

AT&T

April 1985 Vol. 64 No. 4

TECHNICAL
JOURNAL

A JOURNAL OF THE AT&T COMPANIES

Microwave Radio

Robotics

Undersea Lightguide Cable

Microphone Arrays

EDITORIAL BOARD

	M. IWAMA, ¹ <i>Board Chairman</i>	
D. H. BOYCE ¹	A. FEINER ³	J. S. NOWAK ¹
M. M. BUCHNER, JR. ¹	R. C. FLETCHER ¹	L. C. SIEFERT ²
R. P. CREAM ²	D. HIRSCH ⁴	W. E. STRICH ⁵
B. R. DARNALL ¹	S. HORING ¹	J. W. TIMKO ³
B. P. DONOHUE, III ³	J. F. MARTIN ²	V. A. VYSSOTSKY ¹

¹AT&T Bell Laboratories ²AT&T Technologies ³AT&T Information Systems

⁴AT&T Consumer Products ⁵AT&T Communications

EDITORIAL STAFF

B. G. KING, <i>Editor</i>	L. S. GOLLER, <i>Assistant Editor</i>
P. WHEELER, <i>Managing Editor</i>	A. M. SHARTS, <i>Assistant Editor</i>
B. VORCHHEIMER, <i>Circulation</i>	

AT&T TECHNICAL JOURNAL (ISSN 8756-2324) is published ten times each year by AT&T, 550 Madison Avenue, New York, NY 10022; C. L. Brown, Chairman of the Board; T. O. Davis, Secretary. The Computing Science and Systems section and the special issues are included as they become available. Subscriptions: United States—1 year \$35; foreign—1 year \$45.

Back issues of the special, single-subject supplements may be obtained by writing to the AT&T Customer Information Center, P.O. Box 19901, Indianapolis, Indiana 46219, or by calling (800) 432-6600. Back issues of the general, multisubject issues may be obtained from University Microfilms, 300 N. Zeeb Road, Ann Arbor, Michigan 48106.

Payment for foreign subscriptions or single copies must be made in United States funds, or by check drawn on a United States bank and made payable to the AT&T Technical Journal and sent to AT&T Bell Laboratories, Circulation Dept., Room 1E335, 101 J. F. Kennedy Pky, Short Hills, NJ 07078.

Single copies of material from this issue of the Journal may be reproduced for personal, noncommercial use. Permission to make multiple copies must be obtained from the Editor.

Printed in U.S.A. Second-class postage paid at Short Hills, NJ 07078 and additional mailing offices. Postmaster: Send address changes to the AT&T Technical Journal, Room 1E335, 101 J. F. Kennedy Pky, Short Hills, NJ 07078.

Copyright © 1985 AT&T.

AT&T TECHNICAL JOURNAL

VOL. 64

APRIL 1985

NO. 4

Copyright© 1985 AT&T, Printed in U.S.A.

A New Approach to Space Diversity Combining in Microwave Digital Radio	885
Y. S. Yeh and L. J. Greenstein	
A Simulation Study of Space Diversity and Adaptive Equalization in Microwave Digital Radio	907
L. J. Greenstein and Y. S. Yeh	
A Controlled Impedance Robot Gripper	937
M. K. Brown	
Determination of Fiber Proof-Test Stress for Undersea Lightguide Cable	971
T. C. Chu and H. C. Chandan	
Beamwidth and Useable Bandwidth of Delay-Steered Microphone Arrays	983
J. L. Flanagan	
PAPERS BY AT&T BELL LABORATORIES AUTHORS	997
CONTENTS, MAY-JUNE ISSUE	1003

A New Approach to Space Diversity Combining in Microwave Digital Radio

By Y. S. YEH and L. J. GREENSTEIN*

(Manuscript received May 30, 1984)

In this paper we describe a new approach to dual-channel space diversity combining in microwave digital radio. This approach features (1) adaptive control of the relative amplitudes and phases of the two branch gains; and (2) a search strategy, based on noncoherent spectrum measurements at the combiner output, that simultaneously accounts for both dispersion and noise. Computer programs have been developed to simulate the search process and to analyze the resulting performance. Eight representative channel response pairs are postulated and performance results are presented for each. They show that the scheme provides a high degree of channel equalization over bandwidths up to at least 40 MHz, and that, in receivers not using adaptive equalizers, it offers major improvements in detection performance over selection diversity.

I. INTRODUCTION

In a terrestrial digital radio link, frequency selective fading caused by multipath propagation presents the major threat to system availability. Efforts to reduce channel dispersion, and thus to increase availability, typically center on the use of adaptive equalization and/or dual-branch space diversity.¹⁻¹¹

Most conventional space diversity schemes use either selection switching or so-called "in-phase" combining of the diversity branches. The latter approach concentrates on maximizing the combiner output power rather than on minimizing channel dispersion. Recent work,

* Authors are employees of AT&T Bell Laboratories.

Copyright © 1985 AT&T. Photo reproduction for noncommercial use is permitted without payment of royalty provided that each reproduction is done without alteration and that the Journal reference and copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free by computer-based and other information-service systems without further permission. Permission to reproduce or republish any other portion of this paper must be obtained from the Editor.

however, has dealt with “out-of-phase” combining, which reduces output dispersion by suitably adjusting the relative phases between the two branches.^{4,8} This approach can completely eliminate dispersion for certain two-path propagation situations, but not under more general and realistic conditions.

In this paper, we describe a combiner in which the relative phases *and* amplitudes of the two branches are controlled. With little increase in complexity, this approach allows the effects of both dispersion and noise to be jointly minimized. We shall consider this type of combining within the context of M-level Quadrature Amplitude Modulation (M-QAM) systems.

Two specific approaches for finding the best amplitude and phase adjustments are described in Section II. One in particular, based on noncoherent spectrum measurements at the combiner output, is identified for further study. The simulation and analysis of this scheme are discussed in Section III, and its performance for several postulated dual-channel response pairs is assessed in Section IV. Comparisons with other diversity and nondiversity approaches are also given and attest to the effectiveness of the new scheme.

II. THE DIVERSITY COMBINER

2.1 Rationale

The idea of using both amplitude and phase adjustments in space diversity combining, while not entirely new, has yet to be fully understood and optimally exploited. We now illustrate the potential power of this form of combining under quite general circumstances. To do so, we will invoke some recent work on the modeling and analysis of multipath fading responses.

Let $H_1(f)$ and $H_2(f)$ be the complex frequency responses of a fading channel as viewed by two vertically spaced receiver antennas. Under nonfading conditions, these functions are flat with frequency at unity amplitude ($|H_1(f)| = |H_2(f)| = 1$). In all that follows, we measure f from the center of the radio channel, so that $H_1(f)$ and $H_2(f)$ are baseband functions. Moreover, we are interested in their variations over just the interval $[-W/2, W/2]$, where W is the channel bandwidth in hertz. Since multipath fading arises from a finite number of discrete propagation paths, we can present $H_1(f)$ and $H_2(f)$ in the following general forms:

$$H_1(f) = \sum_{k=1}^{K_1} R_{1k} \exp\{-j(\omega\tau_{1k} - \theta_{1k})\} \quad (1)$$

and

$$H_2(f) = \sum_{k=1}^{K_2} R_{2k} \exp\{-j(\omega\tau_{2k} - \theta_{2k})\}. \quad (2)$$

In (1), K_1 is the number of paths and τ_{1k} , R_{1k} , and θ_{1k} are the time delay, amplitude, and phase, respectively, associated with the k th path. Similar definitions apply to K_2 , τ_{2k} , R_{2k} , and θ_{2k} in $H_2(f)$.

Typically, multipath propagation on microwave radio links exhibits delay spreads on the order of 10 ns or less, i.e., the largest and smallest τ values differ by amounts small compared to $1/W$, where $W \leq 40$ MHz in the common carrier bands. This observation inspired earlier efforts to approximate fading channel responses using low-order polynomials in $j\omega$.¹²

To be concrete, let $H_1(f)$ and $H_2(f)$ be represented by the infinite power series

$$H_1(f) = e^{-j\omega t_1}[A_1 + j\omega B_1 + (j\omega)^2 C_1 + \dots] \quad (3)$$

and

$$H_2(f) = e^{-j\omega t_2}[A_2 + j\omega B_2 + (j\omega)^2 C_2 + \dots], \quad (4)$$

where t_1 and t_2 are arbitrary, and the A 's, B 's, and so on are complex coefficients. Using the power series expansion for e^{jx} , they can be easily related to the parameters of the functions (1) and (2), e.g.,

$$A_1 = \sum_{k=1}^{K_1} R_{1k} e^{j\theta_{1k}}; \quad B_1 = \sum_{k=1}^{K_1} R_{1k} e^{j\theta_{1k}} (t_1 - \tau_{1k}), \quad (5)$$

and so on.

The essence of *first-order* polynomial fitting is this: Given $H_1(f)$, a value for t_1 can usually be found such that $(A_1 + j\omega B_1)$ is the dominant part of (3) over $[-W/2, W/2]$, all higher-order terms in $j\omega$ being small, in some sense. Similarly, a value for t_2 can usually be found that does the same for $H_2(f)$, (4). That such first-order polynomial fitting is reasonable to do in common carrier channels has been supported by theory,¹³ noncoherently measured data,¹² and (more recently) coherently measured data.¹⁴

Now suppose that a space diversity combiner were used having an adjustable time delay (τ) and complex gain (β) in the second branch. The composite channel response, as viewed at the combiner output, would then be

$$H(f) = H_1(f) + \beta e^{-j\omega\tau} H_2(f). \quad (6)$$

If τ and β were adaptively adjusted to be

$$\tau = t_1 - t_2 \quad \text{and} \quad \beta = -B_1/B_2, \quad (7)$$

we could then write [see (3) and (4)]

$$H(f) = [A_1 - A_2 B_1/B_2] + \left\{ \begin{array}{l} \text{second- and} \\ \text{higher-order} \\ \text{terms in } j\omega \end{array} \right\}. \quad (8)$$

Thus, by proper choice of delay and gain in one branch, the channel response could be made dispersionless except for small higher-order terms in $j\omega$. This is a quite general result for the channels of interest and shows the power of the combining approach under study. If β were a phase-only factor [$\beta = \exp(j\phi)$], such a strong reduction in dispersion would only be possible in those fortuitous situations where $|B_1| = |B_2|$.

2.2 The combiner structure

The above discussion suggests both a particular structure, (6), for the combiner and particular solutions, (7), for the variable parameters. The discussion was intended, however, to provide insight rather than to identify a serious design approach. For one thing, a variable time delay would be difficult to implement and would offer little incremental benefit in most cases. In addition, the solutions of (7) do not properly address either the residual dispersion (i.e., higher orders in $j\omega$) or the receiver noise. If, for example, A_1/B_1 and A_2/B_2 happened by chance to be close in value, the first term in (8) would be severely weakened in the process of eliminating the $j\omega$ term; this would enhance the effects of both the remaining dispersion and the noise.

Accordingly, we propose a combiner in which the relative amplitudes and phases in the two diversity branches are adaptively adjusted, but not the delays. Thus,

$$H(f) = \beta_1 H_1(f) + \beta_2 H_2(f), \quad (9)$$

as indicated in Fig. 1, where β_1 and β_2 are adapted gains. (Because the dominant thermal noise is introduced before the combiner, several ways of adapting β_1 and β_2 would, in theory, yield equivalent performance. For example, one gain could be held fixed, or adapted in amplitude only, with the other being adapted in both amplitude and phase; or both gains could be adapted in both amplitude and phase. Each of these approaches would permit the adaptation of the *relative* complex branch gains, which is all that matters.) Moreover, we propose the use of control strategies that take proper account of both dispersion and noise.

2.3 Control strategies

We shall discuss two distinct approaches for controlling the gain pair (β_1, β_2) in (9). The first approximates the theoretically best way to do combining when there is no post-combiner equalization. The second approach, thought suboptimal, has features that make it attractive both with and without post-combiner equalization. The new scheme reported here incorporates the second approach, whose performance we will compare with that of the first.

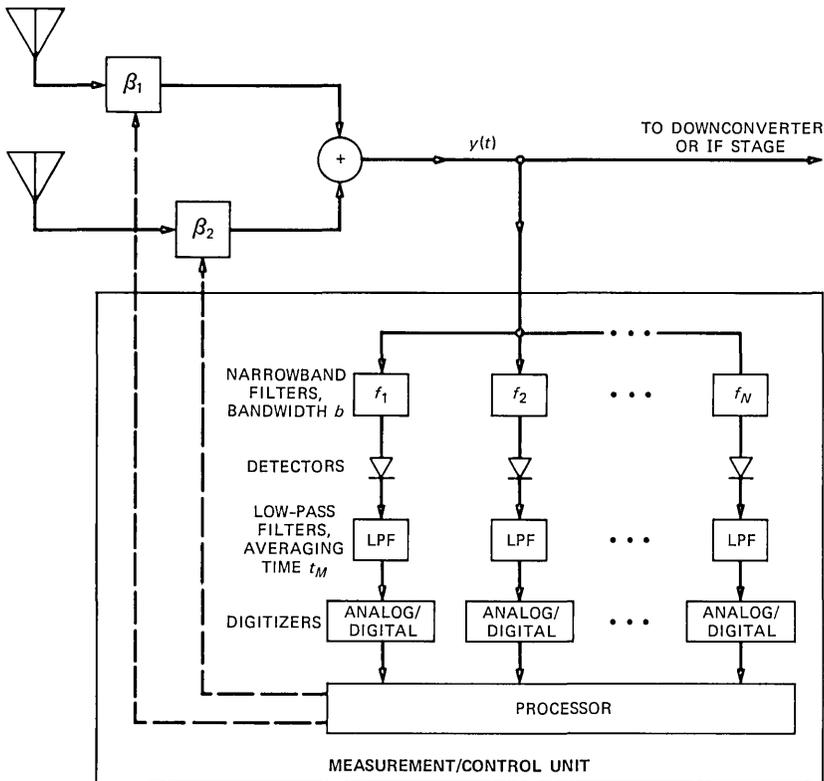


Fig. 1—Block diagram of the space diversity combiner. In this scheme, noncoherent spectral measurements are made on the combiner output at three or more inband frequencies, and the control measure formed from them is used to adjust the complex gains β_1 and β_2 . The combining is shown at Radiofrequency (RF) but could be at Intermediate Frequency (IF) instead.

Approach 1: Assume for now that there is no post-combiner equalization. In that case, the detection bit error rate is virtually minimized by choosing (β_1, β_2) to maximize the ratio of sampled signal to root mean squared (rms) distortion at the receiver output. By “sampled signal,” we mean the half-distance between signal levels, as sampled in every period at the in-phase and quadrature detectors; by “rms distortion,” we mean the rms sum of thermal noise and Intersymbol Interference (ISI) sampled at the detectors. This signal-to-distortion ratio [defined formally by (26) and expressed by (28), below] depends on $H_1(f)$, $H_2(f)$, β_1 , β_2 , and other factors; most important, it is convex in β_1 and β_2 . The result is that there is a unique value of relative complex gain, β_1/β_2 , that maximizes this ratio. We shall regard as *optimal* any gain pair (β_1, β_2) that exhibits this maximizing relative gain.

There is a practical way to realize optimal values for β_1 and β_2 ,

namely, gradient search methods using data decisions.^{15,16} The action is similar to that of an adaptive transversal equalizer, with β_1 and β_2 taking the place of the optimized tap gains. We shall refer to this or any decision-directed scheme that optimizes (β_1, β_2) , in the sense defined above, as Approach 1.

Approach 2: A potential liability of Approach 1 is that it relies on having accurate data decisions, a condition that may not always exist (e.g., during recovery from severe fades). The scheme to be reported here is based on a different strategy, which we designate as Approach 2. It consists of (1) performing certain noncoherent spectral measurements on the combiner output; and (2) sequentially searching over (β_1, β_2) so as to maximize a certain quantity [eq. (18), below] computed from these measurements. By adapting β_1 and β_2 in this way, a close approximation to the “optimal” condition defined above can be achieved, as we show later. That is, Approach 2 should yield near-optimal detection performance for receivers having no post-combiner equalization. For cases where such equalization is used, Approach 2 would serve a different purpose—reducing the signal dispersion as seen by the equalizer input, thereby simplifying the requirements on equalizer design (e.g., number of taps) and improving convergence speed. Thus, Approach 2 has the twin virtues of not relying on data decisions and having utility both with and without post-combiner equalization.

To describe Approach 2, we refer to the block diagram of Fig. 1. The combiner output signal, $y(t)$, has a power spectrum density given by

$$S_y(f) = \underbrace{S(f) |H(f)|^2}_{\text{Signal}} + \underbrace{N_o(|\beta_1|^2 + |\beta_2|^2) |H_R(f)|^2}_{\text{Noise}}, \quad (10)$$

where N_o is the power spectrum density of the receiver input noise,* $H_R(f)$ represents whatever receiver selectivity precedes the combiner output, and $S(f)$ is the spectral density of the signal (excluding channel and combiner effects). More specifically,

$$S(f) = S_o |H_T(f)|^2 |H_R(f)|^2, \quad (11)$$

where $|H_T(f)|^2$ represents the spectral shaping in the transmitter and S_o is a spectral density scale factor.

All functions and parameters in (10) and (11) are design-specified except $|\beta_1|$ and $|\beta_2|$, which are controlled by the combiner circuitry, and $|H(f)|$, which must be measured in real time. Our scheme estimates $|H(f)|$ at N evenly spaced frequencies (N odd) within the

* Though not made explicit by the figure, we assume there is sufficient front-end amplification that the combiner gains β_1 and β_2 have no effect on receiver noise figure.

channel bandwidth by estimating the corresponding values of $S_y(f)$ [see eq. (10)]. Based on these estimates, β_1 and β_2 are adjusted to maximize a computed performance measure, Y , which we introduce shortly. Before doing so, we define the following:

$$\Delta f \triangleq \text{Spacing between estimates of } |H(f)|, \quad (12)$$

where $(N - 1)\Delta f \leq W$;

$$H_n \triangleq |H(f_n)|, \quad f_n = n\Delta f (n = 0, \pm 1, \dots, \pm (N - 1)/2); \quad (13)$$

$$\bar{H} \triangleq \text{Ave}_n\{H_n\} = \text{“Average Signal Gain”}; \quad (14)$$

$$X_{\text{sig}} \triangleq \text{Ave}_n\{S(f_n)\}(\bar{H})^2 W = \text{“Signal Power”}; \quad (15)$$

$$X_{\text{dis}} \triangleq \text{Ave}_n\{S(f_n)(H_n - \bar{H})^2\} W = \text{“Distortion Power”}; \quad (16)$$

$$X_{\text{noise}} \triangleq N_o(|\beta_1|^2 + |\beta_2|^2) \text{Ave}_n\{|H_R(f_n)|^2\} W = \text{“Noise Power”}; \quad (17)$$

where W and T are the channel bandwidth and digital symbol period, respectively. All quantities in these equations are known a priori except the H_n values, which are measured.

We now define the performance measure to be computed and maximized, namely,

$$Y \triangleq X_{\text{sig}} / (X_{\text{dis}} + X_{\text{noise}}). \quad (18)$$

This ratio is an approximation, computed from noncoherent spectral measurements, of the detector output signal-to-distortion ratio defined by (26), below. It is an apt measure to maximize in its own right, for the following reasons: In typical digital radio links, noise will not be a serious factor unless $H_1(f)$ and $H_2(f)$ are strongly faded. Therefore, maximizing Y will, in most cases, amount to minimizing the ratio $X_{\text{dis}}/X_{\text{sig}}$, which is a measure of the dispersion in $H(f)$. Including X_{noise} , however, safeguards against minimizing this ratio at an undue cost in signal (X_{sig}) and thus seriously degrading the signal-to-noise ratio.

The control strategy is therefore as follows (see Fig. 1): At the combiner output, a parallel bank of envelope detectors is used to estimate $|H(f)|$ at N frequencies. The spectral samples are digitized and applied to a microprocessor, which computes Y . This measure drives the search over $|\beta_1|$ (or $|\beta_2|$) and $\phi = \text{Arg}\{\beta_2\}$, i.e., these quantities are adjusted so as to maximize Y . Typically, they are adjusted iteratively, e.g., ϕ is changed in 0.1-radian step until a local maximum is found; then $|\beta_1|$ or $|\beta_2|$ is changed from 1 in steps of 0.1 until a maximum is found; and this process repeats, possibly using smaller steps in successive rounds, until Y can no longer be increased

by varying either β or ϕ . If each measurement (i.e., set of estimates of $|H(f)|$ at N frequencies) and computation for Y takes t_M seconds, and N_A steps are needed to find (β_1, β_2) , then the "solution time" of the combiner will be about $N_A t_M$. This number should be small compared to 1 second to achieve timely adaptation to multipath fades.

2.4 Measurements

We now discuss the scheme for measuring the set of H_n 's in (13). In our simulations (Section III) we treat only the case where $N = 3$ and $\Delta f = W/2$ (i.e., three samples, taken at the channel edges and center). In practice, the outer samples would probably be closer-in so as to minimize errors from adjacent-channel interference. Also, higher values of N (e.g., $N = 5$) might be worthwhile.

To see how accurate estimates of H_n might be obtained, let $G(f)$ represent a low-pass power gain function with bandwidth $b/2 \ll W$. We envision the measurement of H_n as involving a bandpass filter with power response $G(f - n\Delta f)$ followed by envelope detection and t_M -second averaging of the detector output (Fig. 1). Referring to eq. (10), the average power at the output of the bandpass filter will be

$$\bar{P}_n = \int \{S(f) |H(f)|^2 + N_o(|\beta_1|^2 + |\beta_2|^2) |H_R(f)|^2\} G(f - n\Delta f) df, \quad (19)$$

where $S(f)$ is defined by (11). Assuming a square-law detector, the time-averaged detector output will be

$$\langle \bar{P}_n \rangle = \bar{P}_n + \left\{ \begin{array}{l} \text{Fluctuation Noise;} \\ \text{Variance} \sim \bar{P}_n / bt_M \end{array} \right\}. \quad (20)$$

We now define two constants related to the system design functions, namely,

$$\eta_n \triangleq N_o \int |H_R(f)|^2 G(f - n\Delta f) df \quad (21)$$

and

$$\zeta_n \triangleq \int S(f) G(f - n\Delta f) df. \quad (22)$$

Based on (19) and the fact that $H(f)$ changes little over the bandwidth b , a microprocessor can estimate H_n using the formula

$$\hat{H}_n = \sqrt{\frac{\langle \bar{P}_n \rangle - \eta_n [|\beta_1|^2 + |\beta_2|^2]}{\zeta_n}}, \quad (23)$$

where $\langle \bar{P}_n \rangle$ is measured in real time; $|\beta_1|$ and $|\beta_2|$ are controlled parameters of known value; and η_n and ζ_n are predetermined constants.

Equation (23) shows how \hat{H}_n is computed in terms of measured or known quantities. To see what this computed number represents, we insert (19) through (22) into (23) and obtain

$$\hat{H}_n = \sqrt{\frac{\int S(f)G(f - n\Delta f)|H(f)|^2 df}{\int S(f)G(f - n\Delta f)df}} + \left\{ \begin{array}{l} \text{Term due solely to} \\ \text{fluctuation noise} \end{array} \right\}. \quad (24)$$

We can now cite choices for b and t_M that lead to accurate and sufficiently rapid estimations of H_n . As b gets very small, the first term under the radical sign in (24) approaches H_n^2 , so that the major inaccuracy in \hat{H}_n is due to fluctuation noise. To be more precise, the first term is close to H_n^2 so long as $|H(f)|^2$ changes little over the passband of $G(f - n\Delta f)$. Since we are considering propagation media with delay spreads of just a few nanoseconds, the design rule $b \leq 2$ MHz should permit more than adequate resolution in this regard. To achieve low mean-square fluctuation noise as well [second term in (24)], the condition $bt_M \geq 4000$ should be satisfied [see (20)]. Thus, with $b = 2$ MHz and $t_M = 2$ ms, H_n can be approximated with high accuracy by the quantity \hat{H}_n . Moreover, this design choice would permit numerous iterations of the search over β_1 and β_2 before the medium response changes appreciably.

III. PERFORMANCE STUDY

3.1 General

We have written a set of computer programs to simulate the behavior of the combiner scheme described above (Approach 2) and to analyze its performance and that of other receiver techniques. Each simulation is done for a specific pair of fading functions, $H_1(f)$ and $H_2(f)$, and for a specific value of Carrier-to-Noise Ratio (CNR). What is simulated is the sequential search over β_1 and β_2 , as performed by a receiver in real time to maximize the computed measure Y [see eq. (18)].

The analysis programs compute a detection performance measure for a nonequalized receiver using the (β_1, β_2) pairs derived in the simulations for Approach 2. The same measure is also computed for the cases of *optimal* combining (Approach 1) and no combining (non-diversity). Also, the analysis programs examine the signal dispersion at the combiner output for these various cases, which has relevance to receivers with post-combiner equalization.

3.2 Response pairs studied

We have specified eight distinct pairs of $H_1(f)$ and $H_2(f)$ for purposes of study. These pairs are collectively representative of what

Table I—Dual-channel response pairs studied

CASE	PLOTS OF $ H_1(f) , H_2(f) $ (IN dB)	SPACE DIVERSITY CHANNEL	PATH NUMBER, k	τ_k (IN ns)	$R_k \exp(j\theta_k)$
1		1	1	-4.0	$1.0 + j0.1$
		2	2	+1.0	$-0.5 + j0.9$
2		1	1	-4.0	$0.2 + j0.1$
		2	1	-5.0	$-1.1 + j0.6$
3		1	1	-1.0	$1.0 + j0.1$
		2	2	+2.0	$-0.9 + j0.5$
4		1	1	-1.0	$1.0 + j0.0$
		2	2	+2.0	$-0.9 + j0.0$
4		1	1	-1.0	$1.0 + j0.0$
		2	2	+2.0	$-0.9 + j0.0$

might arise in actual radio links using space diversity. In each case, $H_1(f)$ corresponds to a one-, two- or three-path medium, and similarly for $H_2(f)$. [In terms of (1) and (2), $K_1 = 1, 2$, or 3 in each case, and similarly for K_2 .] The corresponding time delays and complex gains are summarized in Table I. Also shown for each case are graphs of $|H_1(f)|$ and $|H_2(f)|$, in decibels, over a 40-MHz bandwidth.

Table I—(Cont.) Dual-channel response pairs studied

CASE	PLOTS OF $ H_1(f) , H_2(f) $ (IN dB)	SPACE DIVERSITY CHANNEL	PATH NUMBER, k	T_k (IN ns)	$R_k \exp(j\theta_k)$
5		1	1 2	-2.0 +2.0	$1.0 + j0.0$ $-0.872 - j0.223$
		2	1 2	-1.0 +3.0	$0.54 + j0.84$ $-0.66 - j0.60$
6		1	1 2	-3.0 +3.0	$1.0 + j0.0$ $-0.5 + j0.0$
		2	1 2	-3.0 +1.0	$-0.068 - j0.998$ $0.48 + j0.76$
7		1	1 2 3	-3.0 0.0 +2.0	$0.0 - j1.0$ $0.5 + j0.5$ $1.0 + j0.0$
		2	1 2 3	-3.0 0.0 +2.0	$0.0 - j1.0$ $-0.5 + j0.5$ $0.5 + j0.0$
8		1	1 2 3	-3.0 -1.0 +2.0	$-1.0 + j0.5$ $0.2 + j0.7$ $0.5 - j1.0$
		2	1 2 3	-2.0 +1.0 +2.0	$1.0 - j0.2$ $1.0 - j0.2$ $0.5 + j0.5$

Because of additive noise, the receiver performance for a given response pair would be affected by any amplitude scaling of $(H_1(f), H_2(f))$. We permit the possibility of such a scaling, for each case in Table I, by including it in the carrier-to-noise ratio parameter discussed below.

3.3 Signal design assumed

We assume throughout the study that the signal spectral density $S(f)$, (11), is rectangular over the channel bandwidth, with magnitude S_o . This corresponds to the use of ideal Nyquist pulses with 0-percent roll-off factor ($\alpha = 0$) and a symbol rate equal to the bandwidth ($1/T = W$). While not realistic, this assumption both simplifies the analysis and leads to somewhat poorer results than for other (more tapered) spectra.¹⁷ It thus serves both convenience and our confidence in whatever favorable outcome the study predicts. Adding to the "worst-case" nature of the results is that they are obtained for the largest common carrier channel bandwidth, $W = 40$ MHz.

3.4 Carrier-to-noise ratio

Finally, we introduce a so-called flat fading gain, g , with which we can amplitude-scale the various pairs ($H_1(f)$, $H_2(f)$) in Table I. Accordingly, we define the flat-fading carrier-to-noise ratio to be

$$\text{CNR} \triangleq g^2 S_o / N_o. \quad (25)$$

A typical system value for this quantity, with $g^2 = 1$, is 10^6 (60 dB). We will consider a wide range of values for CNR, thereby accommodating a wide range of amplitude scaling factors.

3.5 Method of analysis

We consider an M-level QAM system with the previously noted Nyquist signaling. At each baseband detector, data decisions are made every T seconds by comparing the sampled input with a set of decision thresholds. We now introduce the following new quantities:

$P_S \triangleq$ The squared signal sample (excluding ISI and noise) at either baseband detector, when a data value of +1 or -1 is being detected;

$P_I, P_N \triangleq$ The mean-squared ISI and noise, respectively, associated with the periodic samples at either baseband detector.

We define the detection signal-to-distortion ratio to be

$$\rho_D \triangleq P_S / (P_I + P_N) \quad (26)$$

and note that this quantity yields a reasonably tight upperbound to the bit error rate, via¹⁸

$$\text{Bit Error Rate} \leq 2 \exp(-\rho_D/2). \quad (27)$$

We thus regard ρ_D as a proper index for evaluating digital radio receiver performance.

Now consider a space diversity receiver using no post-combiner equalization and exhibiting a channel response $H(f)$, (9), at the

combiner output. For the assumed M-QAM system, with optimal timing and carrier recovery, we can show that

$$\rho_D = \frac{3}{M-1} \frac{\text{CNR}(H_{\max}^2)}{\text{CNR}(\overline{|H|^2} - H_{\max}^2) + (|\beta_1|^2 + |\beta_2|^2)} \equiv \frac{3}{M-1} \Gamma, \quad (28)$$

where

$$H_{\max} \triangleq \text{Max}_{t_0} \left| T \int_{-W/2}^{W/2} H(f) e^{-j\omega t_0} df \right| \quad (29)$$

and

$$\overline{|H|^2} \triangleq T \int_{-W/2}^{W/2} |H(f)|^2 df. \quad (30)$$

Note that the dependence of ρ_D on M resides entirely in the factor $3/(M-1)$. By using Γ as our signal-to-distortion measure, therefore, we can remove M from the set of problem variables.

Since $H(f)$ is linear in β_1 and β_2 , a solution exists for these gains (actually, for their *ratio*) that maximizes Γ . The solution can be found analytically by a variation on the method outlined in Section 5.1 of Ref. 15. In an actual receiver, the solution can be closely realized using practical circuitry, which we have designated as Approach 1 (Section 2.3). This solution for (β_1, β_2) is defined here to be optimal, and we will compute and present Γ results corresponding to it. These will be compared with Γ results for the “suboptimal” (β_1, β_2) , which we define as the pair produced under Approach 2 (Section 2.3) and which are obtained here via computer simulations. For completeness, we will also show Γ results for the case of no diversity, i.e., either $\beta_1 = 1, \beta_2 = 0$ (Branch 1 only is processed) or $\beta_1 = 0, \beta_2 = 1$ (Branch 2 only is processed).

Finally, in consideration of receivers using post-combiner equalization, we define a “dispersion index” for $H(f)$ as follows:

$$R \triangleq \left[\frac{\text{Max}_f \{|H(f)|\}}{\text{Min}_f \{|H(f)|\}} \right] |f| \leq 20 \text{ MHz}. \quad (31)$$

This is just the *range* of $|H(f)|$ over a 40-MHz channel bandwidth. Gersho has shown that adaptive equalizers converge more rapidly when this ratio is close to unity.¹⁹ Another likely benefit of near-unity R is that effective equalization should be attainable using a relatively small number of equalizer taps. We will compute R for the same cases that we compute Γ .

3.6 Simulation of the search algorithm for Approach 2

The simulations assume three spectrum measurements located at -20 MHz, 0 MHz, and $+20$ MHz relative to the band center. As noted before (Section 2.4), we can identify design parameters that drive measurement inaccuracies to negligible levels. For this reason, we assume exact spectral estimates and have not attempted to simulate measurement inaccuracies.

For a given response pair (Table I) and CNR, (24), the simulation proceeds as follows: During the search process one of the β values is always set to unity with the other value less than or equal to unity. Initially, $|\beta_1|$, $|\beta_2|$ and ϕ are set to 1.0, 1.0 and 0, respectively. Then ϕ is varied in $\delta\phi$ -radian steps until Y , (18), is maximized. Next, $|\beta_2|$ is varied in steps of $\delta\beta$ (if the simulation calls for increasing $|\beta_2|$ above unity, we set $|\beta_2| = 1$ and decrease $|\beta_1|$ instead) until Y is again maximized. The procedure is then repeated until Y cannot be further maximized.

In the simulations, we generally started with $\delta\beta = 0.1$ and $\delta\phi = 0.1$ radian. We noticed that a steady-state result usually required about twice the number of steps needed to reach it directly. For example, if the solution were $|\beta_1| = 1.0$, $|\beta_2| = 0.4$, and $\phi = 3.1$ radians, a "direct" path would entail six 0.1-step changes in $|\beta_2|$ and 31 0.1-radian changes in ϕ , or 37 steps in all. The maximizing algorithm described above, however, was found to take about 74 steps. Using $\delta\beta = 0.01$ and $\delta\phi = 0.01$ radian would, of course, require 740 steps. Our approach was to use 0.1 for both increments until a stable solution was reached and then to proceed to a finer solution by changing both increments to 0.01. The end results using this faster two-stage process were identical to those obtained by using 0.01 throughout.

Using the two-stage process, the measurement steps needed to reach a solution should not exceed perhaps 200 for a "cold start" adjustment of the combiner gains. For $t_M = 2$ ms per measurement, this corresponds to a solution time of 0.4 second. In the dynamic situation where the receiver tracks the channel variations, the solution time would be much shorter.

An important question is whether the derived solutions for (β_1, β_2) are unique. We explored this question for each of the eight cases in Table I, plus a few others. For each case, we simulated the search process using 24 different starting points for $|\beta_1|$, $|\beta_2|$ and ϕ . Specifically, $(|\beta_1|, |\beta_2|)$ was initialized to each of three different pairs of values, namely, (0.5, 1.0), (1.0, 1.0), and (1.0, 0.5); and, for each such pair, ϕ was initialized to each of eight values, namely, 0, 45, 90, \dots 315 degrees. In studying the 24 solutions for each case, we observed:

1. The solution for (β_1, β_2) was not unique in every case; for some cases, multiple local maxima for Y were found to exist, and the search

algorithm converged to one or another, depending on the starting point.

2. The multiple-solution condition appears to exist only at high CNR (≥ 40 dB); for CNR = 20 dB, all solutions for (β_1, β_2) were found to be unique.*

3. In those cases where multiple solutions *did* exist, the lowest of the resulting Y values was always quite good (> 30 dB) and usually not much below the highest Y.

From these observations, we tentatively conclude that the proposed scheme converges to stable solutions under all circumstances. Further, the solutions seem to be nonunique only in circumstances where all of the possible solutions are satisfactory.

IV. RESULTS

A qualitative assessment of the combining scheme can be gained using graphical results. Figures 2, 3, and 4 show, for each of Cases 1, 3, and 4 in Table I, plots of $H_1(f)$, $H_2(f)$, and the $H(f)$ obtained when CNR = ∞ . In each figure, the severe dispersions of $H_1(f)$ and $H_2(f)$ are seen to be almost eliminated by the combining: $H(f)$ presents a nearly flat amplitude response and a nearly linear phase response over the channel bandwidth.

To make the assessments more quantitative, we have computed Γ and R , (28) and (31), for each case in Table I; for each of the CNR values 60 dB, 40 dB, and 20 dB; and for each of the following four solution strategies for (β_1, β_2) :

Solution 1: Optimal; based on Approach 1, it leads to a maximum for Γ with respect to β_1 and β_2 .

Solution 2: Suboptimal; based on Approach 2, it leads to a maximum for Y with respect to β_1 and β_2 .[†]

Solution 3: Nondiversity; Branch 1 processed only.

Solution 4: Nondiversity; Branch 2 processed only.

We define Γ_i and R_i ($i = 1, 4$) to be the decibel values of Γ and R , respectively, for Solution 1. The results are given by Tables II, III, and IV, each corresponding to one of the three values of CNR.

V. DISCUSSION AND CONCLUSION

The data of Tables II through IV lead us to the following observations:

1. Comparing Γ_2 with Γ_3 and Γ_4 reveals that an unequalized receiver

* Under low CNR conditions, the noise term dominates. Since the noise term is quadratic in β_1 and β_2 , it leads to a unique maximum for Y.

[†] In all cases where multiple solutions were found to exist, we used the one corresponding to the higher Y.

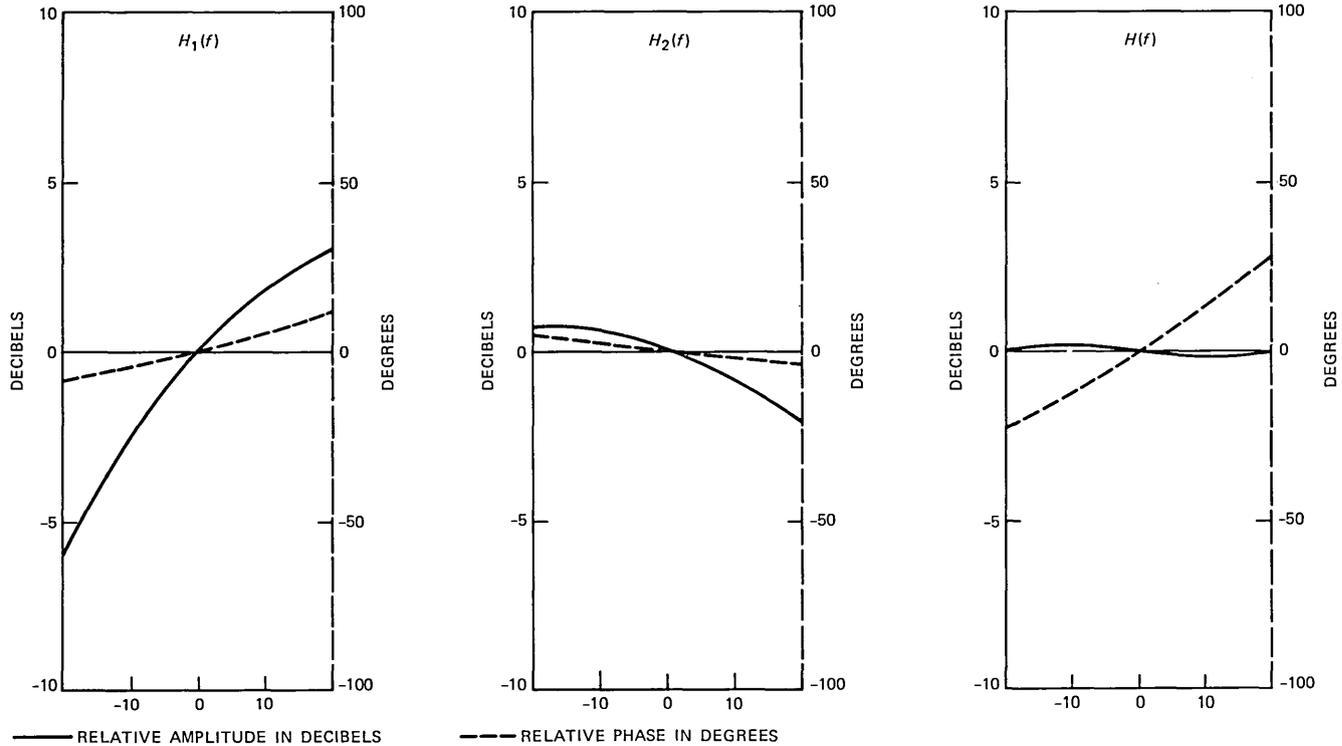


Fig. 2—Relative amplitude and phase plots for $H_1(f)$ and $H_2(f)$, corresponding to Case 1 in Table I, and the resulting combiner output response, $H(f)$, when β_1 and β_2 are derived using the control scheme in Fig. 1; no noise present. The plots cover a 40-MHz bandwidth and, for convenience, all magnitudes and phases are shifted so as to be zero at band center.

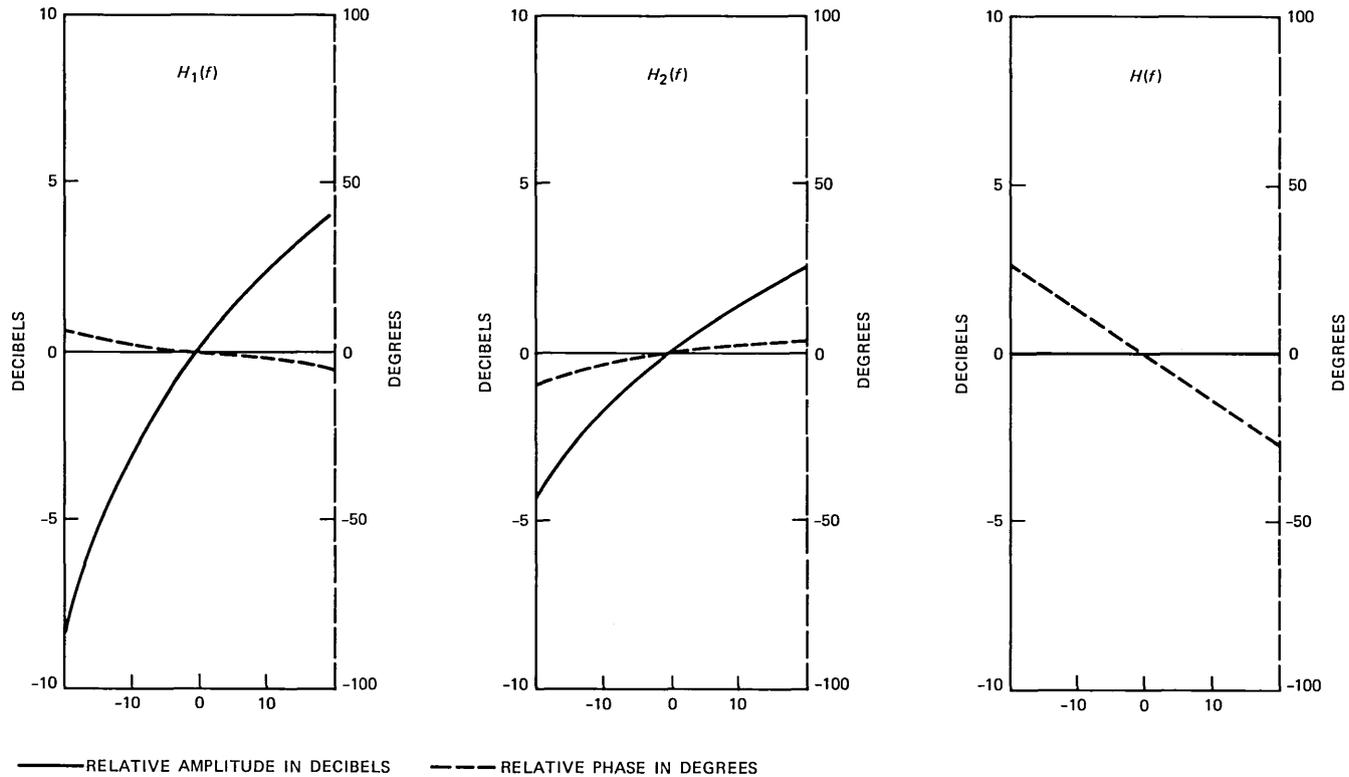
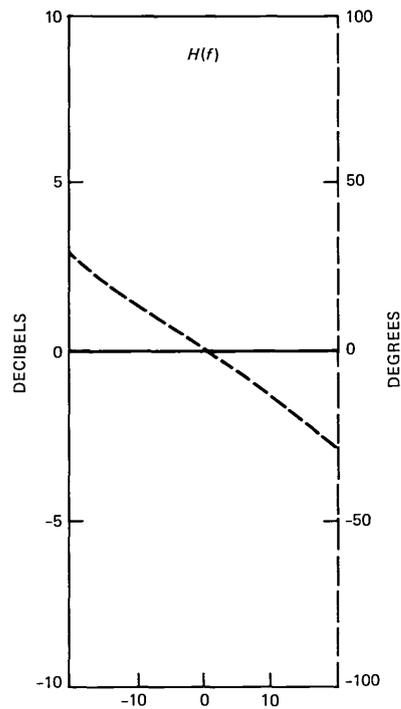
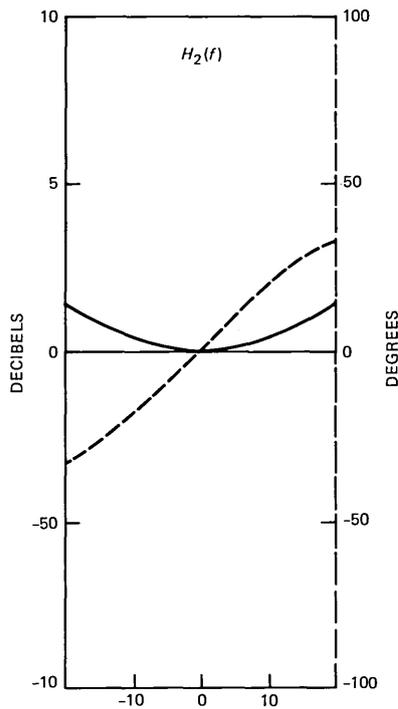
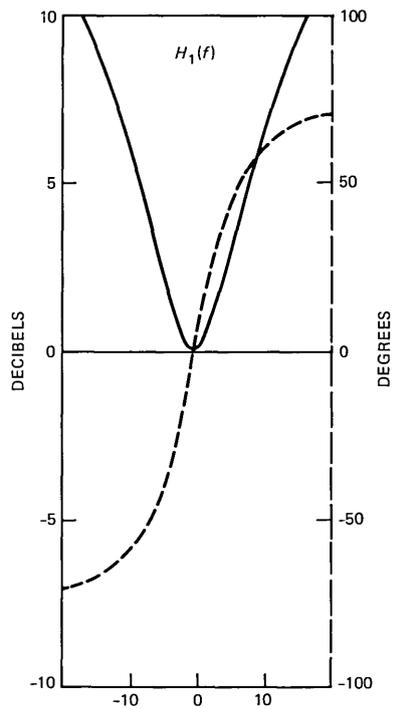


Fig. 3—Same as Fig. 2, for Case 3 in Table I.



— RELATIVE AMPLITUDE IN DECIBELS

- - - RELATIVE PHASE IN DEGREES

Fig. 4—Same as Fig. 2, for Case 4 in Table I.

Table II—Combiner performance results for CNR = 60 dB

Case	Detection Measure (in dB)				Decibel Range of $ H(f) $			
	Γ_1	Γ_2	Γ_3	Γ_4	R_1	R_2	R_3	R_4
1	43.80	42.67	11.12	20.88	0.11	0.062	9.34	2.77
2	46.99	46.99	46.99	11.15	0.006	0.0	0.0	9.79
3	36.06	35.62	9.17	13.39	0.21	0.14	12.25	6.82
4	53.13	53.12	6.82	25.91	0.002	0.001	11.35	1.39
5	29.23	27.60	2.80	3.02	0.37	0.27	16.96	16.49
6	35.32	34.99	18.21	5.38	0.26	0.21	3.19	19.09
7	42.28	40.72	18.31	17.80	0.11	0.13	3.76	3.08
8	63.84	59.24	10.70	29.02	0.006	0.022	7.38	1.06

Table III—Combiner performance results for CNR = 40 dB

Case	Detection Measure (in dB)				Decibel Range of $ H(f) $			
	Γ_1	Γ_2	Γ_3	Γ_4	R_1	R_2	R_3	R_4
1	41.01	40.43	11.12	20.86	0.11	0.06	9.34	2.77
2	27.46	26.99	26.99	11.15	0.44	0.0	0.0	9.79
3	29.91	29.60	9.16	13.39	0.41	0.18	12.25	6.82
4	33.38	33.33	6.77	25.51	0.11	0.05	11.35	1.39
5	27.11	25.70	2.79	3.00	0.40	0.27	16.96	16.49
6	31.29	30.97	18.13	5.37	0.36	0.21	3.19	19.09
7	38.08	37.72	18.30	17.71	0.17	0.13	3.76	3.08
8	46.23	46.00	10.68	28.96	0.022	0.014	7.38	1.06

Table IV—Combiner performance results for CNR = 20 dB

Case	Detection Measure (in dB)				Decibel Range of $ H(f) $			
	Γ_1	Γ_2	Γ_3	Γ_4	R_1	R_2	R_3	R_4
1	25.98	25.76	10.68	19.67	0.58	0.43	9.34	2.77
2	12.91	12.54	6.99	10.96	5.72	4.31	0.0	9.79
3	14.68	14.35	8.29	12.93	3.96	2.85	12.25	6.82
4	15.67	15.63	3.58	15.64	1.54	1.37	11.35	1.39
5	11.72	11.68	1.97	2.19	2.16	1.83	16.96	16.49
6	15.21	15.14	13.49	4.76	1.99	1.62	3.19	19.09
7	23.80	23.74	17.69	12.75	0.75	0.63	3.76	3.08
8	27.63	27.63	9.09	25.41	0.22	0.20	7.38	1.06

with “suboptimal” space diversity combining would generally perform much better than one using no diversity (Γ_3 or Γ_4) or selection diversity (the larger of Γ_3 and Γ_4).

2. Comparing Γ_2 with Γ_1 reveals that the “suboptimal” combiner would perform nearly as well as the “optimal” one, at least for the response pairs considered here.

3. For the “suboptimal” combiner, the range (R_2) of the post-

combiner frequency response is generally much smaller than that for either branch alone (R_3 and R_4). It is also comparable to that for "optimal" combining (R_1), often being even smaller.

4. Comparing Tables II, III, and IV reveals the trends of combiner performance as CNR decreases from high values (≥ 40 dB) to low ones (20 dB). At high values, the combiner emphasizes minimum dispersion, as reflected in the data for R ; at low values, combiner action aims more at minimizing noise and so dispersion reduction is limited.

It thus appears that the noncoherent measurement/control scheme described here would be effective in *any* receiver situation. This includes receivers *without* post-combiner equalization, wherein Γ should be maximal to optimize detection; and receivers *with* equalization, wherein R should be minimal to facilitate equalizer convergence. It is clear that both aims are served by the same algorithm. Moreover, the scheme is simple, fast-acting, and operates in proximity to the combining circuitry.

Some issues remain to be settled. One is whether the search algorithm will always converge to either the near-optimal solution or to one that, in any case, yields very good performance. The results obtained for our limited sampling of response pairs offer encouragement on this score. Another issue concerns the effectiveness of diversity combining over a statistical *ensemble* of fading conditions, that is, we have obtained results here for just eight selected response pairs. In a separate study, however, we invoked a recently developed statistical model for dual-diversity channels^{20,21} to simulate a large population (ensemble) of response pairs ($H_1(f)$, $H_2(f)$), and we obtained probability distributions for Γ and R over that population. The results are reported in a companion paper.²²

REFERENCES

1. C. W. Anderson, S. G. Barber, and R. N. Patel, "The Effect of Selective Fading on Digital Radio," *IEEE Trans. Commun.*, COM-27, No. 12 (December 1979), pp. 1870-75.
2. S. Komaki et al., "Characteristics of a High Capacity 16 QAM Digital Radio System in Multipath Fading," *IEEE Trans. Commun.*, COM-27, No. 12 (December 1979), pp. 1854-61.
3. Y. Y. Wang, "Simulation and Measured Performance of a Space Diversity Combiner for 6 GHz Digital Radio," *IEEE Trans. Commun.*, COM-27, No. 12 (December 1979), pp. 1896-907.
4. S. Komaki, Y. Okamoto, and K. Tajima, "Performance of 16-QAM Digital Radio System Using New Space Diversity," Paper 52.2, Int. Conf. Commun., June 8-12, 1980, Seattle, Washington.
5. T. Murase, K. Morita, and S. Komaki, "200 Mb/s 16-QAM Digital Radio System with New Countermeasure Techniques for Multipath Fading," Paper 46.1, Int. Conf. Commun., June 14-18, 1981, Denver, Colorado.
6. P. Hartmann and B. Bynum, "Design Consideration for an Extended Range Adaptive Equalizer," Paper 46.5, Int. Conf. Commun., June 14-18, 1981, Denver, Colorado.
7. A. J. Giger and W. T. Barnett, "Effects of Multipath Propagation on Digital Radio," *IEEE Trans. Commun.*, COM-29, No. 9 (September 1981), pp. 1345-52.

8. P. D. Karabinis, "Maximum-Power and Amplitude-Equalizing Algorithms for Phase Control in Space Diversity Combining," *B.S.T.J.*, 62, No. 1, Part 1 (January 1983), pp. 63-90.
9. R. Iyer et al., "Effects of Space Diversity Reception on Multipath Induced Distortions in Terrestrial Microwave Channels," Paper 3.B.5, Int. Conf. Commun., June 13-17, 1982, Philadelphia, PA.
10. A. Ricagni and T. Testi, "IF Combining Techniques for Space Diversity in Analog and Digital Radio Systems," Paper 4.B.6, Int. Conf. Commun., June 13-17, 1982, Philadelphia, PA.
11. R. A. Nichols, "An IF Combiner for Digital and Analog Radio Systems," Paper 4.B.7, Int. Conf. Commun., June 13-17, 1982, Philadelphia, PA.
12. L. J. Greenstein and B. A. Czekaj, "A Polynomial Model for Multipath Fading Channel Responses," *B.S.T.J.*, 59, No. 7 (September 1980), pp. 1197-225.
13. L. J. Greenstein and B. A. Czekaj, "Modeling Multipath Fading Responses Using Multitone Probing Signals and Polynomial Approximation," *B.S.T.J.*, 60, No. 2 (February 1981), pp. 193-214.
14. W. C. Wong and L. J. Greenstein, unpublished work.
15. R. W. Lucky, J. Salz, and E. J. Weldon, Jr., *Principles of Data Communication*, New York, NY: McGraw-Hill, 1968, Chapter 6.
16. B. Widrow, "A Comparison of Adaptive Algorithms Based on the Methods of Steepest Descent and Random Search," *IEEE Trans. Ant. and Prop.*, AP-24, No. 5 (September 1976), pp. 615-37.
17. W. C. Wong and L. J. Greenstein, "Multipath Fading Models and Adaptive Equalizers in Microwave Digital Radio," *IEEE Trans. Commun.*, COM-32, No. 8 (August 1984), pp. 928-34.
18. G. J. Foschini and J. Salz, "Digital Communications Over Fading Radio Channels," *B.S.T.J.*, 62, No. 2, Part 1 (February 1983), pp. 429-56.
19. A. Gersho, "Adaptive Equalization of Highly Dispersive Channels for Data Transmission," *B.S.T.J.*, 48, No. 1 (January 1969), pp. 55-70.
20. W. D. Rummmler, "A Statistical Model of Multipath Fading on a Space Diversity Radio Channel," *B.S.T.J.*, 61, No. 9, Part 1 (November 1982), pp. 2185-219.
21. W. D. Rummmler, "A Rationalized Model for Space and Frequency Diversity Line-of-Sight Radio Channels," Paper E2.7, Int. Conf. Commun., June 19-22, 1983, Boston, MA.
22. L. J. Greenstein and Y. S. Yeh, "A Simulation Study of Space Diversity and Adaptive Equalization in Microwave Digital Radio," *AT&T Tech. J.*, this issue.

AUTHORS

Larry J. Greenstein, B.S.E.E., 1958, M.S.E.E., 1961, and Ph.D. (Electrical Engineering), 1967, Illinois Institute of Technology; AT&T Bell Laboratories, 1970—. Mr. Greenstein currently heads the Radio Systems Research Department at Crawford Hill in Holmdel, NJ. His most recent work has dealt with communications satellites, mobile telephony, and microwave digital radio. His previous work was on digital encoding, digital filtering, and, at IIT Research Institute before 1970, airborne radar. Member, Eta Kappa Nu, Tau Beta Pi, and Sigma Xi; Senior Member, IEEE; Senior Technical Editor, *IEEE Communications Magazine*; co-recipient, *IEEE Communications Society's 1984 Prize Paper Award in Communications Systems*.

Yu-Shuan Yeh, B.S.E.E., 1961, National Taiwan University; M.S.E.E., 1964, Ph.D., 1966, University of California, Berkeley, CA; Harvard University, 1966-67; AT&T Bell Laboratories, 1967—. From 1961 to 1962, Mr. Yeh was an electronics officer in the Chinese Navy. He was a Research Fellow at Harvard University from 1966 to 1967 doing antenna research. In 1967 he joined AT&T Bell Laboratories and is currently a Supervisor in the Network Systems Research Department. His research interests include satellite communications, mobile radio, microwave digital radio, and data networks. Mr. Yeh holds over a dozen patents and is the recipient of two best paper awards from *IEEE Transactions*. Fellow, IEEE.

A Simulation Study of Space Diversity and Adaptive Equalization in Microwave Digital Radio

By L. J. GREENSTEIN and Y. S. YEH*

(Manuscript received May 30, 1984)

In this paper we analyze the performance of M -level quadrature amplitude modulation digital radio systems subjected to microwave multipath fading. We consider two kinds of adaptive receiver techniques, either singly or in combination: dual space diversity and adaptive equalization. The space diversity is assumed to be of either the selection type or the continuous-combining type, and the equalization is assumed to be ideal. We describe a specific form of combining which is optimal when no post-combiner equalization is used. A primary aim of the study is to quantify the performance of this combining approach and to compare it with alternate strategies. The study uses Monte Carlo simulations of the dual-channel fading response functions based on a recently published statistical model. For each response pair generated, a receiver detection measure is derived analytically in terms of the system parameters and receiver approach. Probability distributions of this measure, obtained by simulating several thousand response pairs, are then computed. They can be interpreted as displaying the link outage probability as a function of the number of modulation levels (M). We find that the appropriate combining scheme can serve in some cases to avoid the need for adaptive equalization. Also, where post-combiner equalization is used, the same scheme, while no longer optimal, can sharply reduce the dispersion seen by the equalizer input.

I. INTRODUCTION

A continuing challenge in microwave digital radio is to find ways to counter multipath fading. Many recent efforts—theoretical, experi-

* Authors are employees of AT&T Bell Laboratories.

Copyright © 1985 AT&T. Photo reproduction for noncommercial use is permitted without payment of royalty provided that each reproduction is done without alteration and that the Journal reference and copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free by computer-based and other information-service systems without further permission. Permission to reproduce or republish any other portion of this paper must be obtained from the Editor.

mental, and developmental—have addressed the use of space diversity and adaptive equalization, individually or in combination, to maintain stringent outage objectives for increasingly high-level modulations.¹⁻¹⁴ This paper reports on a comparative study of several receiver processing approaches, enlarging on the evaluations presented in a companion paper.¹⁵

The receiver processing approaches studied here consist of three space diversity options, namely, *no* diversity, *selection* diversity, and *continuous-combining* (or *combining*) diversity. For each of these options, we evaluate performance statistics both *with* and *without* the presence of ideal adaptive equalization.

Our investigation is facilitated by a new statistical model for the response functions of two space diversity channels.^{16,17} We describe that model in Section II and introduce some extensions germane to the present study.

We assume throughout this work a system using M -level Quadrature Amplitude Modulation (M-QAM) with cosine roll-off spectral shaping. This system and the various space diversity/equalization configurations to be studied are described in Section III.

For any given dual-channel response pair and receiver processing approach, we can compute one or more relevant performance measures. By combining the published model with Monte Carlo simulation techniques to generate response pairs, we can obtain probability distributions of these measures over an ensemble of multipath fades. In Section IV we define the performance measures to be studied and outline the analytical methods by which they are computed, and in Section V we describe the method of Monte Carlo simulation and present quantitative results. These are in the form of probability distributions and reveal the influence of various system parameters, channel parameters, and receiver processing approaches on link performance.

II. DUAL-CHANNEL FADING MODEL

2.1 Background

We are dealing with propagation media and systems wherein the multipath delay spread generally is small compared to the inverse channel bandwidth. Consequently, fading responses vary smoothly with frequency over a given channel and can be accurately approximated using simple functions. A popular complex function for this purpose is the one introduced by Rummmler, initially for nondiversity channels^{18,19} and, more recently, for dual-diversity channels.^{16,17} In its most general form, Rummmler's function approximates a complex response $H(f)$ on a given frequency interval, $[-W/2, W/2]$, with $W \leq 40$ MHz, by

$$\tilde{H}(f) = [c_1 + c_2 e^{-j\omega\tau}] e^{-j\omega t_o}; \quad |f| \leq W/2, \quad (1)^*$$

where c_1 and c_2 are complex coefficients that vary slowly with time but can be regarded as quasistatic; τ is a fixed parameter (6.3 ns); and t_o is a time delay (nominally, the propagation delay through the medium for which $H(f)$ is the response). If $H(f)$ were precisely known on $[-W/2, W/2]$, $\tilde{H}(f)$ could be fitted to it by first choosing least-mean-squared values for c_1 and c_2 (i.e., by minimizing, for given t_o , the integration of $|H(f) - \tilde{H}(f)|^2$ over $[-W/2, W/2]$), and then minimizing the result with respect to t_o .

An alternate form for $\tilde{H}(f)$ that more closely resembles the form used by Rummler is

$$\tilde{H}(f) = a[1 - b e^{-j\omega\tau}] e^{j(\Phi_o - \omega t_o')}; \quad |f| \leq W/2, \quad (2)$$

where a is real and positive; b is complex; $|b| \leq 1$; and τ is either +6.3 ns ("minimum-phase" response) or -6.3 ns ("nonminimum-phase" response). For each of these two conditions, it is a simple matter to relate a , b , Φ_o , and t_o' to c_1 , c_2 , and t_o in (1). The question "minimum phase or nonminimum phase" has been left open to date, since only amplitude versus frequency data have informed the major attempts at statistical modeling.¹⁸⁻²⁰

In Rummler's work, the phase factor $\exp(j(\Phi_o - \omega t_o'))$ is omitted. Again, this is necessitated by the absence of phase versus frequency information in the available databases. In nondiversity reception, moreover, this factor is immaterial since the phase Φ_o and time delay t_o' would be tracked, and their effects removed, by the carrier and timing recovery circuits of the receiver.

Now, however, consider the dual-diversity case. As before, each of the two responses, $H_1(f)$ and $H_2(f)$, can be approximated on $[-W/2, W/2]$ using functions like (1) or (2). Since combining diversity involves forming the composite response $\beta_1 H_1(f) + \beta_2 H_2(f)$, the Φ_o term for each channel is immaterial; these phases can be regarded as absorbed into the phases of β_1 and β_2 . The difference in t_o' for the two channels, however, is another matter. We will return to this point later, after we summarize the statistical model.

2.2 Joint statistics: Rummler's model

The dual-diversity model is derived from data collected on a 26.4-mile path in Georgia in the 6-GHz band. The details—data, methods, and results—are well documented in Refs. 16 and 17. Here we will merely state some major results, using slightly different symbols where appropriate.

* Throughout this work, f is frequency measured from the center of the RF or IF channel under study, so that all responses are equivalent low-pass functions.

To begin with, the expected number of fading seconds on a microwave hop per year is

$$T_o = 52,800c(F/6)(D/25)^3, \quad (3)$$

where F is the microwave carrier frequency in GHz; D is the hop length in miles; and c is the terrain factor, varying between 0.25 and 4. During the rest of the year (normal propagation), the two response functions are $H_1(f) = H_2(f) = 1 + j0$. During multipath fading, however, they are

$$H_1(f) = a_1[1 - b_1 \exp(\pm j\omega\tau)], \quad (4a)$$

$$H_2(f) = a_2[1 - b_2 \exp(\pm j\omega\tau)], \quad (4b)$$

where $\tau = 6.3$ ns and the proper sign to use before it is an open question. Both functions apply over the limited range $|f| \leq W/2$, where $W \leq 40$ MHz.

The joint statistics of the a 's and b 's in (4) were published initially in Ref. 16. To smooth out some apparent artifacts of the data collections and reductions, Rummler subsequently published a "rationalized" version in Ref. 17. We will not reproduce his mathematical descriptions for the joint statistics, but a few important features should be noted:

- The four quantities $|b_1|$, $\text{Arg}\{b_1\}$, $|b_2|$ and $\text{Arg}\{b_2\}$ are mutually independent random variables, which means that the relative shapes of $H_1(f)$ and $H_2(f)$ are statistically independent.
- The amplitude factors a_1 and a_2 are lognormal random variables, i.e., their decibel values are Gaussian. Moreover, the mean of each Gaussian variable is a function of the magnitude of the corresponding b ; each of the variations about the mean has a standard deviation near 7.0 dB; and the two variations are correlated.
- For the data reductions reported in Ref. 16, the correlation factor (ρ) between these variations was 0.65. In Ref. 17, Rummler generalizes this result by giving a simple empirical relationship between ρ and the vertical spacing of the diversity antennas.

2.3 Extensions and sensitivity considerations

We have modified Rummler's model, for purposes of our study, in three ways. First, we address the minimum-phase/nonminimum-phase question in the following manner: Let

$$H_1(f) = \begin{cases} a_1[1 - b_1 e^{-j\omega\tau}] & \text{with probability } p \\ a_1[b_1 - e^{-j\omega\tau}] & \text{with probability } (1 - p). \end{cases} \quad (5)$$

Assuming that $\tau = 6.3$ ns and $|b| \leq 1$, we see that the first form is minimum phase and the second form is nonminimum phase. Thus, p

is the fraction of fades for which the form of $H_1(f)$ is minimum phase. We use the same description for $H_2(f)$, and assume that the minimum-phase condition occurs independently for the two response functions. In our simulations, we treat p as a parameter and vary it from 0 to 1 to assess its importance.

Second, we assume that $H_2(f)$ contains an additional phase factor, $\exp(-j\omega\delta t)$, where δt is the difference in t'_o for the “best” approximations to the two channel responses. This quantity cannot be known without simultaneous coherent measurements on both channels. We estimate, however, that it can be ± 2 ns just from time delay misalignments between the two diversity branches. Allowing as much as ± 4 ns for propagation differences, we speculate that δt lies between ± 6 ns. We parameterize it accordingly in our simulations, always holding it fixed for a given run of Monte Carlo trials.

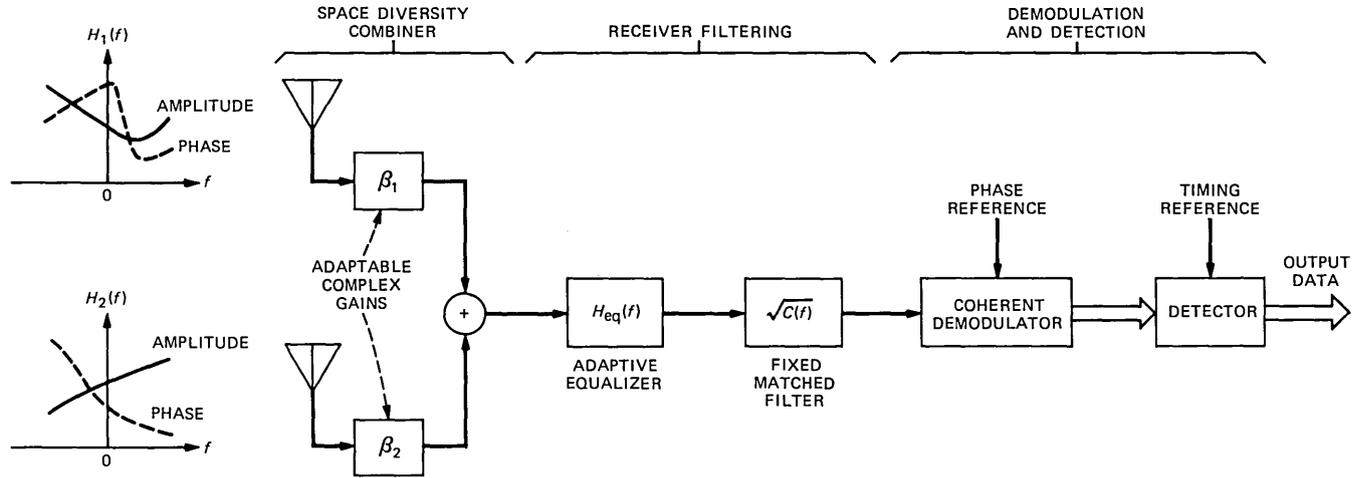
Finally, we allow the possibility that the correlation factor (ρ) between the decibel variations of a_1 and a_2 can differ from one path or time period to another, even for the same antenna spacing. We treat it parametrically in our simulations, varying it from 0 to 1 to assess its importance.

To summarize, we are concerned that our simulation results be “robust”, i.e., not sensitive to—or artifacts of—uncertain features of the model. For this reason, we have identified some key uncertainties, cast them in terms of numerical quantities (p , δt , and ρ), and treated these quantities as parameters. As for the joint statistics assumed for the a 's and b 's in (4),^{16,17} it is conjectural whether these descriptions would be applicable to all paths and time periods. For now, we regard these statistics as representative and useful for study purposes.

III. RECEIVER PROCESSING APPROACHES

3.1 General receiver structure

All the receiver processing approaches to be considered here can be cast as special cases of the general arrangement in Fig. 1. The original transmission is an M-QAM signal whose baseband pulse has the Fourier transform $\sqrt{C(f)}$, with $C(f)$ being the cosine roll-off function.²¹ The two receiver inputs differ only in that the propagation medium delivers a different response into each space diversity antenna. Typically, the antennas would be vertically displaced by 40 feet or more so as to receive relatively uncorrelated responses during fading. Each diversity branch is shown having a separate, adaptable complex gain. The space diversity combiner is terminated by a signal adder that delivers a single-branch signal to the remainder of the receiver. Although shown here at RF, this circuitry could as readily be at IF. Also (though not made explicit by the figure), we assume there is



(NOTE: f IS MEASURED FROM CENTER OF RADIO CHANNEL.)

Fig. 1—Block diagram of a digital radio receiver with dual-channel space diversity. The modulation is M-QAM and the spectral shaping is cosine roll-off.

sufficient front-end amplification that the combiner gains β_1 and β_2 have no effect on receiver noise figure.

The combiner output is shown applied to an adaptive equalizer, with response $H_{\text{eq}}(f)$. For the cases studied that do not involve the use of equalization, we will set $H_{\text{eq}}(f)$ to unity. The second post-combiner filter has a fixed response, $\sqrt{C(f)}$, which provides matched filter reception in the absence of fading. Because the receiver is linear, the order of the adaptive and fixed filters is immaterial, as is the manner in which each is distributed among RF, IF, and baseband stages.

The coherent demodulation stage uses a quadrature pair of local oscillator signals which we assume to be optimally phased.* The baseband outputs are sampled in the detector every T seconds, where T is the symbol period ($1/T$ is the baud), and we assume the sampling phase (timing epoch) to be optimal.

What is *not* shown in Fig. 1 is the means for controlling the adaptable gains, i.e., the tap gains of the equalizer (if used) and the diversity branch gains, β_1 and β_2 . Regarding the former, adaptive equalization is a well-developed art, and so we shall merely assume idealized equalizer responses, confident that they can be closely realized in practice. Regarding the latter, we shall specify idealized criteria for determining β_1 and β_2 , both for selection diversity and combining diversity, and shall apply these criteria analytically in our performance calculations. Only in the case of combining diversity, however, will we address the issue of practical implementation.

3.2 Space diversity options

We delineate the three particular diversity options as follows:

1. *Nondiversity*, wherein only the response $H_1(f)$ (corresponding to the higher antenna) is received; thus, $(\beta_1, \beta_2) = (1, 0)$.

2. *Selection diversity*, wherein (β_1, β_2) is either $(1, 0)$ or $(0, 1)$, i.e., only one branch is processed at a given time, depending on which is deemed "best" in some sense.

3. *Combining diversity*, wherein β_1 and β_2 are continuously variable and are adjusted according to a specified criterion.

We will discuss these options first for the case of no adaptive equalization ($H_{\text{eq}}(f) = 1 + j0$ in Fig. 1). Accordingly, the composite channel response as seen at the fixed filter input is

$$H(f) = \beta_1 H_1(f) + \beta_2 H_2(f), \quad (6)$$

where all quantities are complex.

The nondiversity case is included here for purposes of comparison.

* Optimal carrier and timing phase is meant, in these discussions, to denote phases for which the detection signal-to-distortion ratio is maximal.

It will be shown to yield disastrous performance results in the absence of equalization, as expected.

For the case of selection diversity, the signal adder would, in practice, be a selection switch operated by real-time decisions of some kind. We will obtain results for a theoretically optimal selection strategy, described as follows: At all times, (β_1, β_2) is either $(1, 0)$ or $(0, 1)$, whichever yields the lower bit error rate for the prevailing response pair, $(H_1(f), H_2(f))$. Since selection diversity is included here primarily for purposes of comparison, we ignore the issue of practical implementation.

The combining diversity case is the one that interests us most. We will show that combining diversity alone (i.e., without equalization) can go a long way towards combatting dispersion as well as noise. As explained in Ref. 15, this potential exists for any channel over whose bandwidth the frequency response can be approximated by a first-order polynomial in $j\omega$. This condition applies to the microwave common-carrier channels at hand.²⁰

The criterion we will assume for adapting (β_1, β_2) in the absence of equalization is maximization of a particular signal-to-distortion ratio at the detector, where "distortion" includes Intersymbol Interference (ISI) and noise. This ratio, which we call ρ_D , is defined and analyzed in Section 4.1. We will see that ρ_D is quadratic in β_1 and β_2 ; thus, a unique maximum exists and the (β_1, β_2) pair that produces it can be found analytically. More important, it is well known that this solution for (β_1, β_2) can be realized in a receiver by means of practical circuitry, specifically, via decision-directed gradient search algorithms.^{22,23}

In Ref. 15 we describe an alternative scheme for adapting (β_1, β_2) that does *not* rely on data decisions and yet produces results close to those using the more conventional approach. The scheme involves measuring the combiner output power spectrum density at three or more inband frequencies, computing from these samples a measure that approximates ρ_D , and using this measure to drive the search over β_1 and β_2 .

With all this in mind, we will derive analytical solutions for the (β_1, β_2) pair that maximizes ρ_D in the absence of equalization, and will denote it by $(\tilde{\beta}_1, \tilde{\beta}_2)$. Since maximizing ρ_D minimizes an upper bound on the bit error rate (Section 4.1), we will call this gain pair *optimal*. Finally, we will assume that optimal or near-optimal solutions can always be realized in practical receivers using schemes such as those mentioned above.

3.3 Receivers with adaptive equalizers

A variety of practical approaches could be assumed for the equalizer in Fig. 1. These include the reciprocal equalizer,²⁴ the Minimum Mean

Square Error (MMSE) equalizer,²⁵ the Decision Feedback Equalizer (DFE),²⁶ and adaptive cancellation.²⁷ In terms of detection signal-to-distortion ratio, these four approaches improve in the order cited. However, in terms of outage performance over a large ensemble of fading responses, there seems to be little to differentiate among them.^{28,29} We shall therefore assume an ideal reciprocal equalizer, both because it is the easiest to analyze and because it provides a tight worst-case bound on the performance of the more optimal equalizers.

The reciprocal equalizer has a response

$$H_{\text{eq}}(f) = \frac{1}{H(f)} = \frac{1}{\beta_1 H_1(f) + \beta_2 H_2(f)}. \quad (7)$$

Its obvious effect is to restore the received signal to what it would have been without multipath fading and space diversity. In the process, it eliminates ISI but produces a noise enhancement proportional to the integral of $C(f)/|H(f)|^2$ (see Fig. 1), where $C(f)$ is the cosine roll-off function. We will analyze this equalizer in conjunction with the three space diversity options (nondiversity, selection diversity, and combining diversity) discussed in Section 3.2.

In evaluating selection diversity with equalization, we will again assume a theoretically optimum strategy: At all times, $(\beta_1, \beta_2) = (1, 0)$ or $(0, 1)$, whichever yields the lower bit error rate for the prevailing response pair, $(H_1(f), H_2(f))$. As before, we will not consider the issue of practical implementation.

In evaluating combining diversity with equalization, we will consider three practical strategies for adapting β_1 and β_2 . They are as follows:

1. The relative amplitude and phase of β_1 and β_2 are adjusted to maximize the ratio of signal power to noise power at the combiner output.

2. Only the relative phase of β_1 and β_2 is adjusted to maximize the same ratio, with $|\beta_1| = |\beta_2| = 1$.

3. β_1 and β_2 are adjusted to those values, $\tilde{\beta}_1$ and $\tilde{\beta}_2$, that would maximize ρ_D in the absence of equalization.

Missing from this list of strategies is the (β_1, β_2) pair that maximizes ρ_D in the presence of equalization. Unfortunately, the analytical solution for this case is somewhat intractable. Also, its realization in practice would probably require using data decisions, which might be unreliable when recovering from severe fades. Each of the above three strategies, by contrast, could be implemented using simple power measurements at the combiner output. The first two would involve a single, full-channel power measurement to drive the search over β_1 and β_2 , while the third would involve three or more inband power spectrum density measurements, as described in Ref. 15.

We emphasize that the third strategy is not optimal in the presence

of equalization. Indeed, both of the first two strategies will be shown to yield superior outage statistics. Nonetheless, we will continue to refer to the gain pair $(\tilde{\beta}_1, \tilde{\beta}_2)$ as optimal, but with quotation marks added. Thus, the gain pair (β_1, β_2) is called optimal when there is no equalization and "optimal" when there is.

Why even consider the third strategy in the presence of equalization if it does not yield the best detection? Because it serves another beneficial purpose, namely, sharply limiting the signal dispersion as seen at the input to the equalizer. The virtue in doing this is that it simplifies the circuitry needed to approximate the performance of ideal adaptive equalizers. For example, a digital equalizer might require fewer taps and/or lesser quantizing resolution by operating on input signals with less dispersion. Moreover, Gersho has shown that equalizer convergence speed improves as input signal dispersion decreases.³⁰

Table I summarizes this section by giving, for each combination of space diversity and equalization to be considered, the assumed criteria for adapting (β_1, β_2) .

IV. ANALYSIS

4.1 Receiver detection measure

In current digital radio systems, the symbol rate $(1/T)$ is typically 75 percent of the channel bandwidth (W) and α is typically between 0.25 and 0.5. Our simulations will assume typical symbol rates but, for convenience, will show results for $\alpha = 0$. This case is especially easy to treat because the cosine roll-off function, $C(f)$, reduces to a rectangle on $[-1/2T, 1/2T]$ and this simplifies the analysis and computation of ISI power. At the same time, this approach leads to slightly

Table I—Criteria for adapting (β_1, β_2) under different space diversity/equalization approaches

Space Diversity	Equalization	
	None	Reciprocal Equalizer
None	$(\beta_1, \beta_2) = (1, 0)$	$(\beta_1, \beta_2) = (1, 0)$
Selection diversity	$(\beta_1, \beta_2) = (1, 0)$ or $(0, 1)$, whichever yields larger ρ_D	$(\beta_1, \beta_2) = (1, 0)$ or $(0, 1)$, whichever yields larger ρ_D
Combining diversity	Optimal solution:* $(\beta_1, \beta_2) = (\tilde{\beta}_1, \tilde{\beta}_2)$	<ol style="list-style-type: none"> $\beta_1 = 1 + j0; \beta_2$ and $\text{Arg}\{\beta_2\}$ set to maximize combiner output s/n $\beta_1 = 1 + j0; \beta_2 = 1; \text{Arg}\{\beta_2\}$ set to maximize combiner output s/n "Optimal" solution:* $(\beta_1, \beta_2) = (\tilde{\beta}_1, \tilde{\beta}_2)$

* $(\tilde{\beta}_1, \tilde{\beta}_2)$ is that gain pair for which ρ_D is maximized when there is no equalization.

pessimistic results since, for given $1/T$, ISI distortion decreases somewhat with α .^{29,31} Our assumption for α , then, will serve both to simplify the analysis/computations and to yield worst-case estimates of detection performance.

To begin, we define

$P_S \triangleq$ The squared signal sample (excluding ISI and noise) in either baseband stream when a data value of +1 or -1 is being detected;

$P_I, P_N \triangleq$ The mean square ISI and noise, respectively, associated with the periodic samples in either baseband stream;

and

$$\rho_D \triangleq P_S / (P_I + P_N). \quad (8)$$

The latter is the detection signal-to-distortion ratio and can be related to an upper bound on the bit error rate via²⁸

$$\text{Bit Error Rate} \leq 2 \exp(-\rho_D/2). \quad (9)$$

This upper bound is particularly conservative when ISI dominates because this distortion component is generally peak limited, in contrast to Gaussian noise.

We first consider a receiver with no adaptive equalization ($H_{\text{eq}}(f) = 1 + j0$, in Fig. 1). For $\alpha = 0$, ρ_D can be shown to be

$$\rho_D = \frac{3}{M-1} \cdot \left\{ \frac{\text{Max}_{t_s}(|\bar{H}|^2)}{[|\bar{H}|^2 - \text{Max}_{t_s}(|\bar{H}|^2)] + (|\beta_1|^2 + |\beta_2|^2)/\text{CNR}} \right\} \text{ (No Equalizer)}, \quad (10)$$

where*

$$\bar{H} = \int H(f) e^{j\omega t_s} df T; \quad (11)$$

$$|\bar{H}|^2 = \int |H(f)|^2 df T; \quad (12)$$

CNR is the unfaded carrier-to-noise ratio per diversity branch in a bandwidth $1/T$ (typically close to 10^6 , or 60 dB); $H(f)$ is given by (1); and t_s is the timing epoch. We assume the latter to be optimal, i.e., that value for which ρ_D is maximized. This assumption is implicit in (10), where t_s is specified to maximize $|\bar{H}|^2$.

* In this analysis, all integrals have limits $-1/2T$ and $1/2T$. For convenience, we omit them from the equations.

As for \bar{H} , it represents a complex gain for the signal vector sampled in each data interval. The imaginary part represents interference into each baseband rail from the cross-rail data in the same interval. Optimal carrier recovery amounts to multiplying \bar{H} by a phase factor, $e^{j\theta}$, that makes it real and thus eliminates this interference. The resulting θ is the optimal carrier phase and the resulting signal gain is $|\bar{H}|$, as used in (10). In the case of optimal combining diversity, the optimal θ is implicitly realized in the course of optimizing β_1 and β_2 . For both nondiversity and selection diversity, we will assume an optimal carrier phase without specifying how it is achieved.

Now we consider a receiver with an ideal reciprocal equalizer. The signal-restoring property of this equalizer is such that only thermal noise enhancement modifies ρ_D from its unfaded value. Thus, for the assumed system,

$$\rho_D = \frac{3}{M-1} \left\{ \frac{\text{CNR}}{(|\beta_1|^2 + |\beta_2|^2) \int \frac{dfT}{|H(f)|^2}} \right\} \text{(Reciprocal Equalizer)}. \quad (13)$$

The receiver detection measure we will use, and for which we will find probability distributions, is the bracketed quantity in (10) or (13). Thus, we define

$$\Gamma \triangleq \left(\frac{M-1}{3} \right) \rho_D, \quad (14)$$

which can be viewed as a normalized signal-to-distortion ratio. It is a function solely of the response pair ($H_1(f)$, $H_2(f)$); the gain pair (β_1 , β_2); and the system parameters T and CNR. Combining (14) with (9), we can upperbound the Γ required in an M -level system to achieve a specified bit error rate, BER_o :

$$\Gamma_o = \frac{2}{3} (M-1) \ln(2/\text{BER}_o). \quad (15)$$

Decibel values of Γ_o for various combinations of M and BER_o are given in Table II. We will use these values later in assessing the simulation results.

Table II—Decibel values of Γ_o for various M and BER_o

BER _o	M		
	16	64	256
10 ⁻³	18.81	25.04	31.11
10 ⁻⁴	19.96	26.19	32.26
10 ⁻⁵	20.87	27.10	33.17
10 ⁻⁶	21.62	27.85	33.92

4.2 Formulas for β_1 and β_2

We now present the appropriate gain pair (β_1, β_2) as a function of $(H_1(f), H_2(f))$ for each of the various receiver processing approaches.

In the cases of nondiversity and selection diversity, the specifications are simple: For both equalized and nonequalized receivers,

$$(\beta_1, \beta_2) = \begin{cases} (1, 0); & \text{(No Diversity)} \\ (1, 0) \text{ or } (0, 1), & \\ \text{whichever maximizes } \Gamma; & \text{(Selection Diversity).} \end{cases} \quad (16)$$

In the case of optimal combining diversity without equalization, the analysis is a bit complicated. The initial step is to rewrite (10) as

$$\rho_D = \frac{3}{M-1} \cdot \left\{ \text{Max}_{t_s} \left[\text{Max}_{(\beta_1, \beta_2)} \left\{ \frac{|\bar{H}|^2}{[|\bar{H}|^2 - |\bar{H}|^2] + (|\beta_1|^2 + |\beta_2|^2)/\text{CNR}} \right\} \right] \right\}, \quad (10')$$

where $|\bar{H}|^2$ and $|\bar{H}|^2$ are quadratic in β_1 and β_2 and $|\bar{H}|^2$ is, in addition, a function of t_s . The maximization over β_1 and β_2 can be done analytically using the method outlined in Ref. 25. Briefly, the denominator (quadratic in β_1 and β_2) is minimized subject to the constraint that \bar{H} (linear in β_1 and β_2) equals $1 + j0$. This problem can be solved using Lagrange multipliers and leads to the following result:

$$\tilde{\beta}_1 = \text{NUM}_1/\Delta; \quad \tilde{\beta}_2 = \text{NUM}_2/\Delta, \quad (17)$$

where

$$\text{NUM}_1 = \overline{H_1^* e^{j\omega t_s}} - \text{CNR}(\overline{H_2^* e^{j\omega t_s}} \overline{H_1^* H_2} - \overline{H_1^* e^{j\omega t_s}} \overline{|H_2|^2}); \quad (18)$$

$$\text{NUM}_2 = \overline{H_2^* e^{j\omega t_s}} - \text{CNR}(\overline{H_1^* e^{j\omega t_s}} \overline{H_1 H_2^*} - \overline{H_2^* e^{j\omega t_s}} \overline{|H_1|^2}); \quad (19)$$

$$\Delta = \text{NUM}_1 + \text{NUM}_2; \quad (20)$$

and $\overline{(\quad)}$ is the average over f from $-1/2T$ to $1/2T$. (Because the processing is linear, an equivalent form of the solution is $\tilde{\beta}_1 = 1 + j0$ and $\tilde{\beta}_2 = \text{NUM}_2/\text{NUM}_1$. For convenience only, we will use the solution as presented above.)

These solutions for $\tilde{\beta}_1$ and $\tilde{\beta}_2$ maximize ρ_D for a given value of t_s . In our computations, the above formulas were used to obtain the inner maximum in (10'), and then a numerical search over t_s was performed to find the outer maximum.

Finally, we consider the use of combining diversity *with* adaptive equalization. As noted in Section 3.3, three distinct strategies for adapting β_1 and β_2 have been considered. Mathematical formulas for β_1 and β_2 under these strategies are as follows:

1. $\beta_1 = 1 + j0$ and both the magnitude and phase of β_2 are adjusted to maximize the ratio of signal power to noise power at the combiner output. The solutions are

$$\text{Arg}\{\beta_2\} = \tan^{-1}[-\text{Im}(H_1^*H_2)/\text{Re}(\overline{H_1^*H_2})] \quad (21)$$

and

$$|\beta_2| = \frac{\gamma_a + \sqrt{\gamma_a^2 + 4\gamma_b}}{2\sqrt{\gamma_b}}, \quad (22)$$

where

$$\gamma_a = \overline{|H_2|^2} - \overline{|H_1|^2} \quad (23)$$

and

$$\gamma_b = [\text{Re}(\overline{H_1^*H_2})]^2 + [\text{Im}(\overline{H_1^*H_2})]^2. \quad (24)$$

2. $\beta_1 = 1 + j0$ and $|\beta_2| = 1$, with the phase of β_2 adjusted to maximize the same ratio. The solution for $\text{Arg}\{\beta_2\}$ is again (21).

3. $\beta_1 = \tilde{\beta}_1$ and $\beta_2 = \tilde{\beta}_2$, (17) through (20), where t_s is that timing epoch for which ρ_D in (10') is globally maximized. (Again, $\beta_1 = 1 + j0$ and $\beta_2 = \tilde{\beta}_2/\tilde{\beta}_1$ would yield the same results.)

4.3 Channel outage probabilities

Our simulations lead to probability distributions for Γ over the ensemble of dual-channel response pairs. The probability distribution obtained for a given receiver technique will be denoted by $P(\Gamma)$, representing the expected fraction of multipath fading seconds for which the receiver detection measure lies below Γ . Thus, if Γ_o is the value that Γ must exceed for acceptable performance, then $P(\Gamma_o)$ is the conditional probability of channel outage (i.e., conditioned on the occurrence of fading). Further, if P_{req} is the maximum permissible value of that probability for meeting the system outage objective, the receiver technique can be considered adequate if and only if $P(\Gamma_o) \leq P_{\text{req}}$.

To apply this interpretation to our simulation results requires numerical data for both Γ_o and P_{req} . The former is derived in Section 4.1 and quantified, for various combinations of M and BER_o , in Table II. Derivations of P_{req} are outlined in the Appendix and quantified in Table III. These results are for short- and long-haul systems in the 4-, 6-, and 11-GHz bands, both with and without protection switching.

4.4 A performance measure for equalized receivers

A second performance measure for which we shall obtain probabil-

Table III—Values of P_{req} for various system conditions

Band	System Length	Protection Switching	
		None (Unprotected)	One Protection Channel
4 GHz ($W = 20$ MHz)	Long haul	2.84×10^{-4}	6.88×10^{-3}
	Short haul	4.54×10^{-3}	2.75×10^{-2}
6 GHz ($W = 30$ MHz)	Long haul	1.89×10^{-4}	6.88×10^{-3}
	Short haul	3.03×10^{-3}	2.75×10^{-2}
11 GHz ($W = 40$ MHz)	Long haul	1.03×10^{-4}	5.87×10^{-3}
	Short haul	1.65×10^{-3}	2.35×10^{-2}

ity distributions is the *range* of $|H(f)|$ over the Nyquist bandwidth, $[-1/2T, 1/2T]$. Thus,*

$$R \triangleq \left\{ \frac{\text{Max}_f |H(f)|}{\text{Min}_f |H(f)|} \right\}; \quad |f| \leq 1/2T. \quad (25)$$

The quantity R is of interest in cases where adaptive equalization is used, for reasons given in Section 3.3. It is this index of input signal dispersion that has been shown to provide a measure of convergence speed.³⁰ We will see that its probability distribution, $P(R)$, depends strongly on diversity approach.

V. SIMULATION STUDY AND RESULTS

5.1 Simulation approach

A computer program was written that simulates response pairs $(H_1(f), H_2(f))$ in accordance with the statistical model discussed in Section II and, for each simulation trial, analyzes the receiver techniques described in Section III using the methods of Section IV.

The system parameters studied were channel bandwidth (W) and unfaded carrier-to-noise ratio (CNR). As discussed in Section 4.1, the symbol rate was taken to be 75 percent of the bandwidth (i.e., $1/T = 15, 22.5, \text{ and } 30$ megabaud, respectively, for $W = 20, 30, \text{ and } 40$ MHz) and the cosine roll-off factor was taken to be zero. The values used for CNR ranged from 51 to 67 dB, with 63 dB taken to be the most "typical" value. The channel parameters studied were $\delta t, p$ and ρ , as discussed in Section 3.4, with the "typical" values taken to be 4 ns, 0.5, and 0.65, respectively.

For each combination of receiver technique and parameter set, either or both of the quantities Γ and R were computed for each of four

* We use the Nyquist bandwidth in defining this measure, noting that it is the 6-dB bandwidth of the system response for *any* value of the cosine roll-off factor.

thousand response pairs. From the population of quantities obtained for each such "simulation run," we then computed the appropriate probability distribution(s), i.e., $P(\Gamma)$ and/or $P(R)$.

The response pair for each trial was obtained via Monte Carlo generation of the a 's and b 's in (4), according to the assumed model. The same set of four thousand response pairs was used in each simulation run, except for minor differences related to changed values for δt , p , and ρ . Our purpose was to permit comparisons among receiver techniques for the same fading channel ensemble.

We confirmed, in two ways, that the ensemble of four thousand response pairs generated suitably represents the model, namely, (1) by computing the ensemble statistics of a_1 , a_2 , $|b_1|$, $|b_2|$, etc., and comparing them with those predicted by the model; and (2) by computing $P(\Gamma)$ and $P(R)$ for several different sets of response pairs and examining the agreement between them. We thereby established that the simulations are accurate and consistent down to probability values of 0.005 or lower, which is sufficient for our purposes.

5.2 Results for combining diversity with no equalization

This is the receiver technique of greatest interest to this study, as we wish to determine the multipath-combatting potentialities of those routes equipped with space diversity antennas. We computed $P(\Gamma)$ for this case over a wide range of system and channel parameters.

5.2.1 System parameters: W and CNR

Figure 2 shows curves for $P(\Gamma)$ with W as a parameter. All other parameters are assumed to have "typical" values, i.e., CNR = 63 dB, $\delta t = 4$ ns, $p = 0.5$, and $\rho = 0.65$. The fairly strong influence of W is no surprise. The benefit gained by linearly combining $H_1(f)$ and $H_2(f)$ depends on how effectively this combining cancels their dispersive components. Cancellation is most readily achieved when these components are linear in $j\omega$ over the bandwidth, a condition most nearly approximated at smaller W .¹⁵ As W increases, higher-order terms in $j\omega$ become prominent and a linear sum of two responses is less able to cancel the dispersions. In this light, we see that three or more diversity branches might yield significant gains at larger bandwidths.

To apply the results of Fig. 2 to a concrete example, we assume a protected short-haul 64-QAM system with a BER objective of 10^{-6} . Table II shows that $\Gamma_o \approx 28$ dB for this case, and Table III shows that $P(\Gamma_o)$ must lie below 0.0275, 0.0275, and 0.0235 for $W = 20$, 30, and 40 MHz, respectively. We thus see that using optimal combining diversity without equalization would suffice to meet outage objectives at 4, 6, and 11 GHz (the latter case being somewhat marginal). To meet *long-haul* objectives, however (see Table III), or to operate

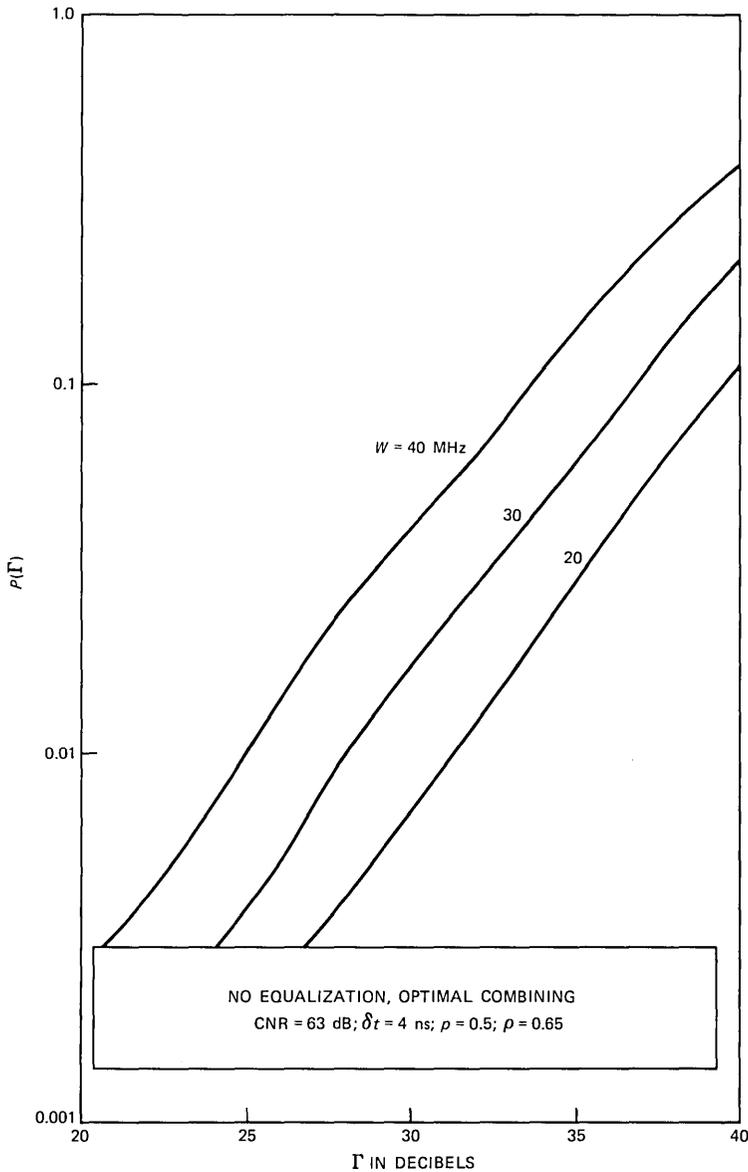


Fig. 2—Probability distributions for Γ in receivers with optimal diversity combining and no equalization. The parameter is channel bandwidth, W . All other system and channel parameters have “typical” values: CNR = 63 dB, $\delta t = 4$ ns, $\rho = 0.5$, and $\rho = 0.65$.

without protection switching or at 256 levels, optimal combining alone would be marginal or downright inadequate, at least for 6- and 11-GHz systems.

To study the influence of CNR, Fig. 3 shows $P(\Gamma)$ for $W = 30$ MHz

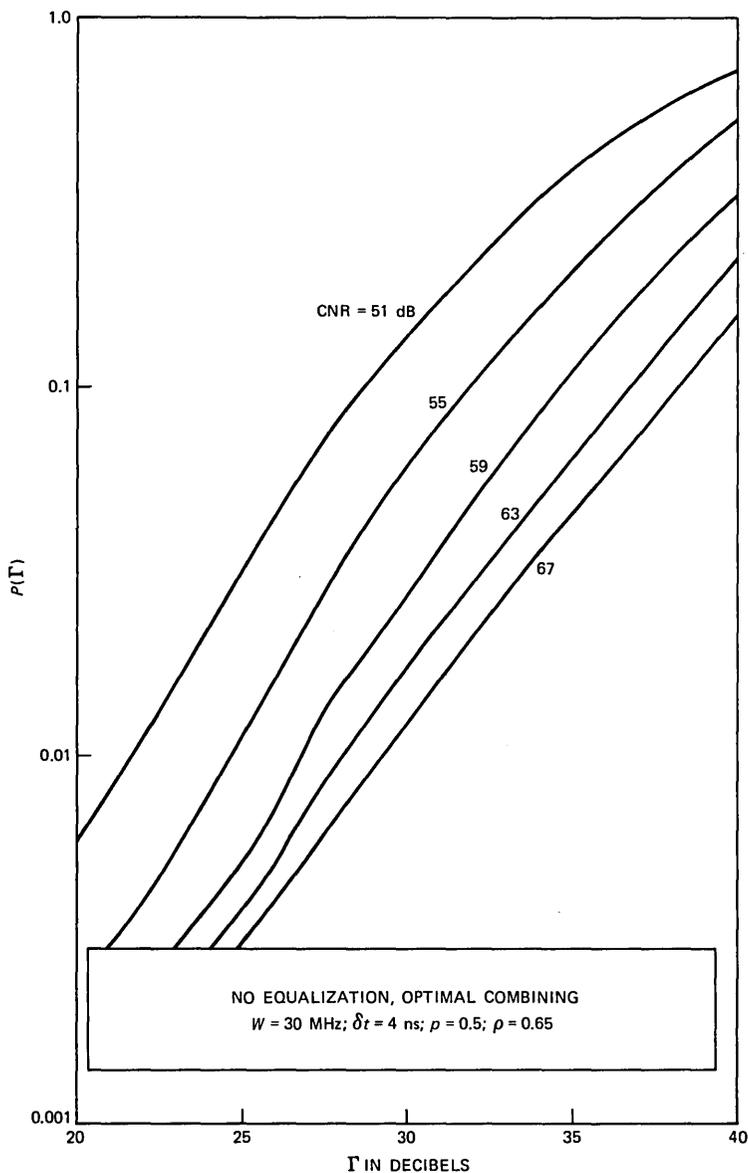


Fig. 3—Probability distributions for Γ in receivers with optimal diversity combining and no equalization. The parameter is CNR. Channel bandwidth $W = 30$ MHz, and all other system and channel parameters have “typical” values.

with CNR parameterized over the range from 51 to 67 dB. The trend is similar for all other W . The sensitivity of the results to CNR is seen to increase as this quantity decreases; this is because noise becomes an increasingly significant factor (as opposed to intersymbol interference) as CNR decreases.

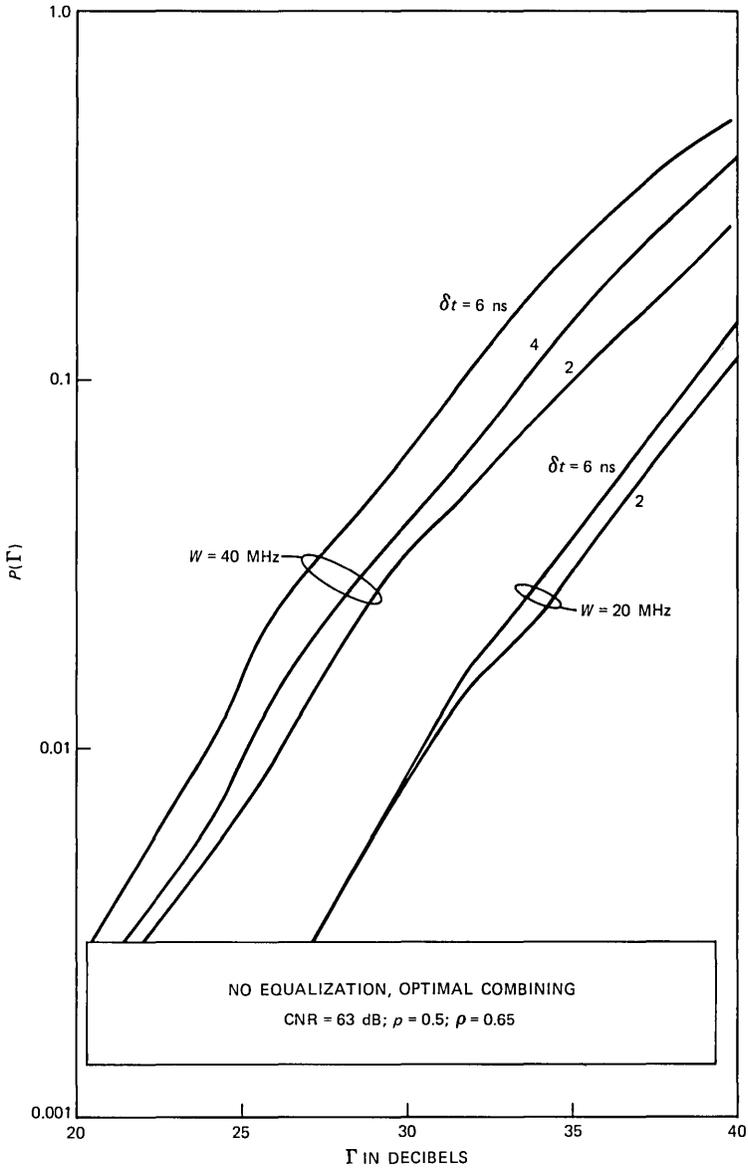


Fig. 4—Probability distributions for Γ in receivers with optimal diversity combining and no equalization. The primary parameter is δt , with results given for $W = 20$ and 40 MHz. All other system and channel parameters have “typical” values.

5.2.2 Channel parameters: δt , ρ , and ρ

Figure 4 shows $P(\Gamma)$ for $\delta t = 2, 4$ and 6 ns for each of two bandwidths ($W = 20$ and 40 MHz). Not surprisingly, the sensitivity of performance to this parameter is greater at larger bandwidths. In the remainder of

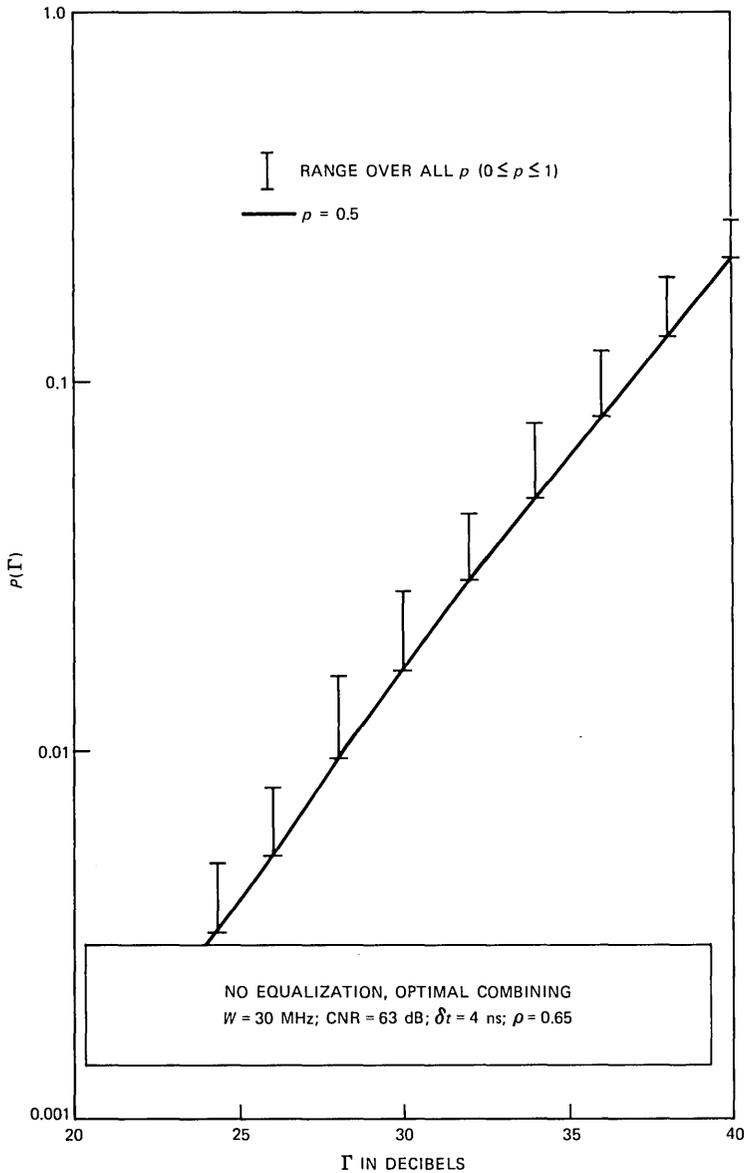


Fig. 5—Probability distributions for Γ in receivers with optimal diversity combining and no equalization. The vertical bars span the range from minimum to maximum over the p -range [0, 1.0] and the solid curve is for $p = 0.5$. Channel bandwidth $W = 30$ MHz and all other system and channel parameters have “typical” values.

this study, we will confine ourselves to the assumed “typical” value of 4 ns.

Figure 5 shows the effect of p , the minimum-phase probability, when $W = 30$ MHz. The vertical bars show the range of $P(\Gamma)$ as p goes from

0 to 1.0, and the solid curve is for the assumed “typical” value of 0.5. Similar results apply for other values of W . The highest point on each bar is for either $p = 0$ (all $H_1(f)$ and $H_2(f)$ minimum phase) or $p = 1.0$ (all $H_1(f)$ and $H_2(f)$ nonminimum phase), and the lowest point is for $p = 0.5$. The explanation for these trends is simple: The conditions $p = 0$ and $p = 1.0$ provide, on the average, for the greatest shape similarities between $H_1(f)$ and $H_2(f)$, while the condition $p = 0.5$ provides, on the average, for the greatest dissimilarities. In any event, Fig. 5 shows that a precise quantification of this parameter is not important for the situation under study.

We also examined the effect of the correlation factor ρ , for $W = 30$ MHz, over the range from 0 to 1.0 (we did not seriously consider anticorrelation). The spread was found to be very small, suggesting that the amplitude factors a_1 and a_2 in (4) are of minor import when CNR is high and that, therefore, the channel parameter ρ need not be precisely known. Clearly, then, poor performance occurs primarily when the relative variations of $H_1(f)$ and $H_2(f)$ are at once highly dispersive and (by random chance) highly similar. We invoked this observation to derive P_{req} in the Appendix.

5.3 Comparisons with non- and selection diversity

Figure 6 shows $P(\Gamma)$ for non-, selection, and optimal combining diversity without equalization. These comparisons are given for two bandwidths ($W = 20$ and 40 MHz). The parameters CNR, δt , p , and ρ are assigned their “typical” values. For non- and selection diversity, the values of δt and p have no influence on $P(\Gamma)$. The values of CNR and ρ were found to have some impact (the statistics improve slightly with increasing CNR for both non- and selection diversity, and improve slightly with decreasing ρ for selection diversity), but we did not study these relationships in detail.

For both bandwidths, receiver performance without space diversity or equalization is seen to be disastrous, as expected. Selection diversity improves matters noticeably, but is still far from adequate. The improvement in going to optimal combining diversity is seen to be dramatic.

5.4 Results for receivers with equalization

For receivers using ideal reciprocal equalizers, we have obtained both $P(\Gamma)$ and $P(R)$. We can thus examine trade-offs between detection performance (measured by $P(\Gamma)$) and dispersion at the equalizer input (measured by $P(R)$). For all diversity arrangements, both $P(\Gamma)$ and $P(R)$ depend weakly (if at all) on δt and p . In addition, $P(\Gamma)$ depends very weakly on W and $P(R)$ depends very weakly on CNR. In several cases, the correlation factor ρ was found to have a discernible

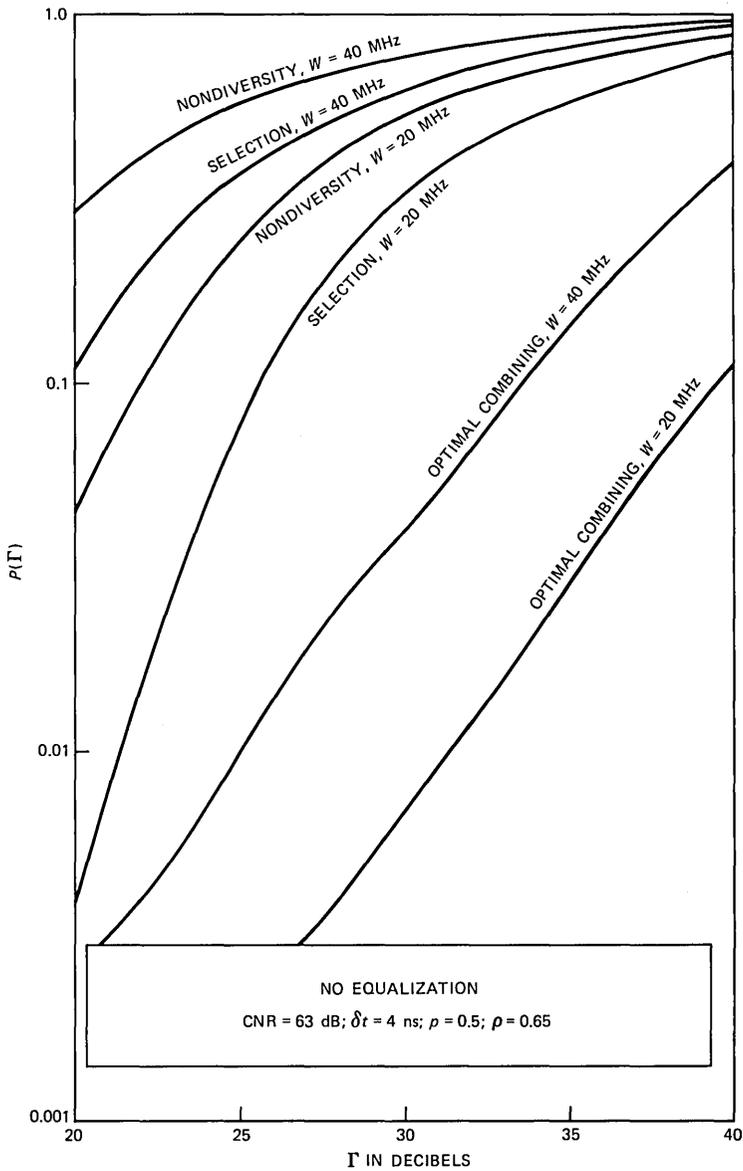


Fig. 6—Probability distributions for Γ in receivers with no equalization. Results are for non-, selection, and optimal combining diversity, for $W = 20$ and 40 MHz. All other system and channel parameters have “typical” values.

impact. This is because ideal equalizers tend to capitalize on total signal power and lower values of ρ (greater independence between a_1 and a_2) correspond to statistically higher values of total signal power. For this part of the study, we confined ourselves to the “typical” and somewhat conservative value of 0.65.

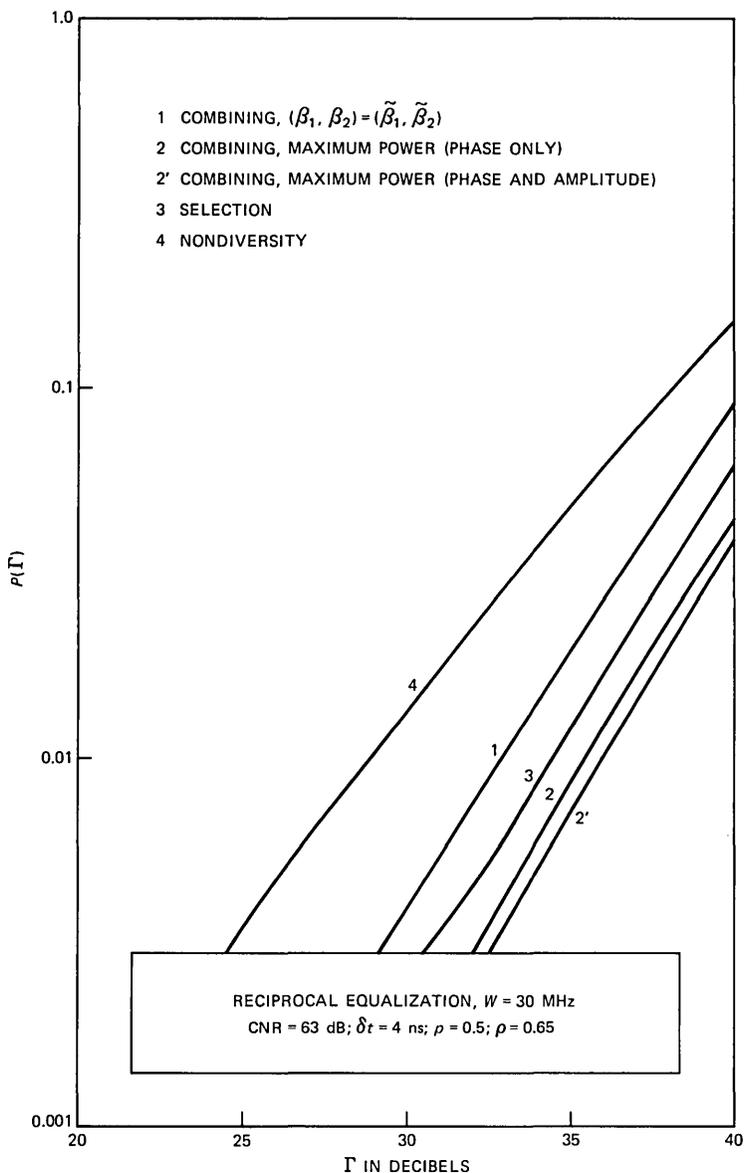


Fig. 7—Probability distributions for Γ in receivers with ideal reciprocal equalization. Results are for non-, selection, and combining diversity, with three gain-setting algorithms considered for the latter. Channel bandwidth $W = 30$ MHz and all other system and channel parameters have “typical” values.

5.4.1 Probability distributions for Γ

Figure 7 shows $P(\Gamma)$ for non-, selection and combining diversity, with three distinct gain-setting algorithms considered for the combining case. Some observations are as follows:

1. While the results shown are for $W = 30$ MHz, they would barely change for other bandwidths. With this in mind, we can see from Figs. 7 and 2 that a reciprocal equalizer alone would be less effective than an optimal diversity combiner alone for $W = 20$ MHz, but more effective for $W \geq 30$ MHz.

2. For reciprocal equalizers, Γ is related strictly to thermal noise enhancement. Thus, this quantity scales directly with CNR (63 dB in Fig. 7), that is, for every decibel of reduction (or increase) in CNR, all curves would shift by one decibel to the right (or left).

3. For combining diversity, the best results shown are for the (β_1, β_2) pairs that maximize combiner output power. Controlling both amplitude and phase is only slightly superior to controlling phase alone, which is only slightly superior to using selection diversity. Using the gain pair $(\tilde{\beta}_1, \tilde{\beta}_2)$, while optimal in the absence of equalization, is seen to be fourth best in this case. We now examine the potential benefit of using these suboptimal gains.

5.4.2 Probability distributions for R

Figure 8 shows $P(R)$ for $W = 20$ and 40 MHz. For each bandwidth, the results shown are for nondiversity, selection diversity, maximum-power combining diversity (phase-only adjustments), and "optimal" combining diversity, $(\beta_1, \beta_2) = (\tilde{\beta}_1, \tilde{\beta}_2)$. The disparities among approaches are seen to be strong, the best results corresponding to "optimal" combining.

VI. CONCLUSION

The results obtained in this study are subject to the usual uncertainties associated with finite simulations and a less-than-universal statistical model of dual-channel fading. Nevertheless, some general conclusions are possible, as follows:

1. The use of appropriate space diversity combining (e.g., the scheme described in Ref. 15) could eliminate the need in some links for post-combiner equalization.

2. In links where adaptive equalization is used, such a scheme could sharply reduce the signal dispersion at the equalizer input, thereby simplifying the equalizer design.

3. For purposes of assessing various receiver approaches, the dual-channel multipath fading model reported by Rummler is, for the most part, well specified. The precise value of the minimum-phase probability (p) appears to be of minor consequence, as is the precise value of the correlation factor (ρ) between the decibel values of a_1 and a_2 . The delay parameter (δt) introduced here could have a modest impact if larger than 4 ns, particularly for the larger channel bandwidths.

4. Finally, the numerical results affirm that space diversity combin-

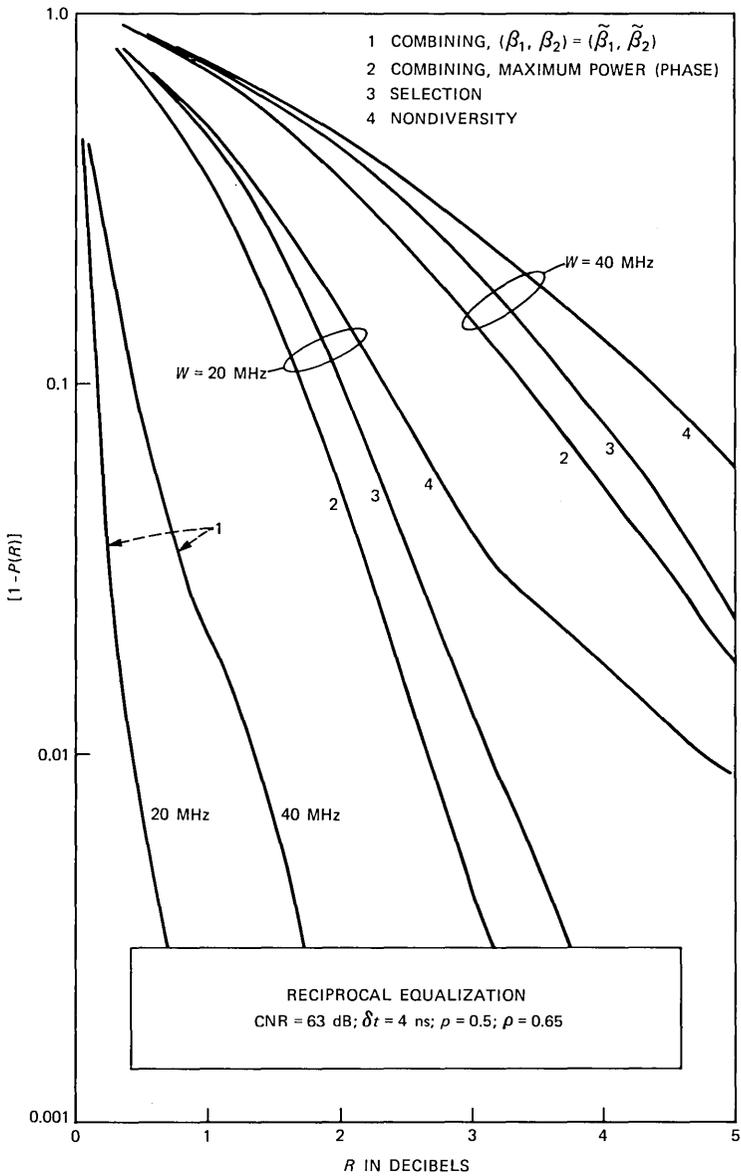


Fig. 8—Probability distributions for R for several diversity approaches, for $W = 20$ and 40 MHz. All other system and channel parameters have “typical” values.

ing and adaptive equalization in tandem comprise a formidable combination of receiver techniques in microwave digital radio.

VII. ACKNOWLEDGMENTS

Lisa J. (Domenico) Case made a major contribution to this study

with her help in programming and executing the computer simulations. We are also grateful to Gerard J. Foschini for his valuable suggestions on methods of analysis.

REFERENCES

1. C. W. Anderson, S. G. Barber, and R. N. Patel, "The Effect of Selective Fading on Digital Radio," *IEEE Trans. Commun.*, COM-27, No. 12 (December 1979), pp. 1870-5.
2. S. Komaki et al., "Characteristics of a High Capacity 16 QAM Digital Radio System in Multipath Fading," *IEEE Trans. Commun.*, COM-27, No. 12 (December 1979), pp. 1854-61.
3. Y. Y. Wang, "Simulation and Measured Performance of a Space Diversity Combiner for 6 GHz Digital Radio," *IEEE Trans. Commun.*, COM-27, No. 12 (December 1979), pp. 1896-907.
4. S. Komaki, Y. Okamoto, and K. Tajima, "Performance of 16-QAM Digital Radio System Using New Space Diversity," Paper 52.2, Int. Conf. Commun., June 8-12, 1980, Seattle, Washington.
5. T. Murase, K. Morita, and S. Komaki, "200 Mb/s 16-QAM Digital Radio System With New Countermeasure Techniques for Multipath Fading," Paper 46.1, Int. Conf. Commun., June 14-18, 1981, Denver, Colorado.
6. P. Hartmann and B. Bynum, "Design Consideration for an Extended Range Adaptive Equalizer," Paper 46.5, Int. Conf. Commun., June 14-18, 1981, Denver, Colorado.
7. A. J. Giger and W. T. Barnett, "Effects of Multipath Propagation on Digital Radio," *IEEE Trans. Commun.*, COM-29, No. 9 (September 1981), pp. 1345-52.
8. P. D. Karabinis, "Maximum-Power and Amplitude-Equalizing Algorithms for Phase Control in Space Diversity Combining," *B.S.T.J.*, 62, No. 1, Part 1 (January 1983), pp. 63-89.
9. R. Iyer et al., "Effects of Space Diversity Reception on Multipath Induced Distortions in Terrestrial Microwave Channels," Paper 3.B.5, Int. Conf. Commun., June 13-17, 1982, Philadelphia, PA.
10. A. Ricagni and T. Testi, "IF Combining Techniques for Space Diversity in Analog and Digital Radio Systems," Paper 4.B.6, Int. Conf. Commun., June 13-17, 1982, Philadelphia, PA.
11. R. A. Nichols, "An IF Combiner for Digital and Analog Radio Systems," Paper 4.B.7, Int. Conf. Commun., June 13-17, 1982, Philadelphia, PA.
12. C. P. Bates and M. A. Skinner, "Impact of Technology on High Capacity Digital Radio Systems," Paper F2.3 Int. Conf. Commun., June 19-22, 1983, Boston, MA.
13. T. Noguchi et al., "6 GHz 135 MBPS Digital Radio System With 64 QAM Modulation," Paper F2.4, Int. Conf. Commun., June 19-22, 1983, Boston, MA.
14. P. R. Hartmann and J. A. Crossett, "135 MBS-6 GHz Transmission System Using 64-QAM Modulation," Paper F2.6, Int. Conf. Commun., June 19-22, 1983, Boston, MA.
15. Y. S. Yeh and L. J. Greenstein, "A New Approach to Space Diversity Combining in Microwave Digital Radio," *AT&T Tech. J.*, this issue.
16. W. D. Rummler, "A Statistical Model of Multipath Fading on a Space Diversity Radio Channel," *B.S.T.J.*, 61, No. 9, Part 1 (November 1982), pp. 2185-219.
17. W. D. Rummler, "A Rationalized Model for Space and Frequency Diversity Line-of-Sight Radio Channels," Paper E27, Int. Conf. Commun., June 19-22, 1983, Boston, MA.
18. W. D. Rummler, "A New Selective Fading Model: Application to Propagation Data," *B.S.T.J.*, 58, No. 7 (May-June 1979), pp. 1037-71.
19. W. D. Rummler, "More on the Multipath Fading Channel Model," *IEEE Trans. Commun.*, COM-29, No. 3 (March 1981), pp. 346-52.
20. L. J. Greenstein and B. A. Czekaj, "A Polynomial Model for Multipath Fading Channel Responses," *B.S.T.J.*, 59, No. 7 (September 1980), pp. 1197-225.
21. R. W. Lucky, J. Salz, and E. J. Weldon, Jr., *Principles of Data Communication*, New York, NY: McGraw-Hill, 1968. (In particular, see pp. 50-51 for a description of cosine roll-off pulses and spectra.)
22. B. Widrow, "A Comparison of Adaptive Algorithms Based on the Methods of Steepest Descent and Random Search," *IEEE Trans. Ant. and Prop.*, AP-24, No. 5 (September 1976), pp. 615-37.

23. R. W. Lucky, J. Salz, and E. J. Weldon, Jr., *Principles of Data Communication*, Chapter 6.
24. L. J. Greenstein and D. Vitello, "Digital Radio Receiver Responses for Combating Frequency-Selective Fading," *IEEE Trans. Commun.*, COM-27, No. 4 (April 1979), pp. 671-81.
25. R. W. Lucky, J. Salz, and E. J. Weldon, Jr., *Principles of Data Communication*, Section 5.1.
26. P. Monsen, "Feedback Equalization for Fading Dispersive Channels," *IEEE Trans. Inform. Theory*, IT-17 (January 1971), pp. 56-64.
27. M. S. Mueller and J. Salz, "A Unified Theory of Data-Aided Equalization," *B.S.T.J.*, 60, No. 9 (November 1981), pp. 2039-63.
28. G. J. Foschini and J. Salz, "Digital Communications Over Fading Radio Channels," *B.S.T.J.*, 62, No. 2, Part 1 (February 1983), pp. 429-56.
29. W. C. Wong and L. J. Greenstein, "Multipath Fading Models and Adaptive Equalizers in Microwave Digital Radio," *IEEE Trans. Commun.*, COM-32, No. 8 (August 1984), pp. 928-34.
30. A. Gersho, "Adaptive Equalization of Highly Dispersive Channels for Data Transmission," *B.S.T.J.*, 48, No. 1 (January 1969), pp. 55-70.
31. N. Amitay and L. J. Greenstein, "Multipath Outage Performance of Digital Radio Receivers Using Finite-Tap Adaptive Equalizers," *IEEE Trans. Commun.*, COM-32, No. 5 (May 1984), pp. 597-608.

APPENDIX

Derivation of P_{req}

We consider a specific W -Hz radio channel carrying digital traffic. It may be protected or unprotected and it may have dual space diversity or no diversity. Discernible multipath fading is assumed to exist, on one or both of the possible diversity branches, for T_o seconds per hop per year, with T_o given by (3). We define the following probabilities within that subset of time:

$P_o \triangleq$ The probability of outage per hop as experienced by a user of the channel;

$P_1 \triangleq$ The probability of outage per hop as experienced by the channel itself (note that $P_o < P_1$ when protection switching is used);

$P_{\text{req}} \triangleq$ The value of P_1 required to achieve a specified value of P_o .

We now make the following assumptions:

1. The yearly outage objective on a total system route is 0.005 percent (one way), apportioned uniformly among the route hops.
2. The path length and terrain factor for each hop are taken to have average values: $D = 25$ miles and $c = 1$ in (3).
3. Within any protection switching section, discernible multipath fading occurs on at most one hop at a time.
4. The subset of multipath fading seconds on each hop is congruent over all W -Hz channels within the same common carrier band.
5. Within that subset of multipath fading seconds, outage events are statistically independent from one W -Hz channel to another.

The third, fourth, and fifth assumptions are germane to calculations

of P_{req} for systems with protection switching. The third assumption is slightly liberal, in that it leads to overestimation of P_{req} ; the fourth assumption, by the same token, is somewhat conservative; and the fifth assumption is discussed later.

In systems without protection switching, P_o and P_1 are the same, and so P_{req} is just the value of P_o dictated by the first two assumptions. For a long-haul system (maximum length = 4000 miles), the outage budget will be met if each hop averages 10 seconds of outage per year; for a short-haul system (maximum length = 250 miles), that number is 160 seconds. Dividing each number by T_o yields the required P_{req} . The results are shown in Table II for the 4-, 6- and 11-GHz bands.

For systems *with* protection switching, the issue is more complicated, involving as it does the joint outage probabilities for channels within the same band. For most receiver techniques, outages in neighboring channels during multipath fading are *not* independent events, in which cases the fifth assumption cited above is too liberal. It might be quite valid, however, for the special case of optimal combining diversity without adaptive equalization. The reasoning is as follows: In such receivers, outage occurs primarily when $H_1(f)$ and $H_2(f)$ have similar shapes and, thus, dispersion cannot be reduced without excessive signal loss. There is no reason, however, why such a similarity between two uncorrelated responses would occur in *two* frequency channels at the same time, except by chance; hence, the independence assumption.

The case of optimal combining diversity without adaptive equalization is of major interest here, as we mean to explore the limits of its applicability in digital radio. Thus, we use the independence assumption in deriving P_{req} , speculating that the result is accurate for that type of receiver and somewhat elevated for most others.

To proceed, let N be the number of channels in a given band and assume that one of them is used for protection. Given an outage in a traffic-bearing channel, the probability that it will not find the protection channel available is

$$P_2 = P_1 + \left(\frac{N-2}{2} \right) P_1 + \text{smaller terms.}$$

The first term is the probability that the protection channel itself is out; the second term is the probability that the protection channel is *not* out but one of the other $N-2$ channels *is* and switches first; and the smaller terms (neglected here) have to do with three or more channels being out at the same time. Outage for a given user occurs if its original channel is out (probability P_1) *and* the protection channel is not available (probability P_2). Thus,

$$P_o = P_1 P_2 = \frac{N}{2} P_1^2.$$

We can now find the P_1 required to achieve a specified P_o . The latter values, as before, are $10/T_o$ and $160/T_o$, respectively, for long- and short-haul systems. The values of N are 12, 8, and 6, respectively, for the 4-, 6- and 11-GHz bands. The resulting values for P_{req} for each of various systems are shown in Table III.

AUTHORS

Larry J. Greenstein, B.S.E.E., 1958, M.S.E.E., 1961, and Ph.D. (Electrical Engineering), 1967, Illinois Institute of Technology; AT&T Bell Laboratories, 1970—. Mr. Greenstein currently heads the Radio Systems Research Department at Crawford Hill in Holmdel, NJ. His most recent work has dealt with communications satellites, mobile telephony, and microwave digital radio. His previous work was on digital encoding, digital filtering, and, at IIT Research Institute before 1970, airborne radar. Member, Eta Kappa Nu, Tau Beta Pi, and Sigma Xi; Senior Member, IEEE; Senior Technical Editor, IEEE Communications Magazine; co-recipient, IEEE Communications Society's 1984 Prize Paper Award in Communications Systems.

Yu-Shuan Yeh, B.S.E.E., 1961, National Taiwan University; M.S.E.E., 1964, Ph.D., 1966, University of California, Berkeley, CA; Harvard University, 1966–67; AT&T Bell Laboratories, 1967—. From 1961 to 1962, Mr. Yeh was an electronics officer in the Chinese Navy. He was a Research Fellow at Harvard University from 1966 to 1967 doing antenna research. In 1967 he joined AT&T Bell Laboratories and is currently a Supervisor in the Network Systems Research Department. His research interests include satellite communications, mobile radio, microwave digital radio, and data networks. Mr. Yeh holds over a dozen patents and is the recipient of two best paper awards from IEEE Transactions. Fellow, IEEE.

A Controlled Impedance Robot Gripper

By M. K. BROWN*

(Manuscript received June 5, 1984)

The design of robot grippers has traditionally involved pneumatics or gear reduction electric drives that do not allow for the control of the stiffness of the mechanism. The manipulation of fragile objects such as eggs or light bulbs requires the gripper to be able to close on the object with minimal impact forces, and yet maintain a static grip force sufficient to firmly handle the object. This paper describes a two-fingered controlled impedance gripper where the impedance is imparted by electrical means. The analog control system allows independent control of the effective mechanical mass and damping of each finger, as well as additional control of common-mode versus differential-mode response. The gripper is modeled by a computer simulation that consists of a set of nonlinear differential equations with time-varying feedback parameters. In the final form the model has ten degrees of mechanical freedom and eight electrical poles, i.e., an 18th-order nonlinear differential equation. The dynamic equations are described and results given are compared with actual results obtained from laboratory experiments with the gripper. An Appendix describes the method of solution used for the nonlinear differential equations. The accuracy of the simulation has been verified by measurements on the actual gripper. This analysis has resulted in the design of a gripper control system capable of providing controlled compliance and reduced finger impact forces while maintaining a quick response and firm grasp.

I. INTRODUCTION

Many compliant mechanisms have been built using mechanical means for obtaining the compliant properties. For example, the Remote Center Compliant manipulator (RCC), developed by Whitney

* AT&T Bell Laboratories.

Copyright © 1985 AT&T. Photo-reproduction for noncommercial use is permitted without payment of royalty provided that each reproduction is done without alteration and that the Journal reference and copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free by computer-based and other information-service systems without further permission. Permission to reproduce or republish any other portion of this paper must be obtained from the Editor.

and Nevins¹ at Draper Labs, utilizes specially designed springs in a configuration that allows free angular movement about a radial center while maintaining high stiffness along the radius. This device is particularly useful for inserting a pin in a hole or for analogous tasks, but the amount of compliance is fixed while in operation.

Similar devices have been developed by the Lord Corporation and others. One interesting device was designed by M. Cutkowski.² It replaces the fixed springs of the RCC device with fluid-filled elastomer bubbles. The compliance thus formed is nonlinear and adjustable by changing the fluid pressure. This feature makes the compliant wrist a potentially active device with bandwidth limited only by the rapidity with which fluid pressure may be changed.

Another compliant manipulator was designed by Hanafusa and Asada, who developed a three-fingered gripper with spring-driven fingers.³ The fingers are arranged in a circle and have rollers on the fingertips. Much of their work has involved determining stable gripping orientations for irregularly shaped objects.

Dynamic computer control of the amount of compliance and damping imparted to the gripper is desirable since this allows the gripper to adapt to various grasping tasks as they proceed without the need to exchange mechanical parts such as springs.⁴ For a two-fingered gripper, a "common-mode" compliance and "differential-mode" compliance can be made usefully different so that a gripper may use a delicate grasp and still hold an object firmly in place.

Figure 1 illustrates the construction of the two-fingered gripper. The fingers are mounted on linear bearing slides and driven through a rack and pinion by two dc servomotors. The two fingers are driven independently by two dc servomotors. Thus, the gripper can close from either side and need not be centered on an object before closing.

A particularly important feature of this mechanical design is the rack and pinion servomotor device. This mechanism is mechanically back drivable, meaning that a force applied to the finger will cause the motor to move. A typical worm drive mechanism, for example, would not have this characteristic because of high internal friction. This feature allows the motor drive mechanism to function as a mechanical impedance (under computer control) applying all of the appropriate effects of mass, damping, and compliance directly to the fingers.

To control the position of the gripper fingers, a means of measuring their positions is necessary. A capacitive measuring system was developed (see Ref. 5) that has the properties of low noise ($s/n > 90$ dB), low mass, and fast transient response. It consists of a metal rod sliding inside an insulated metal tube. The capacitance of the rod and tube varies as a function of the position of the finger. A capacitance measurement circuit produces a voltage proportional to

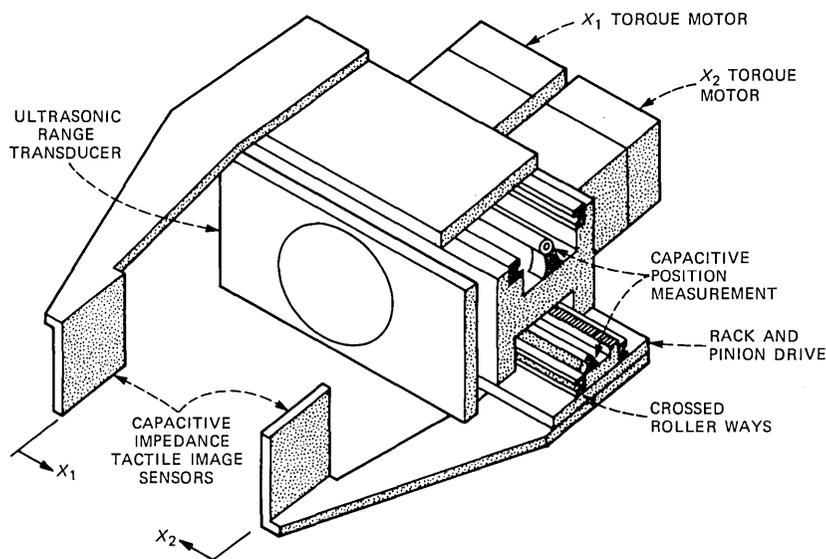


Fig. 1—Instrumented controlled impedance gripper.

the capacitance. The position measurement error developed by this system is less than 1 percent, while the step response rise time is about 150 μ s.

The dc servomotor performance can be modeled by an electrical and mechanical system consisting of an armature resistance and inductance, a back Electromotive Force (EMF), armature inertia, and bearing friction. This forms a third-order differential equation describing the current-voltage relationship as well as the torque output of the motor. Since torque is a quantity that must be controlled and torque is an essentially linear function of motor current, it is desirable to drive the motor with a current source. This reduces the motor equations to second order and greatly simplifies the design of a controller. Motor voltage is then described by an auxiliary equation.

All dynamic systems can be represented by a set of states or state variables which describe the system energies.⁶ The energies can be divided into two basic classes: static/potential and dynamic/kinetic, or analogous terms. For example, the motor state can be described by armature position and rotational velocity. State equations are typically written in matrix notation. In particular, most forms of dynamic systems can be represented by two matrix equations:

$$\dot{\mathbf{x}}(t) = \mathbf{A}(\mathbf{x}, t) \mathbf{x}(t) + \mathbf{B}(\mathbf{x}, t) \mathbf{u}(t) \quad (1)$$

$$\mathbf{y}(t) = \mathbf{C}(\mathbf{x}, t) \mathbf{x}(t) + \mathbf{D}(\mathbf{x}, t) \mathbf{u}(t) \quad (2)$$

where \mathbf{x} is the state variable vector and (1) is the "state equation." Matrix $\mathbf{A}(\mathbf{x}, t)$ describes the feedback part of the system and matrix $\mathbf{B}(\mathbf{x}, t)$ describes the feed forward part from the input vector $\mathbf{u}(t)$. The "output equation," (2), describes all auxiliary variables that may depend on the system state or input vector but do not affect the system state.

As an example, let us model the motor described earlier in state variable form, with a voltage source as the input. The mechanical state variables are motor shaft position and velocity, and electrical energy is best described by the current flowing in the inductance of the armature. Given these three state variables, the coefficients of the motor equation are

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -P/M & K_t/M \\ 0 & -K_e/L & -R/L \end{bmatrix} \quad (3)$$

$$\mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ 1/L \end{bmatrix}, \quad (4)$$

where K_t is the motor torque constant, M is the armature moment of inertia, P is the viscous friction coefficient, K_e is the inducted back EMF coefficient (EMF = $K_e \times \text{Vel}$), L is armature inductance, the state variable vector is

$$\mathbf{x} = \begin{bmatrix} \text{armature position } (\theta) \\ \text{armature velocity } (\dot{\theta}) \\ \text{armature current } I \end{bmatrix}, \quad (5)$$

and the input $u(t) = V_s(t)$. The motor torque T is a direct function of the armature current, giving for the output equation

$$T = y = \mathbf{C}\mathbf{x}(t) + \mathbf{D}u(t) = [0 \ 0 \ K_t]\mathbf{x}(t), \quad (6)$$

where $\mathbf{D} = 0$ in this case.

The third equation in (3) can be eliminated by driving the motor with a current source. The coefficients of the state equation become

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & -P/M \end{bmatrix} \quad (7)$$

$$\mathbf{B} = \begin{bmatrix} 0 \\ K_t/M \end{bmatrix}, \quad (8)$$

where the state variable vector consists of armature position and velocity only, and the input $u(t) = I_s(t)$. The output equation becomes

$$T = \mathbf{D}u(t) = K_t I_s(t), \quad (9)$$

where $\mathbf{C} = 0$ in this case.

The control system is described in the next section. Details of the computer simulation are discussed in Section III. In Section IV the results of various parameter studies are presented. Section V describes some experiments with the completed gripper system. Finally, some concluding remarks concerning certain significant results are offered along with some plans for future study.

II. GRIPPER CONTROL SYSTEM

The principal function of the controller is to position each gripper finger appropriately for approaching an object and to close the gripper fingers on the object. Another goal of the gripper control system design is to provide a means for affecting the mechanical properties so that gripper compliance and dynamic performance can be modified under computer control. For example, it is desirable to minimize impact forces due to the inertia of the gripper finger and motor armature masses while allowing rapid response in the mechanism. One way to reduce impact forces is to reduce the mass of the fingers. Since the mechanical mass of the mechanism has already been minimized by design, it is desirable to further reduce the effective mass of the system by appropriate control variable feedback.

The effective mass of the gripper mechanism can be controlled electrically by feeding back an acceleration term in the control system. It is well known that effective damping in the system is controlled by velocity feedback. This was the basis for an early design of the gripper control system. Subsequently, after nonlinear computer analysis of several system configurations, a system design evolved which is depicted in Fig. 2.

The control system for the gripper consists of an analog section and a digital interface. Analog feedback methods were chosen after numerous computer simulations indicated that a microprocessor would not be fast enough to perform the complicated control calculations necessary at a sufficiently high sample rate. The control circuitry consists of 24 operational amplifiers, six 8-bit Multiplying Digital-to-Analogue Converters (MDACs), two 10-bit DACs, and a digital interface. The digital interface communicates with a SUN 68000 microcomputer running under New Real-Time Executive (NRTX) operating system.⁷ Mechanical compliance is controlled by position feedback, damping by velocity feedback, and effective mass by acceleration feedback. The velocity and acceleration signals are obtained from the position signal by electrical differentiation. In order for this technique to work well, the position signal must have very low noise. The

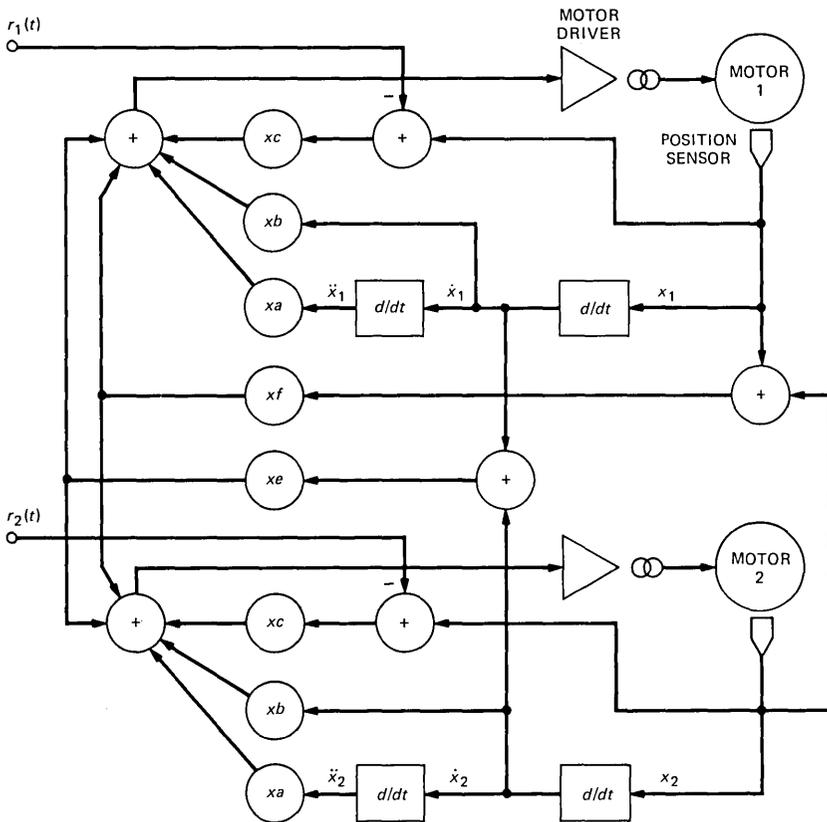


Fig. 2—Gripper control system.

amounts of position and velocity feedback are controlled by computer through the digital interface and 8-bit MDACs indicated by coefficients b and c in Fig. 2. Finger positioning stiffness is controllable from zero to about 23,000 N/m in increments of about 90 N/m. Acceleration feedback gain is fixed by a trimmer potentiometer and is set as high as possible consistent with stable operation of the gripper. Command positions $[r_1(t)$ and $r_2(t)]$ are controlled by 10-bit DACs for a positioning resolution of about 40 μm .

An additional feature of the control system is a means of cross-coupling the two finger positions so that an effect on common-mode compliance is obtained. Common-mode damping is provided by common-mode velocity feedback. Both common-mode position and velocity gains are controlled by 8-bit MDACs (coefficients e and f in Fig. 2). This common-mode feedback allows the gripper fingers to, for example, hold an object with a soft grip (differential mode) while stiffly holding the object's position with respect to the gripper body

(common mode). Alternatively, the common-mode stiffness can be made zero so that the finger positions can conform to an off-center object when they are closing (such as a peg in a hole).

Common-mode compliance control can be illustrated by a few examples. Ignoring velocity and acceleration gains, the feedback equations for the gripper fingers are

$$I_1 = a_1(r_1 - x_1) + f(x_1 + x_2) \quad (10)$$

$$I_2 = a_2(r_2 - x_2) + f(x_1 + x_2), \quad (11)$$

where I_1 is the motor current control signal for motor 1 (which may be scaled by fixed gain constants), a_1 is position gain associated with finger 1, f is the common-mode position gain, r_1 is the commanded position for finger 1, x_1 is the actual position of finger 1 (positive direction is motion toward the center of the gripper), and x_2 is the actual position of finger 2, etc. The gripper fingers always tend to move so that I_1 and I_2 are zero. The amount of deviation of the values of I_1 and I_2 from zero is proportional to the reactive force generated by the servomotor.

By appropriate choice of a_1 , a_2 , and f , several different modes of operation can be obtained. Under normal operating circumstances, the values for finger 1 (a , b , and c) are chosen to be the same as those of finger 2 for balanced operation. If f is set to one half the value of a , for example, then the common-mode stiffness is very high and differential-mode stiffness is zero. The control equations become

$$I_1 = ar_1 + a(x_2 - x_1)/2 \quad (12)$$

$$I_2 = ar_2 + a(x_1 - x_2)/2, \quad (13)$$

where $a = a_1 = a_2$. This results in an effect where each finger individually has no stiffness with respect to the gripper body as long as the other finger is free to move. However, if both fingers are forced to move in the same direction, considerable stiffness is encountered. In this mode the fingers tend to stay equally distant from a position value centered between the two-finger command values r_1 and r_2 . If the values of the a coefficients are increased, a soft grip can be established with very firm centering of the object. This is useful when the robot arm is in motion as it keeps the fingers and object centered without unnecessary gripping force.

Alternatively, if the a coefficients are made zero and f is negative, common-mode stiffness is nullified and all stiffness appears as differential-mode stiffness. The control equations reduce to

$$I_1 = -a(x_1 + x_2) \quad (14)$$

$$I_2 = -a(x_1 + x_2). \quad (15)$$

This means that the gripper fingers are stiff with respect to each other, but exhibit no stiffness with respect to the gripper body. This mode of operation is particularly useful in picking up objects that may not be positioned exactly in the center of the gripper. The fingers will easily conform to the position of the object without applying significant lateral forces. Varying amounts of common-mode versus differential-mode stiffness can be obtained with values of f between zero and one-half a .

We will introduce the state feedback control equations by first considering a single finger with position, velocity, and acceleration feedback with a current-driven servomotor. Initially, we will assume that the velocity and acceleration are available directly from the finger by appropriate sensors (later we will remove this requirement). We will also assume a linear system until later when nonlinearities are added. Then we can write a feedback equation for the motor current as follows:

$$I_s(t) = -a\ddot{y}(t) - b\dot{y}(t) - c[y(t) - r(t)], \quad (16)$$

where $r(t)$ is the reference position input function and a , b , and c are gain coefficients. Substituting (16) into (7) through (8) and combining like terms yields

$$\mathbf{A} = \begin{bmatrix} 0 & 0 \\ -cK_t/(M + aK_t) & -(P + bK_t)/(M + aK_t) \end{bmatrix} \quad (17)$$

$$\mathbf{B} = \begin{bmatrix} 0 \\ cK_t/(M + aK_t) \end{bmatrix}, \quad (18)$$

where, again the state vector, \mathbf{x} , consists of finger position y and velocity \dot{y} , and the input function $u(t) = r(t)$.

The lower left element of \mathbf{A} is now nonzero, indicating that a position-sensitive term is present. This term is effectively a spring constant divided by a mass. For comparison, examine the state equations for a classical elementary mass, spring, and damper problem:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -k/m & -p/m \end{bmatrix} \quad (19)$$

$$\mathbf{B} = \begin{bmatrix} 0 \\ 1/m \end{bmatrix}, \quad (20)$$

where k is the spring constant, p is the damping coefficient, and m is the mass. The input function $u(t)$ applies a force to the mass, which is attached to a stationary support by the spring and damper. Note that in (17) through (18) the spring constant is replaced by (cK_t) , the damping by $(P + bK_t)$, and the mass by $(M + aK_t)$. Thus, these

mechanical parameters can be independently controlled by the electrical feedback gains (a , b , and c) in the control system.

The second finger is controlled in a similar manner. The two fingers thus operate independently in this case. Figure 2 also shows control signal passing from each finger to the other. These signal paths provide additional common-mode position and velocity feedback, which increases the common-mode stiffness. It was found, during computer simulation, that common-mode acceleration was not useful and that it tended to destabilize the control system. The complete feedback equation for motor drive current is

$$I_s(t) = -(a + d)\ddot{y}_1(t) - d\ddot{y}_2(t) - (b + e)\dot{y}_1(t) - e\dot{y}_2(t) \\ - (c + f)y_1(t) - fy_2(t) + (c + f)u_1(t) + fu_2(t). \quad (21)$$

Closed-form solutions of the above equations are not readily obtainable for a nonlinear system. The construction of our gripper involves servomotors with power limitations. The motor driver has both voltage and current limitations (which must be treated separately). The rack and pinion have backlash, and the mechanical parameters change abruptly when an object is grasped. The rack and pinion gear backlash has turned out to be a small effect, but it was not initially clear if this hysteretic effect would cause system instability with the electrical feedback.

III. COMPUTER SIMULATION

A fourth-order Runge-Kutta algorithm was used to study the time response of the nonlinear system just described. The algorithm used is modified somewhat from standard routines found in the literature and is described in Appendix A. The state equations must be expanded to include the various nonlinearities. After doing so, the two-fingered model requires an eight-dimensional state vector. Figure 3 is a diagram of the model and includes an object to grasp (in this case a rubber ball). For the time being, we will still assume that the finger velocity and acceleration are directly available. Later, we will study the influence of electrical time constants on the mechanical performance.

The model shown in Fig. 3 consists of five mechanically independent members. They are the two fingers, two motor armatures, and the rubber ball. Each mechanical member will require two state variables to describe its position and velocity. The ten-dimensional state vector is

1. Left finger position
2. Left finger velocity
3. Left armature position

4. Left armature velocity
5. Right finger position
6. Right finger velocity
7. Right armature position
8. Right armature velocity
9. Ball position
10. Ball velocity.

The state equation describes forces that exist on each mechanical member. These forces may be due to acceleration, frictional drag, or spring forces. Also included are forces applied to each member by another member when they are in contact. To introduce the state feedback matrix, let us first consider the case where no contact is being made between any of the mechanical members and the common-mode feedback is ignored. Then the coefficients of the state equation are

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0^* & -P/M^* & 0^* & 0^* & 0 & 0 & 0 & 0 & 0^* & 0^* \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -c K_t/m^* & -b K_t/m^* & 0^* & -p/m^* & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0^* & -P/M^* & 0^* & 0^* & 0^* & 0^* \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -c K_t/m^* & -b K_t/m^* & 0^* & -p/m^* & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0^* & 0^* & 0 & 0 & 0^* & 0^* & 0 & 0 & 0^* & 0^* \end{pmatrix} \quad (22)$$

$$\mathbf{B} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ K_t & 0 & -a K_t/m & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & K_t & 0 & -a K_t/m & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad (23)$$

where P is the damping coefficient of the finger slide mechanism, M is the mass of the finger, p is the damping coefficient of the motor armature, and m is the mass of the armature. The rubber ball rolls on a table with negligible friction.

The input function $\mathbf{u}(t)$ is a vector of four dimensions. The first two dimensions are the left and right finger position reference inputs. The last two dimensions are inputs from the implied left and right accelerometer sensors. This variation of the feedback equation is useful since acceleration is not one of the state variables available. Alternatively, the time derivative operator could be used in the $\mathbf{A}(\mathbf{x}, t)$ matrix.

The $\mathbf{A}(\mathbf{x}, t)$ matrix becomes more complicated when any of the

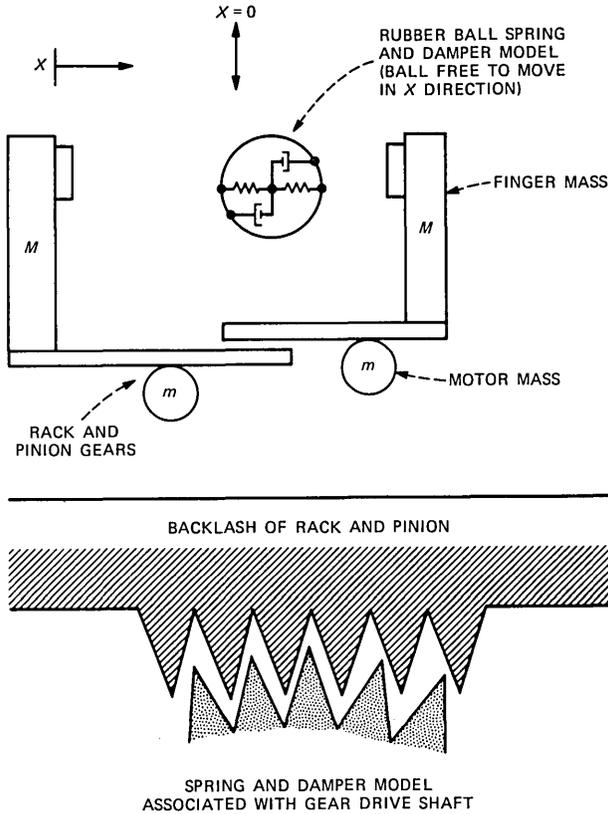


Fig. 3—Model of two-fingered gripper.

mechanical elements come into contact. Some or all of the elements of the matrix indicated with an asterisk in (22) become nonzero, and additional forces are coupled to the various mechanical parts in a symmetric manner. Equal and opposite forces are introduced into two rows of $A(\mathbf{x}, t)$ when contacts are made. Thus, the matrix form of (22) is helpful in illustrating the symmetry in the problem. For example, when the left finger contacts the ball, elements 1, 2, 9, and 10 of row 4 are affected. At the same time, the corresponding elements of row 10 are also affected with identical magnitude, but with opposite signs. The motor drive nonlinearities are saturation limits on both voltage and current. These enter directly into the motor torque equations.

The actual nonlinear state equation would consume more than a single page. A C program listing of the complete state equation is given in Appendix B. This listing of the state equation also includes the electrical time constants, which deserve some discussion.

Heretofore, we had implied that the acceleration of the left and right fingers would be available via some sensor, such as an accelerometer. In the actual gripper design, the acceleration and velocity are obtained by electrical differentiation of signal from the position sensor. The design of an electrical differentiator purposely incorporates a pole in the transfer function to limit the high frequency gain and thus reduce the amount of noise developed by or passed through the circuit. It is desirable to make this frequency limit as low as possible to minimize the circuit noise. However, the accuracy of the differentiation may suffer if the pole location is too low in frequency. An analysis of the effects of differentiator pole position was therefore necessary. The motor drive amplifier also has a dominant pole at a moderately low frequency (about 1500 Hz) and its effect is also included in the state equation.

IV. RESULTS OF THE SIMULATION

Transient response analysis was performed for various parameter settings to determine the effects of different amounts of feedback, the effects of various ball masses and compliances, and the effects of the differentiator time constants. A fourth-order Runge-Kutta algorithm was employed with modifications of the step size correction algorithm as described in Appendix A. The time response output is presented in graphical form that is easily read and gives a good indication of the performance.

Several parameters were initially established for the analysis. These include the mechanical specifications for the dc servomotors and the finger mechanisms. The servomotors were selected for this application based on their physical size, torque capability, and the fact that they are completely characterized by the manufacturer. The published specifications for the motors are

Armature resistance	(<i>Ra</i>)	6.12	Ω
Armature inductance	(<i>La</i>)	0.00275	H
Motor EMF constant	(<i>Ka</i>)	0.0446	V-s/rad
Motor torque constant	(<i>Km</i>)	0.0446	N-m/A
Armature inertia	(<i>m</i>)	1.54×10^{-6}	N-m-s ² /rad
Armature damping	(<i>p</i>)	2.74×10^{-6}	N-m-s/rad.

The symbols in parentheses refer to variables in the computer program (Appendix B). In addition, specifications for the rack and pinion gear mechanism used are

Pinion gear radius	(<i>rg</i>)	0.0081	m
Pinion gear backlash	(<i>H</i>)	5.08×10^{-5}	m
Finger mass	(<i>M</i>)	0.6	kg
Finger damping	(<i>P</i>)	0.5	N-s/m.

The following parameters are variable in the simulation and can be set to establish interesting initial conditions:

(a) Mechanical parameters:

Initial ball position	($y[8]$)
Ball radius	(R)
Ball stiffness	($k\beta$)
Ball mass	($m\beta$)
Ball damping	($p\beta$)

(b) Electrical parameters:

RC time constants	(RC)
Position feedback gain for each finger	(c)
Velocity feedback gain	(b)
Acceleration feedback gain	(a)
Common-mode position gain	(f)
Common-mode velocity gain	(e)
Common-mode acceleration gain	(d)
Motor driver voltage limit	(V_{max})
Motor driver current limit	(I_{max}).

Several of these items introduce nonlinearities into the state equation. Two voltage programmable power supplies were used for motor drivers. These power supplies are linear units capable of operating either as voltage or current sources. The supplies used in this study are limited to 50 volts and/or 8 amperes. They can operate as either a power source or sink. They can also be controlled by either a voltage input or via a digital port, making them useful for microprocessor control.

Initially, the transient analysis is conducted with step position command inputs. Later, we will discuss exponential input functions and their effect on performance. To perform a typical simulation the initial conditions are set so that the gripper is in the fully open position. Thus, the gripper is opened to a width of 6.35 cm. The step input functions are then applied and the gripper closes on an object, in this case the ball. Time responses of each finger and the ball position are computed along with the associated motor voltages and currents. Also calculated are the ball position (and velocity), the impact force being applied to the ball by each finger, and the estimation errors of the electrical differentiators. This information is all displayed on a single page in graphical form for each test response.

The gripper was designed to allow each finger to move 3.8 cm. The fingers have about 1.25 cm of travel overlap, so in the fully opened position they are about 6.35 cm apart. The positional reference for the

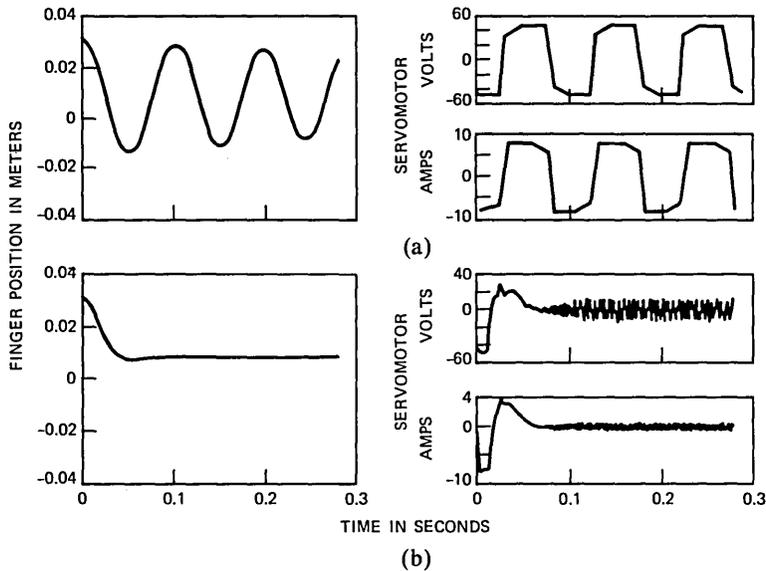


Fig. 4—(a) Step response with position feedback. (b) Step response with position and velocity feedback.

fingers was established so that the closed position of the gripper with the fingers centered is considered the zero position of each finger. The left finger opens in the minus x direction and the right finger opens in the plus x direction.

Figure 4 illustrates the effects of power supply limits, gear backlash, and velocity feedback on the time response. For these responses the gear backlash was set to 1 mm to make the hysteresis effect obvious. This is roughly equivalent to removing every other tooth from the gear. Figure 4a shows a response with position feedback only. Without velocity feedback the mechanical damping is obviously inadequate. The gear backlash is only slightly visible in the plot of position versus time and appears as a slight jaggedness in the response. The voltage and current plots show the effects of the motor driver limits and the inductance of the motor. Note that initially the voltage is saturated at -50 volts. The current goes to nearly -8 amps but then decreases until the driver switches to $+8$ amps. This ramp effect is due to the induced back EMF of the motor as it gains velocity. The voltage response shows similar characteristics during current saturation.

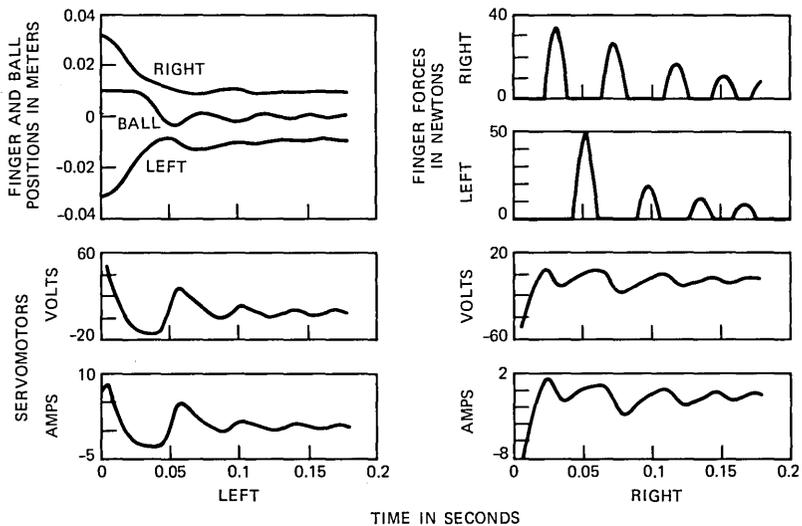
Figure 4b shows a response with viscous damping applied by velocity feedback. In this case the velocity gain is set to properly complement the position gain so that settling time of the finger is somewhat less than "critically damped." This amount of damping results in minimum settling time to the final position since a very slight amount of

overshoot in the response (a few percent) results in much faster convergence to the desired control point. The effects of gear backlash can now be seen in the voltage and current plots. During the dynamic part of the response there is very little evidence of gear backlash. The gear is first in contact on the forward thrusting (accelerating) face and then on the decelerating face. After the position of the finger becomes stable, the gear must continually switch faces to maintain the finger in position. This backlash effect shows up as a small oscillation in the drive current and voltage. The actual gear backlash of the gripper is much smaller than this and its effect will not be apparent in the time responses that follow.

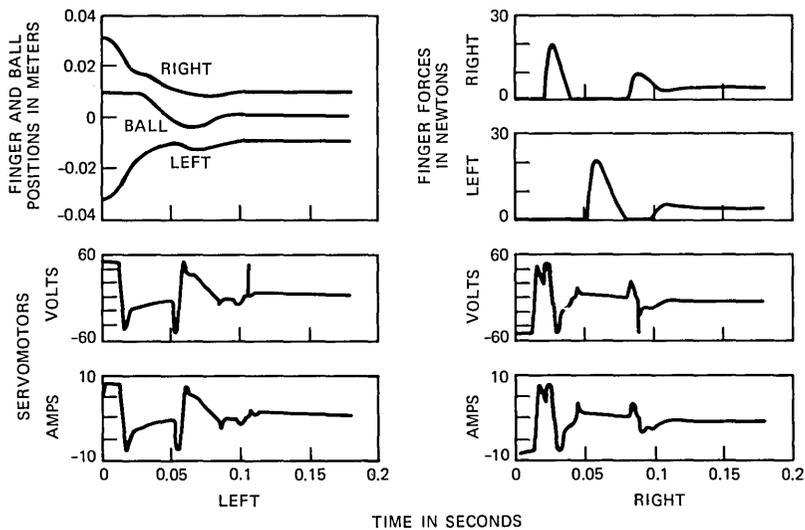
When the fingers come into contact with the ball, impact energy and forces are exerted that may cause damage to the ball (or object) and/or gripper mechanism. These impact forces can be affected significantly by acceleration feedback. Figure 5 illustrates this effect. A typical damped closure of the gripper is shown in Fig. 5a. The gripper is set in the fully open position and closes on a 2-cm diameter ball weighing 0.5 kg. The impact forces are quite high, several times the magnitude of the final static gripping force applied by the servomotors. Figure 5b shows what happens when the mass is reduced by positive acceleration feedback. The amount of positive feedback must be carefully controlled or effective mass will become negative and result in oscillations. For the results of Fig. 5 the acceleration gain has been adjusted to reduce the effective mass to about 10 percent of the true mass. The impact forces have been reduced by acceleration feedback, but the amount of reduction is dependent upon properties of the rubber ball and other feedback parameters. Under good conditions, impact forces can be reduced by almost 50 percent.

There are two significant differences in the responses of Figs. 5a and b. First, the initial peak impact force has been reduced from 50N to about 20N with acceleration feedback. Furthermore, the gripper stabilizes the ball position much more rapidly. This is due to the increased responsiveness of the gripper fingers to the ball dynamics as a result of the reduced effective mass. Increased motor drive bandwidth is also required in order to control the effective mass. Note the high frequency components present in the drive currents of Fig. 5b.

A series of test simulations were conducted for various rubber ball stiffnesses and damping coefficients to investigate the effectiveness of acceleration feedback. Again, simulations were performed using a step function position control input. The ball is positioned 1 cm off-center toward the right finger in order to test the ability of the gripper to center and stabilize the position of the ball. The position feedback gain was set to 500 amps per meter (which is equivalent to a spring



(a)



(b)

Fig. 5—(a) Damped gripper closure with position and velocity feedback. (b) Damped gripper closure with position, velocity, and acceleration feedback.

rate of about 2750 N/m) for the first two test series, and critical damping was established by a velocity feedback of 14.95 A-s/m.

The ball damping was set to 1.0 N-s/m for the first test series. The ball stiffness was varied from 10^2 N/m to 10^6 N/m. Acceleration feedback was set to reduce effective mass to about 5 percent of

mechanical mass at $0.11 \text{ A-s}^2/\text{m}$. The ratio of peak impact force applied by the first contact between a gripper finger and the ball with and without acceleration feedback was taken to be the measure of the effectiveness of this feedback. The results of this test series are plotted in Fig. 6a. The plot shows that the ball stiffness has a substantial influence on the effectiveness of acceleration feedback. Peak impact force is only slightly reduced for very large or very small stiffnesses. The effect is substantial at moderate stiffness (10^4 N/m) but only over a narrow range.

A second series of tests were conducted with the stiffness fixed at 10^4 N/m . The ratio of impact forces was plotted again for ball damping from 0.01 to 10^4 N-s/m . Figure 6b shows that damping has little effect on the impact ratio until it exceeds about 10 N-s/m . After this point is reached, electrical damping predominates over the stiffness. At very large damping values the acceleration feedback becomes totally ineffective.

The position feedback was varied in a third set of test simulations. The ball stiffness was fixed at 10^4 N/m and the damping was set at 1 N-s/m . Velocity feedback was adjusted in each instance so that critical damping was maintained. Figure 7 shows a plot of the results for position feedback gains varying from 10 to 5000. Both very low and very high gains appear to adversely influence the effectiveness of acceleration feedback. Position feedback gains in the range of 50 to 200 A/m have the greatest effect.

Since position feedback produces an effect similar to the stiffness of the ball, it is not surprising that the plots have similar characteristics. Loss of acceleration feedback effectiveness at high stiffnesses can be attributed to relatively slow response time of the servomotors in comparison to the mechanical resonance of the ball and finger system. Since it takes approximately $100 \mu\text{s}$ for the servomotors to reverse the gear force (because of backlash), the mechanism simply does not have enough time to counteract the effects of the real mass of the fingers. Poor acceleration feedback effectiveness at very low stiffnesses appears to be due to mismatch between the stiffnesses of the finger and the ball. If one of the stiffnesses overwhelms the other, particularly by several orders of magnitude, the acceleration feedback becomes useless.

Another means for reducing impact forces is to bring the fingers into contact with the ball at low velocity. This could be accomplished by simply bringing the fingers slowly together by applying a ramp function to the position control inputs; however, this makes the gripper operation slow. Since it is only necessary for velocity to be low just before contact is made with the ball, another possibility is to increase the negative velocity feedback so that the finger system is overdamped.

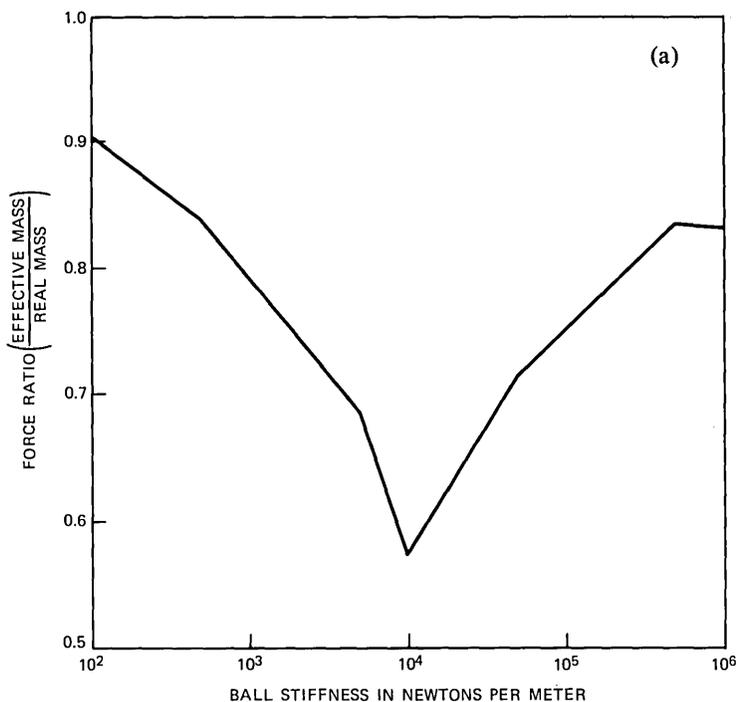


Fig. 6—(a) Peak impact force ratio versus ball stiffness. (Cont.)

This results in an exponential velocity response. However, computer simulation shows that this does not really reduce impact forces, since the fingers become very resistive to springy impacts. The series of Fig. 6b illustrates this. Another alternative is to apply an exponential position control input to each finger. This takes advantage of the higher allowed velocity and slows the fingers down as they approach the contact points. Furthermore, these input functions are very easy to generate in hardware. Figure 8 shows an example of combined positive acceleration feedback and exponential control input functions with 50-ms time constants. The impact forces are now reduced to little more than the final static gripping force. Several time constants for the input exponentials were investigated, and the 50-ms exponential appeared to yield the best overall results (low impact force consistent with rapid closure) for this particular mechanical gripper.

An analysis of sampling period was performed on a discrete time control architecture to determine if the gripper could be controlled adequately by a microprocessor. This was accomplished by modeling the motor drivers as zero-order sample and hold devices. A minimum sample rate of at least 5000 samples per second was indicated by numerous computed responses. This rate would be needed to control

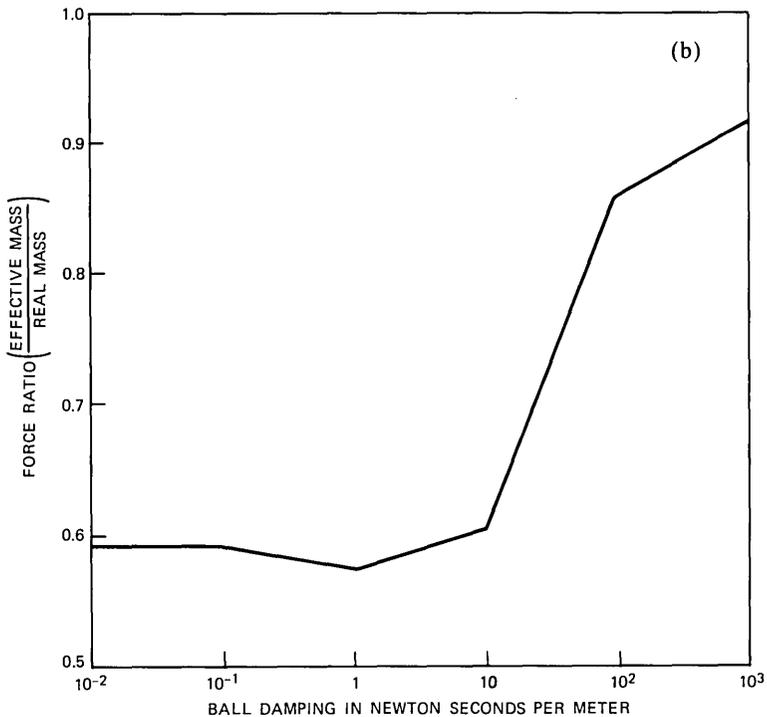


Fig. 6—(b) Peak impact force ratio versus ball damping.

the gripper without loss in performance. This requirement could be relaxed somewhat if certain erroneous dynamic response characteristics could be tolerated; however, the characteristics observed in the simulation were not considered acceptable. This sample rate would only be possible for a microprocessor if a very simple control strategy were being implemented. Computing the motor currents necessary every 200 μ s for the more sophisticated feedback strategy being planned is more than a normal microprocessor can handle. (It is possible to perform the operations with a digital signal processor, however.) Consequently, the controller design is primarily analog, with digital-to-analog converters to allow a microprocessor to control the various feedback gains and apply position control inputs.

A conference on robot manipulators⁸ triggered some thinking on the concept of differing common-mode and differential-mode compliance. Active compliance in our gripper is useful as long as the robot arm is not moving. When the robot arm does move, however, this previously useful compliance becomes a liability because the fingers of the gripper will swing from side to side under the acceleration influence of the arm motion. It is desirable to maintain the compliance relative to the grasped object and simultaneously keep the fingers and object from

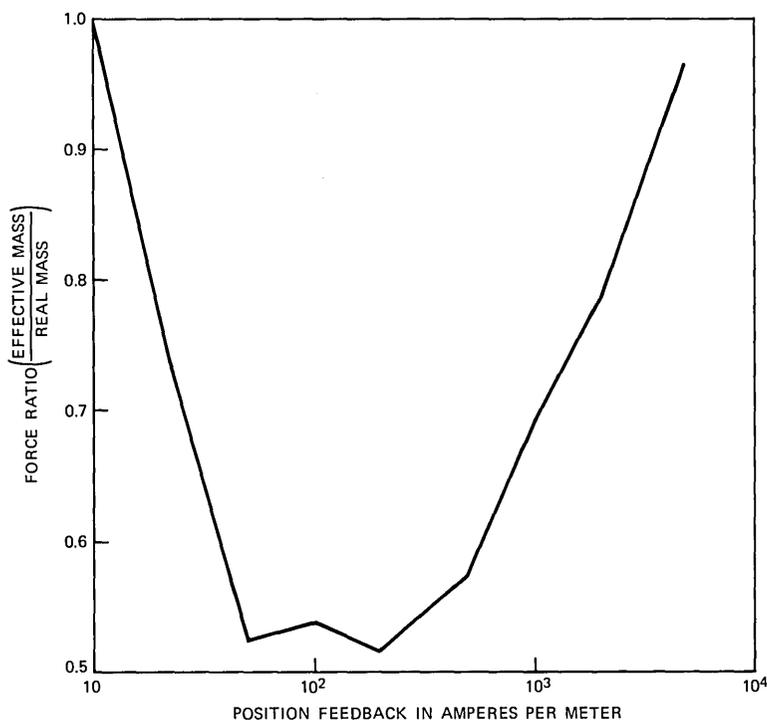


Fig. 7—Peak impact force ratio versus position feedback gain (critically damped).

swinging. This can be accomplished with appropriate common-mode feedback as discussed in Section II.

The effect of additional common-mode stiffness is illustrated in Fig. 9. This should be compared to Fig. 5a. Both common-mode position and common-mode velocity feedback are necessary for suitable common-mode performance. The initial conditions have been set to the same values as in previous tests. Inertia of the ball causes the gripper to accommodate somewhat to the position of the object and close slightly off-center. The object is then forced to the center by the common-mode stiffness. Close inspection of the position plot of Fig. 5a shows some common-mode finger oscillation under the influence of the ball vibration. This common-mode component is almost entirely eliminated in Fig. 9. Further experiments simulating arm motion by applying common-mode variations to the position control input functions indicate that the increased common-mode stiffness significantly reduces the amount of arm motion influence on the fingers.

We had previously assumed that velocity and acceleration signals were available from appropriate sensors such as tachometers and accelerometers. The actual gripper uses a position sensor that consists

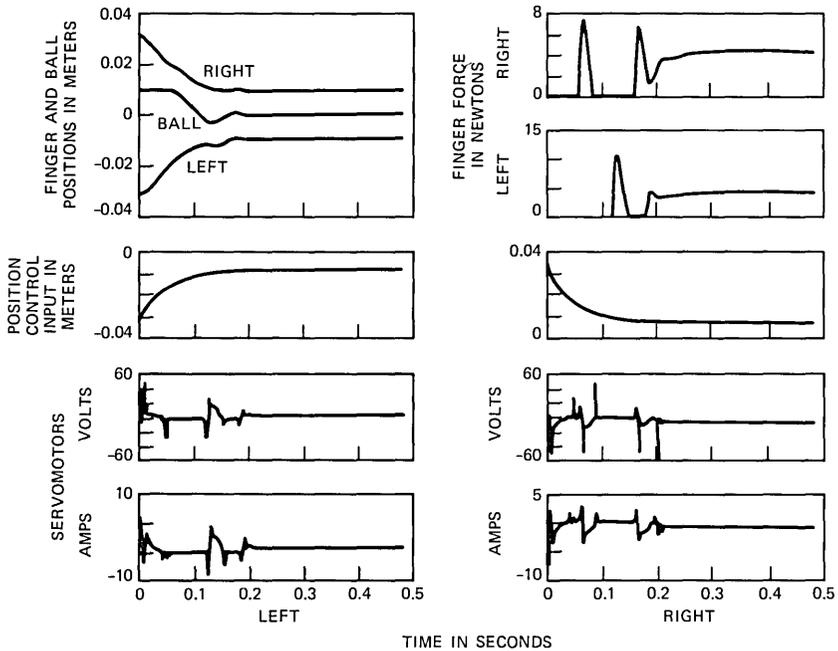


Fig. 8—Damped closure with 50 milliseconds time constant.

of a capacitive measurement device described by G. L. Miller.⁵ The velocity and acceleration signals are obtained electronically from the position signal by electrical differentiation. It is desirable to limit the high-frequency gain of the differentiator circuits used to estimate these parameters, so electrical poles are purposely designed into the differentiator transfer functions. Extensive computer simulation has shown that the electrical time constants should not exceed $100 \mu\text{s}$. Figure 10 shows a time response for circuit time constants of 10 ms. This should be compared to Fig. 9. The obvious distortion of the mechanical response is further indicated in the log error plots of velocity and acceleration estimates for the right finger. The errors are typically no more than about 10 percent for circuit time constants of $100 \mu\text{s}$, and the mechanical response is virtually identical to responses obtained with smaller time constants.

V. TESTING THE GRIPPER

Initial tests conducted on the gripper showed that high-frequency noise was present in the position signal. Although the level of this noise was low (about 50 dB s/n), it became important when used to determine the velocity and acceleration signals. At the acceleration output of the differentiators, the wideband noise level was only 30 dB

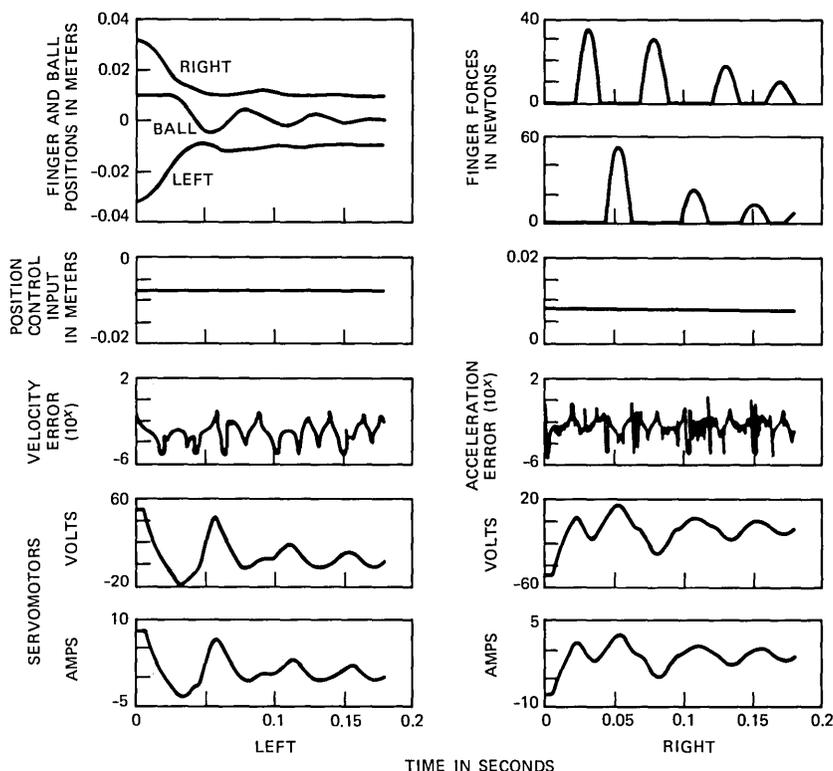


Fig. 9—Damped gripper closure with common-mode feedback.

below the signal, a level high enough to cause gear chatter in the rack and pinion drive. Subsequent modifications of the position sensing circuitry reduced this noise level to about 90 dB below signal, which made the acceleration signal useful for feedback purposes.

To verify the simulation results, measurements of the mechanical finger mass and damping were necessary. This was accomplished by performing underdamped step response tests on the gripper. Figure 11 shows a typical underdamped step response. This response is very much like that of a simple mass, spring, and damper. The mass can be determined from the resonant frequency and the damping from the rate of decay in the response. Undamped resonant frequency is given by

$$\omega^2 = \frac{k}{m}, \quad (24)$$

where ω is frequency in radians per second, k is the spring constant in Newtons per meter, and m is mass in kilograms. Damping (N-s/m) can be determined from

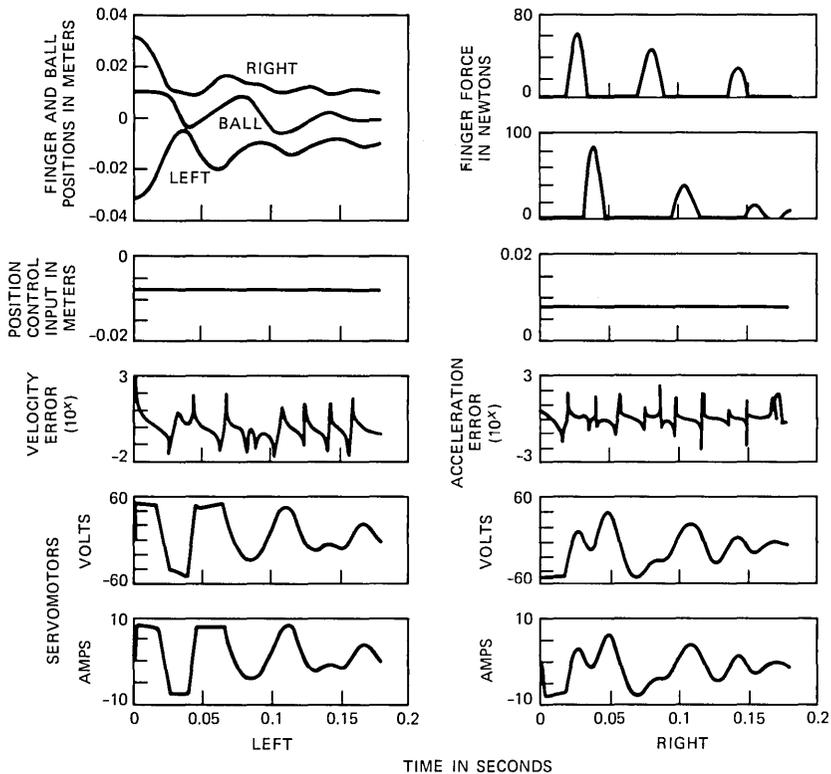


Fig. 10—Damped gripper with common-mode feedback and 10 milliseconds electrical time constants.

$$b = \frac{2m}{T} \ln \frac{x_1}{x_2}, \quad (25)$$

where b is the damping in Newton-seconds per meter, m is the mass in kilograms, x_1 and x_2 are successive peak amplitudes in the decaying response, and T is the cycle period in seconds.

There remains the problem of verifying the spring constant, which, in this case, is due to the position feedback through the servomotor. Although the motor is completely characterized by the manufacturer, it was deemed important to measure the spring constant actually obtained from the servomotor. This was accomplished by performing a step response analysis with two different masses attached to the motor through the rack and pinion, as follows:

Let m_1 = finger mass

m_2 = slide and motor mass.

Then,

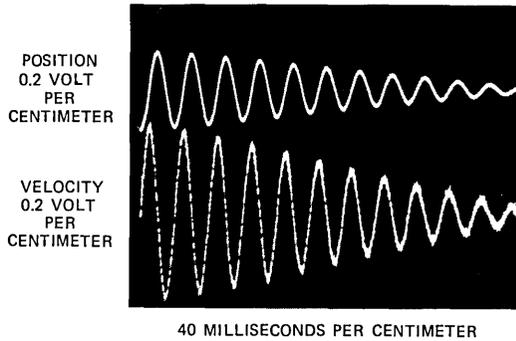


Fig. 11—Underdamped step response.

$$\omega_1^2 = \frac{k}{m_1 + m_2} \quad (26)$$

gives the undamped resonant frequency of the complete motor, slide bearing, and finger masses for a given equivalent spring constant k . The finger mass m_1 can be removed, leaving m_2 and a second resonant frequency ω_2 . Thus an equation for k can be derived:

$$k = \frac{m_1}{\frac{1}{\omega_1^2} - \frac{1}{\omega_2^2}} \quad (27)$$

Mass m_1 is measurable since it can be removed from the rest of the mechanism. Mass m_2 consists of a combination of the slide mechanism, the rack mass, and the gear and motor inertia. Once k is determined, then m_2 can be calculated from (26).

Measurements in the laboratory yielded the following data:

$$\begin{aligned} k &= 3900 \text{ N/m} \\ m_1 &= 33.12 \text{ g} \\ \omega_1 &= 154.2 \text{ rad/s} \\ \omega_2 &= 172.6 \text{ rad/s} \\ m_2 &= 131 \text{ g.} \end{aligned}$$

A calculation of the gear inertia was made from the size of the gear and the density of steel (equivalent mass: 3.5 grams). Adding this to the published specification for the motor inertia and translating into the linear (rack) units yields a slide mass of 104 grams. Thus the total slide and finger mass is about 137 grams. The servomotor inertia with the gear attached is about $1.77 \times 10^{-6} \text{ kg} - \text{m}^2$. With these parameters determined, the simulation program was used to verify the resonant frequencies.

Simulation results compared closely with the actual performance of the gripper. Typical step response settling times of 30 ms were obtained for high-position feedback gains, as predicted by the computer simulations. Absolute positioning accuracy is dependent upon the position sensor accuracy and mechanical imperfections. The currently employed sensor has a maximum error of about 0.5 mm.

Several slight differences in response between the actual gripper and the simulation were noted. The computer simulation did not consider the effect of static friction on the dynamic response. In this case, the frictional force appears to be a constant independent of velocity. With viscous friction, the underdamped response decays exponentially. Static friction results in a linear decay. The actual gripper responses indicate that a slight amount of static friction is present in the mechanism. The amount of static friction present is negligible for normally damped responses (near critical damping) since the viscous friction is dominant under these conditions. However, static friction does become significant when the finger compliance is high. This results in a positioning error, since the force necessary to overcome the static friction is proportional to the spring constant applied to the finger. One well-known method of solving this problem is to introduce high-frequency dither into the motor drive. The amplitude of the dither must be just sufficient to exceed the force necessary to break the static friction free. If the dither frequency is sufficiently high, its effect will not be seen in the finger position due to the inertia of the mechanism. Experiments with dither in the laboratory indicate that static friction is not the only source of positioning error.

Dither was found to be an effective means for obtaining higher positioning accuracy in the presence of static friction at low-position feedback gains (low spring constant); however, magnet detent action in the servomotor due to the magnetic reluctance of the armature poles also causes positioning error. The effect of this motor phenomenon is the appearance of virtual potential energy wells at uniform intervals along the finger track. Thus, the positioning error can be positive or negative, depending on the commanded position and the relative position of the nearest potential energy well. This effect can be compensated by adjusting the command position slightly, but it is nearly impossible to position the finger at a potential energy peak since any slight change in position can cause a change in sign of the detent force, which requires dynamic compensation. The only practical solutions for this problem are to increase the position feedback (which provides the dynamic compensation) of the system or reduce the motor gear radius so that the detents are closer together. The difficulty in electronically compensating for the detent action arises from the fact that there is no way of distinguishing between internally and externally

generated finger displacements with the present sensors. The current version of the gripper exhibits positioning detents at about 4 mm intervals.

The effect of acceleration feedback was tested by studying the underdamped step response of the gripper fingers with a known position feedback gain. The frequency of the response is an indication of the effective mass of the system. It was found that the resonant frequency could be increased by about 20 percent, indicating an effective mass change of

$$\frac{m_2}{m_1} = \frac{\omega_1^2}{\omega_2^2} = 0.64, \quad (28)$$

or about a 36 percent decrease in mass. The simulation predicted a somewhat higher mass reduction (about 55 percent), but it had not considered the effects of position measurement noise, which limited stable feedback gain.

Position measurement noise also becomes significant in the velocity feedback term when velocity gain is high (high damping). When position feedback gain is set to a high level, high-velocity gain is needed to obtain critical damping. Under these conditions the noise present in the position signal causes servomotor gear chatter. This gear chatter has no significant effect on the positioning accuracy or dynamic response of the gripper, but it does produce annoying sounds in the rack and pinion mechanism when the finger is not in contact with an object and is motionless. The high damping in the finger is not generally needed except when the finger is moving. Under computer control, it was found that the dynamic response of the finger is not affected if the high-velocity gain is present for only about 50 ms after the "move finger" command is initiated. Then the velocity gain can be reduced to a safe level for noise-free operation.

To make control of the gripper easier for the user a set of computer routines were written in the C language. These procedures allow the user to set all feedback parameters directly or, at a higher level, to move the fingers without having to be concerned with determining the appropriate velocity feedback for critically damped response. Velocity gain for critical damping can be determined from an equation of the form

$$b = k_1\sqrt{c} - k_2, \quad (29)$$

where b is velocity gain, c is position gain, and k_1 and k_2 are constants. Constant k_2 is due to the mechanical damping already present. Another slower procedure provides a method for opening and closing the fingers slowly by incrementally stepping the fingers to the command position. More sophisticated algorithms allow the user to specify a compliance

and gripping force to be applied when closing. The gripper is closed on the object and the applied force is increased to the specified amount.

VI. SUMMARY

A robot gripper and mechanical impedance control system has been constructed based on the results of computer simulations. The computer simulation is an accurate and useful tool for studying the dynamic behavior of the gripper. It allows one to investigate various performance properties of the mechanical system (i.e., conduct computer experiments) without taking the risk of damaging components of the system or incurring the expense of building experimental equipment. The simulation could be improved somewhat by including the effects of static friction, the servomotor detent action, and the control system generated noise.

The controlled impedance gripper has an adjustable finger positioning stiffness that can be set in steps of about 90 N/m up to 23,000 N/m. Damping can be set accordingly by computer so that near critically damped response is obtained. The effective finger mass can be reduced by acceleration feedback to about 64 percent of its true value. This helps to reduce impact forces when the gripper fingers are closed on an object.

A computer algorithm library was devised to provide a user interface to the gripper control system. With this library a user can establish operating parameters for the gripper and move the fingers with critically damped responses without having to know the necessary velocity feedback parameters. Because of noise in the position measurement system, high-velocity feedback gains cause servomotor gear chatter. An algorithm has been developed to allow critically damped responses at high-velocity feedback gains by applying the velocity gain only during the transient period, after which the velocity feedback gain is reduced to a statically stable level. In this way the chatter was eliminated.

Improvements can be made to the gripper and control system to reduce the effects of static friction and motor detent action. Dither (high-frequency perturbation) of sufficient amplitude can reduce the effect of static friction. Smaller gear diameters will allow more accurate positioning of the fingers since detent positions will be more closely spaced in finger coordinates.

Adaptive control strategies are likely to be useful in future studies of the problems of robotic gripping. This becomes particularly important for manipulating objects of varying masses and physical dimensions. Knowledge of the size of objects being grasped may not, in general, be known in advance. This means that the gripper will have to adjust position feedback and/or control inputs to maintain a desired

gripping pressure on an unknown object. Further adaptations may improve the ability of the gripper to maintain the object in a stable position as the robot arm is moved, even if the mass of the object is unknown. In the future these operations will probably be performed by a microprocessor because of the relatively sophisticated techniques involved, but the analog controller designed here should provide the microprocessor with the control interface necessary to generate the appropriate gripper response.

VII. ACKNOWLEDGMENTS

The gripper mechanical design was initially conceived by R. A. Boie. He and M. S. Sibilia constructed the gripper from this design. The linear position sensor is due to G. L. Miller. W. J. Kropfl designed and programmed the computer interface. The robot and gripper systems are now running under an executive called NRTX developed by D. A. Kapilow. Without their efforts the gripper could not have come to life.

REFERENCES

1. J. L. Nevins and D. E. Whitney, "Computer Controlled Assembly," *Sci. Amer.*, 236, No. 2 (February 1978), pp. 62-74.
2. M. R. Cutkowsky, "Position Sensing Wrists for Industrial Manipulators," Carnegie Mellon Tech. Rep. CMU-RI-TR-82-9, July 1982.
3. H. Hanafusa and H. Asada, "Stable Prehension of Objects by a Robot Hand with Elastic Fingers," *Proc. 7th Int. Symp. Ind. Robots, Tokyo, October 1977*, pp. 361-68.
4. M. K. Brown, "Computer Simulation of Controlled Impedance Robot Hand," *IEEE Int. Conf. Robotics, Atlanta, GA, March 1984*, pp. 442-50.
5. G. L. Miller, private communication.
6. C. T. Chen, *Introduction to Linear System Theory*, New York: Holt, Rinehart, and Winston, 1970.
7. D. A. Kapilow, unpublished work.
8. N. Hogan, Outlook for the Implementation of Robot Technology Symp., M.I.T. Industrial Liaison Program, February 1983.
9. B. Carnahan, H. A. Luther, and J. O. Wilkes, *Applied Numerical Methods*, New York: Wiley, 1969.
10. L. Collatz, *The Numerical Treatment of Differential Equations*, third ed., Berlin: Springer-Verlag, 1960.

APPENDIX A

Runge-Kutta Method for the Solution of Differential Equations

Runge-Kutta methods for solving general differential equations are well known and probably the most widely used algorithms for this purpose. The particular algorithm employed in this study is a fourth-order formula with coefficients attributed to Kutta (see Ref. 9). The algorithm solves the basic first-order equation:

$$\dot{x}(t) = f(x, t) \quad (30)$$

for $x(t)$, where t is the independent variable. Since any high-order

differential equation can be written as a set of first-order differential equations, an equation of arbitrary order may be solved by repetitive application of the first-order algorithm.

Equation (1) is integrated over a small interval h of the independent variable, which is termed the integration interval or step size. Each interval starts at an initial condition (e.g., $x(0)$) and concludes at an estimated value for $x(t)$ at the end of the interval, [e.g., $x(h)$], which provides the initial conditions for the next integration interval. This process is repeated, yielding $x(kh)$ for $k = 1, 2, \dots, n$. For systems of first-order differential equations, the integration is applied to each equation in parallel for each step.

The fourth-order integration formula used is

$$\hat{x}(t+h) = \hat{x}(t) + h(k_1 + 2k_2 + 2k_3 + k_4)/6 \quad (31)$$

$$k_1 = f(x, t) \quad (32)$$

$$k_2 = f(x + hk_1/2, t + h/2) \quad (33)$$

$$k_3 = f(x + hk_2/2, t + h/2) \quad (34)$$

$$k_4 = f(x + hk_3, t + h). \quad (35)$$

Four evaluations of $f(x, t)$ are required for each integration step. This formula reduces to Simpson's rule if $f(x, t)$ is a function of t only.

The integration truncation error committed at each integration interval is dependent upon the size of the interval h and for an m th-order Runge-Kutta method the error is, in general, $O(h^{m+1})$. The actual error committed is further dependent in a complicated manner upon the dynamics of the equation. More specifically, a smaller interval will be required to maintain the same error bounds during rapidly varying periods of the solution function $x(t)$ than will be required when the solution function is slowly varying. Since $f(x, t)$ must be evaluated four times for each integration interval and $f(x, t)$ may be quite complicated and computationally expensive, it is desirable to keep h as large as possible without incurring excessive error accumulation. For this reason the integration interval is adaptively increased and decreased as the integration proceeds.

To determine the proper integration interval during the integration process, an estimate of the truncation error being committed at each step is necessary. Several methods for estimating truncation error and adjusting step size have been reported in the literature (e.g., Collatz¹⁰); however, these methods are often nearly as time-consuming as the integration itself or suffer from certain serious failure modes. The Collatz method, for example, is only a qualitative measure, which is quite efficient, but suffers from a divide-by-zero problem when the slope of $x(t)$ becomes zero. A new method employed in this implemen-

tation of the algorithm was developed by the author to avoid excessive computation and eliminate the possibility of failure due to special cases of the solution. This method is also a qualitative measure and does not give a very accurate estimate of the true truncation error, but empirical study indicates it does yield a figure of merit that is within about an order of magnitude of the true error. By specifying figure-of-merit bounds that are an order of magnitude smaller than the maximum error required, we can easily maintain a sufficiently small integration interval without excessive computation or other complications.

This error estimator is based on a measure of the difference in truncation error expected between a second-order Runge-Kutta method and the fourth-order Runge-Kutta. Equation (35) is a second-order predictor for the solution at the end of the integration interval, $x(t + h)$. Recall that the truncation error committed at each step by an m th-order integration is $O(h^{m+1})$. By comparing k_4 to the solution obtained from the fourth-order integration and scaling the result by h^2 , an error estimate proportional to the fourth-order truncation error is obtained. That is,

$$\text{error} = \alpha(h^2) |k_4 - \hat{x}(t + h)|, \quad (36)$$

where $\hat{x}(t + h)$ is the fourth-order solution for $f(x, t)$, h is the integration interval, and α is a constant (usually between 0.1 and 10). If the error estimate exceeds a predetermined bound, then the integration interval is reduced to one-half its current value. Likewise, if the error estimate is more than (say) a thousand times less than the error bound, then the interval is increased to twice its previous value.

APPENDIX B

State Equation for the Two-Fingered Gripper

The following listing is written in the C programming language. It describes the complete state equation for the two-fingered gripper with common-mode feedback and the electrical time constants.

```
int gripper (t, y, dy)
double t, y[], dy[];
{
    double u1(), u2();
    double fabs(), log10();
    double kg, pg, f, f1;
    /* state variables for the robot gripper:

    0. left finger position
    1. left finger velocity
    2. left motor angle
```

```

3. left motor angular velocity
4. right finger position
5. right finger velocity
6. right motor angle
7. right motor angular velocity
8. rubber ball position
9. rubber ball velocity
10. left vel. RC network
11. left acc. RC network
12. right vel. RC network
13. right acc. RC network
14. common-mode vel. RC network
15. common-mode acc. RC network
16. left motor driver RC
17. right motor driver RC
*/
kg = 4.16e6; /* spring constant and damping coefficient
for gears */
pg = 322.32;
/* compute state derivative function for left hand finger
system */
dy [0] = y[1];
if(fabs((f = rg * y[2] - y[0])) > H) { /*gear force */
    f1 = (f > 0. ? (f-H) : (f+H)) * kg + (rg * y[3] - y[1]) *
        pg;
    if(f > 0.) {
        f1 = (f1 > 0. ? f1 : 0.);
    } else {
        f1 = (f1 < 0. ? f1 : 0.);
    }
} else {
    f1 = 0.;
}
if((if = y[0] - y[8] + R) > 0.) { /* rubber ball force */
    f2 = (f * k3 + (y[1] - y[9]) * p3);
    f2 = (f2 > 0. ? fs : 0.);
} else {
    f2 = 0.;
}
dy [1] = (-P * y[1] + f1 - f2)/M;
/* compute left motor current and voltage */
Ia1 = y[16];
if(fabs(Ia1) > Imax) {
    Ia1 = (Ia1 > 0. ? Imax : -Imax);
}

```

```

    }
    V1 = Ra * Ia1 + La * (Ia1 - Ia1sv)/dt + Ka * y[3];
    if(fabs(V1) > Vmax){
        V1 = (V1 > 0. ? Vmax : -Vmax);
        Ia1 = (V1 + La * Ia1sv/dt - Ka * y[3]) / (Ra + La/dt);
    }
    dy[2] = y[3];
    dy[3] = (Km * Ia1 - p * y[3] - f1 * rg)/m; /*torque */
/* compute state derivative for right hand finger system */
    dy[4] = y[5]
    if (fabs ((f = rg * y[6] - y[4])) > H){ /* gear force */
        f1 = (f > 0. ? (f - H) : (f + H)) * kg + (rg * y[7] - y[5]) *
            pg;
        if (f > 0.){
            f1 = (f1 > 0. ? f1 : 0.);
        }else{
            f1 = (f1 < 0. ? f1 : 0.);
        }
    }else{
        f1 = 0.;
    }
    if ((f = y[8] + R - y[4]) > 0.){ /*rubber ball force */
        f3 = f * k3 + (y[9] - y[5]) * p3;
        f3 = (f3 > 0. ? f3 : 0.);
    }else{
        f3 = 0.;
    }
    dy[5] = (-P * y[5] + f1 + f3)/M;
/* compute right motor current and voltage */
    Ia2 = y[17];
    if(fabs(Ia2) > Imax){
        Ia2 = (Ia2 > 0. ? Imax : -Imax);
    }
    V2 = Ra * Ia2 + La * (Ia2 - Ia2sv)/dt + Ka * y[7];
    if(fabs(V2) > Vmax){
        V2 = (V2 > 0. ? Vmax : -Vmax);
        Ia2 = (V2 + La * Ia2sv/dt - Ka * y[7]) / (Ra + La/dt);
    }
    dy[6] = y[7];
    dy[7] = (Km * Ia2 - p * y[7] - f1 * rg)/m;
/* compute state derivative for rubber ball */
    dy[8] = y[9];
    dy[9] = (f2 - f3)/m3;
/* state equations for electronic circuitry */

```

```

dy[10] = (y[0] - y[10])/RC;
dy[11] = (dy[10] - y[11])/RC;
dy[12] = (y[4] - y[12])/RC;
dy[13] = (dy[12] - y[13])/RC;
dy[14] = ((y[0] + y[4]) - y[14])/RC;
dy[15] = (dy[14] - y[15])/RC;
dy[16] = (-a * dy[11] - b * dy[10] + (c + f) * (u1(t) - y[0])
  - d * dy[15] - e * dy[14] + f * (u2(t) - y[4]) - y[16])/
  1.2e - 4;
dy[17] = (-a * dy[13] - b * dy[12] + (c + f) * (u2(t) - y
  [4]) - d * dy[15] - e * dy[14] + f * (u1(t) - y[0]) -
  y[17])/1.2e - 4;
return (18);
}

```

AUTHOR

Michael K. Brown, B.S., 1973, M.S., 1977, and Ph.D., 1981 (Electrical Engineering), University of Michigan, Ann Arbor; Burroughs Corporation, 1973-1980; AT&T Bell Laboratories, 1980—. From 1973 to 1976 Mr. Brown was employed with the Burroughs Corporation and was involved in the development of ink jet printer systems. From 1976 to 1980 he continued his work with Burroughs as a consultant while pursuing his Ph.D. degree at the University of Michigan. His thesis was in the area of image processing and pattern recognition. In 1980 he joined the Speech Processing Group at AT&T Bell Laboratories, where he was involved in research in speech recognition and synthesis techniques. Since 1983, he has been with the Robotics Principles Research Department.

Determination of Fiber Proof-Test Stress for Undersea Lightguide Cable

By T. C. CHU and H. C. CHANDAN*

(Manuscript received July 24, 1984)

A method is developed to determine the fiber strength requirements for the undersea lightguide cable. This method utilizes the theories involved in fiber proof testing, undersea cable dynamics, and nonlinear tensile behavior of cable. The method is illustrated with a sample cable currently being developed for the transatlantic system. This method provides great flexibility in designing undersea lightguide cable. It can be applied to justify a new cable design at an early stage or to select proper fiber proof-test levels for an existing cable under different operating conditions, i.e., deeper ocean, worse sea state, or faster recovery.

I. INTRODUCTION

Because of the difficulty and high cost of repair, undersea cable requires high reliability. Operations such as laying, recovery, and holding impose high tensions on the cable.¹ The high tension causes high strain and permanent deformation in the cable, and it may reduce or impair system reliability. Much experience has been gained over the past decades in designing conventional undersea cable with metallic conductors to withstand the high tension. In general, cable strains well in excess of 1 percent can be tolerated for the coaxial cables with metallic conductors due to the ductile nature of metals.

For the new undersea cable that has lightguides as the signal-carrying medium, the problems of high tension and strain cause greater

* Authors are employees of AT&T Bell Laboratories.

Copyright © 1985 AT&T. Photo reproduction for noncommercial use is permitted without payment of royalty provided that each reproduction is done without alteration and that the Journal reference and copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free by computer-based and other information-service systems without further permission. Permission to reproduce or republish any other portion of this paper must be obtained from the Editor.

concerns because of the static and dynamic fatigue of optical fibers.^{2,3} To ensure that the fibers survive the high-tension cable operations (laying, recovering, and holding) and last for 25 years, the entire length of each fiber is proof tested^{2,3} to a given stress before cabling. The proof-test level is chosen by taking into account the load-time history of the fibers during the entire service period. An appropriate proof-test level must be chosen since an underestimated proof-test stress reduces the system reliability, and an overestimated one increases system cost due to reduced fiber yield or increased fiber loss caused by more frequent repair splices during proof testing.

In this paper a method is developed to determine the fiber proof-test stress for undersea lightguide cable. In Section II the theory of proof testing is discussed, with special attention to the material constants used in the calculations. In Section III the cable tensions as function of time are analyzed. In Section IV the nonlinear tensile property of the cable is discussed and is used to obtain the cable and fiber strain. The fiber strain is then used to determine the fiber proof-test stress. In Section V the method is applied to a sample cable that is currently under development for transatlantic systems.

II. PROOF-TEST LEVEL

It is well established that the strength of glass optical fibers is determined by micrometer- or submicrometer-size surface or subsurface flaws. These flaws grow with time under the combined action of tensile stress, temperature, and ambient moisture. These fracture-initiating flaws are randomly distributed along the length of the fiber. Therefore, the entire length of the fiber is proof tested to screen out flaws larger than those that correspond to the proof-test level. For example, the largest initial flaw present in a given length of 200 kpsi, proof-tested, fused-silica fiber is $0.212 \mu\text{m}$.⁴

Subsequently, the proof-tested fiber is subjected to varying stresses for different time periods during cable manufacture, laying, recovery, and repair. During these operations, the preexisting flaws in the fiber grow, and this flaw growth is cumulative. At present, no commercial scheme is available that will prevent crack growth in fibers. Therefore, we choose the proof-test level that restricts crack growth during expected worst-service conditions to yield a desired design life for the fibers. In other words, by proof testing we ensure that the initial crack size is sufficiently small so that the full range of stresses to which the fiber is subjected in service will not cause the fiber to rupture.

Assuming a power law for crack growth, the largest crack, a_p , in a proof-tested fiber will grow at a rate $(da)/(dt)$ given by⁵

$$\frac{da}{dt} = AK_f^N, \quad (1)$$

where K_I is the stress-intensity factor (the subscript I refers to the crack-opening mode) at the crack tip and is given by

$$K_I = Y\sigma a^{1/2}, \quad (2)$$

where Y is a geometrical factor related to the shape of the crack and the specimen (≈ 1.241 for semicircular part-through crack, normal to the axis of a cylindrical fiber); σ is the applied stress. The characteristic crack-growth parameters A and N are coupled and are empirically determined from dynamic and static fatigue tests.⁶ From a collection of experimental data covering several independent investigations, we have observed that the relationships between $\log A$ and N are different for dynamic and static fatigue situations. For short-term loads A_d , increasing with time, our studies have indicated that $\log A_d = 3.22N - 8.516$. For a static load A_s , $\log A_s = 3.289N - 10.05$. The choice of N depends on the Relative Humidity (RH) in the ambient environment: $N = 14$ to 16 for wet environment (97-percent RH) and $N = 17$ to 20 for typical laboratory environment (40- to 60-percent RH).

From eqs. (1) and (2), an applied stress of σ_1 for a time t_1 will cause the crack to grow from the proof-test level of a_p to a_1 and can be calculated by

$$a_1^{-\frac{N}{2}+1} - a_p^{-\frac{N}{2}+1} = \frac{2-N}{2} \int_0^{t_1} A Y^N [\sigma_1(t)]^N dt, \quad (3)$$

where a_1 is the flaw size after the fiber has been under an applied stress $\sigma_1(t)$ for a time t_1 . The fused silica optical fibers behave in a linear elastic manner up to the breaking point; therefore, the stress $\sigma_1(t)$ in eq. (3) can be expressed in terms of strain, $\epsilon_1(t) = [\sigma_1(t)]/E$, where E is the Young's elastic modulus of the fiber ($E = 71.9 \text{ GPa}$ or $10.4 \times 10^6 \text{ psi}$). Furthermore, when the fibers are subjected to a sequence of stress loadings, the flaw growths are additive.⁴ From these considerations, eq. (3) can be rewritten as

$$a_p^M = a_f^M - \sum_{i=1}^Q A M E^N Y^N \int_0^{t_i} [\epsilon_i(t)]^N dt, \quad (4)$$

where $M = (2 - N)/2$ and a_f is the final flaw size after the sequence of stress loadings. If the flaw growth is considerable, say $a_f/a_p \geq 10$, and we assume that N is in the range of 14 to 20 ($M \approx -7$), then

$$\frac{a_f^M}{a_p^M} \approx 10^{-7}.$$

Thus, the term a_f^M can be neglected and

$$a_p^M \cong \sum_{i=1}^Q (-AM) E^N Y^N \int_0^{t_i} [\epsilon_i(t)]^N dt. \quad (5)$$

The required proof-test level is determined from the fracture mechanics relation,

$$\sigma_p = \frac{K_{IC}}{Y a_p^{1/2}}, \quad (6)$$

where K_{IC} is the fracture toughness of the material. For fused silica fibers,

$$K_{IC} = 0.789 \text{ (unit in MPa } \sqrt{m}\text{)},$$

and as previously mentioned, $Y \cong 1.241$. Hence,

$$\sigma_p = \frac{0.6357}{a_p^{1/2}}, \quad (7)$$

where a_p is determined from eq. (5) in meter and σ_p is in MPa. Thus, the fibers proof tested at σ_p will endure the stress loadings during cable laying, recovery, repair, and long-term service without fracture.

III. CABLE TENSIONS DURING THE OPERATIONS

To evaluate the fiber proof-test stress, the fiber strain ϵ_f in eq. (5) has to be determined. The amount of fiber strain is related to the cable tension, cable tensile property—i.e., tension versus strain relation—and the coupling between the fibers to the cable structure. In a cable with its fibers fully coupled to the cable structure, the fiber strain is equal to the cable strain, and this is the worst case. By knowing the cable tensions as functions of time during each operation, the cable strain as a function of time can be obtained from the cable tensile property. The cable tension as a function of time can be derived from the maximum tension at the ship, as shown below.

3.1 Laying

In a normal cable-laying operation, the cable forms a straight-line configuration, as shown in Fig. 1a. The cable tension at the ocean bottom is theoretically equal to zero. The cable tension at the ship is equal to wh ,¹ where w is the cable weight per unit length in seawater and h is the ocean depth. By following a cable element during the laying operation, the time-dependent tension experienced by the element can be found as

$$T(t) = wh - \frac{wh}{\left(\frac{h}{v_l \sin(\alpha_s)}\right)} t, \quad 0 \leq t \leq \frac{h}{v_l \sin(\alpha_s)}, \quad (8)$$

where v_l is the cable-laying speed, t is the elapsed time after the cable element leaves the ship and enters the water, and α_s is the angle between the cable and sea level and is

$$\alpha_s = \frac{H}{v_l},$$

where H is the hydrodynamic constant of the cable. For most undersea cable used in deep water, H ranges between 30 to 50 degree-knots. The cable-laying speed is usually about 7 to 8 knots, thus, the angle α_s is less than 10 degrees. For such small angles, the contributions of wave motion and repeater weight to the tension are negligible. Equation (8) is sketched in Fig. 2a.

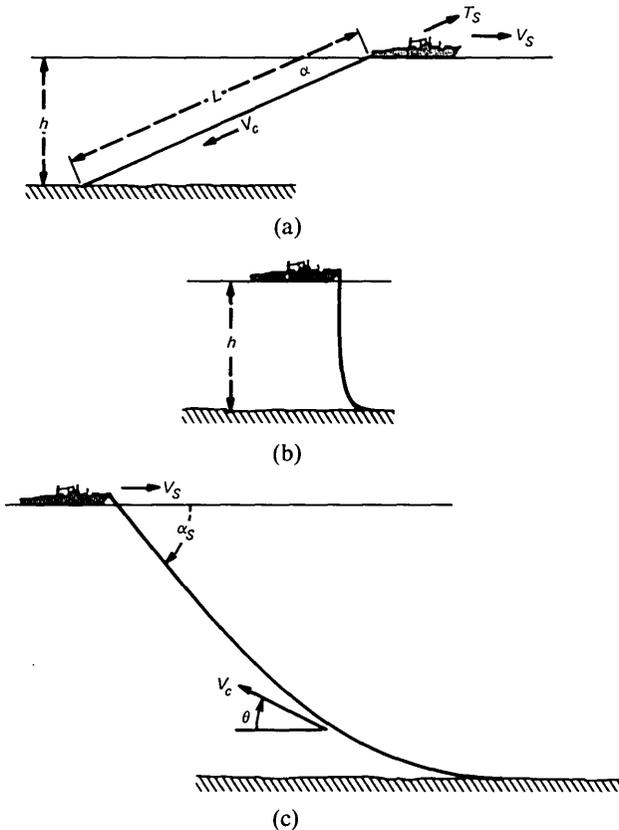


Fig. 1—Cable operations. (a) Laying. (b) Testing or splicing. (c) Recovery.

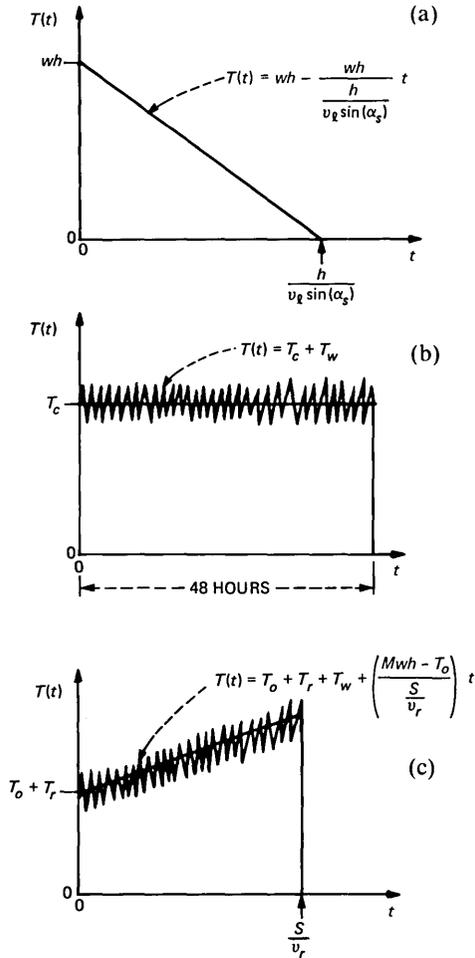


Fig. 2—(a) The time-dependent tensions experienced by the cable element during the laying operation. (b) Upper bound for cable tension during testing. (c) Catenary formed by the cable during recovery.

3.2 Recovery

During recovery, the cable forms a catenary, as shown in Fig. 1b, and the length of the catenary S can be evaluated.¹ For a given recovery speed v_r , the recovery time of a cable element is given by

$$t = \frac{S}{v_r}.$$

Since the cable forms a catenary, the cable tension at the ocean bottom, T_o , is not zero. The cable tension at the ship, not including

the effects of wave motion and repeater weight, is $\bar{M}wh$, where \bar{M} is a magnification factor that indicates the difference in tension between laying and recovery. The factor \bar{M} has the following form:¹

$$\bar{M} = \frac{1}{1 - \tan^2 \frac{\alpha}{2} \left(\frac{\cos \alpha + \cos \alpha_s}{1 - \cos \alpha \cos \alpha_s} \right)^{\frac{1}{\gamma}}},$$

$$\gamma = \frac{\sin^2 \alpha}{2 - \sin^2 \alpha},$$

where α_s is the angle between the cable and sea level and α is the ratio between the cable hydrodynamic constant H and the recovery speed v_r . Assuming that the tension experienced by a cable element during recovery varies linearly with time, the cable tension as a function of time—including the wave motion and repeater weight—turns out to be

$$T(t) = T_0 + \left(\frac{\bar{M}wh - T_0}{\frac{S}{v_r}} \right) t + T_w + T_r,$$

$$0 \leq t \leq \frac{S}{v_r}, \quad (9)$$

where T_r is the weight of a cable-carried point mass (such as a repeater) in seawater, and T_w is the additional tension due to the wave motion. T_w is expressed by

$$T_w = \sqrt{\overline{EA}} \rho_c v_y \cos \frac{2\pi t}{\tau} \cos \alpha_s,$$

where \overline{EA} is the cable tensile stiffness, ρ_c is the cable mass per unit length, v_y is the maximum ship vertical velocity, τ is the period of wave motion, and α_s is the recovery angle. In normal recovery, the angle α_s is kept close to 90 degrees to reduce the tension, thus $\cos \alpha_s \approx 1$. Equation (9) is expressed in Fig. 2b.

3.3 Holding

After its recovery, the cable is held for a period of time for testing and repair. At present, the required splicing time for the lightguide cable is estimated at 16 hours. Assuming (pessimistically) that three attempts are made before a successful repair is executed, then the total holding time is 48 hours. During holding period, the cable also forms a catenary that is essentially stationary except for the effect of wave motion, as shown in Fig. 1c. The tension experienced by the cable element at the ship can be shown as

$$T(t) = T_c + T_w \quad 0 \leq t \leq t_0, \quad (10)$$

where t_0 is the holding time, T_w is the additional tension due to the wave motion, T_c is the tension due to the catenary and can be evaluated from the theory of a catenary as⁷

$$T_c = wC \cos h(z),$$

where w is the cable weight per unit length in water, and C and z are constants to be determined from the ocean depth h and the cable angle α_s at holding. Because the cable sections suspended near the ocean bottom experience lower tension than the cable elements near to the ship while holding, the survival of cable elements at the ship guarantees the survival of the remaining cable. In short, $T(t) = T_c + T_w$ expresses the upper bound for cable tension while loading. This equation is represented in Fig. 2c.

IV. CABLE TENSILE PROPERTIES

By knowing the cable tensions during the operations, the cable strain can be found if the tensile property of the cable is known. A method has been developed to accurately predict the tensile property of undersea cable in both elastic and plastic regions.⁸ The method relates the cable tension T to cable strain ϵ_c by a fifth-degree polynomial as

$$T = C_0 + C_1\epsilon_c + C_2\epsilon_c^2 + C_3\epsilon_c^3 + C_4\epsilon_c^4 + C_5\epsilon_c^5,$$

where C_0 through C_6 are constants depending on the mechanical properties of the constituent cable components. An example of such a polynomial is shown in Fig. 3. Since our present need is to find the

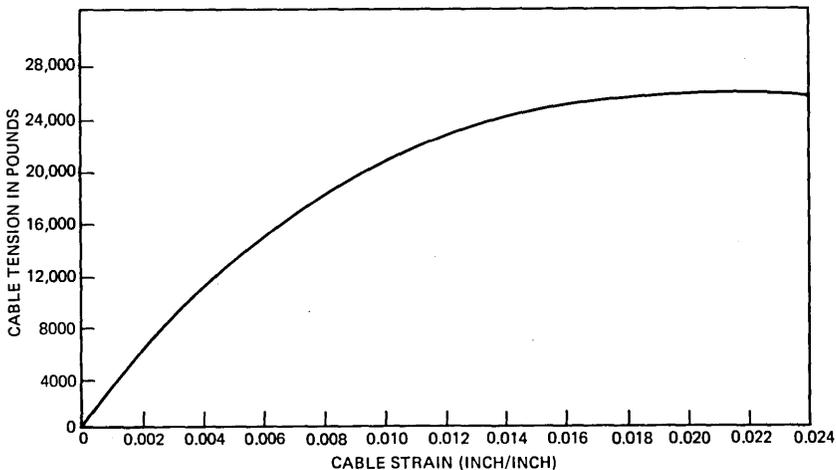
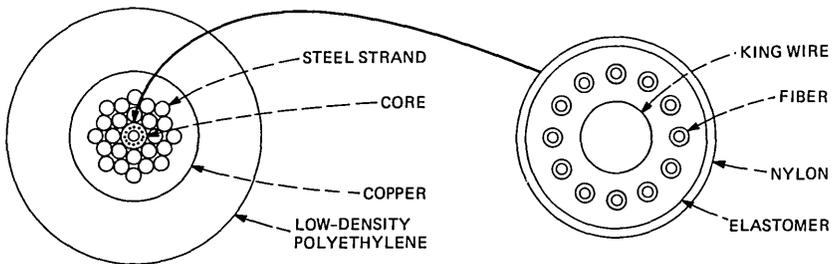


Fig. 3—Tensile properties of the TAT-8 transatlantic cable.



CABLE STRUCTURE:

STRAND DIAMETER = 9.47 mm (0.373 in.)
 CONDUCTOR OD (COPPER) = 10.46 mm (0.412 in.)
 INSULATION OD = 21 mm (0.827 in.)

CABLE CORE:

KING WIRE OD = 0.71 mm (0.028 in.)
 NUMBER OF FIBERS = 12
 FIBER OD (COATED) = 250 μm (0.010 in.)
 NYLON THICKNESS = 0.1 mm (0.004 in.)
 CORE OD = 2.97 mm (0.117 in.)

Fig. 4—Cable structure and dimensions for the TAT-8 transatlantic cable.

cable strain for a given tension, an inverse function of the above relation is required. This can be done by the numerical regression method. The resulting relation is

$$\epsilon_c = \bar{C}_0 + \bar{C}_1 T + \bar{C}_2 T^2 + \dots + \bar{C}_k T^k, \quad (11)$$

where k and \bar{C}_0 through \bar{C}_k are determined with desired accuracy from the numerical regression.

Since the cable tension during its operations is a function of time $T(t)$, eq. (11) gives the cable strain as a function of time, i.e., $\epsilon(t)$. If the fibers are fully coupled to the cable structure, the fiber strain is equal to the cable strain $\epsilon(t) = \epsilon_c(t)$, and this is the strain required in eq. (5) to evaluate the fiber proof-test level.

V. APPLICATIONS

This method of testing fiber strength is applied to the cable that is currently being developed for the TAT-8 transatlantic system. The cable construction and dimensions are shown in Fig. 4. The cable and fiber properties required in the proof-test stress evaluation are listed in Table I. The cable tension during each operation is evaluated under realistic operating conditions.

Table I—Cable and fiber properties

Cable weight per unit length in seawater w	0.326 lb/ft
Cable hydrodynamic constant H	43.5 degree-knots
Cable mass per unit length ρ_c	$0.0174 \frac{\text{lb} - \text{s}^2}{\text{ft}^2}$
Cable tensile stiffness \overline{EA}	$2.9 \times 10^6 \text{ lb}$
Fiber "N" number	17
Fiber tensile modulus E	$10.4 \times 10^6 \text{ psi}$

5.1 Laying

Following are the conditions under which cable was laid:

1. Laying speed $v_l = 7$ knots
2. Ocean depth $h = 18,000$ ft (transatlantic installation)
3. Cable angle $\alpha_s = H/v_l = 6.22$ degrees
4. Cable touchdown time $= h/[v_l \sin(\alpha_s)] = 3.93$ hours.

Considering the above conditions and the cable properties listed in Table I, eq. (8) becomes

$$T(t) = 5875 - \frac{5875}{3.93 \times 60 \times 60} t$$
$$0 \leq t \leq 3.90 \times 60 \times 60 \text{ seconds.} \quad (12)$$

5.2 Recovery

Following are the conditions under which the cable was recovered:

1. Recovery speed $v_r = 1$ knot
2. Recovery angle $\alpha_s = 85$ degrees
3. Vertical ship velocity due to wave motion $v_y = 17$ ft/s (sea state 6)
4. Ocean depth $h = 18,000$ ft
5. $wh = 5875$ lb
6. Additional tension due to wave motion $T_w = 3829$ lb. (To simplify the integration, the tension due to the wave motion is assumed to be equal to $\sqrt{EA\rho_c}v_y$, the amplitude of the oscillating tension. This simplification results in a more conservative fiber proof-test stress.)
7. $\bar{M}wh = 13,488$ lb
8. Cable bottom tension $T_0 \equiv \bar{M}wh - wh = 7613$ lb
9. Suspended cable length¹ $= 1.65 \times 18,000 = 29,700$ ft
10. Recovery time $t = S/v_r = 4.89$ hours
11. Repeater weight in seawater $T_r = 1000$ lb (presently, the repeater weight has been reduced to 300 lb in water).

With the above conditions and cable properties, eq. (9) becomes

$$T(t) = (7613 + 3829 + 1000) + 0.33373t$$
$$0 \leq t \leq 4.89 \times 60 \times 60 \text{ seconds.} \quad (13)$$

5.3 Holding

Following are the conditions during holding:

1. Ocean depth $h = 18,000$ ft
2. Cable angle at holding $\alpha_s = 75$ degrees
3. Additional tension due to wave motion $T_w = 3829$ lb
4. Cable weight per unit length in water $= 0.326$ lb/ft.

By knowing the ocean depth h and cable angle α_s , the constant C and z can be evaluated from the catenary theory as

$$C = 6299.5 \text{ ft}$$

$$z = 2.0563$$

$$T_c = WC \cos h(z) = 8133 \text{ lb.}$$

Thus, eq. (10) becomes

$$\begin{aligned} T(t) &= 8133 + 3829 \\ 0 \leq t &\leq 48 \times 60 \times 60 \text{ seconds.} \end{aligned} \quad (14)$$

5.4 Cable tensile properties

The theoretical tensile property of the submarine cable evaluated by the method previously discussed is shown in Fig. 4.⁸ The inverse polynomial is found to have the following form:

$$\epsilon = \bar{C}_0 + \bar{C}_1 T + \bar{C}_2 T^2 + \bar{C}_3 T^3 \quad (15)$$

for $0 \leq T \leq 21,300$ lb, where

$$\bar{C}_0 = -2.9758111668 \times 10^{-5}$$

$$\bar{C}_1 = +3.96935593146 \times 10^{-7}$$

$$\bar{C}_2 = -8.1189437448 \times 10^{-12}$$

$$\bar{C}_3 = +6.40675624028 \times 10^{-6}$$

and the standard error $\chi = 6.9 \times 10^{-5}$. It is noted that eq. (15) applies for cable tension below 21,300 lb. It is adequate because the maximum tension expected by the cable during its operations is about 17,500 lb.

5.5 Proof-test stress evaluation

The cable tension $T(t)$ in each of eqs. (12), (13), and (14) is substituted into eq. (15) to obtain $\epsilon_1(t)$, $\epsilon_2(t)$, and $\epsilon_3(t)$. Substitution of $\epsilon_1(t)$, $\epsilon_2(t)$, and $\epsilon_3(t)$ into eq. (6) for $\epsilon_i(t)$ gives, upon numerical integration, the initial flaw size a_p :

$$a_p^{-7.5} = 5.5685763 \times 10^{39} + 5.139996 \times 10^{48} + 5.0337450 \times 10^{45}$$

(laying) (recovery) (holding)

or

$$a_p = 3.2 \times 10^{-7} \text{ meter} = 0.32 \text{ } \mu\text{m.}$$

The corresponding proof-test stress is obtained from eq. (7):

$$\sigma_p = 1.124 \text{ GPa} = 163 \text{ kpsi.}$$

Thus, a proof-test stress greater than 163 kpsi should guarantee the survival of the fibers in the cable. Generally, however, a somewhat higher proof-test level (200 kpsi) is specified for transatlantic cable to ensure that adequate margin is provided to offset any subtle effects

brought about by the cable manufacturing and handling processes that are not covered by the model.

VI. CONCLUSIONS

A method is developed to evaluate the required proof-test stress for the fibers used in the undersea cable. The method provides great flexibility in designing the cable. It can be applied to justify a cable design and verify the adequacy of a fiber proof-test level at an early stage, or to select proper fiber proof-test levels for an existing cable under different operating conditions, i.e., deeper ocean, worse sea state, or faster recovery.

VII. ACKNOWLEDGMENTS

The authors would like to thank R. D. Tuminaro and D. Kalish for their careful review and useful comments on the work.

REFERENCES

1. E. E. Zajac, "Dynamics and Kinematics of the Laying and Recovery of Submarine Cable," *B.S.T.J.*, 35, No. 5 (September 1957), pp. 1129-207.
2. S. E. Miller and A. G. Chynoweth, eds., *Optical Fiber Telecommunications*, New York: Academic Press, 1979, Chapter 12.
3. A. S. Tetelman and A. J. McEvily, *Fracture of Structural Materials*, New York: Wiley, 1967.
4. D. Kalish, B. K. Tariyal, and H. C. Chandan, "Effect of Moisture on the Strength of Optical Fibers," 27th International Wire and Cable Symposium, Cherry Hill, New Jersey.
5. A. G. Evans and S. M. Wiederhorn, "Proof Testing of Ceramic Materials—An Analytical Basis for Failure Prediction," *Int. J. Fract.*, 10, 1974, p. 379.
6. David Kalish and B. K. Tariyal, "Static and Dynamic Fatigue of a Polymer-Coated Fused Silica Optical Fiber," *J. Amer. Ceram. Soc.*, 61, No. 11-12 (November-December 1978), pp. 518-23.
7. T. Baumeister, E. A. Arallone, and T. Baumeister III, *Mark's Standard Handbook for Mechanical Engineers*, New York: McGraw-Hill, 8th Ed., pp. 2-36-38.
8. T. C. Chu, "A Method to Characterize the Mechanical Properties of Undersea Cables," *B.S.T.J.*, 62, No. 5 (March 1983), pp. 703-15.

AUTHORS

Harish C. Chandan, B. Tech. and M. Tech. (Metallurgical Engineering), in 1969 and 1971, respectively, Indian Institute of Technology, Kanpur, India; M.S. (Solid State Science), 1974, Syracuse University; Ph.D. (Ceramics Science), 1977, Pennsylvania State University; AT&T Bell Laboratories, 1977—. Since joining AT&T Bell Laboratories, Harish has worked in the area of strength, fatigue, and reliability of optical fibers. He is currently involved in dehydration and sintering of soot preforms for optical fibers made by the outside deposition process. Member, American Ceramic Society, National Institute of Ceramic Engineers.

Tek-Che Chu, B.S., 1964, Cheng Kung University (Taiwan); M.S., 1967, Syracuse University; Ph.D., 1971 (Aerospace Engineering), Cornell University; AT&T Bell Laboratories, 1972—. Mr. Chu has worked on the physical design of T4M digital transmission system and the development of an optical fiber connector. Since 1981 he has been working on the design of undersea lightguide cable.

Beamwidth and Useable Bandwidth of Delay-Steered Microphone Arrays

By J. L. FLANAGAN*

(Manuscript received October 31, 1984)

Automatic delay steering of microphone arrays improves sound pickup in large rooms. But steering to wave-arrival directions away from broadside degrades the acuity of the beam and diminishes the useable bandwidth of the array. This paper derives quantitative relations for the variation in beamwidth and bandwidth for one-dimensional, uniform, unweighted arrays composed of $(2N + 1)$ receivers spaced by distance d . The results show how the upper and lower useful frequencies and the beam acuity are conditioned not only by receiver spacing and frequency, but also by wave-arrival direction and steering direction. The relations developed permit detailed design of steerable arrays for specified frequency range and spatial coverage.

I. MICROPROCESSOR CONTROL OF MICROPHONE ARRAYS

Speech transduction or sound pickup in large rooms—such as auditoria or classrooms—traditionally is plagued by the distortions of room reverberation and noise interference from unwanted sources. The problem is minimized by obtaining, at the transducer, the greatest intensity possible for the signal travelling the direct path from source to receiver, and the least intensity possible for multipath room reflections of that signal and any additive interfering noise. Arrays of microphones, especially two-dimensional arrays designed for high directivity, are useful for achieving this high ratio of direct to distorted sound.

* AT&T Bell Laboratories.

Copyright © 1985 AT&T. Photo reproduction for noncommercial use is permitted without payment of royalty provided that each reproduction is done without alteration and that the Journal reference and copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free by computer-based and other information-service systems without further permission. Permission to reproduce or republish any other portion of this paper must be obtained from the Editor.

But in an auditorium or conference room the desired sound source may shift position. The directional array must be steered, or pointed, to the shifting source to gain the full benefit. Rapid, automatic, electronic steering of the array is therefore attractive. Currently emerging microprocessors have the computational capability to perform beam forming and beam steering dynamically, and can be programmed to seek and track talkers in a room.¹

Beam steering is implemented conveniently by digital delay adjustment of the signals received at each microphone of the array. But forming a beam and steering it to angles away from the normal to the array (i.e., away from "broadside") have two effects that must be taken into account when designing the system. One is that the useful bandwidth diminishes as the steering direction is made acute to the axis of the array (or to the plane of the array for a two-dimensional system). Another is that the beam pattern becomes asymmetric and less sharp. This report considers the variation in beamwidth and useable bandwidth under conditions of delay steering.

II. TIME, FREQUENCY, AND SPATIAL RESPONSES FOR THE UNIFORM DELAY-STEERED ARRAY

The most convenient expository vehicle is the one-dimensional uniform array, unweighted in amplitude. Related analyses apply for two-dimensional and three-dimensional arrays.

Figure 1 shows a line of $(2N + 1)$ receivers, spaced uniformly by distance d . All receivers have the same sensitivity, taken to be real and equal to unity. For beam forming and steering, the individual receiver outputs are passed through controllable delays and are summed to produce the array output.

For a plane, impulsive sound wave of unit amplitude, arriving from the direction of polar angle ϕ , the time-domain and frequency-domain responses of the array are, respectively,

$$h(t) = \sum_{n=-N}^N \delta(t + nT),$$

and

$$H(j\omega) = \sum_{n=-N}^N e^{j\omega nT}, \quad (1)$$

where

t is time,

$\omega = 2\pi f$ is the radian frequency,

$\delta(\cdot)$ is the delta function, and

$T = (\tau - \tau')$ is the time difference between the interelement wave transit delay and the interelement steering delay, in which

$$\tau = \frac{d}{c} (\cos \phi) \quad \text{and} \quad \tau' = \frac{d}{c} (\cos \phi'),$$

where ϕ' is the steering angle, and c is the sound velocity (taken here as 3.4×10^4 cm/s). For notational convenience, a constant delay term that ensures formal causality has been omitted from (1). Further deserving of emphasis, the impulse response $h(t)$ and its Fourier transform $H(j\omega)$ are both implicit functions of the wave-arrival direction ϕ and of the steering angle ϕ' . For specific frequencies ω , the spatial directivity is conveniently revealed in $H(j\omega)$. An elaboration of the expression for $H(j\omega)$ is therefore useful for our purposes.

The finite geometric series of (1) can be reformulated as

$$\begin{aligned} H(j\omega) &= \sum_{n=-N}^N e^{j\omega nT} \\ &= \frac{(e^{j\omega(N+1)T} - e^{-j\omega NT})}{(e^{j\omega T} - 1)} \\ &= \frac{\sin[(2N + 1)(\omega T/2)]}{\sin(\omega T/2)}. \end{aligned} \quad (2)$$

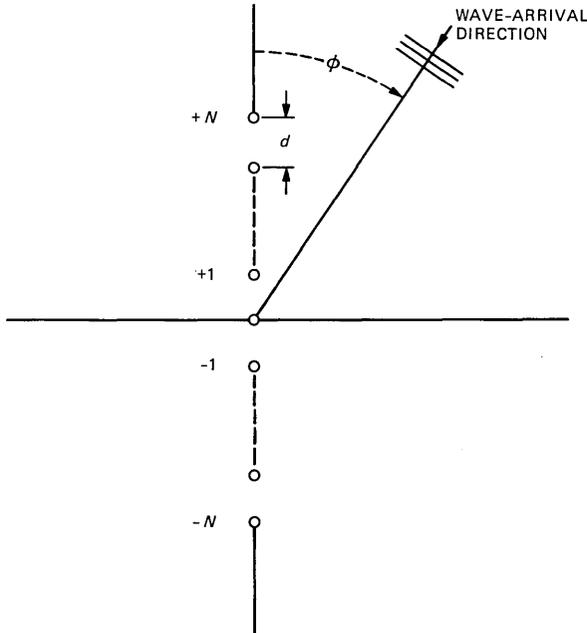


Fig. 1—Geometry for a one-dimensional microphone array of $(2N + 1)$ elements. The microphone spacing is d , and the polar coordinate for wave-arrival direction is ϕ . The array output is a summation of the outputs of individual microphones after delay adjustment.

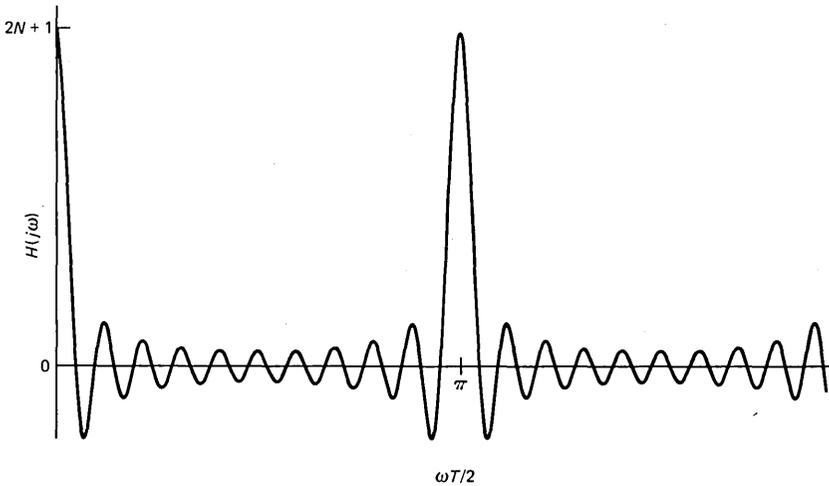


Fig. 2—Output of the uniform array characterized by $\sin[(2N + 1)(\omega T/2)]/\sin(\omega T/2)$.

Given values of the spatial coordinate ϕ and the steering parameter ϕ' , one notices that $H(j\omega)$ is periodic in ω with period equal to $2\pi/T$. This relation is illustrated in Fig. 2. Low amplitude values of $H(j\omega)$ represent conditions where the array is discriminating against sources in direction ϕ and for frequency ω . At very low frequencies, by virtue of the characteristics of (2), the array cannot discriminate as a function of wave-arrival direction. At high frequencies, near $1/T$ and multiples thereof, a similar lack of spatial discrimination must occur. This "spatial aliasing" of the discriminatory capabilities fixes an upper limit to the useful frequency bandwidth of the array. This upper limit is

$$f_{\text{upper}} = 1/|T|_{\text{max}},$$

where

$$|T|_{\text{max}} = \frac{d}{c} |\cos \phi - \cos \phi'|_{\text{max}}. \quad (3)$$

(Unlike time-domain sampling, where signal and impulse train are multiplied and the signal spectrum translated to sidebands about the sampling frequency and its multiples, this spatial aliasing acts only as a filter for the original frequency components of the source signal.)

The first zero of the response $H(j\omega)$ can be taken as a measure of the lowest frequency below which the array cannot provide spatial discrimination. This first zero, as seen from (2), corresponds to

$$(2N + 1)(\omega T/2) = \pi,$$

or

$$f_{\text{lower}} = \frac{f_{\text{upper}}}{(2N + 1)}. \quad (4)$$

As apparent from (3), given the directions ϕ , ϕ' , the upper useful frequency of the array is conditioned by the receiver spacing d . Similarly from (4), the lowest useful frequency for spatial discrimination is dictated by the array size N . Alternatively, (4) gives the required N as

$$N = \left[\frac{1}{2} \left(\frac{f_{\text{upper}}}{f_{\text{lower}}} - 1 \right) \right]_{\text{nhi}},$$

where *nhi* equals the next highest integer.

Further, using the relevant zeros of (2) (i.e., the first, and the first less than f_{upper}), an approximation of the useful discrimination bandwidth of the array is

$$BW = (f_{\text{upper}} - 2f_{\text{lower}}). \quad (5)$$

Additionally, one notes that the value of $|T|$ is conditioned by the values ϕ , ϕ' and has the limits

$$0 \leq |T| \leq 2d/c,$$

corresponding, respectively, to

$$\phi = \phi' \quad (\text{i.e., beam steered to the wave-arrival direction}),$$

and to

$$\begin{aligned} \phi = 0^\circ & \quad \text{for } \phi' = 180^\circ, \text{ or} \\ \phi = 180^\circ & \quad \text{for } \phi' = 0^\circ \end{aligned} \quad (6)$$

(i.e., the beam steered to one axial direction with the wave arrival from the opposite axial direction). As a consequence of the trigonometric relations, the useful upper and lower frequencies do not vary linearly with ϕ' . This variation is plotted in Fig. 3 for a value of $N = 10$.

Finally, if ω and ϕ' are prescribed as parameters, then (2) gives the amplitude response of the array as a function of the wave-arrival direction ϕ . The response can then be considered an explicit function of the spatial coordinate, or $H(\phi)$. Wave arrival from the steering direction, $\phi = \phi'$, gives an infinite period to (2). Also for this condition, $H(\phi) = |H(\phi)|_{\text{max}} = (2N + 1)$, and $f_{\text{upper}} \rightarrow \infty$, and the array passes all frequencies ideally. Wave arrival from a direction away from the steering direction diminishes the period of (2), as well as the output amplitude of $H(\phi)$ for frequencies between the upper and lower values. This diminution in output reflects the beam discrimination, or beam-width, of the array. The half-power output (i.e., $|H(\phi)|_{-3\text{dB}} =$

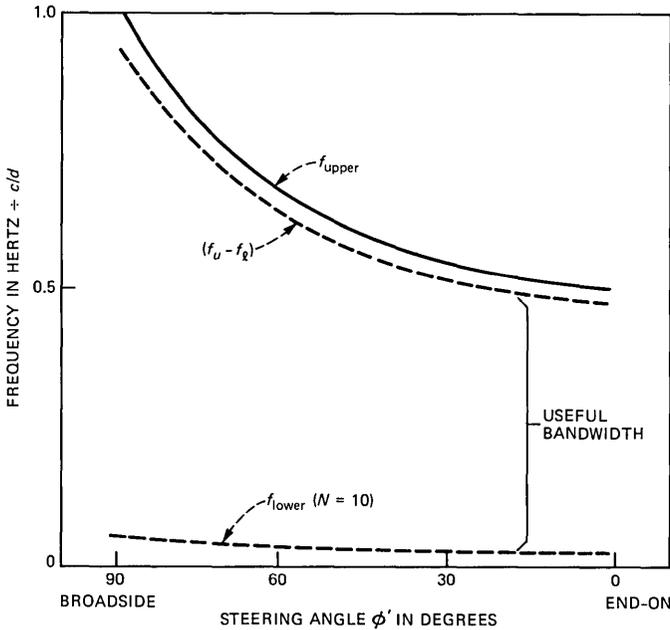


Fig. 3—Variation of the useable array bandwidth as a function of the steering direction ϕ' .

$0.707|H(\phi = \phi')|$ is frequently used to define the beamwidth. Typically, two such half-power directions exist, ϕ_1 and ϕ_2 , where $\phi_2 > \phi_1$. Their difference, $\Delta = (\phi_2 - \phi_1)$, is taken as the beamwidth.

III. BEAMWIDTH RELATIONS

Returning to the response function (2) and viewing its dependence upon the spatial coordinate, we can examine the behavior of $H(\phi)$ in the vicinity of $\phi = \phi'$. This corresponds to a value $T = 0$, and hence to $(\omega T/2) = 0$ in the argument of (2). For small magnitudes of the argument we may approximate the denominator term of (2) as

$$\sin(\omega T/2) \cong (\omega T/2), \quad |\omega T/2| \ll \pi/2. \quad (7)$$

As a consequence,

$$H(\phi) \cong (2N + 1) \left\{ \frac{\sin[(2N + 1)(\omega T/2)]}{[(2N + 1)(\omega T/2)]} \right\}. \quad (8)$$

The half-power output occurs for a value of the argument of (8) that makes the $(\sin x/x)$ term equal to 0.707, which requires

$$[(2N + 1)(\omega T/2)] = \pm 1.4 \text{ radians}. \quad (9)$$

Equation (9) therefore implies the half-power wave-arrival direction as

$$\phi_{-3\text{dB}} = \cos^{-1} \left[\cos \phi' \pm \frac{2.8c}{(2N + 1)\omega d} \right]. \quad (10)$$

Typically, there will be two half-power directions, ϕ_1 and ϕ_2 , corresponding respectively to the + and - signs in (10). It is clear from (10) that the bracketed term must have a magnitude no greater than unity to produce a valid value for $\phi_{-3\text{dB}}$. As the steering direction is made acute to the axis of the array, the value of $|\cos \phi'| \rightarrow 1$, which indicates a limiting steering direction beyond which (for a given frequency and array design) the array will cease to exhibit two directions (in the hemisphere) where its output drops to half power. This limiting direction, $\phi' = \phi'_{\text{lim}}$, also means that one value of $\phi_{-3\text{dB}}$ must be the axial direction (for example, $\phi_1 = 0^\circ$, or $\phi_2 = 180^\circ$), which implies

$$\phi'_{\text{lim}} = \cos^{-1} \pm \left[1 - \frac{2.8c}{(2N + 1)\omega d} \right]. \quad (11)$$

The positive sign for the bracketed term corresponds to $\phi_1 = 0^\circ$ with ϕ'_{lim} in the first quadrant, and the negative sign corresponds to $\phi_2 = 180^\circ$ with ϕ'_{lim} in the second quadrant.

Furthermore, the beamwidth for steering to these limits is simply

$$\Delta_{\text{lim}} = \cos^{-1} \left[1 - \frac{2(2.8c)}{(2N + 1)\omega d} \right]. \quad (12)$$

Recall that the original approximation requires $|(\omega T)/2| \ll \pi/2$, which holds for ϕ near ϕ' , and for frequencies substantially lower than $(f_{\text{upper}}/2)$ for any ϕ, ϕ' .

The spatial directivity pattern for the one-dimensional array is a figure of revolution about the array axis. (In contrast, a two-dimensional array produces a cigar-shaped beam, confined in two dimensions.) Equations (10), (11), and (12) describe the conditions for finding two half-power beamwidth directions (in the hemisphere $0 \leq \phi \leq 180^\circ$) for steering directions no more acute to the axis than ϕ'_{lim} . The complete (figure of revolution) spatial directivity pattern is therefore cone shaped, with the interior of the cone remaining hollow. These limits may be particularly appropriate to the case of the array mounted on an infinite baffle, such as the wall of a room.

If the array is unbaffled and freely suspended, steering to angles more acute than ϕ'_{lim} may be useful. In the limit ϕ' equals 0° or 180° , which is the end-fire condition. For all steering directions more acute than ϕ'_{lim} , only one half-power direction exists in the hemisphere, and

the interior of the cone-shaped directivity pattern "fills in" until the directivity pattern merges into a single main lobe that is axially directed. In this case the half-power beamwidth is simply

$$\Delta = 2\phi_2, \quad 0^\circ \leq \phi \leq 90^\circ, \quad \phi' < \phi'_{lim},$$

or

$$\Delta = 2(180^\circ - \phi_1), \quad 90^\circ \leq \phi \leq 180^\circ, \quad \phi' > \phi'_{lim}. \quad (13)$$

Additionally, the end-fire beamwidth ($\phi' = 0^\circ$ or 180°) is

$$\Delta_{end} = 2 \left\{ \cos^{-1} \left| 1 - \frac{2.8c}{(2N + 1)\omega d} \right| \right\}. \quad (14)$$

IV. REPRESENTATIVE DESIGNS

To fix the utility of these relations, let us consider a practical set of conditions. Suppose the bandwidth of interest is approximately the telephone bandwidth. This bandwidth is included within a range which, for example, we take as 170 to 3500 Hz. Further, suppose the desired range of steering is $60^\circ \leq \phi' \leq 120^\circ$, or $\pm 30^\circ$ from broadside. For this condition $|\cos \phi - \cos \phi'|_{max} = 1.5$ and the relation

$$f_{upper} = 1/|T|_{max} = 3500 \text{ Hz}$$

suggests

$$d = 6.5 \text{ cm.}$$

Similarly, the relation

$$N = \left[\frac{1}{2} \left(\frac{f_{upper}}{f_{lower}} - 1 \right) \right]_{nhi}$$

requires

$$N = 10.$$

These values can be used in eqs. (2) and (10) to examine the spatial discrimination and beamwidth under conditions of steering, both inside and outside the prescribed range of ϕ' . Figures 4 and 5 illustrate, on linear plots, the values of $|H(\phi)|$ versus ϕ for $\phi' = 90, 60, 30,$ and 0 degrees, and for frequencies two octaves apart, namely, 500 and 2000 Hz. As mentioned previously, the complete pattern of spatial selectivity for the one-dimensional array is a figure of revolution about the array axis.

Consistent with (10), Figs. 4 and 5 show that the acuity of the steered beam diminishes with frequency, and with steering directions away from the normal to the array axis. In addition, using (11) through (14) provides a summary characterization of Δ versus ϕ' , with fre-

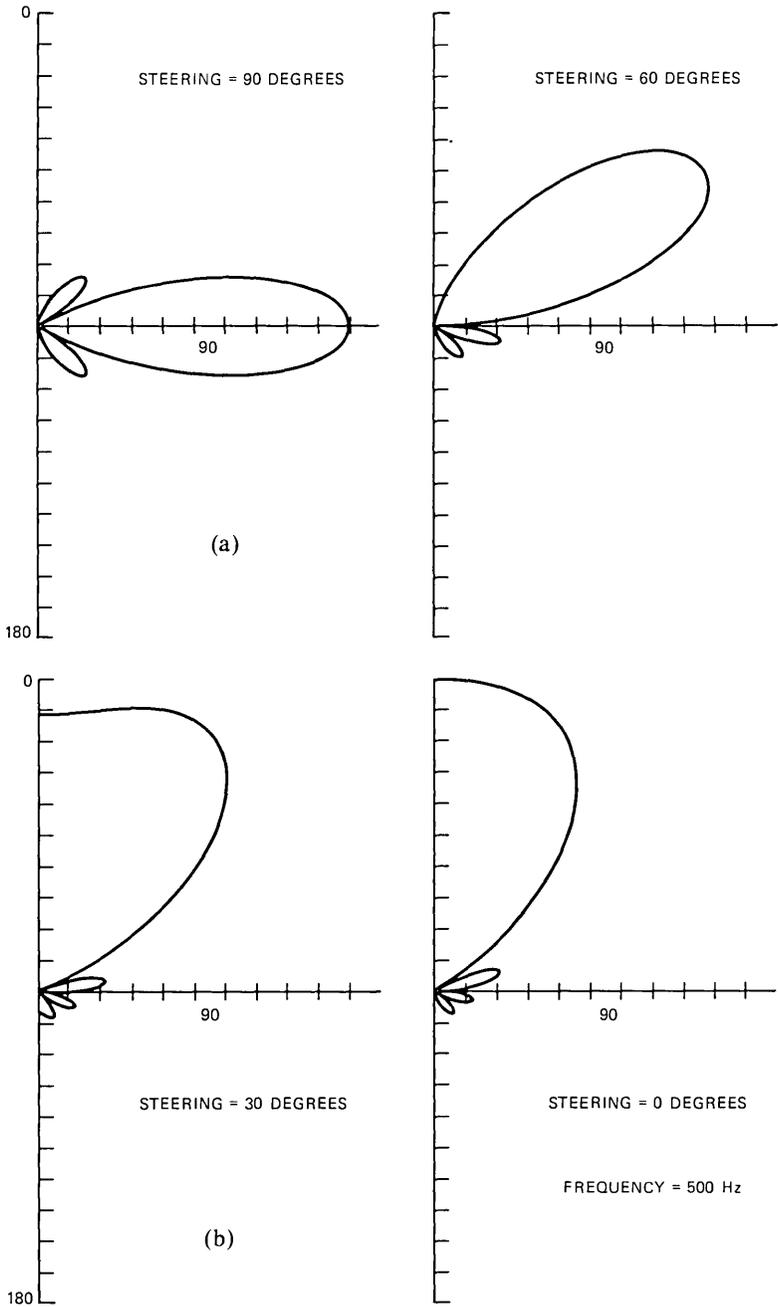


Fig. 4—Spatial response for the uniform array. The parameters are $N = 10$, $d = 6.5$ cm, and $f = 500$ Hz for (a) $\phi' = 90^\circ$, 60° and (b) $\phi' = 30^\circ$, 0° . The polar plots show $|H(\phi)|$ versus ϕ on a linear amplitude scale. For the one-dimensional array, the complete spatial directivity pattern is a figure of revolution about the vertical axis.

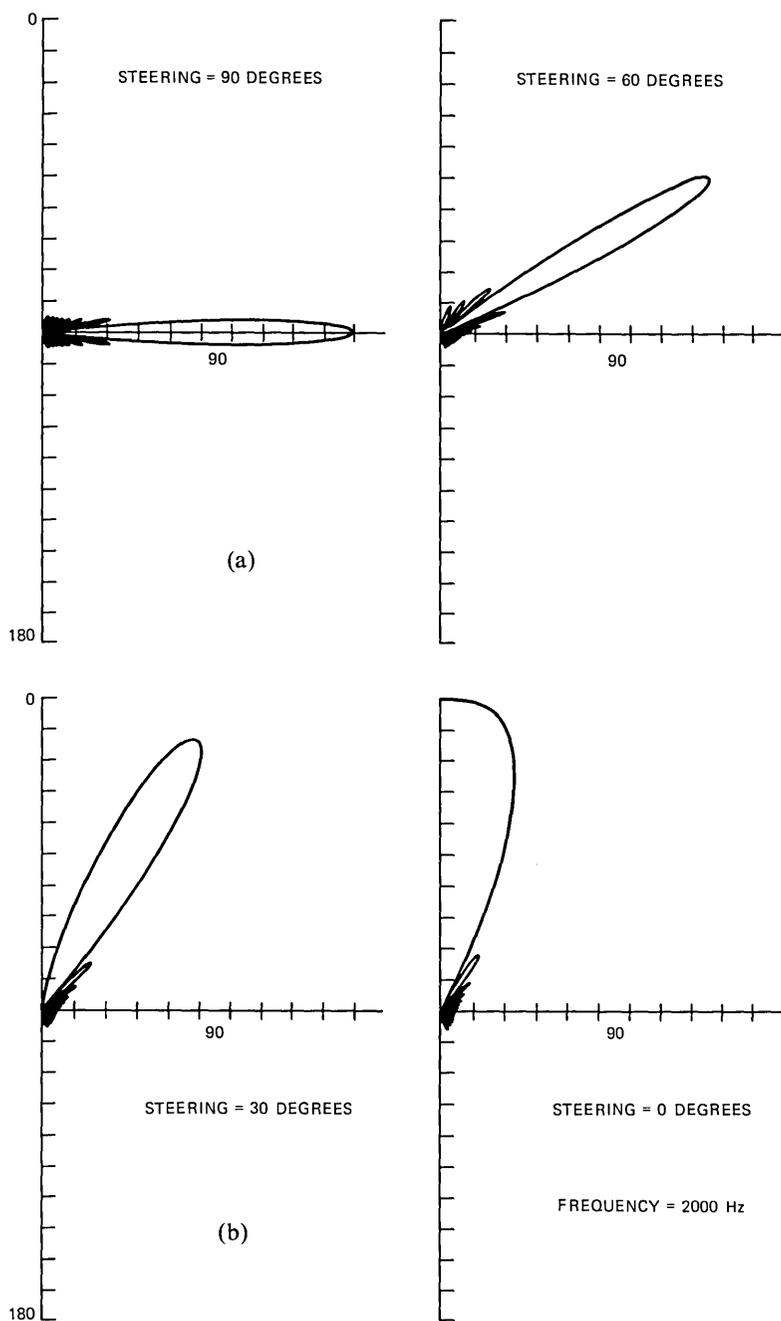


Fig. 5—Spatial response on a linear scale for the uniform array. The parameters are $N = 10$, $d = 6.5$ cm and $f = 2000$ Hz for (a) $\phi' = 90^\circ, 60^\circ$ and (b) $\phi' = 30^\circ, 0^\circ$.

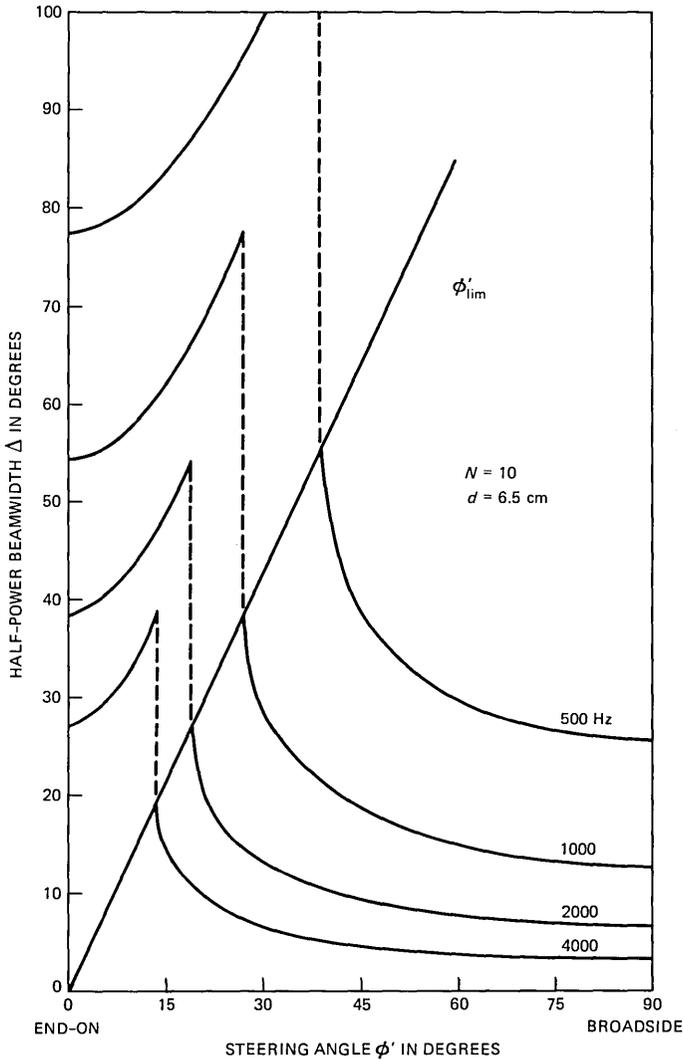


Fig. 6—Variation of half-power beamwidth Δ as a function of steering angle ϕ' . Frequency is the parameter. The limiting steering angle, ϕ'_{lim} , is calculated from eq. (11). For steering angles more acute to the array axis than ϕ'_{lim} , only a single half-power response direction exists in the right hemisphere. The beam pattern merges to an axially directed lobe, as specified in eq. (14), and as indicated by the dashed portions of the curves.

quency the parameter, as shown in Fig. 6. The plot is shown for the first quadrant, $0^\circ \leq \phi' \leq 90^\circ$. The second quadrant, $90^\circ \leq \phi' \leq 180^\circ$, exhibits mirror-image curves, symmetric about $\phi' = 90^\circ$.

Another aspect of interest is the asymmetry of the beam about the

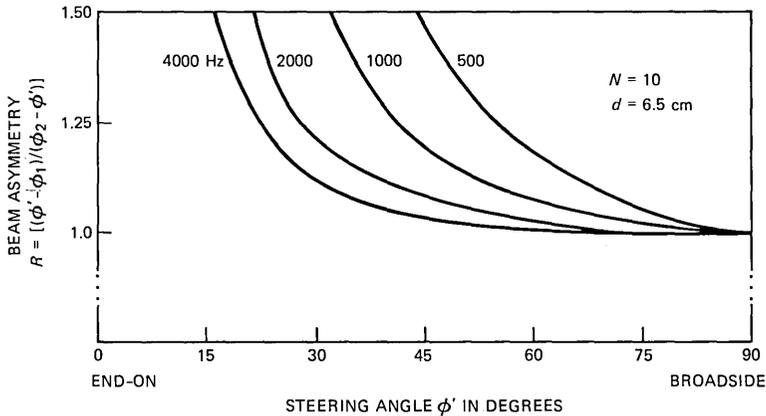


Fig. 7—Variation of beam asymmetry R as a function of steering angle ϕ' , with frequency as the parameter.

steering direction ϕ' . This is reflected by the values $(\phi' - \phi_1)$ and $(\phi_2 - \phi')$, and by their ratios, R . This asymmetry, as a function of the steering angle and with frequency as a parameter, is plotted in Fig. 7.

V. CONCLUDING COMMENT

Automatic, electronic steering of microphone arrays is an attractive means for improving sound pickup in large rooms. This improvement makes more feasible interactive communication between sizable groups of people seated in auditoria or meeting rooms remote from one another. Recently emerging microprocessors have the computational capability for automatic detection of desired sound sources in the room, and for automatic beam forming and beam steering to the desired source.¹ But spatial acuity is affected by steering, typically diminishing as the beam is steered away from the broadside position. Therefore, in addition to the considerations of useful bandwidth (and, hence, array geometry), the design of a beam-steered system must take into account the variations in acuity with steering. The relations derived here provide quantitative means for designing uniform, unweighted arrays for specified conditions of room coverage.

VI. ACKNOWLEDGMENT

I thank Gary Elko of the Acoustics Research Department for a critical reading of this paper, and for a number of constructive comments.

REFERENCES

1. J. L. Flanagan et al., "Computer-Steered Microphone Arrays for Conference Telephony," *J. Acoust. Soc. Am.*, Suppl. 1, 70 (Fall 1981), p. S79.

AUTHOR

James L. Flanagan, B.S. (Electrical Engineering), 1948, Mississippi State University; M.S., 1950, Sc.D., 1955, The Massachusetts Institute of Technology, both in Electrical Engineering; AT&T Bell Laboratories, 1957—. Mr. Flanagan is Director of the Acoustical and Behavioral Research Center. He has project responsibilities in digital speech processing, perception, robotics, and acoustic systems for digital voice encoding, speech recognition and synthesis, and electroacoustic systems. Fellow, IEEE, Acoustical Society of America; member, National Academy of Engineering, National Academy of Sciences.

PAPERS BY AT&T BELL LABORATORIES AUTHORS

COMPUTING/MATHEMATICS

- Bentley J., **Perspective on Performance.** Comm ACM 27(11):1087-1092, Nov 1984.
- Brotman L. S., Badler N. I., **Generating Soft Shadows With a Depth Buffer Algorithm.** IEEE Comp G 4(10):5-12, Oct 1984.
- Calderbank A. R., Coffman E. G., Flatto L., **Optimum Head Separation in a Disk System With Two Read Write Heads.** J ACM 31(4):826-838, Oct 1984.
- Calderbank R., Wales D. B., **The Haemers Partial Geometry and the Steiner System S(5,8,24).** Discr Math 51(2):125-136, Sep 1984.
- Fishburn P. C., **Equity Axioms for Public Risks.** Operat Res 32(4):901-908, Jul-Aug 1984.
- Golumbic M. C., Monma C. L., Trotter W. T., **Tolerance Graphs.** Discr App W 9(2):157-170, Oct 1984.
- Graham R. L., Sloane N. J. A., **Anti-Hadamard Matrices.** Lin Alg App 62(Nov):113-137, Nov 1984.
- Jarvis J. F., **Robotics.** Computer 17(10):283-292, Oct 1984.
- Korner J., Wei V. K., **Odd and Even Hamming Spheres Also Have Minimum Boundary.** Discr Math 51(2):147-165, Sep 1984.
- Lagarias J. C., Mazo J. E., Shepp L. A., **An Inequality for Walks in a Graph (Letter).** Siam Rev 26(4):580-582, Oct 1984.
- Lee T. T., **M/G/I/N Queue With Vacation Time and Exhaustive Service Discipline.** Operat Res 32(4):774-784, Jul-Aug 1984.
- Melamed B., Yadin M., **Numerical Computation of Sojourn-Time Distributions in Queuing Networks.** J ACM 31(4):839-854, Oct 1984.
- Melamed B., Yadin M., **Randomization Procedures in the Computation of Cumulative-Time Distributions Over Discrete State Markov Processes.** Operat Res 32(4):926-944, Jul-Aug 1984.
- Mitra D., Weinberger P. J., **Probabilistic Models of Database Locking—Solutions, Computational Algorithms, and Asymptotics.** J ACM 31(4):855-878, Oct 1984.

ENGINEERING

- Alferness R. C., Eisenstein G., Korotky S. K., Tucker R. S., Buhl L. L., Kaminow I. P., Burrus C. A., Veselka J. J., **Mode-Locking a Ti-LiNbO₃-InGaAsP InP Composite Cavity Laser With an Integrated High-Speed Directional Coupler Switch.** Appl Phys L 45(9):944-946, Nov 1 1984.
- Atwood D. K., Fisanick G. J., Johnson W. A., Wagner A., **Defect Repair Techniques for X-Ray Masks.** P Soc Photo 471:127-134, 1984.
- Barnes Y., Bunin H., **Interference in Frequency-Locked Doppler Tracking Loops.** IEEE Aer El 20(5):574-582, Sep 1984.
- Bowers J. E., Tucker R. S., Burrus C. A., **Six-GHz Direct Frequency Modulation of Cleaved-Coupled-Cavity Channeled-Substrate Buried-Heterostructure Lasers (Letter).** IEEE J Q El 20(11):1230-1232, Nov 1984.
- Chen C. Y., Antreasyan A., Cho A. Y., Garbinski P. A., **GaAs Field-Effect Transistors Prepared on Lattice-Mismatched InP Substrates for Monolithic Optoelectronic Integration.** Electr Lett 20(21):865-866, Oct 11 1984.
- Dingle R., **New High-Speed III-V Devices for Integrated Circuits.** IEEE Device 31(11):1662-1667, Nov 1984.
- Dutta N. K., Napholtz S. G., Wilson R. B., Brown R. L., Cella T., Craft D. C., **High-Power Gain-Guided InGaAsP Laser Array.** Appl Phys L 45(9):941-943, Nov 1 1984.
- Dutta N. K., Nelson R. J., Wilson R. B., Maher D. M., Wright P. D., Sheng T. T., Lin P. S. D., Marcus R. B., **Effect of Active Layer Placement on the Threshold**

- Current of 1.3- μm InGaAsP Etched Mesa Buried Heterostructure Lasers.** Appl Phys L 45(4):337-339, Aug 15 1984.
- Edelson D., Flamm D. L., **Computer Simulation of a CF_4 Plasma Etching Silicon.** J Appl Phys 56(5):1522-1531, Sep 1 1984.
- Franey J. P., Graedel T. E., Gualtieri G. J., Kammlott G. W., Malm D. L., Sharpe L. H., Tierney V., **Conductive Silver Epoxy Pastes—Characteristics of Alternative Formulations.** J Mater Sci 19(10):3281-3286, Oct 1984.
- Georgiou G. E., Jankoski C. A., Palumbo T. A., **DC Electroplating of Sub-Micron Gold Patterns on X-Ray Masks.** P Soc Photo 471:96-102, 1984.
- Glass A. M., **Materials for Optical Information Processing.** Science 226(4675): 657-662, Nov 9 1984.
- Gossard A. C., **New Electronic Phenomena Based on Multilayer Epitaxy.** IEEE Device 31(11):1667-1672, Nov 1984.
- Hayes J. R., Levi A. F. J., Wiegmann W., **Hot-Electron Spectroscopy.** Electr Lett 20(21):851-852, Oct 11 1984.
- Jelinski L. W., **Modern NMR Spectroscopy.** Chem Eng N 62(45):26+, Nov 5 1984.
- Koch T. L., Coldren L. A., Bridges T. J., Burkhardt E. G., Corvini P. J., Miller B. I., Wilt D. P., **Low-Threshold High-Speed 1.55- μm Vapor-Phase Transported Buried Heterostructure Lasers (VPTBH).** Electr Lett 20(21):856-857, Oct 11 1984.
- Korotky S. K., Eisenstein G., Kasper B. L., Alferness R. C., Veselka J. J., Buhl L. L., **Error-Free External Modulation of a Single-Frequency Injection Laser at 1.5 GBit/s Using a Ti-LiNbO₃ Wave-Guide Switch.** Electr Lett 20(21):878-879, Oct 11 1984.
- Lines M. E., **The Search for Very Low-Loss Fiber-Optic Materials.** Science 226(4675):663-668, Nov 9 1984.
- Meloni A., Gregori G. P., Lanzerotti L. J., Medford L. V., **Search for a Possible Electromagnetic Coupling Between a Transatlantic Communication Cable and the Magma Chamber in the Mid-Atlantic Ridge.** J Geophys 55(3):185-190, 1984.
- Olsson N. A., Tsang W. T., Logan R. A., **Active Spectral Stabilization of Cleaved-Coupled-Cavity (C³) Lasers.** J Light T 2(1):49-51, Feb 1984.
- Raack G. A., Sable E. G., Stewart R. J., **Customer Control of Network Services.** IEEE Comm M 22(10): 8-14, Oct 1984.
- Silberberg Y., Smith P. W., Eilenberger D. J., Miller D. A. B., Gossard A. C., Wiegmann W., **Passive-Mode Locking of a Semiconductor Diode Laser.** Optics Lett 9(11):507-509, Nov 1984.
- Tsang W. T., **The C-3 Laser.** Sci Am 251(5):149+, Nov 1984.
- Vanderziel J. P., Mikulyak R. M., Temkin H., Logan R. A., Dupuis R. D., **Optical Beam Characteristics of Schottky-Barrier Confined Arrays of Phase-Coupled Multiquantum Well GaAs Lasers.** IEEE J Q El 20(11):1259-1266, Nov 1984.
- Vanderziel J. P., Temkin H., Logan R. A., Dupuis R. D., **High-Power Picosecond Pulse Generation in GaAs Multiquantum Well Phase-Locked Laser Arrays Using Pulsed Current Injection.** IEEE J Q El 20(11):1236-1242, Nov 1984.

MANAGEMENT/ECONOMICS

- Fishburn P. C., **Dominance in SSB Utility Theory.** J Econ Theo 34(1):130-148, Oct 1984.
- Fishburn P. C., **Probabilistic Social Choice Based on Simple Voting Comparisons.** Rev Econ S 51(4):683-692, Oct 1984.

PHYSICAL SCIENCES

- Abrahams S. C., Ihringer J., Marsh P., Nassau K., **Phase Transition at 434 K, Independent Strain Coupling in Second Transition at 400 K, and Thermal Expansivity in Ferroelastic K₂TeBr₆.** J Chem Phys 81(4):2082-2087, Aug 15 1984.

- Argade P. V. et al., **Confirmation of the Assignment of the Iron Histidine Stretching Mode in Myoglobin.** *J Am Chem S* 106(22):6593-6596, Oct 31 1984.
- Argade P. V., Ching Y. C., Rousseau D. L., **Cytochrome a_3 Structure in Carbon Monoxide-Bound Cytochrome Oxidase.** *Science* 225(4659):329-331, Jul 20 1984.
- Bates F. S., Bair H. E., Hartney M. A., **Block Copolymers Near the Microphase Separation Transition. 1. Preparation and Physical Characterization of a Model System.** *Macromolec* 17(10):1987-1993, Oct 1984.
- Bosch M. A., Haller H. R., Schiavone L. M., Shay J. L., **Reactively Sputtered TeN_x Optical-Recording Media.** *Appl Phys L* 45(3):202-204, Aug 1 1984.
- Brummell M. A., Nicholas R. J., Brunel L. C., Huant S., Baj M., Portal J. C., Razeghi M., Di Forte-Poisson M. A., Cheng K. Y., Cho A. Y., **Cyclotron Resonance and Polaron Effects in a Two-Dimensional Electron-Gas in GaInAs .** *Surf Sci* 142(1-3):380-387, Jul 1984.
- Capasso F., Mohammed K., Alavi K., Cho A. Y., **Impact Ionization Rates for Electrons and Holes in $\text{Al}_{0.48}\text{In}_{0.52}\text{As}$.** *Appl Phys L* 45(9):968-970, Nov 1 1984.
- Chabal Y. J., Patel C. K. N., **Solid Hydrogen in Amorphous Silicon—Phase Transition.** *Phys Rev L* 53(18):1771-1774, Oct 29 1984.
- Childs T. T., Royer W. A., Smith N. V., **Evaluation of LIF Lenses as Monochromators for Inverse Photoemission Spectroscopy.** *Rev Sci Ins* 55(10):1613-1615, Oct 1984.
- Davis G. P., Moore C. A., Gottscho R. A., **Nascent Product Time of Flight Distributions From Laser Initiated Etching of Germanium by Bromine.** *P Soc Photo* 459:115-120, 1984.
- Dimauro L. F., Heaven M., Miller T. A., **Laser-Induced Fluorescence Study of the $B^2A'' \rightarrow X^2A''$ Transition of the Vinyloxy Radical in a Supersonic Free Jet Expansion.** *J Chem Phys* 81(5):2339-2346, Sep 1 1984.
- Duncan T. M., Douglass D. C., **On the ^{31}P Chemical Shift Anisotropy in Condensed Phosphates.** *Chem Phys* 87(3):339-349, Jul 1 1984.
- Dupuis R. D., **Metalorganic Chemical Vapor-Deposition of III-V Semiconductors.** *Science* 226(4675):623-629, Nov 9 1984.
- Dutt B. V., Chin A. K., Camlibel I., Bonner W. A., **Influence of Ambient Media on the Out-Diffusion of S From InP:S .** *J Appl Phys* 56(6):1630-1635, Sep 15 1984.
- Ekin J. W., Hong M., **Electromechanical and Metallurgical Properties of Liquid-Infiltration Nb-Ta/Sn Multifilamentary Superconductor.** *Appl Phys L* 45(3):297-299, Aug 1 1984.
- Evans T. E., Valanju P. M., Benesch J. F., Bengston R. D., Li Y. M., Mahajan S. M., Oakes M. E., Ross D. W., Wang X. Z., Watkins J. G., **Direct Observation of the Structure of Global Alfvén Eigenmodes in a Tokamak Plasma.** *Phys Rev L* 53(18):1743-1746, Oct 29 1984.
- Feldman L. C., **Ion-Scattering Analysis of Interfaces.** *Ultramicros* 14(1-2):51-54, 1984.
- Gerin M., Combes F., Encrenaz P., Linke R., Destombes J. L., Demuyneck C., **Detection of ^{13}CN in Three Galactic Sources (Letter).** *Astron-Astr* 136(2):L 17-L 20, Jul 1984.
- Gibson J. M., **High-Resolution Electron Microscopy of Interfaces Between Epitaxial Thin Films and Semiconductors.** *Ultramicros* 14(1-2):1-10, 1984.
- Gibson J. M., **Uniqueness of Surface Images (Letter).** *Phys Rev L* 53(19):1859, Nov 5 1984.
- Goodby J. W., **Liquid Crystal Phases Exhibited by Some Monosaccharides.** *Molec Cryst* 110(1-4):205-219, 1984.
- Goodby J. W., Leslie T. M., **Ferroelectric Liquid Crystals—Structure and Design.** *Molec Cryst* 110(1-4):175-203, 1984.
- Helfand E., **Dynamics of Conformational Transitions in Polymers.** *Science* 226(4675):647-650, Nov 9 1984.
- Henkel C., Wilson T. L., Johnston K. J., **Ammonia Absorption Toward NGC 7538 IRS 1:2 ARC Second Observations in the (3,3) Line.** *Astrophys J* 282(2):L93-L96, Jul 15 1984.
- Jackel J. L., Rice C. E., **Short-Term and Long-Term Stability in Proton Exchanged Lithium-Niobate Wave-Guides.** *P Soc Photo* 460:43-48, 1984.
- Katz H. E., **Macrocyclic Bipyridines for Alkyltin Trichloride Deactivation.** *Tetrahedr L* 25(43):4905-4908, 1984.

- Kevan S. D., Dubois L. H., **Development of Dispersion Compensation for Use in High-Resolution Electron-Energy-Loss Spectroscopy.** *Rev Sci Ins* 55(10):1604-1612, Oct 1984.
- Kotliar G., Sompolinsky H., **Phase Transition in a Dzyaloshinsky-Moriya Spin Glass.** *Phys Rev L* 53(18):1751-1754, Oct 29 1984.
- Lambert W. R., Cardillo M. J., Trevor P. L., Doak R. B., **Inelastic Helium Scattering From the Ag(001) and Ag(001)C(2×2)Cl Surfaces.** *Surf Sci* 145(2-3): 519-543, Oct 1984.
- Lanzerotti L. J., Brown W. L., Marcantonio K. J., Johnson R. E., **Production of Ammonia-Depleted Surface Layers on the Saturnian Satellites by Ion Sputtering.** *Nature* 312(5990):139-140, Nov 8 1984.
- Leung S. Y., Schumaker N. E., **Layer Thickness Uniformity in Liquid-Phase Epitaxy—A Two-Dimensional Diffusion Analysis.** *J Cryst Gr* 67(3):458-466, Aug 1984.
- Lovinger A. J., Cais R. E., **Structure and Morphology of Poly(Trifluoroethylene).** *Macromolec* 17(10):1939-1945, Oct 1984.
- Macrander A. T., Seidman D. N., **An Atom-Probe Field-Ion Microscope Study of 200-eV⁺H₂⁺ Ions Implanted in Tungsten at 29 K.** *J Appl Phys* 56(6):1623-1629, Sep 15 1984.
- Marcus M. A., **Interaction of Micellar Additions With Thermotropic Liquid Crystals.** *Molec Cryst* 110(1-4):145-152, 1984.
- McRae E. G., Gossmann H. J., Feldman L. C., **Structural Analogy Between GeSi (111)-5×5 and Si (111)-7×7 Surfaces (Letter).** *Surf Sci* 146(1):L540-L546, Oct 1984.
- Milshtein S. K., Joy D. C., Ferris S. D., Kimerling L. C., **Defect Characterization Using SEM-CCM—Relative Contrast Measurements.** *Phys St S-A* 84(2):363-369, Aug 16 1984.
- Narayanamurti V., **Crystalline Semiconductor Heterostructures.** *Phys Today* 37(10):24-32, Oct 1984.
- Nassau K., **Early History of Lithium-Niobate—Personal Reminiscences.** *P Soc Photo* 460:2-5, 1984.
- Nassau K., Lines M. E., **Calculation of Scattering and Other Parameters for Ultra Low-Loss Optical Fibers.** *P Soc Photo* 484:7-13, 1984.
- Rosamilia J. M., Miller B., **Hydrodynamic Modulation Applications to Electroanalysis and Diffusion-Coefficient Measurement With a Commercial Disk Rotator.** *Analyt Chem* 56(13):2410-2413, Nov 1984.
- Ryan J. F. et al., **Time-Resolved Photoluminescence of Two-Dimensional Hot Carriers in GaAs-AlGaAs Heterostructures.** *Phys Rev L* 53(19):1841-1844, Nov 5 1984.
- Schlesinger Z., Hwang J. C. M., Allen S. J., Le H., **Infrared Studies of the GaAs/(GaAl)As Two-Dimensional Electron Gas.** *Surf Sci* 142(1-3):423-426, Jul 1984.
- Schluter M., **Bonding Geometries at Semiconductor Surfaces and Interfaces.** *Ultramicros* 14(1-2):89-96, 1984.
- Schneemeyer L. F., Disalvo F. J., Fleming R. M., Waszczak J. V., **Sliding Charge-Density Wave Conductivity in Potassium Molybdenum Bronze.** *J Sol St Ch* 54(3):358-364, Oct 1984.
- Schwartz G. P., Thiel F. A., Gualtieri G. J., **Thermal Oxidation of InAs_xP_{1-x}.** *J Vac Sci A* 2(3):1252-1256, Jul-Sep 1984.
- Selke W., Huse D. A., Kroll D. M., **Interfacial Adsorption in the Two-Dimensional Blume-Capel Model.** *J Phys A* 17(15):3019-3027, Oct 21 1984.
- Stone J., Burrus C. A., Wiesenfeld J. M., **On the Origin of the Long-Wavelength Loss Anomaly in H₂-Treated Germano-Phospho-Silicate Glass Fibers.** *Appl Phys L* 45(3):212-214, Aug 1 1984.
- Stucki F., Anderson J., Lapeyre G. J., Farrell H. H., **Multiple Vibrational Excitations of H₂O and D₂O on Si (100) (2×1): A Hreels Study.** *Surf Sci* 143(1):84-92, Jul 1984.
- Temkin H., Panish M. B., Logan R. A., van der Ziel J. P., **λ ≈ 1.5 μm InGaAsP Ridge Lasers Grown by Gas Source Molecular Beam Epitaxy.** *Appl Phys L* 45(4):330-332, Aug 15 1984.
- Tonelli A. E., Schillinger F. C., **C-13 NMR Chemical Shifts and the Microstructure**

of Propylene Vinyl-Chloride Copolymers With Low Propylene Content. Macromolec 17(10):1946-1949, Oct 1984.

Vandover R. B., Bacon D. D., Sinclair W. R., **Preparation of NBN by Reactive Magnetron Sputtering Using NE and KR Inert Diluents.** J Vac Sci A 2(3):1257-1260, Jul-Sep 1984.

SOCIAL AND LIFE SCIENCES

Bosch M. A., **Optical Recording.** Sem Semimet 21(PD):173-207, 1984.

Bradley D. R., Dumais S. T., **The Effects of Illumination Level and Retinal Size on the Depth Stratification of Subjective Contour Figures.** Perception 13(2): 155-164, 1984

Brams S. J., Fishburn P. C. **Proportional Representation in Variable-Size Legislatures.** Soc Choice 1(3):211-229, 1984.

Fishburn P. C., Gehrlein W. V., **Powers of Subgroups in Voting Bodies.** Soc Choice 1(2):85-95, 1984.

Starr S. J., **Effects of Video Display Terminals in a Business Office.** Human Fact 26(3):347-356, Jun 1984.

SPEECH/ACOUSTICS

Andrews H. L., **Speech Processing.** Computer 17(10):315-324, Oct 1984.

Lin D. W., **On Digital Implementation of the Fast Kalman Algorithms.** IEEE Acoust 32(5):998-1005, Oct 1984.

CONTENTS, MAY-JUNE 1985

Four-Dimensional Modulation With an Eight-State Trellis Code

A. R. Calderbank and N. J. A. Sloane

100-GHz Measurements of Two Astigmatic Launchers

R. A. Semplak

On the Use of Vector Quantization for Connected-Digit Recognition

S. C. Glinski

Incorporation of Temporal Structure Into a Vector-Quantization-Based Preprocessor for Speaker-Independent, Isolated-Word Recognition

A. F. Bergh, F. K. Soong, and L. R. Rabiner

COMPUTING SCIENCE AND SYSTEMS

A New Light Pen With Subpixel Accuracy

M. Hatamian and E. F. Brown

Sojourn Time Distribution in a Multiprogrammed Computer System

K. M. Rege and B. Sengupta

Application of Decomposition Principle in M/G/1 Vacation Model to Two Continuum Cyclic Queueing Models—Especially Token-Ring LANs

S. W. Fuhrmann and R. B. Cooper

PLACE 2.0—An Interactive Program for PLL Analysis and Design

J. H. Saunders

AT&T TECHNICAL JOURNAL is abstracted or indexed by *Abstract Journal in Earthquake Engineering, Applied Mechanics Review, Applied Science & Technology Index, Chemical Abstracts, Computer Abstracts, Current Contents/Engineering, Technology & Applied Sciences, Current Index to Statistics, Current Papers in Electrical & Electronic Engineering, Current Papers on Computers & Control, Electronics & Communications Abstracts Journal, The Engineering Index, International Aerospace Abstracts, Journal of Current Laser Abstracts, Language and Language Behavior Abstracts, Mathematical Reviews, Science Abstracts (Series A, Physics Abstracts; Series B, Electrical and Electronic Abstracts; and Series C, Computer & Control Abstracts), Science Citation Index, Sociological Abstracts, Social Welfare, Social Planning and Social Development, and Solid State Abstracts Journal*. Reproductions of the Journal by years are available in microform from University Microfilms, 300 N. Zeeb Road, Ann Arbor, Michigan 48106.

