(|x| - d) + \phi \right] \qquad |x| > d. \quad (19)$$

Phase ϕ is defined by

$$\tan \phi = \frac{\rho \sin \sigma d}{\sigma \cos \sigma d} \qquad (20)$$

and the amplitude factor is given as

$$C = \left(\frac{2\rho^2\omega\mu_0 P}{\pi\beta(\rho^2 \sin^2 \sigma d + \sigma^2 \cos^2 \sigma d)} \right)^{\frac{1}{2}}. \quad (21)$$

The coupling coefficient between two modes has the form[5-7]

$$K_{\nu\mu} = \frac{\omega \epsilon_0}{4iP} \int_{-\infty}^{\infty} (n^2 - n_0^2) E_\nu E_\mu^* dx. \quad (22)$$

E_ν and E_μ are the y components of the electric fields of two modes labeled ν and μ. The index distribution $n = n(x, z)$ describes the wave-guide with slight random fluctuations around the average value, and $n_0 = n_0(x)$ is the index distribution that defines the ideal slab wave-guide. It is $n_0 = n_1$ in the core and $n_0 = n_2$ outside. The ensemble average of $n^2 - n_0^2$ vanishes,

$$\langle n^2 - n_0^2 \rangle = 0. \quad (23)$$

The power-coupling coefficients are obtained from the expression[7,8]

$$h_{\nu\mu} = \frac{1}{L} \int_0^L dz \int_0^L dz' \langle K_{\nu\mu}(z) K_{\nu\mu}^*(z') \rangle e^{i(\beta_\mu - \beta_\nu)(z-z')}$$

$$= \frac{\omega^2 \epsilon_0^2}{16LP^2} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx' \int_0^L dz \int_0^L dz' \langle (n^2 - n_0^2)(n'^2 - n_0'^2) \rangle$$

$$\times E_\nu E_\nu'^* E_\mu' E_\mu^* e^{i(\beta_\mu - \beta_\nu)(z-z')}. \quad (24)$$

The prime indicates quantities depending on x' and z'.

   The purpose of this calculation is to study Rayleigh scattering. For this reason we may assume that the correlation of the index fluctuations reaches only over distances that are much smaller than the wavelength $2\pi/\beta_\nu$. The following correlation function is used:

$$\langle (n^2 - n_0^2)(n'^2 - n_0'^2) \rangle = D^2 \langle (n^2 - n_0^2)^2 \rangle \delta(x - x') \delta(z - z'), \quad (25)$$

where D is the correlation length of the index fluctuations. Substitution of (25) into (24) leads to

$$h_{\nu\mu} = \frac{\omega^2 \epsilon_0^2}{16P^2} D^2 \langle (n^2 - n_0^2)^2 \rangle \int_{-\infty}^{\infty} |E_\nu|^2 |E_\mu|^2 dx. \quad (26)$$

The remaining z integration (after integration over the delta function) over the distance L resulted in a factor L that canceled from the equation.

To evaluate the remaining integral in (26) we make the following assumption. All modes are considered sufficiently far from cutoff so that the guided mode fields are very weak at the core boundary, $x = d$, and negligible outside of the core. For a guide supporting very many modes, this assumption is justified for most of them. Thus, the integral in (26) effectively extends only over the region of the core. The integrals are of three different types:

$$I_1 = \int_{-d}^{d} \cos^2 \kappa_\nu x \cos^2 \kappa_\mu x \, dx, \tag{27}$$

$$I_2 = \int_{-d}^{d} \sin^2 \kappa_\nu x \sin^2 \kappa_\mu x \, dx, \tag{28}$$

and

$$I_3 = \int_{-d}^{d} \cos^2 \kappa_\nu x \sin^2 \kappa_\mu x \, dx. \tag{29}$$

Since almost all modes have rapidly oscillating fields in x direction inside of the core, we approximate these integrals by

$$I_1 = I_2 = I_3 \approx \frac{d}{2}. \tag{30}$$

With the help of (11) and (30) we obtain from (26)

$$h_{\nu\mu} = \frac{k^4 \gamma_\nu \gamma_\mu d}{8(1 + \gamma_\nu d)(1 + \gamma_\mu d)\beta_\nu \beta_\mu} \, D^2 \langle (n^2 - n_0^2)^2 \rangle. \tag{31}$$

In the spirit of our approximation, we may assume $\gamma_\nu d \gg 1$ and $\beta_\nu \approx n_1 k$, where n_1 indicates the core index. Thus, the power-coupling coefficient can be approximated as follows:

$$h = h_{\nu\mu} = \frac{k^2}{8 n_1^2 d} \, D^2 \langle (n^2 - n_0^2)^2 \rangle. \tag{32}$$

In this far-from-cutoff approximation, the power-coupling coefficient is independent of mode number. Rayleigh scattering couples with equal strength all of the modes.

With the same type of approximations, we obtain from (11), (15), (26), and (30) the coupling coefficient between a guided mode labeled ν and an even radiation mode with propagation constant β:

$$h_\nu^{(e)}(\beta) = \frac{\rho^2 k^3 D^2 \langle (n^2 - n_0^2)^2 \rangle}{8 n_1 \pi |\beta| (\rho^2 \cos^2 \sigma d + \sigma^2 \sin^2 \sigma d)}. \tag{33}$$

Coupling to odd radiation modes leads to the same type of coupling coefficient, $h_\nu^{(o)}(\beta)$, except that cos σd and sin σd are now interchanged.

The power (scattering) loss coefficient for mode ν is

$$\alpha_\nu = 2 \int_0^{n_2 k} [h_\nu^{(e)}(\beta) + h_\nu^{(o)}(\beta)] d\rho. \tag{34}$$

This expression can be justified as follows. The power-coupling coefficient indicates the amount of power flowing per unit length from the guided mode to each individual radiation mode. The sum of the contributions to all radiation modes gives the total loss. Since radiation modes form a continuum, the sum becomes an integral. The factor 2 in front of the integral indicates the doubling of the loss caused by power flowing not only into forward but also into backward traveling radiation modes. The integral over ρ can be converted to integration over β as follows:

$$\alpha_\nu = 2 \int_0^{n_2 k} [h_\nu^{(e)}(\beta) + h_\nu^{(o)}(\beta)] \frac{\beta}{\rho} d\beta. \tag{35}$$

The integration includes only propagating radiation modes. The contribution of even and odd modes is very nearly the same, so that we use only the coupling coefficient (33) and double the factor in front of the integral:

$$\alpha_\nu = \frac{k^3 D^2 \langle (n^2 - n_0^2)^2 \rangle}{2\pi n_1} \int_0^{n_2 k} \frac{\rho d\beta}{\rho^2 \cos^2 \sigma d + \sigma^2 \sin^2 \sigma d}. \tag{36}$$

To the approximation used in this analysis, the power-radiation-loss coefficient of the guided modes is independent of mode number.

An exact solution of the integral in (36) is hard to obtain. If we consider the fact that for large values of d the sine and cosine functions pass through many periods throughout the range of integration, we can replace the integrand by its average value over a few periods of the periodic functions. This average is

$$\left(\frac{\rho}{\rho^2 \cos^2 \sigma d + \sigma^2 \sin^2 \sigma d} \right)_{\text{average}} = \frac{1}{\sigma}. \tag{37}$$

It now remains to solve the integral:

$$\int_0^{n_2 k} \frac{d\beta}{\sigma} = \int_0^{n_2 k} \frac{d\beta}{\sqrt{n_1^2 k^2 - \beta^2}} = \arcsin \frac{n_2}{n_1}, \tag{38}$$

In most cases of practical interest the ratio n_2/n_1 is very close to unity so that we can approximate the integral by $\pi/2$. We thus obtain the

following equation for the radiation power loss coefficient

$$\alpha = \alpha_\nu = \frac{k^3 D^2 \langle (n^2 - n_0^2)^2 \rangle}{4n_1} = 2n_1 khd. \tag{39}$$

The last part of the equation follows from (32).

III. CALCULATION OF IMPULSE RESPONSE

Pulse propagation in optical fibers can be described by the following equation for the average power[2,7]

$$\frac{\partial P_\nu}{\partial z} + \frac{1}{v_\nu} \frac{\partial P_\nu}{\partial t} = -\alpha_\nu P + \sum_{\mu=1}^{N} h_{\nu\mu}(P_\mu - P_\nu). \tag{40}$$

This system of coupled power equations holds only for modes traveling in the same direction. Rayleigh scattering scatters power in forward as well as backward directions; however, we must ignore the backward scattered power flowing into guided modes. Physically, it appears that this approximation should pose no difficulty, since only those modes that travel in near synchronism have a chance to interact thoroughly. Backward scattered power travels away from the pulse that created it; thus, it cannot alter the shape of the impulse response except, perhaps, by repeated reflections. Backward scattered power contributes mainly to the scattering losses. We have taken backward scattering into radiation modes into account, but the additional loss caused by backward scattering into guided modes contributes far less loss and is ignored in our treatment. Thus, we recognize that the approximation may lead to a slight underestimation of the total scattering loss.

To solve (40) we use the trial solution

$$P_\nu = A_\nu e^{-\sigma z} e^{i\omega [t - (n_1 z/c)]}. \tag{41}$$

Substitution into (40) leads to

$$A_\nu = h \left(\sum_{\mu=1}^{N} A_\mu \right) \Big/ \left[\alpha - \sigma + Nh + i\omega \left(\frac{1}{v_\nu} - \frac{n_1}{c} \right) \right]. \tag{42}$$

We used the fact that the loss coefficients and the coupling coefficients are independent of the mode number. The quantity N is the total number of modes.

We obtain the group velocity of the modes from an approximation of the propagation constant. Using[9]

$$\kappa_\nu d \approx \nu \frac{\pi}{2}, \tag{43}$$

we obtain from (7)

$$\beta \approx \left[n_1^2 k^2 - \left(\nu \frac{\pi}{2d} \right)^2 \right]^{\frac{1}{2}}. \tag{44}$$

The inverse group velocity of mode ν is

$$\frac{1}{v_\nu} = \frac{d\beta}{d\omega} = \frac{1}{c} \frac{d\beta}{dk} = \frac{n_1^2 k}{c\beta}. \tag{45}$$

Using $n_1 k \gg \nu\pi/2d$ we obtain approximately

$$\frac{1}{v_\nu} \approx \frac{n_1}{c} (1 + G\nu^2), \tag{46}$$

with

$$G = \frac{\pi^2}{8(n_1 kd)^2}. \tag{47}$$

The solution of the equation system (42) is accomplished with ease, since we realize that the sum term in the numerator is independent of the mode label. Thus, the coefficients A_ν must be of the form

$$A_\nu = \frac{C}{\alpha - \sigma + Nh + i\omega \dfrac{n_1}{c} G\nu^2}. \tag{48}$$

Substitution of (48) into (42) leads to an eigenvalue equation for the determination of σ:

$$h \sum_{\mu=1}^{N} \frac{1}{\alpha - \sigma + Nh + i\omega \dfrac{n_1}{c} G\mu^2} = 1. \tag{49}$$

The sum can be approximated by the integral

$$\int_0^N \frac{dx}{\alpha - \sigma + Nh + i\omega \dfrac{n_1}{c} Gx^2} = \frac{1}{\left[i\omega \dfrac{n_1}{c} G(\alpha - \sigma + Nh) \right]^{\frac{1}{2}}}$$
$$\times \arctan \left(\frac{i\omega \dfrac{n_1}{c} GN^2}{\alpha - \sigma + Nh} \right)^{\frac{1}{2}}. \tag{50}$$

Thus, we obtain from (49) and (50) the eigenvalue equation

$$\left(\frac{i\omega \dfrac{n_1}{c} GN^2}{\alpha - \sigma + Nh} \right)^{\frac{1}{2}} = \tan \left\{ \frac{1}{h} \left(i\omega \frac{n_1}{c} G(\alpha - \sigma + Nh) \right)^{\frac{1}{2}} \right\}. \tag{51}$$

Fortunately, we need only the lowest-order eigenvalue since it has the

significance of the steady-state loss of the system of coupled modes and also determines the shape of the impulse-response function.[2,7] The solution of (51) is accomplished by using the fact that the lowest-order eigenvalue must be close to the loss coefficient α. Thus, we set

$$\sigma = \alpha + \eta. \tag{52}$$

Next, we expand the tangent function in series and solve for η. In this way we obtain the approximate solution

$$\sigma = \alpha + \frac{4}{45} \left(\frac{n_1}{c} G \right)^2 \frac{N^3}{h} \omega^2 + i\omega \frac{n}{3c} GN^2. \tag{53}$$

For our purposes the coefficient ρ of ω^2 is of most importance. We obtain the general pulse shape by substituting (53) into (41) and integrating over ω from $-\infty$ to ∞. Neglecting the ω dependence of A_ν, we find a Gaussian-shaped pulse whose width is[2,7]

$$\Delta t = 4\sqrt{\rho L} = \frac{8}{\sqrt{45}} \frac{n_1}{c} G \frac{N^{\frac{3}{2}}}{\sqrt{h}} \sqrt{L}. \tag{54}$$

The width of the signal in the absence of mode coupling is

$$\Delta T = \left(\frac{1}{v_N} - \frac{1}{v_1} \right) L = \frac{n_1}{c} GN^2L. \tag{55}$$

The relative improvement of the width of the steady-state pulse in the presence of mode coupling is expressed by the factor[2,7]

$$R = \frac{\Delta t}{\Delta T} = \frac{8}{\sqrt{45NhL}}. \tag{56}$$

Using (39) and (56), we define the loss penalty by the expression[2,7]

$$R^2\alpha L = 2.8 \frac{n_1 kd}{N}. \tag{57}$$

The number of modes is obtained from (44) with the help of the cutoff condition $\beta = n_2 k$ for $\nu = N$; thus,

$$N = \frac{2kd}{\pi} \sqrt{n_1^2 - n_2^2}. \tag{58}$$

The expression for the loss penalty thus assumes the form

$$R^2\alpha L = \frac{4.4}{\sqrt{1 - \frac{n_2^2}{n_1^2}}} \approx \frac{3.1}{\sqrt{1 - \frac{n_2}{n_1}}}. \tag{59}$$

IV. DISCUSSION

We can now answer the question that was asked in the introduction: Is Rayleigh scattering significantly beneficial because of its ability to shorten the width of the impulse response? Let us assume that we have a slab waveguide with a core-to-cladding index ratio of $n_1/n_2 = 1.01$. From (59) we obtain in this case

$$R^2 \alpha L = 31.2 = 135 \text{ dB}. \tag{60}$$

We may now ask how much loss is associated with a relative decrease of the width of the impulse response by a factor 2, or $R = 0.5$. We see from (60) that the amount of scattering loss associated with this "improvement" is

$$\alpha L = 540 \text{ dB}. \tag{61}$$

This shows that if we are hoping for a reduction in the width of the impulse response with the help of Rayleigh scattering, we have to pay an intolerably high price in added loss. Since Rayleigh scattering losses are known to be quite small, (59) indicates that this mechanism does not help to reduce the width of the impulse response under ordinary conditions.

We are thus forced to consider Rayleigh scattering as detrimental to light transmission in optical fibers. Fortunately, it is a small effect that does not provide prohibitively high losses at visible or infrared wavelength.

It is easy to understand why Rayleigh scattering is not more effective in reducing the width of the impulse response. It has been shown that a very carefully shaped power spectrum of the function describing fiber irregularities is required to reduce the loss penalty for pulse width reduction.[2] Rayleigh scattering is particularly poorly suited for this purpose since its power spectrum is flat. Only a very small fraction of the total amount of scattering is used for mode mixing, most of it is used for light scattering into radiation modes leading to scattering losses.

Our calculation was based on a slab waveguide model. However, the result is expected to be representative of round optical fibers. Experience has shown that estimates of the performance of round fibers can be obtained from scattering data calculated on the basis of a slab waveguide model.

V. ACKNOWLEDGEMENT

REFERENCES

1. S. D. Personick, "Time Dispersion in Dielectric Waveguides," B.S.T.J., *50*, No. 3 (March 1971), pp. 843–859.
2. D. Marcuse, "Pulse Propagation in Multimode Dielectric Waveguides," B.S.T.J., *51*, No. 6 (July-August 1972), pp. 1199–1232.
3. D. Marcuse, *Light Transmission Optics*, New York: Van Nostrand-Reinhold Co., 1972, pp. 307–309.
4. Ref. 3, pp. 314–318. The notation for the field in the cladding is written in more compact form in this paper.
5. A. W. Snyder, "Coupled Mode Theory for Optical Fibers," J. Opt. Soc. Am., *62*, No. 11 (November 1972), pp. 1267–1277.
6. D. Marcuse, "Coupled Mode Theory of Round Optical Fibers," B.S.T.J., *52*, No. 6 (July-August 1973), pp. 817–842.
7. D. Marcuse, *Theory of Dielectric Optical Waveguides*, New York: Academic Press, 1974.
8. D. Marcuse, "Derivation of Coupled Power Equations," B.S.T.J., *51*, No. 1 (January 1972), pp. 229–237.
9. Ref. 3, p. 328.

Coupling of Nearly Degenerate Modes in Parallel Asymmetric Dielectric Waveguides

By L. O. WILSON and F. K. REINHART

The coupling of modes in two parallel dielectric waveguides is studied. The individual waveguides are assumed to be asymmetric and unlike each other. If the individual waveguides support modes with nearly equal propagation constants β_2 and $\beta_4 = \beta_2 + 2\Delta$, then the double waveguide system will support two new modes with propagation constants $\beta_- = \beta_2 - \bar{\delta}$ and $\beta_+ = \beta_4 + \bar{\delta}$. The shift $\bar{\delta}$ is related to Δ and to the shift δ which would occur if the original modes were degenerate; δ is expressed in terms of the parameters describing the asymmetric double waveguide system. The field distributions of the new modes are approximately even and odd combinations of those of the original modes in the isolated waveguides; the relative amplitudes with which they are combined depend upon the amount of mismatching Δ. As the modes travel down the waveguide system, they partially cancel and add, thus transferring power. A power transfer ratio F is defined and is shown to decrease rapidly as Δ/δ increases. The beat length L depends upon both δ and Δ/δ; it also decreases as Δ/δ increases. A numerical example is given to illustrate the effects of mismatching and to demonstrate the feasibility of constructing a mode-coupling device. Possibilities of tuning the device to reduce mismatching are discussed.

I. INTRODUCTION

Coupling of degenerate modes of parallel optical waveguides has been discussed by Kapany[1] and, to a greater extent, by Marcuse.[2] Such coupling is of particular interest in the field of fiber optics, since it may cause undesirable crosstalk between adjacent optical fibers used for light transmission. Marcuse[2] has applied the theory of degenerate mode coupling to the problem of crosstalk between cladded optical fibers embedded in a lossy medium and between cladded dielectric slab waveguides. The fabrication of devices which would actually take advantage of mode coupling, such as for light switching, modula-

tion, or power transferral,[3] is fraught with practical difficulties, since the specification of physical parameters must necessarily be stringent. These difficulties require us to view the theory of optical waveguide coupling from a new vantage point.

Let us first sketch briefly what is known. If two optical waveguides each have a mode with the same propagation constant β, then when the two waveguides are placed parallel to each other, the double waveguide system supports two new modes whose propagation constants are $\beta_+ = \beta + \delta$ and $\beta_- = \beta - \delta$. These two modes are approximately symmetric and anti-symmetric combinations of the original modes in the isolated waveguides. The shift in propagation constant, δ, is related to the coupling coefficients involved in a description of the modes by means of general coupled line equations. It can also be expressed via a perturbation treatment of Maxwell's equations. Since the superimposed modes travel down the double waveguide system at different phase velocities, they alternately add and cancel. If the waveguides are lossless, power is transferred back and forth over a beat length $L = \pi/(2\delta)$. On the other hand, Marcuse shows that, if the waveguides are lossy, they tend to equalize the power they carry, provided the modes travel far enough. A lossy external medium also causes mode loss. Marcuse further states that only degenerate modes exchange a significant amount of power if their coupling mechanism is independent of length.

With this abbreviated version of the present theory in mind, we see several criteria which a mode coupling device should satisfy: (*i*) the core and cladding of each waveguide should be lossless, (*ii*) the medium external to the waveguides should be lossless, and (*iii*) the two waveguides should have a degenerate mode. (There are also other criteria, such as that the waveguide walls be free from imperfections, but they are not discussed.)

The first criterion is an important one and certainly merits further study. In this paper, however, we avoid the issue by assuming that the device we fabricate has lossless waveguides. A subtler way to put this is to say that the device is short enough that losses can be ignored.

The second criterion is satisfied by assuming that the claddings of the two waveguides are contiguous and that there is no medium external to them. Instead of thinking in terms of two optical fibers, we consider two dielectric slab waveguides placed next to each other. Fabrication would be similar to that currently used in the production of double heterostructure lasers and modulators.[4,5] Each waveguide will consist of a slab of high refractive index surrounded by two slabs

of lower index. Since the double waveguide device will have a central slab common to both waveguides, the device can be modelled by a 5-dielectric-slab model.

The third criterion, that the two waveguides have a degenerate mode, motivates our present study. In practice, it is very difficult to fabricate a device with degenerate modes. It is therefore quite important to know how well the device will operate if the propagation constants for the modes are slightly mismatched. We study the effect of mismatching on the beat length and on the capability of the device to transfer power. We also discuss methods of tuning the device after it is fabricated. The tuning could be used to match the propagation constants more closely. It might also be used dynamically, thus offering the possibility of utilizing the double waveguide system as a light switch or a modulator.

II. FORMULATION

We adopt the standard slab model of an absorptionless dielectric medium. The optical dielectric $K(x)$, i.e., the square of the refractive index, is assumed to vary only with x and to take the piecewise constant form shown in Fig. 1. If the waves are assumed to travel in the
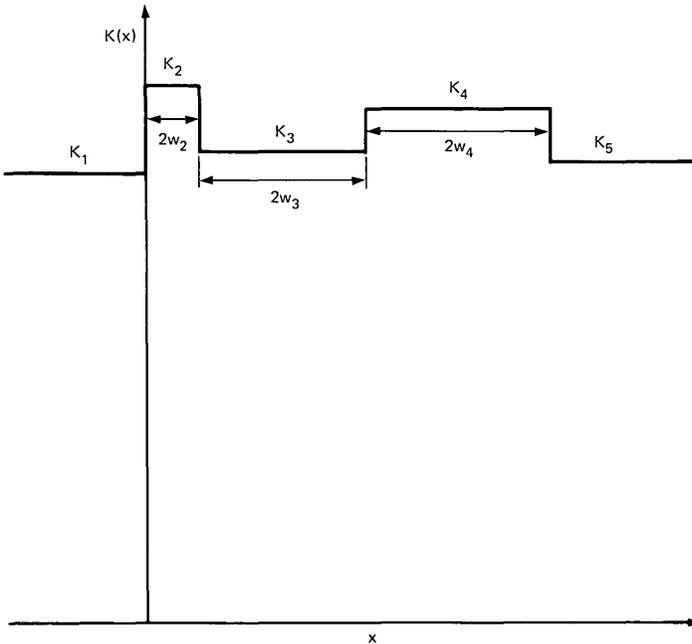
Fig. 1—The optical dielectric profile $K(x)$ for the five-slab model.

z-direction with propagation constant β, then the electric and magnetic fields are independent of y and can be expressed as

$$E = e(x) \exp i(\omega t - \beta z),$$
$$H = h(x) \exp i(\omega t - \beta z),$$

where ω is the angular frequency of the light and t is the time. Both TE and TM modes exist. It follows from Maxwell's equations that the electric field $e_y(x)$ of a TE mode is described by

$$\frac{d^2 e_y}{dx^2} + [k^2 K(x) - \beta^2] e_y = 0, \tag{1}$$

and that the magnetic field $h_y(x)$ of a TM mode is described by

$$K(x) \frac{d}{dx} \left(\frac{1}{K(x)} \frac{dh_y}{dx} \right) + [k^2 K(x) + \beta^2] h_y = 0. \tag{2}$$

It is required that

$$e_y, \frac{de_y}{dx},$$

$$h_y, \frac{1}{K(x)} \frac{dh_y}{dx}$$

be continuous. Since $K(x)$ is piecewise constant, solution of eqs. (1) and (2) subject to the above conditions is straightforward. The solution of (1) is

$$
\begin{aligned}
e_y(x) &= A \exp p_1 x \qquad x < 0 \\
&= A[(p_1/p_2) \sin p_2 x + \cos p_2 x] \qquad 0 < x < 2w_2 \\
&= AC_2[1 + (p_1/p_2)T_2][-X \sinh p_3(x - 2w_2) \\
&\qquad + \cosh p_3(x - 2w_2)] \qquad 2w_2 < x < 2(w_2 + w_3) \\
&= AC_2 C_3[1 + (p_1/p_2)T_2][1 - XT_3] \\
&\qquad \times [(p_3/p_4)Y \sin p_4(x - 2w_2 - 2w_3) + \cos p_4(x - 2w_2 - 2w_3)] \\
&\qquad\qquad\qquad 2(w_2 + w_3) < x < 2(w_2 + w_3 + w_4) \\
&= AC_2 C_3 C_4[1 + (p_1/p_2)T_2][1 - XT_3] \\
&\qquad \times [1 + (p_3/p_4)YT_4] \exp p_5(2w_2 + 2w_3 + 2w_4 - x) \\
&\qquad\qquad\qquad 2(w_2 + w_3 + w_4) < x, \quad (3)
\end{aligned}
$$

with

$$
\begin{aligned}
p_i(\beta) &= (\beta^2 - k^2 K_i)^{\frac{1}{2}} \qquad i = 1, 3, 5, \\
p_i(\beta) &= (k^2 K_i - \beta^2)^{\frac{1}{2}} \qquad i = 2, 4,
\end{aligned} \tag{4}
$$

$$
\begin{aligned}
C_2(\beta) &= \cos 2p_2 w_2, & T_2(\beta) &= \tan 2p_2 w_2, \\
C_3(\beta) &= \cosh 2p_3 w_3, & T_3(\beta) &= \tanh 2p_3 w_3, \\
C_4(\beta) &= \cos 2p_4 w_4, & T_4(\beta) &= \tan 2p_4 w_4,
\end{aligned} \tag{5}
$$

$$X(\beta) = \frac{1}{p_3}\left[\frac{p_2 T_2 - p_1}{1 + (p_1/p_2)T_2}\right], \tag{6}$$

$$Y(\beta) = \frac{1}{p_3}\left[\frac{p_4 T_4 - p_5}{1 + (p_5/p_4)T_4}\right]. \tag{7}$$

The amplitude A is arbitrary. Equation (3) satisfies the continuity condition on $e_y(x)$ everywhere, and that on de_y/dx at all but the point $x = 2(w_2 + w_3)$. The continuity condition at this point leads to the eigenvalue equation

$$T_3(\beta) = \frac{X(\beta) + Y(\beta)}{1 + X(\beta)Y(\beta)}, \tag{8}$$

which determines the values of the propagation constant β for which discrete modes can exist.

For TM modes, the analogous equations are formed by replacing p_i by $\tilde{p}_i$ and w_i by $\tilde{w}_i$, where

$$K_i \tilde{p}_i(\tilde{\beta}) = (\tilde{\beta}^2 - k^2 K_i)^{\frac{1}{2}} \qquad i = 1, 3, 5,$$
$$K_i \tilde{p}_i(\tilde{\beta}) = (k^2 K_i - \tilde{\beta}^2)^{\frac{1}{2}} \qquad i = 2, 4,$$
$$\tilde{w}_i = K_i w_i \qquad i = 2, 3, 4,$$

and $\tilde{\beta}$ denotes the propagation constant for a TM mode. For simplicity of exposition, we have mainly confined our analysis to that of TE modes. It should be clear how to do the corresponding analysis for TM modes.

We remark that the above analysis is quite general and makes no assumptions about the relative heights or widths involved in the dielectric profile $K(x)$ sketched in Fig. 1. By making appropriate choices of the parameters, we could deduce from eqs. (3) to (8) the corresponding equations for a single asymmetric or symmetric waveguide, for example. Two cases which interest us particularly are: (i) $K_3 = K_4 = K_5$ with $K_2 > K_1$ and $K_2 > K_3$, and (ii) $K_1 = K_2 = K_3$ with $K_4 > K_3$ and $K_4 > K_5$. Each of these models an isolated waveguide. We call the first of these (with high dielectric region K_2) guide II and the other (with high dielectric region K_4) guide IV. The eigenvalue equation (8) then reduces to

$$X(\beta) = 1 \qquad \text{guide II},$$
$$Y(\beta) = 1 \qquad \text{guide IV}.$$

These may appear more familiar to the reader when cast in the standard

form for a single asymmetric guide[6]

$$\tan 2w_2 p_2 = \frac{(p_1/p_2) + (p_3/p_2)}{1 - (p_1/p_2)(p_3/p_2)} \qquad \text{guide II,} \qquad (9)$$

$$\tan 2w_4 p_4 = \frac{(p_3/p_4) + (p_5/p_4)}{1 - (p_3/p_4)(p_5/p_4)} \qquad \text{guide IV.} \qquad (10)$$

For the model of two adjacent waveguides, we place guides II and IV adjacent to each other as illustrated in Fig. 1, with $K_2 > \max \{K_1, K_3\}$, $K_4 > \max \{K_3, K_5\}$, and $w_i > 0$ ($i = 1, 2, 3, 4$).

It is also a relatively straightforward procedure to write down precisely how many modes can exist with a given dielectric profile $K(x)$.[7] Although we omit such expressions here, we do comment that, just as a single asymmetric guide may not be able to support a propagating mode, so also an asymmetric double waveguide structure is not always capable of mode propagation. If $K_1 = K_3 = K_5$, though, so that the structure is composed of two parallel symmetric (but not necessarily identical) waveguides, then there is always at least one mode.

III. NEARLY DEGENERATE MODES

Let β_2 and β_4 denote solutions of $X(\beta) = 1$ and $Y(\beta) = 1$, respectively; β_2 and β_4, then, are propagation constants for modes in guides II and IV if the guides were isolated from each other. We need make no assumption about the order of each mode. In practice, though, both propagation constants are likely to be associated with zeroth order modes. For definiteness in notation, we assume that $\beta_4 \geqq \beta_2$ and write[*]

$$\beta_4 - \beta_2 = 2\Delta.$$

We now assume that Δ is "small," i.e., that the two modes are nearly degenerate. This assumption, which is fundamental to the remainder of the analysis, is stated more explicitly later [eq. (18)]. Frequently, it does not matter (to the order of approximation used) whether β_2 or β_4 is used in the evaluation of an expression. In such instances, it sometimes is helpful to use the notation $\beta_0 (\doteq \beta_2 \doteq \beta_4)$.

Our task is now to determine values of β which satisfy (8). Our experience with the degenerate case ($\Delta = 0$) leads us to expect that

---

[*] In the numerical example of Section VI, we relax this notation to read $|\beta_4 - \beta_2| = 2\Delta$, where it is not known a priori whether β_2 or β_4 is larger. This should not be confusing when taken in context.

there will be two solutions β_+ and β_- close to β_0. A study of the coupled line equations[8] for the two modes would demonstrate that $\beta_+ = \beta_4 + \bar{\delta}$ and $\beta_- = \beta_2 - \bar{\delta}$, where $\bar{\delta}$ is expressed in terms of the (unknown) coupling coefficients.[6] We prefer to attack (8) directly; we shall verify the expressions for β_+ and β_-, prove that $\bar{\delta} > 0$, and give an explicit formula for $\bar{\delta}$.

We first show that, if β_2 and β_4 are close enough together, then (8) has no solution $\bar{\beta}$ such that $\beta_2 \leqq \bar{\beta} \leqq \beta_4$. Since $X(\beta_2) = 1$ and $X'(\beta) \leqq 0$ for all β, we know that if $(\bar{\beta} - \beta_2)/\beta_2 \ll 1$, then $0 < X(\bar{\beta}) \leqq 1$. Similarly, $Y(\bar{\beta}) \geqq 1$. Thus,

$$\frac{X(\bar{\beta}) + Y(\bar{\beta})}{1 + X(\bar{\beta})Y(\bar{\beta})} = \frac{1 + X(\bar{\beta})Y^{-1}(\bar{\beta})}{X(\bar{\beta}) + Y^{-1}(\bar{\beta})} \geqq 1.$$

But $T_3(\beta) = \tanh 2w_3 p_3 < 1$ for all β, so (8) is not satisfied.

Next, suppose (8) has a solution $\beta_+ = \beta_4 + \bar{\delta}$, with $\bar{\delta} > 0$. Then if $\bar{\delta}$ is small, we know that $0 < X(\beta_+) < 1$ and $0 < Y(\beta_+) < 1$, so (8) may be written as

$$T(\beta) \equiv \tanh w_3 p_3 = \tanh \left[\tfrac{1}{2}(\tanh^{-1} X + \tanh^{-1} Y)\right]$$

$$= \frac{X[1 + (1 - Y^2)^{\frac{1}{2}}] + Y[1 + (1 - X^2)^{\frac{1}{2}}]}{XY + [1 + (1 - X^2)^{\frac{1}{2}}][1 + (1 - Y^2)^{\frac{1}{2}}]}. \tag{11}$$

We have

$$\begin{aligned}
X(\beta_+) &= X(\beta_2 + 2\Delta + \bar{\delta}) \doteq 1 + (2\Delta + \bar{\delta})X'(\beta_0), \\
Y(\beta_+) &= Y(\beta_4 + \bar{\delta}) \doteq 1 + \bar{\delta}Y'(\beta_0).
\end{aligned} \tag{12}$$

If we substitute (12) in (11) and perform a perturbation analysis under the two assumptions,

$$[\bar{\delta}(2\Delta + \bar{\delta})X'(\beta_0)Y'(\beta_0)]^{\frac{1}{2}} \ll 1, \tag{13}$$

$$[-(2\Delta + \bar{\delta})X'(\beta_0)]^{\frac{1}{2}} + [-\bar{\delta}Y'(\beta_0)]^{\frac{1}{2}} \ll 1, \tag{14}$$

we find

$$T(\beta_0) = 1 - [-(2\Delta + \bar{\delta})X'(\beta_0)]^{\frac{1}{2}}[-\bar{\delta}Y'(\beta_0)]^{\frac{1}{2}},$$

so that

$$\bar{\delta} = -\Delta + (\Delta^2 + \delta^2)^{\frac{1}{2}}, \tag{15}$$

where

$$\delta = \frac{1 - T(\beta_0)}{[X'(\beta_0)Y'(\beta_0)]^{\frac{1}{2}}}. \tag{16}$$

On the other hand, if we suppose that (8) has a solution $\beta_- = \beta_2 - \hat{\delta}$, with $\hat{\delta} > 0$, then $X(\beta_-) > 1$, $Y(\beta_-) > 1$, and (8) becomes

$$T(\beta) \equiv \tanh w_3 p_3 = \tanh \left[\tfrac{1}{2}(\tanh^{-1} X^{-1} + \tanh^{-1} Y^{-1}) \right]$$

$$= \frac{X + (X^2 - 1)^{\frac{1}{2}} + Y + (Y^2 - 1)^{\frac{1}{2}}}{1 + [X + (X^2 - 1)^{\frac{1}{2}}][Y + (Y^2 - 1)^{\frac{1}{2}}]}.$$

By a procedure very much like that used to determine $\bar{\delta} = \beta_+ - \beta_4$, we find that $\hat{\delta} = \bar{\delta}$, as anticipated. Here, the roles of X' and Y' must be interchanged in (14).

The effect, then, of placing guides II and IV next to each other is to shift their (isolated) propagation constants β_2 and β_4 symmetrically outward by $\bar{\delta}$ to β_- and β_+. The physical meaning of δ in (16) is clear: it is the magnitude of the shift which would occur if guides II and IV had degenerate modes ($\Delta = 0$). We shall call δ the "degenerate shift."

Let us consider assumptions (13) and (14) in more detail. By means of (15) and (16), (13) becomes

$$1 - T(\beta_0) = 1 - \tanh w_3 p_3 \ll 1. \tag{17}$$

This, then, is essentially a restriction on the separation between the two waveguides. If they are too close, our approximations will break down. Assumption (14) and its counterpart with the roles X' and Y' reversed are, by (15), satisfied if

$$[\Delta + (\Delta^2 + \delta^2)^{\frac{1}{2}}]^{\frac{1}{2}}\{[-X'(\beta_0)]^{\frac{1}{2}} + [-Y'(\beta_0)]^{\frac{1}{2}}\} \ll 1. \tag{18}$$

This tells us how large Δ can get without invalidating the approximations.

By using eqs. (4), (6), and (7) cleverly, we see that the expressions for $X'(\beta_0)$ and $Y'(\beta_0)$ reduce to the simple forms

$$X'(\beta_0) = \frac{-\beta_0(p_2^2 + p_3^2)}{p_1 p_2^2 p_3} [2p_1 w_2 + 1 + (p_1/p_3)], \tag{19}$$

$$Y'(\beta_0) = \frac{-\beta_0(p_3^2 + p_4^2)}{p_3 p_4^2 p_5} [2p_5 w_4 + 1 + (p_5/p_3)]. \tag{20}$$

Thus, by (16),

$$\delta = \frac{p_2 p_3 p_4 [1 - \tanh w_3 p_3]}{\beta_0 \left\{ \dfrac{1}{p_1 p_5}(p_2^2 + p_3^2)(p_3^2 + p_4^2)\left[2p_1 w_2 + 1 + \dfrac{p_1}{p_3}\right]\left[2p_5 w_4 + 1 + \dfrac{p_5}{p_3}\right]\right\}^{\frac{1}{2}}}. \tag{21}$$

If we take the case of two identical symmetric waveguides $(p_1 = p_3 = p_5,\ p_2 = p_4,\ w_2 = w_4)$ and use the approximation

$$\tanh w_3 p_1 \doteq 1 - 2 \exp\,(-2w_3 p_1),$$

then (21) reduces to

$$\delta = \frac{p_1^2 p_2^2 \exp\,(-2w_3 p_1)}{\beta_0 (p_1^2 + p_2^2)(1 + p_1 w_2)} \tag{22}$$

which is in agreement with results of Marcuse.[2] For TM modes, we arrive at

$$\tilde{\beta}_+ = \tilde{\beta}_4 + \tilde{\delta}, \qquad \tilde{\beta}_- = \tilde{\beta}_2 - \tilde{\delta}, \qquad \tilde{\Delta} = \tilde{\beta}_4 - \tilde{\beta}_2,$$

where

$$\tilde{\delta} = -\tilde{\Delta} + (\tilde{\Delta}^2 + \hat{\delta}^2)^{\frac{1}{2}},$$

$$\hat{\delta} = \frac{1 - \tanh \tilde{w}_3 \tilde{p}_3}{[\tilde{X}'(\tilde{\beta}_0)\,\tilde{Y}'(\tilde{\beta}_0)]^{\frac{1}{2}}},$$

$$\tilde{X}'(\tilde{\beta}_0) = \frac{-\tilde{\beta}_0(\tilde{p}_2^2 + \tilde{p}_3^2)}{K_1^2 K_2^2 K_3^2 \tilde{p}_1 \tilde{p}_2^2 \tilde{p}_3} \left[2K_1^2 K_3^2 \tilde{p}_1 \tilde{w}_2 \right.$$
$$\left. + K_3^2 \left(\frac{K_1^2 \tilde{p}_1^2 + K_2^2 \tilde{p}_2^2}{\tilde{p}_1^2 + \tilde{p}_2^2} \right) + \frac{\tilde{p}_1}{\tilde{p}_3} \left(\frac{K_2^2 \tilde{p}_2^2 + K_3^2 \tilde{p}_3^2}{\tilde{p}_2^2 + \tilde{p}_3^2} \right) \right],$$

$$\tilde{Y}'(\tilde{\beta}_0) = \frac{-\tilde{\beta}_0(\tilde{p}_3^2 + \tilde{p}_4^2)}{K_3^2 K_4^2 K_5^2 \tilde{p}_3 \tilde{p}_4^2 \tilde{p}_5} \left[2K_3^2 K_5^2 \tilde{p}_5 \tilde{w}_4 \right.$$
$$\left. + K_3^2 \left(\frac{K_4^2 \tilde{p}_4^2 + K_5^2 \tilde{p}_5^2}{\tilde{p}_4^2 + \tilde{p}_5^2} \right) + \frac{\tilde{p}_5}{\tilde{p}_3} \left(\frac{K_3^2 \tilde{p}_3^2 + K_4^2 \tilde{p}_4^2}{\tilde{p}_3^2 + \tilde{p}_4^2} \right) \right].$$

IV. A LOOK AT THE MODES

We now discuss what the modes $e_\pm(x)$ associated with $\beta_\pm$ are like in guides II and IV. The expressions for $e_\pm(x)$ are given by (3). Both the shapes of the modes and their relative amplitudes will be of interest.

We see from (3) that the *shape* of $e_\pm(x)$ in guide II is given by

$$f(\beta_\pm,\ x) = [(p_1/p_2) \sin p_2 x + \cos p_2 x]|_{\beta_\pm}. \tag{23}$$

Since

$$f(\beta_+,\ x) \doteq f(\beta_2,\ x) + (2\Delta + \bar{\delta}) \left. \frac{\partial f}{\partial \beta} \right|_{\beta_2},$$

$$f(\beta_-,\ x) \doteq f(\beta_2,\ x) - \bar{\delta} \left. \frac{\partial f}{\partial \beta} \right|_{\beta_2},$$

the shapes of both modes differ just slightly from the unperturbed shape $f(\beta_2, x)$; furthermore, the shifts for the two modes are unequal

and are in opposite directions. The unperturbed shape can be determined with the aid of (9). We find

$$\frac{p_1}{p_2} = \tan\left[w_2 p_2 + \frac{1}{2}\left(\tan^{-1}\frac{p_1}{p_2} - \tan^{-1}\frac{p_3}{p_2} \right) \right] \equiv \tan U$$

for even-numbered modes and

$$\frac{p_1}{p_2} = -\cot U$$

for odd-numbered modes. Thus,

$$f(\beta_2, x) = \tan U \sin p_2 x + \cos p_2 x$$
$$= \sec U \cos (p_2 x - U)$$
$$= \frac{(p_1^2 + p_2^2)^{\frac{1}{2}}}{p_2} \cos \left[p_2 x - w_2 p_2 \right.$$
$$\left. - \frac{1}{2}\left(\tan^{-1}\frac{p_1}{p_2} - \tan^{-1}\frac{p_3}{p_2} \right) \right] \quad (24)$$

for even-numbered modes and, similarly,

$$f(\beta_2, x) = \frac{(p_1^2 + p_2^2)^{\frac{1}{2}}}{p_2} \sin \left[p_2 x - w_2 p_2 \right.$$
$$\left. - \frac{1}{2}\left(\tan^{-1}\frac{p_1}{p_2} - \tan^{-1}\frac{p_3}{p_2} \right) \right] \quad (25)$$

for odd-numbered modes. The mode shapes in guide IV can be determined in an analogous manner, with perturbations performed about β_4 instead of β_2. We leave the details to the reader.

The above results are not surprising. If the double waveguide system has a mode with propagation constant β_+ or β_- which is close to the propagation constants β_2 and β_4 of modes that can travel in the individual isolated waveguides, then we would indeed expect the shape of that double waveguide mode to deviate only slightly in each waveguide from the shape of the mode that could propagate in the isolated waveguide.

The amplitudes of $e_\pm(x)$ in guides II and IV prove to be more interesting. Let the arbitrary amplitude A in (3) be written as $A_\pm$ for $e_\pm(x)$, and let $B_\pm^{\text{II}}$ and $B_\pm^{\text{IV}}$ denote the amplitudes of the modes in guides II and IV, respectively. Then (3) and (24) or (25) show that the mode amplitudes in guide II are given (to our order of approximation) by

$$B_\pm^{\text{II}} = A_\pm \frac{(p_1^2 + p_2^2)^{\frac{1}{2}}}{p_2}. \quad (26)$$

In guide IV, the amplitudes are, by (3),

$$B_\pm^{IV} = A_\pm \left[C_2 C_3 \left(1 + \frac{p_1}{p_2} T_2 \right) (1 - X T_3) \frac{(p_3^2 + p_4^2)^{\frac{1}{2}}}{p_4} \right] \Big|_{\beta_\pm} . \quad (27)$$

Care must be taken in the evaluation of $C_3(1 - XT_3)$ at $\beta_\pm$. Since $X(\beta_+) < 1$, $Y(\beta_+) < 1$, and by (8),

$$T_3 = \frac{X + Y}{1 + XY} = \tanh(\tanh^{-1} X + \tanh^{-1} Y),$$

we have

$$C_3 = \cosh(\tanh^{-1} X + \tanh^{-1} Y) = \frac{1 + XY}{(1 - X^2)^{\frac{1}{2}}(1 - Y^2)^{\frac{1}{2}}},$$

so that

$$C_3(1 - XT_3)|_{\beta_+} = \left(\frac{1 - X^2}{1 - Y^2} \right)^{\frac{1}{2}} \Big|_{\beta_+} \doteq \left[\frac{(2\Delta + \bar{\delta})X'}{\bar{\delta}Y'} \right]^{\frac{1}{2}}.$$

We find in a similar manner that

$$C_3(1 - XT_3)|_{\beta_-} = - \left[\frac{\bar{\delta}X'}{(2\Delta + \bar{\delta})Y'} \right]^{\frac{1}{2}}.$$

Thus the mode amplitudes in guide IV are

$$B_+^{IV} = A_+ C_2 \left(1 + \frac{p_1}{p_2} T_2 \right) \frac{(p_3^2 + p_4^2)^{\frac{1}{2}}}{p_4} \left[\frac{(2\Delta + \bar{\delta})X'}{\bar{\delta}Y'} \right]^{\frac{1}{2}}.$$

$$B_-^{IV} = - A_- C_2 \left(1 + \frac{p_1}{p_2} T_2 \right) \frac{(p_3^2 + p_4^2)^{\frac{1}{2}}}{p_4} \left[\frac{\bar{\delta}X'}{(2\Delta + \bar{\delta})Y'} \right]^{\frac{1}{2}}. \quad (28)$$

We observe that the mode amplitudes will have the *same* signs in one waveguide and the *opposite* signs in the other. Thus, $e_+(x)$ might be termed quasi-even and $e_-(x)$ quasi-odd. More startling, however, is the realization that, if $\Delta > 0$, then the ratio $|B_+^{II}/B_-^{II}| = A_+/A_-$ may be quite different from the ratio $|B_+^{IV}/B_-^{IV}| = (A_+/A_-)[(2\Delta + \bar{\delta})/\bar{\delta}]$. As we see in the next section, this will have serious implications when we consider the double waveguide system as a device for transferring power.

For future reference, we write $e_\pm(x)$ for a system consisting of two symmetric, but not necessarily identical, waveguides ($K_1 = K_3 = K_5$). In this instance, (9) and (10) imply that

$$C_2[1 + (p_1/p_2)T_2] = 1,$$
$$C_4[1 + (p_1/p_4)T_4] = 1,$$

so that we have by (3) and our previous analysis

$$e_+(x) = A_+ \exp p_1 x \qquad x < 0$$

$$= A_+ \frac{(p_1^2 + p_2^2)^{\frac{1}{2}}}{p_2} \cos p_2(x - w_2) \qquad 0 < x < 2w_2$$

$$= A_+[-X \sinh p_1(x - 2w_2) + \cosh p_1(x - 2w_2)]\big|_{\beta_+}$$
$$2w_2 < x < 2(w_2 + w_3)$$

$$= A_+ \left[\frac{(2\Delta + \bar{\delta})}{\bar{\delta}} \frac{X'}{Y'} \right]^{\frac{1}{2}} \frac{(p_1^2 + p_4^2)^{\frac{1}{2}}}{p_4} \cos p_4(x - 2w_2 - 2w_3 - w_4)$$
$$2(w_2 + w_3) < x < 2(w_2 + w_3 + w_4)$$

$$= A_+ \left[\frac{(2\Delta + \bar{\delta})}{\bar{\delta}} \frac{X'}{Y'} \right]^{\frac{1}{2}} \exp p_1(2w_2 + 2w_3 + 2w_4 - x)$$
$$2(w_2 + w_3 + w_4) < x, \qquad (29)$$

$$e_-(x) = A_- \exp p_1 x \qquad x < 0$$

$$= A_- \frac{(p_1^2 + p_2^2)^{\frac{1}{2}}}{p_2} \cos p_2(x - w_2) \qquad 0 < x < 2w_2$$

$$= A_-[-X \sinh p_1(x - 2w_2) + \cosh p_1(x - 2w_2)]\big|_{\beta_-}$$
$$2w_2 < x < 2(w_2 + w_3)$$

$$= -A_- \left[\frac{\bar{\delta} X'}{(2\Delta + \bar{\delta})Y'} \right]^{\frac{1}{2}} \frac{(p_1^2 + p_4^2)^{\frac{1}{2}}}{p_4} \cos p_4(x - 2w_2 - 2w_3 - w_4)$$
$$2(w_2 + w_3) < x < 2(w_2 + w_3 + w_4)$$

$$= -A_- \left[\frac{\bar{\delta} X'}{(2\Delta + \bar{\delta})Y'} \right]^{\frac{1}{2}} \exp p_1(2w_2 + 2w_3 + 2w_4 - x)$$
$$2(w_2 + w_3 + w_4) < x. \qquad (30)$$

V. BEAT LENGTH AND POWER TRANSFER

Suppose the two modes $E_\pm = e_\pm(x) \exp i(\omega t - \beta_\pm z)$ travel down the double waveguide device. Since they travel at different phase velocities, the quasi-even and quasi-odd modes will alternately add and (partially) cancel in each waveguide. Hence, power is transferred between the two waveguides.

The beat length L over which this transfer takes place is given by

$$L = \frac{\pi}{2(\bar{\delta} + \Delta)} = \frac{\pi}{2\bar{\delta}[1 + (\Delta/\bar{\delta})^2]^{\frac{1}{2}}}. \qquad (31)$$

Note that, if the degenerate shift $\bar{\delta}$ is fixed, then as the mismatching Δ increases, the beat length L decreases. We can conceive of ways to

tune the double waveguide device and thus to change the beat length. This might be useful for light switching or modulation.

It is important to learn just how much power can be transferred in the waveguide system. Suppose, for definiteness, that we excite just one waveguide at $z = 0$ (say, guide IV), with the intent of transferring power to guide II via the mode-coupling mechanism. If guides II and IV have degenerate modes ($\Delta = 0$), then as the modes travel down the waveguide system, they will alternately add and then cancel (to order δ^2) in each waveguide, with addition occurring in one waveguide when at the same position cancellation occurs in the other. If guides II and IV have nondegenerate modes ($\Delta > 0$), however, then complete cancellation cannot take place in *both* waveguides: by (27) and (28), we see that if the amplitudes of $e_+(x)$ and $e_-(x)$ are adjusted so that the modes cancel at $z = 0$ in guide II, then they will never cancel fully in guide IV.

From a practical point of view, a parameter which is likely to be of interest in this matter is the fraction of the *total* power introduced into the system which can be transferred into guide II. If the modes are poorly confined, an appreciable fraction of the power carried by a waveguide may actually be outside the high dielectric guiding region. If the reader is interested in the fraction of the power which can be transferred not only to the guiding region of guide II, but also to its vicinity, we would need a power transfer ratio G to be defined by

$$ G = \int_{-\infty}^{a} [e_+(x) + e_-(x)]^2 dx \bigg/ \int_{-\infty}^{\infty} [e_+^2(x) + e_-^2(x)]dx, $$

where a is some number between $2w_2$ and $2(w_2 + w_3)$ which defines the "boundary" between guides II and IV. The numerator of this expression, then, is proportional to the power carried by the entire guide II.

Unfortunately, for a general asymmetric waveguide system, it is not at all clear how to define the position of the "boundary" between the two waveguides. If the system consists of two symmetric waveguides which are nearly identical (except for a small deviation if the modes are slightly mismatched), then it seems clear that the boundary should be midway between the two dielectric regions, i.e., at $a = 2w_2 + w_3$. By using (29) and (30), we find in this instance that we have to first order

$$ G = [1 + (\Delta/\delta)^2]^{-1}. $$

Thus for perfectly matched waveguides ($\Delta = 0$), the power transfer

is complete, to first order. As the mismatching increases, the power transfer ratio decreases rapidly.

Complete power transfer (to first order) is a direct consequence of assumption (17), which implies little overlap between the field associated with guide II and that associated with guide IV. A higher order perturbation analysis would show that in fact there is some field overlap and that, even if $\Delta = 0$, the power transfer is not complete. As the waveguide separation increases, there would be less field overlap and the power transfer would be more nearly complete.

For a general asymmetric waveguide system, we might define the "boundary" between the two waveguides to be, say, at the position where the "quasi-even" field attains its minimum. Such a definition can be cumbersome to apply mathematically. In general, though, we would expect results similar to those obtained for the symmetric system. If the modes are degenerate and one waveguide is excited, then virtually all the power can be transferred to the vicinity of the other waveguide. The power transfer ratio decreases as the mismatching increases.

It will be instructive to introduce a second power transfer ratio F, which can be defined precisely. It will be the fraction of the total power introduced into the system which can be transferred into the *high dielectric region* of guide II, the waveguide which was originally unexcited. If terms of order Δ/β_0 are neglected, this power transfer ratio is defined by

$$F = \int_0^{2w_2} [e_+(x) + e_-(x)]^2 dx \bigg/ \int_{-\infty}^{\infty} [e_+^2(x) + e_-^2(x)] dx, \quad (32)$$

where the mode amplitudes A_+ and A_- are equal.

If the modes are poorly confined in guide II, the power transfer ratio F may be considerably less than unity even if the waveguides are perfectly matched ($\Delta = 0$). The definition of F is concerned only with the power which can be transferred into the high dielectric region of guide II; hence, F depends upon the confinement factor of the waveguide (to be defined below) as well as upon the amount of mismatching Δ.

Evaluation of (32) can be very messy for the general case of two asymmetric waveguides. We simplify the subsequent analysis and yet retain its essential flavor by assuming that the double waveguide structure is composed of two symmetric, but not necessarily identical, waveguides ($K_1 = K_3 = K_5$). The modes $e_\pm(x)$ are then given by (29)

and (30). For the numerator of F, we find

$$\int_0^{2w_2} [e_+(x) + e_-(x)]^2 dx = 4A_+^2 \frac{(p_1^2 + p_2^2)}{p_2^2} \int_0^{2w_2} \cos^2 p_2(x - w_2) dx$$

$$= 4A_+^2 \frac{(p_1^2 + p_2^2)}{p_2^3} [p_2 w_2 + \tfrac{1}{2} \sin 2p_2 w_2].$$

But since (9) implies that

$$\sin 2p_2 w_2 = \frac{2p_1 p_2}{p_1^2 + p_2^2}$$

for a symmetric waveguide, we have

$$\int_0^{2w_2} [e_+(x) + e_-(x)]^2 dx = \frac{4A_+^2}{p_2} \left[p_2 w_2 \left(\frac{p_1^2 + p_2^2}{p_2^2} \right) + \frac{p_1}{p_2} \right].$$

Similar procedures can be used in the evaluation of the denominator of F. Special care must be taken in evaluating the integral in the interval $2w_2 < x < 2(w_2 + w_3)$, where $e_+(x)$ and $e_-(x)$ have the same functional description, but the function is evaluated at β_+ and β_-, respectively. In the resulting analysis, C_3 and S_3 must be evaluated at $\beta_\pm$. Procedures similar to those used following (27) are helpful. The power transfer ratio turns out to be

$$F = \frac{4(P_{II}/p_2)}{2[(P_{II}/p_2) + (1/p_1)] + 2[(P_{IV}/p_4) + (1/p_1)](1 + 2\Delta^2/\delta^2)X'/Y'},$$

where

$$P_{II} = p_2 w_2 \left(\frac{p_1^2 + p_2^2}{p_2^2} \right) + \frac{p_1}{p_2},$$

$$P_{IV} = p_4 w_4 \left(\frac{p_1^2 + p_4^2}{p_4^2} \right) + \frac{p_1}{p_4}. \tag{33}$$

The expression for F can be simplified. By using (19), (20), and (33), we find that

$$\frac{P_{II}}{p_2} + \frac{1}{p_1} = \frac{X'}{Y'} \left(\frac{P_{IV}}{p_4} + \frac{1}{p_1} \right), \tag{34}$$

so that

$$F = \frac{P_{II}/p_2}{[(P_{II}/p_2) + (1/p_1)][1 + (\Delta^2/\delta^2)]}. \tag{35}$$

Both (34) and (35) have interesting physical interpretations. In order to discuss them, we make a brief digression. Suppose that, instead of the double waveguide system, we just have an isolated

guide II, which for generality we assume is not necessarily symmetric. (In the notation of Section II, we have $K_3 = K_4 = K_5$.) If the electric field is given by $e(x) \exp i(wt - \beta_2 z)$, then the power carried by the entire waveguide is proportional to

$$P = \int_{-\infty}^{\infty} e^2(x) dx$$

and the fraction of the total power which is confined to the high dielectric region is given by

$$C = \int_0^{2w_2} e^2(x) dx \Big/ \int_{-\infty}^{\infty} e^2(x) dx.$$

It is not difficult to show that

$$P = \frac{p_1^2 + p_2^2}{2p_1 p_2^2} [2p_1 w_2 + 1 + (p_1/p_3)],$$

where $e(x)$ was assumed to have unit amplitude at $x = 0$. Upon comparing this expression with (19), we find that

$$X' = -\frac{2\beta_0}{p_3} \frac{(p_2^2 + p_3^2)}{(p_1^2 + p_2^2)} P,$$

so that X' is related in a simple manner to the power carried by the waveguide with which it is associated. Now if the isolated guide II happens to be symmetric, then it is also true that

$$P = \frac{P_{II}}{p_2} + \frac{1}{p_1},$$

$$X' = -\frac{2\beta_0}{p_1} P.$$

Since a similar relationship holds for an isolated guide IV, (34) follows. The physical implication of (34) is that, if a single waveguide is excited in a matched double waveguide system ($\Delta = 0$), then the power is distributed between the two modes in such a manner that

$$\int_{-\infty}^{\infty} e_+^2(x) dx = \int_{-\infty}^{\infty} e_-^2(x) dx.$$

If an isolated guide II is symmetric, then the confinement factor C can be shown to be given by

$$C = \frac{P_{II}/p_2}{(P_{II}/p_2) + (1/p_1)}. \tag{36}$$

Hence, the power transfer ratio F is simply given by

$$F = C[1 + (\Delta/\delta)^2]^{-1}. \tag{37}$$

Note that F is completely independent of the parameters of guide IV, as well as of w_3 because of assumption (17).

We remark once more that these results are first-order approximations. If a higher order perturbation analysis were undertaken, it would show that F is also dependent upon the amount of field overlap between the two waveguides.

It is important to note that the amount of mismatching Δ can have a significant effect on the power transfer ratio: as Δ increases, F decreases rapidly. We might remark that, while the beat length L depends upon δ and the ratio Δ/δ, the power transfer ratio F depends only upon Δ/δ. Thus, by proper device design, it may be possible to adjust δ and Δ/δ to get both appreciable power transfer and a desirable beat length.

VI. A NUMERICAL EXAMPLE

To illustrate our results, let us give an example, using parameters which could be realized in a $GaAs - Al_xGa_{1-x}As$ heterostructure. We consider two waveguides which are each symmetric, but which have different widths and dielectric step heights. Our intent is to excite one waveguide and then to transfer power into the other by means of mode coupling. The parameters of one waveguide, say guide II, are taken to be fixed. The width of the other waveguide is considered a variable. For any given width $2w_4$, we adjust the dielectric height K_4 so that the propagation constants β_2 and β_4 for the zeroth order TE modes of the two waveguides match.

We learn how the degenerate shift δ varies with the spacing $2w_3$ between the two waveguides and with the width $2w_4$ of guide IV. We look at the beat length L and the power transfer ratio F for an idea of how much mismatching of the propagation constants β_2 and β_4 can be tolerated. We then discuss the amount of mismatching $2\Delta = |\beta_4 - \beta_2|$ which might occur in a practical situation and show how tuning can reduce this.

To be specific, suppose that

$$K_2 = 11.868,$$
$$K_1 = 0.9K_2 = 10.681,$$
$$w_2 = 0.1 \ \mu\text{m},$$
$$k = 5.4494 \times 10^4 \ \text{cm}^{-1}.$$

The dielectric constant K_2 corresponds to an index of refraction $K_2^{\frac{1}{2}} = 3.445$ of GaAs at a wavelength $\lambda = 2\pi\, k^{-1} = 1.153\ \mu\text{m}$. The halfwidth w_4 of guide IV will be assumed to vary between 0.1 μm and 0.8 μm.

For the zeroth order TE mode of a symmetric waveguide, we find that (9) reduces to

$$\tan p_2 w_2 = p_1/p_2,$$
$$p_1 = (\beta_2^2 - k^2 K_1)^{\frac{1}{2}},$$
$$p_2 = (k^2 K_2 - \beta_2^2)^{\frac{1}{2}}. \tag{38}$$

The above parameters for guide II give us

$$\beta_2 = 1.8049 \times 10^5 \text{ cm}^{-1},$$
$$p_1 = 2.9311 \times 10^4 \text{ cm}^{-1},$$
$$p_2 = 5.1631 \times 10^4 \text{ cm}^{-1}.$$

Equation (38) holds for guide IV if each subscript 2 is replaced by a 4. For any given value of w_4, we adjust K_4 so that β_4 is the same as β_2. Some values of K_4 and p_4 are given in Table I.

The parameters for K_1, K_2, w_2 and any given pair K_4, w_4 then define two waveguides which have a degenerate mode at the given wavelength λ. The degenerate shift δ is given by (21) for symmetric waveguides:

$$\delta = \frac{p_1^2 p_2 p_4 (1 - \tanh w_3 p_1)}{2\beta_2 [(p_1^2 + p_2^2)(p_1^2 + p_4^2)(1 + p_1 w_2)(1 + p_1 w_4)]^{\frac{1}{2}}}.$$

Figure 2 shows δ as a function of w_3 for various values of w_4. We see that the coupling decreases rapidly as the waveguide separation increases or as the width of guide IV increases. The beat length L, then, will increase rapidly with w_3 or w_4 if Δ is small enough. If, however, the

Table I — Values of K_4 (and hence p_4) needed to match β_4 for the zeroth order TE mode in guide IV (of halfwidth w_4) to the β_2 for the corresponding mode in guide II

| $w_4\ (\mu\text{m})$	K_4	$p_4\ (10^4\text{ cm}^{-1})$
0.1	11.868	5.1631
0.2	11.381	3.4917
0.3	11.222	2.7340
0.4	11.145	2.2763
0.5	11.100	1.9619
0.6	11.071	1.7295
0.7	11.051	1.5494
0.8	11.037	1.4048
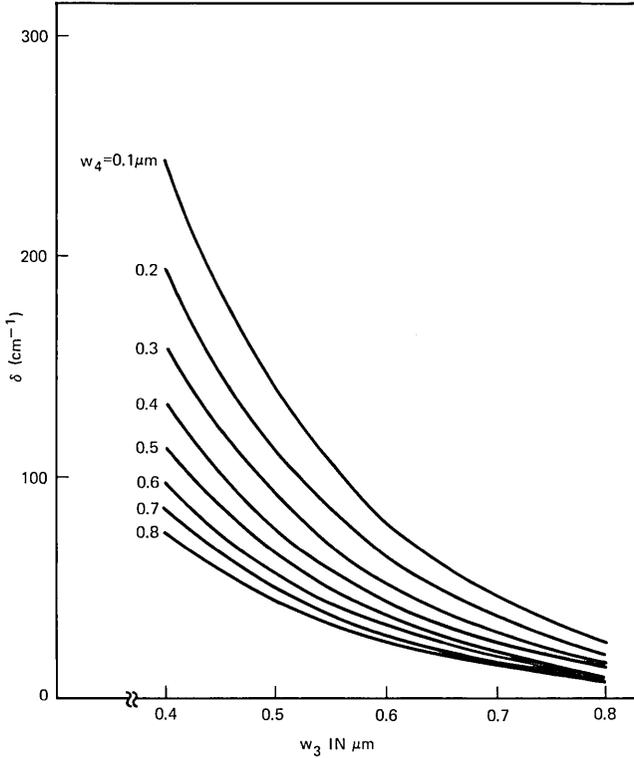
Fig. 2—The degenerate shift δ as a function of half the distance between the two waveguides w_3, with the halfwidth w_4 of guide IV as a parameter.

waveguides cannot be fabricated to match as closely as desired, the story can be different. If we take, quite arbitrarily, $\Delta = 100$ cm^{-1}, we would find for example from (31) and the data in Fig. 2 that if $w_4 = 0.6$ μm,

$$L = \left\{ \begin{array}{c} 0.113 \\ 0.137 \\ 0.150 \\ 0.157 \end{array} \right\} \text{mm} \quad \text{if} \quad w_3 = \left\{ \begin{array}{c} 0.4 \\ 0.5 \\ 0.6 \\ \infty \end{array} \right\} \mu\text{m} \qquad (\Delta = 100 \text{ cm}^{-1})$$

while if $\Delta = 0$,

$$L = \left\{ \begin{array}{c} 0.162 \\ 0.279 \\ 0.490 \end{array} \right\} \text{mm} \quad \text{if} \quad w_3 = \left\{ \begin{array}{c} 0.4 \\ 0.5 \\ 0.6 \end{array} \right\} \mu\text{m} \qquad (\Delta = 0).$$

Thus, L is reduced significantly and changes less rapidly with w_3 if Δ is large enough. Similar results are found if w_3 is fixed and w_4 varies.

Next, suppose we wish to excite one waveguide and to transfer power to the other one. If the propagation constants for the two waveguides were perfectly matched ($\Delta = 0$), then by (35) to (37) the power transfer ratio F would be the confinement factor C which is plotted in Fig. 3. The upper curve is used if guide II is excited and power is transferred to guide IV; the lower curve is used if guide IV is originally excited and power is transferred to guide II. The power transfer ratio is larger if power is transferred from a narrow guide II to a wide guide IV than if the power is being transferred the other way, since the power is more tightly confined within the guiding region in the wider waveguide.

Lest the reader become confused, we recall that F is defined as the fraction of the *total* power introduced into the system which can be transferred into the high dielectric region of the guide which was originally unexcited. F does *not* concern itself with how much power in the high dielectric region of one guide can be transferred into the high dielectric region of the other guide.
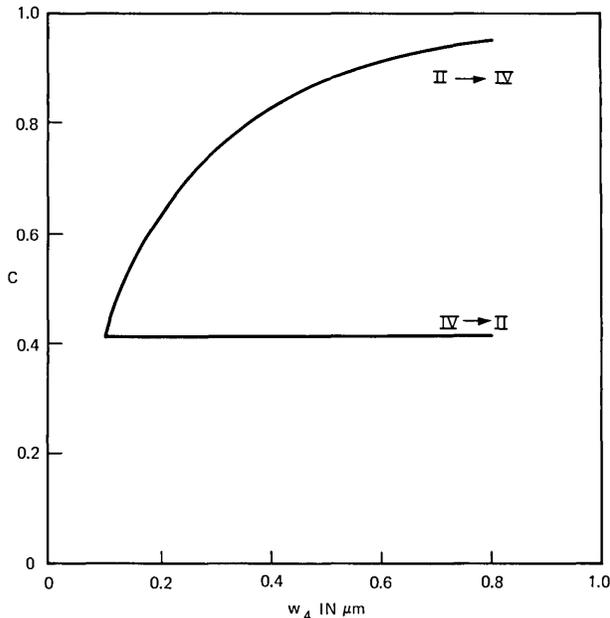

Fig. 3—The confinement factor C for degenerate modes. Guide II is fixed, and the halfwidth w_4 of guide IV varies.

Again, the mismatching Δ can have a significant effect. If, for example, we assume that $w_4 = 0.6$ μm and transfer power from guide II to guide IV, we find from (35) (evaluated with the parameters for guide IV) that if $\Delta = 100$ cm^{-1},

$$F = \begin{Bmatrix} 0.66 \\ 0.44 \\ 0.22 \end{Bmatrix} \quad \text{if} \quad w_3 = \begin{Bmatrix} 0.4 \\ 0.5 \\ 0.6 \end{Bmatrix} \mu\text{m}.$$

(The maximum value of F is 0.91 for $\Delta = 0$.)

If we know that the fabrication procedure will likely make Δ of significant size, then the only way (for given waveguide parameters) to get good power transfer is to make δ large enough. A trade-off thus must be made between good power transfer and a long beat length.

Fortunately, tuning can be a viable alternative to making such a trade-off for badly matched waveguides. To tune a device which has already been fabricated, we would need to alter one or more slab widths or dielectric heights.

Some possible methods of tuning are to change K_2 or K_4 by altering the free carrier density or using the electro-optic effect, or to change the outer slab levels K_1 or K_5 by diffusion or ion implantation. This can be achieved, in principle, by growing at least one waveguide with a small gradient in the slab width. Phase matching then can be achieved by lateral positioning of the light beam, which travels approximately perpendicular to the gradient of the slab width.

We compute a possible value of Δ for a specific example and then see how much tuning is needed to reduce Δ to zero. Suppose that (for any given w_4) the double waveguide device is fabricated according to the specifications for matched propagation constants, but that there are slight errors in w_2, w_4, $K_2 - K_1$, and $K_4 - K_1$. Assume the following errors, which are probably reasonable if the device is fabricated by molecular beam epitaxy:[5] the *ratio* w_2/w_4 is nearly constant, and w_2 varies by ± 0.02 μm; the *ratio* $(K_2 - K_1)/(K_4 - K_1)$ is nearly constant, and $K_2 - K_1$ varies by ± 0.10. We shall take K_2 to be fixed. Then the extreme cases would be given by $w_2' = 0.12$ (0.08) μm, $K_1' = 10.781$ (10.581), and

$$w_4' = (w_4/w_2)w_2',$$

$$K_4' = K_1' - \left(\frac{K_2 - K_1'}{K_2 - K_1}\right)K_1 + \left(\frac{K_2 - K_1'}{K_2 - K_1}\right)K_4,$$

where the primed parameters refer to the values in the fabricated
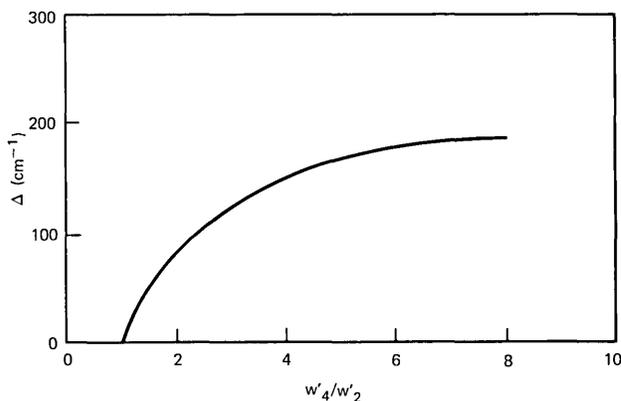
Fig. 4—The mismatching Δ as a function of the ratio of the guide widths.

device. We just treat the extreme case for which $w_2' = 0.12$ μm and $K_1' = 10.781$, since the other gives changes of comparable size.

We find from (38) that, with these fabrication errors, $\beta_2 = 1.8156 \times 10^5$ cm^{-1} and β_4 varies from 1.8156×10^5 cm^{-1} for $w_4' = w_2'$ to 1.8119×10^5 cm^{-1} for $w_4' = 8w_2'$. The resulting values of $\Delta = \frac{1}{2}|\beta_4 - \beta_2|$ are plotted in Fig. 4. Although both β_2 and β_4 have changed significantly from the value 1.8049×10^5 cm^{-1} for which the device was designed, they both change in the same direction, so Δ reflects a less radical change.

In tuning the system, suppose we consider guide IV, and hence β_4, as fixed for any given w_4; we shall alter either the dielectric height or the symmetry of guide II to adjust β_2.

If we lower K_2 to make β_2 match β_4, we find by (38) that the altered values K_2' are those given in Table II. The change in K_2 is thus less than 0.4 percent. Such a change is feasible with free carrier injection

Table II — Values of K_2 needed to tune the mismatched waveguide system

w_4'/w_2'	K_2'	% change
1	11.868	0
2	11.846	0.19
3	11.840	0.24
4	11.831	0.31
5	11.829	0.33
6	11.825	0.36
7	11.824	0.37
8	11.823	0.38

Table III — Values of K_1 needed to tune the mismatched waveguide system

w_4'/w_2'	K_1'	% change
1	10.781	0
2	10.716	0.6
3	10.697	0.8
4	10.670	1.0
5	10.663	1.1
6	10.657	1.2
7	10.647	1.2
8	10.643	1.3

and is about an order of magnitude larger than can be handled completely by the electro-optic effect.

Another possible method of tuning would be to make guide II asymmetric. If we keep $K_2 = 11.868$, $K_3 = 10.781$, $w_2' = 0.12$ μm, and alter K_1 to match β_2 to β_4, we use (9) in the form

$$p_1 = \frac{p_2 \tan 2w_2'p_2 - p_3}{1 + (p_3/p_2) \tan 2w_2'p_2}$$

to get p_1 and then use (4) to find K_1. The altered values K_1' are given in Table III. The change is thus no more than 1.3 percent. This can be handled by diffusion or ion implantation.

REFERENCES

1. N. S. Kapany, *Fiber Optics*, New York: Academic Press, 1967.
2. D. Marcuse, *Light Transmission Optics*, New York: Van Nostrand Reinhold Company, 1972.
3. S. Somekh, E. Garmire, A. Yariv, H. L. Garvin and R. G. Hunsperger, "Channel Optical Waveguide Directional Couplers," Appl. Phys. Lett., *22*, 46 (1973).
4. M. B. Panish, S. Sumski, and I. Hayashi, "Preparation of Multilayer LPE Heterostructures with Crystalline Solid Solutions of $Al_xGa_{1-x}As$: Heterostructure Lasers," Met. Trans. *2*, 795 (1971).
5. A. Y. Cho, "Film Deposition by Molecular-Beam Techniques," Vac. Sci. Techn. 531 (1971).
6. D. F. Nelson and J. McKenna, "Electromagnetic Modes of Anisotropic Dielectric Waveguides at p-n Junctions," J. Appl. Phys., *38*, 4057 (1967).
7. L. O. Wilson and F. K. Reinhart, "Phase Modulation Nonlinearity of Double Heterostructure p-n Junction Diode Light Modulators," J. Appl. Phys., to be published.
8. S. E. Miller, "Coupled Wave Theory and Waveguide Applications," B.S.T.J., *33*, No. 3 (May 1954), pp. 661-719.

Contributors to This Issue

Jacques A. Arnaud, Dipl. Ing., 1953, Ecole Supérieure d'Electricité, Paris, France; Docteur Ing., 1963, University of Paris; Docteur es Science, 1972, University of Paris; Assistant at E.S.E., 1953–1955; CSF., Centre de Recherche de Corbeville, Orsay, France, 1955–1966; Warnecke Elec. Tubes, Des Plaines, Illinois, 1966–1967; Bell Laboratories, 1967—. At CSF., Mr. Arnaud was engaged in research on high-power traveling-wave tubes and supervised a group working on noise generators. He is a supervisor, currently studying microwave quasi-optical devices and the theory of optical wave propagation. Senior Member, IEEE; Member, Optical Society of America.

J. E. Goell, B.E.E., 1962, M.S., 1963, and Ph.D. (E.E.), 1965, Cornell University; Bell Laboratories, 1965–1974. While at Cornell, Mr. Goell was a teaching assistant and held the Sloan Fellowship and the National Science Cooperative Fellowship. At Bell Laboratories, he worked on solid-state repeaters for millimeter wave communication systems and optical integrated circuits, and repeaters for optical fiber communication systems. Member, Tau Beta Pi, Eta Kappa Nu, Sigma Xi, Phi Kappa Phi, IEEE.

Joel Goldman, B.E.E., 1965, The Cooper Union; M.S., 1967, and Ph.D. (E.E.), 1970, Cornell University; Bell Laboratories, 1970—. Mr. Goldman has been concerned with the analysis of the effects of interference on various communications systems and with the foundations of estimation and probability theory. He is presently engaged in the statistical modeling of ocean noise. Member, Institute of Mathematical Statistics, Society for Industrial and Applied Mathematics, IEEE, Eta Kappa Nu, Tau Beta Pi.

Chao-Kai Liu, B.S. (E.E.), 1967, National Taiwan University; M.S. (E.E.), 1969, University of Waterloo, Ontario; Ph.D. (E.E.), 1972, University of Illinois; Bell Laboratories, 1972—. Mr. Liu has been doing exploratory network design and common control studies for customer switching systems. He is presently working on exploratory solid-state electronic switching systems. Member, IEEE.

Enrique A. J. Marcatili, Aeronautical Engineer, 1947, and E. E., 1948, University of Córdoba (Argentina); research staff, University of Córdoba, 1947–54; Bell Laboratories, 1954—. Mr Marcatili has been engaged in theory and design of filters in multimode waveguides and in waveguide systems research. More recently he has concentrated on the study of optical transmission media and circuitry for long-distance communication and for integrated optics. Fellow, IEEE.

Dietrich Marcuse, Diplom Vorpruefung, 1952, Dipl. Phys., 1954, Berlin Free University; D.E.E., 1962, Technische Hochschule, Karlsruhe, Germany; Siemens and Halske (Germany), 1954–57; Bell Laboratories, 1957—. At Siemens and Halske, Mr. Marcuse was engaged in transmission research, and studying coaxial cable and circular waveguide transmission. At Bell Laboratories, he has been engaged in studies of circular electric waveguides and work on gaseous masers. He spent one year (1966–1967) on leave of absence from Bell Laboratories at the University of Utah. He is presently working on the transmission aspect of a light communications system. Mr. Marcuse is the author of three books. Fellow, IEEE; member, Optical Society of America.

V. Ramaswamy, B.Sc. (Physics), 1957, Madras University, India; D.M.I.T. (Electronics), 1960, Madras Institute of Technology, Chromepet, Madras, India; M.S., 1962, and Ph.D., 1969, Northwestern University; Zenith Radio Corporation, 1962–1965; Bell Laboratories, 1969—. Mr. Ramaswamy's work has included microwave components, diode parametric amplifiers, and wave propagation in semiconductor plasmas. At Bell Laboratories, he is engaged in research related to thin film optical waveguides and devices. Member, Sigma Xi, IEEE.

F. K. Reinhart, Dipl. El. Eng., 1958, Dr. Sc. Techn., 1962, Swiss Federal Institute of Technology, Zurich, Switzerland; Bell Laboratories, 1963—. Mr. Reinhart has been engaged in device feasibility studies. Among his special interests are guided optical wave interactions in semiconductors, light modulation, and injection lasers.

Carlo H. Séquin, Ph.D. (Physics), 1969, University of Basel, Switzerland; Bell Laboratories, 1970—. Since joining Bell Laboratories, Mr. Séquin has been engaged in the design and investigation of charge-coupled image sensors. Member, Swiss Physical Society, American Physical Society, IEEE.

Tse Lin Wang, B.S.E.E., 1964, National Taiwan University; M.S., 1969, and Ph.D., 1970, University of South Carolina; Bell Laboratories, 1970—. Mr. Wang has been doing exploratory network design and common control studies for customer switching systems. He is presently working on exploratory key telephone systems. Member, IEEE, Eta Kappa Nu.

Lynn O. Wilson, A.B. (Physics), 1965, Oberlin College; Ph.D. (Applied Mathematics), 1970, University of Wisconsin; Bell Laboratories, 1970—. Ms. Wilson has pursued research in various areas of applied mathematics. She has worked on problems concerning *Picturephone* ® demand, electromagnetic theory, elastic surface waves, dielectric waveguides, and operations research. Member, Sigma Xi, American Physical Society, SIAM.

Bell System