THE BELL SYSTEM TECHNICAL JOURNAL

Impact of Microwave Depolarization During Multipath Fading on Digital Radio Performance

By S. H. LIN

Experimental data describing the statistics of microwave depolarization during multipath fading have been obtained from a propagation experiment conducted near Atlanta, Georgia. The experiment included 6- and 11-GHz reception on a 26.4-mile path, and 11-GHz reception on a 15.9-mile path. A theoretical model, suggested by T. O. Mottl, indicates that the interference occasioned by depolarization for a given copolarized signal level is Rice-Nakagami distributed. The theoretically calculated distribution agrees well with the data. The cross-polarization interference consists of a signal-level-dependent component as well as a residual that is independent of the in-line signal level. The residual is Rayleigh distributed with an rms value about 40 dB below the non-faded in-line signal level, and limits the multipath fade margin of a cochannel dual-polarized digital radio to approximately 30 dB. Calculated multipath outage probabilities for cochannel, dual-polarized, 11-GHz, quaternary-coherent-phase-shift-keyed digital radios with and without space-diversity protection are presented.

I. INTRODUCTION

Maintaining adequate cross-polarization discrimination (XPD) is important to both analog and digital radio transmission systems. For analog radio and single polarization digital radio, adequate XPD allows reduction of frequency separation between cross-polarized, adjacent channels to increase the transmission capacity. This is known as inter-
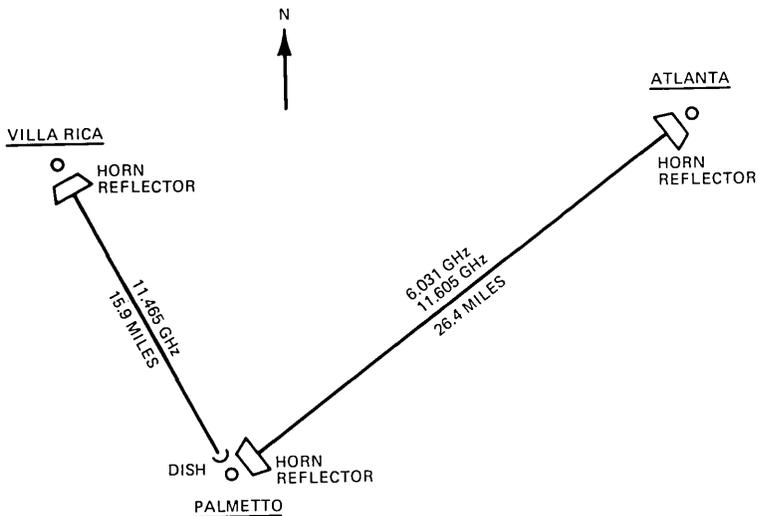
Fig. 1—Path layout, frequencies, and antennas for microwave propagation experiments near Atlanta, Georgia.

stitial operation in channel allocation of analog radio. Some digital radio may also rely on XPD to achieve high transmission capacity by a two-channel-per-frequency allocation in which both orthogonal linear polarizations in the same frequency band are employed as two independent transmission paths.

References 1 to 5 indicate that XPD can degrade significantly during multipath fading. Therefore, statistics of XPD during multipath fading are needed to assess the performance and reliability of both analog and digital radio. Section II of this paper describes a propagation experiment, the measured data, and a theoretical model for microwave depolarization during multipath fading. Section III applies these results to calculate the multipath-caused outages of dual-polarization 11-GHz quaternary-coherent-phase-shift-keyed (QCPSK) digital radio. A companion paper[6] by T. O. Mottl gives greater details on outage probabilities of QCPSK digital radio.

II. MICROWAVE DEPOLARIZATION DURING MULTIPATH FADING

2.1 Introduction

Section 2.2 describes microwave propagation experiments; Section 2.3 presents multipath fading and associated depolarization statistics including cumulative amplitude distributions, number of fades, and average fade durations. Section 2.4 discusses a theoretical model, suggested by T. O. Mottl,[6] to describe the behavior of XPD during multipath fading.

Table I — Path parameters of microwave propagation experiments near Atlanta, Georgia

Path	Path Length Miles	Km	Freq. (GHz)	(v_{xp}) rms when $V_{IL} = 0$ dB (dB)	K (dB)	ϵ_{rms} (dB)	Transmitter Polarization	r
Atlanta-Palmetto	26.4	42.5	6.031	−25.9	−25.9	−47.6	Vertical	0.13
Atlanta-Palmetto	26.4	42.5	11.605	−19.5	−19.6	−36.3	Vertical	0.4
Villa Rica-Palmetto	15.9	25.6	11.465	−42.7	−46.2	−45.3	Vertical	0.015

r = Multipath fade occurrence factor defined in eqs. (32) and (33).
$v_{xp} = |kv_{IL} + \epsilon e^{j\phi}|$

2.2 Microwave propagation experiments

Figure 1 displays the path layout for the Palmetto propagation experiment. Two vertically polarized signals with frequencies of 6.031 and 11.605 GHz are transmitted over the 26.4-mile path from Atlanta to Palmetto, and one vertically polarized signal with a frequency of 11.465 GHz is transmitted over the 15.9-mile path from Villa Rica to Palmetto. At Palmetto, the common receiving site, the levels of both vertically and horizontally polarized received signals are recorded.

Both the transmitter and the receiver on the Atlanta-Palmetto path employ standard Bell System horn reflector antennas, waveguides, and channel separation networks.[7,8] Circular waveguide (WC281) simultaneously supports both polarizations of the 6- and 11-GHz signals. On the Villa Rica-Palmetto path, a standard Bell System horn reflector antenna, waveguide, and network are used to transmit the 11-GHz signal. The 11-GHz receiver at Palmetto utilizes a 6-foot dish antenna with two elliptical waveguides to separately carry two orthogonally polarized received signals.

The measured cross-polarization discrimination (XPD) obtaining on these paths during nonfading periods is listed in Table I. The best performance is 42.7 dB on the 11-GHz Villa Rica-Palmetto path; the worst, 19.5 dB on the 11-GHz Atlanta path. Poor XPD on the 11-GHz Atlanta-Palmetto path is believed due to the 4- and 6-GHz channel-separation networks at both transmitting and receiving ends, as well as the fact that this link requires quite long waveguides (see Table II).* The WC281 waveguide is an overmoded guide at the 11-GHz frequency, which supports 21 higher-order modes in addition to the desired fundamental.[9] Slight imperfections on this long (see Table I) waveguide run can cause mode coupling with resultant depolarization.[10] The imper-

---

* The high fill of radio traffic on this link limits opportunity to study the hardware impact on 11-GHz XPD.

Table II — Waveguide types and lengths

Station	Waveguide Length (ft)	Waveguide Type
	A. Atlanta-Palmetto Path	
Atlanta	75	EW-107
	15	WR-90
	20	WC-281
Palmetto	54	EW-107
	5	WR-90 Flex-guide
	300	WC-281
	B. Villa Rica-Palmetto Path	
Villa Rica	85	Elliptical
	21	WR-90
	170	WC-281
Palmetto	80	Elliptical

fections in antennas, antenna misalignment, and channel-separation networks also contribute to depolarization.

In the following, the in-line signal refers to the received (vertically polarized) signal which is "in-line" with the transmitted signal; the cross-polarization interference refers to the received horizontally polarized signal which is orthogonal to it.

2.3 Experimental Data

The experimental data obtained during the 6.5-month period from August 15, 1974 to February 28, 1975 have been processed.

2.3.1 Statistics of in-line signals

The measured statistics of multipath fading of the in-line signals are shown in Figs. 2, 3, and 4 for the cumulative amplitude distribution, the number of fades, and the average fade durations, respectively, as functions of fade depth.

In the deep-fade region ($\geq$20 dB), the slopes of the distributions in Figs. 2, 3, and 4 are consistent with the theoretical distribution for deep fades.[11-14] The cumulative amplitude distribution has an inverse slope of 10 dB per decade of probability, the number of fades has an inverse slope of 20 dB per decade of number of fades, and the average fade duration has an inverse slope of 20 dB per decade of duration.

2.3.2 Statistics of cross-polarization interference

Let $v_{IL}(t)$ and $v_{xp}(t)$ be the time varying amplitudes of the in-line signal voltage and the depolarized (interference) voltage, respectively, both normalized to the nonfaded in-line signal voltage. The cross-polarization discrimination (XPD) can be written as

$$\text{XPD} = 20 \log_{10}(v_{IL}/v_{xp}), \text{ dB} \tag{1}$$
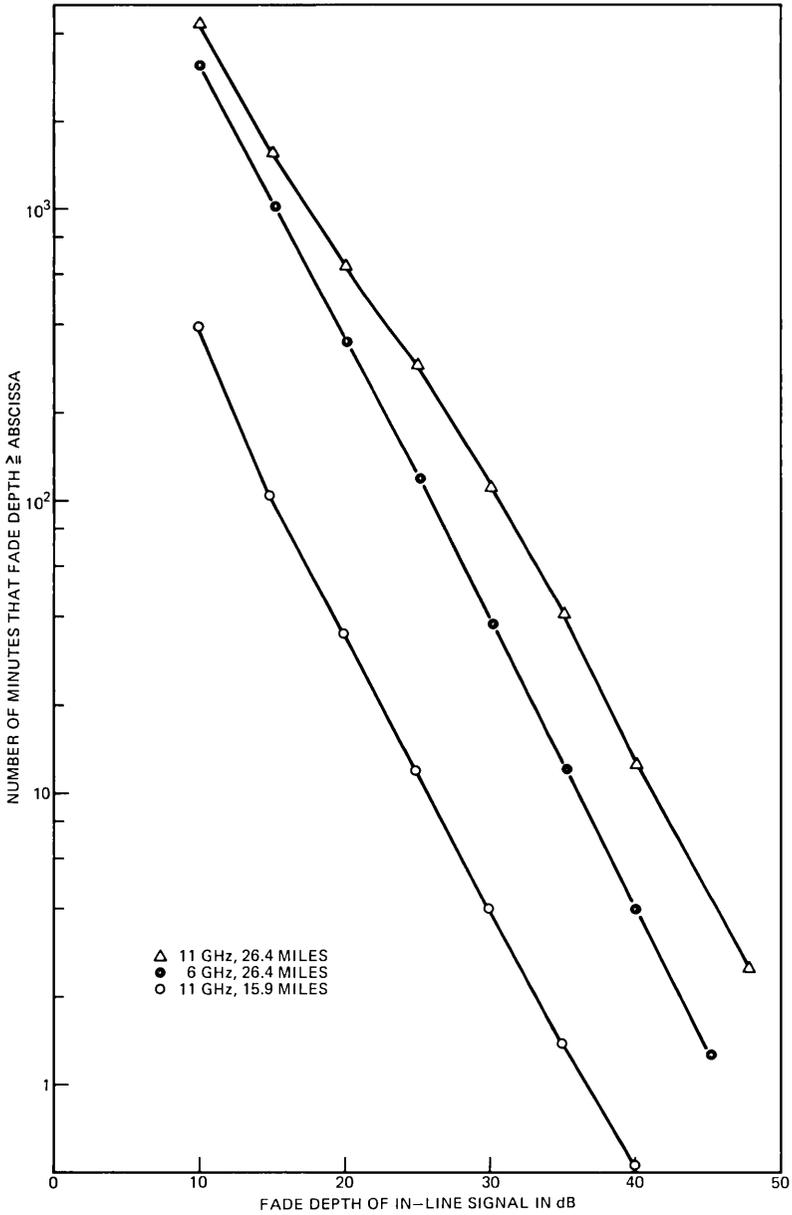
Fig. 2—Measured 6.5-month data (August 15, 1974 to February 28, 1975) on the cumulative amplitude distributions of in-line signals during multipath fading.

$$= V_{\mathrm{IL}} - V_{xp}, \tag{2}$$

where

$$V_{\mathrm{IL}} \equiv 20 \log_{10} v_{\mathrm{IL}} \ \mathrm{dB} \tag{3}$$
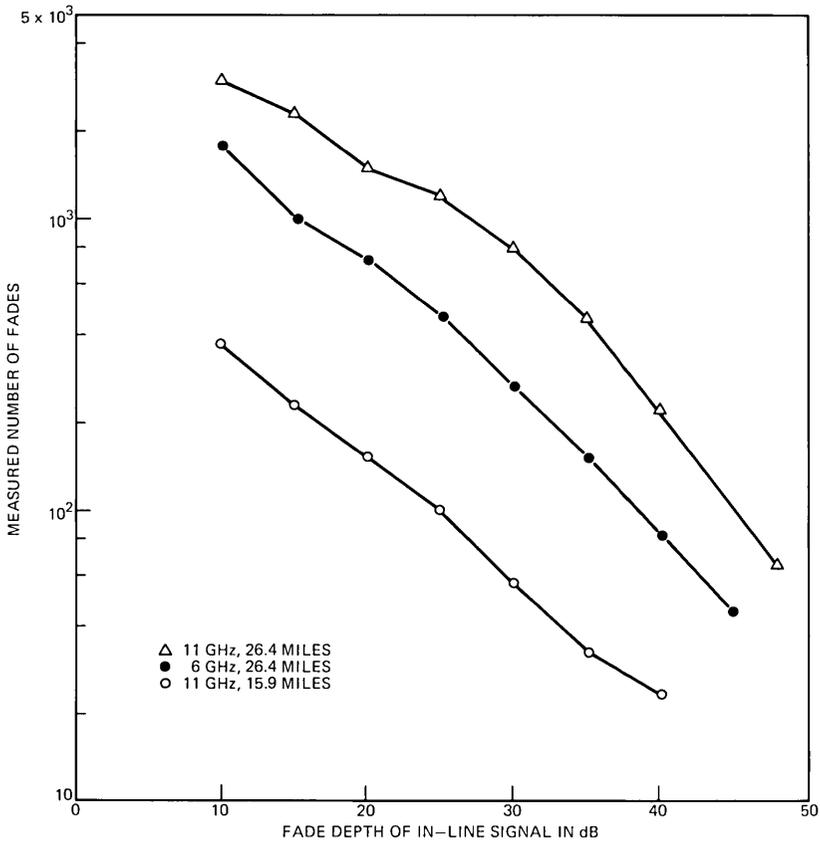
Fig. 3—Measured 6.5-month data (August 15, 1974 to February 28, 1975) on number of fades of in-line signals during multipath fading.

is the time-varying, in-line signal level in dB with respect to its nonfaded level, and

$$V_{xp} \equiv 20 \log_{10} v_{xp} \text{ dB} \tag{4}$$

is the time-varying, depolarized component of the signal measured (in dB) with respect to the nonfaded level of in-line signal.

Figure 5 shows fluctuations of V_{IL} and V_{xp} measured on the 11-GHz, Atlanta-Palmetto path during a typical nonfading hour. Both V_{IL} and V_{xp} scintillate but approximately 20 dB of XPD is maintained. Figure 6 shows the behavior of V_{IL} and V_{xp} during a multipath fading hour. In the 40-minute period from 8:00 a.m. to 8:40 a.m., the variations of V_{IL} and V_{xp} appear to be relatively well correlated. However, at 8:44 a.m., V_{IL} suffers a 35-dB fade, whereas V_{xp} undergoes only a 15-dB fade (from -20 dB to -35 dB), thereby degrading the XPD to 0 dB. In other words, at this moment, the signal received at the cross-polarized feed (inter-
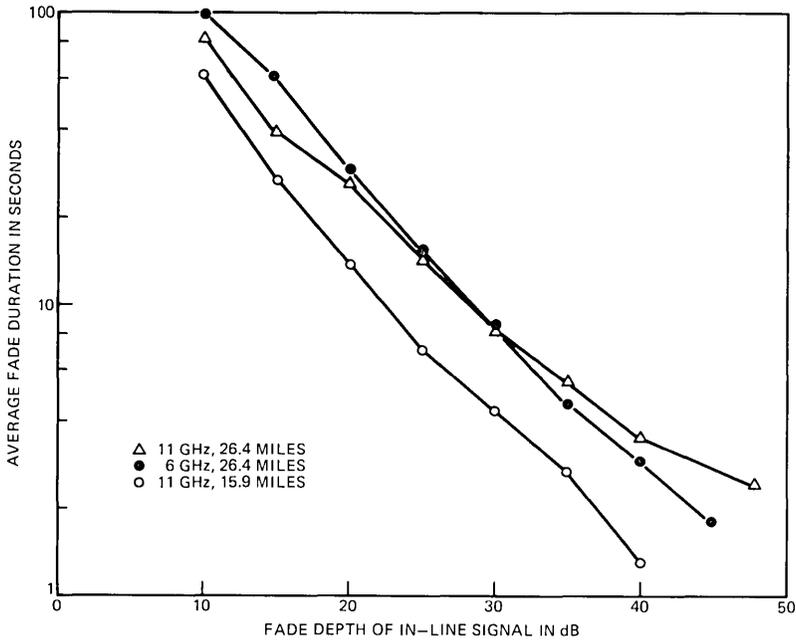
Fig. 4—Measured 6.5-month data (August 15, 1974 to February 28, 1975) on average fade durations of in-line signals during multipath fading.

ference to a signal normally transmitted with horizontal polarization that has faded with the same amount as V_{IL}) is as strong as the desired horizontally polarized in-line signal and will prevent useful transmission of data over the channel; hence, an outage is caused to a 2-channel-per-frequency-assignment digital radio. Figure 7 shows another example of depolarizing fading on the 11-GHz, Villa Rica-Palmetto path.

Although rain also causes depolarization, the measured XPD is generally better than 10 dB, even during rain fades in excess of 40 dB. Depolarization during multipath is therefore considerably more serious than rain-caused depolarization in Western U.S.A., where rain-caused outage is not the dominant controlling factor on radio-system reliability.

Figures 8 through 10 display the rms value of v_{xp} conditioned to the fade depth of the in-line signal v_{IL}. In the shallow fade region (i.e., $V_{IL} \geqq -10$ dB), $(V_{xp})_{rms}$ decreases almost linearly (dB by dB) with V_{IL}; whereas in the deep-fade region (i.e., $V_{IL} < -20$ dB), $(V_{xp})_{rms}$ approaches a residual level, becoming independent of V_{IL} as V_{IL} decreases. The residual depolarized component in these three sets are between 37 and 50 dB below the nonfaded in-line signal level.

Figures 11 through 13 show the probability distributions of V_{xp}, conditioned to a given fade depth of in-line signal, plotted on Rayleigh
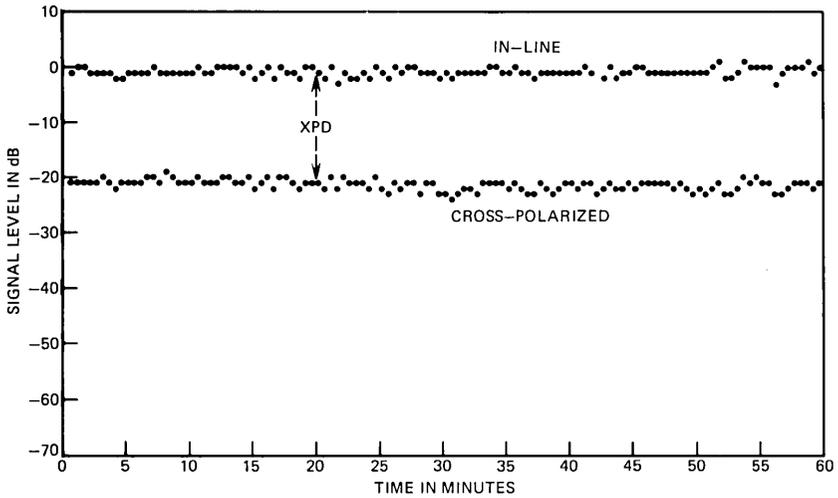
Fig. 5—Scintillations of in-line signal and cross-polarization interference on the 11-GHz 26.4-mile path (Atlanta-Palmetto) during a nonfading hour—August 28, 1974; 12:00 A.M. to 1:00 P.M.

probability paper. For deep fades of in-line signal (i.e., $V_{IL} \leqq -30$ dB), the probability distributions of V_{xp} are approximately Rayleigh. This Rayleigh distribution of V_{xp}, conditioned to deep fades of in-line signal, was reported earlier in Ref. 1.

The solid lines in Figs. 8 through 13 are calculated results from a theoretical model discussed in the next section.
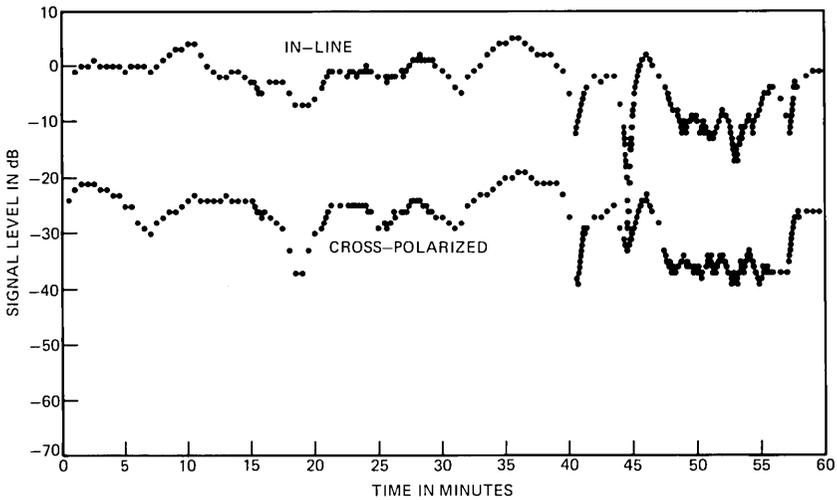
Fig. 6—Fading behavior of in-line signal and cross-polarization interference on the 11-GHz 26.4-mile path (Atlanta-Palmetto) during a multipath fading hour—August 23, 1974; 8:00 A.M. to 9:00 A.M.
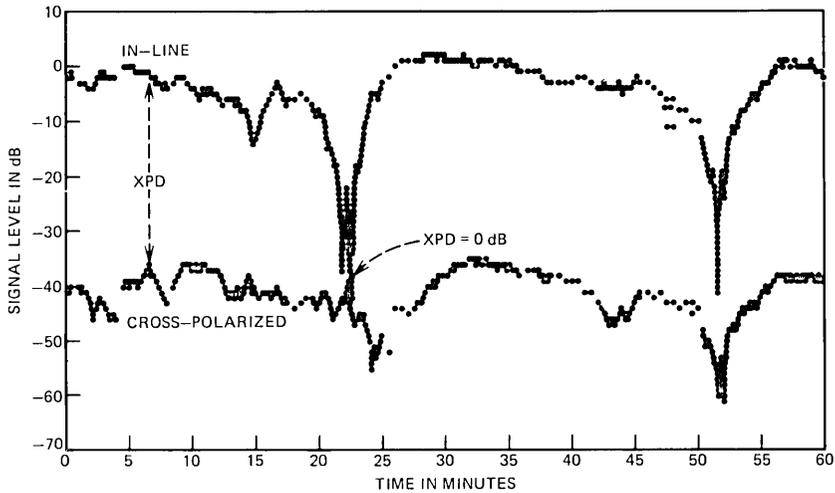
Fig. 7—Fading behavior of in-line signal and cross-polarization interference on the 11-GHz 15.9-mile path (Villa Rica-Palmetto) during a multipath fading hour—November 2, 1974; 9:00 A.M. to 10:00 A.M.

2.4 Theoretical model

Based on early observations similar to those displayed in Figs. 6 through 10, Mottl[6] suggests decomposition of the cross-polarization interference v_{xp} into two components:

$$v_{xp} \equiv |kv_{\text{IL}} + \epsilon e^{j\phi}|, \tag{5}$$

where k is a proportionality constant, and ϕ is the relative phase between the proportionate component and the residual component.

For shallow fades of the in-line signal, v_{xp} is dominated by kv_{IL}, and decreases linearly with v_{IL}. During deep fades of the in-line signal, v_{xp} is dominated by ϵ, which is independent of v_{IL}. During deep fades of in-line signal, the probability distribution of v_{xp} is essentially that of ϵ. The data in Figs. 11 through 13 indicate the conditional distribution of v_{xp}. It is seen that ϵ is approximately Rayleigh, a single parameter distribution uniquely determined by its rms value. The rms values, ϵ_{rms}, are obtained from the measured distribution of v_{xp} during deep fades of v_{IL} and they range from -36 to -48 dB, as listed in Table I.

For a given in-line signal level, eq. (5) indicates that v_{xp} consists of a constant vector [i.e., kv_{IL}] plus a Rayleigh vector. Such interpretation immediately implies that, for a given in-line signal level, the cross-polarization interference, v_{xp}, is Rice-Nakagami distributed.[15-17] In other words, the conditional distribution of v_{xp} can be written as

$$P(v_{xp} \geq a | v_{\text{IL}}) = \int_a^\infty p(v_{xp} | v_{\text{IL}}) dv_{xp}, \tag{6}$$
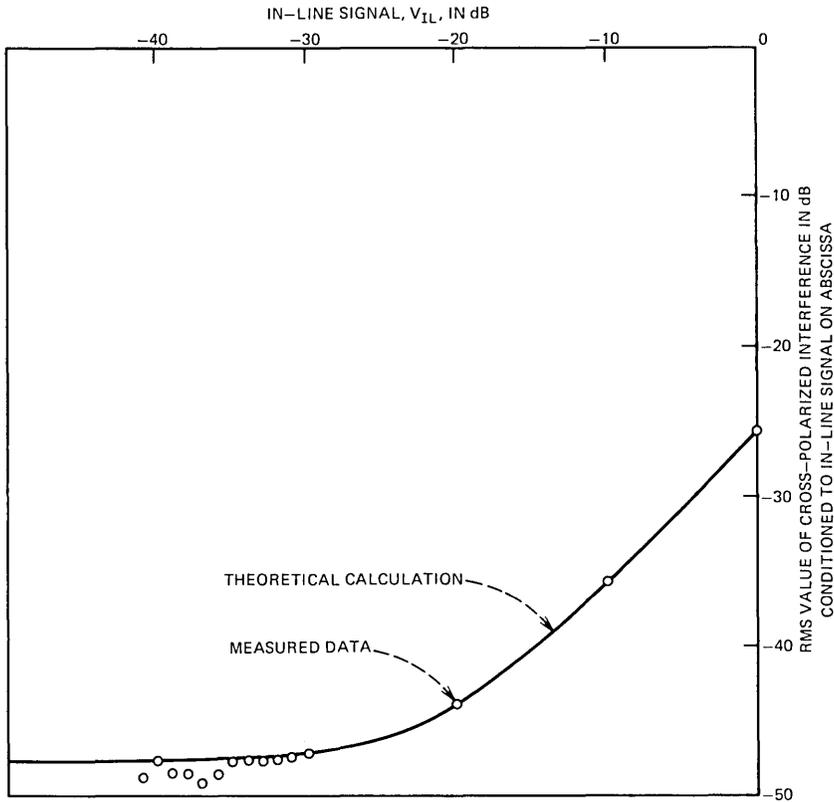
Fig. 8—Dependence of rms value of cross-polarization interference on in-line signal level measured on the 6-GHz 26.4-mile path (Atlanta-Palmetto); measured 6.5-month data (August 15, 1974 to February 28, 1975).

where

$$p(v_{xp}|v_{IL}) = \left[\frac{2v_{xp}}{\epsilon_{rms}^2} \exp - \left(\frac{v_{xp}^2 + k^2 v_{IL}^2}{\epsilon_{rms}^2}\right)\right] \left[I_0 \left(\frac{2v_{xp}k v_{IL}}{\epsilon_{rms}^2}\right)\right] \quad (7)$$

and $I_0(\sim)$ denotes the modified Bessel function of zeroth order. Based on the well-known properties of this Rice-Nakagami distribution,[15-17] it is easily shown that

$$[v_{xp}(v_{IL})]_{rms} = [k^2 v_{IL}^2 + \epsilon_{rms}^2]^{1/2} \quad (8)$$

from which,

$$k = \frac{1}{v_{IL}} \{[v_{xp}(v_{IL})]_{rms}^2 - \epsilon_{rms}^2\}^{1/2} \quad (9)$$

for any v_{IL}, where $v_{xp}(v_{IL})$ denotes v_{xp} conditioned to a given v_{IL}. This equation shows that the proportionality constant k in eq. (5) can be
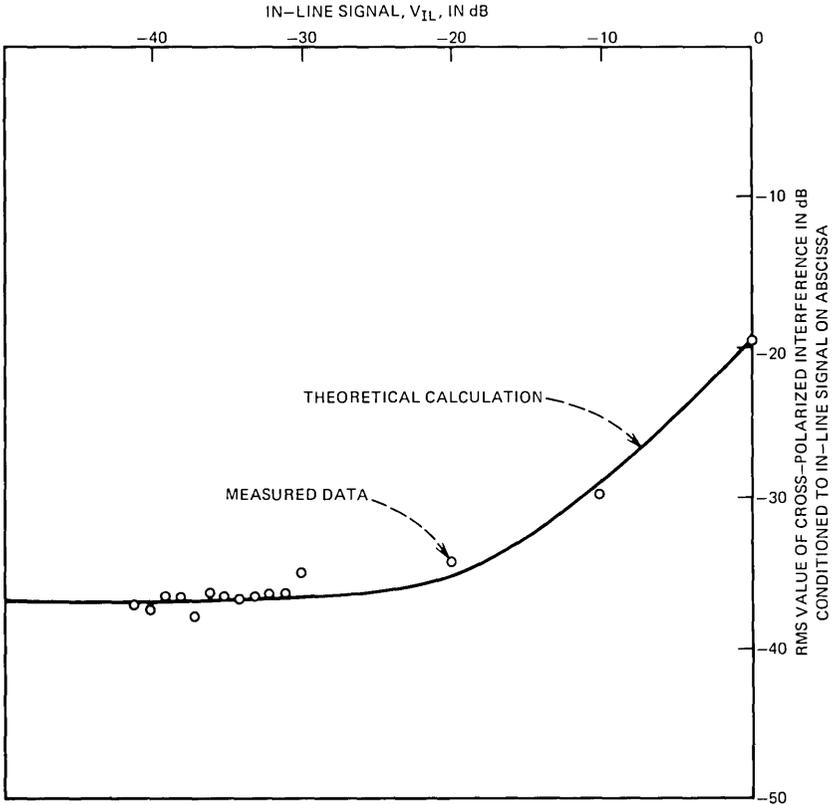
Fig. 9—Dependence of rms value of cross-polarization interference on in-line signal measured on the 11-GHz 26.4-mile path (Atlanta-Palmetto); measured 6.5-month data (August 15, 1974 to February 28, 1975).

calculated from measured values of $[v_{xp}(v_{IL})]_{rms}$ and ϵ_{rms}. For example, the data in Fig. 9 for the 11-GHz Atlanta-Palmetto path indicate

$$[v_{xp}(v_{IL} = 0 \text{ dB})]_{rms} = -19.5 \text{ dB} \tag{10}$$

and

$$\epsilon_{rms} = [v_{xp}(v_{IL} \leqq 30 \text{ dB})]_{rms} = -36.3 \text{ dB}. \tag{11}$$

Substituting (10) and (11) into (9) yields

$$k = 0.105 \tag{12}$$

from which

$$K = 20 \log_{10} k = -19.6 \text{ dB}. \tag{13}$$

Table I lists estimated values of K for all three sets of data.

The Rice-Nakagami distribution (7) for v_{xp} is completely determined
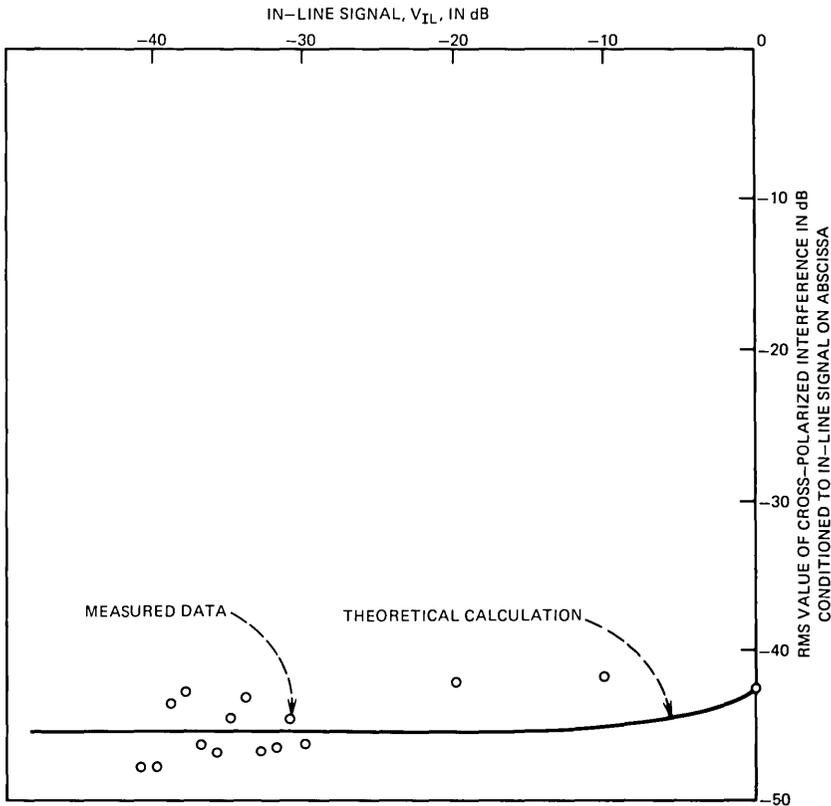
Fig. 10—Dependence of rms value of cross-polarization interference on in-line signal level measured on the 11-GHz 15.9-mile path (Villa Rica-Palmetto); measured 6.5-month data (August 15, 1974 to February 28, 1975).

by two parameters: ϵ_{rms} and k. Since these two parameters are obtained from the measured data, the Rice-Nakagami distribution of v_{xp} for any v_{IL} can be calculated. Figures 11 through 13 show that the theoretical distributions (solid lines) calculated from the two parameters (ϵ_{rms} and k) agree very well with the measured data.

Similarly, eq. (8) indicates that $[v_{xp}(v_{IL})]_{rms}$ for any v_{IL} is also completely determined by the same parameters: ϵ_{rms} and k. Figures 8 through 10 show that the theoretical dependence of $[v_{xp}(v_{IL})]_{rms}$ calculated by eq. (8) also agrees closely with the data.

At the present time, the physical interpretation of the theoretical model (5) is speculative. The proportionality constant k may be related to antenna alignment, imperfections in the antennas, the waveguides, or the channel-separation networks (designed for transmission of the fundamental mode). In other words, the proportional component, kv_{IL}, may be controllable by reducing these imperfections. On the other hand,
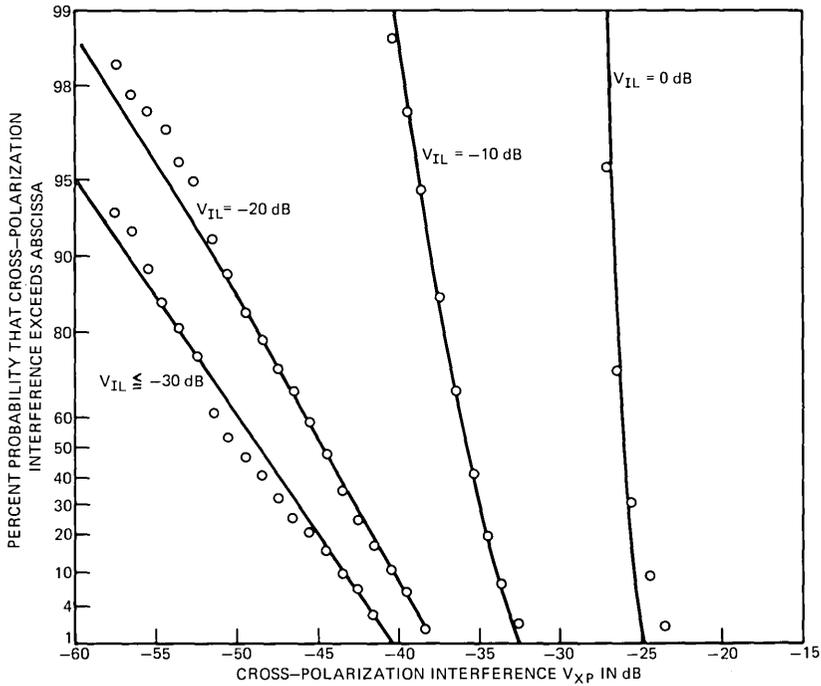
Fig. 11—Rice-Nakagami distribution (with Rayleigh probability coordinates) of cross-polarization interference, V_{xp}, conditioned to a given in-line signal level, V_{IL}, and measured on the 6-GHz 26.4-mile path (Atlanta-Palmetto). The circles are measured 6.5-month data (August 15, 1974 to February 28, 1975).

the Rayleigh distribution of the residual cross-polarized component, ϵ, suggests that ϵ may represent the sum of many small depolarized components due to, say, foreground scattering,* the antenna cross-polarization response to off-axis incoming rays, or the excitation of higher-order modes in the waveguides by the off-axis incoming rays.†

For successful isolation of the two information channels of a dual-polarized digital radio system, it is reasonable to expect that

$$K = 20 \log_{10} k \leq -20 \text{ dB}. \tag{14}$$

Since the correlated component, $k v_{IL}$, of cross-polarization interference always fades simultaneously with the desired signal and maintains a carrier-to-interference ratio (CIR) of 20 dB or better, the degradation of radio fade margin due to the correlated components, $k v_{IL}$, is quite

---

* In the multipath propagation condition, the relative amplitudes and phases of the foreground scattered components are quite different from those of direct paths. Therefore, the sum of foreground scattered components is decorrelated from the sum of the direct paths.

† Waveguide imperfections may couple depolarized components through higher-order modes. Since these are dispersed[18,19] in the circular waveguide and antenna responses to off-axis incoming rays are mode-dependent, their sum is decorrelated from v_{IL}.

MICROWAVE DEPOLARIZATION   657
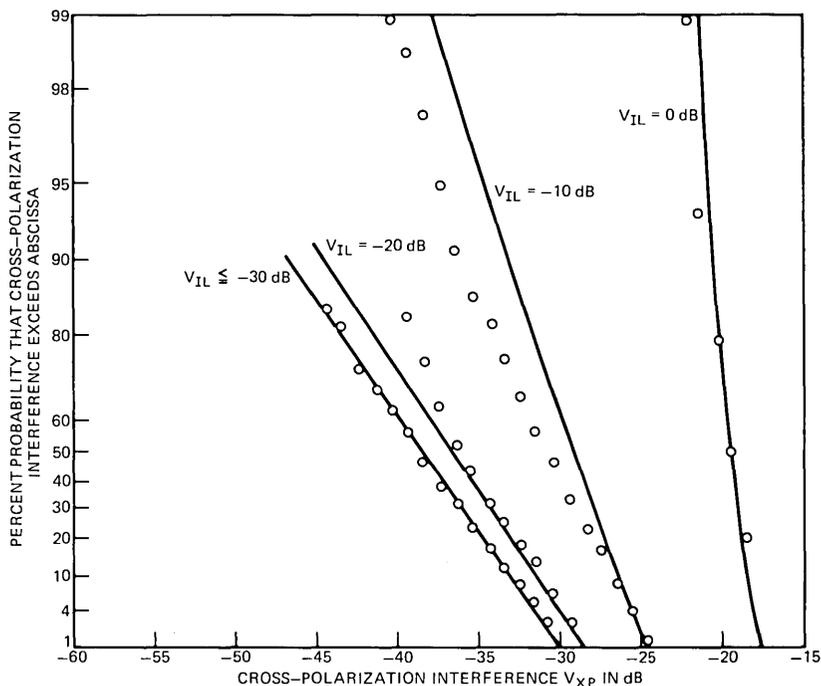
Fig. 12—Rice-Nakagami distribution (with Rayleigh probability coordinates) of cross-polarization interference, V_{xp}, conditioned to a given in-line signal level, V_{IL}, and measured on the 11-GHz 26.4-mile path (Atlanta-Palmetto). The circles are measured 6.5-month data (August 15, 1974 to February 28, 1975).

small, say 2 dB or less. On the other hand, the residual cross-polarization interference ϵ, being independent of the desired signal, will limit fade margin. For example, for a typical 11-GHz digital radio, the thermal noise is about 60 dB below the nonfaded signal, whereas ϵ_{rms} is only about 40 dB below the nonfaded signal (see Table I). Therefore, the residual cross-polarization interference may reduce the fade margin by as much as 20 dB and greatly increase the "multipath caused" outages.

III. MULTIPATH OUTAGES OF DUAL-POLARIZATION, 11-GHz, PHASE-SHIFT-KEYED DIGITAL RADIO

3.1 Introduction

Section 3.2 presents a model for the performance of quaternary-coherent-phase-shift-keyed (QCPSK) digital radio subjected to interference and noise. Section 3.3 outlines the procedure and states the assumptions needed for calculation of multipath-caused outage probabilities of dual-polarization, 11-GHz, QCPSK digital radio. Section 3.4 summarizes the results of outage calculations.
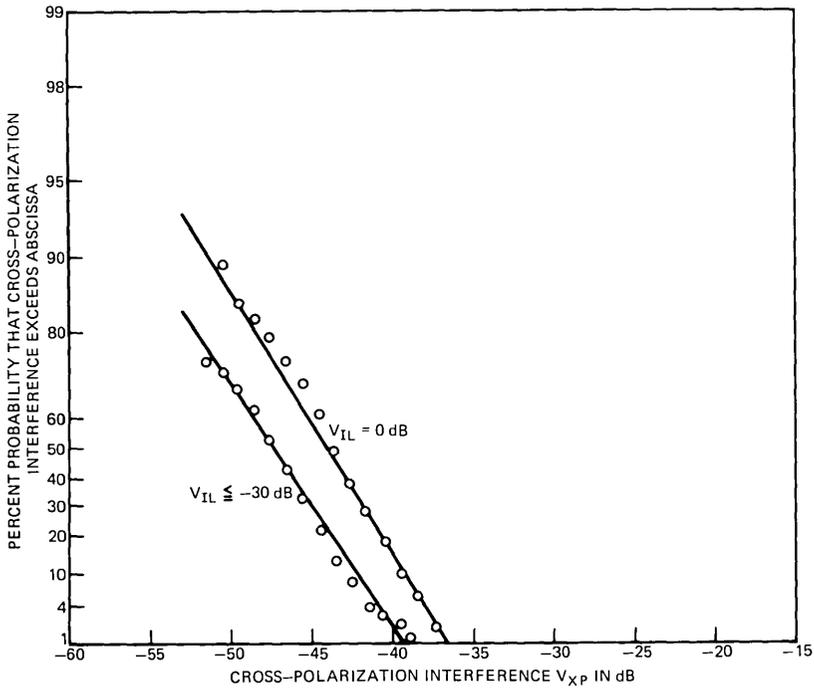
Fig. 13—Rice-Nakagami distribution (with Rayleigh probability coordinates) of cross-polarization interference, V_{xp}, conditioned to a given in-line signal level, V_{IL}, measured on the 11-GHz 15.9-mile path (Villa Rica-Palmetto). The circles are measured 6.5-month data (August 15, 1974 to February 28, 1975).

Rain-caused outage probabilities of 11-GHz radio is treated elsewhere.[20,21] This paper, therefore, treats only multipath-caused outages.

3.2 Digital radio model

We model the cochannel, dual-polarization, 11-GHz digital radio as a QCPSK system corrupted by complex gaussian noise with one major interference representing the depolarized component of the cochannel cross-polarized signal.

Both noise and interference cause digital transmission errors. Figure 14 shows the relationship between the carrier-to-noise ratio (CNR) and carrier-to-interference ratio (CIR) for a fixed bit-error-rate (BER). In Fig. 14 the circles represent the measured performance[22] of a prototype QCPSK digital radio; the solid lines are approximations described by

$$10^{-\frac{CNR}{10}} + 10^{-\frac{CIR + \Delta}{10}} = 10^{-\frac{CNR_a}{10}}, \tag{15}$$
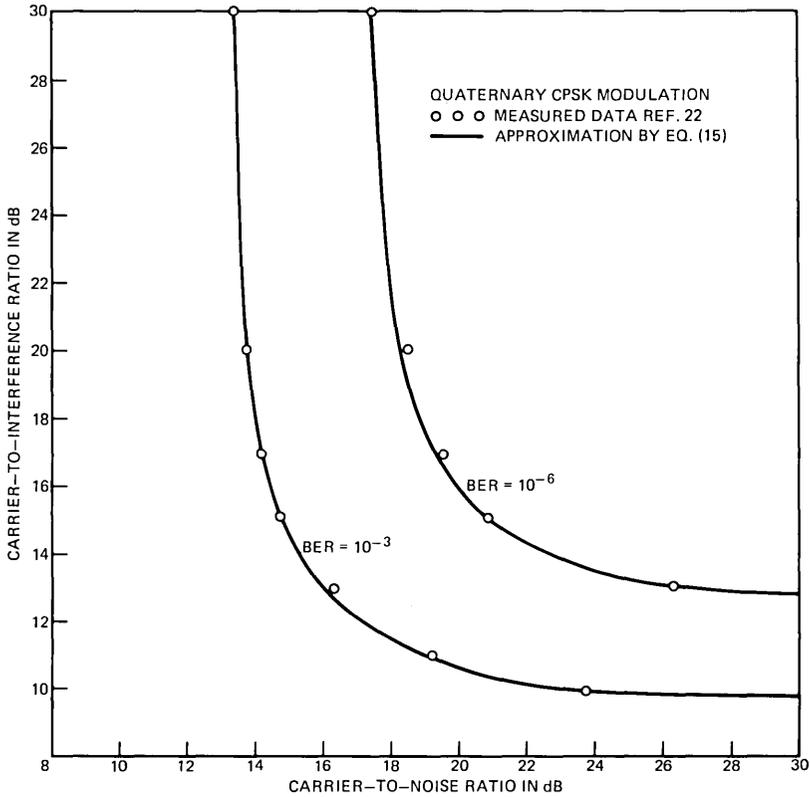
Fig. 14—Relationship between carrier-to-interference ratio and carrier-to- noise ratio for a given bit-error-rate (BER) of QCPSK radio.

or equivalently,

$$\text{CIR} = -10 \log_{10} \left\{ 10^{-\frac{\text{CNR}_a}{10}} - 10^{-\frac{\text{CNR}}{10}} \right\} - \Delta, \qquad (16)$$

where

$$\text{CNR}_a = \begin{cases} 13.4 \text{ dB for BER} = 10^{-3} & (17) \\ 17.4 \text{ dB for BER} = 10^{-6} & (18) \end{cases}$$

and

$$\Delta = \begin{cases} 3.8 \text{ dB for BER} = 10^{-3} & (19) \\ 4.8 \text{ dB for BER} = 10^{-6} & (20) \end{cases}$$

Equation (15) means that for a given BER threshold, the interference power reduced by Δ dB plus noise power is a constant. In other words,

the QCPSK radio system is more resistant to interference than to noise by Δ dB.

The probability of outage of the QCPSK radio for a given BER threshold is simply the integral of the joint CNR, CIR probability density function over the two-dimensional region lying below the BER threshold curve in Fig. 14. This double integration is carried out by eqs. (27) and (31) in the next section.

3.3 Outline of the outage estimation procedure

(a) The nonfaded carrier-to-noise ratio, CNR_{NF}, for a reference 25-mile hop length is assumed to be 67.4 dB.[22]

(b) The dependence of CNR_{NF} on hop length D is

$$CNR_{NF}(D) = CNR_{NF}(D_0) - 20 \log \frac{D}{D_0}, \qquad (21)$$

where

$$D_0 = 25 \text{ miles}. \qquad (22)$$

(c) With the desired, in-line signal faded α dB, the effective $CNR(D)$ for path length D is:

$$CNR(D) = CNR_{NF}(D) - \alpha, \qquad (23)$$

where

$$\alpha = -20 \log_{10} v_{IL}, \text{ dB}. \qquad (24)$$

(d) For a given BER (say 10^{-3}) at outage threshold, the CNR versus CIR curve in Fig. 14 means that the cross-polarization interference, v_{xpo}, at outage threshold is a function of in-line signal level, v_{IL}. By combining eqs. (15), (23), and (24), it can be shown that

$$v_{xpo}(v_{IL}) = v_{IL} \left\{ 10^{-\dfrac{CNR_a}{10}} - 10^{-\dfrac{CNR_{NF} - \alpha}{10}} \right\}^{1/2} 10^{\dfrac{\Delta}{20}}. \qquad (25)$$

(e) For a given in-line signal level, the probability of outage is

$$P(\text{outage}|v_{IL}) = P(v_{xp} \geq v_{xpo}|v_{IL}) \qquad (26)$$

$$= \int_{v_{xpo}}^{\infty} p(v_{xp}|v_{IL}) dv_{xp}, \qquad (27)$$

where $p(v_{xp}|v_{IL})$ is given by eq. (7).

(f) The distribution of the in-line signal, v_{IL}, without diversity protection is assumed to be Rayleigh, i.e.,

$$p(v_{IL}) = 2v_{IL} e^{-v_{IL}^2}. \qquad (28)$$

(g) The distribution of the in-line signal, v_{IL}, with space diversity and selective switching is[23,24]

$$p(v_{\text{IL}}) = \frac{8v_{\text{IL}}}{(1 - \rho^2)} \int_0^{v_{\text{IL}}} z \left[I_0 \left(\frac{2\rho z v_{\text{IL}}}{1 - \rho^2} \right) \right] \exp - \left(\frac{z^2 + v_{\text{IL}}^2}{1 - \rho^2} \right) dz,$$

(29)

where ρ^2 is approximately equal to the correlation coefficient between the two input signals received from the space-diversity antenna pair.[25] The dependence of ρ^2 on antenna spacing S is obtained empirically as[26,27]

$$\rho^2 = 1 - 7 \times 10^{-5} \frac{S^2 f}{D},$$

(30)

where

$\quad S \;=\;$ antenna center-to-center separation in feet,
$\quad f \;=\;$ radio frequency in GHz,
$\quad D \;=\;$ path length in miles.

Strictly speaking, assumptions (f) and (g) are valid only for the deep-fade region, but not the shallow-fade region. However, outages occur mostly in the deep-fade region (see Figs. 15 and 16), therefore, assumptions (f) and (g) are acceptable for outage calculations.

Notice that the selective-switching scheme always connects the receiver to the better one of the two input signals and is an idealized switch. In practice, there are many different ways of utilizing the two input-diversity signals. For example, threshold-blind switching provides less improvement, whereas equal-gain combining or maximum-ratio combining provides more improvement than the selective switching.[12] Furthermore, a digital radio may be caused to switch at a given error-rate threshold rather than through the signal amplitudes of the two input signals. In this paper, we use selective switching for ease of computation and thus provide only an estimation of feasible diversity improvement. A full-scale investigation of various diversity-protection schemes is beyond the scope of this paper.

(h) The total outage probability is

$$P_{\text{outage}} = \int_0^\infty P(\text{outage}|v_{\text{IL}})[p(v_{\text{IL}})dv_{\text{IL}}]$$

$$= \int_0^\infty P(v_{xp} \geqq v_{xpo}|v_{\text{IL}})[p(v_{\text{IL}})dv_{\text{IL}}]$$

(31)

and the total two-way outage time is

$$T_{\text{outage}} = 2rT_0 P_{\text{outage}} \text{ minutes/year,}$$

(32)
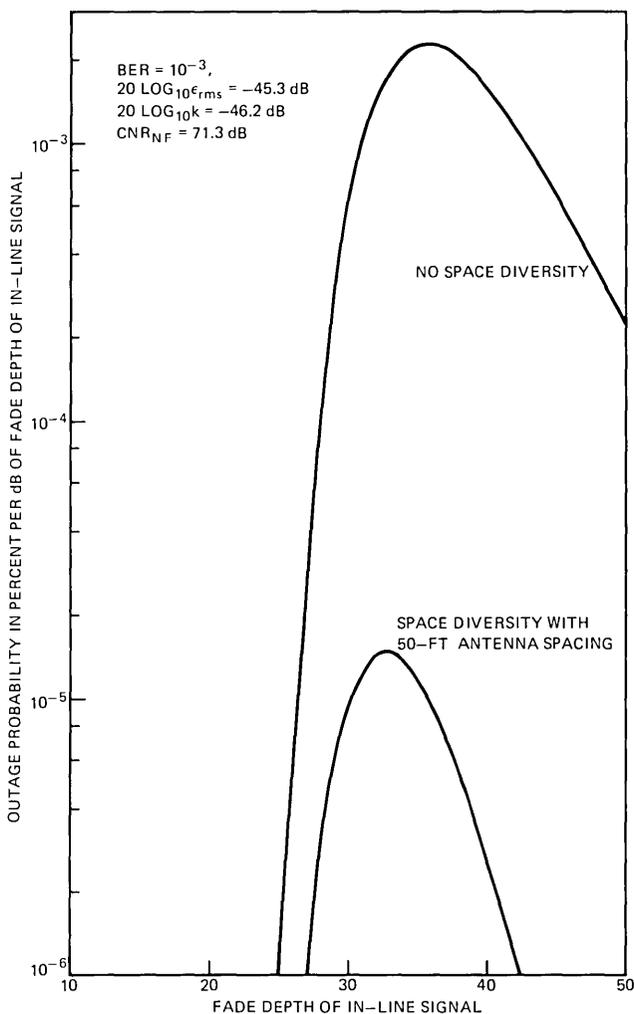
Fig. 15—Dependence of one-way outage probability density function on multipath fade depth of in-line signal of 11-GHz QCPSK radio on a 15.9-mile path with and without space-diversity protection.

where $T_0 = 525,600$ minutes is the total annual time and r is the multipath occurrence factor,[14,27] which depends on many factors such as radio frequency, path length, path terrain, and geographic location. Figures 15 and 16 show the integrand, $[P(\text{outage}|v_{\text{IL}})p(v_{\text{IL}})]$, as a function of fade depth of the in-line signal. Most outage occurs at about 12 dB above the rms power level (i.e., $20 \log_{10}\epsilon_{\text{rms}}$) of the residual cross-polarization interference ϵ. This is because QCPSK radio needs at least 10 dB of CIR at BER $= 10^{-3}$ and the correlated component, kv_{IL}, of cross-polarization interference adds another 2-dB requirement. Therefore, the "effective

MICROWAVE DEPOLARIZATION   663
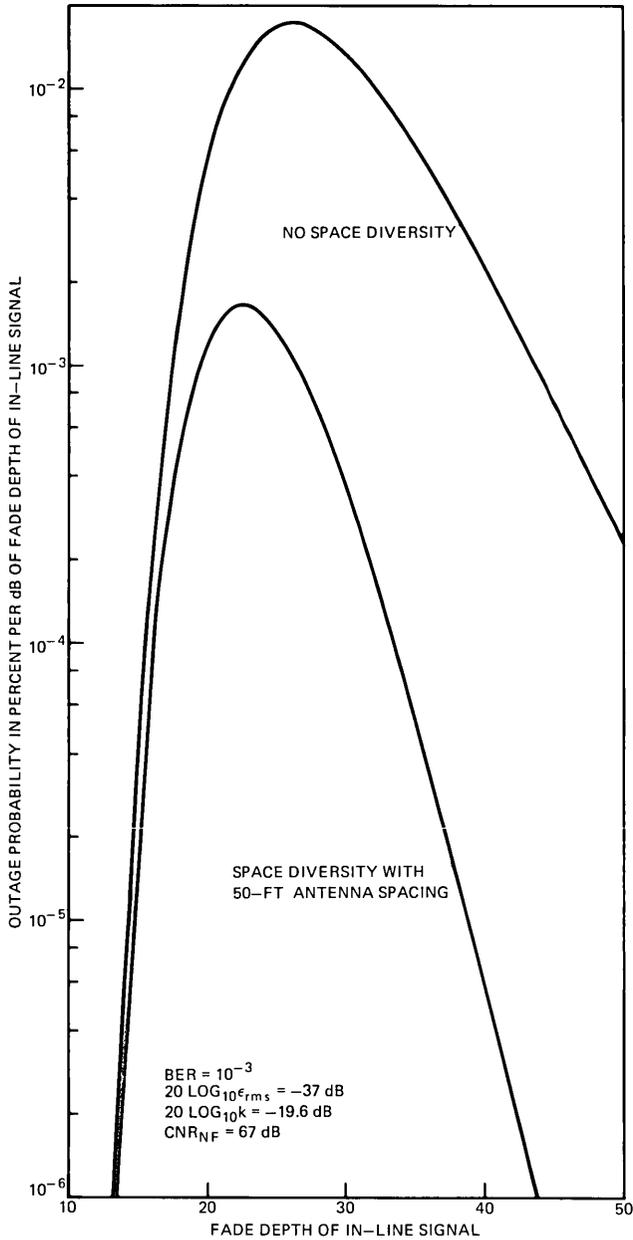
Fig. 16—Dependence of one-way outage probability density function on multipath fade depth of in-line signal of 11-GHz QCPSK radio—26.4-mile path with and without space diversity protection. Data based on multipath-caused one-way outage.

outage threshold" of dual-polarization QCPSK radio is approximately 12 dB* above ϵ_{rms}. For example, the 26.4-mile path has an effective fade margin of only 25 dB.

Notice that the thermal noise is 67 dB below the nonfaded signal level whereas this outage threshold for the 26.4-mile path is only 25 dB below the nonfaded signal level. Therefore, "multipath caused" outage of dual-polarization radio is strongly controlled by ϵ_{rms}.

(i) According to assumptions (f) and (h), the distribution of v_{IL} in the deep-fade region is

$$T(v_{IL} \leq L) \cong rT_0L^2 \text{ for } L \leq 0.1, \tag{33}$$

where $T(v_{IL} \leq L)$ denotes the accumulated time per year that v_{IL} fades below L. The value of multipath occurrence factor, r, can be determined empirically by fitting eq. (33) to the measured data on $T(v_{IL} \leq L)$. The estimated values of r for a 15.9-mile path and a 26.4-mile path near Atlanta, Georgia are given in Table I.

3.4 Results

To calculate outage probability of a radio path, we need the parameters: r, k, and ϵ_{rms} on that path. At the present time, these parameters are available only for two paths (15.9 and 26.4 miles—see Table I) measured near Atlanta, Georgia. This limits our calculations to these two particular path lengths only. The linear interpolation between these two paths lengths in Figs. 17 through 19 gives a crude estimate of outage probabilities for intermediate path lengths.

3.4.1 Effect on dual-polarization transmission

Figure 17 displays the calculated outage probabilities without diversity protection. The upper curve represents cochannel dual-polarization transmission, whereas the lower curve represents idealized single-polarization performance without cross-polarization interference and without adjacent channel interference. It is seen that the impact of a dual channelization is three orders of magnitude increase in multipath outage time. This is because the residual cross-polarization interference ϵ_{rms} is 30 dB higher than the thermal noise.

However, the idealized single-polarization performance, represented by the lower curve in Figure 17, is academic in that the impacts of adjacent channel interferences and channel dispersion during multipath fading are neglected, resulting in more than 54 dB of multipath fade margin. For a practical single-polarization-per-frequency digital system,

---

* 16 dB for BER = 10^{-6}.
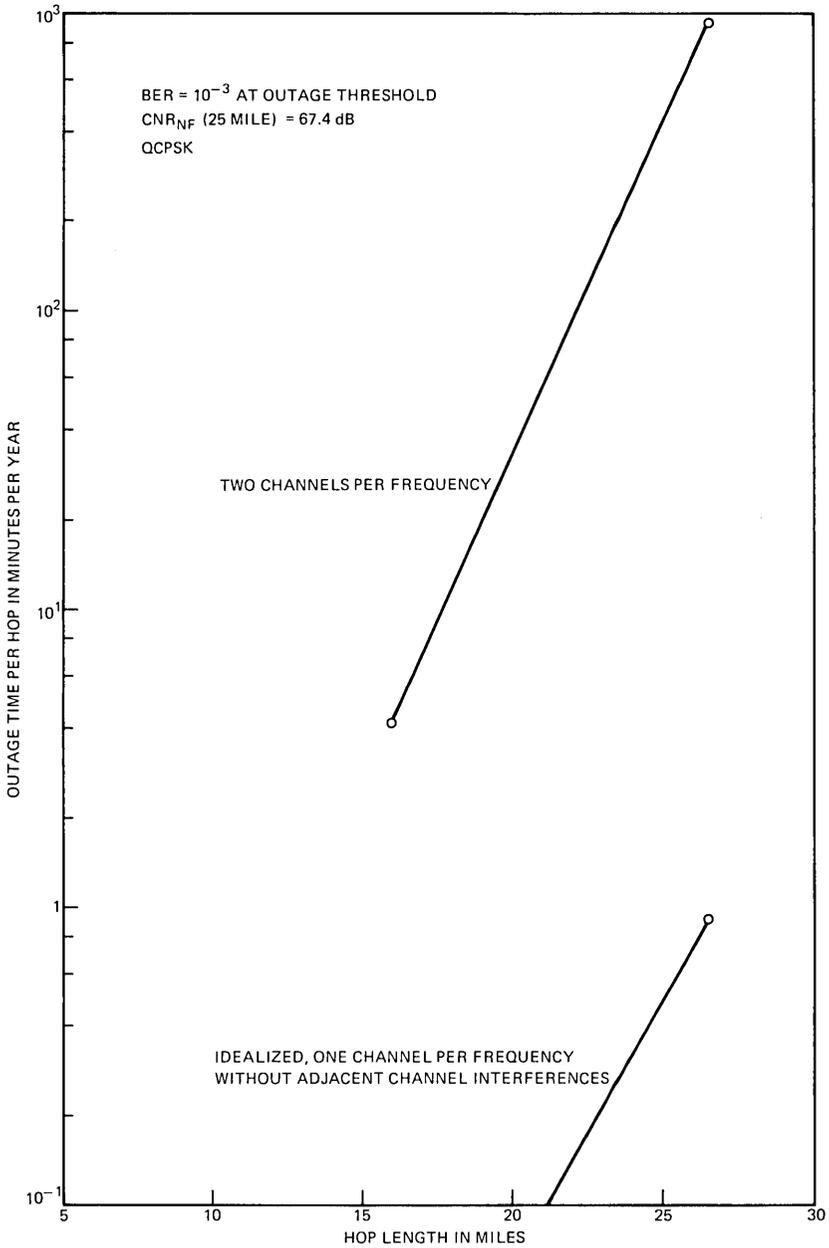
Fig. 17—The impact of cochannel dual-polarization transmission on multipath two-way outage probability of 11-GHz QCPSK radio. No space-diversity protection was provided.

the adjacent channel interferences and channel dispersion will limit the multipath fade margin to less than 35 dB and significantly increase the multipath outage time with reference to the idealized curve.

3.4.2 Space-diversity improvement

Figure 18 displays the reduction of multipath outage time through the use of space diversity. With 50-foot antenna spacing, the outage time is reduced by one order of magnitude for the 26.4-mile path and two orders of magnitude for the 15.9-mile path. The diversity improvement factor achieved by the dual-polarization digital radio is generally much smaller than we familiarly associate with a single-polarization analog radio.[24],[26],[27] This is because the dual-polarization digital radio has an effective fade margin of only 20 to 30 dB, whereas single-polarization analog radio generally provides a fade margin of 35 dB or more. Diversity improvement is proportional to fade margin.

3.4.3 Effect of error-rate requirement

Figure 19 shows that tightening the BER requirement at outage threshold from 10^{-3} to 10^{-6} increases the outage time by a factor of 6, even with diversity protection. This sensitivity is related to the steep inverse slope, 5 dB per decade, of fading probability of a dual-diversity signal.[12] Tightening the BER from 10^{-3} to 10^{-6} is equivalent to 4 dB (i.e., 16 to 12 dB) loss of effective fade margin.

3.4.4 Effects of ϵ_{rms} and k of interference

Figure 20 indicates that the outage probability is extremely sensitive to the power level, ϵ_{rms}, of the residual cross-polarization interference. A 10-dB decrease in ϵ_{rms} will reduce outage probability by two orders of magnitude. Again, this is related to the steep, inverse slope of 5 dB per decade of fading probability for a dual-diversity signal.

On the other hand, Figure 21 shows that the outage probability is practically independent of the correlated components (kv_{IL}) of cross-polarization interference as long as $20 \log_{10}k \leqq -30$ dB. This is intuitively obvious because this component of CIR is constant. As long as CIR $\geqq 30$ dB, the interference has negligible effect on QCPSK radio (see CNR versus CIR curve in Fig. 14). In practice, this means it is of questionable benefit to improve XPD of a radio link beyond 30 dB, during periods of normal propagation. The controlling factor on outage time is the residual cross-polarization, ϵ, which is typically 10 dB or more below kv_{IL} and is almost unobservable during a nonfading period. Therefore, the under-
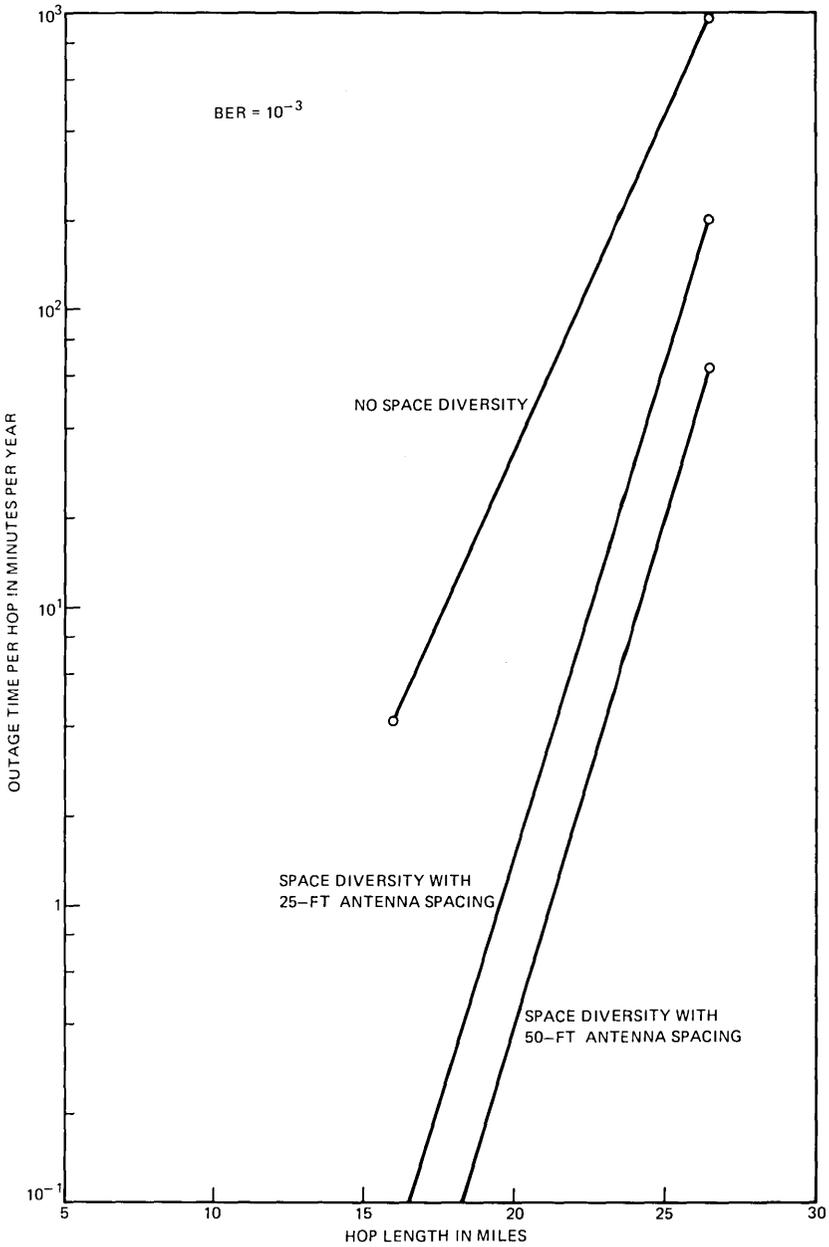
Fig. 18—Effect of space-diversity protection on multipath two-way outage probability of 11-GHz dual-polarization QCPSK radio.
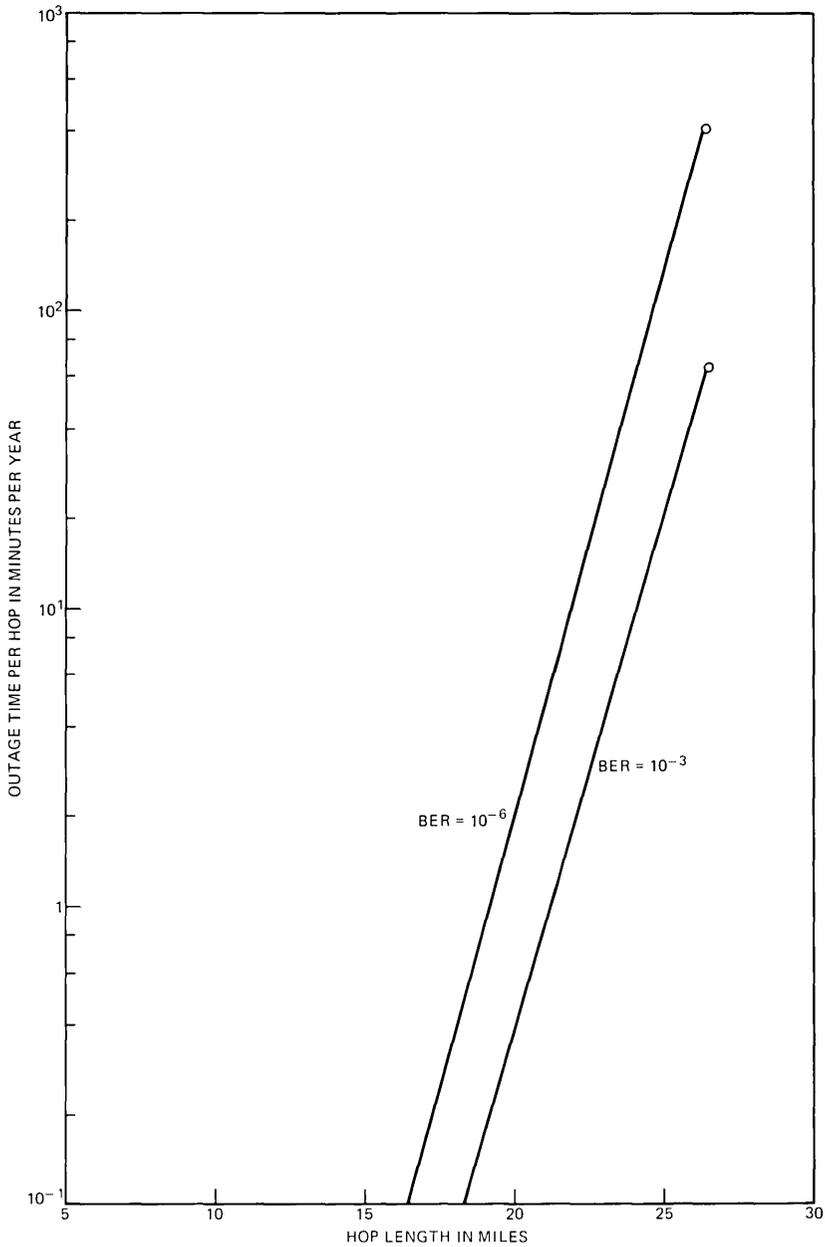
Fig. 19—Effect of bit-error-rate requirement at outage threshold on multipath two-way outage probability of 11-GHz dual-polarization QCPSK radio—with space diversity and 50-foot antenna spacing.
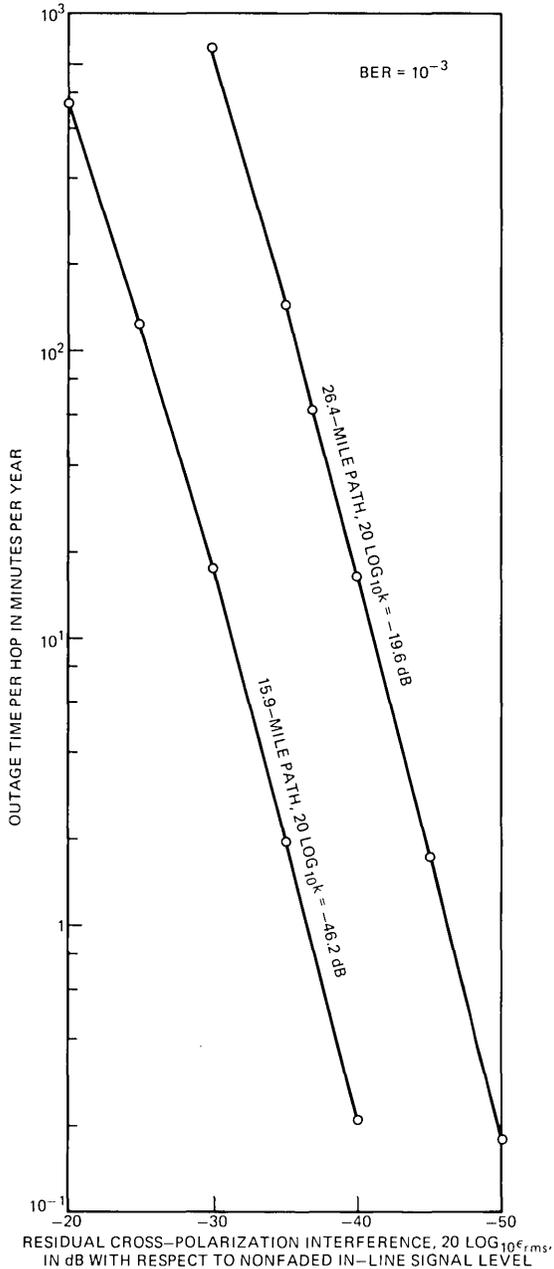
Fig. 20—Sensitivity of multipath two-way outage probability of 11-GHz dual-polarization QCPSK radio on the rms value, ϵ_{rms}, of the residual component of cross-polarization interference—with space diversity and 50-foot antenna spacing.
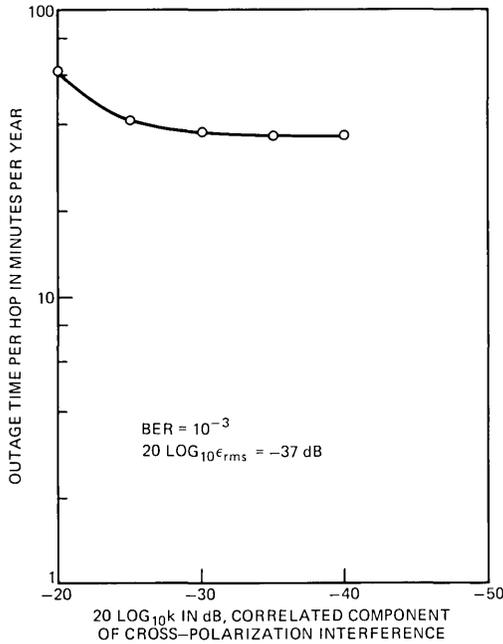
Fig. 21—Dependence of multipath two-way outage probability of 11-GHz dual-polarization QCPSK radio on the correlated component, kv_{IL}, of cross-polarization interference—with space diversity, 50-foot antenna spacing, and 26.4-mile path.

standing and identification of major contributors of the residual cross-polarization interference ϵ are important subjects for future study.

3.4.5 Effect of thermal noise

The effects of varying CNR_{NF} on outage probability are shown in Fig. 22. Since $CNR_a = 13.4$ dB, at $BER = 10^{-3}$, assumption (a) implies a 54-dB fade margin* in the absence of the interference.[22] Figure 22 shows it is also not of benefit to suppress the thermal noise level below the level, $\epsilon_{rms} - \Delta$, of residual cross-polarization interference as far as multipath outage† is concerned. As long as the noise level is below $\epsilon_{rms} - \Delta$, the multipath outage is controlled by ϵ_{rms} and is independent of noise level.

3.5 Some qualifications

3.5.1 Effect of channel dispersion during multipath fading

A quaternary PSK signal is equivalent to two streams of binary PCM signals, each phase modulating a carrier of the same frequency, but with

---

   * The fade margin is reduced to 48 dB during rain because of 4-dB wet radome attenuation and 2-dB degradation by rain-induced depolarization.
   † On the other hand, reducing the noise level will reduce rain-caused outage time.
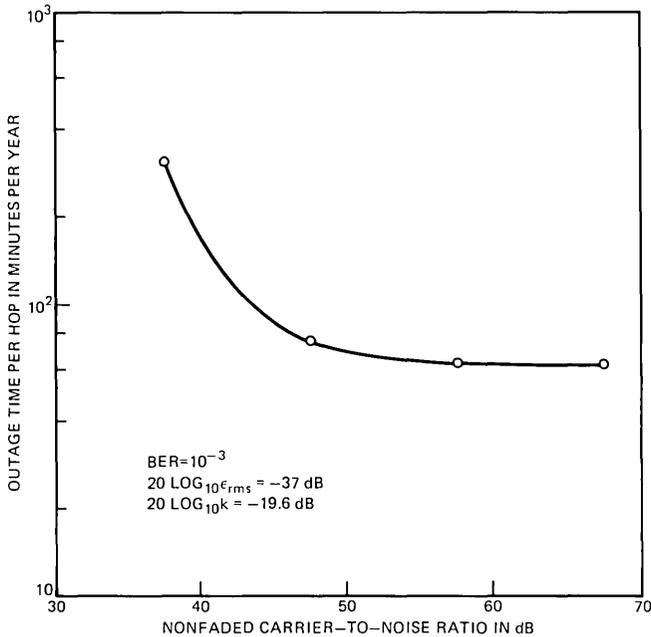
Fig. 22—Effect of reducing thermal noise level on the multipath two-way outage probability of 11-GHz dual-polarization QCPSK radio—with space diversity, 50-foot antenna spacing, and 26-mile path length.

90-degree phase shift.[22] Therefore, cochannel dual-polarization QCPSK radio requires the following two orthogonalities for successful transmission:

(*i*) Polarization orthogonality between the cochannel pair.

(*ii*) Phase orthogonality between the two streams of binary PSK signals.

Multipath fading, with channel dispersion, will degrade both polarization and phase orthogonalities. At the present time, our understanding of degradation of phase orthogonality during multipath fading is insufficient for outage calculations. Therefore, the possible outages due to crosstalk between the two orthogonally phased carriers are unaccounted for in this paper.

3.5.2 Limitation of available path parameters

All the calculations in this report are based on measured path parameters: r, k, and ϵ_{rms}, pertinent to two paths near Atlanta, Georgia. The variations of k and ϵ_{rms} parameters with time base, path length, and geographic location are not understood.

3.5.3 *Relationship between CNR and CIR*

The CNR versus CIR relationship for QCPSK radio (Fig. 14) is based on presently available hardware technology. This behavior is several dB poorer than the idealized relationship in Refs. 28 and 29. Future advances may improve this behavior and reduce the outage time.

IV. CONCLUSION

A microwave propagation experiment near Atlanta, Georgia has provided statistics describing depolarization during multipath fading at 6 and 11 GHz on two paths (15.9 and 26.4 miles). Data gathered over a 6.5-month period have been processed and presented. These data are supportive of a model in which the depolarized signal (interference) consists of a component correlated with the in-line signal, and a residual component independent of the in-line signal. The residual component is Rayleigh distributed with a typical rms value about 40 dB below the nonfaded in-line signal. The cross-polarization interference, conditioned to a given in-line signal level, is Rice-Nakagami distributed. The calculated Rice-Nakagami distribution fit these 6.5-month data very closely.

The experimental data and theoretical model are applied to estimate the multipath outage probability of cochannel dual-polarization 11-GHz QCPSK digital radio affected by thermal noise and cross-polarization interference during multipath fading. Detailed results are given in Section 3.4.

V. ACKNOWLEDGMENTS

REFERENCES

1. W. T. Barnett, "Degradation of Cross-Polarization Discrimination During Rain and Multipath Fading at 4 GHz," IEEE 1974 International Conference on Communications, June 17–19, 1974, Minneapolis, Minnesota, Conference Record, pp. 12D-1—12D-4, IEEE Cat. No. 74, CHO 859-9-CSCB.
2. K. Morita, "Fluctuation of Cross-Polarization Discrimination Ratio Due to Fading," Rev. of Elec. Commun. Lab., *19*, No. 5–6 (May-June 1971), pp. 649–652.
3. T. Inoue and T. Akiyama, "Propagation Characteristics of Line-of-Sight Oversea Paths in Japan, "IEEE Trans. Ant. Propag., *AP-22*, No. 4 (July 1974), pp. 557–565.
4. T. Inoue, "Microwave Oversea Circuit Cross-Polarization Distortion Characteristics," Rev. Elec. Commun. Lab., *23*, No. 3–4 (March-April 1975), pp. 374–386.
5. G. Lefrancois, L. Martin, and M. Rooryck, "Influence of the Propagation on the Value

of the Decoupling of Two Orthogonal Polarizations," Ann. Telecommun., *28* (1973), pp. 316–324.

6. T. O. Mottl, "Dual-Polarized Channel Outages During Multipath Fading," B.S.T.J., this issue, pp. 675–701.

7. R. W. Friis and A. S. May, "A New Broadband Microwave Antenna System," Commun. Electron., No. 35 (March 1958), pp. 97–100.

8. E. T. Harkless, "A Network for Combining Radio Systems at 4, 6 and 11 kmc," B.S.T.J., *36* (September 1957), pp. 1253–1267.

9. M. A. Gerdine and H. F. Lenzing, "Reduction of Delay Distortion in a Horn-Reflector Antenna System Employing Overmoded-Waveguide Feeder," IEEE Trans. Commun. Technol., *COM-18,* No. 1 (February 1970), pp. 21–26.

10. P. I. Sandsmark, "Effect of Ellipticity on Dominant-Mode Axial Ratio in Nominally Circular Waveguides," IRE Trans. Microw. Theory Tech., *MTT-3,* No. 5 (October) Vol. 1955, pp. 15–20.

11. S. H. Lin, "Statistical Behavior of a Fading Signal," B.S.T.J., *50,* No. 10 (December 1971), pp. 3211–3270.

12. S. H. Lin, "Statistical Behavior of Deep Fades of Diversity Signals," IEEE Trans. Commun., *COM-20,* No. 6 (December 1972), pp. 1100–1107.

13. A. Vigants, "Number and Duration of Fades at 6 and 4 GHz," B.S.T.J., *50* (March 1971), pp. 815–841.

14. W. T. Barnett, "Multipath Propagation at 4, 6 and 11 GHz," B.S.T.J., *51,* No. 2 (February 1972), pp. 321–361.

15. S. O. Rice, "Mathematical Analyses of Noise," B.S.T.J., *23* (July 1944), pp. 282–332, and *24* (January 1945), pp. 46–256.

16. M. Nakagami, "The m-Distribution—A General Formula of Intensity Distribution of Rapid Fading," in W. C. Hoffman (ed.), *Statistical Methods in Radio Wave Propagation,* Oxford: Pergamon, 1960.

17. P. Beckmann, *Probability in Communication Engineering,* New York: Harcourt, Brace & World, 1967.

18. E. T. Harkless and H. F. Lenzing, "Excitation of Higher Order Antenna Modes by Multipath Propagation," IEEE Trans. Commun. Technol., *COM-15,* No. 4 (August 1967), pp. 567–603.

19. H. F. Lenzing, "Higher Order Mode Excitation in Larger Aperture Receiving Antennas," Microw. J., *12,* No. 12 (December 1969), pp. 61–65.

20. S. H. Lin, "A Method for Calculating Rain Attenuation Distributions on Microwave Paths," B.S.T.J., *54,* No. 6 (July-August 1975), pp. 1051–1086.

21. Unpublished work.

22. A. J. Giger and T. L. Osborne, "3A-RDS 11 GHz Digital Radio System," 1976 International Conference on Communications, Philadelphia, Pennsylvania, June 14–16, Conference Record, pp. 18-1—18-7, *IEEE Cat.,* No. 76CH1085-0, CSCB, Vol. II.

23. H. Staras, "Diversity Reception with Correlated Signals," J. Appl. Phys., *27,* No. 1 (January 1956), pp. 93–94.

24. A. Vigants, "Space Diversity Performance for Antennas of Unequal Size and Correlated Rayleigh Fading," unpublished work.

25. R. H. Clarke, "A Statistical Theory of Mobile-Radio Reception," B.S.T.J., *48* (August 1969), pp. 957–1000.

26. A. Vigants, "Space Diversity Performance as a Function of Antenna Separation," IEEE Trans. Commun. Technol., *COM-16,* No. 5 (December 1969), pp. 831–836.

27. A. Vigants, "Space Diversity Engineering," B.S.T.J., *54,* No. 1 (January 1975), pp. 103–142.

28. A. S. Rosenbaum, "PSK Error Performance with Gaussian Noise and Interference," B.S.T.J., *48,* No. 2 (February 1969), pp. 413–442.

29. 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.

Dual-Polarized Channel Outages During Multipath Fading

By T. O. MOTTL

A statistical model for outages that occur in a dual-polarized-frequency radio channel during periods of multipath fading is proposed, and an analytical expression for outage time obtained. The model, which results in Rice-Nakagami statistics for the cochannel interference signal, describes channel outage time as a function of several environmental and radio-system parameters. The formulation obtained allows for efficient parameteric studies to evaluate the importance of these parameters to channel outage time and to examine parameter sensitivity questions. Results of practical significance relative to hardware XPD requirements, maximum hop length, system gain, dependence on geographic environment, and digital terminal performance characteristics are obtained for present 11-GHz QCPSK (quaternary-coherent-phase-shift-keyed) digital radio systems. Estimates of dual-polarized-frequency channel outage time are obtained for a variety of representative system parameter values and compared with expected outage times for a conventional channel. Particular attention is given to the mechanism of channel outages during multipath fading, and several potential means for control and/or reduction of channel outage time are discussed.

I. INTRODUCTION

1.1 DPF radio systems

Federal regulatory requirements for the channel capacity of an 11-GHz digital radio system are achieved in some short-haul radio systems by a diplex approach that uses each available radio-channel frequency band twice. Such designs, denoted here as dual-polarized-frequency (DPF) radio systems, are based on the high isolation theoretically obtainable between the two linear polarizations (vertical and horizontal) available for propagation at each radio-wave frequency.

Practical implementation of a DPF radio system relies on the attainability of high polarization isolation in the radio system, in the antenna system, and over the propagation path, and on the maintainability of these conditions. At any given frequency, degradation in the isolation between the orthogonally polarized signals can result in cochannel interference, with attendant degradation in the channel bit-error rate (BER).

This paper presents a theoretical modeling study of DPF radio channel reliability during periods of multipath fading. Using channel bit-error rate as a criterion, the effects of the polarization-isolation properties of system hardware and the propagation path on channel-outage time are evaluated. The sensitivity of bit-error rate to system hardware and propagation-path characteristics is of particular interest because of their direct impact on feasibility and cost of DPF radio systems.

A complete listing of all symbols and notation used in this paper is provided in Appendix A.

1.2 Linear polarization isolation

During periods of multipath fading activity, measurements on 11-GHz radio waves propagated along line-of-sight paths display statistical fluctuations in the isolation between orthogonal linearly polarized waves at the same frequency.[1-4] A convenient measure of the isolation between two cross-polarized signals—i.e., signals transmitted via such orthogonally polarized waves—is the cross-polarization-discrimination ratio or XPD. The XPD is defined as the ratio (usually expressed in dB) of the energy received on a reference polarization that results from a signal transmitted with the same polarization (the copolarized, desired signal) to the energy received on this same polarization but transmitted with the orthogonal polarization (the cross-polarized or interference signal).

1.3 Multipath outage estimates

Estimates of single-hop (point-to-point) outage probabilities for 11-GHz QCPSK digital radio systems were first obtained by Lin in 1973 based in part on a one-parameter model for the RMS-XPD ratio during multipath fading. In this model, at any fixed fade depth in the deep-fade region, the cross-polarized voltage envelope varies as a Rayleigh distributed random variable, and the RMS-XPD behaves as a linear function of the fade depth in dB, as shown by curve "L" in Fig. 1.

Some early experimental data[4] showed a nonlinear dependence of RMS-XPD vs fade depth, with linear proportionality between RMS-XPD and fade depth only during deep multipath fading, e.g., fades greater than 10 to 15 dB. Curve NL in Fig. 1 is a plot of one sample data set that
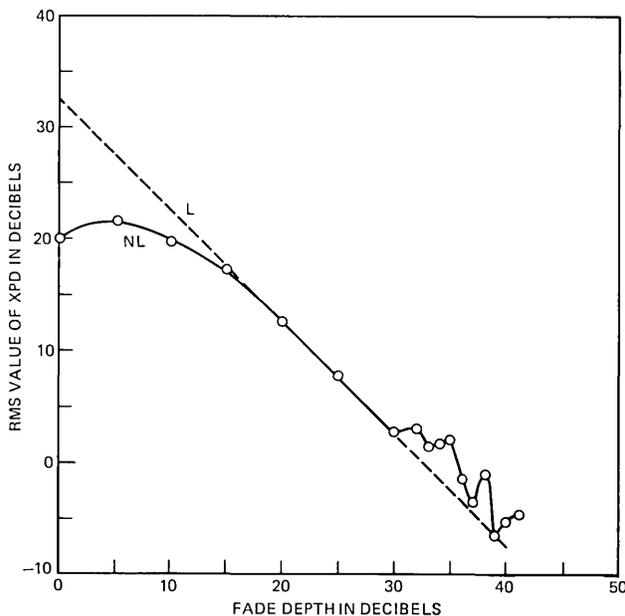
Fig. 1—RMS value of cross-polarized interference voltage as a function of fade depth. (Atlanta to Palmetto, Georgia, June 1974, 11 GHz.)

shows this type of behavior. A model for the cross-polarized signal that accounts for this observed behavior is proposed and developed in detail in this paper.

This proposed model for cross-polarized interference describes the signal as a sum of two components

$$V_{TX} = V_{DX} + V_{RX}, \tag{1}$$

where

V_{TX} = peak amplitude of the voltage envelope of the total signal received due to the transmitted cross-polarized signal.

V_{DX} = peak amplitude of the envelope of a "direct" received component, due to the transmitted cross-polarized signal.

V_{RX} = peak amplitude of the envelope of a random received component, due to the transmitted cross-polarized signal.

The component V_{DX} derives from imperfect polarization isolation in system hardware, and is subject to fading behavior identical to that of the copolarized signal. The random component V_{RX} is attributed to random spatial and temporal conditions that characterize the propagation path, and it is not subject to fading, thereby giving rise to the observed behavior of curve NL in Fig. 1. This random component is described here as a gaussian random process, which, as shown below, results

in V_{TX} becoming a Rice-Nakagami distributed random variable. This result is supported by experimental data taken at Palmetto, Georgia, as discussed by Lin.[4]

Use of a gaussian model for V_{RX} is based on the multiplicity of effects that can contribute to this component and on the conjecture that random scattering is a dominant mechanism. The generalized parametric modeling study of outage statistics carried out in this paper is predicated in part on this latter assumption. Without a model for the dependence of the random component of the interference signal on path length, outage-time estimates are limited to those few hop lengths where sufficient experimental data is available to specify the magnitude of this component.

In a companion paper in this issue, Lin[4] also uses the proposed Rice-Nakagami model for cross-polarized interference voltage to compute, by essentially the same method, estimates of DPF channel multipath outage time for QCPSK systems for two existing radio hops with and without space-diversity protection. To provide perspective on some of the principal differences between Lin's study and ours, the following comments are provided:

(i) Lin's analysis of experimental propagation data provides support for the Rice-Nakagami model for cross-polarized interference.

(ii) Estimates of outage time for two specific hop lengths are provided by Lin, whereas in this study hop length is a free model parameter. As a result, in Lin's study, outage time for intermediate hop lengths must be inferred through linear interpolation between the two point estimates. The physical conditions corresponding to such interpolated values are not clear, since the two (experimental radio path) interpolation end points correspond to different hardware and propagation path characteristics. Also, this linear interpolation, performed on semilog outage time plots, corresponds to a simple power law dependence of outage time on path length.

(iii) Some effects of space diversity on multipath outage time are considered in Ref. 4, but not here.

(iv) The sensitivity of multipath outage time to differences between theoretical and experimental DPF channel performance characteristics, and the implications for potential outage time reductions, are presented in this study, but not in Ref. 4.

II. DPF RADIO-CHANNEL PERFORMANCE DURING MULTIPATH FADES

2.1 Multipath-fading outage probability

During multipath fading, a DPF radio-channel outage occurs whenever the channel BER exceeds some defined threshold for acceptable service.

Two service outage thresholds are considered here: BER $= 10^{-3}$ and BER $= 10^{-6}$. A service outage results if

(*i*) a multipath fade occurs and exceeds the thermal-noise fade margin (conventionally defined fade-margin), or

(*ii*) the polarization isolation (XPD) of the DPF-channel degrades below that necessary to maintain the required error rate (assuming both polarizations are in service).

The effect on outage time of spectral distortions due to the frequency selectivity of multipath fades is not considered here. As a result, optimistic estimates of multipath fading outages are obtained, and can be viewed as lower bounds for actual outage times.

Under the condition that multipath fading is present, the probability of a DPF digital-radio channel outage occurring as a result of a multipath fade on a single radio hop is given by the expression

$$P_{\text{OUT}} = \int_0^\infty P_e(\text{BER} \geqq \text{BER}_0 | L) p_F(L) dL, \qquad (2)$$

where $L \equiv$ peak envelope voltage normalized to the nominal (unfaded) value—i.e., the ratio of faded to unfaded peak amplitude of the voltage envelope, and where $P_e(\text{BER} \geqq \text{BER}_0 | L)$ is the complement of the conditional probability distribution for the channel bit-error rate, (i.e., the probability that the BER of the channel exceeds that for acceptable service (BER$_0$), at a given fade level L), and the function $p_F(L)$ is the probability density function for the amplitude of the multipath fading signal.

The conditional probability P_e for exceeding a given bit-error rate BER$_0$, given a fixed fade level L, is defined by the relation

$$P_e(\text{BER} \geqq \text{BER}_0 | L) = \begin{cases} \int_{V_M(L)}^\infty p_{VX}(\xi | L) d\xi & \text{for } L \geq L_m(D) \\ 1.0 & \text{for } L < L_m(D) \end{cases}, \qquad (3)$$

where p_{VX} is the conditional density function for the cross-polarized interference-signal envelope, $V_M(L)$ is the maximum value of V_{TX} for which the desired BER$_0$ can be obtained with the channel signal faded to L, and $L_m(D)$ is the minimum-allowable normalized amplitude of the received-signal envelope for a path length D required to maintain some specified maximum channel BER$_0$—i.e., the fade level at the system thermal-noise threshold. [See Section 3.1 for a graphical interpretation and explanation of (3) and (4).]

It is instructive to rewrite (2) in the following form, using the definition in (3):

$$P_{\text{OUT}} = \int_0^{L_m(D)} p_F(L)dL + \int_{L_m(D)}^{\infty} \int_{V_m(L)}^{\infty} p_{VX}(\xi|L)p_F(L)d\xi dL.$$

$$(4)$$

The first term on the right-hand side of eq (4) represents the outage time resulting from fades below the system thermal-noise threshold $L_m(D)$, while the second term represents outage time resulting from degradation of cross-polarization isolation during fades down to the threshold $L_m(D)$.

To estimate total outage probability (or outage time), which results during multipath fades, statistical models are required for both the co-polarized and cross-polarized signals as received on the copolarized receiving channel during multipath fading conditions. In the deep-fade region (fades in excess of 10 dB) an accepted statistical model for the peak envelope voltage of a signal is given by the probability distribution function,[1,8]

$$P_F(l \leq L) = (1 - e^{-L^2}) \cong L^2. \qquad (5)$$

This distribution function results in the density function

$$p_F(L) = \frac{d}{dl} P_F(l \leq L) = 2Le^{-L^2} \cong 2L \qquad (6)$$

required in the outage probability expression, eq (2).

It is shown below that restriction of the statistical model for the peak envelope voltage L to the deep-fade region does not restrict results of practical interest. This fortunate circumstance results because outage statistics accumulate almost entirely during time intervals that correspond to this deep-fade region.

With the proposed model for the cross-polarized interference voltage, it can easily be shown that the conditional probability density function of the envelope of the cross-polarized interference voltage at a fixed fade depth L is given by the expression[9]

$$p_{VX}(V|L) = \frac{V}{\sigma_{RX}^2} I_0 (C_0LV/\sigma_{RX}^2)\exp\left(-\frac{V^2 + (C_0L)^2}{2\sigma_{RX}^2}\right), \qquad (7)$$

where σ_{RX}^2 = variance of the gaussian component of the cross-polarized interference voltage, $I_0(\cdot)$ is the modified Bessel function of order zero, and $0 \leq C_0 \leq 1.0$ is (approximately) the ratio of the cross-polarized interference to copolarized envelope voltages received during unfaded propagation conditions, i.e.,

$$C_0 = 10^{-(\text{XPD}_0/20)},$$

where XPD_0 is the cross-polarization discrimination ratio with no multipath fading. (Voltages used here and in all subsequent expressions are
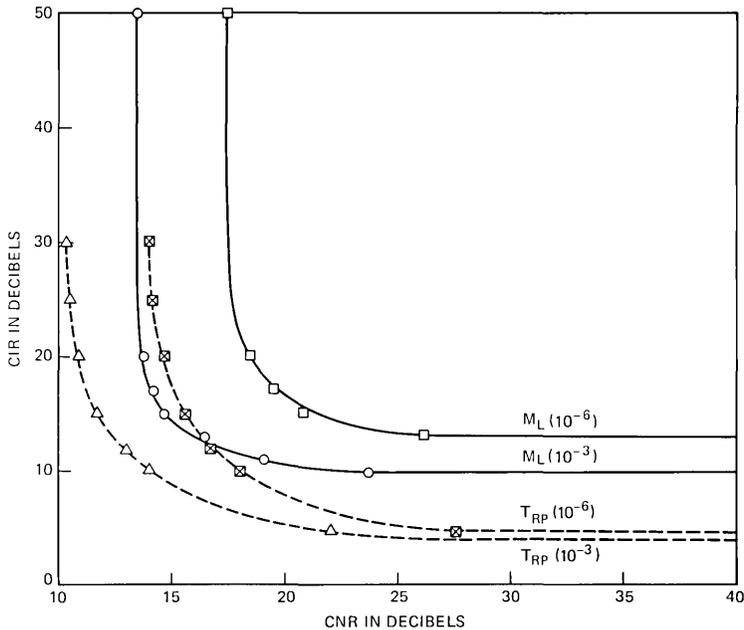
Fig. 2—DPF-channel digital radio performance curves plotted for constant bit-error rate.

given normalized to their unfaded levels.) Expression (7) is known as the Rice-Nakagami probability density function.[5-7]

2.2 Characterization of channel performance at constant bit-error rate

A graphical representation of the relation between carrier-to-interference ratio (CIR) and carrier-to-(thermal) noise ratio (CNR) at constant bit error rate for a DPF channel can be derived from recent theoretical studies.[10,11] The two dashed curves shown in Fig. 2 were obtained from results given in the above referenced works and represent the relationship between CIR and CNR at bit-error rates of 10^{-3} and 10^{-6}. For convenient reference, such curves are referred to as DPF-channel performance curves in this study.

For a single interferer (the cross-polarized interference signal of the DPF channel), the relation between V_{TX} and CIR is

$$\text{CIR} = 20 \log_{10}(V_c/V_{TX}). \qquad (8a)$$

Note that, in terms of the normalized voltage $V_M(L)$, which appears in (3), the values of CIR that correspond to a DPF-channel performance curve are given by the expression

$$\text{CIR}_{\text{BER}_0} = -20 \{\log_{10}[V_M(L)]\}. \qquad (8b)$$

The relationship of L to CNR is given by

$$\mathrm{CNR}(D) = \mathrm{CNR}_0(D) - \mathrm{FD}(L), \tag{8c}$$

where $\mathrm{CNR}_0(D)$ is the nominal unfaded design value of CNR for a path of length D, and

$$\mathrm{FD}(L) = -20 \log_{10}(L) \tag{8d}$$

is the channel fade depth (L as defined above). The quantity $\mathrm{CNR}_0(D)$ can also be expressed as

$$\mathrm{CNR}_0(D) = \mathrm{CNR}_{th} + \mathrm{FM}(D), \tag{9}$$

where CNR_{th} is the thermal-noise threshold value of CNR for a given BER_0 with no interference present, and $\mathrm{FM}(D)$ is the system thermal-noise fade margin for path length D.

The fade margin can be written as

$$\mathrm{FM}(D) = -20 \log_{10}[L_m(D)], \tag{10}$$

where $L_m(D)$ is the normalized thermal-noise threshold signal level as previously defined.

The solid curves shown in Fig. 2 present data from laboratory measurements of 3A-RDS performance for 10^{-3} and 10^{-6} bit-error rates. Note that threshold values for CIR are apparent from this figure—i.e., a value CIR_0 below which a given BER is exceeded independent of CNR. Use of both theoretical and experimental channel performance curves in this modeling study is intended to provide some insight into those non-ideal channel performance factors that significantly influence multipath outage time. These comparisons are unique to this paper.

2.3 Multipath-fading outage time

Experimental results on multipath fading statistics at 11 GHz provide the basis for obtaining estimates of expected DPF radio-channel outage time. From these measurements the expression for multipath-caused outage time in minutes per year can be written as

$$T_{\mathrm{OUT}} = r T_0 P_{\mathrm{OUT}}, \tag{11}$$

$$r = \text{multipath occurrence factor}$$

$$T_0 \equiv (t/50)\, 1.33 \times 10^5 \text{ minutes} \qquad 35° \leq t \leq 75°,$$

where t is average temperature in °F.

The variable T_0 represents the time period during which fading outages can accumulate (see Vigants[12]). The multipath occurrence factor r is a normalization coefficient required to account for the influence of climate, terrain, signal frequency, and path-length on the quantitative
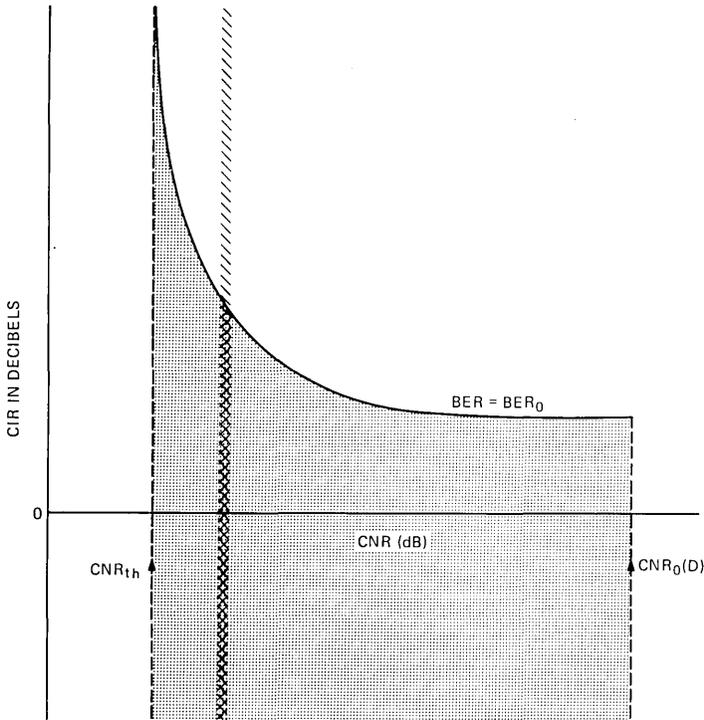
Fig. 3—Graphic representation of outage-probability computation.

statistical behavior of a signal during multipath fades—i.e., it is required to account for variations in the fraction of time that multipath occurs. This factor is defined by the expression.[1]

$$r = c(f/4)D^3 10^{-5}, \tag{12}$$

where

c = geographic factor accounting for terrain and humidity effects (typically $0.25 \leqq c \leqq 4$, with $c = 1.0$ for average terrain and climate).

f = frequency in GHz.

D = path length in miles.

III. OUTAGE ESTIMATES

3.1 Graphical interpretation

The expressions given in Section II for outage probability during multipath fading can be interpreted graphically in terms of a DPF channel performance curve. In the discussion that follows, refer to Fig. 3, which shows a hypothetical performance curve for some fixed BER_0.

Associated with each point in the plane defined by the CNR and CIR axes is a probability density for the occurrence, during multipath fading conditions, of that pair of values (CIR, CNR) that define this point. All points lying along a line perpendicular to the CNR axis correspond to a fixed fade depth (cross-hatched strip), while all those points lying below the performance curve correspond to the condition BER $\geqq$ BER$_0$ (shaded region).

In terms of Fig. 3, the expression for the probability P_e (BER $\geqq$ BER$_0|L$), eq. (3), represents the integration of the probability density of occurrence of all points along a line (fixed value L) perpendicular to the CNR axis and lying in the region below the performance curve (double cross-hatched portion of strip). The expression for total outage probability P_{OUT} is then the integration of the probabilities along all such strips between CNR$_{th}$ and CNR$_0(D)$ (shaded region). The entire region to the left of the vertical asymptote of the DPF-channel performance curve (at CNR$_{th}$) is associated with the outage probability component which results from fades below the system thermal noise threshold.

3.2 Analytic results

The integration of the density function p_{VX} $(V|L)$ indicated in eq. (3) can be carried out to give the analytic expression

$$P_e(\text{BER} \geqq \text{BER}_0|L) = \begin{cases} \exp\left\{-\dfrac{\alpha^2(L) + \beta^2(L)}{2}\right\} \\[2em] \times \displaystyle\sum_{k=0}^{\infty} \dfrac{\left[\displaystyle\sum_{j=0}^{k} \left(\dfrac{\beta^2(L)}{2}\right)^j \Big/ j!\right] \alpha^{2k}(L)}{2^k k!} \; ; \; \text{CNR} \geqq \text{CNR}_{th} \\ \hspace{6cm} ; \text{CNR} < \text{CNR}_{th}, \\[1em] 1.0 \end{cases} \tag{13}$$

where

$$\alpha(L) = \frac{C_0 L}{\sigma_{RX}} \tag{14}$$

$$\beta(L) = \frac{V_M(L)}{\sigma_{RX}} . \tag{15}$$

For convenience in computations, a DPF-channel performance curve is represented analytically by the expression

$$\text{CIR}(D) = K_1 \, \text{Log}_{10}[1.0 - 10^{-K_2\{(\text{CNR}-\text{CNR}_{th})/K_3-\text{CNR})\}}] + \text{CIR}_0, \tag{16}$$

Table I — Empirical DPF-channel performance curve parameters

Curve	K_1	K_2	K_3	CNR_{th}	CIR_0
T_{RP} BER $= 10^{-3}$	−12.90	1.00	35.0	10.25	4.1
M_L BER $= 10^{-3}$	−7.55	1.09	30.0	13.399	9.95
T_{RP} BER $= 10^6$	−13.79	0.95	35.0	13.85	4.85
M_L BER $= 10^6$	−8.70	0.85	30.0	17.3998	12.9

T_{RP} signifies DPF-channel performance characteristics derived from theoretical results given by Rosenbaum[10] and Prabhu.[11]
M_L signifies DPF-channel performance characteristics derived from laboratory measurements on 3A-RDS.

where K_1, K_2, and K_3 are fitting parameters determined empirically to approximate closely the data points as shown in Fig. 2. Table I gives sets of values for the various parameters in eq. 16 used to approximate the performance curves required for subsequent calculations.

Table I also establishes the notation that denotes the DPF-channel performance curve used to obtain specific results: T_{RP} denotes use of the DPF-channel performance curves derived from results obtained by Rosenbaum[10] and Prabhu;[11] M_L denotes use of the DPF-channel performance curves derived from 3A-RDS laboratory measurements.

Table II presents other expressions and parameter values used in subsequent calculations.

3.3 Propagation path and system hardware effects

From the previous expressions for outage probability (eq. 2, and eqs. 11 through 15), it is apparent that outage time statistics for a DPF-channel are dependent on the parameters D (hop length), C_0 (equipment polarization isolation), σ_{RX} (propagation path phenomena), and r (multipath occurrence factor).

Measures of the sensitivity of DPF-channel outage statistics to system hardware characteristics (as represented by C_0) and to propagation path characteristics (as represented by σ_{RX}) are important for evaluating system design and cost.

Table II — Definitions and parameter values pertinent to numerical calculations

Parameter	Definition
Section loss (SL)	SL $= 27.0 + 20.0 \log_{10}(D)$ where D is in miles
System gain (SG)	SG $=$ FM $+$ SL, where FM $\equiv$ fade margin
SG for BER $\leqq 10^{-6}$	$SG_{10^{-6}} = 108$ dB
SG for BER $\leqq 10^{-3}$	$SG_{10^{-3}} = 112$ dB
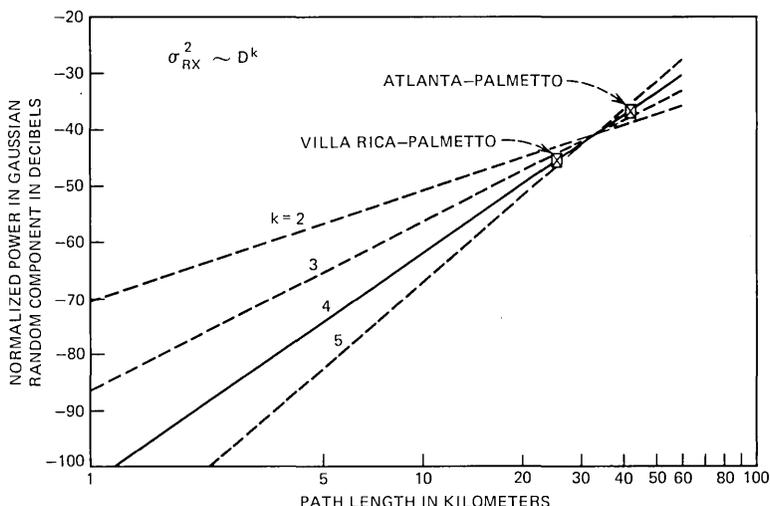
Fig. 4—RMS value of random component of cross-polarized interference voltage as a function of path length: power law approximations to experimental data points.

3.3.1 Propagation-path scattering hypothesis

The parameter σ_{RX}^2, a measure of the energy in the random component of the cross-polarized signal, is probably dependent on a variety of propagation-path parameters. At present, this conjectured path dependence is not clearly understood theoretically, nor is sufficient experimental data available to establish an empirical model. Experimental measurements[4] at Palmetto Ga., taken on two propagation paths, provide the only directly applicable data for σ_{RX}^2. These two data points are shown plotted in Fig. 4, along with various power law relationships between σ_{RX}^2 and distance D.

To compute outage statistics via the expressions for P_{OUT} [eq. (2)], a functional dependence of σ_{RX}^2 on path length D is required. Using the two available data points, we find that an approximate path-length power-law dependence of D^4 would provide a reasonable fit (12 dB per double distance). However, this rate of increase in energy of the random cross-polarized signal component seems too large from the standpoint of physical intuition to apply over any substantial path length interval.

Based on the hypothesis that terrain scattering is a primary contributor to the energy in V_{RX}, a very simple, idealized, theoretical estimate of σ_{RX}^2 as a function of path length D was obtained for three antenna tower heights (see Appendix B). Figure 5 summarizes results of these estimates scaled to the experimental value for σ_{RX}^2 obtained on the longer of the two propagation paths at Palmetto (26.4 mi/42.5 km). This empirical scaling of the estimate for σ_{RX}^2 is equivalent to use of a scattering
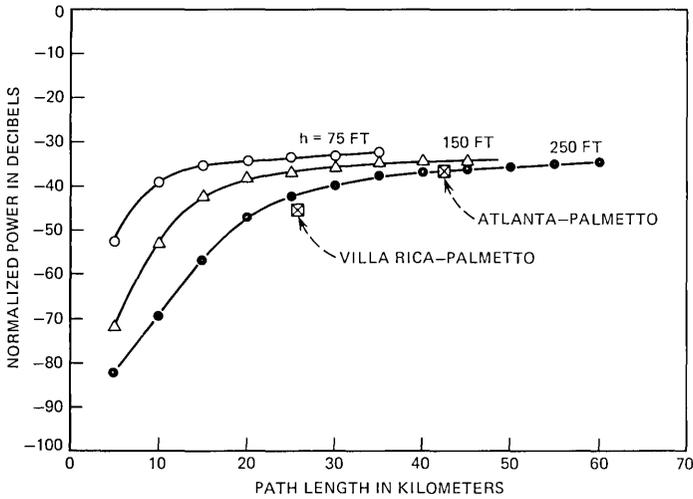
Fig. 5—Prediction for power in random component of cross-polarized interference signal based on terrain-scattering model.

cross section of about -28 dB, in general agreement with experimental data from radar terrain-backscatter measurements.[14,15] Note that while this scaling also produces reasonable agreement between these predictions and the measured value of σ^2_{RX} on the shorter path at Palmetto, the different antennas, widely different antenna heights, and uneven terrain characterizing this latter path prohibit any detailed comparison.

In the absence of experimental data for other path lengths, and in view of the general agreement between the terrain scattering estimates and the Palmetto data points shown in Fig. 5, theoretical estimates of σ^2_{RX} are used here to define a functional relation between this parameter and the path length D for a fixed tower height. The curve in Fig. 5 for the tower height $h = 250$ ft will be used for this purpose as a baseline in calculating outage statistics.

3.3.2 System hardware XPD requirements

Variations in P_{OUT}, which result from changes in the value of the parameter C_0, provide a measure of the sensitivity of a DPF radio channel to the level of polarization isolation (XPD) in the radio and antenna system hardware. For the two bit-error rates used here, Figs. 6 and 7 summarize calculations of channel outage statistics as a function of radio path length for several values of C_0 spanning a range of practical concern.

Two reference curves, "A" and "B," are shown in Fig. 6 and on all subsequent figures. Curve "A" represents the multipath fading outage time for a conventional single-polarization radio channel with identical
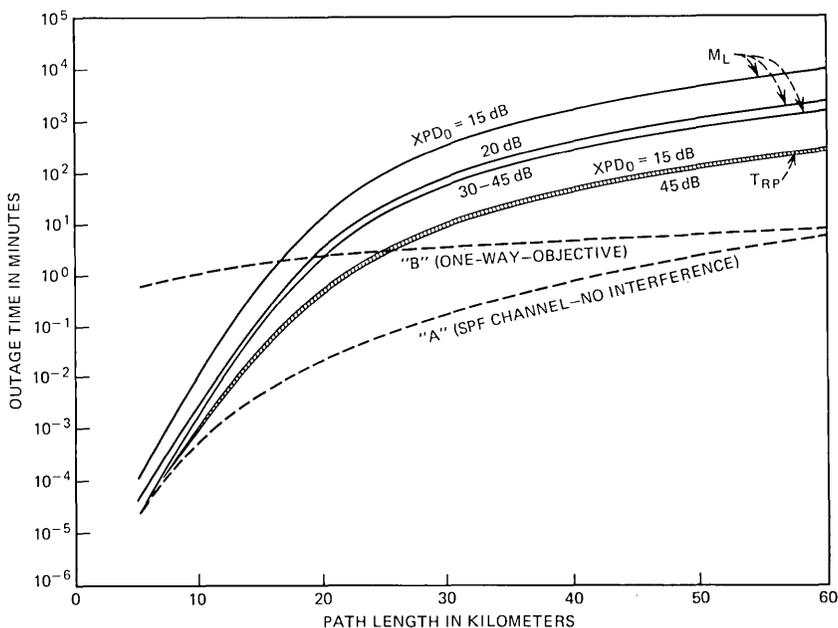
Fig. 6—Outage-time statistics as a function of hardware XPD for $BER_0 = 10^{-6}$.

thermal-noise fade margin and no interference, and curve "B" a per-hop one-way outage-time objective typical of that used in the design of short-haul radio systems. Reference curve "A" allows for comparisons to show the increase in outage probability suffered by a DPF radio channel over a conventional channel as a result of fade-dependent degradation in cross-polarization discrimination.

Outage-objective reference curve "B" shows, by its intersection with a DPF-channel outage curve, an upper limit on radio-path hop length at a given BER_0. This limit on path length is termed a maximum-reliable-path-length (MRPL) and is used here to illustrate implications of outage-time variations. The limit actually imposed on hop length for a DPF channel in practice must be determined by the combined outage probabilities resulting from rain, equipment failures, and multipath fading.[16]

These calculations indicate that for current DPF-channel performance characteristics, channel outage during multipath fading is quite insensitive to the polarization isolation of system hardware as long as a minimum isolation of 20 dB is maintained, i.e., $XPD_0 \geq 20$ dB. This lower bound, to an acceptable range for hardware XPD, can be extended to 15 dB in theory (see T_{RP} curves in Figs. 6 and 7), and can also be extended in practice for the larger bit-error-rate criterion ($BER_0 \leq 10^{-3}$). However, for the measured DPF-channel performance curve at $BER_0 = 10^{-6}$, a
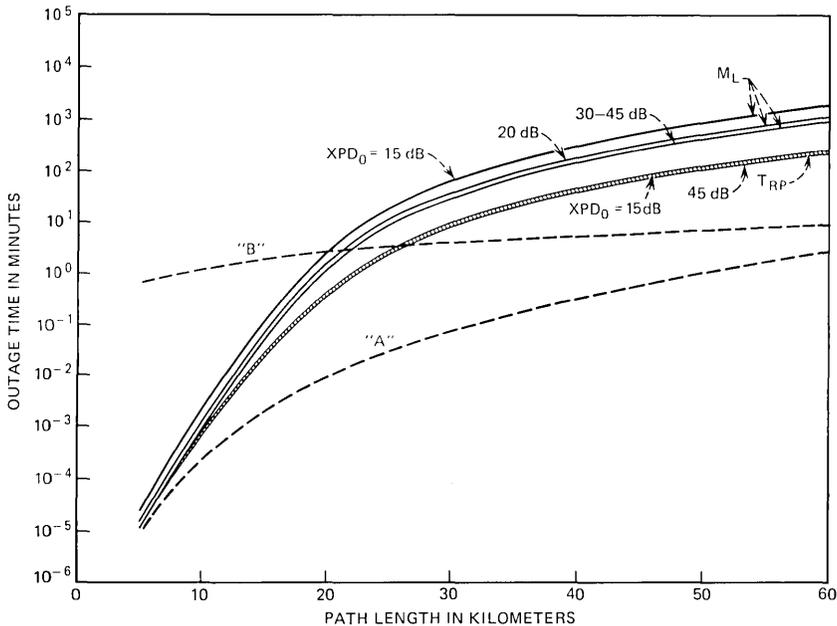
Fig. 7—Outage-time statistics as a function of hardware XPD for $BER_0 = 10^{-3}$.

system hardware $XPD_0 = 15$ dB degrades outage performance by almost an order of magnitude at any particular path length, and decreases the intercept of this outage curve and the one-way-objective curve "B" by 15 to 20 percent.

These results have significance from two standpoints. First, the XPD requirements on system hardware are moderate and are therefore less costly to implement. This is particularly important for the radio antenna system. Secondly, maintenance requirements for XPD characteristics are relatively loose, again producing a positive cost factor from an operational standpoint.

3.4 Numerical results: estimates of outage statistics

As discussed in Section 3.3.2, the outage probability is not sensitive to changes in the value of the parameter C_0. Consequently, all computations are performed for the upper and lower limits on the practical range of interest for C_0, i.e., 20 dB $\leq 20 \log_{10}(C_0^{-1}) \leq 45$ dB. Since these computed results bracket those for all intermediate values of C_0, a strip or band of values is shown, in general, rather than a single curve in the graphical results presented below. Note that in some cases this "band" has, effectively, zero width and appears to be only a single curve.

The simple relationship between the outage time (T_{OUT}) and outage probability (P_{OUT}) for a single radio hop during multipath fading is given

by eq (11) in Section 2.3. All results given below for T_{OUT} are computed according to eq. (11) with an average temperature of $t = 55°F$, which gives a value of $T_0 = 1.467 \times 10^5$ minutes.

Note that all results obtained below are for an "average" multipath environment corresponding to the value $c = 1.0$ in the expression (12) for the multipath occurrence factor. All results can be applied directly to other environments by direct scaling of outage time via a simple multiplication by the appropriate value of c.

3.4.1 Deep-fade restriction on statistical model

Figures 8 and 9 show representative sets of curves that display the dependence of outage probability on the faded signal level L. The curves in Fig. 8 represent the integrand in the expression for P_{OUT} (eq. 2), and are computed for constant path length D, for the value $XPD_0 = 20$ dB. From these curves, it is apparent that with adequate hardware XPD ($XPD_0 \geq 20$ dB), significant contributions to P_{OUT} occur only for fades $L \leq 0.32$ ($-20 \log_{10}(L) \geq 10$ dB). This result supports the adequacy of the statistical model for the signal envelope peak value used here in calculating outage statistics for a DPF channel.
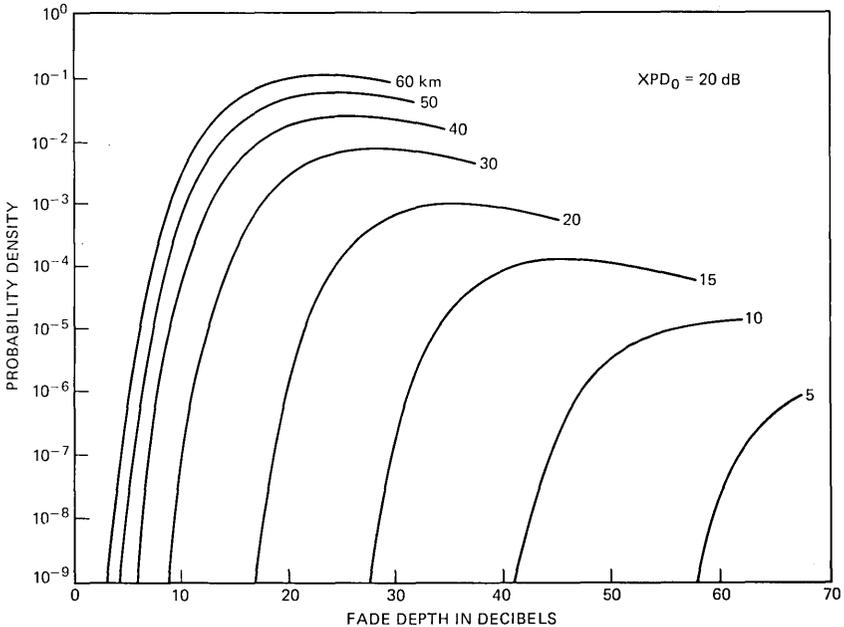
Fig. 8—Functional dependence of $\{P_e (\text{BER} \geq 10^{-6} | L) \, p_F(L)\}$ on fade depth $[-20 \log_{10}(L)]$ for M_L channel characteristics.
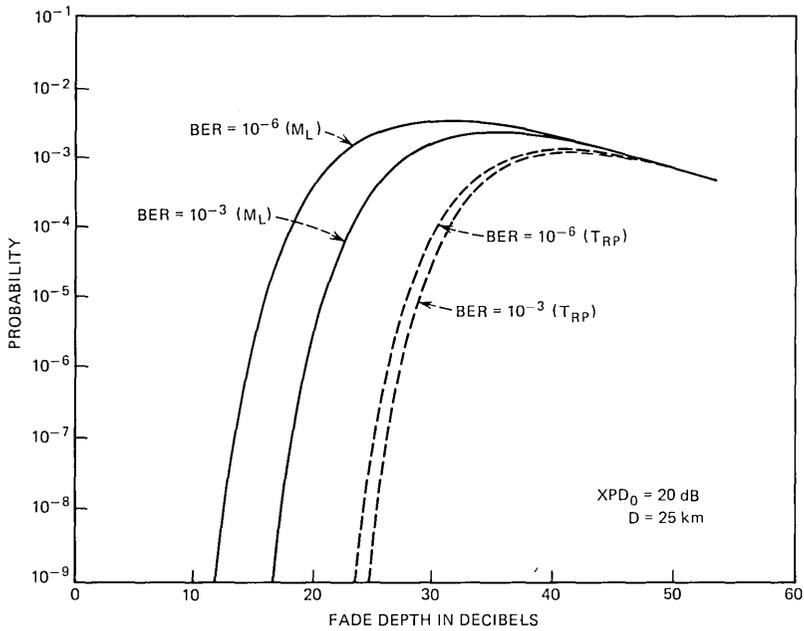
Fig. 9—Functional dependence of $[P_e(\text{BER} \geqq \text{BER}_0 | L)\ p_F(L)]$ on fade depth $[-20\log_{10}(L)]$ for M_L and T_{RP} channel-performance curves.

3.4.2 Path-length dependence of T_{OUT}

Figure 10 summarizes results obtained by computing outage time for average terrain and climate conditions ($c = 1.0$ in eq. 12) on a single hop for a DPF channel. Curves are presented for two bit-error rates: $\text{BER}_0 = 10^{-6}$ and $\text{BER}_0\ 10^{-3}$.

Comparison of the curves computed using theoretical and measured DPF-channel performance curves show the effect of degradations introduced by intersymbol interference and other non-ideal conditions in the digital terminal and associated radio equipment. Most important in this respect is the threshold value CIR_0, which appears to dominate other parameters.

From Figure 10, it is apparent that for the model parameters used here the error-rate criterion used to specify an MRPL (or, for a particular path length of interest, the expected outage time) is not critical. Theoretically, the change in MRPL that results in going from 10^{-3} BER to 10^{-6} BER is less than 5 percent, while for results based on measured DPF-channel performance, the change is only about 10 percent.

At a particular path length, the increase in outage time incurred by going from 10^{-3} to 10^{-6} BER is at most a factor of 2 for results based on either the theoretical or measured DPF-channel performance curves.
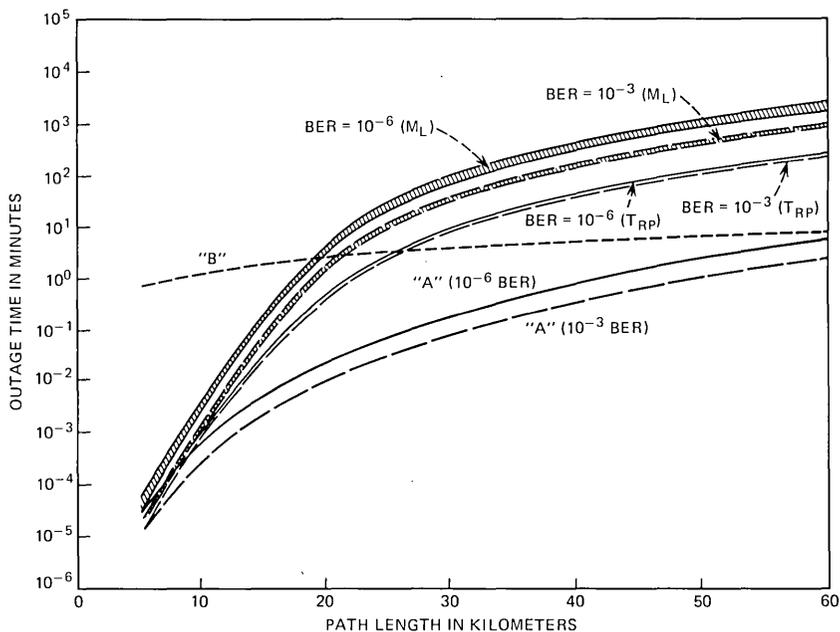
Fig. 10—Outage time vs path length for T_{RP} and M_L channel-performance curves.

The outage time estimates obtained for measured DPF-channel performance result in a net decrease of about 20 to 25 percent in MRPL from those estimates obtained using theoretical performance.

Figure 11 draws a comparison between outage-time estimates obtained under the terrain-scattering hypothesis for two different antenna tower heights for $BER_0 = 10^{-6}$. The difference in the estimated outage time for these two cases decreases with increasing path length, varying from about three orders of magnitude at 10 km to about a factor of 5 at 35 km.

The MRPL decreases by about 60 percent in going from the 250-ft tower to the 75-ft tower, indicating a large increase in cross-polarized energy contributed by terrain scattering (see Fig. 5). Assuming that the terrain-scattering model is valid (or at least dominant), this result suggests that use of large tower heights would be desirable for DPF-channel radio systems.

3.4.3 Dependence of T_{OUT} on system gain

An index of performance for a radio system is the quantity known as system-gain (SG). System gain is defined as the required difference between transmitted and received signal-power levels necessary to achieve some standard level of performance. As above, the two performance standards considered here are $BER_0 = 10^{-3}$ and $BER_0 = 10^{-6}$. Typical
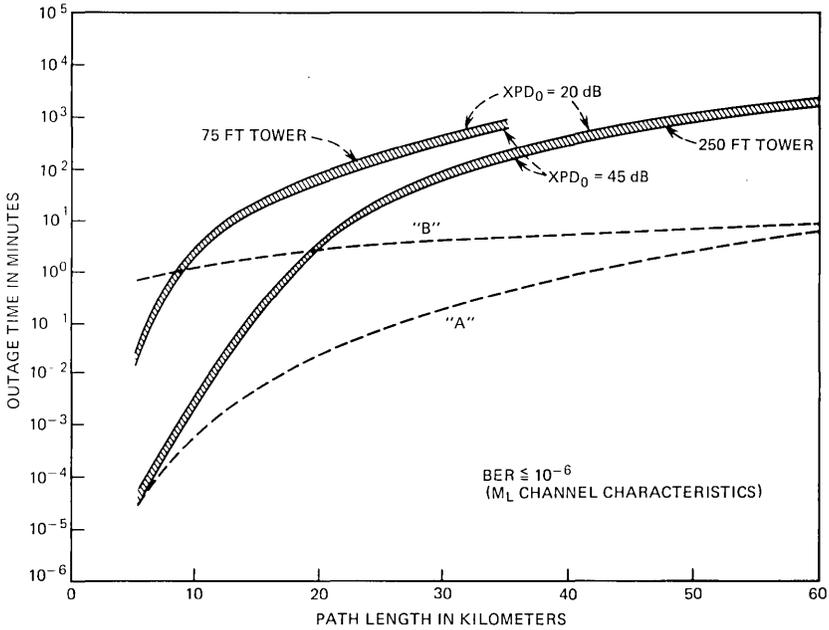
Fig. 11—Outage-time statistics for two antenna tower heights under terrain-scattering hypothesis.

system-gain values for commercially available DPF-channel digital radio systems vary from 95 to 108 dB for a bit-error rate limit of 10^{-6}.

To evaluate the effect of system gain on outage time statistics, calculations of T_{OUT} have been made using SG as a parameter (with XPD_0 = 20 dB). Figures 12 and 13 summarize results obtained for the two bit-error rate standards, with SG varying from 95 to 115 dB in 5 dB steps.

Results for the two different error rates are quite similar. Based on theoretical DPF channel performance, outage-time statistics vary somewhat (less than a factor of 2) as the system gain increases from 95 to 100 dB, but remain virtually unchanged in the range $100 \leq SG \leq 115$ dB.

Using measured DPF-channel performance for the model parameters used here, outage time is almost totally insensitive to system gain for path lengths in excess of 18 to 20 km. For shorter path lengths, some differences in outage time do result as system gain is changed, but since the values of outage time for these short paths are very small in comparison to the system objective, such variations are not of practical concern.

For path lengths in the vicinity of the MRPL, system gain does not appear to affect significantly the outage performance of a DPF channel.
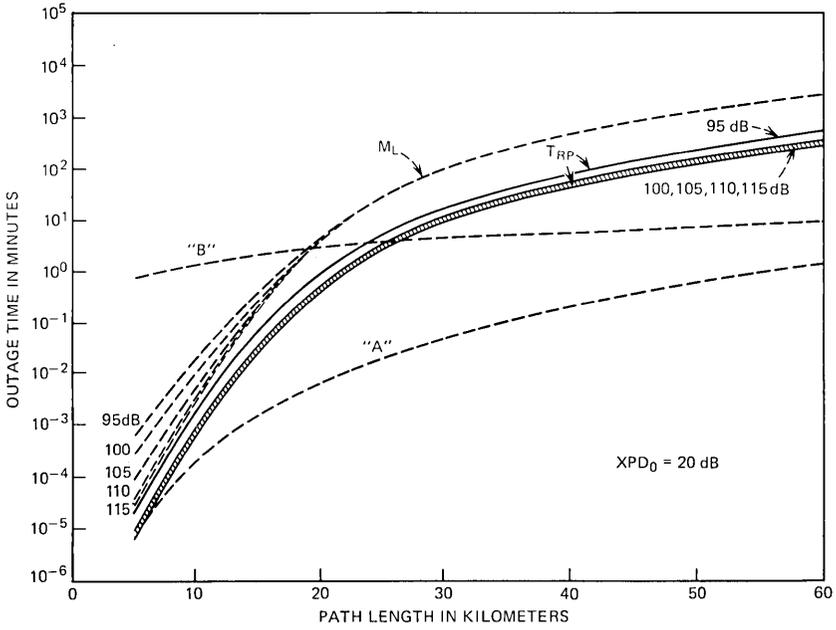
Fig. 12—Outage-time statistics as a function of system gain for DPF radio channel with $\text{BER}_0 = 10^{-6}$.
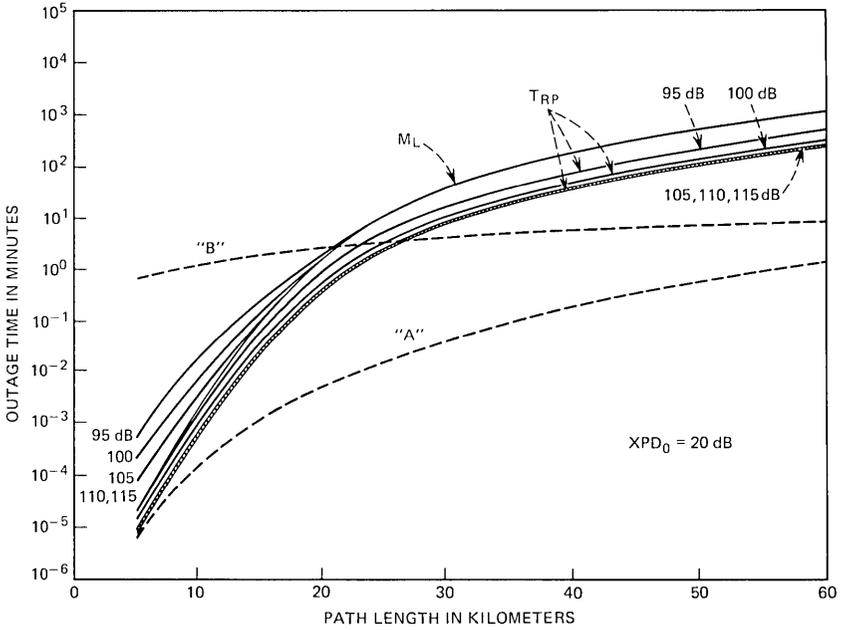
Fig. 13—Outage-time statistics as a function of system gain for a DPF channel with $\text{BER}_0 = 10^{-3}$.
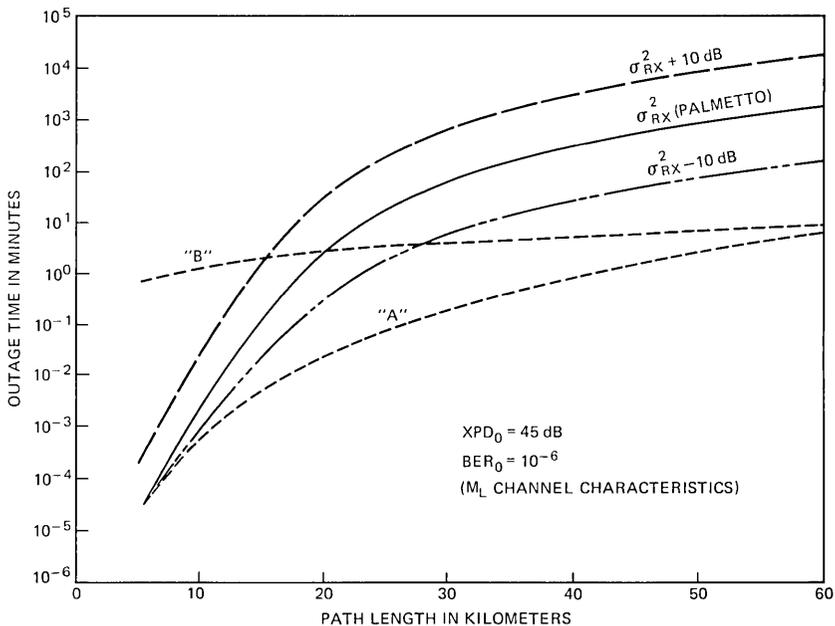
Fig. 14—Sensitivity of outage-time statistics to power level of random component of cross-polarized signal.

3.4.4 Environmental sensitivity of T_{OUT}

As discussed in Section 3.3.1, outage statistics are dependent on propagation path characteristics expressed through the parameter σ_{RX}^2. Various characteristics of the geographic location of a radio propagation path will influence the value of σ_{RX}^2. To explore the sensitivity of results obtained here to change in the magnitude of σ_{RX}^2, calculations for T_{OUT} were carried out with σ_{RX}^2 as a parameter.

Figure 14 summarizes results obtained by computing T_{OUT} as a function of path length for three values of σ_{RX}^2; in each case, σ_{RX}^2 has the same relative variation with D—e.g., the curve shown in Fig. 5 for $h = 250$ ft. These computations show that T_{OUT} varies by about one order of magnitude as σ_{RX}^2 varies by one order of magnitude (10 dB). (The midrange value of σ_{RX}^2 corresponds to that obtained from the Atlanta-to-Palmetto, Georgia data set at 42.5 km.)

These results indicate the importance of obtaining data on σ_{RX}^2 from a variety of locations to establish its range of variability. Depending on the magnitude of such variations, the allowable path length for reliable digital communication via DPF radio-channels could vary significantly with geographic location—e.g., the results shown in Fig. 14 show potential changes of ±25 to ±30 percent as σ_{RX}^2 varies by ±10 dB.

IV. SUMMARY

4.1 Radio-system XPD requirements

An important consideration for the practical implementation of DPF digital radio systems is the degree of isolation required between the two orthogonal, linearly polarized waves transmitted at each frequency. Severe requirements on such isolation (XPD) imply increased system costs resulting both from hardware design and system maintenance.

In general, a minimum requirement of 20 to 25 dB for hardware XPD is not severe from either a design or maintenance standpoint. This level of XPD results in DPF channel error-rate performance well above the minimum service standards used here for unfaded signals (see Fig. 2).

The estimates of outage statistics obtained in this study with a variety of model parameters indicate that outage performance for current DPF channels during multipath fading is not materially affected by the system-hardware XPD as long as it remains better than 20 dB (see Figs. 6, 7, and 10). Multipath-related outages do not, therefore, impose any severe hardware design or maintenance requirements.

4.2 Maximum reliable path length

The example outage estimates computed in the above sections for representative system parameter values suggest that for a 0.02 percent two-way reliability objective for a 400-km route, and with a high antenna tower (250 feet), the maximum reliable path length (MRPL) for a DPF channel is about 21.5 km (13.4 mi) for a bit-error rate of 10^{-3} in the absence of other outage effects, *and for an average multipath environment* (see Fig. 10). This MRPL can vary from 17.5 km (10.9 mi), on Gulf Coast or overwater paths, to 31.0 km (19.3 mi) in dry, mountainous regions (Albuquerque), reflecting changes in the magnitude of the multipath occurrence factor for these environmental extremes.

For a bit-error rate BER $\leq 10^{-6}$, the range of MRPL in these example calculations is 17.0 km (10.6 mi) to 26.0 km (16.2 mi) with a value of 20.0 km (12.4 mi) *for an average multipath environment.*

The rather steep slope of an outage-time curve of Fig. 10 in the vicinity of its intersection with the outage-time objective curve "B" indicates that the changes in MPRL resulting from inclusion of rain and equipment outages will not be severe, i.e., a small decrease in MRPL allows for a substantial contribution to outage time from these other sources of channel outage. Note that the preceding statement does not hold for rain-limited propagation paths (e.g., in the Miami, Florida area) where multipath outages are strongly dominated by rain-induced outages. The actual magnitude of such changes will, however, depend on the rain statistics for a given geographic location.

4.3 Dependence of path length on system gain

For digital radio systems with high system gain, the exact value of system gain, while significant for the average BER performance of a DPF channel, does not significantly affect the MRPL. In the examples considered here, changes on the order of 20 percent in system gain (in dB) result in very small changes in MRPL—i.e., less than 1.0 km.

4.4 General conclusions

The model proposed and developed here for DPF-channel outages during multipath fading, together with results obtained from the example calculations carried out for representative system and environmental parameter values, lead to the following general conclusions:

(*i*) Outage times during multipath fading conditions are between two and three orders of magnitude larger for a DPF channel of current design than for a conventional (single-polarization-per-frequency) radio channel with no interference.

(*ii*) The threshold carrier-to-interference ratio (CIR$_0$) below which a given bit-error rate cannot be obtained, independent of the signal-to-noise (thermal) ratio, is the channel performance characteristic of greatest importance to multipath outages. Up to an order of magnitude improvement in outage time could be obtained if a near theoretical value of CIR$_0$ could be achieved through hardware/design improvements.

(*iii*) Multipath outage time is dependent on geographic and environmental conditions expressed through the multipath occurrence factor r and the statistical parameter σ_{RX}^2. Outage-time estimates computed here scale almost linearly with changes in σ_{RX}^2, other model parameters remaining approximately constant.

(*iv*) Under the hypothesis that terrain scattering is the dominant mechanism for the random component of the interference signal in a DPF channel, multipath outage time for such channels can be minimized through use of high antenna towers and, when possible, through judicious choice of radio-hop terrain.

APPENDIX A

Symbol Table

$\alpha(L) = C_0 L / \sigma_{RX}$.

$\beta(L) = V_M(L) / \sigma_{RX}$.

σ_{RX} = Standard deviation of random component of cross-polarized signal.

σ_N = Normalized, cross-polarizing, bistatic, terrain-scattering cross section.

BER = Bit-error rate.

BER_0 = Fixed maximum value for BER.

c = Geographic multipath-fading normalization factor.

C_0 = Ratio of cross-polarized to copolarized signal voltages during periods of no multipath fading—i.e., $C_0 = 10^{-(XPD_0/20)}$.

CIR = Carrier-to-interference ratio in dB for single interferer arising from cross-polarized DPF-channel signal.

CNR = Carrier-to-(thermal)-noise ratio in dB.

CNR_0 = Nominal unfaded design value of CNR.

CIR_0 = Value of CIR below which BER_0 is exceeded independent of CNR.

CNR_{th} = thermal-noise threshold value of CNR at a given BER_0 with no interference present.

D = Path length (miles or kilometers).

f = Frequency in GHz.

FD = Channel fade depth $[-20 \log_{10}(L)]$.

FM = System thermal-noise fade margin.

h = Tower height for radio antennas.

I_0 = Modified Bessel function of order zero.

$\{K_1,$ = Empirical fitting parameters for DPF-channel performance
$K_2,$ curves.
$K_3\}$

L = Ratio of faded to unfaded peak envelope voltage.

L_m = Minimum allowable normalized amplitude of the received-signal envelope (for a path length D) required to maintain some specified maximum channel BER_0.

MRPL = Maximum-reliable-path-length. (The path length determined by the intersection of an outage time curve and the system outage-time-objective curve).

M_L = DPF-channel performance characteristics derived from laboratory measurements on 3A-RDS.

P_e = Conditional probability of DPF-channel BER exceeding some specified value BER_0 at a given, fixed multipath fade level L.

$p_F(L)$ = Probability density function of multipath fading signal.

P_F = Probability distribution function of multipath fading signal.

P_{out} = Probability of DPF-channel BER exceeding some specified value BER_0 during periods of multipath fading.

P_r = Receiving antenna radiation pattern.

P_t = Transmitting antenna radiation pattern.

p_{VX} = Conditional probability density of cross-polarized interference envelope voltage.

r = Multipath occurrence factor.

R_1 = Distance from transmitting antenna to terrain-scatter point.

R_2 = Distance from terrain-scatter point to receiving antenna.

SG = System-gain (the difference between transmitted and received signal power levels necessary to achieve some standard level of performance, i.e., $BER = BER_0$).

t = Average temperature in °F.

T_0 = Base time period for multipath fading activity during one year.

T_{out} = Total time (in minutes per year) during which the DPF-channel BER can be expected to exceed some BER_0 as a result of multipath fading activity.

T_{RP} = DPF-channel performance characteristics derived from theoretical results given by Rosenbaum[10] and Prabhu.[11]

V = Peak envelope voltage.

V_M = Maximum normalized value of V_{TX} for which $BER \leq BER_0$.

$\langle v_{RS}^2 \rangle$ = Energy received as a result of terrain scattering, including conversion from one linear polarization to the orthogonal linear polarization.

V_{DX} = Peak envelope voltage of direct component of cross-polarized signal.

V_{RX} = Peak envelope voltage of random component of cross-polarized signal.

V_c = Peak envelope voltage of copolarized signal.

V_{TX} = Peak envelope voltage of total cross-polarized signal.

XPD = Cross-polarization-discrimination ratio (ratio of copolarized to cross-polarized received signal in dB).

XPD_0 = Cross-polarization discrimination ratio during period of no multipath fading.

APPENDIX B

B.1 Estimation of σ_{RX}^2 Based on Terrain-Scattering Hypothesis

Terrain scattering along the radio-hop propagation path is one physical phenomenon that can contribute energy to the random component of the cross-polarized interference signal in a DPF radio channel.
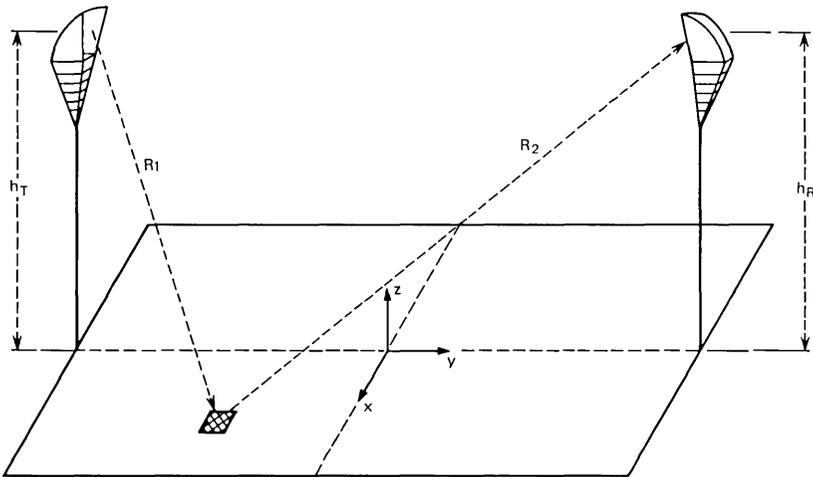
Fig. 15—Idealized radio-hop geometry used in terrain-scattering computation.

The magnitude of this component is not easily determined experimentally or theoretically, but an estimate can be obtained through an idealized numerical calculation scaled to available data.

Figure 15 shows the geometry assumed as a basis for calculating the energy contributed to the cross-polarized-interference signal by terrain scattering. The expression used to compute this random-scatter component is

$$\langle v_{RS}^2 \rangle = \int\int dx dy \left\{ \sigma_N(x,y) \frac{P_t(x,y)}{R_1^2} \frac{P_r(x,y)}{R_2^2} \right\}, \tag{17}$$

where $\sigma_N(x,y)$ is a factor proportional to the cross-polarized, bistatic scattering cross section of the ground and various normalization factors. Interest here is centered on the functional dependence of $\langle v_{RS}^2 \rangle$ on radio-hop path length, rather than the absolute magnitude of this quantity. The functions P_t and P_r are the antenna response patterns at the transmitter and receiver, respectively.

Given this restricted interest and the very sharply directive radiation pattern of typical transmitting and receiving antennas at 11 GHz (a pyramidal horn-reflector-antenna response pattern was used for calculations performed here), the variation of $\sigma_N(x,y)$ with location is neglected and assumed to have unity value. The area over which the double integration shown in eq. (17) is performed is limited by an error criterion for the incremental contributions that occur during evaluation of this expression. The error criterion used corresponds to neglecting contributions from areas illuminated by very low-level side-lobes of the transmitting antenna radiation pattern, and sensed by the low-level side-lobe region of the receiving antenna response pattern.

Normalizing the computed estimates to the experimental values obtained at Palmetto, Georgia (see Section 3.3.1), the normalized scattering cross section has a value of about -28 dB, which agrees within a few dB with radar cross-section measurements on various types of terrain.[14,15]

The results of the calculations of scattered cross-polarized energy are summarized in Fig. 5 for three different tower heights. These curves display an approximate path-length dependence varying from D^2 for long paths to D^6 for short paths.

REFERENCES

1. W. T. Barnett, "Multipath Propagation at 4, 6, and 11 GHz," B.S.T.J., *51,* No. 2 (February 1972), pp. 321–361.
2. W. T. Barnett, "Determination of Cross-Polarization Discrimination During Rain and Multipath Fading at 4 GHz," IEEE 1974 International Conference on Communications, June 17–19, 1974, Minneapolis, Minnesota, Conference Record, pp. 12D-1–12D-4, IEEE Cat. No. 74, CHO 859-9-CSCB.
3. G. LeFrancois, L. Martin, and M. Rooryck, "Influence of the Propagation on the Value of the Decoupling of Two Orthogonal Polarizations," Ann. Telecommun., 28 (1973) pp. 316–324.
4. S. H. Lin, "Impact of Microwave Depolarization During Multipath Fading on Digital Radio Performance," B.S.T.J., this issue, pp. 645–674.
5. S. O. Rice, "Mathematical Analyses of Random Noise," B.S.T.J., *23,* No. 3 (July 1944), pp. 282–332, and *24,* No. 1 (January 1945), pp. 46–256.
6. M. Nakagami, "The M-Distribution—A General Formula of Intensity Distribution of Rapid Fading," in W. C. Hoffman (Ed.), *Statistical Methods in Radio Wave Propagation,* Oxford: Pergaman, 1960.
7. P. Beckmann, *Probability in Communication Engineering,* New York: Harcourt, Brace and World, 1967.
8. S. H. Lin, "Statistical Behavior of a Fading Signal," B.S.T.J., *50,* No. 10 (December 1971), pp 3211–3270.
9. W. B. Davenport and W. L. Root, *An Introduction to the Theory of Random Signals and Noise,* New York: McGraw-Hill, 1968, pp. 165–166.
10. A. S. Rosenbaum, "PSK Error Performance with Gaussian Noise and Interference," B.S.T.J., *48,* No. 2 (February 1969), pp. 413–442.
11. V. K. Prabhu, "Error Rate Considerations for Coherent Phase-Shift Keyed Systems with Co-Channel Interference," B.S.T.J., 48, No. 3 (March 1969), pp. 743–767.
12. A. Vigants, "Space Diversity Engineering," B.S.T.J., *54,* No. 1 (January 1975), pp. 103–142.
13. K. A. Norton, L. E. Vogler, W. V. Mansfield, and P. J. Short, "The Probability Distribution of the Amplitude of a Constant Vector Plus a Rayleigh Distributed Vector," Proc. IRE, *43* (October 1955), pp. 1354–1361.
14. M. W. Long, *Radar Reflectivity of Land and Sea,* Lexington Books, Lexington, Mass.: D. C. Heath, 1975.
15. C. R. Grant and B. S. Yaplee, "Back Scattering from Water and Land at Centimeter and Millimeter Wavelengths," Proc. IRE (London), *47,* No. 7 (July 1957), pp. 976–982.
16. M. S. Levetin, "A New Approach to Setting Reliability Objectives for Microwave Radio Systems," IEEE International Conference on Communications, June 14–16, 1976, Philadelphia, Pa.

Loss Analysis of Single-Mode Fiber Splices

By D. MARCUSE

This paper analyses losses caused by the misalignment of two fibers joined in a splice. We consider the possibility that the two fibers of different dimensions are separated in longitudinal direction and are tilted or offset with respect to each other. Central to our discussion is the observation that the modes of single-mode fibers are very nearly gaussian in shape regardless of the fiber type—step-index or graded-index. The splice losses are thus related to the corresponding losses of gaussian beams. We specify the relation between the actual mode field and the gaussian beam that matches this field optimally. The trade-off between slice tolerances with respect to tilt and offset is expressed as an "uncertainty principle." Because of the near-gaussian nature of single-mode fiber fields, our results are immediately applicable to the excitation of single-mode fibers by gaussian-shaped laser beams.

I. INTRODUCTION

Light transmission losses of single-mode fiber splices depend on the alignment accuracy of the fiber ends relative to each other.[1] We assume that the fibers are immersed in index-matching fluid to minimize reflection losses at the fiber ends. Most troublesome are transverse misalignments (offsets) and angular misalignments (tilts). Fiber splices are surprisingly tolerant of longitudinal misalignment.

We begin our discussion by showing that the fields of single-mode, step-index fibers are very nearly gaussian in shape. This observation holds with even more assurance for parabolic-index fibers, because the modes of the infinitely extended parabolic-index medium are themselves gaussian and are changed only slightly by the truncation of the index profile at the core boundary. Once it is established that fiber modes may be closely approximated by gaussian field distributions, the evaluation of splice losses reduces to the computation of transmission losses between misaligned gaussian beams.[2,3]

We present formulas for relating the width of the gaussian field distribution to the fiber parameters. An implicit relation for all types of

fibers is given and explicit formulas are derived for the important cases of step-index and parabolic-index fibers. Next, we present simple analytical expressions for the transmission coefficient of fiber splices for the case of longitudinal, transverse, and angular misalignment for fibers of different dimensions.

Transverse splice tolerances become less stringent for fibers whose mode fields extend farther in transverse directions. Wide mode fields can be obtained by selecting a core index that is very nearly equal to the refractive index of the cladding. However, a wide fiber mode is less tolerant of angular misalignments. The relative tolerance of fiber splices with respect to offsets and tilts is expressed as an "uncertainty principle."

We limit our discussion to "weakly guiding" fibers[4] defined by the relation $n_1/n_2 - 1 \ll 1$, where n_1 is the maximum value of the refractive index of the fiber core and n_2 is the value of the cladding index. The transmission coefficient of weakly guiding fiber modes can be obtained by matching only the transverse components of the electric field vector of the two modes; their transverse magnetic field components are automatically matched approximately. We designate the electric field vectors of the modes (guided and radiation modes) of the fiber by the symbol $\mathbf{E}_\nu$. The incident electric field $\mathbf{E}$ at the input end of the fiber can then be expressed in terms of fiber modes as follows[5]:

$$\mathbf{E} = \sum_\nu c_\nu \mathbf{E}_\nu. \tag{1}$$

The summation symbol indicates symbolically summation over guided modes (only one for single-mode fibers) and integration over radiation modes. The symbol ν labels the modes (we use $\nu = 0$ as the label of the guided mode of the single-mode fiber). Mode orthogonality allows us to obtain c_0 from (1),

$$c_0 = \frac{1}{2P} \int_0^{2\pi} d\phi \int_0^\infty (\mathbf{E} \times \mathbf{H}_0) \, \mathbf{e}_z \, rdr. \tag{2}$$

$\mathbf{H}_0$ is the magnetic-field vector of the guided mode, $\mathbf{e}_z$ is a unit vector in the direction of the fiber axis, and r and ϕ are cylindrical coordinates in the plane at right angles to the axis of the fiber. We assume that the fiber is receiving radiation from either an input fiber at a splice or from a free-space gaussian laser beam. Correspondingly, $\mathbf{E}$ represents the field that the first fiber of the splice generates at the input end of the second fiber or, alternatively, the gaussian beam mode of a laser.

The power transmission coefficient, finally, is obtained from (2) by the relation,

$$T = |c_0|^2. \tag{3}$$

II. REPRESENTATION OF THE FIBER MODE AS A GAUSSIAN BEAM

The guided modes of weakly guiding fibers are very nearly transverse and linearly polarized.[4] The electric field vector of the input field consists likewise of one dominant transverse component.[3] Let us assume that the input field is gaussian

$$E_y = [4\sqrt{\mu_0/\epsilon_0}\, P/\pi n_2 w^2]^{1/2} \exp\left(-\frac{r^2}{w^2}\right) e^{-i\beta z}. \qquad (4)$$

The refractive index n_2 equals the cladding index of the fiber, P is the power carried by the field and is identical to the P parameter in (2), w is the width parameter of the gaussian field, β is its propagation constant, μ_0 and ϵ_0 are the magnetic susceptibility and the dielectric permittivity of vacuum.

We wish to compare the gaussian field to the mode of the step-index fiber,[5]

$$\mathbf{H}_{x0} = -\sqrt{\frac{2}{\pi}} \left(\frac{\epsilon_0}{\mu_0}\right)^{1/4} \frac{W}{aVJ_1(U)} \sqrt{n_2 P} \begin{cases} J_0\left(U\dfrac{r}{a}\right) & r \leqq a \\[2mm] \dfrac{J_0(U)}{K_0(W)} K_0\left(W\dfrac{r}{a}\right) & r \geqq a. \end{cases} \qquad (5)$$

The P parameter is identical to those in (2) and (4), W and U are related to the important V parameter by the equation,

$$U^2 + W^2 = V^2 = (n_1^2 - n_2^2)k^2 a^2. \qquad (6)$$

The free space propagation constant of plane waves is $k = 2\pi/\lambda$ and a is the core radius of the fiber. J_0 and J_1 are Bessel functions and K_0 is the modified Hankel function. The parameter U can be related to the propagation constant β_s [omitted from (5)] as follows:

$$U = (n_1^2 k^2 - \beta_s^2)^{1/2} a. \qquad (7)$$

By substitution of (4) and (5) into (2) and (3), we obtain the transmission coefficient of a gaussian input beam exciting the HE_{11} mode of a fiber. The r-integral in (2) must be evaluated numerically. It is clear that the value of T depends on the width parameter w of the gaussian beam; T assumes a maximum as a function of w. The maximum value of T is plotted as a function of V in Fig. 1. It is remarkable how closely T approaches unity over the range of V-values shown in the figure. At the important point $V = 2.4$, we have $T = 0.9965$. $V = 2.4$ is close to the largest value at which the fiber supports only one mode. The next higher mode comes in at $V = 2.405$. It is apparent that at $V = 2.4$ the field distribution of the fiber mode matches the gaussian field almost perfectly. The best match is achievable at $V = 2.8$; T decreases very slowly for larger values of V. For smaller V-values, the decrease and consequently
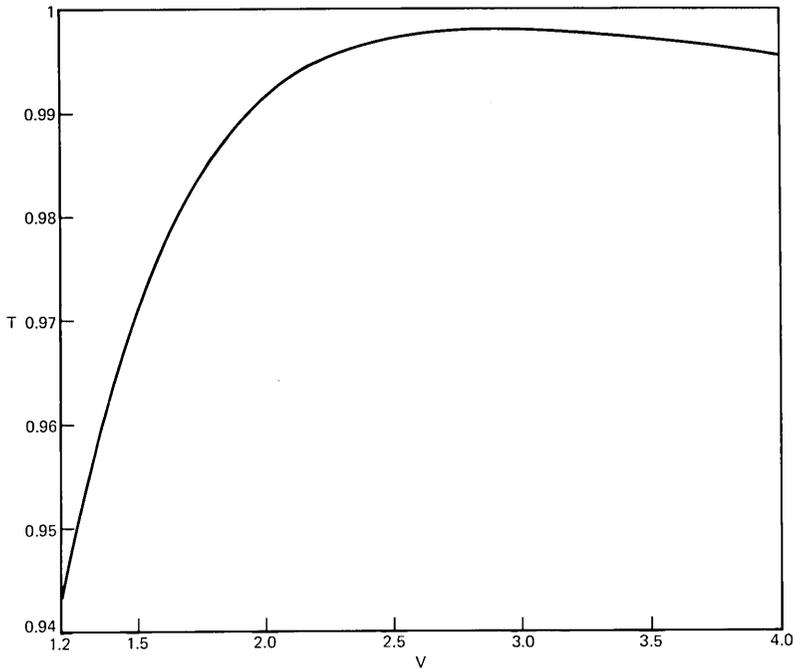
Fig. 1—Maximum value of transmission coefficient between a perfectly aligned and optimally adjusted gaussian beam and a single-mode step-index fiber.

the mismatch between gaussian field and fiber mode is more pronounced; but even at $V = 1.2$, we have $T = 0.946$, a value very close to unity so that, even for such small values of V, the gaussian beam is a reasonably good approximation of the fiber mode. It can be shown that the optimum value of w divided by the core radius is only a function of V. Figures 2 and 3 show the optimum values of w/a as a function of V as the solid line. This function can be approximated very closely (to within a fraction of 1 percent) by the empirical formula,

$$\frac{w}{a} = 0.65 + \frac{1.619}{V^{3/2}} + \frac{2.879}{V^6}. \tag{8}$$

This equation holds, of course, only for step-index fibers. The meaning of the dotted curve in Fig. 2 will be explained later.

It is desirable to have similar relations for graded-index fibers because this would enable us to predict their splice losses. The fields of general graded-index fibers are not known explicitly, so that we cannot use eqs. (2) through (4) to optimize the width of the corresponding gaussian beams. However, we can use a different approach. If we insist that (4) should be used to approximate the guided mode of a single-mode fiber with refractive index distribution $n(r)$ for $r < a$ and $n(r) = n_2$ for $r > a$,
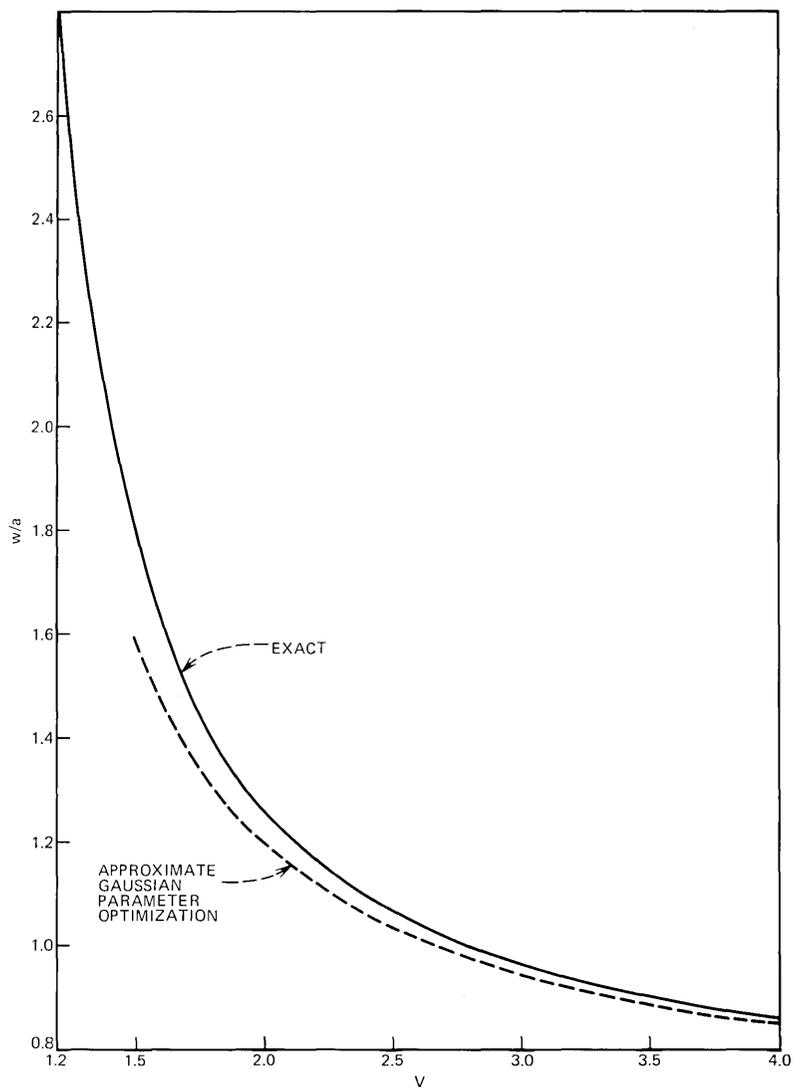
Fig. 2—Normalized optimum width parameter w/a as a function of the V-parameter for step-index fibers. The dotted line is obtained from the approximate procedure expressed in eqs. (19) and (20).

we may substitute (4) into the wave equation

$$\frac{d^2E_y}{dr^2} + \frac{1}{r}\frac{dE_y}{dr} + [n^2(r)k^2 - \beta^2]E_y = 0, \tag{9}$$

and obtain

$$\left\{\frac{4}{w^2}\left(\frac{r^2}{w^2} - 1\right) + n^2(r)k^2 - \beta^2\right\}E_y = 0. \tag{10}$$
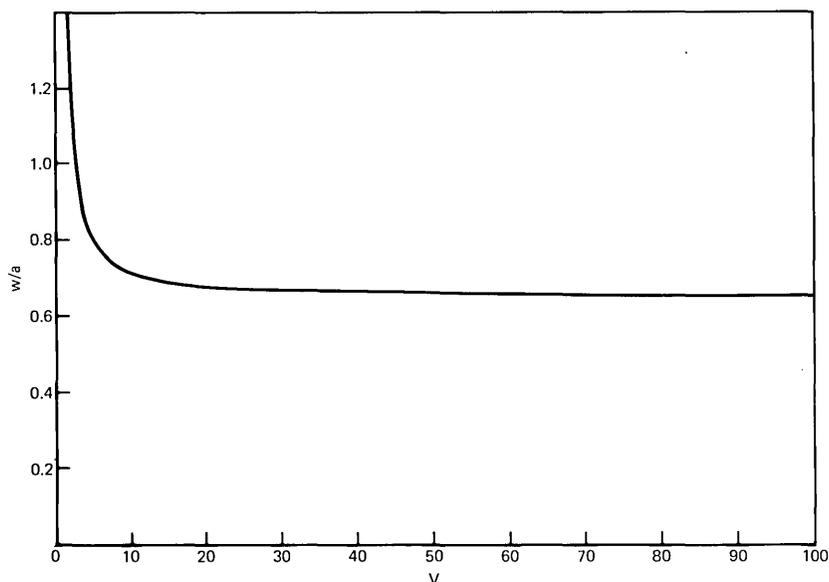
Fig. 3—Same as Fig. 2 for wider range of V-values.

For a graded-index distribution,

$$n(r) = n_1 \left[1 - \left(\frac{r}{a}\right)^g \Delta \right] \tag{11}$$

with $g = 2$, (10) can be satisfied exactly. Note that (11) is an infinitely extended parabolic-index profile. In this case, we find

$$w = \sqrt{\frac{2}{V}}\, a \tag{12}$$

and

$$\beta = [n_1^2 k^2 - 2V/a^2]^{1/2}. \tag{13}$$

We define the V-parameter for any value of g by the equation

$$V = n_1 k a \sqrt{2\Delta}. \tag{14}$$

This expression is also a good approximation of (6), if we use

$$\Delta = 1 - \frac{n_2}{n_1} \ll 1. \tag{15}$$

Equations (12) and (13) are not correct for actual parabolic-index fibers whose refractive index distributions are given by (11) (with $g = 2$) only for $r < a$, but assume the form $n(r) = n_2$ for $r > a$. We refer to profiles of this kind as truncated index distributions.

It appears reasonable to attempt an approximate evaluation of w and β by squaring (10) and integrating over the entire infinite cross-section. The parameter optimization is then achieved by requiring that the expression so obtained assumes a minimum value. After dividing by the constant 2π, we obtain

$$J = \int_0^\infty \left[\frac{4}{w^2} \left(\frac{r^2}{w^2} - 1 \right) + n^2(r)k^2 - \beta^2 \right]^2 E_y^2 \, r dr = \text{min.} \quad (16)$$

This empirical extremum principal leads to two equations:

$$\frac{\partial J}{\partial w} = 0, \quad (17)$$

and

$$\frac{\partial J}{\partial \beta} = 0. \quad (18)$$

It is actually the expression in brackets that should vanish if the equation could be satisfied exactly. E_y^2 serves only the purpose of a weighting function. For this reason, we consider the width parameter of E_y as constant, unaffected by the differentiation in (17). (It was found empirically that this procedure leads to more accurate results.) Substitution of (4) into (16) allows us to obtain from (17) and (18)

$$\beta^2 = \left\{ \frac{4k^2}{w^2} \int_0^\infty n^2(r) \exp\left(-2r^2/w^2\right) r dr \right\} - \frac{2}{w^2} \quad (19)$$

and

$$\int_0^\infty \left[\frac{4}{w^2} \left(\frac{r^2}{w^2} - 1 \right) + n^2(r)k^2 - \beta^2 \right] \left(2\frac{r^2}{w^2} - 1 \right)$$
$$\exp\left(-2r^2/w^2\right) r dr = 0. \quad (20)$$

Equations (19) and (20) must be solved simultaneously. Analytical solutions are impossible to obtain so that we resort to numerical solutions.

It might be expected that the optimization procedure works best for parabolic-index fibers, since it yields the exact result (12) and (13) for infinitely extended parabolic index profiles. We expect that the step-index profile presents the worst possible case. It is also a member of the class of profiles given by (11) for $r < a$ and by $n(r) = n_2$ for $r > a$ and is obtained in the limit $g \to \infty$. The result of solving (19) and (20) for $g = 100$ is shown as the dotted line in Fig. 2. At $V = 2.4$ the approximate optimization procedure is in error by 3 percent. The percentage error decreases for larger values of V (but for very large values of V, the agreement becomes poorer once more). At $V = 1.5$ we have an error of 14 percent; smaller values of V are of little practical interest. We see that the method works surprisingly well for the step-index fiber. Comparisons
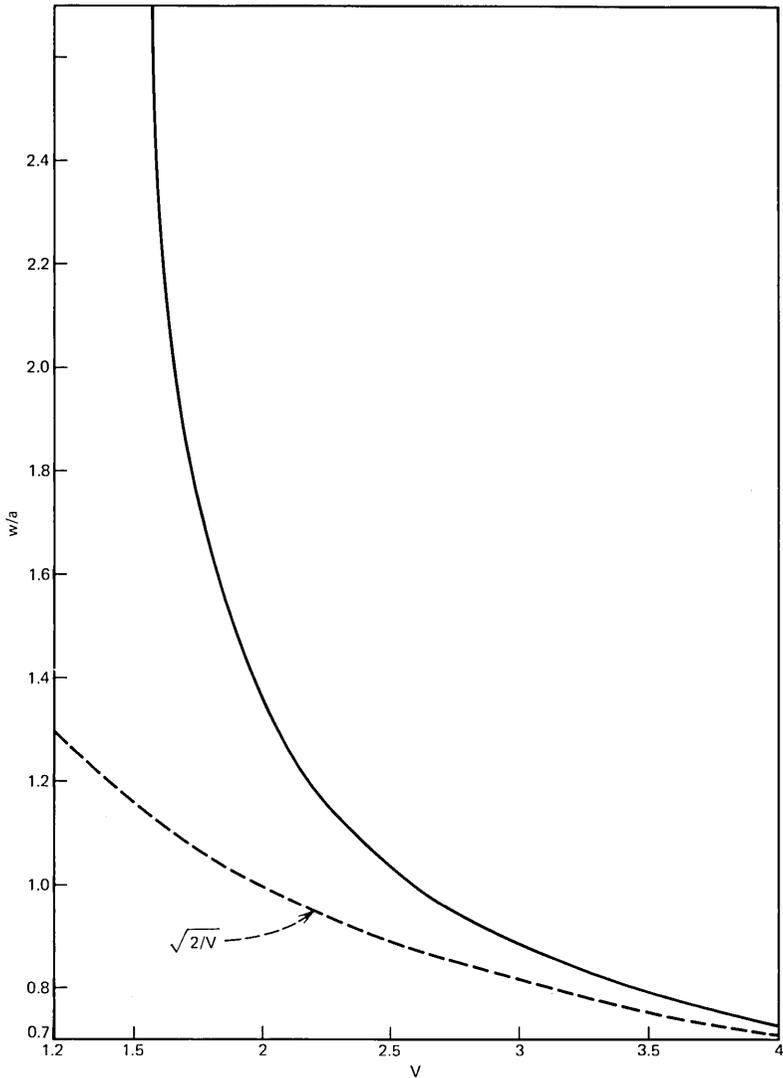
Fig. 4—Normalized optimum beam width w as a function of V for the parabolic-index fiber. The dotted line applies to an infinitely extended parabolic-index profile.

of the accuracy of the method for other values of g are harder to make since the exact field distributions are hard to obtain; no attempts were made to evaluate the performance of the parameter optimization procedure for smaller values of g. However, there is little doubt that the results for values near $g = 2$ will be much better than the comparison shown in Fig. 2.

Figure 4 shows the numerical solution of (19) and (20) for the trun-
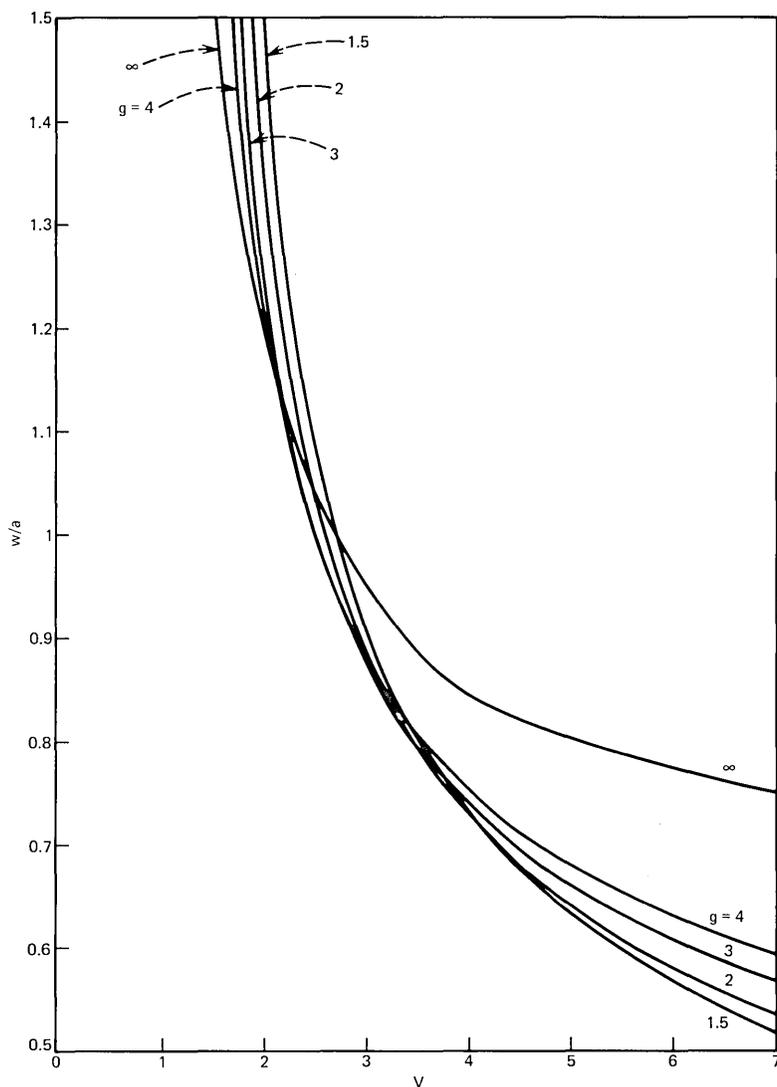
Fig. 5—w/a as a function of V for several values of the power law parameter g defined by (11).

cated parabolic-index profile—that is, for $n(r)$ given by (11) with $g = 2$ for $r < a$ and by $n(r) = n_2$ for $r > a$. The dotted line shown in Fig. 4 applies to the infinitely extended parabolic-index profile and represents the solution (12). It is clear that the field distribution of the truncated parabolic-index profile is wider than the field of the infinitely extended profile because that part of the field that extends into the cladding is no longer under the focusing influence of the graded-index distribution.
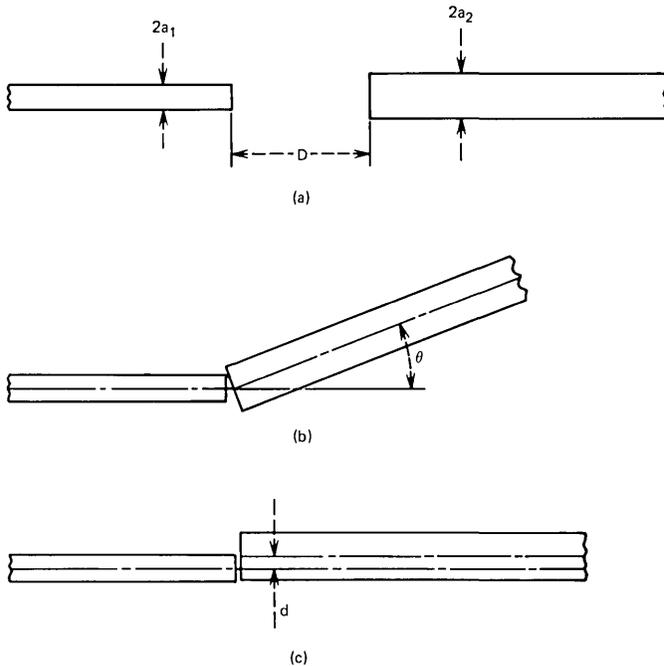
Fig. 6—Several types of splice imperfections.

For large V values both curves coincide; at $V = 5$ the difference of the two curves is already reduced to 1 percent. The solid line in Fig. 4 can be expressed by the empirical approximation

$$\frac{w}{a} = \sqrt{\frac{2}{V}} + \frac{0.23}{V^{3/2}} + \frac{18.01}{V^6}. \qquad (21)$$

This equation gives the optimum width of the gaussian field profile that best approximates the actual field distribution of a parabolic-index fiber.

Figure 6 shows the optimum width w/a computed from (19) and (20) for truncated graded-index profiles for several values of the exponent g appearing in (11).

III. SPLICE LOSSES

Henceforth, we represent the fields of single-mode fibers by gaussian field distributions of the form (4) keeping in mind that the optimum width parameters w of the gaussian can be obtained as solutions of (19) and (20) for general graded-index fibers, or, explicitly, by (8) or (21) for step-index and parabolic-index fibers.

The different types of splice defects are shown in Fig. 6. We allow both fibers joined by the splice to have different parameters shown as different

core diameters in Fig. 6. The actual differences may consist of different refractive index distributions as well as different core diameters. For our analysis, each fiber is represented by the width parameter of the optimum gaussian field distribution, w_1 belongs to the fiber with radius a_1, and w_2 belongs to the fiber with a_2.

.   The relevant formulas can all be derived by using (2) and (3) with the fields of both fibers represented by gaussian field distributions of the form (4). The E field in (3) is understood to be the gaussian field of the first fiber transformed to the input plane of the second fiber. Such mode matching calculations involving gaussian fields are not new, we present here only the results.[2,3]

3.1 Longitudinal fiber separation

For the splice shown in Fig. 6a, we find the power transmission coefficient

$$T = \frac{4\left[4Z^2 + \dfrac{w_1^2}{w_2^2}\right]}{\left[4Z^2 + \dfrac{w_2^2 + w_1^2}{w_2^2}\right]^2 + 4Z^2\dfrac{w_2^2}{w_1^2}} . \tag{22}$$

The normalized fiber separation distance is defined as

$$Z = \frac{D}{n_2 k w_1 w_2} . \tag{23}$$

Two special cases are of interest. At $D = 0$, we have

$$T_0 = \left(\frac{2w_1 w_2}{w_1^2 + w_2^2}\right)^2 . \tag{24}$$

For $D \to \infty$, we obtain asymptotically

$$T_\infty = \frac{1}{Z^2} = \left(\frac{n_2 k w_1 w_2}{D}\right)^2 . \tag{25}$$

3.2 Splices with tilt

For the fiber tilt shown in Fig. 6b, we obtain the power transmission coefficient

$$T = \left(\frac{2w_1 w_2}{w_1^2 + w_2^2}\right)^2 \exp\left[-\frac{2(\pi n_2 w_1 w_2 \theta)^2}{(w_1^2 + w_2^2)\lambda^2}\right] . \tag{26}$$

When the tilt angle θ becomes large enough to make the exponent of the exponential function in (26) unity, the transmitted power decreases to

$1/e$ of its maximum value. This angle is given by the expression,

$$\theta_e = \left(\frac{w_1^2 + w_2^2}{2}\right)^{1/2} \frac{\lambda}{\pi n_2 w_1 w_2}. \tag{27}$$

3.3 Splices with fiber offset

The power transmission coefficient through the fiber splice shown in Fig. 6c assumes the form

$$T = \left(\frac{2w_1 w_2}{w_1^2 + w_2^2}\right)^2 \exp\left[-\frac{2d^2}{w_1^2 + w_2^2}\right]. \tag{28}$$

The amount of offset that reduces the transmitted power to $1/e$ of its maximum value can be defined as

$$d_e = \left(\frac{w_1^2 + w_2^2}{2}\right)^{1/2}. \tag{29}$$

For identical fibers with $w_1 = w_2$, we obtain a very useful and interesting relation by combining (27) and (29)

$$d_e \theta_e = \frac{\lambda}{n_2 \pi}. \tag{30}$$

This expression is reminiscent of the uncertainty principle of quantum mechanics, because it states that as one of two variables becomes smaller, the other must become larger. If a single-mode fiber is designed with a small value of Δ, to allow the field to spread out in transverse direction, w becomes large and, consequently, d_e may be large indicating that a large offset can be tolerated. Equation (30) states that for large values of d_e the tilt angle tolerance decreases. A fiber that is tolerant of large offsets is intolerant with respect to tilts and vice versa.

IV. DISCUSSION AND NUMERICAL EXAMPLES

Throughout our discussion, we are using the width parameter w of the field distribution (4). Experimental observations of the light field of a single-mode fiber detect the light power instead of the field intensity. The power may also be approximated by a gaussian distribution of the form

$$P = P_0 \exp(-r^2/w_p^2). \tag{31}$$

The power width parameter w_p is related to the field intensity width parameter w by the expression

$$w_p = \frac{w}{\sqrt{2}}. \tag{32}$$
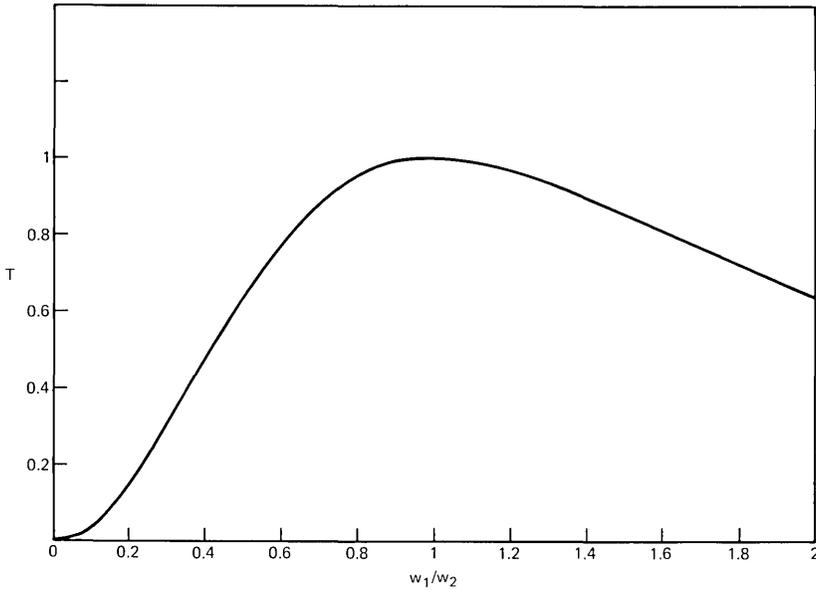
Fig. 7—Power transmission coefficient T as a function of the ratio w_1/w_2 of the beam width parameters of the two fibers joined by a perfectly aligned splice.

This relation is important if the width of the mode field is known from measurements instead of being inferred from the known V-value of the fiber.

We begin our discussion of splice losses by considering two perfectly aligned fibers with different dimensions. Figure 7 shows a plot of the power transmission coefficient T as a function of w_1/w_2. This function is, of course, identical if plotted versus w_2/w_1. A ratio $w_1/w_2 = 1.4$ (or 0.71) causes a power loss of 10 percent. If we assume that we are dealing with step-index fibers, we see from Fig. 2 that a reduction of the V-value from $V = 2.4$ to $V = 1.68$ causes w/a to increase by a factor of 1.4. (An increase of the V-value has far less influence on the beam size.) These changes of V translate directly into changes of w only if a is kept constant and V is changed by varying Δ. Now let us keep Δ constant and change V from 2.4 to 1.68 by decreasing the value of the core radius a. This change increases w/a by 1.4, which means that the beam width is actually decreased by a factor of 0.98. This example shows that a change of the core radius does not cause a proportional change of the beam width.

In the remainder of our discussion, we assume that the beam widths of the guided modes of both fibers joined by the splice are identical, $w_1 = w_2$. Figure 8 is a graph of (22) for a step-index fiber splice with $\lambda = 1$ μm, $n_2 = 1.457$, and $V = 2.4$. The figure illustrates how insensitive a fiber splice is to longitudinal separation of the fiber ends. However, this figure
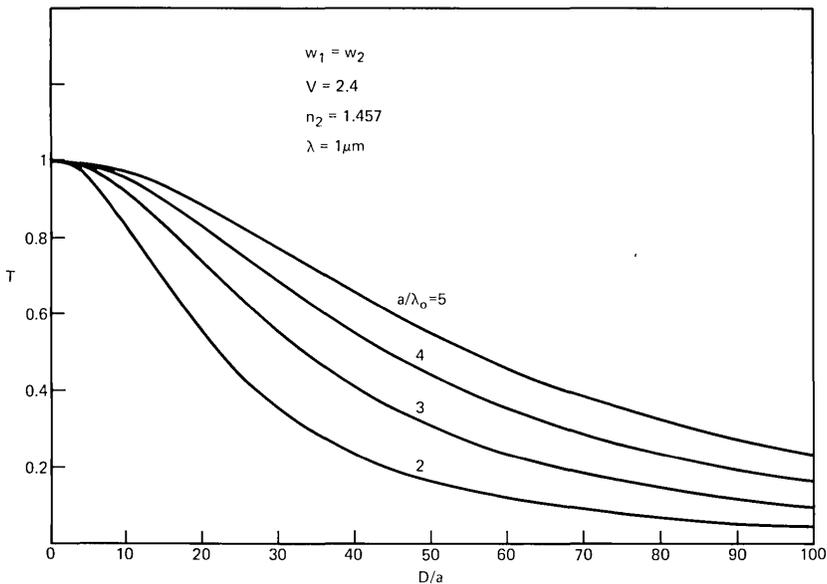
Fig. 8—Power transmission coefficient as a function of normalized longitudinal fiber displacement for identical step-index fibers. The curve parameter is the ratio a/λ of the core radius to free space wavelength.

was drawn under the assumption that both fibers of the splice are immersed in index-matching fluid. For splices in air, we must set $n_2 = 1$ in (22) and (23), which leads to lower values of T. The figure shows that larger core radii result in lower splice losses. However, it must be remembered that the figure is drawn for a fixed value of $V = 2.4$, fibers with larger core radii, thus have smaller values of Δ.

The transmission coefficients for tilts and offsets are gaussian functions of the tilt angle θ or the amount of offset d. Using normalized variables, $\pi n_2 w \theta / \lambda$ for the tilt and d/w for the offset, we can represent both cases in Fig. 9. For $w_1 = w_2 = w$ (27) simplifies to

$$\theta_e = \frac{\lambda}{\pi n_2 w}, \tag{33}$$

and the amount of offset (29) that causes T to drop to $1/e = 0.368$ of its maximum value reduces to

$$d_e = w. \tag{34}$$

We illustrate the meaning of these expressions with a specific example. Let $V = 2.4$, $\lambda = 1~\mu$m, and $\Delta = 0.002$ so that we obtain a core radius of $a = 4.15~\mu$m for $n_2 = 1.457$. For the step-index fiber, we find from (8) or Fig. 2, $w/a = 1.1$ or $w = 4.56~\mu$m. (The corresponding value for the parabolic index fiber would be $w = 4.48~\mu$m.) The power transmission
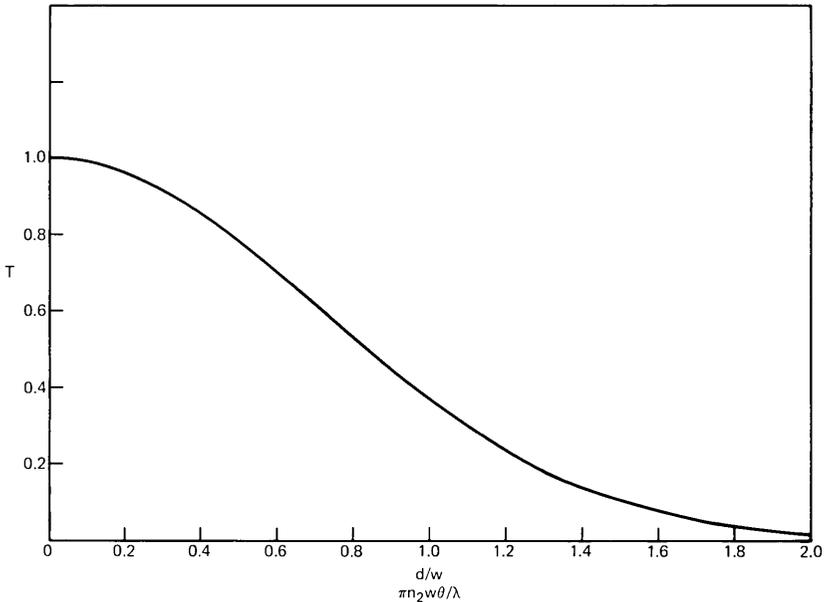
Fig. 9—Power transmission coefficient as a function of normalized offset or tilt angle.

coefficient would be $T = 0.368$ for $d = w = 4.56$ μm, or it would be $T = 0.9$ for $d = 1.5$ μm. If the axis of the fibers (joined by the splice) are laterally aligned, but if there is a tilt, a tilt angle $\theta = 0.048$ radians $= 2.7°$ causes $T = 0.368$, while $\theta = 0.91°$ reduces T from its maximum value to $T = 0.9$. A fiber with a narrower width parameter w would be less tolerant of offsets, but correspondingly more tolerant of tilts. This mutual relationship is expressed by the "uncertainty relation" (30).

V. CONCLUSION

Using the close match between gaussian beams and the field distributions of single-mode fibers, we have presented formulas and graphs for the power transmission coefficient of light through a fiber splice. The fibers on either side of the splice need not be identical. Splice losses occur for mismatched fiber parameters, transverse fiber displacement (offset), angular displacement (tilt), and longitudinal displacement[6]. All four cases have been discussed. Splice tolerances with respect to tilt and offset are mutually exclusive. This relationship has been expressed by an "uncertainty principle".

The results presented in this paper are immediately applicable to the excitation of single mode fibers by gaussian-shaped laser beams.[7,8]

REFERENCES

1. J. S. Cook, W. L. Mammel, and R. J. Grow, "Effect of Misalignment on Coupling Efficiency of Single-Mode Optical Fiber Butt Joints," B.S.T.J., *52*, No. 8 (October 1973), pp. 1439–1448.
2. H. Kogelnik, "Coupling and Conversion Coefficients for Optical Modes in Quasi-Optics," Microwave Research Institute Symposia Series, *14*, New York: Polytechnic Press, 1964, pp. 333–347.
3. J. A. Arnaud, *Beam and Fiber Optics*, New York: Academic Press, 1976.
4. D. Gloge, "Weakly Guiding Fibers," Appl. Opt., *10*, No. 10 (October 1971), pp. 2252–2258.
5. D. Marcuse, *Theory of Dielectric Optical Waveguides*, New York: Academic Press, 1974.
6. D. Gloge, "Offset and Tilt Losses in Optical Fiber Splices," B.S.T.J., *55*, No. 7 (September 1976), pp. 905–915.
7. J. R. Stern, M. Peace, and R. B. Dyott, "Launching into Optical-Fibre Waveguide," Electr. Lett., *6*, No. 6 (March 1970), pp. 160–162.
8. D. Marcuse, "Excitation of the Dominant Mode of a Round Fiber by a Gaussian Beam," B.S.T.J., *49*, No. 8 (October 1970), pp. 1695–1703.

Minimum Impulse Response
in Graded-Index Fibers

By J. S. COOK

(Manuscript received October 15, 1976)

A straightforward analysis slightly extending the work of Kawakami and Nishizawa and of Gloge and Marcatili provides some insight about the extent to which the mode differential delay in a graded-index fiber can be minimized. We show that an ideal fiber (no mode mixing) with uniform mode excitation and loss and uniform material dispersion can theoretically have an rms pulse broadening due to mode differential delay as small as about $0.02L\Delta^2 n_0/c$. We suggest that further improvement can result through recognition of differential mode loss and by accurate control of the (non-zero) rate of change of dispersion with fiber index.

I. INTRODUCTION

It is known through simple first-order analysis that differential delay between the propagational modes in multimode optical fibers can be greatly reduced by grading the optical index of the core so that the index

$$n = n_0(1 - \Delta R^2), \tag{1}$$

where R is the fiber radius normalized to unity at the core-cladding boundary. It is also known that even less differential delay can be realized theoretically by slightly perturbing the gradient from this parabolic shape.

We have taken some direct steps based on existing analyses to determine how much further improvement might be realized if the optimum gradient could be realized in an "ideal" fiber, where geometry is invariant over its length (no mode mixing) and where material dispersion is invariant with radius. The approach is very simple and will be so stated, but the algebra is tedious and what little has been included will be found in the appendix.

In their very nice analysis, Kawakami and Nishizawa[1] showed that an improvement in fiberguide impulse response could be obtained by perturbing the parabolic profile through the addition of a small fourth-order variation in the index gradient. They suggested that the minimum pulse width would be obtained when the fourth-order coefficient, δ, lies between the values $\frac{2}{3}$, where all meridional modes are synchronous, and 1, where circular spiral modes are synchronous. Minimum total pulse width, τ, in fact occurs when $\delta = \frac{2}{3}$, and minimum rms width, σ, occurs when $\delta = \frac{5}{7}$. Their expression for n can be written:

$$n = n_0[1 - 2\Delta R^2 + \delta(2\Delta)^2 R^4]^{1/2}. \tag{2}$$

It will be seen presently that when the index gradient is near optimum it takes some care to keep track of which propagating modes are the fastest and which are the slowest. The overall pulse width, τ, is simply the difference in arrival time between the slowest and fastest modes. If we plot τ vs δ [found by solving (5) for all mode numbers at each δ and taking the difference between the extremes], the curve is continuous but its derivative is not (Fig. 3 below). Personick[2] has pointed out that the rms pulse width, σ, (the second moment of the received pulse) is more useful for fiberguide system analysis than is τ.

$$\sigma^2 = \frac{\int_\tau t^2 p(t)dt}{\int_\tau p(t)dt} - \left[\frac{\int_\tau tp(t)dt}{\int_\tau p(t)dt}\right]^2, \tag{3}$$

where $p(t)$ is the power arriving at a given point at time, t. Note that $d\sigma/d\delta$ is continuous, hence minimum σ can be found by direct computation.

Gloge and Marcatili have shown in their analysis[3] that minimum total pulse width occurs when

$$n = n_0[1 - 2\Delta R^{2(1-\Delta)}]^{1/2}. \tag{4}$$

For convenience, we introduce a multiplier, ρ, in the exponent of R in (4), namely,

$$n = n_0[1 - 2\Delta R^{2(1-\rho\Delta)}]^{1/2}. \tag{4'}$$

By definition, minimum total pulse width occurs for $\rho = 1$; minimum rms width, σ, however, occurs when $\rho = 1.2$.

These results are found by determining in each case the time of arrival of energy propagating in the μ, ν mode (radial, azimuthal mode number) with respect to the arrival time of the zero-order mode:

$$t(\delta) = \frac{L\Delta^2 n_0}{2cM^2}[(1 - 3\delta/2)(2\mu + \nu + 1)^2 + \delta/2(\nu - 1)^2] \tag{5}$$
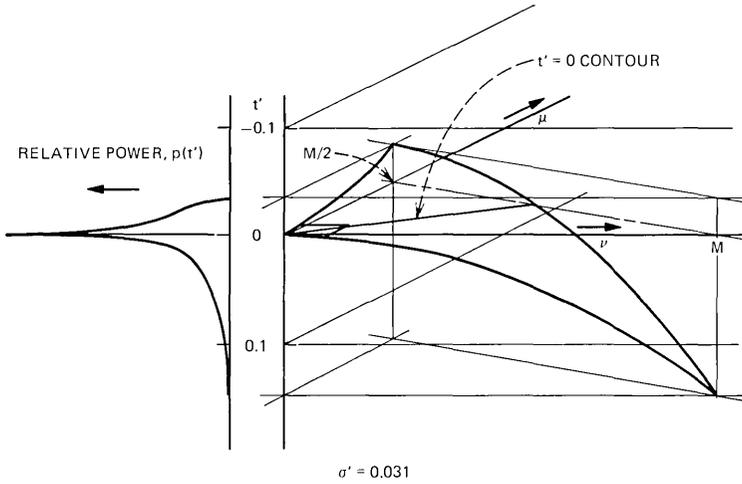
Fig. 1—Normalized time of arrival vs mode number (right) and mode count vs time of arrival (left) for $n = n_0[1 - 2\Delta R^2 + (5/7)(2\Delta)^2 R^4]^{1/2}$; i.e., $\delta = 5/7$.

$$t(\rho) = \frac{L\Delta^2 n_0}{2cM^2}\left[(2\mu + \nu + 1)^2 - \rho M(2\mu + \nu + 1)\right], \tag{6}$$

where L is the length of the fiber, c is free-space light velocity, and M is the largest guided-mode number.

In both cases,

$$M = (2\mu + \nu + 1)_{\max} \approx \pi n_0 (a/\lambda_0)\sqrt{2\Delta}, \tag{7}$$

where a is the core radius and λ_0 is the free-space wavelength of light. Also, since we must include both E and H waves and two polarizations, a fiber with near-parabolic core gradient carries a total of about M^2 modes.

Equation (3) can be solved easily if we assume all modes are equally excited $[p(\mu, \nu) = 1]$ and integrate over all modes. So

$$\sigma^2 = \frac{\int_0^M \int_0^{(M-\nu)/2} t^2(\mu, \nu)d\mu d\nu}{M^2/4} - \left[\frac{\int_0^M \int_0^{(M-\nu)/2} t(\mu, \nu)d\mu d\nu}{M^2/4}\right]^2. \tag{8}$$

Substitution of (5) and (6) into (8) (assuming $M \gg 1$ to simplify the algebra) and minimization with respect to δ and ρ, respectively, produce the results already stated.

We can substitute these minimum values back into (5) and (6) and plot arrival time as a function of μ and ν, as shown in Figs. 1 and 2. Also

shown in these figures (on the left) is a plot of relative power vs time of arrival, again assuming uniform excitation of the modes.

Figures 3 and 4 show τ' and σ' vs δ and ρ, respectively. The prime denotes normalization with respect to $L\Delta^2 n_0/c$.

$\sigma' = 0.029$
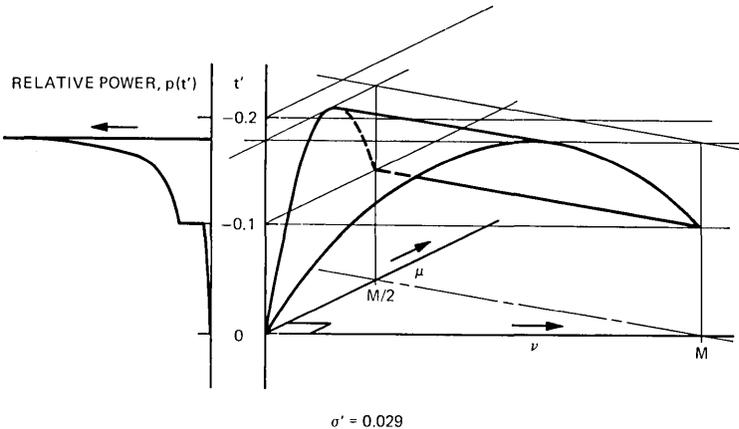
Fig. 2—Normalized time of arrival vs mode number (right) and mode count vs time of arrival (left) for $n = n_0[1 - 2\Delta R^{2(1-6\Delta/5)}]^{1/2}$; i.e., $\rho = \frac{6}{5}$.
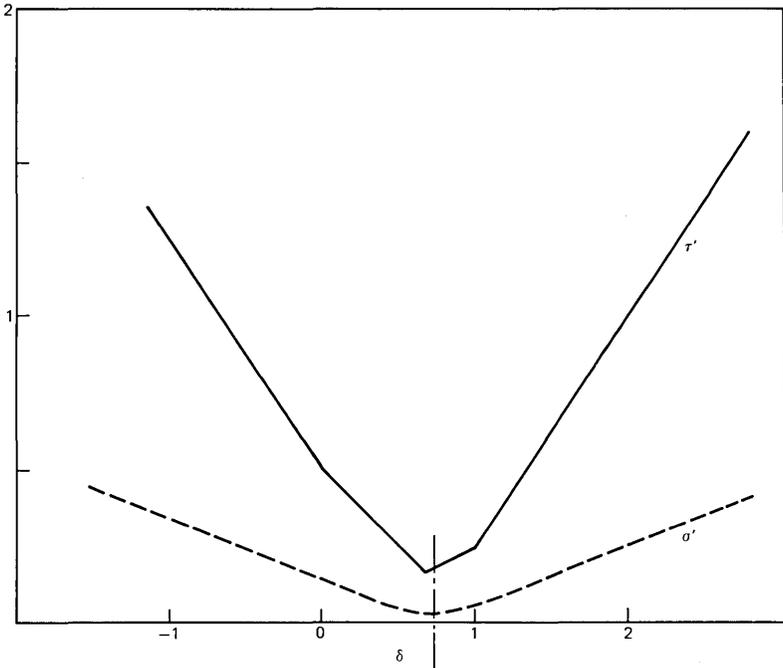
Fig. 3—Normalized total pulse width, τ', and rms pulse width σ' vs δ for $n = n_0[1 - 2\Delta R^2 + \delta(2\Delta)^2 R^4]^{1/2}$.
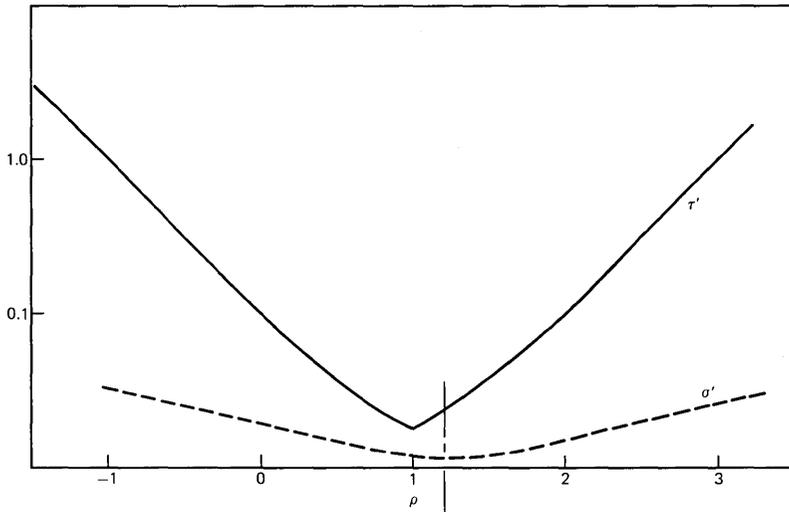
Fig. 4—Normalized total pulse width, τ', and rms pulse width σ' vs ρ for $n = n_0[1 - 2\Delta(R)^{2(1-\rho\Delta)}]^{1/2}$.

It is natural to question whether we can improve on these independent functional perturbations from the parabola by an appropriate combination thereof. Combination of (5) and (6) yields:

$$t(\delta, \rho) = \frac{L\Delta^2 n_0}{2cM^2} \left[\left(1 - \frac{3\delta}{2}\right)(2\mu + \nu + 1)^2 \right.$$

$$\left. + \frac{\delta}{2}(\nu - 1)^2 - \rho M(2\mu + \nu + 1) \right]. \quad (9)$$

Substituting (9) into (8), as before, and minimizing with respect to both δ and ρ produces: $\delta = \frac{1}{3}$, $\rho = \frac{2}{3}$. Substituting these into (9) and neglecting the 1s (large mode count) yields:

$$t \approx \frac{L\Delta^2 n_0}{cM^2} \left[\mu^2 + \mu\nu + \left(\frac{1}{3}\right)\nu^2 - \left(\frac{1}{3}\right)M(2\mu + \nu) \right]. \quad (10)$$

This is plotted in Fig. 5.

Now combine (2) and (4′) and let $\delta = \frac{1}{3}$ and $\rho = \frac{2}{3}$ to find

$$n = n_0[1 - 2\Delta R^{2(1-2\Delta/3)} + (\frac{1}{3})(2\Delta)^2 R^4]^{1/2} \quad (11)$$

as the near-optimum index gradient. Equation (11) can also be written

$$n \approx n_0[1 - \Delta R^2 - \Delta^2 \epsilon(R)], \quad (12)$$

which may be more convenient if we are seeking the excursion of the

RELATIVE POWER, p(t')

t' = −0.08 CONTOUR

σ' = 0.0216
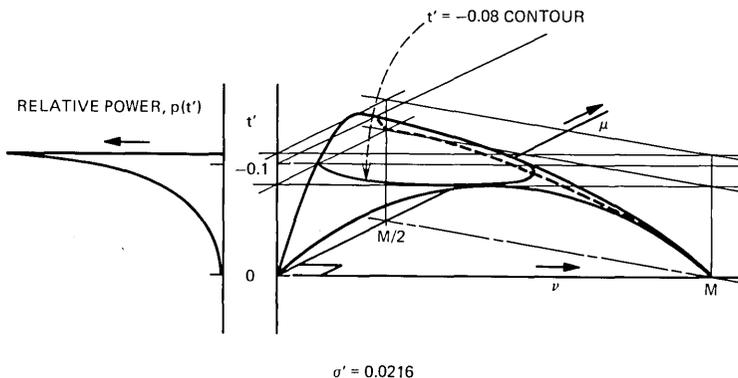
Fig. 5—Normalized time of arrival vs mode number (right), and mode count vs time of arrival (left), for $n = n_0[1 - 2\Delta R^{2(1-2\Delta/3)} + (1/3)(2\Delta)^2 R^4]^{1/2}$; i.e., $\rho = \frac{2}{3}, \delta = \frac{1}{3}$.

improved gradient from the inverse parabola of (1):

$$\epsilon(R) \approx \left(\frac{2}{3}\right) R^2 \ln (1/R^2) - \frac{R^4}{6}. \tag{13}$$

This is plotted in Fig. 6.

II CONCLUSION

The rms pulse width in an ideal fiber with parabolic gradient is

$$\sigma \approx 0.145 \frac{L\Delta^2 n_0}{c}. \tag{14}$$

For the profile determined in this paper,

$$\sigma = 0.0216 \frac{L\Delta^2 n_0}{c}. \tag{15}$$

Hence, nearly an order-of-magnitude improvement in fiberguide impulse response over that produced by an inverse parabolic gradient could be realizable in an ideal multimode fiber.

Uniform excitation of the modes was assumed for this analysis. If the low-order modes tend to carry more power,[2] even smaller σ could be realized and Fig. 1 would suggest that $\rho \to 0$ and $\delta \to \frac{5}{7}$ might be more nearly optimum.

Olshansky and Keck[4] have shown that when material dispersion varies with radius, the optimum gradient is significantly different from parabolic and from those discussed here. On the other hand, it can be shown that procedures similar to those indicated here can be carried out; and if the dependence of the dispersion on optical index could be controlled to a predetermined practical value, even less differential delay than that
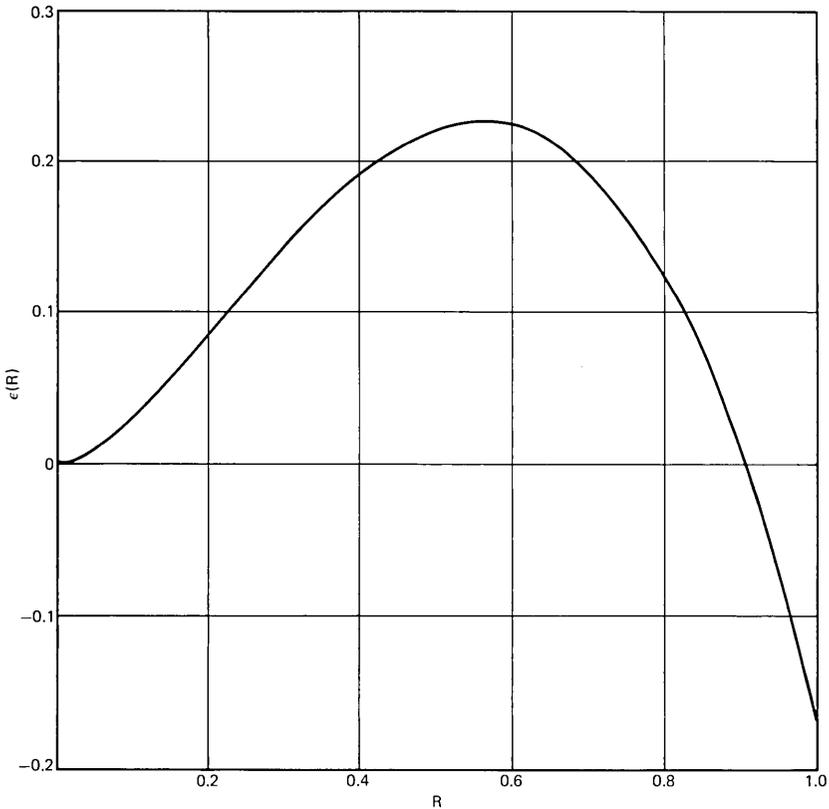
Fig. 6—Index perturbation from inverse parabola for minimum rms pulse width [see eqs. (12), (13)].

indicated here could be realized. This does not necessarily contradict the results of the calculation by Arnaud and Fleming[5] based on an analysis[6] which includes tunneling modes at equal weight with the more clearly guided modes. They showed that if we assume a very specific non-zero dependence of dispersion on index, namely, that resulting from the inclusion of germania in silica, the differential delay of the optimum graded-index fiber is considerably degraded. We suggest, however, that if we carefully choose materials to enhance the index (a propitious mix of germania and phosphorous oxide, for example), we might provide a dispersion vs index dependence that would improve rather than degrade the mode differential delay, at least at a particular light wavelength.

This goes well beyond our present ability to make measurements and control materials, however, and only leads us to conclude that improvement in technology can potentially bring significant improvements in the information-carrying capacity of low-loss graded-index fibers.

APPENDIX

It is our purpose first to justify (5) and (6). Equation (5) comes from Ref. 1, eq. (29), which translates directly to

$$\frac{\omega^2 n_0^2}{c^2} - \beta^2 = \frac{2\omega n_0}{c} (2\mu + \nu + 1) \frac{\sqrt{2\Delta}}{a}$$

$$- \delta[6\mu^2 + 6\mu(\nu + 1) + (\nu + 1)(\nu + 2)] \frac{2\Delta}{a^2}, \quad (16)$$

where a is the core radius.

This can be solved for β to find

$$\beta \approx \frac{\omega n_0}{c} - (2\mu + \nu + 1) \frac{\sqrt{2\Delta}}{a}$$

$$- \frac{c\Delta}{\omega n_0 a^2} \left[\left(1 - \frac{3\delta}{2}\right) (2\mu + \nu + 1)^2 + \frac{\delta}{2} (\nu^2 - 1) \right]. \quad (17)$$

The largest propagating mode numbers can be found by recognizing that the phase constant, β, can be no lower than the phase constant, k_a, of an unbound wave in the cladding, where

$$n_a \approx n_0[1 - 2\Delta]^{1/2} \approx n_0 (1 - \Delta). \quad (18)$$

So

$$\beta_{\min} \approx \frac{\omega n_0}{c} - (2\mu + \nu + 1) \frac{\sqrt{2\Delta}}{a} = k_a \approx \frac{\omega n_0}{c} (1 - \Delta). \quad (19)$$

Eq. (19) yields

$$(2\mu + \nu + 1)_{\max} \equiv M = \frac{\omega n_0 a \sqrt{\Delta}}{c\sqrt{2}} \quad (20)$$

$$M^2 = \frac{\omega^2 n_0^2 a^2 \Delta}{2c^2}. \quad (21)$$

Substitution of (20) into (17) shows that each term on the right-hand side thereof is of order Δ smaller than the previous.

The time of arrival of energy carried in any mode through length, L, of fiber relative to the arrival of energy in the zero-order mode is

$$t = L \left[\frac{\partial \beta}{\partial \omega} - \frac{\partial \beta_0}{\partial \omega} \right]. \quad (22)$$

Differentiation of (17) and substitution of that [utilizing (20)] into (22) produces eq. (5).

Equation (6) is found from substitution of GM(17) [i.e., eq. (17) of Gloge and Marcatilli in Ref. 2] into GM(19), recognizing that the α of Gloge and Marcatili is our $2(1 - \rho\delta)$. First, we must identify m/M of
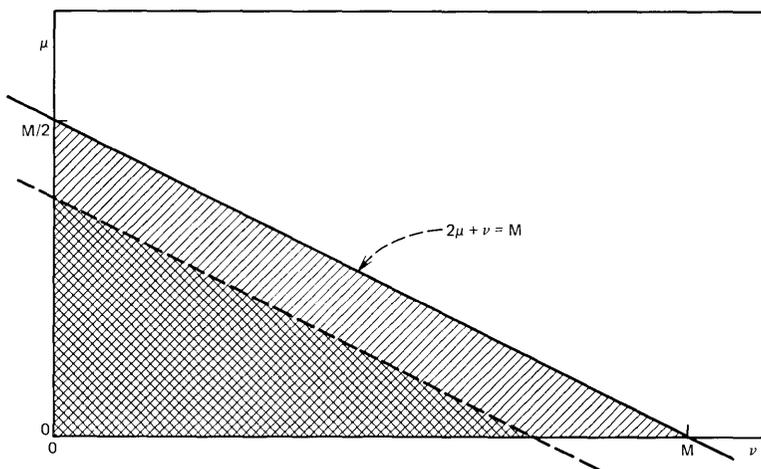
Fig. 7—All modes that lie under the line intercepting the μ axis at $M/2$, the ν axis at M, are propagating modes.

GM(17). The symbols are unfortunate, but we find this equates to our $(2\mu + \nu + 1)^2/M^2$.

This is easier to justify than derive. All modes are associated with positive mode numbers that lie in the μ, ν plane, as shown in Fig. 7. We can plot the identity of (20) on the plane to bound the crosshatched area in Fig. 7 within which lie all propagating mode numbers. The total number of propagating modes is four times the total mode-number combinations in that area, which can be seen by inspection to be M^2. Gloge and Marcatili[2] showed that for their assumed functional index variation, the phase constant depends on the total mode number independent of the ratio of μ to ν. If, then, we draw a dotted line parallel to the limit line, as in Fig. 7, the sum of all modes with phase constant greater than that represented by those on the line are identified by numbers that lie within the double crosshatched area. This Gloge and Marcatili call m, and we call $(2\mu + \nu + 1)^2$. Their M is the total number of propagating modes, which is our M^2. [We have neglected the extra (1) since we have assumed $M \gg 1$. We would have to be more precise if only a small number of modes were involved.]

We can now translate GM(17) to be

$$\frac{(2\mu + \nu + 1)^2}{M^2} \approx (\delta_{GM}/\Delta)^2$$

or

$$\delta_{GM} = \frac{\Delta(2\mu + \nu + 1)}{M}. \tag{23}$$

This in GM(19) results in (6).

Performing the integrations for (8) is simple but tedious. A helpful formula, however, is

$$\int_0^{(M-\nu)/2} (2\mu + \nu)^n d\mu = \frac{M^{n+1} - \nu^{n+1}}{2(n+1)}. \tag{24}$$

The evaluations of (8) for the two cases of (5) and (6) are

$$\sigma(\delta) = \frac{L\Delta^2 n_0}{c} \frac{1}{2\sqrt{3}} \left[\frac{1}{4} - \frac{2\delta}{3} + \frac{7\delta^2}{15} \right]^{1/2} \tag{25}$$

$$\sigma(\rho) = \frac{L\Delta^2 n_0}{c} \frac{1}{2\sqrt{3}} \left[\frac{1}{4} - \frac{2\rho}{5} + \frac{\rho^2}{6} \right]^{1/2}. \tag{26}$$

For the combined case,

$$\sigma(\delta, \rho) = \frac{L\Delta^2 n_0}{c} \frac{1}{2\sqrt{3}} \left[\frac{1}{4} - \frac{2\delta}{3} + \frac{7\delta^2}{15} - \frac{2\rho}{5} + \frac{\rho^2}{6} + \frac{8\delta\rho}{15} \right]^{1/2}. \tag{27}$$

Minimizing these is straightforward.

REFERENCES

1. S. Kawakami and J. I. Nishizawa, "An Optical Waveguide with the Optimum Distribution of the Refractive Index with Reference to Waveform Distortion," IEEE Trans. Microw. Theory Tech., *MTT-16*, No. 10 (October 1968), pp. 814–818.
2. S. D. Personick, "Receiver Design for Digital Fiber Optic Communication Systems," B.S.T.J., *52*, No. 9 (July–August 1973), pp. 843–886.
3. D. Gloge and E. A. J. Marcatili, "Multimode Theory of Graded-Core Fibers," B.S.T.J., *52*, No. 9 (November 1973), pp. 1563–1578.
4. R. Olshansky and D. B. Keck, "Pulse Broadening in Graded-Index Optical Fibers," Applied Optics, *15*, No. 2 (February 1976), pp. 483–491.
5. J. A. Arnaud and J. W. Fleming, "Pulse Broadening in Multimode Optical Fibers With Large Δn/n: Numerical Results," Electron. Lett., *12* (1976), pp. 167–169.
6. J. A. Arnaud, "Pulse Broadening in Multimode Graded-Index Fibers," Electron. Lett., *11* (1975), pp. 8–9.

Simplified Theory of the Multimode Fiber Coupler

By K. OGAWA

We simplify the coupling theory between two contiguous, parallel, multimode step-index fibers, describe the coupling concept, and derive an upper estimate for the overall coupling efficiency between the two fibers. The maximum coupling obtainable, according to this estimate, is less than 72 percent (−1.5 dB). The coupling efficiency derived for short coupling lengths shows good agreement with experimental results.

I. INTRODUCTION

A multimode fiber tap-coupler is a useful component for certain optical communication systems such as the optical data bus. Unfortunately, it is not an easy matter to evaluate the simultaneous coupling process between the hundreds or thousands of modes. An analysis of the problem has been given by Snyder and others.[1-3] These authors analyzed only the coupling between certain mode pairs. Snyder[4] recently reported the total power transition between multimode fibers. However, his conclusion is based on HE_{1m} modes for analyzing the crosstalk.

We have derived a similar simplified expression for the total coupling between identical, contiguous, parallel, step-index, multimode fibers, which can expand to all modes under the restriction that two fiber cores are touching each other. We predict that the maximum coupling efficiency is less than 1.5 dB when all modes carry an equal amount of power. The distance between two fibers affects the coupling efficiency very seriously when fibers have large numerical apertures.

Our simple formula agrees very well with experimental results, in spite of a large number of approximations made.

II. COUPLING COEFFICIENT

2.1 Simplified coupling coefficient

Figure 1 shows the geometry of the fiber coupler. The cores are parallel
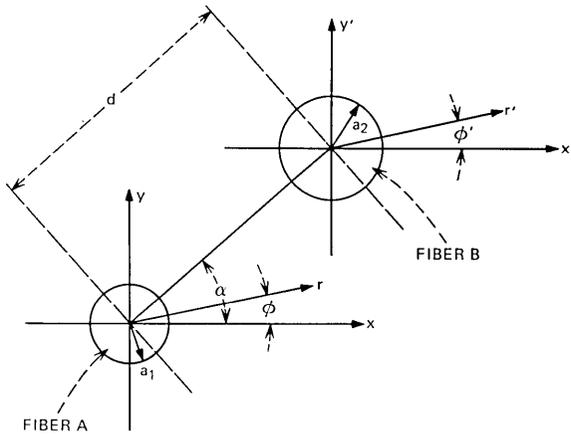
Fig. 1—Geometry used to derive the coupling coefficients.

to each other and surrounded by a medium that has the same refractive index as the cladding.

A general description of strongly coupled multimode fibers is given in Appendixes A and B. The coupling coefficient in eq. (24) has a complicated form. It involves the modified Bessel functions and many different parameters that depend upon the eigenvalue equations. We will simplify the coefficient under the following assumptions: (i) the two fibers are identical, (ii) only coupling between modes having the same propagation constant is considered (see discussion in Section 2.2), and (iii) the distance, d, between the two fiber axes is nearly equal to $2a$, where a is the core radius of the fiber.

We can rewrite the coupling coefficient by using assumptions (i) and (ii). The result is

$$|C_{A_iB_i}| = |C_{B_iA_i}|$$

$$= \gamma \frac{\sqrt{2\Delta}}{a} \frac{u^2}{v^3} \frac{|K_{2l}(wd/a) \cos(2l\alpha) \pm K_o(wd/a)|}{K_{l-1}(w)K_{l+1}(w)}, \quad (1)$$

where

$$\gamma = \tfrac{1}{2} \text{ when } l = 0$$

$$\gamma = 1 \text{ when } l > 0$$

$$\alpha = \text{(defined in Fig. 1)}$$

$$a = \text{radius of core}$$

$$n_0 = \text{refractive index of core}$$

$$\Delta = \text{relative refractive index difference}$$

$$d = \text{distance between the two fibers}$$
$$l = \text{azimuthal order number}$$
$$v^2 = w^2 + u^2$$
$$v^2 = w^2 + u^2$$
$$v^2 = a^2 k^2 n_0^2 (2\Delta).$$

The coupling coefficient (1) is still complicated because it includes the modified Bessel functions. Further simplification is necessary to present a simple physical picture for the coupling process.

We use assumption (*iii*) to simplify (1) using the asymptotic expansions of the modified Bessel function. As discussed in Appendix C, the modified Bessel function term in (1) can be expressed by a very simple expression for a fairly large range of azimuthal order numbers. We use the average value of $|C_i|$ as shown in Appendix C. The result is as follows:

$$|\overline{C_{A_i B_i}}| = |\overline{C_{B_i A_i}}| = \left(\gamma \frac{\sqrt{2\Delta}}{a} \frac{u^2}{v^3} \right) \left(\frac{\sqrt{2w}}{\sqrt{\pi d/a}} e^{-w(d/a-2)} \right), \qquad (2)$$

where

$$d/a \cong 2$$
$$\gamma = 1 \text{ for all modes except } LP_{0m}; \gamma = \tfrac{1}{2} \text{ for } LP_{0m}.$$

We let γ equal 1 for all modes. This assumption does not seriously affect the results of this analysis.

The coupling coefficient is thus,

$$|\overline{C_{A_i B_i}}| = \frac{\sqrt{2\Delta}}{a} \left(\frac{u^2}{v^3} \frac{\sqrt{2w}}{\sqrt{\pi d/a}} e^{-w(d/a-2)} \right), \qquad (3)$$

where

$$d/a \cong 2.$$

The coupling coefficient expressed by (3) does not explicitly depend upon the azimuthal order number l. However, it is still dependent upon the solution of the eigenvalue equation. We simplify eq. (3) further by introducing simple expressions for u, w, and v.

The parameter v is expressed by the total number of modes in the fiber when the total number of modes is large. The results are[5,6]

$$v^2 = 2N, \qquad (4)$$

where

$$v \gg 1$$
$$N = \text{total mode number of fiber } A \text{ and } B.$$

We order all LP modes according to their z-propagation constants from the largest to the smallest. We label them by the sequential numbers, $i = 1 \cdots N$. For example, the two orthogonal LP_{01} modes are designated as the first and the second mode. The LP_{11} modes assume the orders 3, 4, 5, and 6. For $i \gg 1$, the cutoff value of u for the ith mode is approximately[5]

$$u_{\text{cutoff}} = (2i)^{1/2} \quad (i \gg 1).\tag{5}$$

We replace u by u_{cutoff} and we obtain

$$u = (2i)^{1/2} \quad (i \gg 1).\tag{6}$$

Equation (3) can be simplified by (4) and (6). The result is,

$$|\overline{C_{A_i B_i}}| = |\overline{C_i}| = \frac{2^{3/4}}{\sqrt{\pi}} \frac{\sqrt{\Delta}}{a N^{1/4}} \left(\frac{i}{N}\right) \left(1 - \frac{i}{N}\right)^{1/4}$$
$$\times \frac{\exp[-(2N - 2i)^{1/2} (d/a - 2)]}{\sqrt{d/a}}\tag{7}$$

or

$$= \frac{2^{3/4}(\Delta)^{1/4}}{\sqrt{\pi \cdot k \cdot n_0} a^{3/2}} \cdot \left(\frac{i}{N}\right) \left(1 - \frac{i}{N}\right)^{1/4} \frac{\exp[-(2N - 2i)^{1/2} (d/a - 2)]}{\sqrt{d/a}},$$

where

$$a = \text{radius of core}$$

$$k = 2\pi/\lambda$$

$$n_0 = \text{refractive index of core}$$

$$\Delta = \text{relative refractive index difference}$$

$$d = \text{distance between the two fibers.}$$

When $d/a = 2$, the coupling coefficient becomes,

$$|\overline{C_{A_i B_i}}| = |\overline{C_i}| = \frac{2^{1/4} \Delta^{1/4}}{\sqrt{\pi k n_0} a^{3/2}} \left(\frac{i}{N}\right) \left(1 - \frac{i}{N}\right)^{1/4}.\tag{8}$$

This simple expression for the coupling coefficient does not require the eigenvalue solutions. Figure 2 shows that the coupling coefficient reaches a maximum for the mode number $i = (4/5)N$. The reason is the following: the coupling coefficient as expressed by (24) is based on the field interaction of the evanescent field tail of a mode of fiber A and the core field of the same mode order in fiber B. Generally speaking, the higher-order modes have a stronger field in the cladding relative to their fields in the core. Therefore, the field interaction between the field tail of a mode of fiber A and the core field of the same order in fiber B in (16)
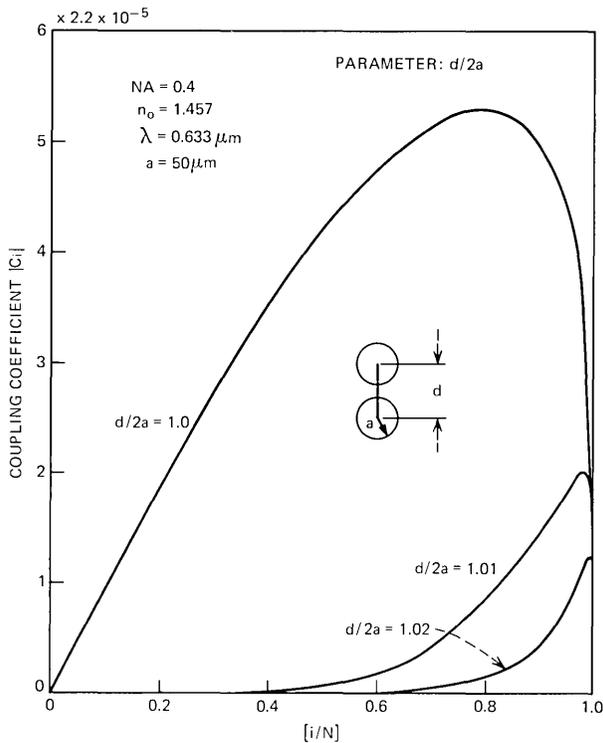
Fig. 2—Coupling coefficient vs mode order.

can be expected to increase with mode order until it reaches a maximum. Very high mode orders have very small fields in the core and therefore the interaction between very-high-order modes decreases again.

Figure 2 shows the effect of the gap between two fiber cores. Coupling occurs only between the higher-order modes as the gap increases. We will discuss the coupling efficiency using (7) and (8) in the next sections.

2.2 Coupling efficiency

The power coupled from the ith mode of fiber A to the ith mode of fiber B can be obtained from (25) when the ith mode of fiber A carries unit power and the ith mode of fiber B carries zero power at $z = 0$. The coupled power is

$$(P_{A \to B})_i = \sin^2 [|\overline{C_i}|z], \tag{9}$$

where

$$z = \text{the coupling length.}$$

The total coupled power is obtained by the summation of (9). (The coupling coefficients for some of the modes are nearly equal to zero; especially equal to zero are the coupling coefficients for one of the modes with $l = 0$. So the coupling efficiency defined here is an upper bound.) If all modes of fiber A carry equal power, then the total power from fiber A to fiber B is,

$$\frac{P_{A \to B}}{P_{in}} \leqq \frac{1}{N} \sum_{i=1}^{N} [\sin^2 (|C_i| z)]. \tag{10}$$

If we treat the mode number i as a continuous variable, the coupling efficiency η becomes

$$\eta = \frac{P_{A \to B}}{P_{in}} \leqq \frac{1}{N} \int_0^N [\sin^2 (|\overline{C}_i| z) di]. \tag{11}$$

The coupling between the near synchronous modes is not negligible, especially when $d/2a = 1$. However, the upper bound of the coupling efficiency is expressed by (11) when $|C_{ij}||C_{ji}| \approx |C_i|^2$ for $i - \sigma \leqq j \leqq i + \sigma$ $(\sigma/N) \ll 1$, a defined by (25). When the two fiber cores are touching each other, the coupling efficiency η is

$$\eta = \frac{P_{A \to B}}{P_{in}} \leqq \frac{1}{N} \int_0^N \sin^2 |\overline{C}_i| z \, di$$

$$= \int_0^1 \sin^2 \left(\frac{(2^{1/4} \Delta^{1/4}) z}{\sqrt{\pi k n_0} a^{3/2}} t (1-t)^{1/4} \right) dt, \tag{12}$$

where

$$t = \frac{i}{N}.$$

Figures 3 and 4 show several examples.

III. DISCUSSION

As a first example, consider two fibers having the following parameters: NA $= 0.2$, $a = 25$ μm (radius), $n_0 = 1.457$, $\gamma = 0.633$ μm. This fiber carries about 1,250 modes ($N = 1,250$).

The coupling coefficient is

$$|\overline{C}_i| = \left(\frac{1}{\sqrt{\pi}} \frac{\sqrt{NA}}{\sqrt{ka} \, n_0} \right) \frac{(t)(1-t)^{1/4}}{a} \frac{\exp[-\sqrt{2N}(\sqrt{1-t})(d/a - 2)]}{\sqrt{d/2a}}$$

$$= 0.011 \left(\frac{1}{a} (t)(1-t)^{1/4} \frac{\exp[-2\sqrt{2N}(\sqrt{1-t})(d/2a - 1)]}{\sqrt{d/2a}} \right). \tag{13}$$

According to Fig. 2, the maximum coupling efficiency (1.5 dB) is achieved at $d/2a = 1$ and 0.011 $z/a = 3.67$. Thus, the coupling length z is about $334a$ or $z = 8.36$ mm.
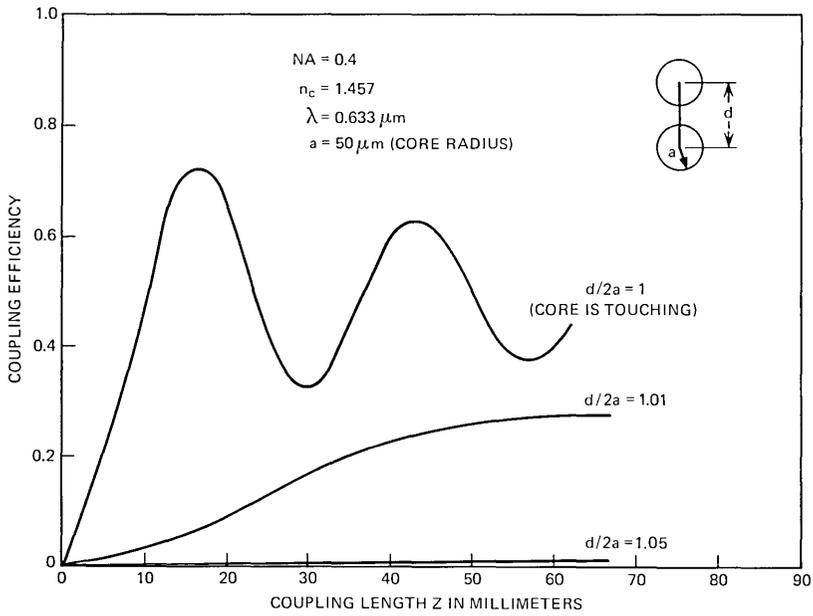
Fig. 3—Coupling efficiency vs coupling length.

Fig. 4—Coupling efficiency vs coupling length.

MULTIMODE FIBER COUPLER   735
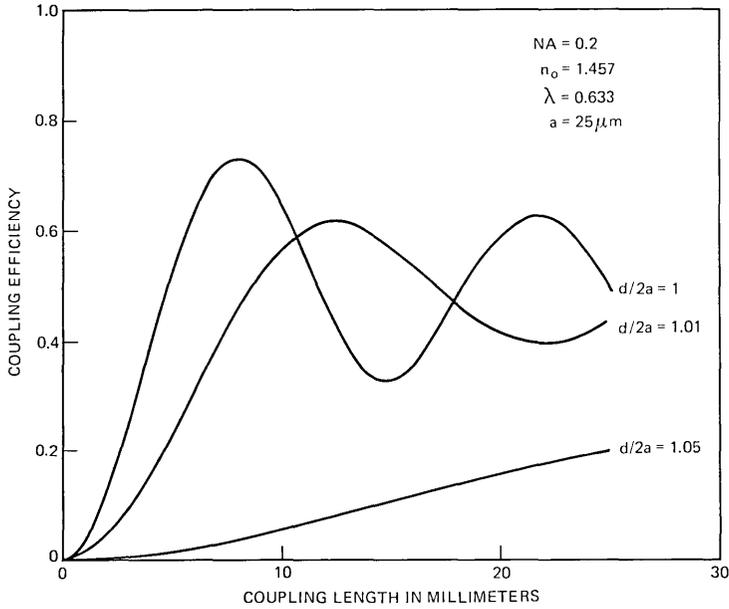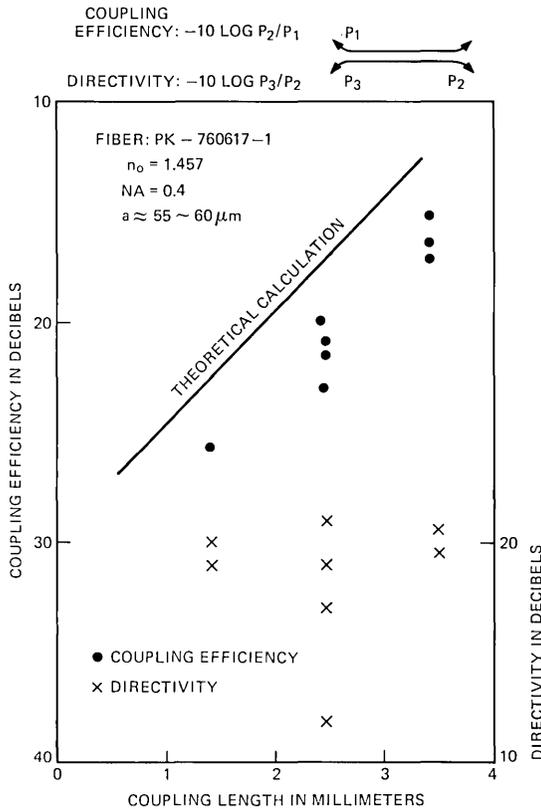
Fig. 5—Coupling efficiency vs coupling length; points indicate experimental results.

When $d/2a = 1.01$ so that the gap between the two fibers is about 0.5 μm, the coupling efficiency drops to 3.2 dB with the same coupling length (about 8.36 mm). When $d/2a = 1.05$ so that the gap between the two fibers is about 2.5 μm, the coupling efficiency becomes about 13.7 dB with the same coupling length. Thus, the distance between two fibers affects the coupling efficiency very strongly.

We now look at another example with respect to the following fiber parameters: NA = 0.4, $a = 50$ μm, $n_0 = 1.457$, $\lambda = 0.633$ μm. This fiber carries about 20,000 modes. The coupling coefficient is expressed by

$$|\overline{C_i}| = 0.01 \frac{1}{a} (t)(1 - t)^{1/4}$$

$$\times \frac{\exp\left[-2\sqrt{2N}(\sqrt{1-t})(d/2a - 1)\right]}{\sqrt{d/2a}}. \quad (14)$$

The maximum coupling efficiency (about 1.5 dB) is achieved when $d/2a$

x 10^{-4}

14

12

10

8

|C|

6

4

2

0

SIMPLIFIED COUPLING COEFFICIENT
LIMITATION OF COUPLING COEFFICIENT

ℓm : LP$_{\ell m}$

Λ = 0.00235
n_o = 1.457
a = 12.10310 [μm]
λ = 0.633 [μm]
v = 12
d/2a = 1.0

cos$\ell\,\phi$, α = 0
sin$\ell\,\phi$, $\alpha = \pi/2$

cos$\ell\,\phi$, $\alpha = \pi/2$
sin$\ell\,\phi$, α = 0   ($\ell > 0$)

α = 0

$\alpha = \pi/2$ ($\ell = 0$)

$$-|C_i| = \frac{2^{1/4}\sqrt[4]{\Lambda}}{\sqrt{\pi k n_o a^{3/2}}}$$

$$(i/N)\,(1-i/N)^{1/4}$$

0   0.2   0.4   0.6   0.8   1.0   1.2   1.4

(u/v)^2 OR [i/N]
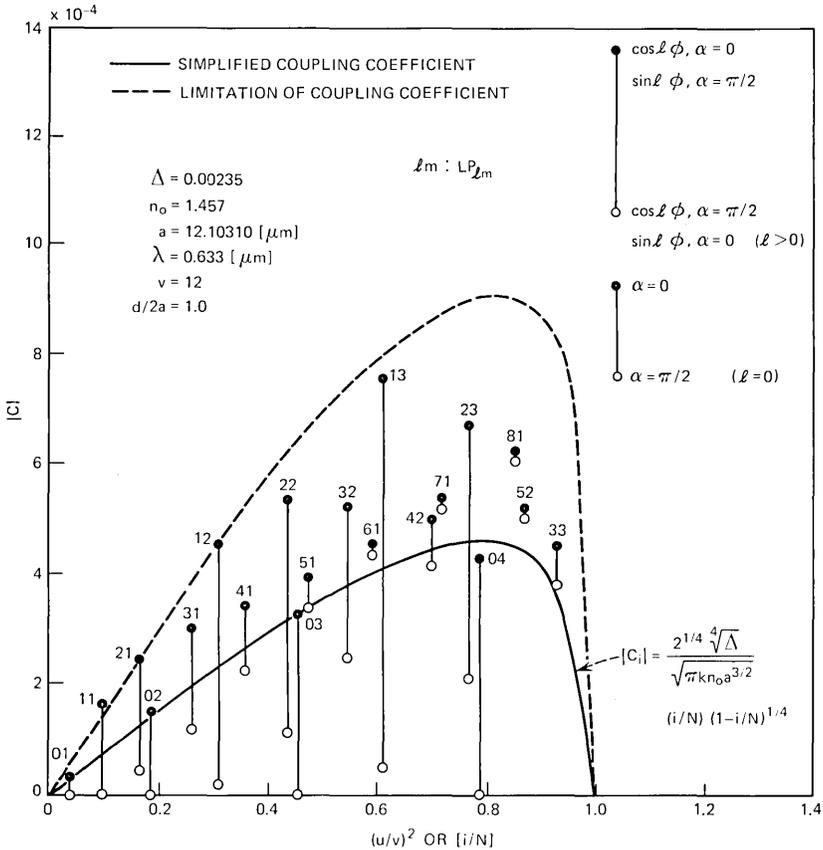
Fig. 6—Coupling coefficient vs mode order; points indicate numerical results.

= 1 and the coupling length is 16.7 mm. However, if the ratio $d/2a$ changes by only 1 percent so that the gap between the fibers becomes about 1 μm, the coupling efficiency becomes about 11.7 dB. The gap between two fibers affects the coupling efficiency more seriously when the fiber has a large numerical aperture and large radius.

Practically, it is not easy to produce a fiber coupler that has a uniform interaction gap over a long coupling length. Figure 5 shows experimental results that only cover very short coupling lengths. Plastic-clad fiber is attached to the acrylic base and its cladding is peeled off over the required coupling length. The exposed cores are pushed together and form a parallel coupling region. Silicone of the same type as the fiber cladding is injected to form a common cladding around the coupling region. Figure 4 shows that this simple theoretical approach yields good agreement. Figures 6 and 7 show results obtained from a numerical calculation of the coupling coefficient (23) and for fibers with the following parameters:
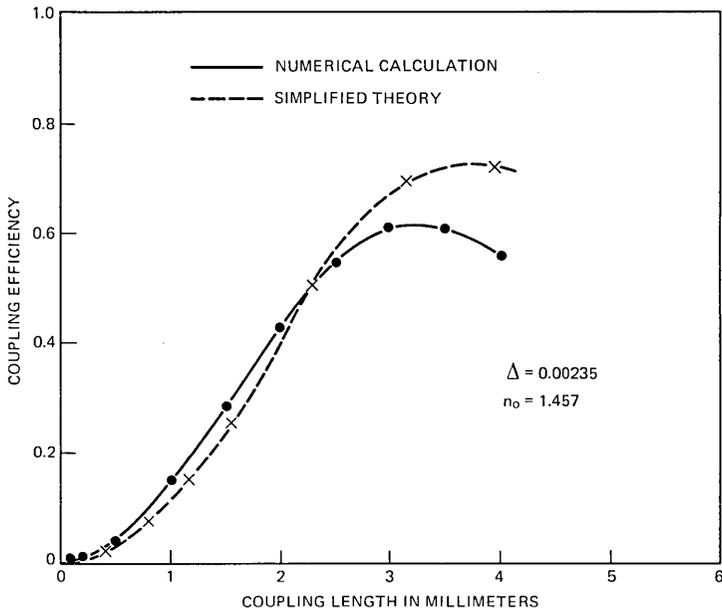
Fig. 7—Coupling efficiency vs coupling length; points indicate numerical results.

$\Delta = 0.00235$, $a = 12.1$ μm, $n = 1.457$. The simplified theory shows good agreement.

IV. CONCLUSION

We have derived a very simple formula for the coupling between two multimode fibers and have discussed the coupling mechanism and the coupling efficiency. We emphasize that the formula obtained involves some rather drastic approximations. However, this coupling formula explains the coupling mechanism very clearly and agrees with experimental and numerical results.

V. ACKNOWLEDGMENT

The author would like to thank D. Gloge for his comments and help, and D. Marcuse and J. A. Arnaud for helpful discussions. The author also would like to thank P. Kaiser for providing the fibers used in the experiments.

APPENDIX A

Coupling Coefficient

Coupling between two modes of different fibers was discussed by Snyder.[1,2] We use the formula derived by Snyder.

We assume the two fibers to be nearly identical so that the pth mode of fiber A can only couple to the qth mode of fiber B, provided the two modes have the same z-direction propagation constant.

In this case, the coupling equations have the following form

$$\frac{dA(z)}{dz} + j\beta_A A(z) = -jB(z)C_{AB}$$

$$\frac{dB(z)}{dz} + j\beta_B B(z) = -jA(z)C_{BA}, \qquad (15)$$

where $A(z)$ and $B(z)$ are the amplitudes of the modes of fiber A and fiber B respectively. The coupling coefficient is defined by

$$C_{AB} = \frac{\omega}{2} \int_{\substack{\text{core area} \\ \text{of fiber } B}} \epsilon_0 (n_0^2 - n_c^2) E_A E_B dS, \qquad (16)$$

where

$$n_0 = \text{core index of fiber } B$$

$$n_c = \text{cladding index of fibers } A \text{ and } B$$

$$E_A = \text{the normalized electric field of fiber } A$$

$$E_B = \text{the normalized electric field of fiber } B.$$

Weakly guiding fibers have simple field expressions, which are called the linearly polarized modes (LP). Each LP_{lm} mode represents a set of four modes when $l > 0$. (When $l = 0$, LP_{0m} represents two modes.) The four modes differ in polarization and azimuthal field distribution (the $\sin l\phi$ or $\cos l\phi$ term in the field-expansion equation). The field components can be described by[3]

$$E_y = H_X \begin{vmatrix} z_0/n_0 \\ z_0/n_c \end{vmatrix} = A_l \begin{vmatrix} J_l(ur/a)J_l(u) \\ K_l(wr/a)K_l(w) \end{vmatrix}$$

$$[\cos l\phi \text{ or } \sin l\phi]_{r \geq a}^{0 < r \leq a}, \qquad (17)$$

where

$$Z_0 = \sqrt{\frac{\mu}{\epsilon}}$$

$$A_l = \left(\frac{\mu}{\epsilon}\right)^{1/4} \left(\sqrt{\frac{2}{\pi n_0}}\right) \frac{uK_l(w)}{a[v\sqrt{K_{l-1}(w)K_{l+1}(w)}]} \quad l \neq 0$$

$$A_0 = \left(\frac{\mu}{\epsilon}\right)^{1/4} \left(\frac{1}{\sqrt{\pi n_0}}\right) \frac{uK_0(w)}{avK_1(w)} \quad l = 0.$$

These fields satisfy the following eigenvalue equation

$$u[J_{l-1}(u)/J_l(u)] = -w[K_{l-1}(w)k_l(w)]$$

or

$$u[J_{l+1}(u)/J_l(u)] = w[K_{l+1}(W)/K_l(W)]. \qquad (18)$$

We calculate the coupling coefficient between the pth mode of fiber A and the qth mode of fiber B, by using (16), (17) and (18). The pth mode field of fiber A is defined by l_1, w_1, a_1, r, ϕ, and the qth mode of fiber B is defined by l_2, w_2, a_2, r', ϕ'. The coupling coefficient is defined by (17).

$$C_{A_p B_q} = \frac{\omega}{2} \frac{\epsilon_0(n_0^2 - n_c^2)}{K_{l_1}(w_1)K_{l_2}(w_2)} (A_{l1})\, A_{l2}$$

$$\times \int_0^{2\pi} d\phi' \begin{Bmatrix} \cos l_1\phi \, (\cos l_2\phi') \\ \sin l_1\phi \, (\sin l_2\phi') \end{Bmatrix}$$

$$\times \int_0^{a_2} r'dr'K_{l_1}\left(w_1 \frac{r}{a_1}\right) J_{l_2}\left(u_2 \frac{r'}{a_2}\right). \qquad (19)$$

We introduce the following theorem to change the coordinate $(r\ \phi)$ to the coordinate (r', ϕ') shown in Fig. 1:

$$K_{l_1}\left(w_1 \frac{r}{a_1}\right) \begin{Bmatrix} \cos l_1\phi \\ \sin l_1\phi \end{Bmatrix} = \sum_{k=-\infty}^{\infty} (-1)^k K_k\left(w_1 \frac{d}{a_1}\right) I_{l_1+k}\left(w_1 \frac{r'}{a_1}\right)$$

$$\times \begin{Bmatrix} \cos\left[(l_1+k)\phi' - k\alpha\right] \\ \sin\left[(l_1+k)\phi' - k\alpha\right] \end{Bmatrix}. \qquad (20)$$

The coupling coefficient is,

$$C_{A_p B_q} = \frac{\omega}{2} \frac{\epsilon_0(n_0^2 - n_c^2)}{K_{l_1}(w_1)K_{l_2}(w_2)} A_{l_1}(A_{l_2})\pi[Q(l_1,l_2)]$$

$$\times \int_0^{a_2} r'dr'I_{l_2}\left(w_1 \frac{r'}{a_1}\right) J_{l_2}\left(u_2 \frac{r'}{a_2}\right)$$

$$= \frac{\dfrac{\omega}{2} \dfrac{\epsilon_0(n_0^2 - n_c^2)\pi}{K_{l_1}(w_1)K_{l_2}(w_2)} A_{l_1}A_{l_2}[Q(l_1,\, l_2)]}{\dfrac{\omega}{2} \dfrac{\epsilon_0(n_0^2 - n_c^2)\pi}{K_{l_1}(w_1)K_{l_2}(w_2)} A_{l_1}A_{l_2}[(Q(l_1,\, l_2)][R(l_1,\, l_2)]}, \qquad (21)$$

where

$$Q(l_1,\, l_2) = (-1)^{l_1-l_2}K_{l_1-l_2}\left(w_1 \frac{d}{a_1}\right) \cos(l_1 - l_2)\alpha$$

$$\pm (-1)^{l_1+l_2}K_{l_1+l_2}\left(w_1 \frac{d}{a_1}\right) \cos(l_1 + l_2)\alpha$$

$$R(l_1, l_2) = \frac{a_2}{\left(\dfrac{u_2}{a_2}\right)^2 + \left(\dfrac{w_1}{a_1}\right)^2}\left[\frac{u_2}{a_2}J_{l_2+1}(u_2)I_{l_2}\left(w_1\frac{a_2}{a_1}\right)\right.$$

$$\left. + \frac{w_1}{a_1}J_{l_2}(u_2)I_{l_2+1}\left(w_1\frac{a_2}{a_1}\right)\right].$$

If we substitute the eigenvalue equation into $R(l_1, l_2)$, we obtain

$$R(l_1, l_2) = \frac{a_2}{\left(\dfrac{u_2}{a_2}\right)^2 + \left(\dfrac{w_1}{a_1}\right)^2}\frac{J_{l_2}(u_2)}{K_{l_2(w_2)}}\times\left|\frac{w_2}{a_2}K_{l_2+1}(w_2)I_{l_2}\left(w_1\frac{a_2}{a_1}\right)\right.$$

$$\left. + \frac{w_1}{a_1}K_{l_2}(w_2)I_{l_2+1}\left(w_1\frac{a_2}{a_1}\right)\right|. \quad (22)$$

The final result for the coupling coefficient is

$$C_{A_pB_q} = \frac{kn_0a_2}{a_1}\frac{2\Delta_2u_1u_2}{v_1v_2}\left\{w_2K_{l_2+1}(w_2)I_{l_2+1}\left(w_1\frac{a_2}{a_1}\right)\right.$$

$$\left. + w_1\frac{a_2}{a_1}K_{l_2}(w_2)I_{l_2+1}\left(w_1\frac{a_2}{a_1}\right)\right\}$$

$$\times\frac{\left\{(-1)^{l_1-l_2}K_{l_1-l_2}\left(w_1\dfrac{d}{a}\right)\cos(l_1-l_2)\alpha\right.}{\left\{(u_2)^2 + \left(w_1\dfrac{a_2}{a_1}\right)^2\right\}}$$

$$\frac{\left. \pm (-1)^{l_1+l_2}K_{l_1+l_2}\left(w_1\dfrac{d}{a}\right)\cos(l_1+l_2)\alpha\right\}}{} . \quad (23)$$

$$\left(K_{l_1-1}(w_1)K_{l_1+1}(w_1)K_{l_2+1}(w_2)K_{l_2-1}(w_2)\right)^{1/2}$$

APPENDIX B

General Coupling Equation

The coupling between two fibers, as shown in Fig. 1, is expressed by the following coupling equation with the coupling coefficient defined by (23).

$$\frac{dA_p(z)}{dz} + j\beta_{A_p}A_p(z) = -j\sum_q B_q(z)C_{A_pB_q}$$

$$\frac{dB_q(z)}{dz} + j\beta_{B_q}B_j(z) = -j\sum_p A_p(z)C_{B_qA_p}$$

$$|C_{A_p B_q}| = \frac{(\gamma_1 \gamma_2)}{2} \frac{\sqrt{2\Delta_1}}{a_1 v_1} u_1 u_2$$

$$\times \left| \frac{K_{l_1-l_2}(w_1 d/a_1) \cos (l_1 - l_2)\alpha \pm K_{l_1+l_2}(w_1 d/a_1) \cos (l_1 + l_2)\alpha}{\left[u_2^2 + w_1^2 \left(\frac{a_2}{a_1}\right)^2 \right]} \right|$$

$$\times \frac{(w_2)K_{l_2+1}(w_2)I_{l_2}(w_1 a_2/a_1) + w_1(a_2/a_1)K_{l_2}(w_1)I_{l_2+1}(w_1 a_2/a_1)}{[K_{l_1-1}(w_1)K_{l_1+1}(w_1)K_{l_2-1}(w_2)K_{l_2+1}(w_2)]^{1/2}}, \quad (24)$$

where

$A_p(z)$ = the amplitude of the pth mode of fiber A

$B_q(z)$ = the amplitude of the qth mode of fiber B

$C_{A_p B_q}$ = coupling coefficient

β_{A_p} = the z-direction propagation constant of the pth mode in fiber A

β_{B_q} = the z-direction propagation constant of the qth mode in fiber B

l_1 = the azimuthal order number of the pth mode in fiber A

l_2 = the azimuthal order number of the qth mode in fiber B

$K_l(z)$ = the modified Bessel function

a_1 = the core radius of fiber A

a_2 = the core radius of fiber B

Δ_1 = the normalized index difference between the core and the cladding of fiber A

Δ_2 = the normalized index difference between the core and the cladding of fiber B

γ_1, γ_2 = when $l_1 = 0$, $l_2 = 0$, $\gamma_1 = \gamma_2 = 1$; when $l_1 = 0$, $l_2 = 0$, $\gamma_1 = \gamma_2 = 2$.

The power transfer from the pth mode of fiber A to qth mode of fiber B is given by the following equation where the pth mode of fiber A carries unit power and the qth mode of fiber B carries no power.

$$|A_p(z)|^2 = 1 - \kappa_{AB} \sin^2 \beta_{AB} z$$
$$|B_q(z)|^2 = \kappa_{BA} \sin^2 \beta_{AB} z, \quad (25)$$

where

$$\kappa_{AB} = \left[1 + \frac{(\beta_{A_p} - \beta_{B_q})^2}{4|(C_{A_p B_q})(C_{B_q A_p})|} \right]^{-1}$$

$$\kappa_{BA} = \left| \frac{C_{B_q A_p}}{C_{A_p B_q}} \right| \kappa_{AB}$$

$$\beta_{AB} = \left[\frac{|(C_{A_p B_q})(C_{B_q \cdot A_p})|}{\kappa_{AB}} \right]^{1/2}$$

z = the coupling length of two fibers.

Strong coupling between two modes is expected only when the propagation constants are matched. Therefore the coupling between two multimode fibers, each of which carries hundreds of modes, can be analyzed by the coupling of modes whose propagation constants are matched.

APPENDIX C

Simplification of the Modified Bessel Function

The asymptotic expansion for the modified Bessel function is given as follows

$$K_l(z) = \sqrt{\frac{\pi}{2z}}\, e^{-z} \sum_{n=0}^{\infty} \frac{(l, n)}{(2z)^n}, \tag{26}$$

where

$$|z| \text{ is large}$$

$$|\arg z| < 3\pi/2$$

$$(l,n) = \frac{(4l^2 - 1^2)(4l^2 - 3^2) \cdots (4l^2 - (2n-1)^2)}{n!2^{2n}} = \frac{\Gamma\left(l + n + \dfrac{1}{2}\right)}{n!\,\Gamma\left(l - n + \dfrac{1}{2}\right)}.$$

The parameters w and wd/a are large for most modes except those very close to the cutoff value. Therefore, most of the modified bessel functions can be expressed by the above asymptotic expansions. Especially when both order numbers of the mode are small (i.e., w is large), the modified Bessel function can be expressed by the first term of the asymptotic expansion. We obtain

$$\frac{K_{2l}(wd/a) \pm K_0(wd/a)}{K_{l+1}(w)K_{l-1}(w)}.$$

$$\cong 2\sqrt{\frac{2w}{\pi d/a}}\, e^{-w(d/a-2)} \qquad (\alpha = 0)$$

$$\cong 0 \qquad (\alpha = 90°). \tag{27}$$

We show below that (27) can also be used for large l. The coupling coefficient has the following modified Bessel function term:

$$\frac{K_{2l}(wd/a) \pm K_0(wd/a)}{K_{l+1}(w)K_{l-1}(w)}. \tag{28}$$

We combine (26) into (28); then the first term is,

$$\frac{K_{2l}(wd/a)}{K_{l+1}(w)K_{l-1}(w)} = \sqrt{\frac{2w}{\pi d/a}} \, e^{-w(d/a-2)}$$

$$\times \frac{\displaystyle\sum_{n=0}^{\infty} \frac{(2l, n)}{(2w)^n (d/a)^n}}{\displaystyle\sum_{n=0}^{\infty} \sum_{m=0}^{n} \frac{(l+1, n-m)(l-1, m)}{(2w)^n}}$$

$$\cong \sqrt{\frac{2w}{\pi d/a}} \, e^{-w(d/a-2)} \times \frac{\displaystyle\sum_{n=0}^{\infty} \left(\frac{1}{(2w)^n}\right)\left(\frac{a}{d}\right)^n [4(2l)^2]^n \frac{1}{n!2^{2n}}}{\displaystyle\sum_{n=0}^{\infty} \frac{1}{(2w)^n} (4(l)^2)^n \sum_{m=0}^{n} \frac{1}{(n-m)!m!} \frac{1}{2^{2n}}}$$

$$= \sqrt{\frac{2w}{\pi d/a}} \, e^{-w(d/a-2)} \frac{\displaystyle\sum_{n=0}^{\infty} \frac{1}{(2w)^n} \left(\frac{a}{d}\right)^n \frac{[4(l)^2]^n}{2^{2n} n!} 2^{2n}}{\displaystyle\sum_{n=0}^{\infty} \frac{1}{(2w)^n} \frac{2^n [4(l)^2]^n}{2^{2n} n!}}. \quad (29)$$

If $d/a \cong 2$, the above equation is

$$\frac{K_{2l}(wd/a)}{K_{l+1}(w)K_{l-1}(w)}$$

$$\cong \sqrt{\frac{2w}{\pi d/a}} \, e^{-w(d/a-2)} \frac{\displaystyle\sum \frac{1}{(2w)^n} \left(\frac{a}{d}\right)^n \frac{2^{2n}(4l^2)^n}{2^{2n}n!}}{\displaystyle\sum \frac{1}{(2w)^n} \frac{2^n(4l^2)^n}{2^{2n}n!}}$$

$$\cong \sqrt{\frac{2w}{\pi d/a}} \, e^{-w(d/a-2)}. \quad (30)$$

Therefore, when l is large, the modified Bessel function term is expressed by the above equation. Then,

$$\frac{K_{2l}(wd/a) \pm K_0(wd/a)}{K_{l+1}(w)K_{l-1}(w)} \cong \xi \sqrt{\frac{2w}{\pi d/a}} \, e^{-w(d/a-2)}, \quad (31)$$

where

$$\xi = 1 \text{ when } l \text{ large}$$

$$\xi = 2 \text{ when } l \text{ small, } w \text{ large (even)}$$

$$\xi \cong 0 \text{ when } l \text{ small, } w \text{ large (odd).}$$

Therefore, the coupling coefficient has the following limitation.

$$0 \leqq |\overline{C_i}| \leqq |C_{\text{limit}}| = \frac{2^{5/4}\Delta^{1/4}}{\sqrt{\pi k n_0}a^{3/2}} \left[\frac{i}{N}\right] \left[1 - \frac{i}{N}\right]^{1/4}. \qquad (32)$$

However, we assume that all Bessel function terms are expressed by the case $\xi = 1$ as the average coefficient. The result is

$$|\overline{C_i}| = \frac{2^{1/4}\Delta^{1/4}}{\sqrt{k n_0 \pi}a^{3/2}} \left[\frac{i}{N}\right] \left[1 - \frac{i}{N}\right]^{1/4}, \qquad (33)$$

where

$$\overline{C_i} = \text{average coupling coefficient.}$$

REFERENCES

1. A. W. Snyder, "Coupled-Mode Theory of Optical Fibers," J. Opt. Soc. Am., *62* (November 1972), pp. 1267–1277.
2. P. D. Holntyre and A. W. Snyder, "Power Transfer Between Optical Fibers," J. Opt. Soc. Amer., *63* (December 1973), pp. 1518–1527.
3. J. A. Arnaud, "Transverse Coupling in Fiber Optics—Part IV: Crosstalk," B.S.T.J., *54* (October 1975), pp. 1431–1450.
4. A. W. Snyder and P. McIntyre, "Crosstalk Between Light Pipes," J. Opt. Soc. Am., *66* (September 1976), pp. 877–882.
5. D. Gloge, "Weakly Guiding Fibers," Appl. Opt., *10* (October 1971), pp. 2252–2258.
6. D. Gloge, "Propagation Effects on Optical Fibers," IEEE Trans. Microw. Theory Tech., *MTT-23* (January 1975), pp. 106–120.

On the Design of Sub-band Coders for Low-Bit-Rate Speech Communication

By R. E. CROCHIERE

This paper presents a detailed discussion of issues involved in the design of sub-band coders for low-bit-rate speech communications. Specifically, bit rates in the range of 7.2 to 16 kb/s are emphasized. Design guidelines, based on results of extensive computer simulations and subjective comparisons, are presented for selection of sub-band coder parameters. Practical considerations for selecting sub-bands under integer-band sampling and multiplexing constraints are also discussed, and a method for synchronous multiplexing of the sub-band data, without buffering, is proposed. Several examples of sub-band coders for transmission rates of 7.2, 9.6, and 16 kb/s are presented, and the quality of these coders is compared against that of ADPCM and ADM coders.

I. INTRODUCTION

In recent work by Crochiere, Webber and Flanagan,[1] an approach to speech encoding has been proposed which is based on the partitioning of the speech band into sub-bands and encoding the sub-bands individually. The technique offers attractive possibilities for coding speech economically at bit rates in the range of 7.2 to 16 kb/s. At 16 kb/s good quality encoding, comparable to that of 26.5 kb/s adaptive differential (fixed predictor) PCM (ADPCM) encoding, is possible. Potential applications exist in areas of narrow-band communications, mobile radio, and voice storage applications.

When the bit rate is extended down into the upper data rate range of 9.6 and 7.2 kb/s, moderate quality encoding can be achieved comparable to that of 19 and 18 kb/s adaptive delta modulation (ADM), respectively. Interesting potential applications exist for voice coordination on digital data lines and for secure voice communications by digital encryption and transmission over conventional data lines.

In the design of sub-band coders, a variety of issues and "trade-offs" must be dealt with. The number of sub-bands, the partitioning of sub-
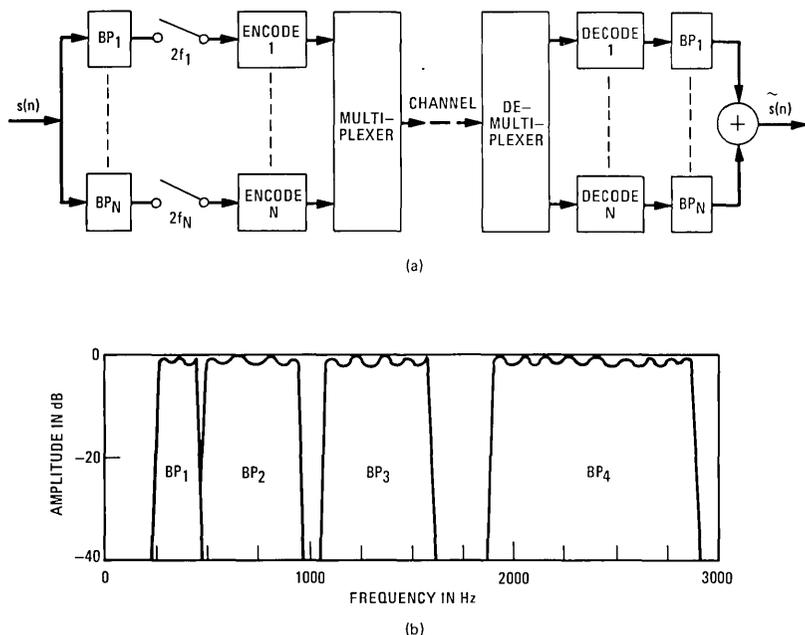
Fig. 1—(a) Implementation of a sub-coder based on integer-band sampling. (b) Frequency-domain illustration of the sub-band partitioning of the speech band.

bands (and gaps between bands), coder parameters, parceling of bits among sub-bands, and compromises between bits/sample and bandwidth are all variables that must be considered. In addition, a number of constraints are introduced by practical considerations of multiplexing the digitized sub-band signals and by considerations of efficient hardware implementation. In this paper, we attempt to clarify these issues and present useful criteria and guidelines for designing sub-band coders. In many respects, the only truly meaningful criterion for selecting parameters of the sub-band coder is a perceptual one. Therefore, design criteria have been supported, as much as possible, by results of extensive computer simulations and listener preference tests.

II. A REVIEW OF SUB-BAND CODERS

In the sub-band coder, the speech band is partitioned into sub-bands by bandpass filters. Each sub-band is low-pass translated, sampled at its Nyquist rate, and digitally encoded. By this process of dividing the speech band into sub-bands, each sub-band can be preferentially encoded according to perceptual criteria for that band. On reconstruction, sub-band signals are decoded and bandpass translated back to their original bands. They are then summed to give a replica of the original speech signal.
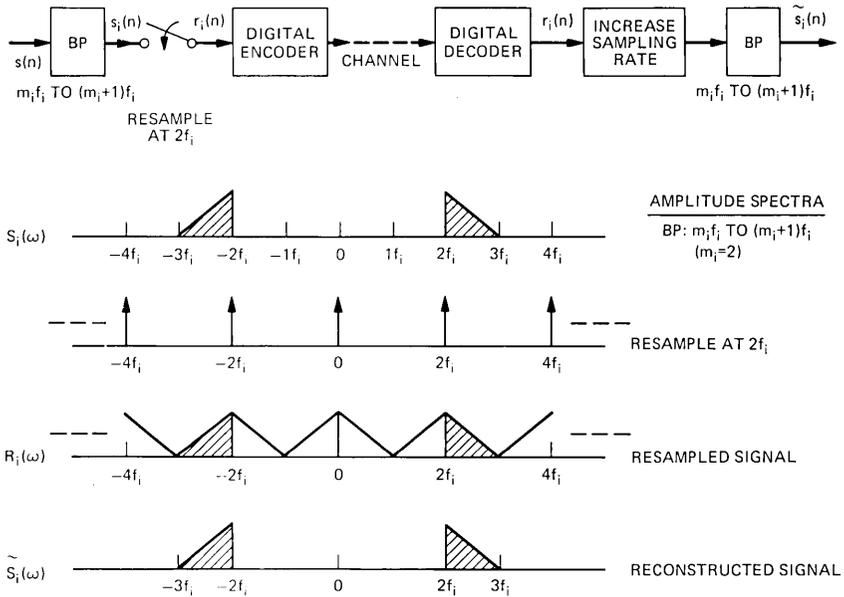
Fig. 2—Integer-band sampling technique and a frequency-domain interpretation.

A variety of techniques exists for performing the low-pass and bandpass translations. However, one approach is particularly attractive for hardware implementation since it eliminates the need for modulators. It is based on the integer-band sampling method proposed in Ref. 1 and will be the method primarily considered in this paper.

The integer-band sampling implementation of the sub-band coder is illustrated in Figs. 1 and 2. The speech band is partitioned into N sub-bands by bandpass filters BP_1 to BP_N. It will be assumed in this paper that the filters are discrete-time (e.g., digital or CCD) filters. Typically four or five bands are used and, at lower bit rates, small gaps are permitted between bands to conserve bandwidth and bit rate, as illustrated in Fig. 1b.

The output of each filter in the transmitter is resampled at a rate of $2f_i$, where f_i is the width of the sub-band and i refers to the ith sub-band. The sampled sub-band signals are digitally encoded and time multiplexed for transmission over the digital channel. At the receiver the digital signals are demultiplexed and decoded. The sub-band signals are reconstructed by filtering the outputs of the decoders with another set of bandpass filters, identical to BP_1 to BP_N, that act as interpolating filters. Prior to this filtering, the sampling rates of the decoder outputs are increased to the original sampling rate of $s(n)$ by filling in with zero-valued samples. The outputs of these filters are then summed to give a reconstructed replica $\hat{s}(n)$ of the original speech signal $s(n)$.

The integer-band sampling scheme imposes certain constraints on the choice of sub-bands, as illustrated in Fig. 2. Sub-bands are required to be between $m_i f_i$ and $(m_i + 1)f_i$, where m_i is an integer. This constraint is necessary to avoid aliasing in the sampling process.

Encoding in sub-bands offers several advantages over full-band coding.[1] Quantization noise can be contained in bands to prevent masking of one frequency range by quantizing noise in another frequency range. Separate quantizer step-sizes are used in each band. Therefore, bands with lower signal energy will have lower quantizer step-sizes and contribute less quantization noise. Finally, the partitioning of the speech band into sub-bands enables the parceling of bits in bands according to perceptual criteria. In lower bands where pitch and formant structure must be accurately preserved, a larger number of bits/sample can be used for encoding, whereas in upper bands where fricatives and noise-like sounds occur in speech, fewer bits/sample can be used.

In the following sections, we focus on the various issues involved in the design of sub-band coders. Section III addresses issues of coder selection for sub-bands and the choice of their parameters. "Trade-offs" involved in the allocation of bits among bands are also discussed. Section IV deals with problems of sub-band partitioning of the speech spectrum under the constraints of integer-band sampling requirements and multiplexing requirements. Section V involves issues in the design of filters for the sub-bands. Finally, Section VI presents further results on comparisons of sub-band coder performance with other waveform coding methods.

III. SELECTION OF CODERS AND CODER PARAMETERS FOR SUB-BANDS

Because encoders are individually tailored to each sub-band, a spectrum of coders and parameters must be considered. For the lower-frequency sub-bands, typically 3 or 4 bits/sample encoders are used, and for upper bands 2 or less bits/sample are used. Since the characteristics of the sub-band signals are considerably different from those of full-band speech, encoding techniques developed for encoding of full-band speech signals do not necessarily lead to good results for encoding of sub-band signals. In this section, we therefore address issues in the design of encoders for sub-band signals and in the parceling of bits among bands.

The choice of encoder parameters is determined in part by the static or long-term spectral characteristics of the speech waveform. Figure 3a illustrates typical long-term speech spectra (averaged over a sentence) based on measurements made by Beranek[2] and Dunn and White.[3] The same spectra are plotted in Fig. 3b with a warped frequency scale based on a constant (5 percent/division) contribution to the articulation index[2] in order to illustrate the relative perceptual importance of the various
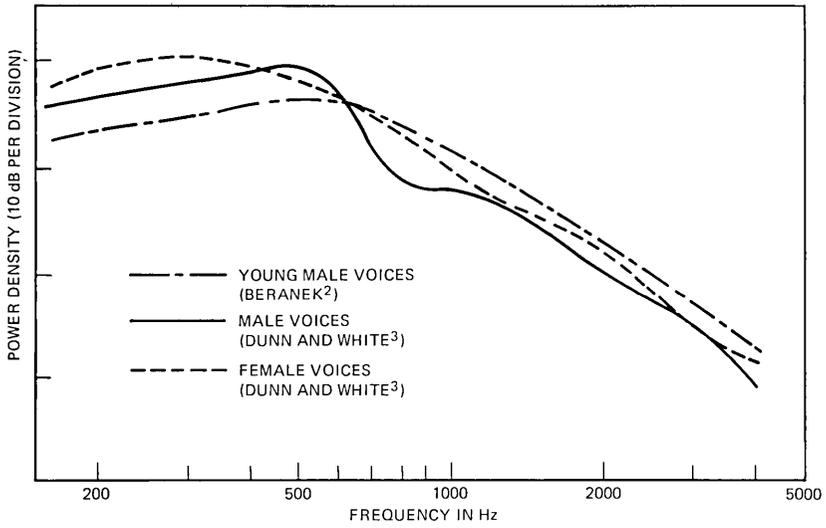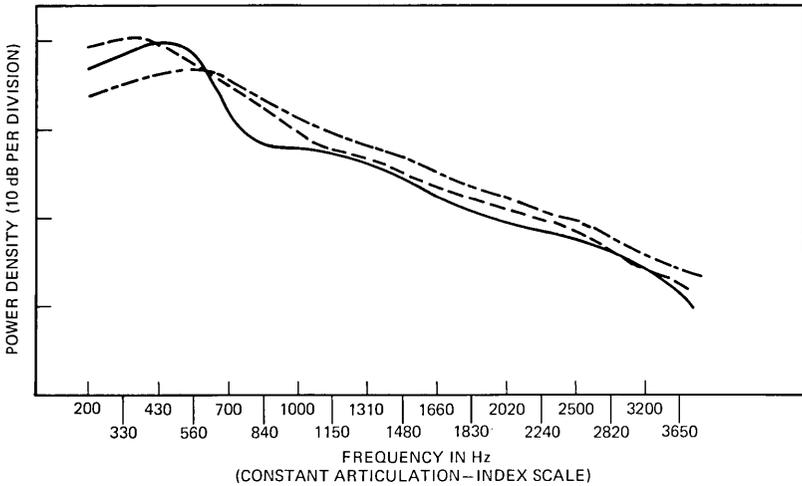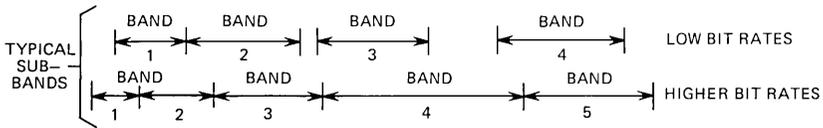
Fig. 3—Long-term spectrum of speech based on measurements by Beranek[2] and Dunn and White.[3] (a) Logarithmic frequency scale. (b) Frequency scale based on a constant contribution to the articulation index.

frequencies. Two possibilities for sub-band selection for low and high bit rates (to be discussed later) are illustrated above Fig. 3b. It is seen that across the entire speech spectrum there is a characteristic drop in

SUB-BAND CODER DESIGN   751

BAND 1

→| |← 1/2f$_i$ = 2.08 ms

BAND 2

→| |← 1/2f$_i$ = 1.04 ms

BAND 3

→| |← 1/2f$_i$ = 0.94 ms

BAND 4

→| |← 1/2f$_i$ = 0.52 ms
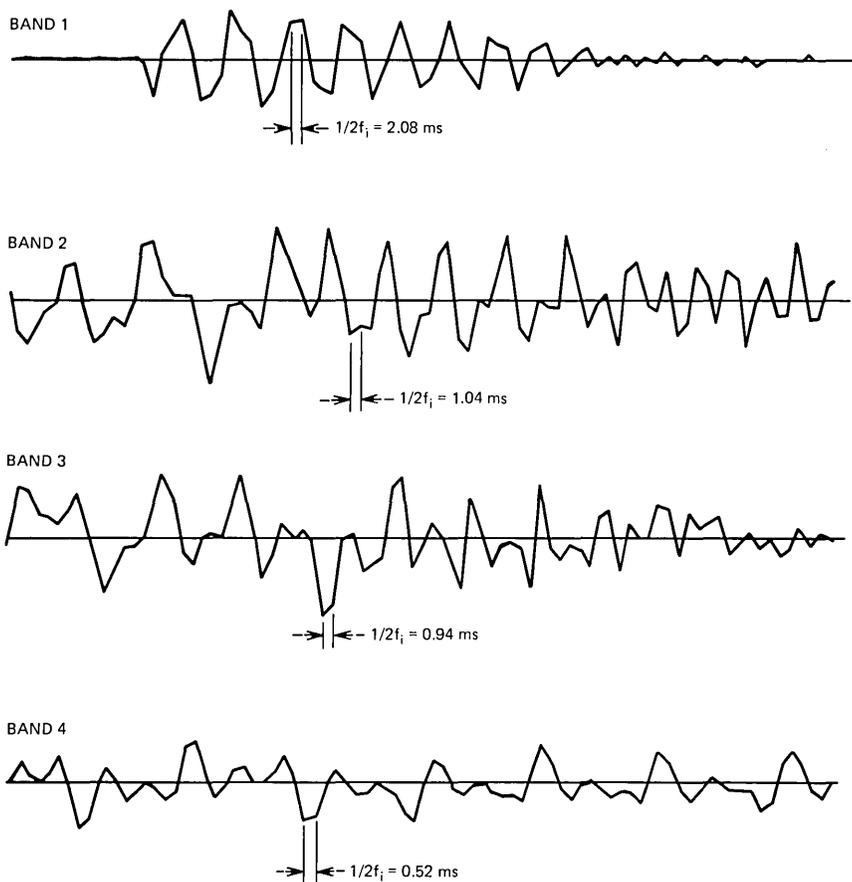
Fig. 4—Typical waveforms of uncoded sub-band signals for bands 1 to 4. Eighty samples are plotted on each line.

power density with increasing frequency. Across any one band, however, the drop in power density is relatively small. Since sub-bands are, in effect, low-pass translated and sampled at their Nyquist rate, they appear essentially as flat spectrum signals at the low sub-band sampling rates and have essentially no sample-to-sample correlation. Figure 4 shows examples of sub-band signals for bands 1 to 4. Because of their low sample-to-sample correlation, encoding is best performed by adaptive PCM (APCM). Encoding based on differential or fixed prediction, commonly used for full-band encoding, does not lead to good results for encoding of sub-band signals.

The step-size adaption strategy used in simulations for the APCM coders is based on the one-word step-size memory approach proposed by Jayant, Flanagan, and Cummiskey.[4,5] The coder input signal, denoted
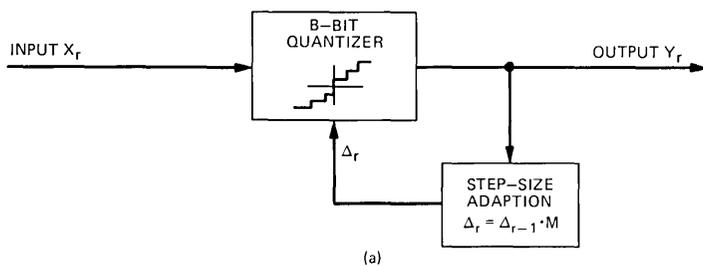
Fig. 5—Step-size adaption algorithm and quantizer characteristics of the APCM coders.

as x_r for the rth sample, is quantized to one of 2^B levels according to the quantizer characteristics shown in Fig. 5, where B is the number of bits in the coder. The step-size adaption circuit examines the quantizer output bits for the $(r-1)$th sample and computes the quantizer step-size, Δ_r, for the rth sample according to the relation

$$\Delta_r = \Delta_{r-1} M(L_{r-1}), \tag{1a}$$

where

$$\Delta_{\text{MIN}} \leqq \Delta_r \leqq \Delta_{\text{MAX}}, \tag{1b}$$

and where Δ_{r-1} is the step size used for the $(r-1)$th sample. $M(L_{r-1})$

Table I — APCM coder parameters

B =	4	3	2	$1\frac{1}{2}$	$1\frac{1}{3}$	$1\frac{1}{4}$
M_1	0.9	0.85	0.85	0.92	0.92	0.92
M_2	0.9	1.0	1.9	1.4	1.4	1.4
M_3	0.9	1.0				
M_4	0.9	1.5				
M_5	1.2					
M_6	1.6					
M_7	2.0					
M_8	2.4					
Typical s/n (dB)	18	11.5	7	4	3.3	2.5

is a multiplication factor whose value depends on the quantizer magnitude level L_{r-1} at time $r-1$. It can take on one of 2^{B-1} values, $M_1, M_2, \cdots M_{2^{B-1}}$. If the lower-magnitude quantizer levels are used at time $r-1$, a value of $M(L_{r-1}) = M_i$ less than one is used to reduce the next step size. If upper-magnitude levels are encountered, a value of M_i greater than one is chosen. In this way, the coder continuously adapts its step size in an attempt to track the short-time variance of the input signal. For practical reasons, the step size, Δ_r, is constrained to be between some minimum and maximum value Δ_{MIN} and Δ_{MAX}, respectively.

Typical values of M_i for 2-, 3-, and 4-bit APCM coders are given in Table I. These values were determined experimentally and were found to agree reasonably well with values reported by Jayant[4] for encoding of full-band speech. As observed by Jayant, small changes in these values do not strongly affect the performance of the coders. Typical signal-to-quantizing noise ratios (s/n) found for encoding sub-band signals are also reported in Table I.

An interesting modification to the above algorithm, proposed by Goodman,[6] allows for encoding at an average bit rate of $1 + 1/K$ bits/sample, where K is an integer. In this approach, the sign of the signal x_r is encoded for each sample, r, and the magnitude of the signal is encoded with one bit every K samples. The step-size adaption is essentially that of (1) with M_i and the quantizer magnitude level repeated for $K-1$ samples at the decoder. For example, if $K = 2$, a sign and a magnitude bit are transmitted on odd numbered samples. On even numbered samples, only the sign bit is transmitted and the magnitude bit is assumed to be that of the previous sample. The sign bit transmits essentially the "zero crossing" or phase information and the magnitude bit conveys the amplitude information in the waveform at a reduced rate.

The $1 + 1/K$ bit coder is found to be useful for encoding the uppermost bands when overall bit rates must be kept low. The upper bands contain

primarily the fricative and noise-like sounds in the speech and can therefore be quantized more coarsely than lower bands without a perceived loss in quality. Typical adaption parameters found to be useful for $1 + 1/K$ bit coders are also given in Table I.

The quantities Δ_{MAX} and Δ_{MIN} in the above algorithms represent practical constraints in the adaption logic. Their ratio determines the dynamic range that the coder can handle and their absolute values determine the center of this dynamic range. In simulations, a ratio $\Delta_{\text{MAX}}/\Delta_{\text{MIN}} = 128$ was consistently used, resulting in a useful dynamic range of about 40 dB for the coders. The actual values of Δ_{MIN} and Δ_{MAX} must be different for each sub-band, however, to match properly the dynamic range characteristics of the sub-band coder to that of the long-term speech spectrum. This is easily seen in Fig. 3. Since upper sub-bands have lower power densities than lower sub-bands, they should have smaller values of Δ_{MAX} and Δ_{MIN} in their coders. A useful criterion for choosing relative values of $\Delta_{\text{MIN}}(\Delta_{\text{MAX}} = 128\Delta_{\text{MIN}})$ can be derived by assuming that the power-density spectrum in sub-band i is approximately flat across the band and has a value S_i. The long-term variance, $\sigma_i{}^2$, of the sub-band signal is then proportional to $S_i f_i$.

To match the center of the dynamic range of the coders in each band, Δ_{MIN} should be selected to be proportional to the square root of the long-term variance of the signal in that band. Therefore, the ratio of $\Delta_{\text{MIN}}(\text{band } i)$ in band i to $\Delta_{\text{MIN}}(\text{band } j)$ in band j can be determined as

$$\frac{\Delta_{\text{MIN}}(\text{band } i)}{\Delta_{\text{MIN}}(\text{band } j)} \cong \frac{\sigma_i}{\sigma_j} \cong \sqrt{\frac{S_i f_i}{S_j f_j}} \tag{2}$$

or if values are expressed in dB, (2) becomes

$$\left. \frac{\Delta_{\text{MIN}}(\text{band } i)}{\Delta_{\text{MIN}}(\text{band } j)} \right|_{\text{dB}} \cong S_i\big|_{\text{dB}} - S_j\big|_{\text{dB}} + 20 \log \sqrt{\frac{f_i}{f_j}}. \tag{3}$$

Equation (3) states that the ratio of minimum step size (in dB) of band i to band j is equal to the difference in power densities (in dB) between band i and band j plus a correction factor to account for the differences in bandwidths. Values of S_i and S_j can be obtained from Fig. 3. Although eq. (3) is only approximate, it serves as a useful criterion for choosing relative values of Δ_{MIN} for coders. Good agreement was found with experimentally derived values.

A final consideration in the selection of coders for sub-bands relates to the questions of how many bits/sample should be allocated to each sub-band under constraints of fixed total transmission rate and how should the sub-band bandwidths and gaps between bands be traded against bits/sample for the coders. The answer to both questions is highly dependent on perceptual criteria and is greatly influenced by the overall

allowed transmission rate. Therefore, we do not propose to answer these questions in detail but simply provide some insight.

A useful measure for assisting in the parceling of bits among sub-bands is the signal-to-quantizing noise ratio (s/n) as a function of frequency. Figure 6 shows typical s/n values as a function of frequency that are found to give preferred signal quality at bit rates of 16, 9.6, and 7.2 kb/s, respectively. At 16 kb/s it is found that good quality coding can be achieved with an allocation of 4 bits/sample ($\approx$18 dB s/n) in the lower sub-bands, 3 bits/sample ($\approx$11.5 dB) in the middle sub-bands, and 2 bits/sample ($\approx$7 dB) in the upper sub-bands. Contiguous sub-bands are used. One possible choice of sub-bands is shown above Fig. 6 and will be discussed in greater detail in the next section.

In the other extreme, moderate quality coding at transmission rates of 7.2 kb/s can be achieved by trimming the lowest band to 3 bits/sample, the second band to 2 bits/sample, and the upper bands to $1\frac{1}{3}$ or $1\frac{1}{4}$ bits/sample. In addition, to conserve bandwidth, gaps may be allowed between sub-bands as shown in the band arrangement above Fig. 6. While these gaps introduce a slightly reverberant quality to the coder, the reverberation is generally preferred at this transmission rate to a further reduction in bits/sample and a corresponding increase in noise in the coders, which would be necessary if gaps were not present.

At the intermediate transmission rate of 9.6 kb/s, a distribution of 3, 2, and $1\frac{1}{2}$ bits/sample is possible across the frequency ranges, as shown in Fig. 6 by the solid line. A second alternative, which is also judged close in quality, is given by the dotted line. In this case, 3 bits/sample is used only in the lowest band and 2 bits/sample is used for encoding all upper bands. In both cases, gaps are allowed between bands, as shown above the figure. In listener preference comparisons, 63 percent of the listeners preferred the quality of the first bit/sample distribution (solid line) and 37 percent preferred the quality of the second distribution (dotted line). A third approach was also tried at 9.6 kb/s, which involved 3 bits in the lowest band, 2 bits in the second band, and $1\frac{1}{2}$ bits in the two upper bands, with no gaps appearing between bands. In this way, the reverberant quality of the coder was traded for slightly lower overall s/n. This approach was preferred by only 13 percent of the listeners over that of the first distribution (solid line) and by only 37 percent of the listeners over that of the second distribution (dotted line). Therefore, at 9.6 kb/s, a slight reverberant quality in the coder is preferred by listeners over the lower s/n obtained if no gaps between sub-bands are used.

As observed in the above discussion, many "trade-offs" are possible and the only meaningful criterion for comparing them is a perceptual one. Often it is a matter of trading one type of distortion for another with the hope of finding a compromise that is most acceptable to the majority of listeners.
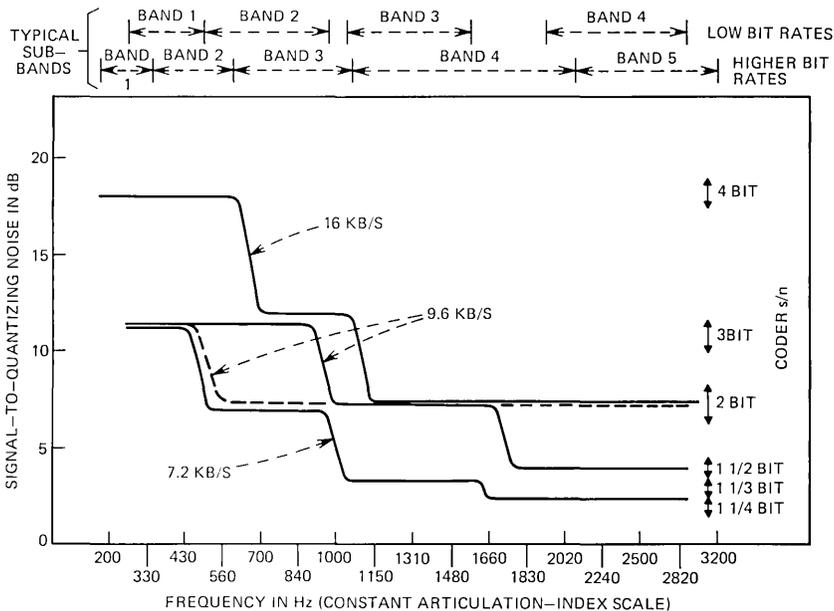
Fig. 6—Signal-to-quantizing noise ratio (s/n) as a function of frequency for bit allocations for 16-μ, 9.6-μ, and 7.2-kb/s coders.

IV. PARTITIONING OF THE SPEECH BAND INTO SUB-BANDS AND MULTIPLEXING OF DATA

The selection of sub-bands involves a variety of considerations. Of preliminary interest is the number of bands. Next, bandwidths and locations of sub-bands must be chosen. This choice is strongly influenced by constraints imposed by the integer-band sampling technique and multiplexing requirements. In this section, we discuss these issues and present candidates for sub-band coders at various bit rates.

Through simulations, a good compromise in the number of bands necessary for sub-band coding was generally found to be about four or five bands. When less than four bands are used, bandwidths become too wide and do not allow for full utilization of the advantages of sub-band encoding. Designs with more than four or five bands tend to consume bandwidth in transition bands of filters in addition to requiring more hardware for practical implementation.

The partitioning of the speech band into sub-bands presents a more difficult problem. A useful preliminary guideline for choosing sub-bands, suggested in Ref. 1, is to partition the speech band into sub-bands that represent approximately equal contributions to the articulation index (AI) under noiseless conditions. In this way each sub-band contains a significant portion of the important frequencies of the speech band. Lower sub-bands should have narrower bandwidths and bandwidths

Table II — Choice of bands for integer-band sampling
and 9.6-kHz sampling rate

Decimation Ratio	f_i	$2f_i$	$3f_i$	$4f_i$
1	4800	9600	14400	19200
2	2400	4800	7200	9600
3	1600	3200	4800	6400
4	1200	2400	3600	4800
5	960	1920	2880	3840
6	800	1600	2400	3200
7	686	1371	2057	2743
8	600	1200	1800	2400
9	533	1067	1600	2133
10	480	960	1440	1920
11	436	873	1309	1745
12	400	800	1200	1600
13	369	738	1108	1477
14	343	686	1029	1371
15	320	640	960	1280
16	300	600	900	1200
17	282	565	847	1129
18	267	533	800	1067
19	253	505	758	1011
20	240	480	720	960
21	229	457	686	914
22	218	436	655	873
23	209	417	626	835
24	200	400	600	800
25	192	384	576	768

should become progressively wider with increasing frequency. Gaps between sub-bands can also be determined by this criterion. The allocation of bits in sub-bands, however, is made according to subjective quality considerations, as discussed in the previous section.

The integer-band sampling scheme imposes the constraint that the ratio of upper to lower band edges of sub-bands be $(m_i + 1)/m_i$, where m_i is an integer that may be different for different bands (see Fig. 2). For hardware considerations, it is required that the sampling rates for sub-bands be derivable from a common clock. Furthermore, for digital or CCD hardware implementations, it is desirable to relate these sampling rates to the sampling rate of the bandpass filters by ratios that are integers. Finally, the requirements for multiplexing digitally encoded sub-band signals dictate that the transmission bit rates of each sub-band be a rational fraction of the total bit rate so that the data can be framed and synchronized. Also, a small fraction of this total bit rate must be reserved for synchronizing and framing information.

This multitude of constraints greatly restricts the choices for sub-bands. To assist in the selection of sub-bands, it is helpful to construct tables such as Table II. It is assumed in Table II that the sampling rate of the bandpass filters is 9.6 kHz. Column 1 indicates the integer deci-

mation (reduction) ratios that relate sub-band sampling rates to 9.6 kHz. Column 2 gives bandwidths, f_i, and column 3 gives $2f_i$ sampling rates for the possible sub-bands. Columns 2 through 4 specify choices for band edges $m_i f_i$ ($m_i = 1,2,3, \cdots$). Therefore, all choices for sub-bands are discernible from the tables once the sampling rate for the filters is chosen. Considerations in selecting sub-bands on the basis of articulation index, the distribution of bits/sample across bands, and the total transmission rate quickly reduce the choices of sub-bands further to only a few possibilities. The final choice is still not complete, however, without an analysis of multiplexing requirements. Practically, the transmission rate of each sub-band must be a rational fraction of the total bit rate so that the sub-band data can be multiplexed into a repetitive framed sequence. The lowest common denominator of these rational fractions, including the fraction of transmission rate reserved for synchronization, determines the smallest frame size.

To illustrate these points more clearly, it is helpful to analyze several examples of coders. Table III shows one choice of sub-bands that can be used for 9.6 and 7.2 kb/s four-band coders. The selection of sub-bands is obtained from Table II and corresponds to the low-bit-rate sub-band arrangement illustrated in Figs. 1(b), 3(b), and 6. As seen in Fig. 3(b) or Fig. 6, the bands all have approximately equal width on the warped frequency (constant AI) scale. The lowest sub-band is slightly narrower due to constraints imposed by integer-band sampling. A 107-Hz gap appears between sub-bands 2 and 3 and a 320-Hz gap appears between sub-bands 3 and 4, giving the coders a slightly reverberant quality.

Coder examples A and B represent 9.6 kb/s coders with bit parceling among sub-bands according to distributions shown in Fig. 6 by solid and dotted lines for 9.6 kb/s. Example C is a 7.2 kb/s coder with the bit allocation in Fig. 6. Also included in Table III are sampling rate reduction (decimation) ratios and sampling rates for sub-bands. Relative values of minimum coder step-size (expressed in dB) that match the long-term speech spectrum, as discussed in Section III, eq. (3), are given in column 5. Finally, typical s/n values observed for the examples are given at the bottom of the table. They were measured by comparing simulations with and without coders and represent distortions only contributed by coders and not due to band gaps or filtering.

A fourth coder, example D, was designed for 16 kb/s. The design is based on a filter sampling rate of 10.67 kHz ($\frac{2}{3} \times 16$), which gives the choice of sub-bands shown in Table IV. This led to a slightly better selection of sub-bands for the 16 kb/s coder and resulted in the five-band coder design given in Table V. The sub-band selection corresponds to that shown above Figs. 3(b) and 6. Lower sub-bands overlap slightly to allow for transition bands of filters so that no gaps appear in this frequency range.

Table III — Sub-band coder designs for 9.6 and 7.2 kb/s

Band	Decimate From 9.6 kHz	Band Edges (Hz)	Sub-band Sampling Rates (Hz)	Δ_{MIN} Ratios (dB)	Example A 9.6 kb/s coder		Example B 9.6 kb/s coder		Example C 7.2 kb/s coder	
					Bits/samp.	kb/s	Bits/samp.	kb/s	Bits/samp.	kb/s
1	20	240–480	480	0 (Ref.)	3	1.44	3	1.44	3	1.44
2	10	480–960	960	−3	3	2.88	2	1.92	2	1.92
3	9	1067–1600	1067	−8.5	2	2.13	2	2.13	$1\frac{1}{3}$	1.42
4	5	1920–2880	1920	−14	$1\frac{1}{2}$	2.88	2	3.84	$1\frac{1}{4}$	2.40
Sync						0.27		0.27		0.02
			Total Bit Rate (kb/s)			9.60		9.60		7.20
			Typical s/n (dB)			10.8		10		8.9

Table IV — Choice of bands for integer-band sampling and 10.67-kHz sampling rate.

Decimation Rate	f_i	$2f_i$	$3f_i$	$4f_i$
1	5333	10667	16000	21333
2	2667	5333	8000	10667
3	1778	3556	5333	7111
4	1333	2667	4000	5333
5	1067	2133	3200	4267
6	889	1778	2667	3556
7	762	1524	2286	3048
8	667	1333	2000	2667
9	593	1185	1778	2370
10	533	1067	1600	2133
11	485	970	1455	1939
12	444	889	1333	1778
13	410	821	1231	1641
14	381	762	1143	1524
15	356	711	2133	1422
16	333	667	1000	1333
17	314	627	941	1255
18	296	593	889	1185
19	281	561	842	1123
20	267	533	800	1067
21	254	508	762	1016
22	242	485	727	970
23	232	464	696	928
24	222	444	667	889
25	213	427	640	853
26	205	410	615	821
27	198	395	593	790
28	190	381	571	762
29	184	368	552	736
30	178	356	533	711

Table V — Sub-band coder design for 16 kb/s

Band	Decimate From 10.67 KHz	Band Edges (Hz)	Sub-band Sampling Rates (Hz)	Δ_{MIN} Ratios (dB)	Example D 16-kb/s Coder	
					Bits	kb/s
1	30	178–356	356	−2	4	1.42
2	18	296–593	593	0 (Ref.)	4	2.37
3	10	533–1067	1067	−6	3	3.20
4	5	1067–2133	2133	−11.5	2	4.27
5	5	2133–3200	2133	−18	2	4.27
Sync						0.47
				Total Bit Rate (kb/s)		16.00
				Typical s/n (dB)		13.6

The analysis of the multiplexing requirements for coder examples A through D is summarized in Table VI. The required frame length for multiplexing is 180 bits for the 9.6-kb/s coders, 405 bits for the 7.2-kb/s coder and 135 bits for the 16-kb/s coder. The frame length corresponds

Table VI — Multiplexing and framing information for sub-band coder examples

Band	Fraction of Total Bit Rate	Samples/Frame
Example A (9.6 kb/s)	Frame Length = 180 Bits	
1	27/180	9
2	54/180	18
3	40/180	20
4	54/180	36
Sync	5/180	—
Example B (9.6 kb/s)	Frame Length = 180	
1	27/180	9
2	36/180	18
3	40/180	20
4	72/180	36
Sync	5/180	—
Example C (7.2 kb/s)	Frame Length = 405 Bits*	
1	81/405	27
2	108/405	54
3	80/405	60
4	135/405	108
Sync	1/405	—
Example D (16 kb/s)	Frame Length = 135 Bits	
1	12/135	3
2	20/135	5
3	27/135	9
4	36/135	18
5	36/135	18
Sync	4/135	—

* See text.

to the number of bits that must be stored or transmitted before the multiplexing pattern repeats itself. It is determined by the lowest common denominator of the fractions of total bit rate contributed by sub-bands and by the synchronization channel in column 2. If the frame length is too large, a different sub-band arrangement or bit allocation must be chosen. For example, in the 7.2-kb/s coder, only 1 bit in a frame of 405 bits is reserved for synchronization. If, the third sub-band is quantized with $1\frac{1}{4}$ bits/sample, a frame length of 135 bits is possible with 2 bits reserved for synchronization. This is achieved, of course, at a cost of a slightly reduced coder quality. Column 3 in Table VI gives the number of sub-band samples represented by each frame of data.

The fact that the sub-bands are multiplexed in frames does not necessarily imply that a complete frame of data must be stored before transmission. By careful design of the multiplexer, it is possible to synchronously encode the sub-bands and multiplex them without buffering the data. One scheme for doing this, for coder example A, is illustrated in Table VII. The table depicts the bit allocation for one frame (180 bits)

Table VII — Synchronous multiplexing of coder example A

Bit No.	Band 1	2	3	4	S
1	1		1		X
2					X
3					X
4				1	X
5				X	X
6				X	
7		1		X	
8			X		
9			X	2	
10			2	X	
11		X			
12		X			
13		X			
14				3X	
15				X	
16			X		
17		2	X		
18	X				
19	X		3	4	
20	X				
21	2			X	
22			X		
23			X		
24		X		5	
25		X			
26		X			
27		3		X	
28			4	X	
29			X	6	
30			X		

Bit No.	Band 1	2	3	4	S
31				X	
32		X			
33		X			
34		X		7	
35				X	
36				X	
37	X	4	5		
38	X				
39	X			8	
40				X	
41	3		X		
42			X		
43		X			
44		X		9	
45		X			
46			6	X	
47		5		X	
48			X		
49			X	10	
50		X			
51		X			
52		X			
53				X	
54	X			11	
55	X		7		
56	X				
57		6		X	
58				X	
59			X	12	
60			X		

Bit No.	Band 1	2	3	4	S
61	4			X	
62		X			
63		X			
64		X	8	13	
65				X	
66				X	
67		7	X		
68			X		
69		X		14	
70		X			
71		X			
72				X	
73	X		9		
74	X			15	
75	X				
76				X	
77		8		X	
78			X		
79			X	16	
80		X			
81	5	X			
82		X	10		
83				X	
84			X	17	
85			X		
86				X	
87		9		X	
88		X			
89		X		18	
90		X			

Table VII — Continued

Bit No.	Band 1	Band 2	Band 3	Band 4	S	Bit No.	Band 1	Band 2	Band 3	Band 4	S	Bit No.	Band 1	Band 2	Band 3	Band 4	S
91			11	X		121	7	X				151		X			
92	X					122		X				152		X			
93	X					123				X		153				X	
94	X			19		124	X			25		154			18X	31	
95				X		125	X					155			X		
96				X		126	X					156				X	
97		10	X			127		13	15	X		157		16		X	
98			X			128				X		158				X	
99		X		20		129			X	26		159		X		32	
100		X	12			130			X			160		X			
101	6	X				131				X		161	9			X	
102				X		132		X				162	X				
103			X			133		X				163	X		19		
104			X	21		134		X		27		164	X			33	
105				X		135				X		165			X		
106				X		136			16	X		166			X		
107		11X				137		14		X		167		17		X	
108		X				138				X		168				X	
109		X	13	22		139		X		28		169		X		34	
110				X		140		X				170		X			
111			X			141	8	X				171		X			
112			X			142				X		172			20	X	
113	X					143	X					173			X		
114	X			23		144	X			29		174			X	35	
115	X					145	X		17			175				X	
116				X		146				X		176				X	
117		12		X		147		15		X		177		18X		X	
118			14X			148			X			178		X			
119			X	24		149			X	30		179		X		36	
120		X				150		X				180				X	

of data. The first column gives the bit or clock number (at a clock rate of 9.6 kb/s), the next four columns represent sub-bands, the last column represents the synchronization channel and the X's represent allocated bits. Numbers and partitions in each sub-band column represent coder sampling times and sampling intervals. For example, in the first sub-band, nine samples of data (see Table VI) are coded with 3 bits/sample at appropriate clock times, $1, 21, 41, \cdots, 161$. This corresponds to one sample every 20 clock times, which is the decimation ratio of sub-band 1 (see Table III). Within each sampling interval three slots (X's) are allocated for transmission of these three bits and, therefore, they do not have to be stored for more than one sampling interval. In the fourth sub-band, bit allocations alternate between two slots and one slot per sampling interval according to the needs of the 1½-bit coder. A frame sequence begins with the transmission of five synchronization bits. The sampling intervals of the sub-bands are offset in time so that these five bits can be transmitted together without conflict. The scheme could easily be implemented with the aid of a read-only memory (ROM).

The synchronous multiplexing scheme is also useful as a means for conveniently ordering bits in a frame even if frames must be buffered for other purposes. Another potentially useful application of synchronous multiplexing occurs in an all-digital implementation, where coder hardware and possibly filter hardware can be shared between sub-bands.

V. DESIGN AND IMPLEMENTATION OF THE FILTERS

The parameters of the bandpass filters are depicted in Fig. 7. The sub-band covers the frequency range from $m_i f_i$ to $(m_i + 1)f_i$. For practical reasons the filter passband must have a slightly narrower frequency range from $m_i f_i + \Delta f$ to $(m_i + 1)f_i - \Delta f$. A transition region, Δf, on the order of 50 to 60 Hz was used in simulations with good results. Filters are 175 to 200-tap FIR designs. If wider transition regions are allowed, lower-order filters can be used at a cost of an increased reverberant quality of the coder. A passband ripple of ±0.5 dB gives satisfactory results in simulations.

Signal frequencies outside of the sub-band are aliased into the sub-band by the decimation process in the transmitter. This aliasing is illustrated by the dotted line in Fig. 7. With a filter stop-band attenuation on the order of 45 dB, this aliasing is not detectable. Near the sub-band edges, a slightly larger amount of aliasing can be allowed, as shown in Fig. 7, in order to keep the filter passbands as wide as possible. Filter attenuations of 12 dB at sub-band edges were used in simulations. Since two such filters are cascaded in the sub-band coder (see Fig. 1), this aliasing is reduced by 24 dB at sub-band edges. It occurs only over a very narrow frequency range (a few Hz) and is not detectable. If lower filter
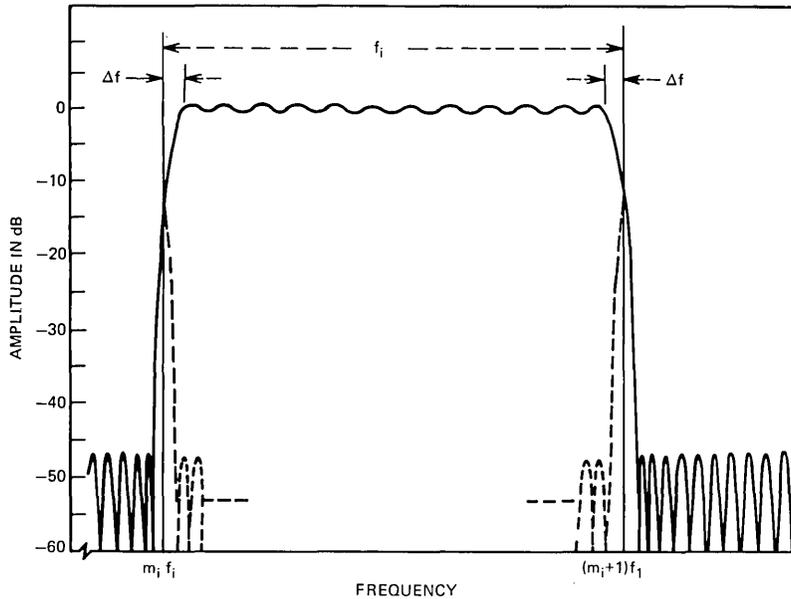
Fig. 7—Parameters of the bandpass filters.

orders (i.e., wider transition bands) are used, correspondingly larger attenuations should be used at band edges to compensate for the smaller slope of the filter roll-off in the transition regions.

The overall frequency response of the sub-band coders was measured by computer simulations. Figure 8a shows results for a 175-tap FIR filter implementation of sub-bands in Table III. Similar results are observed for IIR elliptic filters of order 6, 6, 8, 8 for bands 1 to 4, respectively. Phase distortions introduced by the IIR filters are not perceptible. In fact, the "smearing" of the phase helped to reduce the peak factor of the speech waveforms and led to a slightly improved performance (0.5 dB) in the adaptive coders. Figure 8b shows results of a 200-tap FIR filter implementation of the five-band coder in Table V.

In the receiver, the interpolating filters must have additional passband gain in order to restore the signal energy lost by decimation. The gains are equal to the decimation ratios. For example, if the sampling rate in the transmitter is decimated by 20, the interpolating filter must have a gain of 20 to account for signal energy lost in samples discarded in the decimation process.

Several hardware technologies are amenable to the implementation of sub-band coders. An attractive emerging technology, already mentioned in Ref. 1, is the charge-coupled-device (CCD) technology.[7] It offers possibilities for one or more filters on a chip with analog-to-discrete-time conversion accomplished essentially automatically. Filter outputs can

Table VIII — Comparison of sub-band coders vs ADPCM and ADM
(1-bit ADPCM) coding.

Coder Comparison	Preference for Sub-band Coder (%)	Preference for ADPCM (%)
A. 16 kb/s sub-band coder (Example D)		
(1) 24 kb/s ADPCM (3 bit)	58	42
(2) 32 kb/s ADPCM (4 bit)	34	66
B. 9.6 kb/s sub-band coder (Example B)		
(1) 10.2 kb/s ADM	96	4
(2) 12.9 kb/s ADM	82	18
(3) 17.2 kb/s ADM	61	39
C. 7.2 kb/s sub-band coder (Example C)		
(1) 12.9 kb/s ADM	79	21
(2) 17.2 kb/s ADM	56	44

be offered in a convenient sample-and-hold format. The technology may also be tractable for the implementation of the coders.

All-digital technologies also offer many attractive possibilities for the sharing of hardware between sub-bands. Efficient computational methods are possible for implementing filters for decimating and interpolating digital signals.[8] Since digital or CCD filter cutoff frequencies are normalized to the filter-sampling frequencies, the bit rates of the coders can be varied over a limited range by simply varying the master clock frequency—a feat that cannot easily be accomplished with continuous-time filter technologies.

VI. SUBJECTIVE COMPARISONS WITH OTHER WAVEFORM CODING METHODS

Further subjective comparisons have been made at 16 kb/s and 7.2 kb/s in addition to comparisons reported in Ref. 1. Thirteen listeners were asked to compare pairs of sentences for quality and indicate which was better. Two speakers were used in the experiment and several comparisons of the same sentence pairs were made by each listener at different randomly selected times during the test. The results are summarized in Table VIII.

In part A of Table VIII, the quality of the 16-kb/s sub-band coder (Example D) is compared against the quality of 24- and 32-kb/s ADPCM. It was preferred in 58 percent of the sentence pair comparisons against 24-kb/s ADPCM and in 34 percent of the comparisons against 32-kb/s ADPCM. If the results are linearly extrapolated, the quality of the 16-kb/s sub-band coder can be said to be comparable to approximately 26.5-kb/s ADPCM. This is a significant improvement over earlier results reported in Ref. 1. It was obtained by allowing less overlap of the sub-bands and trading the extra bandwidth for more bits/sample in the lower sub-bands.
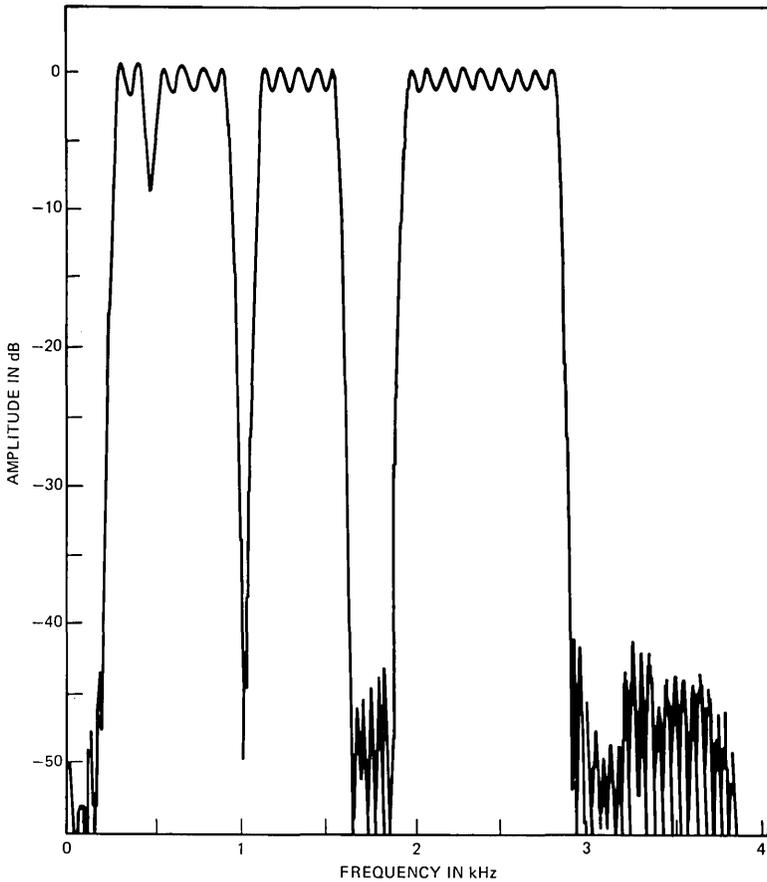
Fig. 8(a)—Measured frequency responses for 7.2- and 9.6-kb/s coders.

The 9.6-kb/s sub-band coder (Example B) is the same coder that was used for comparisons in Ref. 1. It is comparable to 19-kb/s ADM in quality. A slight improvement on this quality was observed from the sub-band coder in Example A.

In part C of Table VIII, the 7.2-kb/s sub-band coder (Example C) is compared against 12.9- and 17.2-kb/s ADM. The quality is preferred over that of 17.2-kb/s ADM and, if the results are linearly extrapolated, it is found to be comparable to approximately 18-kb/s ADM.

As seen by the above comparisons, a consistent advantage of about 10 kb/s in transmission rate is obtained by the sub-band coder over ADPCM or ADM for the same quality. Alternatively, at the same bit rate an improved quality is possible with the sub-band coder.
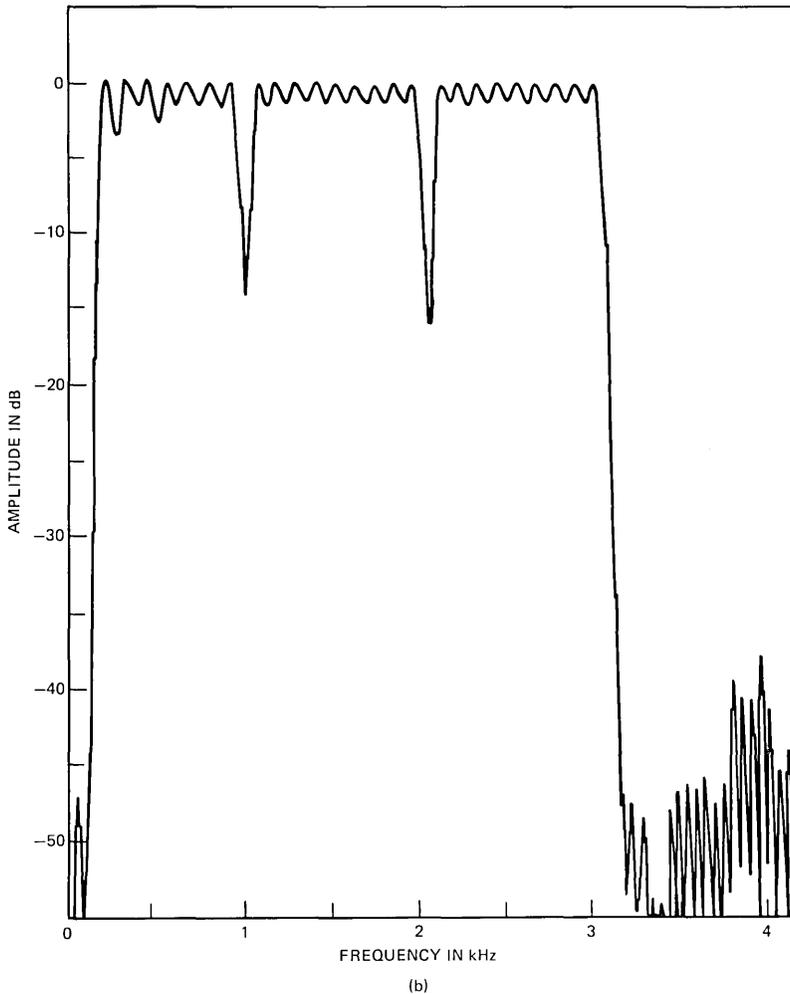
Fig. 8(b)—Measured frequency response for 16-kb/s coder.

VII. CONCLUSIONS

The design of sub-band coders involves the consideration of a large number of parameters and "trade-offs." For many of these parameters, no analytical means exist for choosing them in an optimal way. Consequently, in this paper we have attempted to provide some useful guidelines and insight for selecting parameters of sub-band coders. The guidelines are based on extensive computer simulations and subjective comparisons.

A number of practical considerations involved in selecting sub-bands, multiplexing sub-band data, and implementing the filters have also been

discussed. Several sub-band coder designs have been proposed for bit rates of 7.2, 9.6, and 16 kb/s, and their performances have been compared with those of other waveform coding techniques.

VIII ACKNOWLEDGMENTS

REFERENCES

1. R. E. Crochiere, S. A. Webber, and J. L. Flanagan, "Digital Coding of Speech in Sub-bands," B.S.T.J., *55* (October 1976), pp. 1069–1085. See also Proc. 1976 IEEE Int. Conf. on Acoust. Speech, and Sig. Proc., pp. 233–236.
2. L. L. Beranek, "The Design of Communication Systems," Proc. IRE, *35* (September 1947), pp. 880–890.
3. H. K. Dunn and S. D. White, "Statistical Measurements on Conversational Speech," J. Acoust. Soc. Amer., *11* (1940), 278–288.
4. N. S. Jayant, "Digital Coding of Speech Waveforms: PCM, DPCM, and DM Quantizers," Proc. IEEE, *62* (May 1974), 611–632.
5. P. Cummiskey, N. S. Jayant, and J. L. Flanagan, "Adaptive Quantization in Differential PCM Coding of Speech," B.S.T.J., *52* (September 1973), 1105–1118.
6. D. J. Goodman, private communication.
7. C. H. Sequin and M. F. Tompsett, *Charge Transfer Devices,* suppl. 8 for *Advances in Electronics and Electron Physics,* New York: Academic, 1975.
8. R. E. Crochiere and L. R. Rabiner, "Optimum FIR Digital Filter Implementations for Decimation, Interpolation, and Narrowband Filtering," IEEE Trans. Acoustics, Speech, Signal Process., *ASSP-23,* No. 5 (October 1975), pp. 444–456.

A Variable-Band Coding Scheme for Speech Encoding at 4.8 kb/s

By R. E. CROCHIERE and M. R. SAMBUR

The standard fixed sub-band coding scheme has been modified to allow the center frequency of the two upper bands to vary in accordance with the dynamic movement of the vocal tract resonances F2 and F3. A relatively simple zero-crossing technique is used to measure the formants F2 and F3. Through the use of this variable band coder, it is possible to produce moderate-quality, intelligible speech at 4.8 kb/s (quality is slightly less than that of a 7.2-kb/s fixed sub-band coder and equal to that of about a 16-kb/s ADM coder). The reasonably good intelligibility of the 4.8-kb/s variable-band coded speech can be attributed to the coder's attempt to capture and encode those spectral components of the signal that are perceptually most significant (the region around the formants). The major advantage of the variable-band scheme is that its implementation is considerably less complex than other waveform coding schemes or vocoder systems that can produce intelligible, narrowband speech.

I. INTRODUCTION

Recently, a method for digitally coding speech signals in terms of sub-bands of the total spectrum was introduced that resulted in an improvement in quality of the coded signal over that obtained from a single full-band coding of the total spectrum.[1,2] The rationale for coding the signal in sub-bands is based upon the experimental fact that quantizing distortion is not equally detectable at all frequencies, and hence, the quality of the coded signal can be significantly improved by controlling the distribution of quantizing noise across the signal spectrum. Coding the signal in sub-bands offers the possibility of achieving this control.

In the recent work by Crochiere, Webber, and Flanagan,[1,2] the selection of the appropriate sub-bands was guided by the perceptual data contained in the so-called articulation index (AI).[3] The articulation index denotes, on the average, the contribution of each part of the spectrum

to the overall perception of the spoken sound. For high-quality speech at moderate bit rates (16 kb/s and greater), the frequency range 200 to 3200 Hz was partitioned into four fixed contiguous bands that contributed equally to the AI. The transmission bit rate of the sub-band coder could be lowered gracefully by limiting these sub-bands in width and by tolerating some spectral gaps that did not contribute significantly to the AI. However, carried to excess, the noncontiguous bands produced an unpleasant, reverberant quality in the signal that finally resulted in unacceptable speech. A bit rate of approximately 7.2 kb/s was found to be about the lowest bit rate that still produced acceptable, intelligible speech. (The quality at this bit rate was judged about equal to that of 18-kb/s ADM speech.[2])

In using the AI in selecting sub-bands, it should be noted that this index only indicates the *average* perceptual contribution of each part of the spectrum. Since the speech spectrum is highly variable across a given utterance, it seems appropriate to select sub-bands that do not remain fixed but vary in accordance with the changing character of the speech. One way to achieve this goal is to allow the center frequency of the sub-bands to follow the variation of the formant frequencies across the utterance. The formant frequencies of a particular sound correspond to the resonance frequencies of its short time spectrum, and the frequency bands around these formants are perceptually the most significant regions of the spectrum. It is the purpose of this paper to show that by varying the sub-bands in accordance with the formant frequencies, it is possible to lower the bit rate to 4.8 kb/s and still maintain a speech quality that is approximately comparable to that of the 7.2-kb/s fixed sub-band coder. In addition, it is also shown that the formant frequencies can be sufficiently estimated for use in the variable-band scheme by a simple zero-crossing measurement technique. Thus, the variable-band coder can achieve very low data rates (4.8 kb/s) at considerably less expense than conventional vocoder systems, while still providing an intelligible signal.

II. VARIABLE-BAND CODER

The concept of the variable-band coder is illustrated in Fig. 1. The speech band is divided into four sub-bands and encoded separately in each sub-band. The two lower sub-bands are fixed bands that cover the frequency range from approximately 250 to 820 Hz. This represents the region of primary speech energy for voiced sounds. The two upper sub-bands are variable bands (with fixed bandwidths) centered about the F2 and F3 resonance peaks of the short-time speech spectrum (as illustrated by the dotted line). By varying the center frequencies of these two bands as the short-time spectrum changes, the encoder attempts to capture the maximum amount of speech energy and represent those
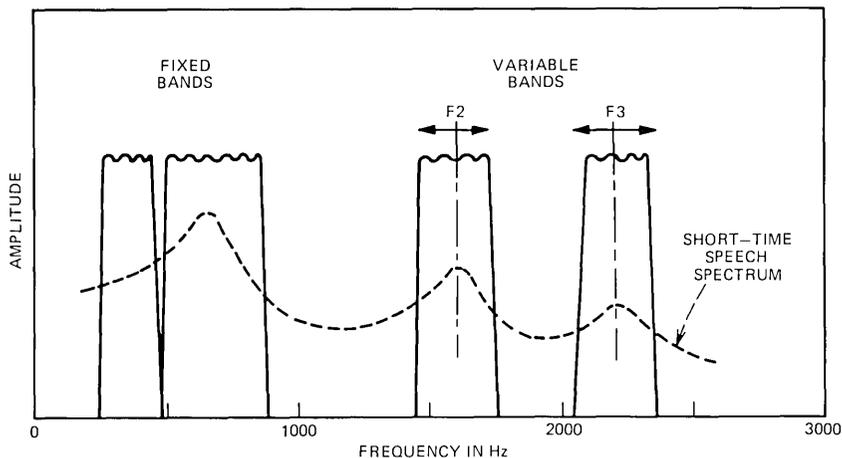
Fig. 1—Frequency domain interpretation of the variable-band coder.

frequencies that are perceptually most significant. Regions between the sub-bands are ignored to conserve bandwidth. While these gaps give a reverberant quality to the coder, the effect, as will be discussed, is not as pronounced as with a fixed-band scheme at the same bit rate.

The implementation of the sub-band coder can be achieved by any of the modulation schemes suggested in Ref. 1. In particular, the most efficient approach for implementing the fixed bands is the integer band sampling method. For the two upper sub-bands, a modulation scheme is required in which the center frequency of the band can be varied. This can be accomplished with the complex modulation method discussed in Ref. 1. In addition, a method for adaptively varying the center frequencies of these bands is required.

The overall configuration of the sub-band coder then takes the form shown in Fig. 2. The formant estimator determines the resonances F2 and F3 in the speech band. This information is encoded at a low bit rate and sent to the receiver. It is also decoded and used to control the variable-band center frequencies in the transmitter. In this way, the variable bands in the transmitter and receiver track identically.

The measurement of the resonances F2 and F3 is accomplished by a simple zero-crossing measurement technique. In this method, the individual resonances are first isolated by filtering the speech signal into frequency ranges appropriate to each formant.[4] After filtering, the resulting signal is ideally a damped sinusoid, and the formant frequency can then be estimated by measuring the axis-crossing rate of the filtered waveform. Figure 3 depicts the structure of the formant frequency extraction system.* To correct for isolated errors in the formant extraction

---

* The formants are measured 50 times per second and can be efficiently coded using less than 300 b/s by ADPCM techniques.[5]
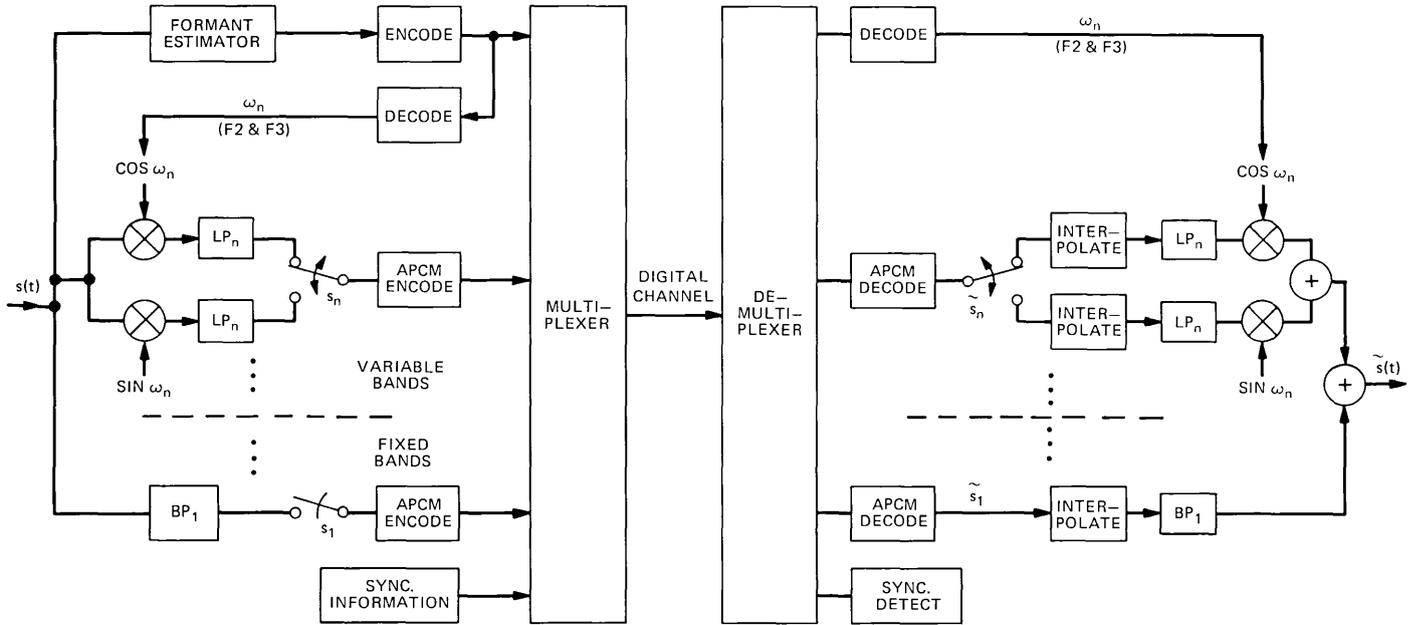
Fig. 2—Implementation of variable-band coder with integer-band sampling for the fixed bands and complex modulation for the variable bands.
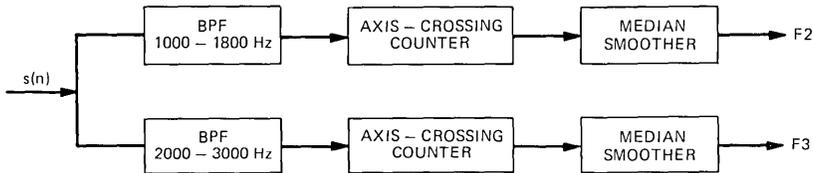
Fig. 3—Formant estimation scheme for estimating F2 and F3.

scheme and to insure that the measured formant trajectory is not excessively rough, a median smoother is employed.[6] It should be noted that although this system of formant measurement is not as accurate as the more elaborate methods of linear prediction[7] or spectral estimation,[4] the zero-crossing technique is computationally far less expensive than these schemes and, moreover, there is no enhancement in quality in the variable-band encoded signal when the more sophisticated measurements of F2 and F3 are used to control the variable sub-bands.

The sub-band signals are encoded with APCM encoders and the data are multiplexed together with the synchronization data and formant data, as illustrated in Fig. 2. Typically, more bits/sample are used for encoding lower sub-bands for the perceptual reasons explained in Refs. 1, and 2. Alternatively, a dynamic allocation of bits/sample can be employed in a manner similar to that used by Noll for transform coding.[7] Also, a slight amount of center-clipping can be used in sub-bands to reduce idle channel noise.

III. RESULTS OF COMPUTER SIMULATIONS

The sub-band coder system in Fig. 2 has been implemented by computer simulation for a transmission bit rate of 4.8 kb/s. Sub-band center frequencies and bandwidths corresponding to those in Table I were used. These bands also correspond to those shown in Fig. 1.

The formants were estimated by the method in Fig. 3 and were used to control the center frequencies of bands 3 and 4. Figure 4 shows the variation of these center frequencies as a function of time for the sentence

Table I—4.8-kb/s variable-band coder

| Band	Center Frequency (Hz)	Bandwidth (Hz)	Fixed	Dynamic Voiced	Unvoiced/ Silence
				Bits/Sample Allocation	
1	356	213	3	3	$1\frac{1}{2}$
2	640	356	2	2	2
3	F2	320	$1\frac{1}{2}$	$1\frac{1}{2}$	2
4	F3	320	$1\frac{1}{4}$	$1\frac{1}{4}$	2
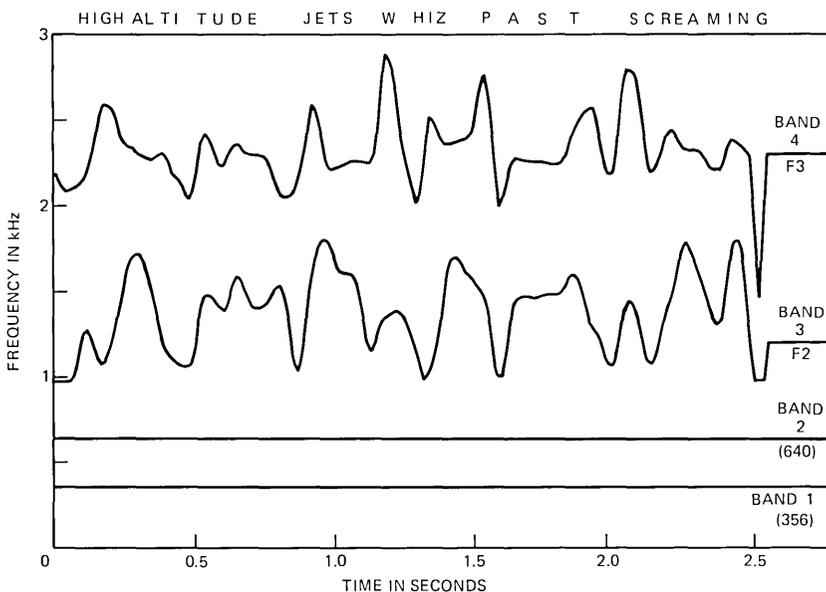
Fig. 4—Center frequencies of bands as a function of time for the sentence "High altitude jets whiz past screaming."

"High altitude jets whiz past screaming." A comparison of this plot to the spectrogram in Fig. 5a shows that these center frequencies do in fact track the major F2 and F3 resonances in the speech signal.

Two different bits/sample allocation schemes were tried in the simulations, a fixed allocation and a simple dynamic allocation scheme. In the fixed allocation 3, 2, 1½, and 1¼ bits/sample were used for encoding sub-bands 1 to 4, respectively. In the dynamic allocation scheme, 3, 2, 1½, and 1¼ bits/sample were used for encoding the voiced regions of the speech signal for bands 1 to 4, respectively. For unvoiced/silence regions, an allocation of 1½, 2, 2, and 2 bits/sample for bands 1 to 4 was used to encode more accurately the stronger energy in the higher frequencies during these intervals. A simple voiced/unvoiced decision was made by observing the variable step size of the APCM coder in the lowest band. If this step size was greater than five times its minimum allowed size, then the speech was assumed to be voiced and the 3, 2, 1½, and 1¼ bits/sample allocation was used. If it was less than five times the minimum step size, then the unvoiced/silence condition was assumed and the bits/sample allocation of 1½, 2, 2, and 2 was used.

Figure 5 shows spectrograms of the resulting computer simulations. The original sentence is represented by the upper spectrogram of Fig. 5a. Figure 5b corresponds to a sentence that was sub-band filtered (without encoding) with a fixed-band scheme (the two upper bands had center frequencies of 1200 Hz and 2300 Hz). In contrast, Fig. 5c shows
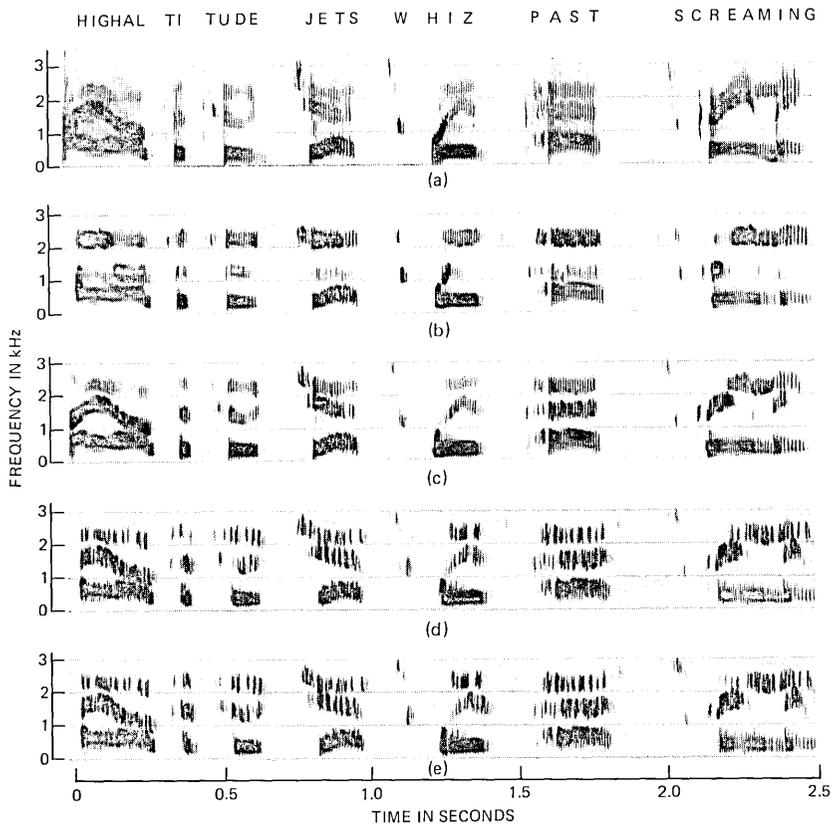
Fig. 5—Spectrograms of computer simulations. (a) Original. (b) Fixed-band filtered speech (uncoded). (c) Variable-band filtered speech (uncoded). (d) Variable-band coded with fixed bits/sample allocation. (e) Variable-band coded with dynamic bits/sample allocation.

the same sub-band arrangement except that the center frequencies of the two upper bands were allowed to vary according to Fig. 4. Again, the sub-bands were not encoded but simply filtered. A comparison of these two spectrograms (Figs. 5b and 5c) shows that the variable-band scheme gives a better representation of the important spectral features of the speech signal than the fixed-band scheme for the same total bandwidths. For example, in the words "high" and "whiz," the F2 resonance is lost in the gap between bands 3 and 4 of the fixed-band scheme; however, it is clearly present in the variable-band scheme.

Figures 5d and 5e show the results of the variable-band coder with the fixed and dynamic bit allocations discussed earlier. By comparison of these sentences with the unquantized sentence of Fig. 5c, the effects of the quantization can be observed. Typically, the quantized sentences have spectrograms that are more "ragged" in appearance due to the

presence of the quantization errors and noise. Little difference was observed between the quality of the fixed and dynamic bit-allocation methods of quantization nor can it be observed on the spectrograms. Only a slight improvement in quality, during unvoiced regions, is gained by the dynamic bit allocation. More substantial improvements might be possible through more sophisticated allocation schemes than the one tried here (see Noll[6]).

The quality of the 4.8-kb/s variable-band coder was observed to be only slightly less than that of the 7.2-kb/s fixed-band coder reported in Ref. 2 (the 7.2 kbps coder was rated equal to that of an 18-kbps ADM coder). The movement of the two upper bands produced a noticeable "swishy" noise in the background. This was more readily observed with earphones than with a loudspeaker. Also, the quantization noise of the coders gave a slightly hoarse sound to the speech. On the other hand, although the quality of the variable-band coder is only moderate, the intelligibility of the coded speech is still reasonably good because it attempts to capture and encode those spectral components of the speech signal that are perceptually most significant.

IV. CONCLUSIONS

The standard fixed sub-band coding scheme has been modified to allow the center frequencies of the two upper bands to vary in accordance with the dynamic movement of F2 and F3. The formants F2 and F3 are measured by a relatively simple zero-crossing technique. Using this variable-band system, it is possible to produce moderate-quality, intelligible speech at 4.8 kb/s.

The variable-band system can be viewed as a hybrid type coder that combines the simplicity of a sub-band coder with the low-bit-rate potential of a vocoder type system. The ability of the variable-band coder to achieve narrowband transmission is directly associated with its vocoder-like utilization of the perceptually significant regions around the formants F2 and F3. But, unlike the vocoder, it is a true waveform coder that does not attempt merely to *model* the signal in terms of such features as pitch and vocal tract resonances.[4] It directly codes the entire 250-Hz to 818-Hz region of the spectrum and two 320-Hz bands centered about the crudely estimated values of F2 and F3. The variable-band coder can thus avoid the computationally expensive analysis-synthesis systems required of a vocoder, and can produce moderate-quality, intelligible speech in a relatively inexpensive manner.

REFERENCES

1. R. E. Crochiere, S. A. Webber, and J. L. Flanagan, "Digital Coding of Speech in Sub-bands," B.S.T.J., *55*, No. 8 (October 1975), pp. 1069–1085. See also Proc. of the 1976 IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, pp. 233–236.

2. R. E. Crochiere, "On the Design of Sub-band Coders for Low Bit Rate Speech Communication," B.S.T.J., this issue, pp. 747–770.
3. L. L. Beranek, "The Design of Communication Systems," Proc. IRE, *35* (Sept. 1947), pp. 880–890.
4. J. L. Flanagan, *Speech Analysis Synthesis and Perception,* New York: Springer-Verlag, 1972.
5. N. S. Jayant, "Delta Modulation of Pitch, Formant and Amplitude Signals for the Synthesis of Voiced Speech," IEEE Trans. Audio Electroacoust., *AU-21,* No. 3 (June 1973), pp. 135–140.
6. L. R. Rabiner, M. R. Sambur, and C. E. Schmidt, "Applications of a Nonlinear Smoothing Algorithm to Speech Processing," IEEE Trans. Acoust. Speech Sig. Process., *ASSP-23,* No. 6 (December 1975), pp 552–557.
7. B. S. Atal and S. L. Hanaver, "Speech Analysis and Synthesis by Linear Prediction of the Speech Wave," J. Acoust. Soc. Amer., *50,* (1971) pp. 637–655
8. R. Zelinski and P. Noll, "Adaptive Transform Coding of Speech Signals," IEEE Trans. Acoust. Speech Sig. Process., *25,* No. 4 (August 1977).

On Speech Encryption Using Waveform Scrambling

By S. C. KAK and N. S. JAYANT

This paper presents a study of speech scramblers based on permutations of samples within an N-block. It has been found that a new family of "uniform" (U) permutations (defined by the address mapping $i \rightarrow ki$ modulo N; k prime to N) is as effective as pseudorandom (PR) permutations in destroying speech intelligibility. Analytical results show the relation between input and scrambled-signal spectra, while computer simulations compare the effects of scrambling on PAM samples and on codes based on ADM, APCM, and ADPCM. Scrambling is increasingly effective in that order, and encoding delays in ADPCM can be as low as 1 to 2 ms. Finally, scrambling has been compared with frequency inversion, which corresponds to sign-inversion in every other Nyquist-rate waveform sample.

I. INTRODUCTION AND SUMMARY OF RESULTS

Waveform scrambling permits a conceptually simple method of speech encryption. In view of a certain ambiguity of definition in the subject, we define scrambling as a reversible temporal rearrangement of waveform samples within a waveform block. Though specific scrambling methods have been well known and documented,[1,2] no general study seems to have been made to relate a temporal permutation to the associated modification of the short-term frequency spectrum. Furthermore, for a scrambling block length N, the total number of permutations is $N!$; while certain pseudorandom (PR) permutations constitute an effective subclass of these $N!$ permutations, other interesting subclasses exist, and none of these has been discussed in the literature to our knowledge. Finally, while it is known that effective speech encryption can result from scrambling several types of speech codes (from waveform coders or from linear predictive vocoders[2]), alternative speech codes have not been compared from the point of view of encryption potential. Such a comparison would be particularly useful if the comparisons included alter-

native methods for encoding speech at comparable bit rates—for example, speech waveform coding[3] using adaptive pulse code modulation (APCM), adaptive DPCM (ADPCM), and adaptive delta modulation (ADM).

This paper discusses (i) the effect of a permutation on the input spectrum shape; (ii) the generation of a new subclass of permutations characterized by a simple algorithm and by desirable spectrum modifications; (iii) comparison of the new permutation subfamily with shift-register-generated PR permutations[2] and with the classical technique of frequency inversion;[1] and (iv) comparison of scrambler performance on ADM, APCM, and ADPCM bits and on pulse-amplitude modulated (PAM) speech samples for block lengths varying from 4 to 64.

For ADM bits the sampling rate was 24 kHz. The APCM and ADPCM coders used 8-kHz samples and 3 bits per sample. Thus, the bit rate of all three coders was 24 kb/s. The PAM samples were sampled at 24 kHz for scrambling and at 8 kHz for frequency inversion; in the time domain, frequency inversion corresponds to reversing the polarity of every other Nyquist-rate PAM sample. The Nyquist frequency for the 200- to 3200-Hz speech sample used in our perceptual experiment was 8 kHz. With block-scrambling of 24-kHz samples or bits, a block length of 64 samples or bits implies a time duration of $64/24 = 2.7$ ms, while a block length of four samples represents a duration of 0.17 ms.

Our study is addressed to issues of casual privacy as well as formal encryption. For privacy, the objective is to render speech as unintelligible as possible by means of one out of many possible transformations, so that only a receiver with knowledge of an inverse transformation can understand the message. There is very little change of the transformation function with time. Recorded messages in a privacy system are amenable to cryptanalysis after sufficient processing. This kind of privacy also obtains in transformations on text[1] (where, however, the secrecy requirement is potentially greater since there are no restrictions such as those in a real-time communication link).

In formal encryption, secrecy is maintained by a repeated change of the transformation procedure (referred to as the key), and the number of keys available becomes a measure of the effectiveness of the system. As a general mathematical study of encryption is already available,[4] we shall devote more attention to the specific issue of speech privacy— specifically, the problem of destroying speech intelligibility using transformations such as permutations. A formal speech encrypter would employ the strategy of switching from one permutation transformation to another in a fashion known only to the intended receiver. The technique of frequency inversion is not useful for formal encryption because there is only one key associated with it. On the other hand, we can use the technique of adding a masking signal,[1] such as modulo 2 addition

of a pseudorandom binary sequence,[2] for formal speech encryption in real time. We have compared scrambling with this technique and have found that for speech applications where some processing delay is acceptable, temporal scrambling may be preferable. This is because the key information in scrambling is very small, and it takes very little information to transmit changes of keys.

The rest of this section summarizes the results of this paper. The novel subclass of permutations proposed in this paper is defined by the simple mapping of sample-position r into position s, with

$$s = k_1 \times r \text{ (modulo } N\text{)}$$
$$r = 1, 2, \cdots N, \tag{1a}$$

where k_1 is prime to N and N is the number of samples in the waveform block that is being scrambled.

We refer to the one-to-one mapping above as uniform, or U, permutations (see Section II). The set of U permutations increases faster with block length N than the corresponding set of PR permutations; for $N = 8$, it is larger by a factor of about 2, while for $N = 128$, this factor is nearly 7.

The analytical results presented in the paper include an algorithmic relation between a permutation and the modification it produces in the DFT spectrum. Thus, if permutations of time samples are characterized by a matrix P (see Section II) and if F is the standard DFT matrix [given by (5)], the transformation T of DFT samples is expressed by

$$T = FPF^{-1}. \tag{2}$$

For example, it has been shown that U permutations produce a uniform scrambling in the DFT spectrum as well. Descrambling is performed by another U permutation:

$$r = k_2 s \text{ (modulo } N\text{)}$$
$$\text{if } k_1 k_2 = 1 \text{ (modulo } N\text{)}. \tag{1b}$$

Thus, for $N = 32$, if scrambling is done with $k_1 = 7$, descrambling will need $k_2 = 23$.

Perceptual comparisons have been made of U permutations, PR permutations, and frequency inversions for PAM, ADM, APCM, and ADPCM samples. Results indicate the following increases in efficiency:

(i) PAM < ADM < APCM < ADPCM for scrambling with a given block length N.

(ii) ADM < PAM < APCM < ADPCM for frequency inversions.

The nature of the encrypted waveform is different between the first two and the last two codes, due conceivably to the fact that in PCM and DPCM, different bits have different weights associated with them. For

scrambling, U permutations performed at least as well as PR permutations, and, for one case of $N = 32$, in fact did better in informal perceptual comparisons. For longer block lengths, no meaningful comparisons could be made between the two scramblers, because intelligibility was totally destroyed in each case (except with PAM samples). We have also found that if the scrambling delay is sought to be less than about 2 ms, it becomes essential to use APCM or ADPCM codes.

This paper is arranged as follows: Section II presents analytical results on permutation transformations and compares U and PR permutations. Section III briefly discusses the classical technique of frequency inversion. Section IV summarizes perceptual results and spectrograms from a computer simulation. Section V provides a heuristic comparison of masking and scrambling as alternative techniques for encryption.

II. PERMUTATION TRANSFORMATIONS

2.1 Permutation matrices

The total number of permutations possible on a block of N samples is $N!$. If we also included the possibility of sign changes of the samples, this total is increased to $2^N N!$. As sign changes are easy to implement, a general study of such transpositions should be useful. In this paper, however, we restrict ourselves to simple permutations only, except for the special case of sign changes of alternate samples (Section III). This corresponds to classical frequency inversion, and is therefore of interest for comparison with permutations (scrambling).

Temporal sample permutations can be characterized by a permutation operator P that is a matrix of ones and zeros. Thus, for a block of $N = 5$ samples, if bit positions 0, 1, 2, 3, and 4 are permuted to 0, 1, 4, 2, and 3, we can denote the permutation by

$$[0, 1, 2 \to 4 \to 3 \to 2]$$

or by the permutation matrix

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Expressing P as $[2 \overset{P}{\to} 4 \overset{P^2}{\to} 3 \overset{P^3}{\to} 2]$ gives the order of the cyclic group defined by P directly; since P^3 brings 2 back to 2, the order is 3. In other words, P^3 is an identity matrix I.

If a permutation is defined in terms of more than one cycle, it can be seen that it has an order equal to the least common multiple (lcm) of (order of cycle 1, order of cycle 2, $\cdots$). For example, the permutation [0

$\rightarrow 3 \rightarrow 5 \rightarrow 0, 2 \rightarrow 4 \rightarrow 6 \rightarrow 1 \rightarrow 2$] has an order = lcm (3, 4) = 12, and this permutation is characterized as a $P(3, 4)$ permutation. The number of repeated $P(3, 4)$ permutation operations that will bring us back to an original unpermuted sequence is 12. (The permutation in the earlier example was of the $P(1, 1, 3)$ type; in that case, three successive $P(1, 1, 3)$ operations would lead to the unpermuted original sequence.) In general, if

$$P: P(p_1, p_2, p_3, \cdots),$$

where

$$p_1 + p_2 + p_3 + \cdots = N, \tag{3a}$$

the number of repeated P permutations that would result in the original unpermuted sequence [equivalently, the number of distinct mappings $L(P)$ that can be generated using P] is

$$L(P) = \text{lcm} \ (p_1, p_2, p_3 \cdots). \tag{3b}$$

2.2 Effect of permutations on frequency spectrum

Let $x(n)$, $(n = 0, 1, \cdots, N - 1)$ be the discrete input block of length N, which is to be permuted. Let $X(m)$, $(m = 0, 1, \cdots, N - 1)$ be its DFT, and let x and X denote corresponding column vectors $[x(0), x(1), \cdots x(N - 1)]'$ and $[X(0), X(1), \cdots X(N - 1)]'$, respectively. Note that

$$X = Fx, \tag{4}$$

where

$$F = \begin{bmatrix} 1 & 1 & 1 & 1 & . & . & 1 \\ 1 & W & W^2 & W^3 & . & . & W^{N-1} \\ 1 & W^2 & W^4 & W^6 & . & . & W^{2(N-1)} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 1 & W^{N-1} & W^{2(N-1)} & . & . & . & W^{(N-1)^2} \end{bmatrix} \tag{5}$$

with the standard DFT identities

$$W = \exp\left(-j\frac{2\pi}{N}\right); \quad W^N = 1$$

and $\hfill (6)$

$$\sum_{r=0}^{N-1} W^r = 0.$$

Let X_P, the DFT of the permuted sequence, be described as a transformation of X:

$$X_P = TX = TFx. \tag{7a}$$

Also, in the manner of (4),

$$X_P = FPx. \tag{7b}$$

From (7a) and (7b),

$$FP = TF, \text{ or}$$

$$T = FPF^{-1} \tag{8a}$$

and

$$P = F^{-1}TF, \tag{8b}$$

where the inverse of F can be seen to be

$$F^{-1} = \frac{1}{N}
\begin{bmatrix}
1 & 1 & 1 & . & . & 1 \\
1 & W^{-1} & W^{-2} & . & . & W^{-(N-1)} \\
1 & W^{-2} & W^{-4} & . & . & W^{-2(N-1)} \\
. & . & & . & . & . \\
. & . & & . & . & . \\
. & . & & . & . & . \\
1 & W^{-(N-1)} & . & & . & W^{-(N-1)^2}
\end{bmatrix}. \tag{9}$$

Theorem: If P is represented in terms of a, b, c, d in the earlier characterization $[0 \rightarrow a, 1 \rightarrow b, 2 \rightarrow c, 3 \rightarrow d, \cdots]$,

$$T_{rs} = \frac{1}{N} [W^{-as} + W^{r-bs} + W^{2r-cs} + W^{3r-ds} + \cdots], \tag{10}$$

where, as usual, further simplifications can be made using the relations in (6).
For a proof of this theorem, see the Appendix.

Example 1: Let P_1: $[0 \rightarrow 2 \rightarrow 1 \rightarrow 3 \rightarrow 0]$. That is, $N = 4$, $W = -j$, $W^2 = -1$, $W^3 = j$, and $W^4 = 1$; and $a = 2$, $b = 3$, $c = 1$, and $d = 0$.

$$T_1 =
\begin{bmatrix}
-1/2 & (1+j)/2 & j/2 & 0 \\
(j-1)/2 & 0 & (-j-1)/2 & 0 \\
-j/2 & (1-j)/2 & -1/2 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}.$$

$$T_1^2 =
\begin{bmatrix}
0 & 0 & -j & 0 \\
0 & -1 & 0 & 0 \\
j & 0 & 0 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix};
\quad
T_1^4 =
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}.$$

It is seen that T_1 produces a transformation where DFT values are added together with different weights and phase shifts; while T_1^2 simply changes the positions of varying DFT values with varying phase shifts without modifying magnitudes. The fact that $T_1^4 = I$ could also be deduced by noting that in the time domain, $P^4 = I$ (because the order of P is 4).

Example 2: Let

$$P_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & . & . & . & 0 & 0 \\ 0 & 0 & 0 & 0 & . & . & . & 0 & 1 \\ 0 & 0 & 0 & 0 & . & . & . & 1 & 0 \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ 0 & 0 & 1 & 0 & . & . & . & 0 & 0 \\ 0 & 1 & 0 & 0 & . & . & . & 0 & 0 \end{bmatrix} . \tag{11a}$$

P_2 corresponds to $[0 \rightarrow 0, 1 \rightarrow N - 1, 2 \rightarrow N - 2 \cdots]$. Hence, from (10),

$$N \cdot T_{rs} = 1 + W^{r-s(N-1)} + W^{2r-s(N-2)} + \cdots$$

$$= 1 + W^{r+s} + W^{2(r+s)} + \cdots = (1 - W^{(r+s)N})/(1 - W^{r+s})$$

using (6); since the numerator in the above sum is always zero, T_{rs} is nonzero if and only if the denominator is zero, or if $s = MN - r; M = 0, 1, 2 \cdots$; in this case, $NT_{rs} = N$, or $T_{rs} = 1$. Consequently,

$$T_2 = P_2. \tag{11b}$$

In words, local (intra-block) time-inversion (excluding the first sample) causes a corresponding DFT inversion (excluding the first sample). Notice that an unchanged first sample in the DFT vector represents an unchanged zero-frequency value of the input spectrum. (We emphasize here that DFT inversion does not represent analog frequency inversion, which will be discussed in Section III. The DFT spectrum is symmetrical about its middle point; inverting it is a perceptually trivial operation (see Section IV), and DFT inversion does not represent a useful means for speech encryption.

2.3 Pseudorandom permutations

A pseudorandom (PR) scrambling of samples within a block of length N is achieved by assigning to each sample a new address $A (A = 1,$ or 2, or 3, $\cdots$, or N) determined by the state of a maximal-length shift-register arrangement. The theory and design of maximal-length sequences is well documented.[5,6,7] We shall therefore only provide a constructive recapitulation in this paper. The approach is to start with a shift register whose length is $D = \log_2 N$ (assume that the block length N is a power of 2 and

| PRIMITIVE POLYNOMIAL			NO.	STATE OF SHIFT REGISTER	DECIMAL EQUIVALENT
$y^5 + y^2 + 1$			1	00001	1
			2	10000	16
CORRESPONDING NETWORK			3	01000	8
			4	00100	4
			5	10010	18
			6	01001	9
			7	10100	20
			↑	-----	↑
MOD 2 ADDER			INPUT ADDRESSES	REPETITION PERIOD = 31	OUTPUT ADDRESSES
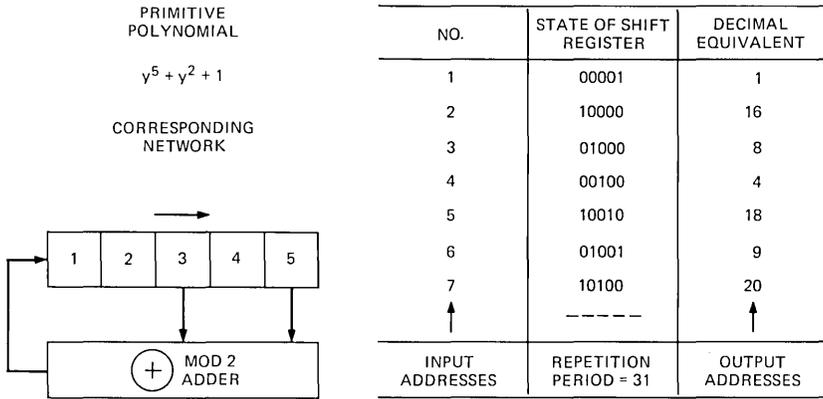
Fig. 1—Pseudorandom scrambler with a five-stage shift register.

that elements in the register are either 1 or 0). The next step is to select a primitive polynomial $Q(y)$ of degree D and to include stage $D - d$ in the register ($d = 0, 1, \cdots D - 1$) in an EXCLUSIVE OR feedback (modulo 2-add) arrangement if the coefficient of y^d in $Q(y)$ is nonzero. The resulting network now generates a succession of $2^D - 1 = N - 1$ nonzero states in the shift register at successive clock times, after which the cycle repeats, starting once again with the original initial state of the shift register. The number of nonzero states in the cycle is identically equal to the repetition period $N - 1$ of the cycle. Consequently, the $N - 1$ states of the shift-register (specifically, their decimal equivalents) can be used as PR addresses for a block of $N - 1$ input samples in a one-to-one mapping of addresses. If the input block has N rather than $N - 1$ samples (because of the frequent requirement that N be a power of 2), the address of the Nth sample is usually left unaltered by the scrambler. Such simplification is, however, not mandatory, and appropriate manipulations that scramble all N samples are quite conceivable.

Figure 1 illustrates the scrambler design for the example of $D = 5$ and $N = 31$, as defined by a primitive polynomial $Q_5(y) = y^5 + y^2 + 1$. We see how input samples ($1, 2, 3, 4, 5, 6, 7, \cdots$) get scrambled into PR positions ($1, 16, 8, 4, 18, 9, 20, \cdots$). We can verify easily that the mapping is one to one for the 31 sample values in an input block. The PR scrambling of this example is illustrated more completely in the permutation matrix in Section 2.5 (Fig. 3a).

It is clear that in the arrangement of Fig. 1, the use of a different initializing sequence (other than 00001) can lead to a totally different mapping of sample addresses. There would be $N - 1$ nonzero initializations corresponding to every given $Q_5(y)$. In addition, the number of primitive polynomials $Q_5(y)$ of degree 5 is 6, and this implies a further increase in the total number of possible mappings.

Table I — List of primitive polynomials

Degree D	Typical Primitive Polynomial $Q_D(y)$	Number of Primitive Polynomials $L(D)$ of Degree D
1	$y + 1$	1
2	$y^2 + y + 1$	1
3	$y^3 + y + 1$	2
4	$y^4 + y + 1$	2
5	$y^5 + y^2 + 1$	6
6	$y^6 + y + 1$	6
7	$y^7 + y + 1$	18
8	$y^8 + y^4 + y^3 + y^2 + 1$	16
9	$y^9 + y^4 + 1$	48
10	$y^{10} + y^3 + 1$	60
11	$y^{11} + y^2 + 1$	176
12	$y^{12} + y^6 + y^4 + y + 1$	144

Table I lists, for $D = 1$ to 12, a typical set of primitive polynomials, and also the number of primitive polynomials $L(D)$ for each D. Note, for example, that a 12-stage shift register with an EXCLUSIVE OR feedback network involving stages 12, 11, 8, and 6 ($D - d; d = 0, 1, 4,$ and 6) provides one of 144 possible bases for a scrambler that would operate on an input block of $2^{12} = 4096$ samples.

The possibility of distinct scrambler mappings (number of keys) as defined by different initializations and/or different primitive polynomials is an important consideration from the point of view of the average descrambling time needed for an eavesdropping code-breaker. Let us note formally, then, that the number of keys K for a PR permutation of N sample blocks ($N = 2^D$) is

$$K_{PR} = (N - 1)L(\log_2 N). \qquad (12)$$

The first term of the product in (12) represents the total number of shift-register initializations, and the second term gives the number of primitive polynomials (Table I) of degree $D = \log_2 N$. Numerical values of K_{PR} are discussed in Section 2.5.

2.4 Uniform permutations

We now discuss the new subclass of scramblers defined by an output/input (s-r) mapping indicated in (1a):

$$s = k_1 r \text{ (modulo } N)$$

$$r = 1, 2, \cdots N; \quad k_1 \text{ is prime to } N. \qquad (13a)$$

It can be verified that k_1 should be prime to the block-length N for the

INPUT ADDRESS	OUTPUT ADDRESS
1	7
2	14
3	21
4	28
5	3
6	10
7	17
––	––
r	7r

UNIFORM ONE–TO–ONE MAPPING

$r \longrightarrow 7r$ (MOD 32)

$r = 1, 2, \cdots 32$

Fig. 2—Uniform scrambler ($r \rightarrow 7r$ mod 32).

mapping (13a) to be one to one. Figure 2 illustrates these permutations for the example of $N = 32$ and $k_1 = 7$. The complete permutation matrix is discussed in Section 2.5 (Fig. 3b), where it may be noted that the 1s in the matrix are spread uniformly in the 32 by 32 matrix—hence, the name uniform permutations, or U permutations.

The descrambling of a U permutation is achieved by means of another (inverse) U permutation so that $k_1 k_2 = 1$ (modulo N):

$$r = k_2 s \text{ (modulo } N)$$

$$s = 1, 2, \cdots N \tag{13b}$$

and $k_1 k_2 = 1$(modulo N). The condition that $k_1 k_2 = 1$ follows from the requirement that $k_2 s = k_2(k_1 r) = r$, for all r. Thus, for $N = 32$, if scrambling is done with $k_1 = 7$, descrambling will require $k_2 = 23$ in order that $k_1 k_2 = 161$ mod 32 = 1.

2.4.1 Choice of k_1

The value of k_1 in (13a) is an interesting issue. We do not offer any rigorous criterion for optimizing k_1, and we believe in general that there is no single optimum for practical applications where, in fact, we use different k_1 values as different keys in encryption. However, there are three observations regarding k_1 that are worth noting.

First, a reasonable scrambling procedure is one that distributes N 1s uniformly in an N by N permutation. For simplicity, let N be a perfect square. The uniform distribution problem is then one of dividing the N by N matrix into a number N of $\sqrt{N}$ by $\sqrt{N}$ submatrices, and to place one and axactly one of the N 1s in the center of each submatrix. Adjacent 1s would then be separated by a distance equal to the side of the square submatrix. Indeed, this distance $\sqrt{N}$ would correspond to the uniform mapping parameter k_1. On the other hand, N in general need not be a

perfect square, and $\sqrt{N}$ need not be an integer. Furthermore, k_1 should be prime to N for one-to-one mapping, and if $\sqrt{N}$ should indeed be an integer, it will not be prime to N. All that can be said in conclusion, then, is that in the absence of a better criterion, a value of k_1 that is close to $\sqrt{N}$, and at the same time is prime to N, would represent a reasonable design value.

A second viewpoint on the value of k_1 relates to adjacent sample correlations. The U permutation causes input samples separated by a distance k_1 to be brought together. Thus, if input samples separated by the distance k_1 are uncorrelated, then the U permutation will convert the original N sequence to one in which adjacent samples tend to be uncorrelated. In a subsequent section, we discuss cross-correlations between input and scrambled sequences (and corresponding cross-spectra).

The third observation is that there is an interesting geometrical interpretation of k_1. It is related to the slopes of the zig-zag straight line loci obtained by joining successive 1s in the permutation matrix while scanning it from top to bottom (increasing r).

Finally, notice that the special value of $k_1 = N - 1$ corresponds to DFT inversion (local time inversion, as discussed earlier).

2.4.2 Number of keys in U permutations

It can be seen that the number of keys (distinct mappings K for U permutation is a product of the form

$$K_{\text{U}} = NG(N), \tag{14}$$

where $G(N)$ is used to denote the number of k_1 values that are prime to the block length N, and N is the number of cyclic shifts (translations by one sample) of an input block prior to permutation via a given permutation matrix (which is equivalently expressed as the number of ways of selecting the first row in vertical cyclic shifts of the permutation matrix). It can be verified that the following values of G apply for the special cases where N is prime or a power of two:

$$G(N) = N - 2 \text{ if } N \text{ is prime}$$

$$G(N) = N - D - 1 \text{ if } N = 2^{\text{D}}. \tag{15}$$

Numerical values of K_{U} are discussed in Section 2.5.

2.4.3 Effect of U permutations on frequency spectrum

The effect of U permutations on input spectrum follows immediately from an earlier relation (10). For U permutations, then, the DFT transposition matrix T is characterized by

$$NT_{rs} = 1 + W^{r-sk_1} + W^{2r-2sk_1} + \cdots. \tag{16a}$$
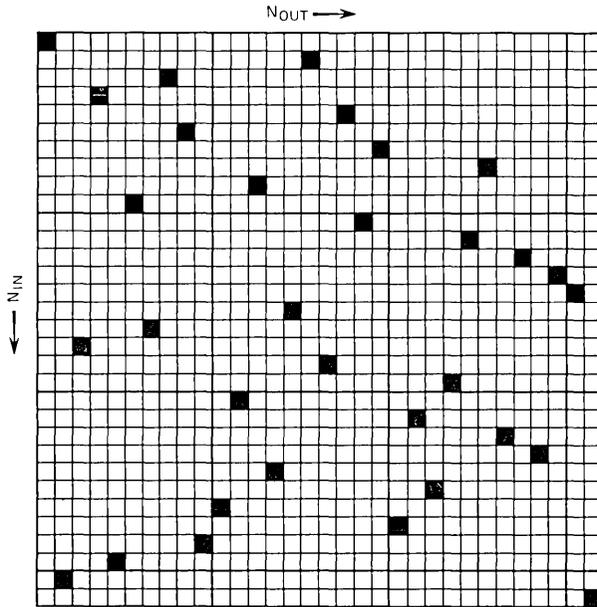
Fig. 3(a)— Pseudorandom permutation matrix.

Summing the geometric series above,

$$T_{rs} = \frac{1}{N}\left[\frac{1 - W^{rN - sk_1N}}{1 - W^{r - sk_1}}\right].$$

Because $W^N = 1$ [from (6)], the numerator above is always zero. A nonzero T_{rs} therefore requires that the denominator is zero as well:

$$W^{r - sk_1} = 1, \quad \text{or}$$

$$r - sk_1 = 0, \quad \text{or}$$

$$s = r/k_1. \tag{16b}$$

It can be verified from (16a) that this condition will make $T_{rs} = 1$. Furthermore, if (16b) should hold for all r, one requires that s/r should be independent of r. That is,

$$s = k_3 r, \tag{16c}$$

so that from (16b),

$$k_3 r = r/k_1 \text{ or}$$

$$k_1 k_3 = 1 \ (\text{modulo } N). \tag{16d}$$

In other words, uniform permutation in the time-domain causes a uniform permutation in the frequency domain. Thus, if the U permu-
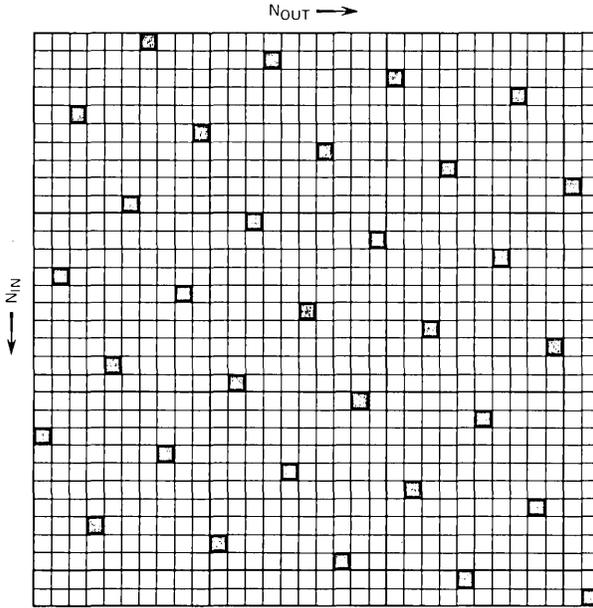
Fig. 3(b)—Uniform permutation matrix.

tations use $k_1 = 7$, $k_3 = 23$ from (16d); and the mapping (16c) (which denotes nonzero T_{rs} values) indicates the following frequency transpositions: $[0 \rightarrow 0, 1 \rightarrow 23, 2 \rightarrow 14, 3 \rightarrow 5, \cdots$ or in general $t \rightarrow 23t$ (modulo 32) for $t = 0, 1, \cdots 31]$. Notice, once again, the special example of DFT inversion and the corresponding local time inversion; these are expressed by the mappings $k_3 = N - 1$ and $k_1 = N - 1$, respectively.

2.5 Comparisons of PR and U permutations

2.5.1 Permutation matrices

Figures 3a and b are typical permutation matrices for PR and U permutations. As in Fig. 1, the PR matrix of Fig. 3a is based on a fifth-order primitive polynomial $Q_5(y) = y^5 + y^2 + 1$, a beginning shift-register state of 00001 and a mapping of position 32 into itself. Following Fig. 2, the U matrix of Fig. 3b is based on a value of $k_1 = 7$.

2.5.2 Number of keys in encryption

Table II lists illustrative values of the number of keys K_{PR} and K_U, as obtained from (12) and (14), respectively. Note that with both PR and U mappings, not all of the total number of permutations can be equally effective in destroying speech intelligibility. A good example of a per-

Table II — Number of keys in PR and U permutations

N (number of samples in block)	$K_{PR} = (N-1)\mathrm{L}(\log_2 N)$	$K_U = NG(N)$
8	14	32
16	30	176
32	186	832
64	378	3648
128	2286	15360
256	4080	63232

ceptually uninteresting permutation is DFT inversion* (see Section IV). In any case, on the basis of Table II, N_U increases much faster with N than N_{PR}. For this reason, U permutations are potentially more attractive candidates, *a priori,* for formal encryption.[†]

2.5.3 Effects on frequency spectrum

Assessments of speech encryption techniques should be really based on perceptual testing; Section IV describes such testing and also provides interesting speech spectrograms. Meanwhile, it is instructive to compare PR and U permutations on the basis of their effects on illustrative input spectra.

Figure 4 provides these comparisons. Figures 4 through 6 use linear low-pass, high-pass, and mid-pass models for input spectra, while the input spectrum of Fig. 7 is a simplified three-pole model that could be an example of the short-term spectrum of a voiced speech sample. Note that both PR and U permutations are effective in distorting input spectra; in fact, the tendency in each case is to whiten the spectrum. However, the whitening in PR permutations is both global and local, while U permutations produce whitening only in a global (average) sense. In other words, PR-scrambled spectra are smoother in terms of adjacent sample transitions in the output spectrum. This is due to the fact that entries in the DFT-transposition matrix T, eqs. (7a) and (8a), have varying weights in PR permutation. With U permutations, on the other hand, the only nonzero entries in T are ones; the effect is simply one of rearranging input DFT samples without any magnitude weighting.

2.6 Cross-correlations between input and scrambled samples

Spectral distortions that provide optimal speech encryption are difficult to specify; and further, they are likely to be input-spectrum de-

---

* It is interesting that DFT inversion is ineffective despite the fact that the associated permutation (time inversion) is characterized by a large input-output "distance" $\sum_{r=1}^{N} |r - s(r)|$.

[†] Another advantage of U permutations is that N need not be a power of 2.
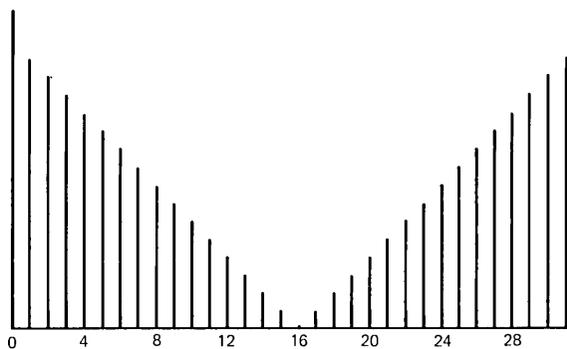
(a) ORIGINAL SPECTRUM
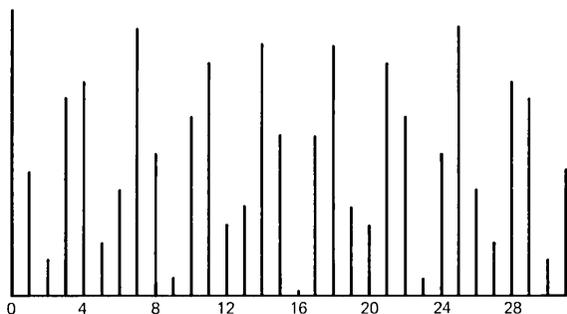
(b) U PERMUTATION SPECTRUM (k = 7)
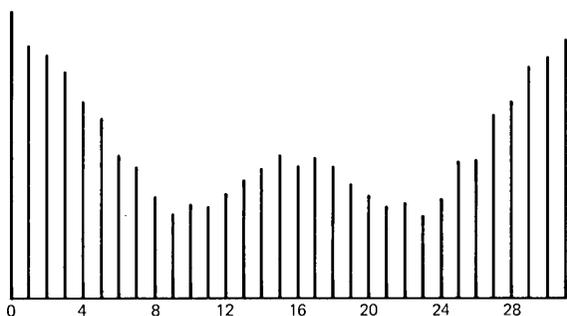
(c) PR PERMUTATION SPECTRUM

Fig. 4—Effects of PR and U permutations on low-pass spectrum.

pendent. It appears reasonable, however, that the whitening effects in Figs. 4 through 7 are very desirable for encryption. A different criterion for encryption is the decorrelation of input and scrambled spectra; spectral whitening is in general a sufficient, but not necessary, condition for such decorrelation.

Let us define a correlation measure C by $x'x$, which represents the dot
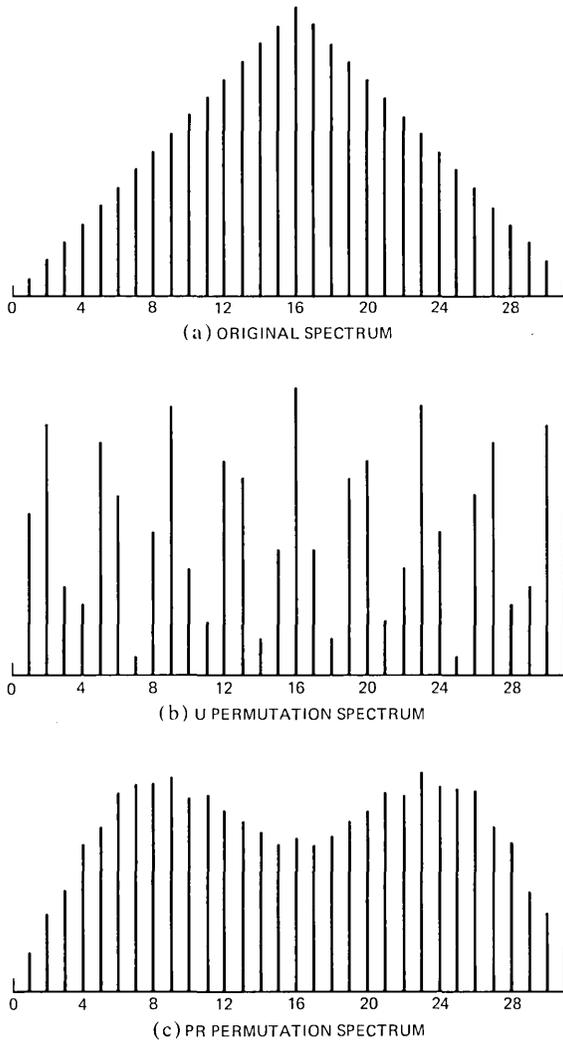
Fig. 5—Effects of PR and U permutation on high-pass spectrum.

product of an input sequence with itself. Correlation between scrambled and original sequences is $x'_P x$, where the subscript P indicates permutation or scrambling. A corresponding correlation between input and scrambled spectra is $X'_P X^*$, where X^* is the complete conjugate of the input spectrum X. Notice that

$$X'_P X^* = (FPx)'(Fx)^* = (Px)'F'F^*x.$$

It can be shown from (5) that $F'F^* = IN$, where I is an identity matrix. Consequently, spectral and time-domain cross-correlations between

(a) ORIGINAL SPECTRUM

(b) U PERMUTATION SPECTRUM

(c) PR PERMUTATION SPECTRUM
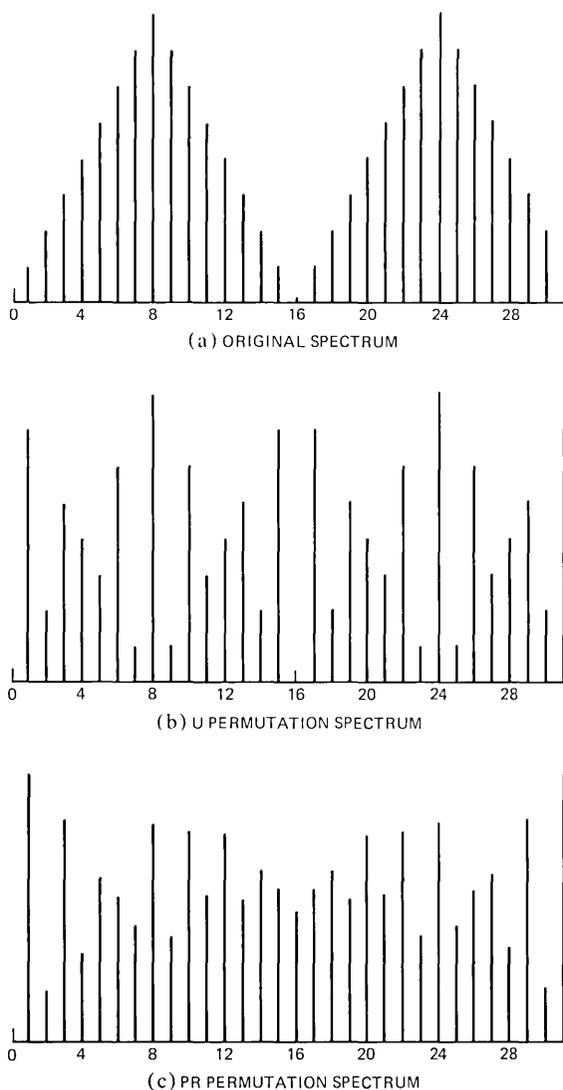
Fig. 6—Effects of PR and U permutations on mid-pass spectrum.

input and scrambled samples are related by

$$X'_P X^* = N(Px)'x. \tag{17}$$

In other words, spectral decorrelation requires that the permuted and the original sequences are themselves orthogonal or uncorrelated. The significance of negative cross-correlations is not always clear. In fact, in at least one instance, a negative cross-correlation of -1 between input and scrambled spectra is known to be perceptually suboptimal. This is
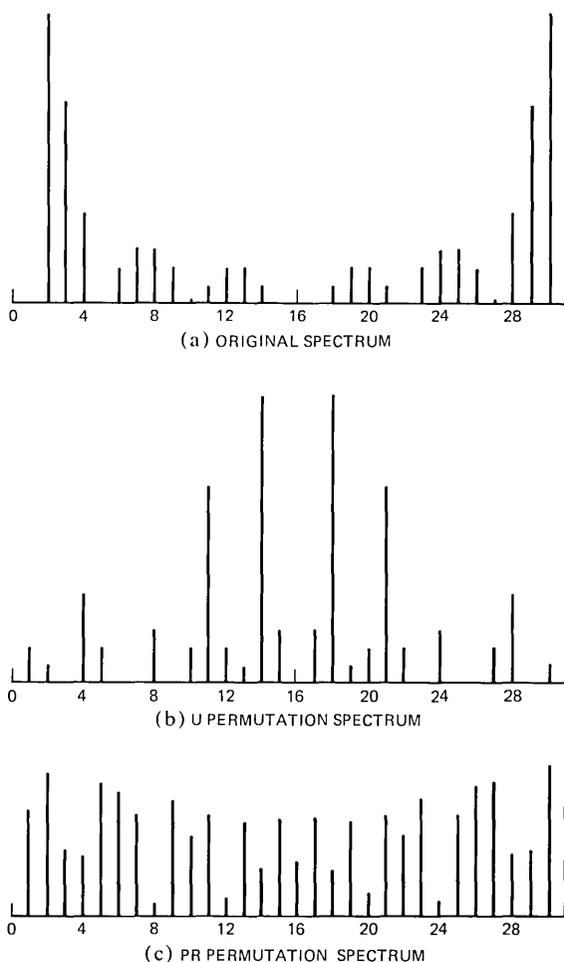
Fig. 7—Effects of PR and U permutations on an example of a voiced speech spectrum.

the example of frequency inversion (Section III); however, it has been reported that listeners can be trained to understand frequency-inverted speech[1]. It would appear then that scramblers that cause the cross-correlations in (17) to approach zero are, in general, very desirable. The case of zero correlation is realized for U permutation, for example, if input message samples that are separated by a distance of k_1 are uncorrelated.

2.7 Permutations of binary sequences

In practical speech-encryption techniques, we scramble speech-carrying bits in digital codes such as PCM, DPCM, and DM[3] rather than
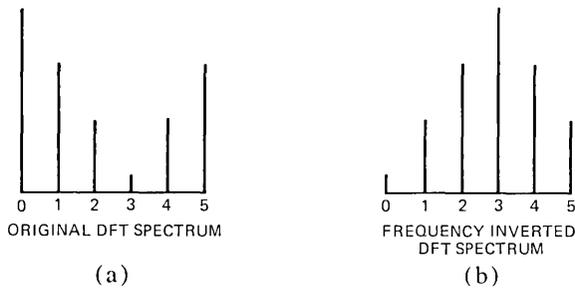
Fig. 8—Frequency inversion.

PAM speech samples. Although the analyses of preceding sections can be applied in principle to the special case of binary samples, the utility of such analyses may be limited because the issue of interest is the effect of scrambling on the decoded speech spectrum rather than the spectrum of the bits itself. Furthermore, there are many other factors, such as the varying significance of PCM (or DPCM) bits depending on their position in a PCM (or DPCM) word. In view of the above problems, we have deferred our observations on speech code scrambling to the section on perceptual experiments (Section IV), rather than attempt analytical predictions. We briefly analyze, on the other hand, the interesting case of permutations with sign changes, before proceeding to summarize results of perceptual experiments.

III. TRANSFORMATIONS WITH SIGN CHANGES—FREQUENCY INVERSION

The total number of permutations of a sequence of N samples is $N!$. If sign changes are allowed in addition to position changes, the total number of possible transformations increases to $2^N N!$. The analytical approaches of Section II can perhaps be extended to study such transformations. The purpose of this section, however, is only to consider a very specific transformation. This transformation involves no explicit permutation, but it introduces sign changes in every other input sample. This is equivalent to the classical encryption technique of analog frequency inversion.[1]

We begin by recapitulating that DFT inversion, as given by (11a) and (11b), does not constitute analog frequency inversion, because the DFT of a real signal is symmetric in the sense $|X(r)| = |X(N-r)|$ for $r = 1$, $2, \cdots N$, and the highest analog frequency corresponds to $[N/2]$ where $[\ ]$ indicates "largest integer in." Analog frequency inversion would result, on the other hand, if the DFT coefficients are inverted about $N/4$ (Fig. 8), assuming that N is even. (Such inversion is achieved if all the DFT coefficients are translated cyclically through a distance of $N/2$. For
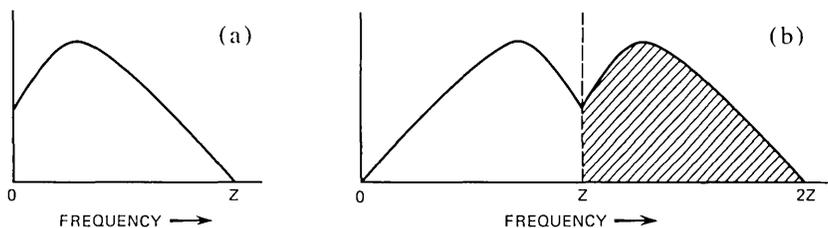
Fig. 9—Frequency inversion by modulation and low-pass filtering. (a) Original band-limited spectrum. (b) Result of modulating spectrum with carrier at Z. The left half of (b) is the inverse of (a).

odd N, there is a modification of these coefficients as well, since a simple translation will render the spectrum asymmetric.)

For a cyclic shift through $N/2$, the N by N frequency transformation matrix (for even N) is

$$T = \begin{bmatrix} 0 & \mathrm{I} \\ \mathrm{I} & 0 \end{bmatrix},$$
(18a)

where the submatrices are each of order $N/2$. From (8b), the corresponding input transformation is

$$P_{\text{FREQ INVERSION}} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdot & \cdot \\ 0 & -1 & 0 & 0 & & \\ 0 & 0 & 1 & 0 & & \\ 0 & 0 & 0 & -1 & & \\ \cdot & & & & \cdot & \\ \cdot & & & & & \cdot \end{bmatrix}.$$
(18b)

Thus, analog frequency inversion is obtained if alternate input samples are multiplied by -1. This is also apparent from the fact that frequency inversion occurs when a band-limited signal is modulated by a carrier at the highest input frequency, and out-of-band frequencies are filtered out from the modulation product (see Fig. 9). In the digital technique, the sequence of $+1$ and -1 elements in T corresponds to the carrier, and its frequency is simply the frequency of $+1$ (or -1) entries.

IV. EXPERIMENTAL RESULTS

In this section, we summarize experimental observations on speech waveform encryption using PAM samples as well as APCM, ADPCM, and ADM codes.[3] The letter A in code notation stands for instantaneously adaptive quantization,[3] and our DPCM and DM codes used simple first-order predictors. Our results are from computer simulations, and the studies have included PR and U permutations, analog frequency inversion, and (the academic case of) DFT inversion. Our conclusions are based on informal perceptual tests and on spectrograms of original and en-

crypted speech. The speech sample used in the experiments was a 2-second band-limited (200 to 3200 Hz) utterance, "The chairman cast three votes."

4.1 Perceptual observations

4.1.1 Scrambling with PR and U permutations

The sampling frequency for PAM and ADM bits was 24 kHz. The APCM and ADPCM bits used Nyquist-sampled (8 kHz) speech with 3 bits of quantization per sample. Thus, the scramblers operated on 24-kHz sequences in all four cases, and the encoding delay due to scrambling was the same for all the four waveform formats as long as the scrambler block length N was the same. Values of N ranged from 4 to 64. The speech-encoding qualities resulting in the identical bit-rate (24 kb/s) ADM, APCM, and ADPCM codes were, of course, noticeably different,[3] but this issue is not strictly relevant to the present discussion. The following were the main observations on encryption efficiency.

(i) In comparing pseudorandom and uniform permutations, we did not notice important perceptual differences: the U-permutation speech scramblers were about as good as the PR scramblers for given N, and, in fact, they were slightly better in some cases than PR scramblers.

(ii) In comparing speech code formats from the point of view of the benefits of scrambling, there was a definite ordering of efficiency:

$$\text{PAM} < \text{ADM} < \text{APCM} < \text{ADPCM}. \tag{19}$$

This means that for a given encoding delay (or scrambling block length N), the least desirable candidate for scrambling is a PAM sequence, while the most desirable format was an ADPCM code. The following specific observations are worth noting:

(a) Scrambling of PAM samples is marginally effective with $N = 64$.

(b) Scrambling of ADM samples is quite effective with $N = 64$. In fact, for very casual encryption using ADM bit-scrambling, even $N = 32$ can be useful; with $N = 32$, the specific U permutations that were used ($k_1 = 7$) were slightly more effective than the PR permutations.

(c) Scrambling of APCM and ADPCM bits destroys speech intelligibility very effectively with $N = 32$, while for casual encryption even $N = 16$ can be useful. Thus, ADPCM at 24 kb/s, with a scrambling delay of $16/24 = 0.67$ ms, can be a very attractive candidate for practical speech-privacy systems (for example, optional facilities for privacy in mobile telephony).

4.1.2 Frequency inversion

The speech formats used here were the same as for the scrambler studies, except in the case of PAM samples. These were sampled at the

Nyquist frequency of 8 kHz (instead of 24 kHz) to provide frequency inversion (Section III and Fig. 9), which is a classical speech-encryption procedure.[1] As mentioned in Section III, the inversion of every other Nyquist-rate PAM sample provides a simple digital technique for speech frequency inversion. The sign-reversal operations on ADM, APCM, and ADPCM codes are less interesting in that they do not provide speech spectrum inversion, but only the inversion of respective bit spectra. In any case, sign reversals of adjacent samples provided effective speech distortions for all the speech-formats studies, with the following ordering of encryption efficiency in informal tests:

$$ADM < PAM < APCM < ADPCM. \qquad (20)$$

The important point, however, is that unlike in scrambling, encryption potential in frequency inversion cannot be increased by such simple means as increasing a block length N ($N = 2$ is the minimum as well as maximum useful block length for frequency inversion). More importantly, owing to the fact that only one key is associated with this technique, frequency inversion is not suitable for formal encryption with time-varying keys. Furthermore, even for casual privacy, the aforementioned utility of frequency inversion (with PAM samples) is to be qualified with the observation that listeners can be trained to follow frequency-inverted speech. Frequency inversions with ADM bits were rather ineffective, for reasons not completely understood by the authors; while the more effective APCM and ADPCM inversions (20) were no more efficient than APCM- and ADPCM-based scramblers. (See, for example, spectrograms in Figs. 10a and b.)

Finally, perceptual tests confirmed our earlier observation that DFT inversions are academic and useless for speech encryption.

4.2 Speech spectrograms

Figure 10 reinforces and supplements the perceptual tests summarized in the previous section.

Figure 10a compares the classical technique of frequency inversion with a scheme that inverts the signs of alternate ADPCM bits. Notice the lack of speech-like patterns in the latter example and the contributions of ADPCM quantization noise that also serve to reduce speech intelligibility. In contrast, the frequency inversion of PAM samples leads to a spectrogram that is informationally equivalent to the original speech spectrogram, being only a mirror image of it across the 4-kHz line.

Figure 10b shows the benefits of increasing the block length N in temporal scrambling for the example of ADPCM codes and PR permutations.

Figure 10c demonstrates the effect of U permutations in scrambling, and compares PAM, ADM, and ADPCM samples as candidates for

"THE CHAIRMAN CAST THREE VOTES"

ORIGINAL

INVERSION OF ALTERNATE
PAM SAMPLES
(FREQUENCY INVERSION)

INVERSION OF ALTERNATE
ADPCM SAMPLES

FREQUENCY IN KILOHERTZ

TIME (APPROXIMATE TOTAL DURATION IS TWO SECONDS)
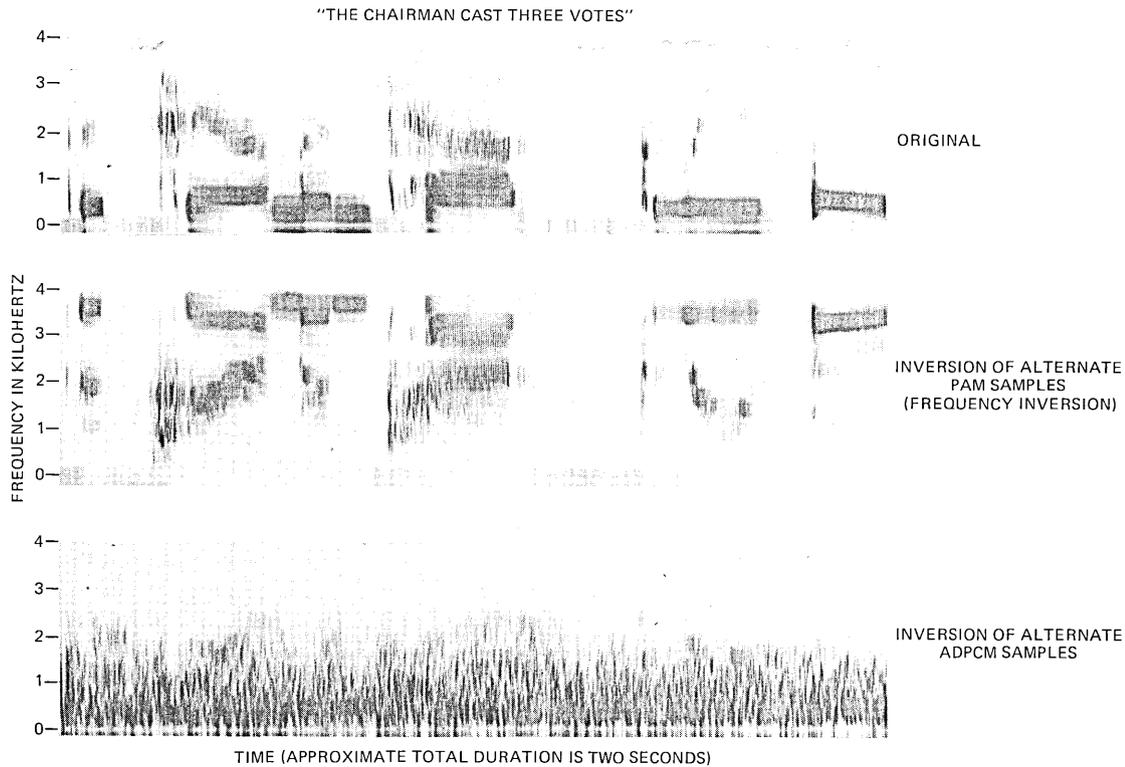
Fig. 10(a)—Spectograms of original and scrambled speech: original speech, frequency inversion, sign changing of alternate ADPCM bits.

"THE CHAIRMAN CAST THREE VOTES"
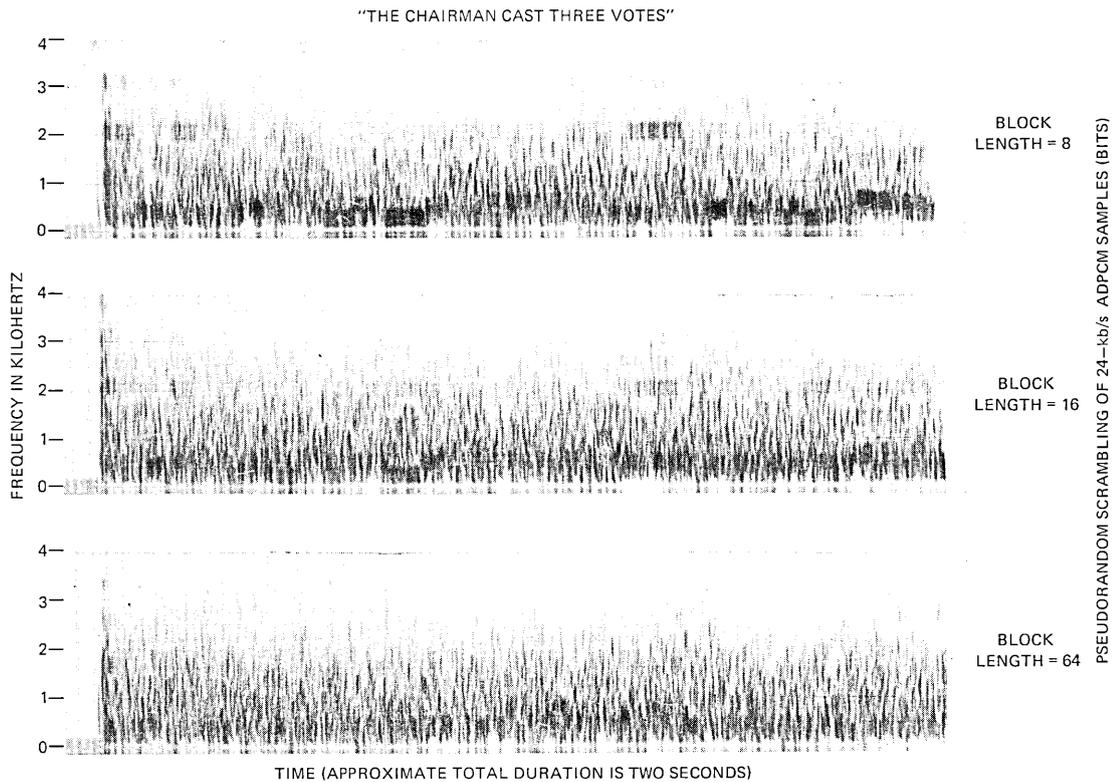
Fig. 10(b)—Spectograms of PR scrambling of ADPCM bits: N = 8, 16, 64.

"THE CHAIRMAN CAST THREE VOTES"
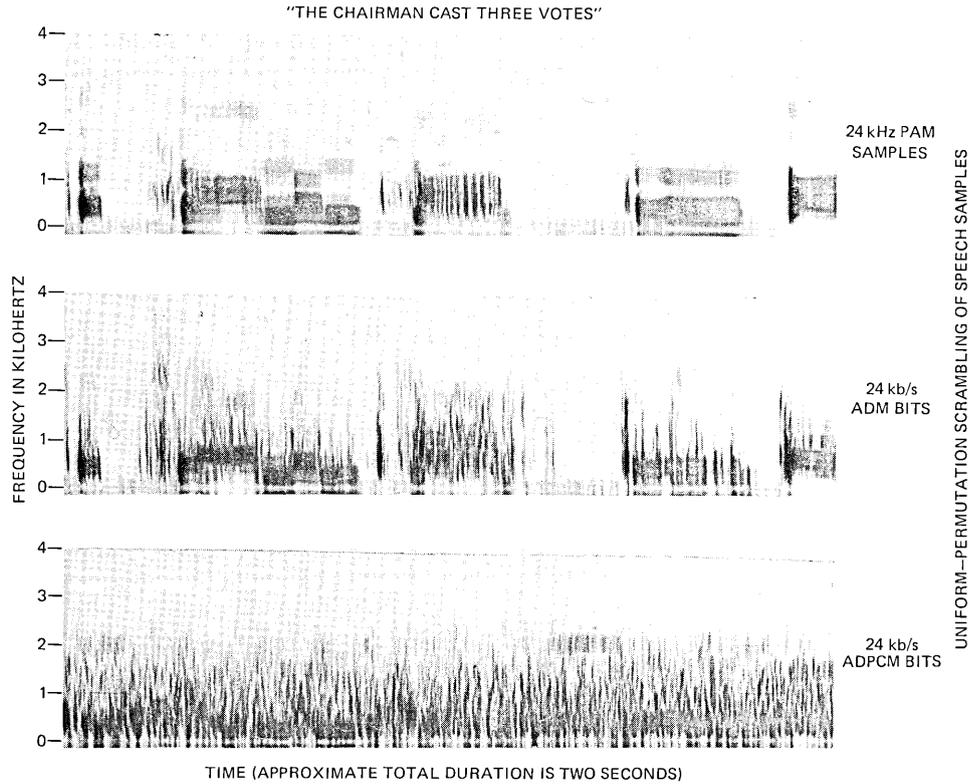
Fig. 10(c)—Spectrograms of U scrambling of speech codes ($N = 16$): PAM, ADM, ADPCM.

scrambling. The spectrogram confirms the ordering of efficiency noted in (19): PAM < ADM < ADPCM.

V. ENCRYPTION, SCRAMBLING, AND MASKING

The preceding sections have looked mainly at effective privacy employing temporal permutations. The effectiveness of scrambling for encryption would depend on the rate at which the key is changed, the distortion produced by each key, and the number of keys available. Assuming that the distortion introduced by a typical key is adequate, a measure of goodness of any subfamily of permutations would be the number of members of that set. On this count, U permutations are better than PR permutations ($K_U > K_{PR}$ for a given N, Table II). We do not have a sufficient understanding of how many of these K_U or K_{PR} permutations are desirable from the point of view of distorting speech.* But we believe that for both PR and U permutations, the numbers of trivial or obviously useless permutations are small fractions of the number of possibilities in Table II. It can also be shown that the degree of complexity in encryption is equal to $\log_2 K$. This number results from statistical procedures that minimize the number of receivers needed for cryptanalysis.[4]

Many so-called speech scramblers encrypt speech not by temporal permutations (as in our definition of scrambling), but by the addition of appropriate masking signals. A typical example is the EXCLUSIVE OR mod-2 addition of a pseudorandom binary sequence, bit by bit, to an ADM or PCM bit stream.

If the masking signal is such that the spectrum of the sum signal is white, we obtain a perfect cipher in some sense. However, the entropy of the key must equal the entropy of the signal to be encrypted, and a perfect cipher is an idealized case in this sense. Realistically, we seek masking signals that change slowly with the message waveform, and typical resultant spectra are non-speech-like although not perfectly white.

The use of modulo-m additions (example modulo-2 adds) in masking ensures that the amplitude range of the encrypted signal is the same as that of the input signal; preservation of properties such as dynamic range serves as a precaution against techniques of cryptanalysis.

Temporal scrambling has the following limitations with respect to masking: (*i*) scrambling introduces encoding delays (= N) and (*ii*) it leaves long ($\gg N$) tell-tale silences unaffected in speech encryption. Masking, on the other hand, is characterized by a greater complexity of

---

* This depends not only on the properties of the permutations, but also on those of the samples being permuted. Highly correlated PAM samples, for example, demand more drastic permutations than less redundant ADM bits for a given N.

the key signal. If delays of a few milliseconds are not objectionable and if the exposure of the on-off (speech-silence) patterns is not a problem, the simplicity of scrambling should make it more attractive.

A final issue is the comparison of temporal scrambling and masking purely from the point of view of reducing the intelligibility of speech sounds. In at least one experiment,[8] performances have been found to be very comparable. This experiment used 24-kb/s ADPCM speech bits. For scrambling, a block length of $N = 16$ was employed, and for masking, a 16-bit PR number (binary sequence) was added (modulo 2), bit-by-bit, to each of 16 contiguous ADPCM bits in the speech code to be masked.

For practical implementations, it is conceivable that sophisticated schemes may run permutation and masking operations in tandem. Also, if the associated keys are time-varying, it may be practical to use certain serial configurations for the shift register, as described in Ref. 9, in place of a bank of conventional hard-wired registers.

APPENDIX

Proof of Theorem (10)

Recall from (8a) that $T = FPF^{-1} = [FP]F^{-1}$. The value of T_{rs} is the dot-product of the sth column of $[F^{-1}]$ and the rth row of $[FP]$.

The sth column in F^{-1} has a typical element W^{-Rs}/N where R is the row number [see (9)].

To evaluate the dot product that gives T_{rs}, we need V_{rR}, the Rth column element in the rth row of $[FP]$. The quantity V_{rR} in turn is the product of the rth row in F and the Rth column in P. The only nonzero elements in P are the 1s that occur in positions $a, b, c, d, \cdots$ in rows 0, 1, 2, 3, $\cdots$ as characterized by $P[0 \rightarrow a, 1 \rightarrow b, 2 \rightarrow c, 3 \rightarrow d, \cdots]$. Furthermore, the rth row in F has a typical element W^{rn}, where n is the column number [see (5)]. Consequently, V_{rR} has the form $W^{rn(R)}$, where $n(R)$ is the column element in the rth row of F, which gets multiplied by the single 1 entry in the Rth column of P. (The other elements in the rth row of F get multiplied by 0 entries in the Rth column of P, and do not contribute to V_{rR}.)

Consequently, the dot product that specifies T_{rs} has the form

$$\frac{1}{N} \sum_{R=0}^{N-1} W^{-Rs} V_{rR} = \frac{1}{N} \sum_{R=0}^{N-1} W^{-Rs} W^{rn(R)}.$$

Because of a one-to-one mapping between R and $n(R)$, the above summation can be rewritten as a summation over $n(R)$:

$$T_{rs} = \frac{1}{N} \sum_{n(R)=0}^{N-1} W^{n(R)\cdot(r)-(R\cdot s)}.$$

In characterizing P, $n(R)$ values of 0, 1, 2, 3, $\cdots$ correspond to R values

of $a, b, c, d, \cdots$. Hence,

$$T_{rs} = \frac{1}{N} (W^{-as} + W^{r-bs} + W^{2r-cs} + W^{3r-ds} + \cdots).$$

VI. ACKNOWLEDGMENTS

REFERENCES

1. D. Kahn, *The Code-Breakers*, New York: The Macmillan Company, 1967.
2. M. R. Sambur and N. S. Jayant, "Speech Encryption by Manipulations of LPC Parameters," B.S.T.J., *55*, No. 9 (November 1976), pp. 1373–1388.
3. N. S. Jayant, "Digital Coding of Speech Waveforms—PCM, DPCM and DM Quantizers," Proc. IEEE, *62* (May 1974), pp. 611–632.
4. C. E. Shannon, "Communication Theory of Secrecy Systems," B.S.T.J., *28*, No. 4 (October 1949), pp. 656–715.
5. S. W. Golomb, *Shift Register Sequences*, San Francisco: Holden-Day, 1967, pp. 62–65.
6. R. G. Gallager, *Information Theory and Reliable Communications*, New York: John Wiley, 1968.
7. F. J. MacWilliams and N. J. A. Sloane, "Pseudo-Random Sequences and Arrays," Proc. IEEE, *64* (December 1976), pp. 1715–1728.
8. N. S. Jayant, unpublished work.
9. S. V. Ahamed, "The Design and Embodiment of Magnetic-Domain Encoders and Single-Error-Correcting Decoders for Cyclic Block Codes," B.S.T.J., *51*, No. 2 (February 1972), pp. 461–485.

Techniques for Coding Dithered Two-Level Pictures

By A. N. NETRAVALI, F. W. MOUNTS, and J.-D. BEYER

This paper considers several methods for the efficient coding of two-level pictures dithered to give the appearance of multiple amplitude levels. In the dithering technique, a multilevel image signal is compared with a position-dependent set of thresholds (called a dither matrix), and, if the image value exceeds the threshold, the two-level output signal is taken to be "white," otherwise it is taken to be "black." Spatial correlation present in the original image is not preserved in the two-level picture due to spatial variation of the value of the threshold; and, therefore, standard techniques for coding two-level pictures, such as run-length coding, lose their efficiency. We show how some of our recently developed techniques for coding two-level pictures can be modified to code two-level dithered images. Our computer simulations on a few representative two-level dithered pictures indicate that an entropy between 0.2 to 0.3 bit/pel is possible using our technique. A comparison with some recently proposed techniques by Judice indicates that those schemes result in about 10 to 60 percent higher entropy than our schemes.

I. INTRODUCTION

Techniques for representing the entire gray scale of a picture by only two levels have been receiving considerable attention[1-10] because many devices are limited to recording or displaying two-level signals. Although these techniques may differ in specific algorithms, they provide the subjective illusion of a wide range of gray shades by controlling the proportion of picture elements in a neighborhood that are in the "on" state.

One of the techniques (called "dither" and studied in detail by Limb,[2] Lippel and Kurland,[3] and Judice et al.[9-10]) consists of comparing the multilevel input image signal with a position-dependent set of thresholds and setting only those picture elements to "white" (or 1) where the image

PEL COLUMN ℓ	ℓ+1	ℓ+2	ℓ+3	
0	128	32	160	LINE m
192	64	224	96	LINE m + 1
48	176	16	144	LINE m + 2
240	112	208	80	LINE m + 3

Fig. 1—A 4 by 4 dither matrix to be used for images with 8-bit quantized samples.

input signal exceeds the threshold. A square matrix of threshold values (elements of a "dither matrix") is repeated as a regular array to provide a threshold pattern for the entire image. Subjective effects of gray shades are achieved by using a dither matrix. The 4 by 4 dither matrix used by Judice et al.[9] for an input image having gray levels between 0 and 255 is shown in Fig. 1. The values of the adjacent elements of the dither matrix were chosen to take advantage of the spatial low-pass filtering present in the human visual system and, at the same time, to reproduce edges accurately and to avoid objectionable patterns. When the input image intensity is compared with spatially varying thresholds, a large amount of spatial correlation present in the input image is suppressed. This loss of correlation in the dithered two-level image makes some of the standard methods of coding two-level signals, such as run-length coding,[11] inefficient.

Several modifications of the standard techniques are possible. Judice has considered two such modifications. In one scheme,[12] called bit interleaving, runs of picture elements corresponding to equal or near equal elements of the dither matrix are coded using standard techniques. The other scheme,[13] called pattern matching, assigns a code to two-dimensional bit patterns of the dithered image and relies for bit-rate reduction on the fact that all possible bit patterns do not occur with the same frequency.

In our schemes, the "state" of the coder when an element is to be coded is a function of the already transmitted values of surrounding elements and the dither threshold at the element to be coded. The value that minimizes the probability of prediction error conditioned on a state is the predicted value of the element to be coded. This is an extension of the predictive coding for two-level pictures discussed in Ref. 14. In one of our schemes, we code run-lengths of the prediction error. In our other schemes, we change the relative order of the picture elements along a scan line in such a way as to increase the average run-length of the black

and/or white elements and then transmit the run-lengths. These schemes are described in detail in Section 2.3.

We have investigated the efficiency of our techniques by computer simulation on a few representative 4- by 5-inch pictures, scanned with an array of 512 by 512. Using run-length coding of prediction errors, it is possible to decrease the bit rate to about 0.22 to 0.42 bit/pel. "Good-bad" state ordering performs best among all the ordering schemes that we have considered, and it brings the bit rates down to between 0.20 to 0.30 bit/pel. As a comparison, results from simulation of the same pictures by Judice show a bit rate of between 0.22 to 0.48 bit/pel.

II. CODING ALGORITHMS

In this section, we describe our coding algorithms in detail and present results of our computer simulations. The pictures used for computer simulation are shown in Fig. 2. These pictures were 4 inches by 5 inches and were scanned with 512 samples along a line and 512 lines. Each picture element was digitized by a uniform PCM coder to an accuracy of eight bits (256 levels). The digitized signal was dithered by using the dither matrix shown in Fig. 1. The coding algorithms were applied to the dithered images. As a measure of performance, we used the sample first-order entropy of run-length statistics. We computed the average black and white run-lengths and the entropy of black and white runs using, for example, the formula

$$E_w = -\sum \frac{n_i}{N} \log_2 \frac{n_i}{N},$$

where E_w is the entropy of white run-lengths, n_i is the number of white runs of length i, and N is the total number of white runs. Using these, and eq. (1), we computed the entropy in bits/pel by:

$$E = \frac{E_w N_w + E_b N_b}{r_w N_w + r_b N_b} \tag{1}$$

$$= \frac{E_w + E_b}{r_w + r_b} \quad \text{(when } N_w \cong N_b\text{)},$$

where
  E_b is the entropy of the black run statistics (bits/run)
  r_w is the average white run-length (pels/run)
  r_b is the average black run-length (pels/run)
  N_w, N_b are the number of white and black runs, respectively
  E is the entropy in bits/pel.

2.1 Prediction algorithm

Consider a dithered picture element S_{ij} at position (i,j) where the dither matrix has value D_{ij}. To develop a predictor for S_{ij}, consider

(a)
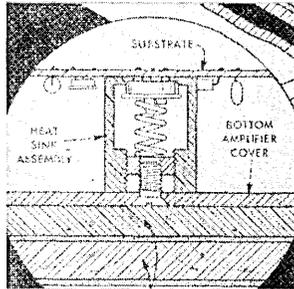
(b)

(c)

Fig. 2—Dithered images used for computer simulation of coding algorithms. (a) Karen. (b) Engineering drawing. (c) House.

surrounding elements W, X, Y, Z as shown in Fig. 3. Note that elements W, X, Y, Z and the value of the dither matrix D_{ij} are known to the receiver when decoding the signal at point (i,j). The state associated with S_{ij} is defined as the five-tuple

$$Q = (D_{ij}, W, X, Y, Z). \tag{2}$$

Since we are using a 4 by 4 dither matrix, D_{ij} can have 16 different values. Each W, X, Y, Z can have two values, and, therefore, the number of possible states is 256. Let these be denoted by the set $\{Q_k\}$, $k = 1, \cdots, 256$.
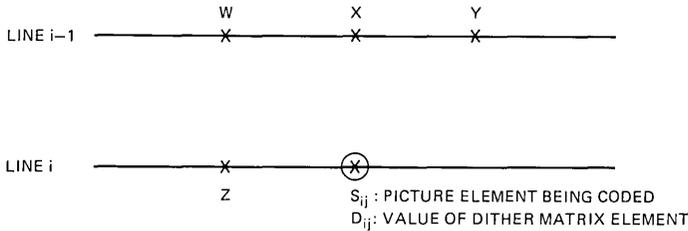
Fig. 3—Configuration for definition of state.

Then a predictor code book is given by:

$$C(Q_k) = \text{``0''} \text{ if } P(S_{ij} = \text{``0''}/Q = Q_k) \geqq 0.5$$
$$= \text{``1''} \qquad \text{otherwise,} \qquad (3)$$

where $C(\cdot)$ is the predictor code book and $P(\cdot|\cdot)$ is the experimentally determined conditional probability that a picture element has a value of "0" or "1" given the state Q_k. Thus, the predictor thus depends upon the previously transmitted values in a neighborhood and some partial information about the present picture element known both to the transmitter and the receiver in the form of the value of the dither matrix. The code book can be designed for each picture and transmitted before actual picture transmission,* or it can be taken to be an average for a class of pictures. A portion of the code book for each of the three pictures that we considered is shown in Table I, where we assign the state number, k, by computing formula (4), and show the number of occurrences of each state and the probability of error.

$$k = D_{ij} + 8W + 4X + 2Y + Z + 1. \qquad (4)$$

We note that the probability of prediction being in error is always less than 0.5 due to our method of prediction. The code book defined by eq. (3) does vary from picture to picture. We examine the effects of such variation in a later section.

2.2 Run-length coding of prediction errors

In this technique, we code the run-lengths of the prediction errors along a scan line using the appropriate code book for each picture. The entropy of the run-length statistics is given in Table II. It varies between 0.22 to 0.42 bit/pel. To evaluate the effects of the variation of code book with respect to pictures, we used the code book of one picture for the prediction of another. The resulting entropies of the run-lengths of

---

* Only 256 bits are needed to transmit a code book; since there are $(512)^2$ picture elements per picture, transmission of one code book per picture corresponds to an additional 0.001 bit/pel.

Table I—State dependent predictors for different pictures. Also shown is the probability of prediction error arid the number of elements in state. Only the first 54 states are shown for brevity.

State Definition						Picture: Karen			Picture: Engineering Drawing			Picture: House		
State Number	Dither Matrix (D_{ij})	W	X	Y	Z	Total Elements in State	Pre-diction	Probability of Prediction Error	Total Elements in State	Pre-diction	Probability of Prediction Error	Total Elements in State	Pre-diction	Probability of Prediction Error
1	0	0	0	0	0	4026	1	0.018	3197	1	0.041	6225	1	0.000
2	0	0	0	0	1	12	1	0.000	197	1	0.000	0	0	0.000
3	0	0	0	1	0	77	1	0.000	760	1	0.011	32	1	0.000
4	0	0	0	1	1	2	1	0.000	97	1	0.000	0	0	0.000
5	0	0	1	0	0	0	0	0.000	0	0	0.000	0	0	0.000
6	0	0	1	0	1	0	0	0.000	0	0	0.000	0	0	0.000
7	0	0	1	1	0	0	0	0.000	0	0	0.000	0	0	0.000
8	0	0	1	1	1	0	0	0.000	0	0	0.000	0	0	0.000
9	0	1	0	0	0	2590	1	0.000	1489	1	0.003	5665	1	0.000
10	0	1	0	0	1	43	1	0.000	429	1	0.000	0	0	0.000
11	0	1	0	1	0	6121	1	0.000	2342	1	0.002	4204	1	0.000
12	0	1	0	1	1	3258	1	0.000	7618	1	0.000	3	1	0.000
13	0	1	1	0	0	0	0	0.000	0	0	0.000	0	0	0.000
14	0	1	1	0	1	0	0	0.000	0	0	0.000	0	0	0.000
15	0	1	1	1	0	0	0	0.000	0	0	0.000	0	0	0.000
16	0	1	1	1	1	0	0	0.000	0	0	0.000	0	0	0.000
17	16	0	0	0	0	2938	1	0.279	3259	1	0.417	4217	1	0.047
18	16	0	0	0	1	4	1	0.000	108	1	0.000	0	0	0.000
19	16	0	0	1	0	61	1	0.000	785	1	0.135	10	1	0.000
20	16	0	0	1	1	0	0	0.000	32	1	0.000	0	0	0.000
21	16	0	1	0	0	0	0	0.000	2	0	0.000	0	0	0.000
22	16	0	1	0	1	0	0	0.000	0	0	0.000	0	0	0.000

23	16	0	1	1	0	0	0	0.000	0	0	0.000	0	0	0.000
24	16	0	1	1	1	1	1	0.000	0	0	0.000	0	0	0.000
25	16	1	0	0	0	2302	1	0.003	1335	1	0.061	5193	1	0.000
26	16	1	0	0	1	40	1	0.025	267	1	0.011	0	0	0.000
27	16	1	0	1	0	9185	1	0.006	5154	1	0.009	6964	1	0.000
28	16	1	0	1	1	1811	1	0.007	5442	1	0.001	0	0	0.000
29	16	1	1	0	0	0	0	0.000	0	0	0.000	0	0	0.000
30	16	1	1	0	1	0	0	0.000	0	0	0.000	0	0	0.000
31	16	1	1	1	0	3	1	0.000	0	0	0.000	0	0	0.000
32	16	1	1	1	1	39	1	0.000	0	0	0.000	0	0	0.000
33	32	0	0	0	0	4038	1	0.403	3096	0	0.193	6317	1	0.113
34	32	0	0	0	1	24	1	0.000	396	1	0.063	9	1	0.000
35	32	0	0	1	0	2562	1	0.002	1327	1	0.251	5641	1	0.000
36	32	0	0	1	1	70	1	0.000	587	1	0.027	24	1	0.000
37	32	0	1	0	0	0	0	0.000	0	0	0.000	0	0	0.000
38	32	0	1	0	1	0	0	0.000	1	1	0.000	0	0	0.000
39	32	0	1	1	0	0	0	0.000	0	0	0.000	0	0	0.000
40	32	0	1	1	1	2	1	0.000	0	0	0.000	0	0	0.000
41	32	1	0	0	0	33	1	0.091	333	1	0.339	42	1	0.000
42	32	1	0	0	1	47	1	0.021	513	1	0.226	16	1	0.000
43	32	1	0	1	0	1996	1	0.001	1187	1	0.127	2319	1	0.000
44	32	1	0	1	1	7090	1	0.000	8774	1	0.002	1888	1	0.000
45	32	1	1	0	0	0	0	0.000	0	0	0.000	0	0	0.000
46	32	1	1	0	1	0	0	0.000	0	0	0.000	0	0	0.000
47	32	1	1	1	0	15	1	0.000	1	1	0.000	0	0	0.000
48	32	1	1	1	1	379	1	0.000	41	1	0.000	0	0	0.000
49	48	0	0	0	0	2943	0	0.271	3024	0	0.082	4136	0	0.487
50	48	0	0	0	1	9	1	0.000	229	1	0.066	0	0	0.000
51	48	0	0	1	0	2274	1	0.047	1349	1	0.375	5146	1	0.008
52	48	0	0	1	1	6	1	0.000	265	1	0.049	0	0	0.000
53	48	0	1	0	0	0	0·	0.000	0	0	0.000	0	0	0.000
54	48	0	1	0	1	0	0	0.000	0	0	0.000	0	0	0.000

prediction errors are also shown in Table II. As expected, when a code book is used that is not specifically matched to the picture, there is a loss of coding efficiency. However, by using the code book for the picture of Fig. 2a for all the pictures, there is a maximum loss of only 0.08 bit/pel. Thus, although better results in terms of entropy can be obtained by using a matched code book, it appears possible to use a general code book that will not degrade the performance significantly.

2.3 Ordering techniques

These ordering techniques are extensions of our techniques for two-level pictures.[15] In these techniques, we order either the elements or the prediction errors of the present line using a reference signal available to both the transmitter and the receiver—for example, the elements of the previous line.

To illustrate the technique, consider a memory containing 512 cells (equal to the number of elements per line). Suppose the cells of this memory are numbered from 1 to 512. If the first element of the previous line is white, then we put the prediction error for the first element of the present line in memory cell 1; if the first element of the previous line is black, then we put the prediction error for the first element of the present line in memory cell 512. We continue in this manner: the prediction error of the ith element of the present line is put in the unfilled memory cell of smallest index or of largest index depending on whether the ith element of the previous line is white or black. When the memory is filled, its cells are read in numerical order and the contents are run-length encoded. It is easy to see that the present line can be uniquely reconstructed from the knowledge of the run-lengths of the ordered line, since the ordering information is known to the receiver.

The efficiency of the simple ordering technique discussed above is given in Table II. It is seen that due to the process of dithering, much of the efficiency of ordering is lost. We overcome the effects of dithering by using an ordering technique based on the "goodness" of the state. We divide the states defined in eq. (2) into two groups. States that have a high probability of correct prediction are called "good" states and the remaining are called "bad" states. Our algorithms can be described as follows: we first evaluate the prediction error for a particular element of the present line, and then, if the state is "good," we put the prediction error in the unfilled memory cell with the smallest index; if the state is "bad," the prediction error is put in the unfilled memory cell with the largest index. Having ordered the prediction errors, we run-length code them as before.

It is easy to see that the line of picture data can be uniquely reconstructed from the coded run-lengths of the prediction error. The entropies obtained by this scheme are given in Table II. The criterion of

Table II—Entropy comparisons for different coding algorithms

			Entropy (Bit/Pel)		
Algorithms			Picture I "Karen"	Picture II "Engineering Drawing"	Picture III "House"
	Coding Algorithms				
I.	Run-length coding of prediction errors using code book of picture I		0.33	0.47	0.30
	using code book of picture II		0.38	0.42	0.44
	using code book of picture III		0.45	0.65	0.22
	Ordering Algorithms				
II.	Run-length coding of ordered elements 'X' with respect to elements 'Y'				
	(a) X: Samples of present line; Y: prediction of present sample		0.48	0.63	0.62
	(b) X: Prediction errors of present line; Y: samples of previous line		0.37	0.44	0.29
	(c) X: Prediction errors of present line; Y: good-bad states	Goodness threshold: 0.02		0.29	0.39	0.22
		0.05	0.26	0.30	0.20
		0.1	0.29	0.30	0.22
		0.2	0.31	0.40	0.22
	(d) X: Prediction errors of present line; Y: good-bad states; code book of picture I	Goodness threshold: 0.05	0.26	0.40	0.23
	Algorithms of Judice				
III.	One-dimensional bit interleaving		0.37	0.50	0.27
IV.	Two-dimensional bit interleaving		0.33	0.48	0.22
V.	Pattern matching		0.40	0.61	0.28

"goodness" of a state is determined by a threshold on the probability of error. Our simulations indicate that a threshold of 5-percent probability of error does better than thresholds of 2 percent, 10 percent, and 20 percent for all three pictures. The entropy is reduced to between 0.20 to 0.30 bit/pel. The advantage of ordering, obtained by comparing these entropies with those obtained from run-length coding of the prediction errors, is about 9 to 29 percent. We also considered the use of a prediction code book from Fig. 2 for all pictures. This resulted in a small increase in entropy over that obtained by using a matched code book.

2.4 Comparisons with the algorithms of Judice

We mentioned earlier that Judice has recently given two algorithms for coding of dithered two-level pictures. In one of them,[12] runs of picture elements corresponding to the same values of the dither matrix are run-length coded. He has discussed this scheme in one dimension (called one-dimensional bit interleaving) as well as two dimensions (called two-dimensional bit interleaving). The bit rates obtained by these two schemes are reproduced from Ref. 12 in Table II. A bit rate of 0.22 to 0.48 bit/pel is possible with these schemes. This is about 10- to 60-percent higher than the entropies obtainable from our "good-bad" state ordering schemes. The other scheme discussed by Judice et al.,[13] called pattern matching, assigns a code to two-dimensional bit patterns of the dithered images. The entropies obtained in this case, also reproduced in Table II from Ref. 13, are generally higher than those achieved by two-dimensional bit interleaving. Thus, our "good-bad" state ordering schemes perform more efficiently than the schemes proposed by Judice.

III. DISCUSSION AND SUMMARY

We have described schemes for efficient coding of dithered two-level signals. We started with the description of a predictor that depends upon already transmitted neighboring elements and the value of the dither matrix at the element being predicted. The predictor minimizes the probability of prediction being in error. We found that the run-length coding of the prediction errors brought the bit rate down to about 0.22 to 0.42 bit/pel for the three pictures we used for simulation. We then discussed several ordering algorithms in which the relative order of transmission of the picture elements in a scan line is changed to increase the average lengths of black and white runs. We found that the ordering scheme based on goodness of the state decreased the bit rate to 0.20 to 0.30 bit/pel. Finally, we compared our results with those obtained by Judice and found that his schemes gave about 10 to 60 percent higher entropy.

It should be mentioned that this is not a definitive coding system study. We have not considered many important factors crucial to the success of any coding system, such as run-length codes and their picture dependence, and the effect of transmission errors.

IV. ACKNOWLEDGMENTS

We would like to thank C. N. Judice of Bell Laboratories for many helpful discussions and for supplying us the picture data used in this paper.

REFERENCES

1. P. Stucki, "Generation of Grey Tones by Computer for Simulation of Visual Information Systems," IEEE Trans. Comput., *18* (July 1969), pp. 642–643.
2. J. O. Limb, "Design of Dither Waveforms for Quantized Visual Signals," B.S.T.J., *48*, No. 7 (September 1969), pp. 2555–2582.
3. B. Lippel and M. Kurland, "The Effect of Dither on Luminance Quantization of Pictures," IEEE Trans. Commun. Technol., *COM-19*, No. 6 (December 1971), pp. 879–888.
4. K. Knowlton and L. Harmon, "Computer-Produced Grey Scales," Computer Graphics and Image Processing, *1*, 1972, pp. 1–20.
5. B. E. Bayer, "An Optimum Method for Two-Level Rendition of Continuous-Tone Pictures," IEEE International Conference on Communications, Conference Record, 1973, pp. (26-11)–(26-15).
6. S. Arneman and M. Tasto, "Generating Halftone Pictures on Graphic Terminals Using Run-Length Coding," Computer Graphics and Image Processing, *2*, 1972, pp. 1–11.
7. J. F. Jarvis, C. N. Judice, and W. H. Ninke, "A Survey of Techniques for the Display of Continuous Tone Pictures on Bilevel Displays," Computer Graphics and Image Processing, *5*, No. 1 (March 1976), pp. 13–40.
8. J. F. Jarvis and C. S. Roberts, "A New Technique for the Bilevel Rendition of Continuous Tone Pictures," IEEE Trans. Commun., *COM-24* (August 1976), pp. 891–898.
9. C. N. Judice, J. F. Jarvis, and W. H. Ninke, "Using Ordered Dither to Display Continuous Tone Pictures on an AC Plasma Panel," Proc. Soc. Inform. Display, *15/4* (Fourth Quarter 1974), pp. 161–164.
10. C. N. Judice, "Dithervision—A Collection of Techniques for Displaying Continuous Tone Still-and-Animated Pictures on Bilevel Displays," International Symposium, Society for Information Display, 1975, pp. 32–33.
11. J. Capon, "A Probabilistic Model for Run-Length Coding of Pictures," IRE Trans. Inform. Theory, *IT-5* (December 1959), pp. 157–163.
12. C. N. Judice, "Data Reduction of Dither Coded Images by Bit Interleaving," Proc. Soc. Inform. Display, *17*, No. 2 (1976), pp. 91–99.
13. C. N. Judice, A. B. White, and R. L. Johnson, "Tranmission and Storage of a Dither Coded Image Using 2-D Pattern Matching," Proc. Soc. Inform. Display, *17*, No. 2 (1976), pp. 85–90.
14. J. S. Wholey, "The Coding of Pictorial Data," IRE Trans. Inform. Theory, *IT-7* (April 1961), pp. 99–104.
15. A. N. Netravali, F. W. Mounts, and E. G. Bowen, "Ordering Techniques for Coding of Two-Tone Facsimile Pictures," B.S.T.J., *55*, No. 9 (December, 1976), pp. 1539–1552.

A Probability Inequality and Its Application to Switching Networks

By F. K. HWANG and A. M. ODLYZKO

(Manuscript received July 26, 1976)

The use of channel graphs to study the blocking probabilities of multistage switching networks was first proposed by Lee and has gained popularity ever since. A channel graph between an input terminal and an output terminal is the union of all paths connecting them in the network. Usually, the assumption is made that links connecting the same two stages, say the ith stage and the (i + 1)st stage, have constant and identical probability p_i of being busy. Let $G(s,\lambda)$ denote the class of channel graphs with s stages and λ paths. We show that for every channel graph in $G(s,\lambda)$ with multiple links, there exists a channel graph in the same class without multiple links that has smaller or equal blocking probabilities for all $\{p_i\}$. We obtain this result by first proving a probability inequality of a more general nature.

I. INTRODUCTION

In this paper we consider a switching network as a directed graph. A vertex is called a *switch* if its in-degree and out-degree are both positive, an *input terminal* if it has in-degree zero and out-degree one, and an *output terminal* if it has in-degree one and out-degree zero. The edges between the switches are called *links*. A switch is said to be of size $n \times m$ if it has in-degree n and out-degree m. Every switch in our network is assumed to be *two-sided nonblocking* in the sense that when the network is in actual use, traffic can be routed from every input link to every output link in a switch, provided the two links involved are not carrying other traffic, and regardless of the traffic carried by other links.

In a multistage switching network, the switches are partitioned into a sequence of stages with the following properties.

(*i*) The sizes of switches in a given stage are identical.

(*ii*) All input terminals are connected to the switches of the first stage; all output terminals are connected to the switches of the last stage.

(*iii*) Links exist only between two switches in adjacent stages. [We call links between the ith stage and the $(i + 1)$st stage the ith-stage links.] The direction of an ith-stage link is from the ith stage to the $(i + 1)$st stage.

A *channel graph* between a given input terminal and a given output terminal is the smallest subgraph containing all paths connecting the two terminals. Since a link in a path is also shared by other paths connecting possibly other pairs of terminals, the actual routing of a path will fail if any link involved has already been used to route some other path. In that case, we say that the path is blocked. The blocking probability of a channel graph is the probability that every path in it is blocked.

Lee[5] first suggested the use of channel graphs to study the blocking performances of switching networks. Usually, the assumption that each ith-stage link has the constant and independent probability p_i of being *busy* (meaning the link is used in routing some other path) is made to simplify the computations of blocking probabilities. Lee's method has gained popularity both in theory and in practice since its proposal.

A class of multistage switching networks that has been widely used but only recently has come under systematic study is the class of *balanced networks*[4]. Balanced networks are characterized by the property that the channel graphs for all pairs of input terminals and output terminals are isomorphic. Thus, the blocking performance of a balanced network can be studied by analyzing just one channel graph.

Let $G(s, \lambda)$ denote the class of channel graphs with s stages and λ paths. Comparisons of channel graphs in a given $G(s, \lambda)$ have been made in Refs. 1 and 3. This paper is a continuation of this study. We are particularly interested in channel graphs with multiple links. Networks with multiple links between a pair of switches have recently been studied by Fontenot.[2] In this paper, we show that for every such channel graph, there is a channel graph in the same class but without multiple links with an equal or smaller blocking probability for any arbitrarily given set $\{p_i\}$. In some cases, a switching network constructed using such a channel graph has a larger number of crosspoints than the corresponding multiple-link network. In other cases, however, our construction produces a network with the same number of crosspoints (and therefore cost), but lower blocking probability. This is illustrated by a simple example at the end of our paper.

II. A PROBABILITY INEQUALITY

We prove a probability inequality which is itself of some interest and has application to our study of channel graphs.

Theorem 1:

$$\prod_{i=1}^{k} (1 - p_i^{c_i}) \leqq 1 - \left\{1 - \prod_{i=1}^{k} (1 - p_i)\right\}^{\prod_{i=1}^{k} c_i},$$

where p_i and c_i are real numbers satisfying $1 \geqq p_i \geqq 0$ and $c_i \geqq 1$ for i = 1, . . . ,k.

Proof: Proof is by induction on k. Theorem 1 is trivially true for $k = 1$. For general k, assume Theorem 1 is true for all $k' = 1, \cdots, k - 1$.

Let $b = \prod_{i=1}^{k-1} c_i$, $y = \prod_{i=1}^{k-1} (1 - p_i)$, $p_k = p$ and $c_k = c$. Then $b \geqq 1$ and $1 \geqq y \geqq 0$. By induction,

$$\prod_{i=1}^{k} (1 - p_i^{c_i}) \leqq (1 - p^c) - (1 - p^c)(1 - y)^b. \tag{1}$$

It is sufficient to prove that

$$(1 - p^c) - (1 - p^c)(1 - y)^b \leqq 1 - \{1 - (1 - p)y\}^{bc},$$

or equivalently,

$$\{1 - (1 - p)y\}^{bc} \leqq p^c + (1 - p^c)(1 - y)^b. \tag{2}$$

Let $z = p^c$. Then, $1 \geqq z \geqq 0$. Ineq. (2) can be written as

$$\{1 - (1 - z^{1/c})y\}^{bc} \leqq z + (1 - z)(1 - y)^b. \tag{3}$$

We first show that

$$f(z,y) = \{1 - (1 - z^{1/c})y\}^c \leqq 1 - (1 - z)y = g(z,y). \tag{4}$$

Clearly $f(1,y) = g(1,y)$. Furthermore,

$$\frac{\partial}{\partial z} f(z,y) = c\{1 - (1 - z^{1/c})y\}^{c-1} y \frac{1}{c} z^{1/c-1}$$

$$\geqq y = \frac{\partial}{\partial z} g(z,y),$$

since

$$\{1 - (1 - z^{1/c})y\}^{c-1} z^{1/c-1} \geqq \{1 - (1 - z^{1/c})\}^{c-1} z^{1/c-1} 1 = 1.$$

Therefore, Ineq. (4) is true. Consequently, to prove Ineq. (3), it suffices to prove

$$h(z,y) = g(z,y)^b = \{1 - (1 - z)y\}^b$$

$$\leqq z + (1 - z)(1 - y)^b = u(z,y). \tag{5}$$

We have $h(z,0) = u(z,0)$. Furthermore

$$\frac{\partial}{\partial y} h(z,y) = -b\{1 - (1 - z)y\}^{b-1}(1 - z)$$

$$\leqq -(1 - z)b(1 - y)^{b-1} = \frac{\partial}{\partial y} u(z,y),$$

since

$$1 \geqq 1 - z \geqq 0$$

and

$$1 - (1 - z)y \geqq 1 - y.$$

Therefore, Ineq. (5) is true. The proof is completed.

It is easy to construct counter-examples of Theorem 1 if the conditions $c_i \geqq 1$ for $i = 1, \ldots, k$ are violated.

III. A THEOREM ON CHANNEL GRAPHS

Consider a channel graph which contains the subgraph of Fig. 1,
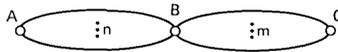
Fig. 1—Graph with multiple links.

where A is an ith-stage switch, B an $(i + 1)$st-stage switch, C an $(i + 2)$nd-stage switch and $\text{Max}\{m,n\} > 1$. Since B is a two-sided nonblocking switch, there are nm paths from A to C. Let p_i, $i = 1, \cdots, s$ be the probability that an ith-stage link is busy. We show that if we replace Fig. 1 with the subgraph of Fig. 2,
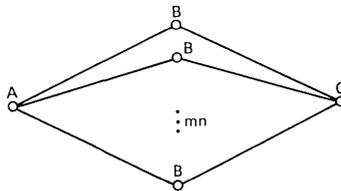
Fig. 2—Graph without multiple links.

then the new channel graph, clearly in the same class $G(s,\lambda)$, will have an equal or smaller blocking probability. It suffices to show that the blocking probability of the graph in Fig. 1 is equal or larger than that of the graph in Fig. 2. Routing from A to C can be realized in Fig. 1 if at least one link from each of the n and m links is available. The probability of this event is

$$(1 - p_i^n)(1 - p_{i+1}^m).$$

The same routing can be realized in Fig. 2 if at least one of the nm two-link paths (see Fig. 3)
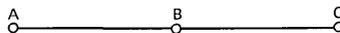

Fig. 3—Two-link path.

is nonblocking. The probability of this event is

$$1 - \{1 - (1 - p_i)(1 - p_{i+1})\}^{nm}.$$

That the first probability is equal or smaller than the second probability is an immediate consequence of the method of Theorem 1 by setting $k = 2$. Therefore we have proved the following theorem.

Theorem 2: For every channel graph with multiple links, there is a channel graph in the same $G(s, \lambda)$ class without multiple links that has equal or smaller blocking probability.

Example. Let us compare the blocking probabilities of the two five-stage balanced networks in Fig. 4 and Fig. 5.
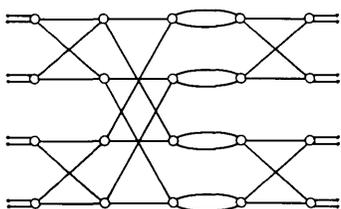
Fig. 4—Network with multiple links.
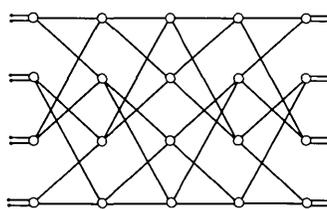
Fig. 5—Network without multiple links.

The two channel graphs are shown in Fig. 6 and Fig. 7, respectively.
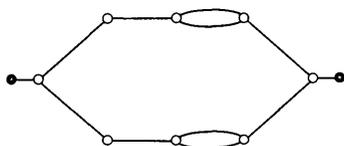
Fig. 6—Channel graph of network with multiple links.
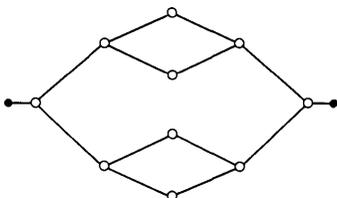
Fig. 7—Channel graph of network without multiple links.

By Theorem 1, the blocking probability of the channel graph in Fig. 6 is equal to or greater than that of the channel graph in Fig. 7. Therefore, we conclude that the network in Fig. 5 has smaller blocking probabilities than the one in Fig. 4. Note that the two networks are identical except for the way the switches are linked.

REFERENCES

1. F. R. K. Chung and F. K. Hwang, "A Problem on Blocking Probabilities in Connecting Networks," unpublished work.
2. M. L. Fontenot, "Optimal Three-Stage Switching Networks," unpublished work.
3. F. K. Hwang, "Link Designs and Probability Analyses for a Class of Connecting Networks," unpublished work.
4. F. K. Hwang, "Balanced Networks," Conf. Record, XII IEEE Int. Conf. on Commun., 1976, Philadelphia, Pennsylvania, *1*, pp. 7-13–7-16.
5. C. Y. Lee, "Analysis of Switching Networks," B.S.T.J. *34* (November 1955), pp. 1287–1315.

Contributors to This Issue

Jean-David Beyer, B.A., 1961, University of Buffalo; Sierra Research Corporation, 1960–1965; Bell Laboratories, 1965—. Mr. Beyer has worked on computer and operating system design for interactive picture processing systems. He has also been concerned with compiler design for simulating sampled data systems. He is currently working on design of an interactive data base management system. Member of ACM.

J. S. Cook, B. E. E., M.S. (Electrical Engineering), 1952, The Ohio State University. Bell Laboratories, 1952—. Mr. Cook has done research in the fields of traveling-wave tubes, microwave propagation and devices, antennas, and satellite communications. He has been working with optical fiber communication systems in recent years and currently heads a department responsible for development of optical fiber connectors and the special technology of optical fiber telecommunication systems. Senior member, IEEE, member, OSA, SPIE, Eta Kappa Nu, Tau Beta Pi, Sigma Xi.

Ronald E. Crochiere, B.S. (E.E.), 1967, Milwaukee School of Engineering; M.S. (E.E.), 1968, and Ph.D. (E.E.), 1974, Massachusetts Institute of Technology; Raytheon, 1968–1970; M.I.T. Research Laboratory of Electronics, 1970–1974; Bell Laboratories, 1974–. Mr. Crochiere has worked on the design of microwave phase shifters, on digital network theory, and digital filter structures. He is presently engaged in research in speech communications and digital signal processing. Associate editor, IEEE G-ASSP Transactions. Member, Sigma Xi and IEEE G-ASSP Technical Committee on Digital Signal Processing.

Frank K. Hwang, B.A., 1960, National Taiwan University; M.B.A., City University of New York; Ph.D. (Statistics), 1968, North Carolina State University; Bell Laboratories, 1967—. Mr. Hwang spent the fall of 1970 visiting the Department of Mathematics of National Tsing-Hua University. He has been engaged in research in statistics, computing science, discrete mathematics, and switching networks.

Nuggehally S. Jayant, B.Sc. (Physics and Mathematics), 1962, Mysore University; B.E., 1965, and Ph.D., 1970 (Electrical Communication Engineering), Indian Institute of Science, Bangalore; Research Associate at Stanford University, 1967–1968; Bell Laboratories, 1968—. Mr. Jayant was a Visiting Scientist at the Indian Institute of Science from January–March 1972 and August–October 1975. He has worked on coding for burst error channels, detection of fading signals, statistical pattern discrimination, spectral analysis, and problems in adaptive quantization and prediction, with special reference to speech signals. He is also editor of an IEEE Reprint Book on *Waveform Quantization and Coding*.

Subhash C. Kak, B.E., 1967, Kashmir University; Ph.D. (E.E.), 1970, Indian Institute of Technology, Delhi. Mr. Kak has been on the faculty of the Indian Institute of Technology since 1971. He spent 1975–1976 as Academic Visitor at Imperial College, London, and Summer 1976 at Bell Laboratories. Mr. Kak has worked on information and system theories and more recently on their applications to speech and pictures. He was the convenor of the First National Systems Conference held at New Delhi in November 1974.

Sing-Hsiung Lin, B.S.E.E., 1963, National Taiwan University; M.S.E.E., 1966, and Ph.D., 1969, University of California, Berkeley; Bell Laboratories, 1969—. At the Electronics Research Laboratory, University of California at Berkeley, Mr. Lin was engaged in research on antennas in plasma media and numerical solutions of antenna problems. Mr. Lin is presently working on wave propagation problems on terrestrial radio systems and earth-satellite radio systems. Member, IEEE, Sigma Xi, AIAA.

Dietrich Marcuse, Diplom Vorpruefung, 1952, Dipl. Phys., 1954, Berlin Free University; D.E.E., 1962, Technische Hochschule, Karlsruhe, Germany; Siemens and Halske (Germany), 1954–1957; Bell Laboratories, 1957–. At Siemens and Halske, Mr. Marcuse was engaged in transmission research and studying coaxial cable and circular waveguide transmission. At Bell Laboratories, he has been engaged in studies of circular electric waveguides and work on gaseous masers. He spent one year (1966–1967) on leave of absence from Bell Laboratories at the University of Utah. He is presently working on the transmission aspect of a light communications system. Mr. Marcuse is the author of three books. Fellow, IEEE; member, Optical Society of America.

Thomas O. Mottl, B.S., 1959, Purdue University; M.S., 1960, Ph.D., 1968, The University of Michigan; Bell Laboratories, 1968–1976; The Analytic Sciences Corp. 1976-. At Bell Laboratories, Mr. Mottl worked on underwater acoustics, crosstalk coupling characteristics of multipair cables, and microwave radio propagation. His work on multipath outage modeling was carried out at the Merrimack Valley Laboratory in Massachusetts.

F.W. Mounts, E.E., 1953, M.S., 1956, University of Cincinnati; Bell Laboratories, 1956—. Mr. Mounts has been concerned with research in efficient methods of encoding pictorial information for digital television systems. Member, Eta Kappa Nu; Senior Member, IEEE.

Arun N. Netravali, B. Tech. (Honors), 1967, Indian Institute of Technology, Bombay, India; M.S., 1969, and Ph.D. (E.E.), 1970, Rice University; Optimal Data Corporation, 1970–1972; Bell Laboratories, 1972—. Mr. Netravali has worked on problems related to filtering, guidance, and control for the space shuttle. At Bell Laboratories, he has worked on various aspects of signal processing. Member, Tau Beta Pi, Sigma Xi.

A. M. Odlyzko, B.S. and M.S. (Mathematics), 1971, California Institute of Technology; Ph.D. (Mathematics), 1975, Massachusetts Institute of Technology; Bell Laboratories, 1975—. Mr. Odlyzko works on coding theory, combinatorics, and related areas.

Kinichiro Ogawa, B.S.E.E. (1966), M.S.E.E. (1968), University of Tokyo; D.Sc. (Electrical Engineering), Washington University; Nippon Telegraph & Telephone Public Corporation, 1968–1976; Bell Laboratories, 1976—. At N.T.T., Mr. Ogawa worked on the long-haul analogue coaxial cable system and the video transmission system. He also worked on the development of the PCM-FDM digital system. He was responsible for the evaluation of the C.A.I. system and contributed to the video system development group. At Bell Laboratories, he has been working on the application of fiber optics in transmission.

Marvin R. Sambur, B.E.E. (1968), City College of New York; S.M. (1969) and Ph.D. (1972), Massachusetts Institute of Technology; Bell Laboratories, 1972-. Mr. Sambur is a member of the Acoustics Research Department and has worked in the areas of speech recognition, speaker recognition, low-bit-rate vocoder systems, encryption techniques, and digital waveform coders. Member, MPA-TC subcommittee on Speech Recognition and Understanding, Eta Kappa Nu, Tau Beta Pi, Sigma Xi.

Abstracts of Papers by Bell System Authors Published in Other Journals

CHEMISTRY

Block Copolymer Theory—IV. Narrow Interphase Approximation. E. Helfand and Z. R. Wasserman, Macromolecules, *9* (November–December 1976), pp. 879–888. A theory of microdomain structure in block copolymers, developed earlier, is simplified by approximations appropriate when the interphase width is narrow compared to the domain size. The free energy, written as an algebraic function, is minimized to obtain predictions of domain sizes. These are compared with experiments.

Gas Phase EPR Linewidths and Intermolecular Potentials V. Quantal Derivation. G. J. Fisanick-Englot and T. A. Miller, J. Chem. Phys., *66* (February 1, 1977), pp. 1175–1182. A fully quantal derivation of the isolated line-broadening parameters applicable to gas phase EPR experiments is presented. Errors inherent in the use of zero field operators are discussed.

HCN from the Reduction of NO over Platinum, Palladium, Ruthenium, Monel and Perovskite Catalysts. R. J. H. Voorhoeve, C. K. N. Patel, L. E. Trimble, R. J. Kerl, and P. K. Gallagher, J. Catal., *45* (1976), pp. 297–304. HCN was produced in mixtures of NO, CO, and H_2 at temperatures from 400–800°C. Most active in HCN production was a supported Pt catalyst, followed by Pd, Cu-Ni and Ru, in that order. Perovskite $La_{0.8}K_{0.2}MnO_3$ yields little HCN, but over $La_{0.8}K_{0.2}Mn_{0.94}Ru_{0.06}O_3$ the yield is higher than over either Ru or the matrix perovskite.

Oxidative Stability of Expanded Polyethylene for Wire and Cable. F. R. Wight, J. Cell. Plast. *12*, No. 6 (November/December 1976), pp. 317–319. Decomposition of azodicarbonamide blowing agent in high density polyethylene has been reported to have an adverse effect on the oxidative stability of the material. The decomposition process itself has been implicated as being responsible rather than cell structure or residues. Data presented here clearly show that at 195°C and only in the presence of copper, blowing agent residues reduce the oxidative stability of high density polyethylene. Furthermore, data produced at 120°C show that neither the decomposition process nor the residues have any effect on oxidative stability at this temperature. Rather, it is argued that any decrease in oxidative stability at 120°C or lower is due to the cell structure which alters the diffusional properties of the finished insulation.

Singlet-Triplet Anticrossings in ^{4}He. III. Separation and Mixing of the n = 3 − 8 ^{1}D and ^{3}D States. J. Derouard, R. Jost, M. Lombardi, T. A. Miller, and R. S. Freund, Phys. Rev. A, *14* (1976), pp. 1025–1035. Measurements of both the zero-field singlet-triplet separation and the antisymmetric part of the spin-orbit coupling between singlet and triplet D states of He with n = 3 − 8 are presented.

Theory of the Concentrated Polymer Solution/Solvent Interface. E. Helfand A. M. Sapse,* J. Polym. Sci. C, Flory Symposium, *54* (1976), pp. 289–297. A theory is presented of the interface between a solvent and a saturated polymer solution. Random-walk statistics are assumed for the macromolecules. Results are presented for a general and a Flory-Huggins form of the free energy. There is good correspondence with a recently developed lattice theory of these systems. *City University of New York.

Tropospheric Halocarbons: Estimates of Atmospheric Chemical Production. T. E. Graedel and D. L. Allara, Atmos. Environ., *10* (1976), pp. 385–388. Selected thermal and photochemical atmospheric reactions have been evaluated as potential sources for the family of halocarbons recently detected in tropospheric air. Formation of CH_3Cl is extremely slow and that of CCl_4, $CHCl_3$, CH_3I, CH_3CCl_3 and the chlorinated ethylenes

is negligible, implying that direct emission is responsible for the presence of these compounds.

ELECTRONIC AND ELECTRICAL ENGINEERING

Frequency-Agile Millimeter-Wave Phase Lock System. P. S. Henry, Rev. Sci. Instrum. *47* (September 1976), pp. 1020–1025.   A frequency-agile phase lock system for millimeter-wave klystrons is described. The system locks the klystron to a crystal-controlled reference signal derived from a frequency synthesizer. By programming the synthesizer, the klystron can be stepped through any sequence of frequencies lying within a band roughly 200 MHz wide.

Optical-Fiber Impulse-Response Measurement System. J. W. Dannwolf, S. Gottfried, G. A. Sargent, R. C. Strum, IEEE Trans. Instrum. Meas.,*IM-25*, V. 4 (December 1976), pp. 401–406.   This paper describes time-domain instrumentation designed to measure impulse response and delay of multimode optical fibers used in an experimental optical communications system at Bell Laboratories. Time-domain data is transformed to frequency-domain by a minicomputer, and the result is displayed as the fiber's baseband frequency response.

Tantalum Thin Film RC Circuit Technology for a Universal Active Filter. W. Worobey and J. Rutkiewicz, IEEE Trans. Parts, Hybrids and Packag., *PHP-12*, No. 4 (December 1976), pp. 276–282.   This paper describes the physical layout, process sequence, and component properties of an RC universal active filter. The high-precision filter is fabricated on a 16-pin dual in-line package ceramic substrate using tantalum thin film technology. It is comprised of 300 $\Omega/\square$ resistors, 190 V anodized tantalum capacitors, and an operational amplifier.

The Use of Echo Time-Weighting to Derive Oscilloscope Graticules for Rating Television Transmission Performance. R. W. Edmonds, SMPTE J., *85*, No. 6 (June 1976), pp. 393–396.   This paper describes a method for designing oscilloscope graticules for measuring the short-time waveform performance of broadcast television systems. The design is based on recently obtained single-echo time-weighting functions for monochrome and color signals.

Using Triangularly Weighted Interpolation to Get 13-Bit PCM from a Sigma-Delta Modulator. J. C. Candy, Y. C. Ching, and D. S. Alexander, IEEE Trans. Commun., *24*, No. 11 (November 1976), pp. 1268–1275.   Accumulating a weighted sum of sigma-delta codes generates a high-resolution PCM signal. Several weighting methods are evaluated with regard to resolution and spectral response; a triangular weighing is near optimum. Implementation of a 13-bit PCM encoder is described and a method for overcoming a threshold phenomenon is presented.

MATERIALS SCIENCE

Electrical, Structural and Optical Properties of Amorphous Carbon. J. J. Hauser, J. Noncrystal. Solids, *23* (January 1977), pp. 21–41.   The planar and transverse electrical resistivity of amorphous carbon (a-C) films is well fitted by the expression $\rho = \rho_0 \exp (T_0/T)^{1/4}$. Films thinner than 600 Å display a two-dimensional hopping conductivity from which one deduces a density of states $N(E_F)$ at the Fermi level of 10^{18} eV^{-1}cm^{-3} and a radius of the localized wave functions (a) of 12 Å.

Epitaxial Structures with Alternate-Atomic-Layer Composition Modulation. A. C. Gossard, P. M. Petroff, W. Weigmann, R. Dingle, and A. Savage, Appl. Phys.Lett.,*29*, No. 6 (15 September 1976), pp. 323–325.   Epitaxial structures grown by alternate monolayer depositions of GaAs and AlAs are reported. As many as 10^4 alternate (100) layers of GaAs and AlAs as thin as 1.0 ± 0.1 and 1.0 ± 0.1 monolayers, respectively, were deposited and studied by transmission electron microscopy and optical techniques.

The Optical Properties of a Soda-Lime-Silica Glass in the Region From 0.006 to 22 eV. B. G. Bagley, E. M. Vogel, W. G. French and G. A. Pasteur, J. N. Gan, and J. Tauc,* J. Non-Crystalline Sol, *22* (November/December 1976), pp. 423–436.   From the measured absorption and reflection spectra, we have determined the optical properties of a well-characterized (with respect to impurities and homogeneity) high-purity 21.3 wt% Na_2O—5.2 wt% CaO—73.5 wt% SiO_2 glass over the energy range 0.006—22 eV. The origins of the absorption spectra are discussed.   *Brown University.

PHYSICS

Dynamic Central Peaks in a Crystalline Solid: $KTaO_3$. K. B. Lyons and P. A. Fleury, Phys. Rev. Lett., *37,* (July 19, 1976), pp. 161–164.   We report two central peaks in the quasielastic-light-scattering spectrum of $KTaO_3$. The polarization and angular dependence of the linewidth indicate that the narrow component (2.3 ± 0.3 GHz at 300 K in right-angle scattering) is due to entropy fluctuations. A tentative identification of the broader component with two-phonon processes is made.

Two-Dimensional Cochlea Fluid Model—New Results. J. B. Allen, JASA, *61,* No. 1 (January 1977), pp. 110–119.   With the recent availability of the physical measurements of Rhode [JASA, *49* (1971), pp. 1218] many theoretical models have proven to be incomplete or inadequate. In this paper, solution of the two-dimensional model of Lesser and Berkley [J. Fl. Mech.,*1* (Jan. 1970)] are shown to be in close agreement with the results of Rhode.

Growth and Continuous Compositional Grading of Ga(As,Sb,P) by Liquid Phase Epitaxy. R. E. Nahory, M. A. Pollack, and J. C. DeWinter, J. Appl. Phys., *48,* No. 1 (January 1977), pp. 320–323.   Liquid phase epitaxial growth of Ga(As,Sb,P) is described. Layers have been continuously graded from a composition with the lattice constant of GaAs to a composition suitable for subsequent growth of a GaAsSb double heterostructure laser. The details of the grading are discussed in terms of a phase diagram calculation.

SYSTEMS ENGINEERING AND OPERATIONAL RESEARCH

Application of New Ridge Regression Methods to a Study of Bell System Scale Economies. H. D. Vinod, JASA, *71* (December 1976), pp. 835–841.   We suggest a new horizontal scaling for the "ridge trace," some new techniques for monitoring ridge solutions including an index of stability of relative magnitudes (ISRM) and numerical largeness of more significant (NLMS) regression coefficients. A method of curtailing the tending of ridge regression coefficients toward zero is also presented. Two examples illustrate these methods and estimate Bell System's scale economies to be large.

ATLAS—An Automated Software Testing System. W. H. Jessop, J. R. Kane, S. Roy, and J. M. Scanlon, Proc. Second Int. Conf. Software Eng. (October 12, 1976), pp. 629–635.   ATS formalizes a concept of model-referenced testing for large software systems. Its object is to certify software under test against a directed-graph model. This objective is met by components of ATLAS, which automatically derive tests from the model and then apply the tests and verify the results.

Transmission Planning in International Telephony. F. T. Andrews, Jr., IEEE Trans. Commun.,   June 1976 , pp. 658–660.   A framework for controlling transmission quality of international connections is the joint responsibility of CCITT Study Groups XII, XV, XVI and Special D. These groups recommend rating methods, objectives, network plans, and equipment characteristics. Still closer cooperation may help to avoid unnecessary regional differences in equipment standards.