

THE APRIL 1981 VOL. 60, NO. 4

BELL SYSTEM

TECHNICAL JOURNAL



G. E. Peterson, A. Carnevale, U. C. Paek, and J. W. Fleming	Numerical Calculation of Optimum α for a Germanium-Doped Silica Lightguide	455
U. Timor	Multistage Decoding of Frequency- Hopped FSK System	471
W. D. Wynn	A Bubble Memory Differential Detector	485
N. S. Jayant	Subsampling of a DPCM Speech Channel to Provide Two "Self- Contained" Half-Rate Channels	501
V. E. Beneš	Blocking States in Connecting Networks Made of Square Switches Arranged in Stages	511
W. Pferd and G. C. Stocker	Optical Fibers for Scanning Digitizers	523
S. H. Francis, H. R. Lunde, T. E. Talpey, and G. A. Reinold	Magnetic Localization of Buried Cable by the SCARAB Submersible	535
	CONTRIBUTORS TO THIS ISSUE	575
	PAPERS BY BELL LABORATORIES AUTHORS	579
	CONTENTS, MAY-JUNE ISSUE	582

THE BELL SYSTEM TECHNICAL JOURNAL

ADVISORY BOARD

D. E. PROCKNOW, *President*, *Western Electric Company*
I. M. ROSS, *President*, *Bell Telephone Laboratories, Incorporated*
W. M. ELLINGHAUS, *President*, *American Telephone and Telegraph Company*

EDITORIAL COMMITTEE

A. A. PENZIAS, *Chairman*

A. G. CHYNOWETH	W. B. SMITH
R. P. CLAGETT	G. SPIRO
T. H. CROWLEY	J. W. TIMKO
I. DORROS	I. WELBER
R. A. KELLEY	M. P. WILSON

EDITORIAL STAFF

B. G. KING, *Editor*
PIERCE WHEELER, *Associate Editor*
JEAN G. CHEE, *Assistant Editor*
H. M. PURVIANCE, *Art Editor*
B. G. GRUBER, *Circulation*

THE BELL SYSTEM TECHNICAL JOURNAL is published monthly, except for the May-June and July-August combined issues, by the American Telephone and Telegraph Company, C. L. Brown, Chairman and Chief Executive Officer; W. M. Ellinghaus, President; V. A. Dwyer, Vice President and Treasurer; F. A. Hutson, Jr., Secretary. Editorial inquiries should be addressed to the Editor, The Bell System Technical Journal, Bell Laboratories, Room WB 1L-331, Crawfords Corner Road, Holmdel, N.J. 07733. Checks for subscriptions should be made payable to The Bell System Technical Journal and should be addressed to Bell Laboratories, Circulation Group, Whippany Road, Whippany, N.J. 07981. Subscriptions \$20.00 per year; single copies \$2.00 each. Foreign postage \$1.00 per year; 15 cents per copy. Printed in U.S.A. Second-class postage paid at New Providence, New Jersey 07974 and additional mailing offices.

© 1981 American Telephone and Telegraph Company. ISSN0005-8580

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

Comments on the technical content of any article or brief are welcome. These and other editorial inquiries should be addressed to the Editor, The Bell System Technical Journal, Bell Laboratories, Room WB 1L-331, Crawfords Corner Road, Holmdel, N.J. 07733. Comments and inquiries, whether or not published, shall not be regarded as confidential or otherwise restricted in use and will become the property of the American Telephone and Telegraph Company. Comments selected for publication may be edited for brevity, subject to author approval.

THE BELL SYSTEM TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING
ASPECTS OF ELECTRICAL COMMUNICATION

Volume 60

April 1981

Number 4

Copyright © 1981 American Telephone and Telegraph Company. Printed in U.S.A.

Numerical Calculation of Optimum α for a Germania-Doped Silica Lightguide

By G. E. PETERSON, A. CARNEVALE, U. C. PAEK, and J. W. FLEMING

(Manuscript received September 11, 1980)

Using an exact numerical solution to Maxwell's equations, we determine the optimum refractive index profile parameter α for a germania-doped silica lightguide. Our results agree closely with the earlier work of Olshansky and Keck and with the work of Marcatili. However, differences exist that may be important in the manufacture of very high bandwidth lightguides.

I. INTRODUCTION

It is now well established that a lightguide with a near parabolic index profile can have single frequency bandwidths of 1 GHz·km or greater.¹⁻³ This is expected to happen if the profile is smooth, the central dip is narrow or small, the profile parameter is properly chosen, and modes that have a significant amount of their energy transported in the cladding are eliminated.⁴⁻⁶ Our understanding of these factors has largely come from approximate solutions to Maxwell's equations for lightguides. The availability of computer software for exact numerical solutions⁷ to Maxwell's equations allows us to reexamine some of these effects and to look at other factors that may influence lightguide bandwidths. In this paper, we determine optimum power law profile parameters (i.e., α parameters) as a function of wavelength for a germania-doped silica lightguide.

II. PROCEDURE

We study a lightguide with a 13.5 mole percent germania-doped silica core and a pure silica cladding. The core is assumed to be

perfectly cylindrical with a diameter of 50 microns. The cladding is considered to be essentially infinite. The doping profile is such that the refractive index is a power law function of distance R . Thus,

$$N = N_1 + (N_2 - N_1) \cdot \left(\frac{R}{R_{cc}} \right)^\alpha, \quad (1)$$

where N_1 is the refractive index at the center of the core, N_2 is the refractive index in the cladding, R_{cc} is the radius of the core, and α is the profile parameter. We assume that this functional form holds at all wavelengths of interest.

We have accurately determined the wavelength variation of the refractive index of silica and some doped silicas. Figure 1 gives the data for 13.5 mole percent germania-doped silica and pure silica. As is well known, the refractive index falls off slowly with wavelength for both of these glasses.⁸ We express the data by the following modified Sellmeier formula:⁹

$$N_i = C_0 + C_1\lambda^2 + C_2\lambda^4 + C_3/(\lambda^2 - 0.035) + C_4/(\lambda^2 - 0.035)^2 + C_5/(\lambda^2 - 0.035)^3, \quad (2)$$

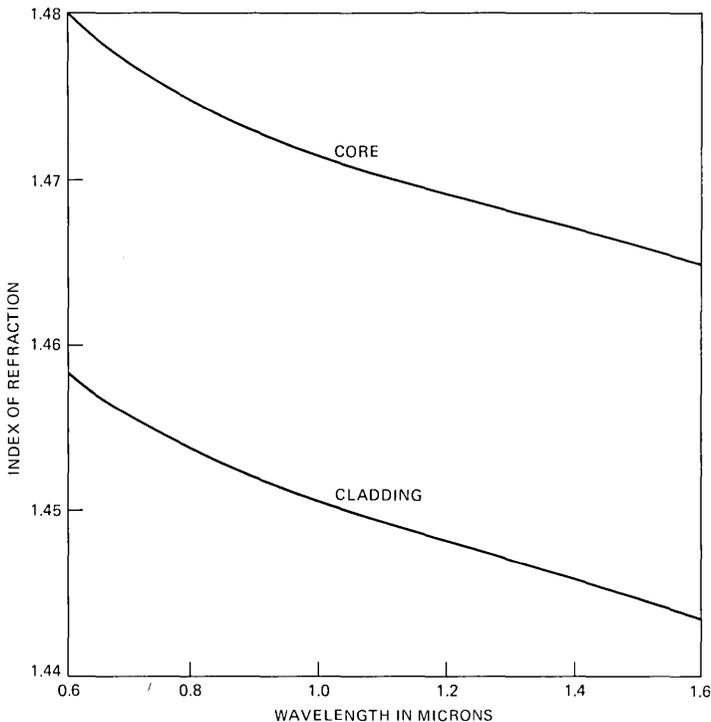


Fig. 1—The wavelength dependence of 13.5 mole percent germania-doped silica and pure silica.

where λ is in microns and $i = 1, 2$. Table I lists the coefficients for eq. (2) for both 13.5 mole percent germania-doped silica and silica.

Our numerical procedure to solve Maxwell's equations has been described in detail in an earlier paper.⁷ Thus, we only give a very brief description here. The method is similar to that described by Vigants and Schlesinger,¹⁰ and Vassel.¹¹ We write Maxwell's equations in cylindrical coordinates as

$$\frac{d\Gamma}{d\rho} = \rho^{-1}A(\rho)\Gamma(\rho), \quad (3)$$

where $\Gamma(\rho)$ is a column vector whose four elements are related to the tangential components of the electromagnetic field. They are as follows:

$$\begin{aligned} \Gamma_1 &= E_z, \\ \Gamma_2 &= -i\rho H_\phi Z_0, \\ \Gamma_3 &= \rho E_\phi, \\ \Gamma_4 &= -iH_z Z_0, \end{aligned} \quad (4)$$

where Z_0 is the wave impedance of free space and the variable ρ is defined as K_0R .

The 4×4 matrix $A(\rho)$ can be written as

$$\begin{vmatrix} 0 & (N_e^2/\kappa) - 1 & 0 & -MN_e/\kappa \\ \rho^2\kappa - M^2 & 0 & MN_e & 0 \\ 0 & MN_e/\kappa & 0 & \rho^2 - (M^2/\kappa) \\ -MN_e & 0 & N_e^2 - \kappa & 0 \end{vmatrix} \quad (5)$$

This matrix describes the properties of the media.

We solve eq. (3) by an optimized fourth-order Runge-Kutta procedure.¹² This is a higher-accuracy formula than used in our earlier work. The most important computational results are the effective indices N_e for the various modes in the lightguide.

The group indices N_g are calculated from the effective indices by the following formula:

$$N_g = N_e - \lambda dN_e/d\lambda. \quad (6)$$

Table I—Coefficients for eq. (2)

13.5 Mole Percent Germania-Doped Core	Silica Cladding
$C_0 = 1.4706868$	1.4508554
$C_1 = -0.0026870$	-0.0031268
$C_2 = -0.0000356$	-0.0000381
$C_3 = 0.0035756$	0.0030270
$C_4 = -0.0000828$	-0.0000779
$C_5 = 0.0000018$	0.0000018

The derivative is obtained from numerical calculations of N_e at three closely spaced wavelengths.

For a lightguide with the materials and profile described earlier, a simple approximate relationship exists between group index N_g and effective index N_e for certain modes. In particular, for a given profile parameter α , angular mode number M , wavelength λ , and for modes far from cutoff, the effective indices are nearly linearly related to the group indices. Evidence for this can be obtained from Fig. 2, curve (a). In this case, we have $\lambda = 0.6328$ micron, $M = 0$, and $\alpha = 2$. The solid line is the linear least-squares fit to the data when the last two modes

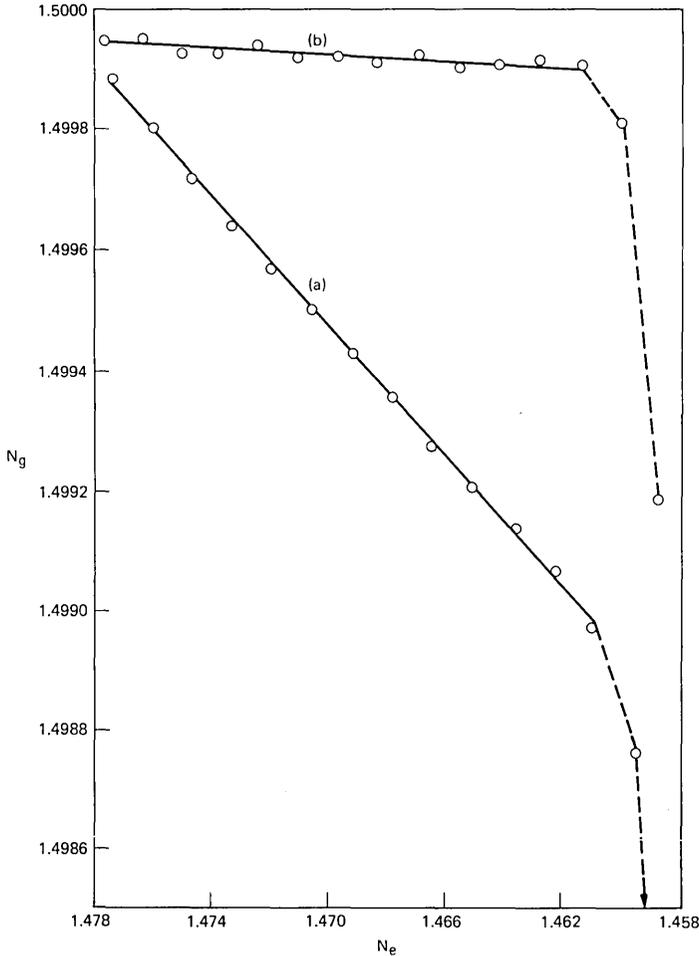


Fig. 2—Plot of group index (N_g) versus effective index (N_e) at a wavelength (λ) of 0.6328 micron with $M = 0$. In (a) the profile parameter α is 2.0 while in (b) α is 2.2. Note that a good linear relationship exists between N_g and N_e if the last few modes are neglected.

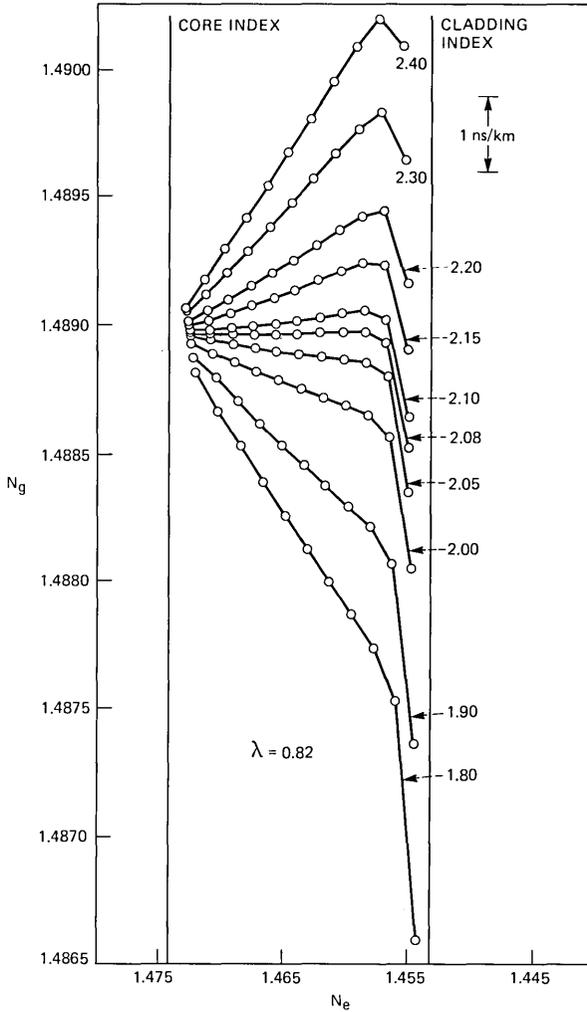


Fig. 3—Plot of group index versus effective index for $\lambda = 0.82$ micron and $M = 0$. When $\alpha = 2.08$ the group indices are all nearly the same except for those modes being influenced by the cladding.

are deleted. (The last mode is too far off scale to be plotted.) Obviously a nearly linear relationship exists between N_g and N_e . In addition, there is a very slight systematic deviation from the least-squares line. This suggests that additional terms are needed to completely describe the data. However, this will be of no consequence to the analysis to be described here.

The last two modes (those that deviate strongly from the line) have a significant fraction of their energy being transported in the cladding.

Thus, they are not tightly bound to the core and normally are not considered in an optimum α calculation.

An optimum profile parameter α , for a particular family of modes, is one that minimizes the spread in group indices for that family at a particular wavelength. For example, we might consider all the modes associated with a specified M value. In this case, the optimum α would be the one that gives the smallest slope to the least-squares line relating N_e and N_g .

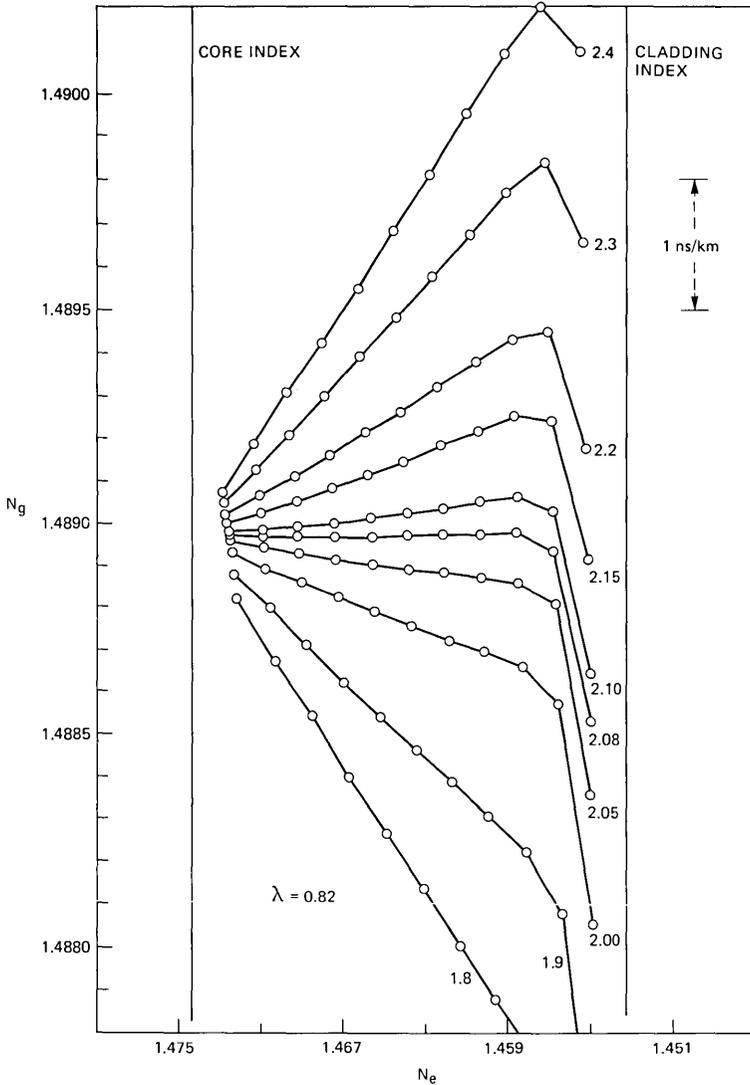


Fig. 4—Part of the data of Fig. 3 on an expanded scale.

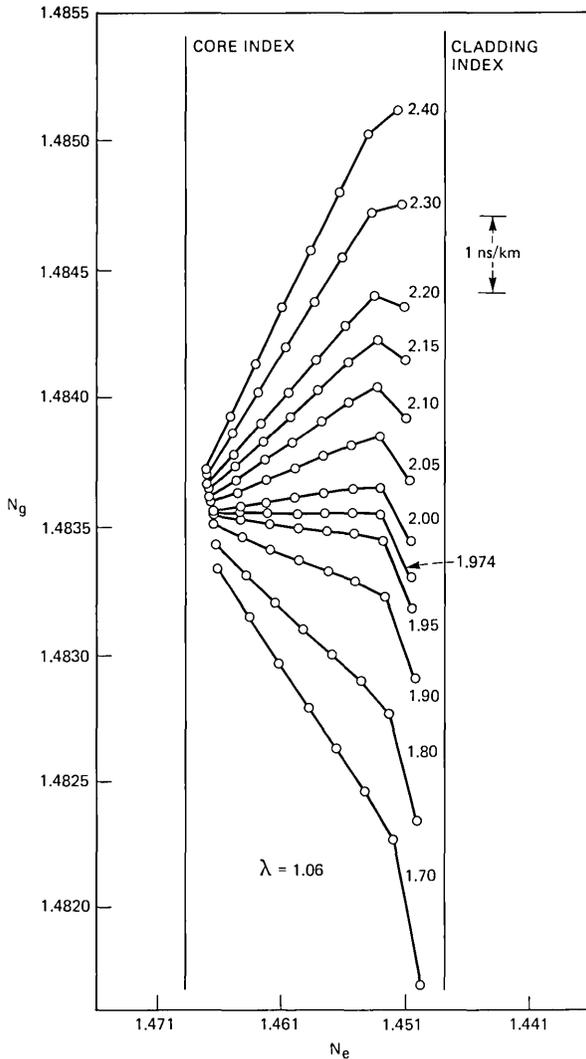


Fig. 5—Plot of group index versus effective index for $\lambda = 1.06$ microns and $M = 0$.

Figure 2, curve (b) shows a near optimum value of $\alpha = 2.2$ for $M = 0$ and $\lambda = 0.6328$ micron. We see that the least-squares line is now nearly horizontal and that evidence exists for the necessity of higher-order terms to exactly fit the data. The strong drop-off of the last few modes is still evident.

A very good overall optimum α for all M values will be nearly identical to the optimum α for $M = 0$. As will be shown shortly, this results in excellent equalization of the lower-order modes and a good

equalization of the higher-order modes. For most engineering applications, optimum α 's in the wavelength range of 0.63 to 1.55 microns are needed.

III. RESULTS

Figures 3 through 7 show plots of N_g versus N_e for various α parameters and wavelengths for $M = 0$. Figure 4 shows part of the

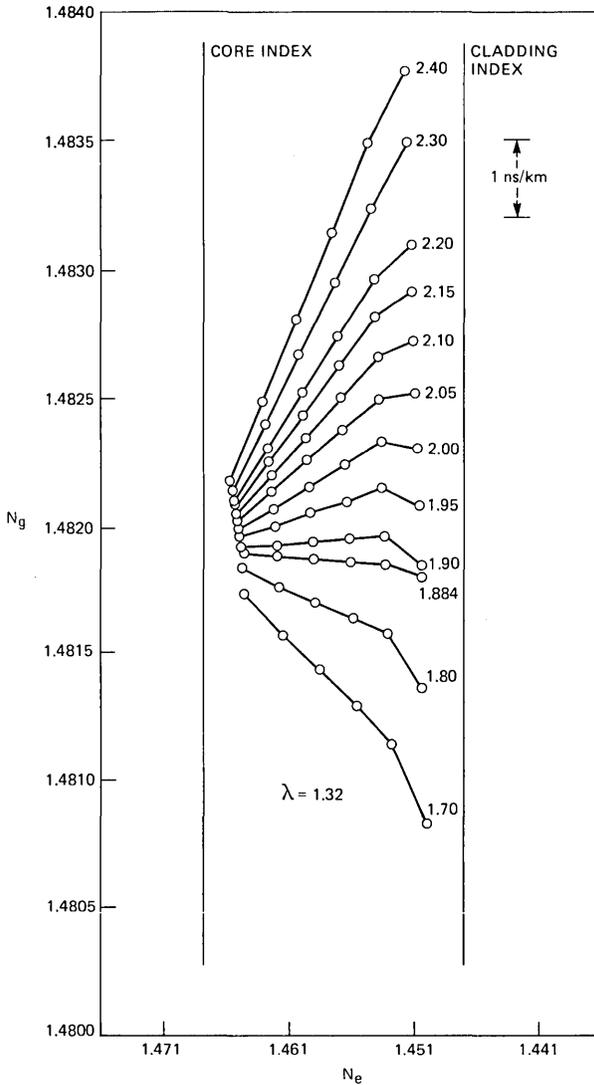


Fig. 6—Plot of group index versus effective index for $\lambda = 1.32$ microns and $M = 0$.

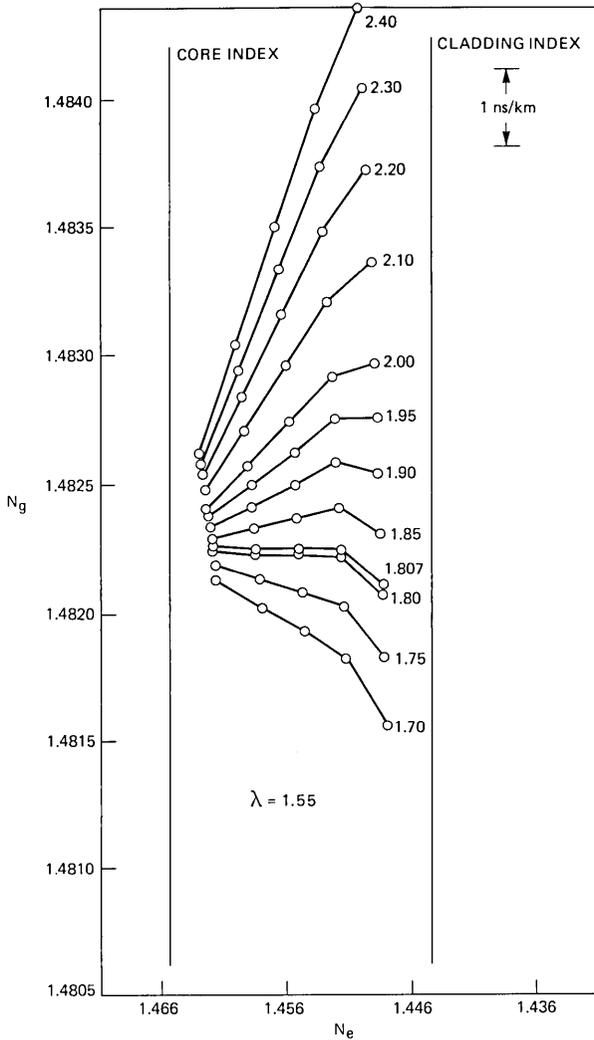


Fig. 7—Plot of group index versus effective index for $\lambda = 1.55$ microns and $M = 0$.

data of Fig. 3 on an expanded scale. Note that in all a nearly linear relationship exists between N_g and N_e , except for the last couple of modes. As mentioned before, these modes are nearing cutoff and are usually not considered in an optimum α determination.

At each wavelength, we see that there is an α value, α_{opt} , that makes the slope of the line nearly zero. Table II lists α_{opt} values for wavelengths of engineering importance and also for a few other wavelengths as well.

The data of Figs. 3-7 pertaining to the infrared wavelengths are

Table II—Optimum alpha values

λ (microns)	α_{opt}
0.6328	2.211
0.73	2.130
0.82	2.081
0.90	2.037
0.983	2.000
1.06	1.974
1.20	1.925
1.32	1.884
1.40	1.856
1.55	1.807

summarized in Figs. 8 and 9. Figure 8 shows the slope of the line relating N_g to N_e as a function of wavelength for various α values. We see that most infrared wavelengths of interest require an α value in the range from 1.8 to 2.1. Figure 9 shows the same data, but this time as a function of α for various wavelengths. The same conclusion is obvious.

Figure 10 displays a calculation of the group indices for all the bound

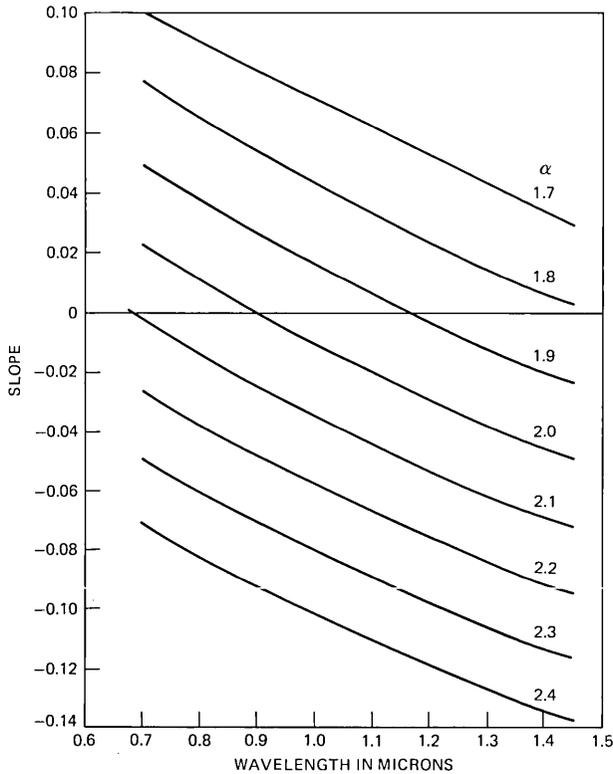


Fig. 8—Plot of the slope of N_g versus N_e as a function of λ for various α values.

modes in the fiber for $\lambda = 0.82$ micron and $\alpha_{\text{opt}} = 2.081$. This plot gives a vivid representation of the fiber characteristics. We see that indeed $\alpha = 2.081$ gives a good equalization for all the modes if those being influenced by the cladding are ignored. A delay-time calibration corresponding to 1 ns/km is also included in the plot.

IV. DISCUSSION

Within the framework of the WKB approximation, Marcatili¹³ has obtained an exact result for α_{opt} that minimizes pulse widths. His formula is as follows:

$$\alpha_{\text{opt}} = (2 - p)(1 + \sqrt{1 - 2\Delta}) - 2, \quad (7)$$

where

$$p = \frac{N_1}{N_g^0} \cdot \frac{\lambda}{\Delta} \cdot \frac{d\Delta}{d\lambda}, \quad (8)$$

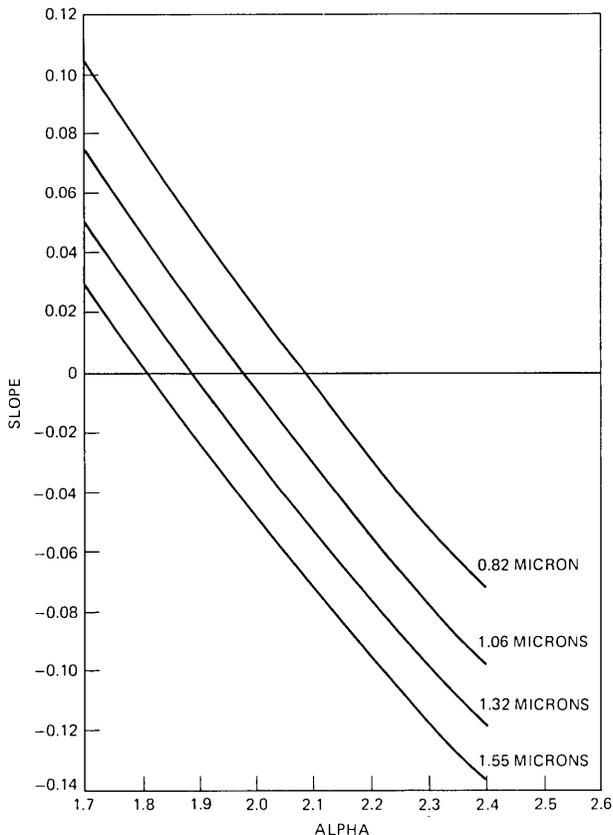


Fig. 9—Plot of the slope of N_g versus N_e as a function of α for various λ values.

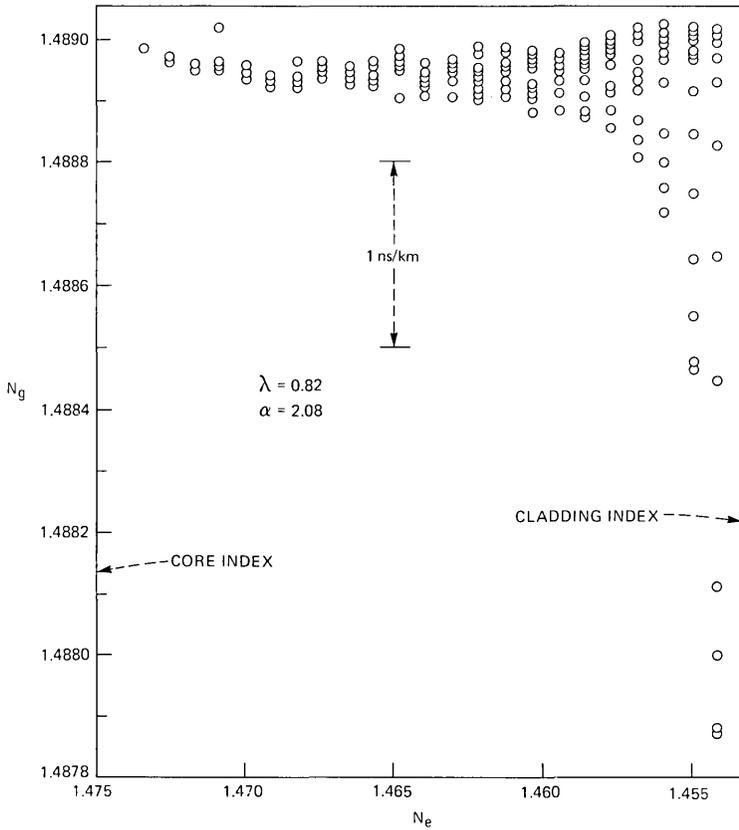


Fig. 10—Plot of group index versus effective index for all the bound modes in the lightguide. We see that for $\lambda = 0.82$ micron and $\alpha = 2.081$ there is a good equalization of all the group indices.

$$N_g^0 \text{ is the group index on axis, and } \Delta = \frac{N_1 - N_2}{N_1}. \quad (9)$$

Also, within the same WKB framework, Olshansky and Keck¹⁴ have obtained an approximate formula for α_{opt} that minimizes the rms width of a pulse. Approximations that coincide with Olshansky and Keck's can also be obtained from the exact result of Marcatili.

It is clear that α_{opt} depends upon the type and number of approximations employed, how the various modes are weighted, and whether we wish to minimize pulse width, rms width, or some other parameter. Another possibility would be to define an optimum α that would include the effects of those modes having significant energy transported in the cladding. This is likely to lead to an α quite different from that obtained in this work.

Figure 11 shows the results of Olshansky, Marcatili, and ourselves. There is great similarity in the shape of the curves and in the actual numbers. The deviation between our results and the WKB results is largest at the longer wavelengths. It is obvious that the agreement between all results is good and, at the present state of the manufacturing art, can be considered the same. However, this may not necessarily be the case in the future. To illustrate (see Fig. 12), we calculate N_g as a function of N_e at 1.55 microns for $M = 0$ and $\alpha_{opt} = 1.775$ (Olshansky) and $\alpha_{opt} = 1.807$ (this work). There is an obvious difference in slope; this difference would be significant in the design of high bandwidth lightguides.

A good representation of our optimum α values is given by the following formulas:

$$0.63 \text{ micron} \leq \lambda \leq 1.00 \text{ micron},$$

$$\alpha_{opt} = 2.9970857 - 1.6647237\lambda + 0.6629181\lambda^2, \quad (10)$$

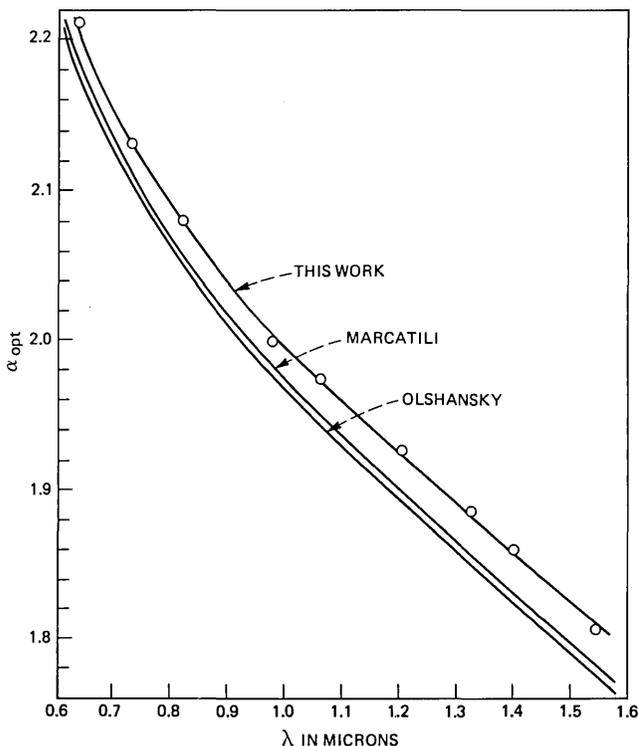


Fig. 11—Plots of optimum profile parameter α versus wavelength for a 13.5 mole percent germania-doped silica lightguide. The index dispersion data employed is from Fleming's work. A clear similarity exists among the results of Marcatili, Olshansky, and this work.

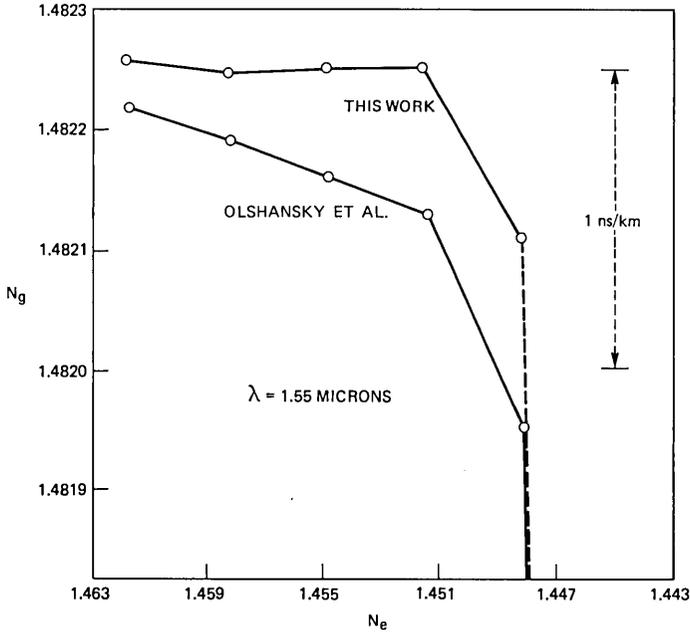


Fig. 12—A plot of N_g versus N_e for $\lambda = 1.55$ microns. $\alpha_{opt} = 1.775$ is calculated from the formula of Olshansky et al., and $\alpha_{opt} = 1.807$ is from this work.

$$1.00 \text{ micron} \leq \lambda \leq 1.55 \text{ microns,}$$

$$\alpha_{opt} = 2.3835734 - 0.4170708\lambda + 0.0290252\lambda^2. \quad (11)$$

It is reasonable to assume that by linear interpolation or extrapolation we can generate the refractive index dispersion curves for germania-doped silica in the range from 0 to 16 mole percent. Using this data, we can then calculate the dependence of N_g on N_e for a number of levels of doping. Figure 13 shows a family of curves for $M = 0$, $\alpha = 1.884$, and $\lambda = 1.32$ microns. There is a good equalization of N_g for all the compositions. We also see that, as expected, the number of modes decreases as the doping level drops. Thus, at the 1 percent doping level for $M = 0$ only one N_e exists. As usual, the last couple of modes for each doping level show a substantial cladding effect.

Finally, the data displayed in Figs. 8 and 9 may have some implications to lightguide engineering. From Fig. 8, we see that the various α curves pass through zero with about the same slope. This suggests that the range of operating wavelengths will be the same for all design wavelengths. Likewise from Fig. 9, we see that the various λ curves

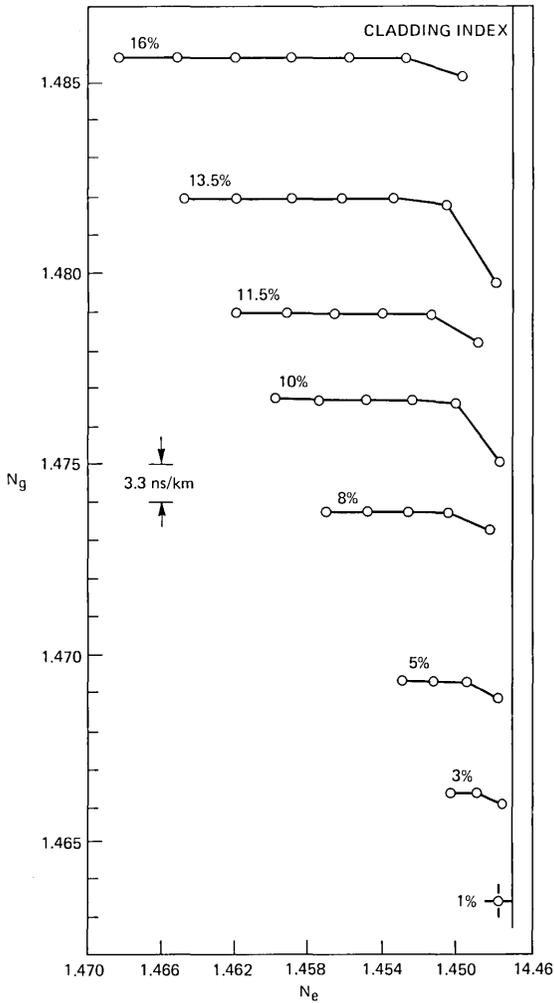


Fig. 13—A family of plots of N_g versus N_e for various doping percentages of germania. The α value is 1.884 and $M = 0$.

pass through zero with similar slopes. This suggests that the allowable error in α will be about the same for all design α 's.

V. ACKNOWLEDGMENTS

We greatly appreciate the enthusiasm for this work expressed by M. I. Cohen, L. S. Watkins, and R. J. Klaiber.

REFERENCES

1. D. B. Keck and R. Bouillie, "Measurements on High-Bandwidth Optical Waveguides," *Opt. Commun.*, *25* (April 1978), pp. 43-8.
2. M. Nakahara et al., "Ultra Wide Bandwidth V.A.D. Fibre," *Electron. Lett.*, *16* (May 1980), pp. 391-2.
3. M. Nakahara et al., "Fabrication of Low-Loss and Wide Bandwidth V.A.D. Optical Fibres at $1.3\ \mu\text{m}$ Wavelength," *Electron. Lett.*, *16* (January 1980), pp. 102-3.
4. D. Marcuse, "Calculation of Bandwidth From Index Profiles of Optical Fibers 1: Theory," *Appl. Opt.*, *18* (June 1979), pp. 2073-80.
5. R. Olshansky, "Propagation in Glass Optical Waveguides," *Rev. Mod. Phys.*, *51* (April 1979), pp. 341-67.
6. S. Ishikawa, K. Furuya, and Y. Suematsu, "Vector Wave Analysis of Broadband Multimode Optical Fibers With Optimum Refractive Index Distribution," *J. Opt. Soc. Am.*, *68* (May 1978), pp. 577-83.
7. G. E. Peterson et al., "An Exact Numerical Solution to Maxwell's Equations for Lightguides," *B.S.T.J.*, *59* (September 1980), pp. 1175-96.
8. J. W. Fleming, "Material Dispersion in Lightguide Glasses," *Electron. Lett.*, *14* (May 1978), pp. 326-8.
9. M. Herzberger, *Handbook of Physics*, New York: McGraw-Hill, 1958, E. U. Condon and H. Odishaw, Eds., pp. 6-42 and 6-43.
10. A. Vigants and S. P. Schlesinger, "Surface Waves on Radially Inhomogeneous Cylinders," *IEEE Trans.*, *MTT-10* (September 1962), pp. 375-82.
11. M. O. Vassel, "Calculation of Propagating Modes in a Graded Index Optical Fiber," *Opto-Electron.*, *6* (July 1974), pp. 271-386.
12. Anthony Ralston, "Runge-Kutta Methods With Minimum Error Bounds," *Math. Comput.*, *16* (October 1962), pp. 431-7.
13. E. A. J. Marcatili, "Modal Dispersion in Optical Fibers With Arbitrary Numerical Aperture and Profile Dispersion," *B.S.T.J.*, *56*, No. 1 (January 1977), pp. 49-63.
14. R. Olshansky and D. B. Keck, "Pulse Broadening in Graded-Index Optical Fibers," *Appl. Opt.*, *15* (February 1976), pp. 483-91.

Multistage Decoding of Frequency-Hopped FSK System

By UZI TIMOR

(Manuscript received July 17, 1980)

A recent paper described an improved decoding scheme for a frequency-hopped multilevel FSK system. We examined this multiple access communication system for possible application in satellite communication and mobile radio telephony. The new decoder, using the known algebraic structure of the users' addresses, reduces mutual interference and achieves a 50 to 60 percent increase in efficiency over conventional decoding. The present paper shows how additional decoding can further increase the efficiency, bringing it very close (within half a percent) to optimum. The scheme makes use of information derived while decoding the messages of other users and thus is especially attractive for the base station, where such information is readily available and does not require a significant increase in complexity. Compared to conventional decoding, the new scheme more than doubles the number of simultaneous users.

I. INTRODUCTION

In a recent paper¹ a new decoding scheme for a frequency-hopped multilevel FSK system was described. This system, where M users share a common frequency band, has been examined for applications in multiple access satellite communication² and mobile radio telephony.³ Each user conveys a K -bit message every T seconds by transmitting a sequence of L tones (chips) chosen from an alphabet of 2^K sinewaves of duration $\tau (=T/L)$. Each user is assigned an address (code) and the message is modulated onto the address. The receiver, knowing the address, decodes the received signal and extracts the message. However, transmissions by other users can combine to cause an erroneous message resulting in an ambiguous reception. Thus, even without channel impairments the number of simultaneous users the system can support at a given error probability is interference limited.

Whenever an ambiguous reception occurs, i.e., more than one mes-

sage value is decoded, the conventional decoder can't identify the correct message. The new decoder¹ makes use of the algebraic structure of the addresses to eliminate erroneous messages that come from interference. Each one of the L chips comprising a message is checked to determine whether it is part of an interference pattern, i.e., a possible sequence transmitted by some other user. An interference message will always be identified as such. The correct message will usually fail to have interference patterns for some chips and thus will be identified and correctly decoded.

This additional decoding (stage 2) results in a substantial improvement in performance over conventional decoding, allowing a 50 to 60 percent increase in the number of users that can simultaneously share the system at a given error probability. Yet it is not optimal and it can be further improved.

In performing stage 2 decoding interference patterns are sought but the (addresses of the) users who might have caused them do not have to be identified. This information is readily available and can be used to eliminate pseudo interference patterns. For example, an interference pattern might "belong" to a user who is not active at the moment. Even if it belongs to an active user it might not be his sequence but a combination of others.

This paper describes how additional decoding (stage 3) eliminates pseudo interference patterns and thus further improves the performance of the system. There is a clear distinction between decoding at the base station and at the mobile unit. The base station decodes the messages of all users and thus any information required for stage 3 decoding is already available. Therefore, additional decoding with improved performance can be accomplished without a significant increase in complexity. In the mobile unit, on the other hand, the additional decoding requires an increase in complexity as the messages of other users have to be decoded.

Upper bounds on the number of simultaneous users that the system can accommodate at a given error probability as well as simulation results are presented. A comparison with a hypothetical decoder, which fails only when the correct message coincides with an interference message (and therefore is undecodable), shows that the stage 3 decoder is very close (within half a percent) to the optimum decoder.

In the noiseless case where the system has a total bandwidth of 20 MHz and the data rate of each user is 32 kilobit/s, the number of users that can simultaneously share the system at a bit error probability of 10^{-3} is increased from 216 (in conventional decoding) through 345 (in stage 2 decoding) to 450 (in stage 3 decoding)—a total improvement of 108 percent. The resulting efficiency (total rate transmitted through the system per unit bandwidth) of stage 3 decoding is 72 percent.

Under noisy and multipath conditions, the number of simultaneous users is reduced but a comparable advantage of the new decoder over the conventional one is maintained.

II. SYSTEM DESCRIPTION

We briefly describe the system here. A detailed description is given in Refs. 2 and 3.

The elementary signals of the system are a set of 2^K sinewaves, which are orthogonal over the chip duration τ . Each user conveys a K -bit message every T seconds by transmitting a sequence of L tones of duration τ chosen from the signal set. The sequence is determined by the user's address and his K -bit message.

Let the address of the m th user be denoted by a vector \mathbf{a}_m ,

$$\mathbf{a}_m = (a_{m1}, a_{m2}, \dots, a_{mL}), \quad (1)$$

where each a_{mi} is a K -bit number corresponding to one of the 2^K frequencies of the system.

The transmitted sequence is

$$\mathbf{Y}_m = \mathbf{a}_m + X_m \cdot \mathbf{1}, \quad (2)$$

where X_m is the K -bit message, and

$$\mathbf{1} = \underbrace{(1, 1, \dots, 1)}_L.$$

Let us denote the 2^K frequencies of the system as elements of $GF(2^K)$, the finite field (Galois field) of 2^K elements $0, 1, \dots, 2^K - 1$. Accordingly, the message X_m , the components of \mathbf{a}_m and \mathbf{Y}_m can be expressed as elements of $GF(2^K)$, and the operations of addition (subtraction) and multiplication (division) are performed according to the rules of $GF(2^K)$. The transmitted sequence \mathbf{Y}_m can be described as a pattern in the $L \times 2^K$ time-frequency matrix A . Simultaneous transmissions by M users will result in up to $L \times M$ entries in A . The receiver performs every τ seconds a spectral analysis of the composite received signal and decides which of the 2^K frequency cells contain energy. Thus, after T seconds, assuming no channel impairments, a duplicate of A is generated at the receiver. To decode X_m , the receiver adds $-a_{mi}$ to each column of A to obtain user m 's decoded matrix A_m . The message appears as a complete row in A_m . Ambiguous decoding occurs when transmissions by other users combine to form other complete rows in A_m .

III. ADDRESS ASSIGNMENT

The interference between users can be minimized by a proper choice of addresses with minimum cross correlation. Schemes for assigning

2^K addresses that guarantee minimum mutual interference between 2^K or fewer users have been proposed⁴ using an algebraic approach.

Let us consider the following address structure: The address of user m is defined to be⁴

$$\mathbf{a}_m = (\gamma_m, \gamma_m\beta, \dots, \gamma_m\beta^{L-1}), \quad (3)$$

where γ_m is the element in $GF(2^K)$ assigned uniquely to user m and β is a primitive element in $GF(2^K)$, which is fixed for the system.

Note that the modulation (2) and address assignment (3) are not unique. Other schemes are possible (see for example Refs. 4 and 5) and the analysis and results that follow can be extended to such systems.

IV. STAGE 2 DECODING

Let M be the number of simultaneous users of the system. Each user i transmits a sequence \mathbf{Y}_i (2) where the address \mathbf{a}_i is assigned according to (3) and the message value is X_i .

In the decoded matrix A_m of user m , X_m appears as a complete row $X_m \cdot \mathbf{1}$. Suppose we have another complete row X' in A_m that is the result of interference. According to the address assignment, each chip in X' must come from a different user, so at least L users i_1, \dots, i_L have combined to cause this interference. We denote by i_n the user that contributes the interference in column n of X' . (If more than one user caused this interference, i_n can be any one of them.) To simplify notation, let us denote the address element (γ_{i_n}) and the message value (X_{i_n}) of user i_n by γ_{in} and X_{in} , respectively.

As mentioned in the previous section, the number (location) of each row in A_m is an element of $GF(2^K)$. We can subtract (the number of) row X' from all rows in A_m to obtain a new matrix $D_{X'}$, where row X' will now be at row zero.

Let $q_n(j)$ denote an entry in row q_n and column j in $D_{X'}$.

It was shown¹ that a necessary condition for row X' in A_m to be caused by interference is the existence of L nonzero numbers δ_n in $GF(2^K)$ such that all entries,

$$q_n(j) = \delta_n(\beta^{j-1} - \beta^{n-1}), \quad j, n = 1, \dots, L, \quad (4)$$

appear in $D_{X'}$. The entries $q_n(j)$ are the contribution of user i_n with address element $\gamma_{in} = \delta_n + \gamma_m$, where γ_m is the address element of user m . From (4) we can derive the following relation:

$$q_n(n+1) = q_n(j)f_{j-n}, \quad n = 1, \dots, L-1, \quad (5)$$

$$q_L(L-1) = q_L(j)f_{j-L}^*, \quad (6)$$

where

$$f_{j-n} = \frac{\Delta}{\beta^{j-n} - 1}, \quad j = 1, \dots, L, \quad j \neq n, \quad (7)$$

$$f_{j-L}^* = \frac{1}{\beta} f_{j-L}. \quad (8)$$

To check the condition (4) for $n (n = 1, \dots, L)$ we multiply all entries (i.e., their row numbers) of column j in $D_{X'}$ ($j = 1, \dots, L, j \neq n, j \neq n + 1$) by f_{j-n} (or f_{j-L}^*) and look for a “complete” row (complete except the n th term). If we find such a row, then there exists a possible interferer contributing to column n of X' .

If the condition is satisfied for all $n = 1, \dots, L$, we assume that row X' in A_m is the result of interference. Using this decoding scheme, all interference rows will be identified as such. The correct row X_m will usually fail to satisfy the condition for some n and thus could be identified and decoded as the correct message.

It was shown¹ that a simple way to perform the column multiplications required during the test is to express each row number by an exponent of β . Since β is a primitive element of $GF(2^K)$ this is equivalent to a (fixed) row permutation of $D_{X'}$. This way the multiplications can be substituted by cycle shifting of columns, which is easy to implement. Thus, stage 2 decoding requires a small increase in complexity (of an order of L^2 cycle shifts) yet achieves a substantial improvement in performance. We show below how additional decoding can further improve the performance.

V. STAGE 3 DECODING

5.1 Discussion

Stage 2 decoding fails to identify the correct message X_m when for every $n (n = 1, \dots, L)$ a possible interferer i_n exists such that all entries $q_n(j), j = 1, \dots, L$, appear in D_{X_m} .

If such L interferers actually exist, that is, if X_m coincides with an interference row, there is no way to distinguish between X_m and any other interference row and no decoder can identify the correct message.

There are however many cases when only $L - j$ chips of X_m are (also) the result of interference and the remaining j chips, although not caused by interference, satisfy the interference condition. That is, entries from other users combine to form interference patterns at those chips. If those pseudo interferences can be identified, for at least one chip, X_m will be identified and correctly decoded. This will be done by making use of available information, derived but not used in stage 2 decoding.

Recall that a possible interference pattern at chip n of X_m was

identified as a "complete" row (complete except for the term in the n th column) in D_{X_m} . From (4) to (6) this row will be at row number q_n , where

$$q_n = \begin{cases} \delta_n(\beta^n - \beta^{n-1}), & n = 1, \dots, L-1, \\ \delta_L(\beta^{L-2} - \beta^{L-1}), & n = L. \end{cases} \quad (9)$$

Knowing q_n (which is found during the test), we can easily compute δ_n and the address $\gamma_{in} = \delta_n + \gamma_m$ of the possible interferer. The message value X^* associated with this interference can be computed as follows: The transmitted sequence of user i_n with message value X^* is (2),

$$Y_{in} = \gamma_{in}\beta^{j-1} + X^* \cdot 1. \quad (10)$$

The entry at column n of A_m will be at row number

$$\gamma_{in}\beta^{n-1} + X^* - \gamma_m\beta^{n-1} = \delta_n\beta^{n-1} + X^*. \quad (11)$$

But by definition i_n is the user who contributes the interference at column n of X_m . Thus,

$$\delta_n\beta^{n-1} + X^* = X_m$$

or

$$X^* = X_m - \delta_n\beta^{n-1}. \quad (12)$$

Substituting δ_n from (9), we can write

$$X^* = \begin{cases} X_m - \frac{q_n}{\beta - 1}, & n = 1, \dots, L-1, \\ X_m - \frac{q_L\beta}{1 - \beta}, & n = L. \end{cases} \quad (13)$$

5.2 Principle of decoding

Let us assume that the list of (the addresses of) the active users (those who are currently using the system) is known. This is certainly the case for the base station that communicates with all its active users, but the list could also be relayed to the mobile unit.

Thus, if an address γ_{in} of a possible interferer belongs to a user who is not active, the interference can be immediately identified as pseudo interference and eliminated.

If the address γ_{in} belongs to an active user, there are two possibilities:

(i) $X^* = X_{in}$. The transmitted message X_{in} and the computed message X^* associated with the interference pattern are the same. In this case the interference was actually caused by user i_n .

(ii) $X^* \neq X_{in}$. In this case the interference pattern was not caused by user i_n , but is the result of a combination of (at least) $L-1$ other users. Thus, this is a pseudo interference and should be eliminated.

Consider the decoded matrix A_{in} of user i_n . His transmitted sequence

will appear as a complete row at row number X_{in} . If $X^* \neq X_{in}$, then we will have another complete row (an interference row) in A_{in} at row number X^* . By decoding A_{in} (stage 2 or stage 3 when necessary) the correct message X_{in} will usually be identified. Thus, if $X^* \neq X_{in}$, it will be identified as pseudo interference of user m and discarded.

If there are several possible interferers at column n of X_m (more than one "complete" row is found in the test) all of them must be checked and discarded before we can conclude that interference is not at that column. If we find (at least) one of the L chips for which no actual interference exists, we can conclude that X_m is the correct message.

If we cannot assume knowledge of the list of active users, then the pseudo interference pattern attributed to nonactive users cannot be identified as such since there is no message X_{in} to be decoded.

VI. DECODING AT THE BASE STATION AND AT THE MOBILE UNIT

6.1 Base station

The base station has a list of all active users and decodes their messages. The decoding procedure can be described as follows.

Step 1 (for each user m)

(i) The decoded matrix A_m is generated and searched for complete rows. If there is only one such row, the message is decoded and listed in a list ($L1$) of users whose decoding is completed.

(ii) If there is more than one complete row, stage 2 decoding is performed. If the message is decoded it is transferred to $L1$.

(iii) If the message is not identified, the address and message values of possible interferers are computed [according to (9) and (13)]. Those are compared with the list of nonactive users and with $L1$ to eliminate a pseudo interference pattern.

At any stage, if the message is decoded it is transferred to $L1$. Similarly, if the message is found to be undecodable (i.e., all interference patterns are found to be true interferences) it is transferred to list $L2$ of undecodable users.

At the end of step 1 (for all users) we have N_1 users in list $L1$, N_2 users in $L2$ and $N_3 = M - (N_1 + N_2)$ users, each with a list of possible interferers and associated message values for some of its chips. If $N_3 > 0$, we proceed to step 2.

Step 2 (for users who have not been decoded)

For each of the N_3 users, we compare his list of possible interferers with $L1$ to eliminate pseudo interferers. As list $L1$ is increased, more possible interferers can be checked and more users decoded.

The procedure terminates when $N_3 = 0$ or when $L1$ remains unchanged after a complete cycle of checking all N_3 users.

6.2 Mobile unit (user m)

(i) The decoded matrix A_m is generated and searched for complete rows. If (up to) stage 2 decoding does not identify the message a list of possible interferers and their associated message values is computed.

(ii) Pseudo interferences from nonactive users (if list is available) are identified and discarded.

(iii) If the message is still undecoded, the decoded matrices of possible interferers have to be generated and decoded (stage 2) until enough pseudo interferences are identified to decode the message.

We define this as stage 3/2 decoding since the messages of possible interferers are decoded up to stage 2.

We could in principle perform stage 3 decoding of the possible interferers; however, the complexity increases exponentially as more and more users have to be decoded.

7. ERROR PROBABILITIES

7.1 Upper bounds

We have an ambiguity in decoding when the following two conditions are satisfied.

(i) There are two or more complete rows in A_m .

(ii) The correct row cannot be identified, i.e., all its chips have interference patterns (actual interference or unidentified pseudo interference).

When this happens, we choose one of the complete rows at random and decode it as the message.

Although not strictly independent, it can be shown using random coding arguments that the two conditions can be assumed to be independent with a negligible effect on the probability of error. Thus, $P_E^{(i)}$, the word-error probability when stage i decoding is performed, is given by

$$P_E^{(i)} = P_1 P_2^{(i)}, \quad (14)$$

where P_1 is the probability of condition 1 and $P_2^{(i)}$ is that of condition 2 when stage i decoding is performed.

If 2^K is the number of frequencies, L is the length of the sequence and M is the number of simultaneous users, P_1 can be upper-bounded² by

$$P_1 < (2^K - 1) p^L, \quad (15)$$

where

$$p = 1 - (1 - 2^{-K})^{M-1}. \quad (16)$$

Let $P_{2,j}^{(i)}$ be the probability of condition 2 (stage i decoding) when j

chips in X_m come from interference and the remaining $L-j$ chips have a pseudo interference pattern. We then have

$$P_{\frac{1}{2}}^{(i)} = \sum_{j=0}^L P_{\frac{1}{2},j}^{(i)}. \quad (17)$$

(a) Stage 2 decoding

For stage 2 decoding we have shown¹ that $P_{\frac{1}{2},j}^{(2)}$ can be upper-bounded by

$$P_{\frac{1}{2},j}^{(2)} < p^L \binom{L}{i} S^j, \quad (18)$$

where

$$S = (2^K - 1)(1 - p)p^{L-2}. \quad (19)$$

Thus,

$$P_{\frac{1}{2}}^{(2)} < p^L(1 + S)^L, \quad (20)$$

and as long as $p(1 + S) < 1$ the word-error probability can be upper-bounded by

$$P_E^{(2)} < (2^K - 1)(1 + S)^L p^{2L} \quad (21)$$

and the bit-error probability by

$$P_b^{(2)} < 2^{K-2}(1 + S)^L p^{2L} \quad (22)$$

(b) Stage 3 decoding

$P_{\frac{1}{2},L}^{(3)}$ corresponds to the case where all interference patterns actually come from other users (i.e., X_m coincides with an interference row in A_m). Thus,

$$P_{\frac{1}{2},L}^{(3)} = P_{\frac{1}{2},L}^{(2)} < p^L \quad (23)$$

Consider the case where $L-1$ chips in X_m actually come from interference and the remaining chip has a pseudo interference pattern. This pseudo interference will correspond to an active user with probability.

$$\bar{p} = \frac{M-1}{2^K-1} \quad (24)$$

The probability that the correct message of this user (and therefore the pseudo interference) will not be identified as such when stage 3 decoding is performed for that user is $P_{\frac{1}{2}}^{(3)}$. Thus we can upper-bound

$$P_{\frac{1}{2},L-1}^{(3)} < P_{\frac{1}{2},L-1}^{(2)} \bar{p} P_{\frac{1}{2}}^{(3)}. \quad (25)$$

Similarly,

$$P_{2,L-j}^{(3)} < P_{2,L-j}^{(2)} (\bar{p} P_2^{(3)})^j. \quad (26)$$

From (17), (18), and (26) we get

$$P_2^{(3)} < p^L (1 + S \bar{p} P_2^{(3)})^L. \quad (27)$$

Let P_2^* be the solution of (27) when an equality is substituted for the inequality. Then

$$P_2^{(3)} < P_2^* = p^L (1 + S \bar{p} P_2^*)^L. \quad (28)$$

(c) Stage 3/2 decoding

In a similar way we can upper-bound $P_2^{(3/2)}$, the probability of error of condition 2 when the messages of possible interferers are decoded up to stage 2. This is relevant only to the mobile unit (see section 6.2).

Case 1: List of active users available

Substituting $P_2^{(2)}$ for $P_2^{(3)}$ in (25) yields

$$P_2^{(3/2)} < p^L (1 + S \bar{p} P_2^{(2)})^L, \quad (29)$$

where $P_2^{(2)}$ is given by (20).

Case 2: List of active users unknown

In this case we will always have an error when a possible interferer is not an active user (which occurs with probability $1 - \bar{p}$) since there is no correct message to be decoded. Thus

$$P_2^{(3/2)} < p^L [1 + S (\bar{p} P_2^{(2)} + 1 - \bar{p})]^L. \quad (30)$$

(d) Coincidence of X_m with an interference

Whenever we have two or more interference rows in A_m and the correct row coincides with one of them, it cannot be identified and decoded. An optimal decoder can do no better than a decoder which fails only when such a coincidence occurs. Thus, it is of interest to compare the various schemes with such a (hypothetical, not necessarily achievable) decoder.

The probability of condition 2 for such a decoder is given by (23)

$$P_2^{(\text{coin})} < p^L. \quad (31)$$

We can summarize the performance of the various schemes as follows: The word-error and the bit-error probabilities are upper-bounded by

$$P_E^{(i)} < (2^K - 1) g_i^L p^{2L}, \quad (32)$$

$$P_b^{(i)} < 2^{K-2} g_i^L p^{2L}, \quad (33)$$

where

$$g_i = \begin{cases} 1 + S, & \text{stage 2 decoding,} \\ 1 + S\bar{p}P_2^*, & \text{stage 3 decoding,} \\ 1 + S\bar{p}P_2^{(2)}, & \text{stage 3/2 decoding} \\ & \text{(list of users known),} \\ 1 + S(\bar{p}P_2^{(2)} + 1 - \bar{p}), & \text{stage 3/2 decoding} \\ & \text{(list of users unknown),} \\ 1, & \text{coincidence fail decoder,} \end{cases} \quad (34)$$

and S , $P_2^{(2)}$, \bar{p} , and P_2^* are given by (19), (20), (22), and (28), respectively.

The upper bounds of the various schemes as a function of the number of simultaneous users (M) are depicted in Fig. 1 for the case $K = 9$ (512 frequencies) and $L = 11$ (which is optimal when the total bandwidth is 20 MHz and each user's rate is 32 kilobit/s).

As can be seen, at small error probabilities the performance of stage 3 decoding is very close to the probability of coincidence, that is, almost all messages which do not coincide with an interference are correctly decoded. As the number of users is increased, a strong threshold effect occurs (at $M = 475$ in the above example) and the performance deteriorates rapidly. This is due to the fact that as M increases, the interference is so high that a large number of users cannot be decoded (stage 2) and thus the information (decoded message values of possible interferers) required for stage 3 decoding is not available.

As M is decreased, all the above bounds converge to the coincidence bound.

7.2 Simulations

The various decoding schemes were simulated on a Digital Equipment Corporation PDP 11 computer. Figure 1 shows simulation results for stage 2, stage 3, and stage 3/2 (when the list of active users is known) decoding as well as for the probability of coincidence.

It can be seen that for bit-error probabilities up to 10^{-3} , the simulated results are within 5 percent of the upper bounds on the number of simultaneous users.

The threshold effect of stage 3 decoding is clearly seen. As M increases there are some K -bit blocks where the majority of users (60 percent and more) are undecodable. The percentage of undecodable blocks increases from 6 percent at $M = 478$ to over 75 percent at $M = 490$.

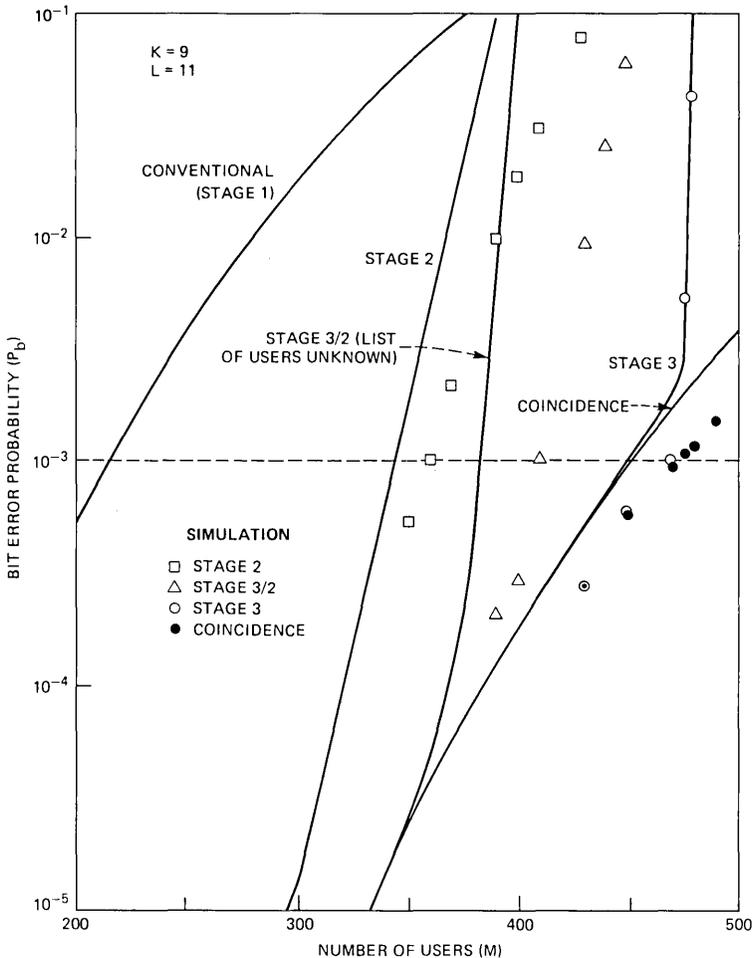


Fig. 1—Upper bounds and simulations of the bit-error probabilities as a function of the number of users M for stage 2, stage 3/2 (list of active user known), and stage 3 decoding and for coincidence of the message with interference. The system parameters are $W = 20$ MHz, $R = 32$ kbit/s, $K = 9$, and $L = 11$.

VIII. CONCLUSIONS

The performance of the frequency-hopped multilevel FSK system can be substantially improved by making use of the known structure of the addresses to perform additional decoding. In stage 2 decoding additional checks are performed on the user's own decoded matrix but no reference to the decoded matrices of other users is needed. The stage 3 decoder achieves a performance that is very close to optimum by referring to results obtained in decoding the sequences of other users.

In the base station, where all users are decoded anyway, stage 3 decoding can be realized without a significant increase in complexity. This is not the case at the mobile unit; therefore, an intermediate stage of decoding (stage 3/2) was proposed, requiring a moderate increase in complexity.

When the total bandwidth is 20 MHz and the transmission rate of each user is 32 kilobit/s, the number of simultaneous users that the system can accommodate at bit-error probability of 10^{-3} is increased from 216 (conventional decoding) through 345 (stage 2) to 383 (stage 3/2) and 450 (stage 3). The corresponding efficiencies of the system (total rate transmitted through the system per unit bandwidth) are 35 percent (conventional), 55 percent (stage 2), 61 percent (stage 3/2), and 72 percent (stage 3). As can be seen, stage 3 decoding more than doubles the efficiency. Simulation results show that an efficiency up to 75 percent can be obtained with stage 3 decoding.

REFERENCES

1. U. Timor, "Improved Decoding Scheme for Frequency Hopped FSK System," *B.S.T.J.*, 59, No. 10 (December 1980), pp. 1839-55.
2. A. J. Viterbi, "A Processing Satellite Transponder for Multiple Access by Low-Rate Mobile Users," *Digital Satellite Communications Conference*, Montreal, October 23-25, 1978.
3. D. J. Goodman, P. S. Henry, and V. K. Prabhu, "Frequency Hopped Multilevel FSK for Mobile Radio," *B.S.T.J.*, 59, No. 7 (September 1980), pp. 1257-75.
4. G. Einarsson, "Address Assignment for a Time-Frequency-Coded Spread-Spectrum System," *B.S.T.J.*, 59, No. 7 (September 1980), pp. 1241-55.
5. B. G. Haskell, "Computer Simulation Results on Frequency Hopped MFSK Mobile Radio—Noiseless Case," *IEEE Trans. Commun.*, COM-29, No 2 (February 1981), pp. 125-32.



A Bubble Memory Differential Detector

By W. D. WYNN

(Manuscript received September 5, 1980)

In the application of digital memories, significant variations can occur in the memory sensor analog responses processed by detection circuitry to recover stored binary data. Across an ensemble of memory devices of the same type, the variations are due to manufacturing deviations, and for a particular device, they occur with changes in operating conditions. With magnetic bubble memories, binary data are recovered by sensing the presence or absence of a magnetic domain in each bit interval. In general, a nonadaptive bubble detector does not accommodate the variations that may occur in the bubble response when each decision is based only on the sensor response in the current bit interval. Here we present a differential detector design that is reasonably simple to implement as one integrated circuit and is tolerant of signal amplitude drifts in bubble memory sensor outputs.

I. INTRODUCTION

In typical field-accessed magnetic bubble memories (MBM), stored binary data on memory chips are detected by processing millivolt-level signals that are measured differentially across a matched pair of permalloy magnetoresistive sensor strips.¹⁻³ Ideally these sensors are subjected to the same magnetic, electrical, and environmental conditions with the exception that only one, the active sensor, is subjected to H fields of traversing bubble domains. In the drive field period of a field-accessed MBM chip, by design at least one time interval exists, defined here as the detection interval (DI), in which the sensor differential response is relatively free of unavoidable signal interferences such as MBM function drive crosstalk and sensor magnetoresistive switching noise. Within the DI a detectable difference should exist between responses for the bubble and no-bubble cycles of the memory. Even after carefully choosing the DI, undesirable sensor response variations can still occur. Variations within the DI that have been of major concern in the use of previous nonadaptive MBM detectors are

the device processing variation in magnetoresistive sensor response, the negative temperature coefficient of this response, the phase shift of bubble response transition that occurs with chip manufacturing and drive field changes, and the adjacent symbol interaction of bubbles in the sensor array. In a detector where the DI is fixed and where the individual bubble response in the DI is compared to a fixed threshold these variations have been intolerable.

An accurate analytical parametric description of the family of sensor signals that are observed for any MBM chip type has been intractable because of the nonlinear field effects of permalloy magnetoresistive sensor arrays and a lack of models for statistical data about the manufacturing variations for MBM devices. Figure 1 shows qualitative examples of differenced sensor responses within a DI that may exist for a chip type when the manufacturing and operational variations are considered. Each signal pair is shown in the DI for some particular chip of an ensemble of the same chip type at a particular operating condition. All signals have been referenced to a voltage V_r at the start of the DI. The amplitude and time scales for Fig. 1 are representative of field-accessed MBMs with field drive frequencies in the 100-kHz range. Not only can response amplitude change with operating conditions and selected chip, but also the shape of the desired bubble response and intersymbol interference contributions may vary.

Detectors could be designed that adapt to variations in the bubble and no-bubble responses where such variations preclude a fixed threshold detector. For example, detector amplifier gains could be temperature compensated to track the temperature sensitivity of the magnetoresistive sensor. Also, the phase of the DI could be adaptive, and detection thresholds could be adjusted dynamically from responses in a known data pattern. Unfortunately, adaptive bubble detectors have disadvantages such as increased circuit complexity, increased cost, and system limitations such as detector adaptation time before reading data.

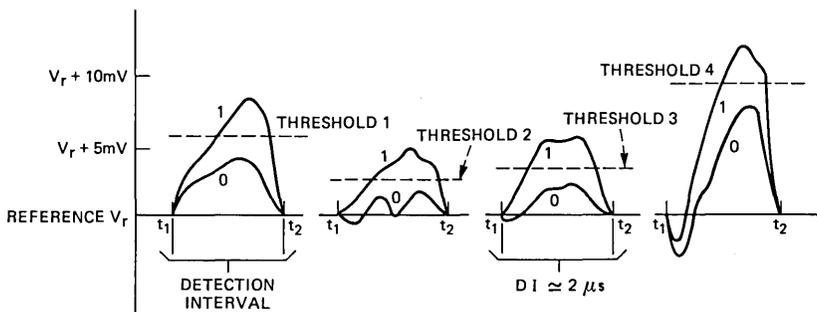


Fig. 1.—Dc restored bubble (denoted by "1") and no-bubble (denoted by "0") differential sensor responses for a memory chip ensemble.

We describe a differential detection approach, and a circuit realization of it, which are insensitive to some signal amplitude, phase and shape variations in the DI, if these variations occur slowly when compared with the memory drive field period. The detector is based on the hypothesis that some signal function, such as peak-to-peak transition or the signal energy, is available in a DI such that the minimum functional change between unlike data pairs is greater than the maximum functional change between like data pairs in memory detection cycles that are contiguous. This condition must hold over the range of manufacturing and environmental changes in the sensor outputs for the chip ensemble, including noise and intersymbol interference effects. This proposed detector can be developed as a small inexpensive IC. It is easily incorporated in MBM circuit designs, and it is compatible with memory system requirements.

II. DIFFERENTIAL MBM SIGNAL DETECTION

2.1 *Signal and interference characteristics of memory output*

As indicated by the qualitative results shown in Fig. 1, the dynamic range of bubble and no-bubble responses in the DI may be significant across the desired operating range for an ensemble of a MBM chip type. However, for a given MBM chip and stationary operating environment, the responses for two data 1's (bubbles) or for two data 0's in successive DI's differ only through random noise and intersymbol interference (ISI) effects.

Sensor responses for bubble and no-bubble traversals depend on the memory sensor design, the chip technology, memory manufacturing quality control, and operating environment. Typically, the bubble response transitions of the MBM sensor in the DI have been in the 2-mV to 20-mV range, with a more restricted dynamic range over the ensemble and operating range of any particular memory type. Although unit-to-unit variations may exist in both bubble and no-bubble responses, as indicated by Fig. 1, a typical memory product gives a minimum transition difference of about 2 mV between the two responses in the DI.

When the phase of the DI is properly selected relative to sporadic random interference, such as magnetoresistive switching noise from the sensor or control signal cross talk, residual noise within the DI becomes a second-order effect. For example, the important frequency band of bubble signals in the DI is below $10 f_f$, where f_f is the MBM field drive frequency. For MBM drive frequencies of less than 200 kHz, the signal processing band for a detector is then less than 2 MHz. For a 2-MHz noise equivalent bandwidth in a detector preamplifier, the total random noise within the DI is less than $50 \mu V$ rms referred to the output contacts of the MBM chip sensor. This includes random noise

from the MBM sensors under active memory conditions, the sensor bias current sources, and the detector electronics.

Fringe field coupling of bubble domains can exist across elements in a sensor array. This effect is observed as ISI in the DI response. Like the desired DI response, the ISI depends on the sensor array design and the memory operating conditions, but it is not well understood. Because of the nonlinear magnetic interaction of permalloy sensor arrays and bubble domains, accurate modeling and analysis of ISI as well as the desired bubble response have been intractable. In memory design, arrays have been selected to reduce the ISI, as well as random noise, to acceptable levels relative to the minimum difference between isolated bubble and no-bubble transitions in the DI. This condition is indicated by Fig. 2a for contiguous DIs for two signal amplitude cases. If ISI is severe for a memory design, this design can be changed or signal equalization can be considered for the detector.^{4,5} However, because of the variable nature of ISI as well as the desired signal's shape in the DI, an equalizer should be adaptive. For typical useful DI durations of a few μs and very short setup times of the memory system, adaptive equalization of ISI is not cost effective in a bubble memory detector. By careful sensor design and processing control for the memory, it has been practical to hold ISI and noise to levels that permit differential detection with probability of error $P_E < 10^{-8}$ without using ISI equalization.

2.2 Data precoding and differential recovery from the MBM

Figure 1 indicates that variations in the absolute levels, phases, and shapes of bubble and no-bubble responses along with a close match of like symbol responses in contiguous DIs make differential amplitude detection attractive for field-accessed MBM. This is analogous to differential phase detection used in communication channels.⁶ If the binary data "1" is stored in the MBM chip as a bubble-to-no-bubble transition, or the reverse transition in contiguous memory cycles, then recovering data from memory is possible when the absolute values of the greatest DI response differences for pairs of like symbols are less than the smallest absolute values of differences for unlike symbol transitions. Figure 2b summarizes these prerequisite conditions for successful difference detection.

Direct difference encoding of the original data has the potential for propagating errors since a single error in decision can result in the loss of phase reference and cause ambiguity in the detector. This problem is eliminated by a simple precoding step for the data prior to recording. The precoding does not reduce the data throughput or storage capacity of the memory.

An original data sequence defined by $\{a_k\}$, where $a_k \in \{0, 1\}$ for k

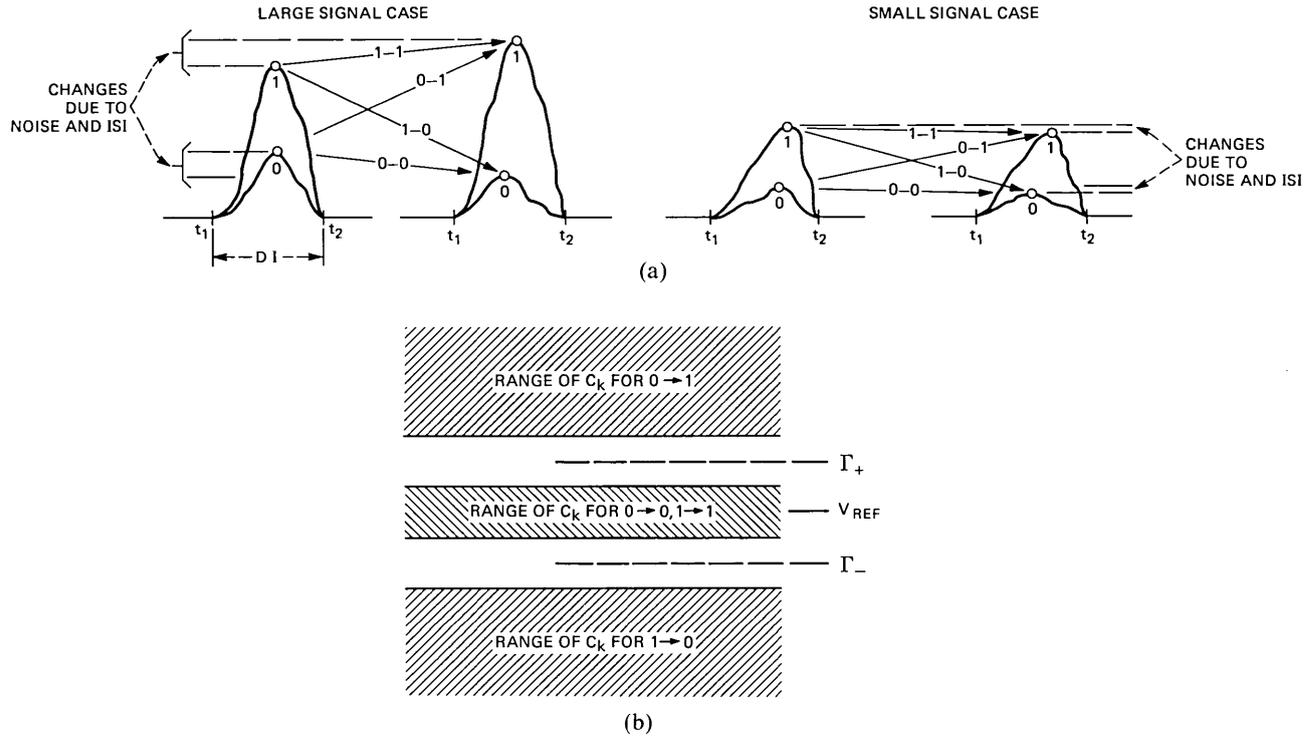


Fig. 2—(a) Dc restored bubble sensor response examples showing effects of dynamic range, ISI, and noise. (b) Response peak difference C_k ranges and Γ_+ , Γ_- threshold placement.

≥ 0 , is to be recorded as transitions between bubble and no-bubble states in consecutive field cycles in a way that eliminates propagating detection errors through loss of reference. This is accomplished by precoding $\{a_k\}$ into an equivalent sequence $\{b_k\}$, where $b_k \in \{0, 1\}$ for $k \geq 0$, by the algorithm

$$b_k = a_k \oplus b_{k-1}, \quad \text{where } b_{-1} = 0, \quad (1)$$

and \oplus denotes modulo 2 addition. Then $\{b_k\}$ is recorded in the memory with a bubble state for $b_k = 1$ and a no-bubble state for $b_k = 0$. The precoding in (1) is implemented with a shift register and an EX-OR logic stage in the memory recording circuit.

If, in response to b_k , a functional B_k is available from the response in the DI and a decision of bubble or no-bubble in cycle k is to be based on B_k , difference detection is performed as follows. In the k th data cycle the analog difference $C_k \triangleq B_k - B_{k-1}$ is formed. Ideally, there are three possible analog levels for C_k corresponding to the four permutations of (B_{k-1}, B_k) , and a ternary decision on C_k results in a three-state logic output $c_k \in \{-1, 0, +1\}$. In terms of the binary sequence $\{b_k\}$,

$$c_k = b_k - b_{k-1}, \quad k \geq 0. \quad (2)$$

From expression (1), the modulo 2 subtraction of b_{k-1} from both sides of (2) gives

$$(c_k)_{\text{mod } 2} = (b_k - b_{k-1})_{\text{mod } 2} = a_k. \quad (3)$$

It follows then that $\{a_k\}$ is available by simple logic operations on the ternary sequence $\{c_k\}$ obtained from $\{C_k\}$ without the problem of error propagation due to phase ambiguity if a detection error occurs.

In practice, the differential detector outputs are estimates of the components of $\{c_k\}$ that are based on the available analog differences $\{C_k\}$. Then the detector outputs form the estimate sequence $\{\hat{c}_k\}$ and the estimate of the original data sequence is

$$\{\hat{a}_k\} = \{(\hat{c}_k)_{\text{mod } 2}\}, \quad k \geq 0. \quad (4)$$

If in the k th cycle there is no detection error, $\hat{c}_k = c_k$ and hence $\hat{a}_k = a_k$. Figure 3 gives an example of the operations discussed in this section, where the functionals $\{B_k\}$ are the positive peaks of dc restored bubble and no-bubble responses in the detection intervals.

2.3 Choice of functional B_k

As discussed above, many variations can occur in the sensor responses from a given type of MBM chip which effect the shape, amplitude, noise content, and position of the responses in a DI. To successfully detect differences, any signal processing of the responses

in modeling the statistical characteristics of the MBM signal and interference components, the analytical solution for a useful optimum detector has been intractable. The selection of any detector, and, in particular any signal processing approach to obtain $\{B_k\}$ for the differential detector considered here, has been based primarily on the performance observed in experiments with a representative ensemble of devices for each MBM design. The optimum choice of a detector is complicated additionally by practical considerations of its circuit realization, cost, and flexibility of operation.

Section 3 describes a differential detector that has been successful with signal characteristics of present bubble memories. This realization uses peak detection of dc restored signals in the DI to form $\{B_k\}$. We have found that signal equalization to reduce ISI and special noise filtering are not necessary in this detector for the bubble memory families of interest.

2.4 Detection errors and monitoring

In Section 2.2 the detected data sequence is

$$\{\hat{a}_k\} = \{(\hat{c}_k)_{\text{mod } 2}\}, \quad (5)$$

where \hat{c}_k is the ternary logic sequence obtained by a bipolar threshold comparison of $C_k = B_k - B_{k-1}$. With the exception of noise and intersymbol effects in the generation of $\{B_k\}$, only two values occur for B_k under the conditions of static operation between memory detection cycles. Then of the nine combinations of c_k, c_{k+1} , the pairs (1, 1) and (-1, -1) can occur only through detection error. The sequence $\{\hat{c}_k\}$ can be monitored for the disallowed patterns as a simple error check, but this is not a complete error test.^{7,8} Detection errors that yield valid c_k, c_{k+1} pairs cannot be detected by this simple monitoring. In general, an error within the memory that results in a recovered bubble sequence other than the $\{b_n\}$ that was recorded cannot be detected with this pattern monitoring. The altered bubble sequence in the MBM appears at the detector as a different $\{b_n\}$ sequence but one that could be generated by an admissible $\{a_n\}$ memory sequence.

Since the difference detector considered here involves the threshold comparison of $C_k = B_k - B_{k-1}$, it would seem that a detection error in recovery of $a_k = (c_k)_{\text{mod } 2}$ caused by a disturbance in acquiring a B_k would be highly correlated with a detection error in $a_{k+1} = (c_{k+1})_{\text{mod } 2}$ since $C_{k+1} = B_{k+1} - B_k$. This apparent coupling exists regardless of the data formatting discussed in Section 2.2 that is used for preventing error strings caused by detector reference loss. Whether it is more probable in difference detection to have an error pair given that one error occurs will depend on the method of generating functional B_k

and the interference that causes detection errors. It is known that in the case of additive Gaussian noise and matched filtering of signals without intersymbol interference, differential detection does not enhance the probability of double errors given single errors.⁹

III. A PARTICULAR DETECTOR CIRCUIT REALIZATION

The circuit shown in Fig. 4 performs the functions discussed in Section 2.2, including forming the functional $C_k = B_k - B_{k-1}$ required in difference detection. The system characteristics outlined below are compatible with available monolithic bipolar IC technology. Speed requirements implied by memory field frequencies up to 200 kHz seem to make bipolar technology essential at present for any analog difference detector type of IC. The realization of the functional C_k is simple; it requires a capacitor, a clamping switch, and a buffer amplifier. Then to obtain C_k it is not necessary to construct a true sample-and-hold for B_{k-1} , which is an advantage in bipolar circuit integration. All capacitors can be integrated and the complete bubble detector implemented as a single inexpensive IC with three synchronous input logic controls.

Figure 5 shows the key voltage waveforms for the circuit for the example data sequence $\{a_k\}$ in Fig. 3. To simplify the operating description, any voltage offsets in switches $SW1$, $SW2$, and $SW3$ are neglected as are the droop effects due to leakage currents from capacitors C_2 , C_3 , and C_4 when these are in their hold states. Adverse effects of switch offsets in an IC realization of the circuit can be eliminated by component trims during manufacture. At points in the circuit where leakage currents from capacitors are critical to performance, bipolar field effect transistor (BIFET) amplifier isolation is assumed; these are shown in Fig. 4. BIFET amplifiers have bias requirements on the order of 1 nA within a circuit ambient range of $[-10^\circ\text{C}, +70^\circ\text{C}]$.

The data cycle period in Fig. 5 is $T = 1/f_f$, where f_f is the MBM drive field frequency. The signal detection interval (DI) is $[t_1, t_2]$, modulo T . The other intervals of interest are $[t_3, t_4]$, modulo T , where switch $SW2$ is closed, and $[t_5, t_6]$, modulo T , where C_3 is discharged through $SW3$. The maximum operating field frequency is limited by the bubble memory technology. The period T considered for the differential detector design is in the range $[5 \mu\text{s}, 40 \mu\text{s}]$.

In each DI, the B_k acquired is the maximum of the peak catcher input

$$e_i^* = e_1 - e_{c2}(t_1) + V_r, \quad \text{for } t \in [t_1, t_2], \text{ modulo } T, \quad (6)$$

where $e_{c2}(t_1)$ is the initial condition of the voltage across C_2 at the start of the DI, and V_r is the switch voltage reference. The value of $e_{c2}(t_1)$ is a function of $e_1(t)$, for $t \leq t_1$, as well as the filter action provided by the R_{02}, C_2, R_{s1} circuit when $SW1$ is closed. The noise that

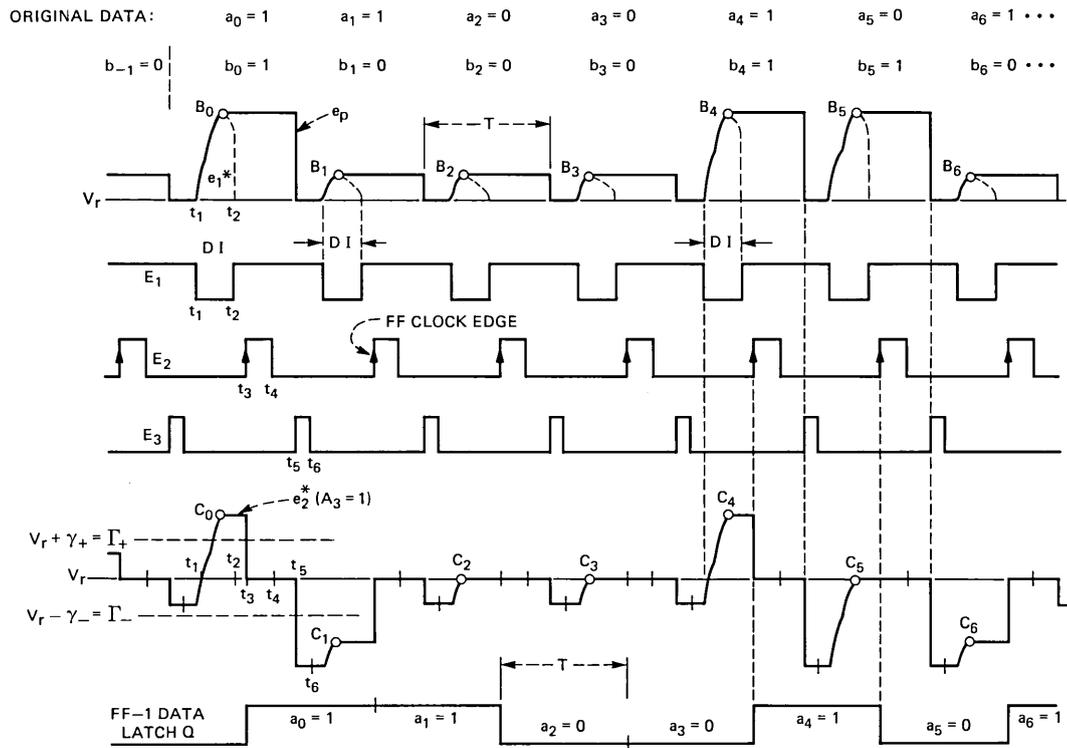


Fig. 5—Ideal signal and timing sequence example for differential detector.

is present with the desired signal input e_1 of the dc Restore I circuit is enhanced in the restored signal e_1^* . Because of this, a nonzero time constant, $\tau_a = (R_{02} + R_{s1})C_2$, can be advantageous. If $\tau_a \rightarrow 0$, then $e_{c2}(t_1) \rightarrow e_1(t_1)$, and $e_1^*(t_1) \rightarrow V_r$, which gives an exact restoration of e_1^* to reference V_r at t_1 . But then the steady state noise variance of e_1^* in DI is twice the noise variance of e_1 .

The functional B_k is retained on the storage capacitor C_3 until discharged by a HI state of E_3 during $[t_5, t_6]$. With E_2 held LOW during the DI as well as the PEAK RESET interval, the Restore Switch II output e_2^* follows the acquisition of B_k and the reset transition that occurs in the peak catcher output e_p when C_3 is reset by SW3. When E_2 is HI, SW2 is closed, and a voltage e_{c4} is stored across C_4 that is a function of both $e_2(t)$ for $t \in [t_3, t_4]$ and the filter time constant $\tau_b = (R_{03} + R_{s2})C_4$. Then for $t \in [t_4, T + t_3]$, modulo T , SW2 is opened, and the comparator (+) input is

$$e_2^* = e_2 - e_{c4}(t_4) + V_r. \quad (7)$$

If $\tau_b \rightarrow 0$, then $e_{c4}(t_4) \rightarrow e_2(t_4)$ and an exact restoration of e_2^* to V_r occurs. But if noise is present in e_2 , the filtering provided when $\tau_b \neq 0$ may result in improved performance to a degree that depends on the signal and noise characteristics of e_2 .

For the time sequence of controls, E_1 , E_2 , and E_3 shown in Fig. 5, the value of $e_2^*(t)$ when $t = (kT + t_3^-)$ and $k = 0, 1, 2, \dots$, is $A_3 C_k = A_3 (B_k - B_{k-1})$, where A_3 is the gain of the second interstage amplifier. When, in response to like and unlike symbols in the data pair (b_{k-1}, b_k) , the analog values $C_k = B_k - B_{k-1}$ lie in three nonoverlapping bands, as shown in Fig. 2b, the recovery of the original sequence $\{a_k\}$ follows from the trilevel comparison of each $A_3 C_k$ value with a threshold pair (Γ_-, Γ_+) . If $A_3 C_k \geq \Gamma_+ = V_r + \gamma_+$, the trilevel logic response is $c_k = +1$; if $A_3 C_k \leq \Gamma_- = V_r - \gamma_-$, the decision is $c_k = -1$; and for all other values of $A_3 C_k$, $c_k = 0$. Then from eq. (3), the original sequence $\{a_k\}$ is the output of the EX-NOR gate in Fig. 4. Amplifier A_3 in conjunction with the dc Restore II circuit provides the desired differencing C_k in each detection cycle. A complete sample-and-hold function is unnecessary since the difference C_k and not the absolute levels of B_k and B_{k-1} are required for differential detection.

In the k th data cycle, a_k is latched in FF1 by E_2 . The same positive edge of E_2 clocks FF2 to latch the logic level out of comparator CP1. The DATA and DATA MONITOR outputs are sufficient to detect invalid (c_k, c_{k+1}) pairs $(+1, +1)$ and $(-1, -1)$ that may occur if there are detection errors. This was described in Section 2.4.

IV. DETECTOR OPERATION EXAMPLE

The detector shown in Fig. 4 was constructed from discrete BIFET and bipolar transistor devices that are typical of those that could be

used in making a monolithic linear IC. Table I lists a summary of the conditions for this discrete component circuit. We chose the parameters listed in Table I, including values of the two thresholds Γ_+ and Γ_- , for an MBM type that seems to be typical of the field-accessed bubble-memory art. For other MBM device designs, it may be desirable to have different parameter values than those considered here to maximize detection performance.

Fig. 6a shows superimposed bubble and no-bubble responses for e_1 and the restored signals e_1^* and e_2^* for a recorded $\{b_n\}$ sequence in the MBM that is the periodic hexadecimal pattern ABEF5410. One of the two orthogonal coil currents, I_x , that comprise the bubble memory square wave drive field, is shown also. In the e_2^* trace of Fig. 6a, the detector "EYE" diagram is formed by the three levels at the right side of the peaks of e_1^* . The levels of this EYE correspond to the trilevel logic outputs $c_k \in \{-1, 0, +1\}$. The thresholds Γ_+ and Γ_- , at the negative inputs of comparators $CP1$ and $CP2$, are positioned within the EYE to get minimum detection error probability over the memory ensemble and its range of operating conditions. Nonadaptive detection is possible if Γ_+ and Γ_- can be placed between the center band and the upper and lower bands of the e_2^* EYE. The difference detector performance is immune to the absolute levels of the peaks of e_1^* that occur within the detection intervals, but it does depend on the minimum change in peaks for unlike contiguous data pairs and the maximum change for like data pairs. The widths of the EYE are determined by

Table I—Detector Parameters

-
1. Sensor resistances: $R_D, R_{DD} = 1600 \Omega$
 2. Bias currents: $I_{B1}, I_{B2} = 3.75 \text{ mA}$
 3. SW1 filter time constant: $[R_{02} + R_{s1}]C_2 \cong 60 \text{ ns}$
 4. SW2 filter time constant: $[R_{03} + R_{s2}]C_4 \cong 60 \text{ ns}$
 5. Amplifier A_1 : gain = 10, 3 dB frequency = 3 MHz, critically-damped two-pole response.
 6. Amplifier A_2 : gain = 5, 3 dB frequency = 3 MHz, two-pole response with 1 dB overshoot.
 7. Amplifier A_3 : gain = 3, frequency response same as A_2 .
 8. Capacitors: $C_1, C_2, C_3, C_4 \cong 50 \text{ pF}$.
 9. Resistors:
 $R_{01}, R_{0p} \cong 200 \Omega$,
 $R_{02}, R_{03} \cong 10^3 \Omega$,
 $R_{s1}, R_{s2}, R_{s3} \cong 200 \Omega$,
 $R_b \cong 10^3 \Omega$.
 10. Thresholds*:
 $\Gamma_+ = V_r + \gamma_+ + [\text{closed SW2 offset}]$,
 $\Gamma_- = V_r - \gamma_- + [\text{closed SW2 offset}]$,
 $\gamma_+, \gamma_- = 345 \text{ mV}$.
 11. Voltages: $V_b = 10V, V_r = 4V, V_{cc} = 12V$
 12. Leakage currents: For e_1^*, e_p , and e_2^* nodes, current leakages are approximately 1 nA.
-

* The thresholds Γ_+ and Γ_- are offsets from the reference voltage $\{V_r + [\text{Closed SW2 Offset}]\}$ of the comparator input e_2^* . This is shown in Figs. 4, 5, and 6. The optimum values for the offsets, γ_+ and γ_- , are determined by the particular properties of the bubble memory family for which differential detection is considered. For the memory code under investigation here, the values selected were $\gamma_+ = \gamma_- = 345 \text{ mV}$.

the difference between bubble and no-bubble signal transitions in the DI and by intersymbol interference and noise. The effect of noise has been of secondary importance when DI is properly selected. Also, through proper MBM sensor design, ISI effects due to bubble coupling in the sensor array can be held to an acceptable fraction of the minimum isolated bubble and no-bubble difference. The signal, noise, and ISI are all functions of the MBM design.

Figure 6b shows the timing sequence of E_1 , E_2 , and E_3 relative to the response e_2^* out of the Restore Switch II for the sequence $\{b_n\}$ at an isolated 1 subpattern $\dots, 0, 0, 1, 0, 0, \dots$. This e_2^* waveform contains small voltage steps that coincide with the logic level shifts in E_1 , E_2 , and E_3 . These steps are not shown in the idealized e_2^* response in Fig. 5, but they can occur in a real circuit that has nonzero switch offset voltage or stray capacitance feedthrough of switch control logic into the signal path. Because the detector bases decisions on the difference $C_k = B_k - B_{k-1}$ each cycle, the effects of switch offsets and feedthroughs can be eliminated by equating the contributions from SW1 and SW3 while trimming thresholds Γ_+ and Γ_- during IC manufacture to account for SW2 offset.

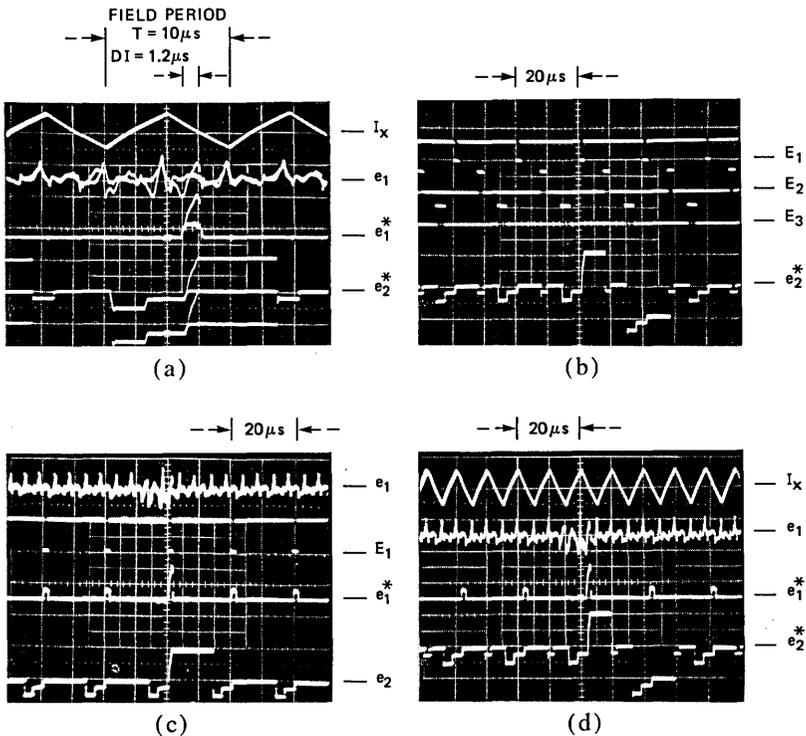


Fig. 6—Signal and control waveforms for differential detector.

Figure 6c shows the peak catcher response e_2 after amplification by A_3 for the $\{b_n\}$ subpattern $\dots, 0, 0, 1, 0, 0, \dots$. The trilevel response e_2^* in Fig. 6b corresponds to this e_2 pattern, and is obtained by clamping e_2 at SW_2 when E_2 is HI.

Figure 6d gives another comparison of the detector signals for the $\{b_n\}$ sequence $\dots, 0, 0, 1, 0, 0, \dots$. The test patterns in Fig. 6 were obtained for a bubble memory device where no-bubble cycles were interleaved with the data cycles. The interleaved no-bubble responses are present in the e_1 waveforms but have been suppressed in e_1^* , e_2 , and e_2^* .

V. CONCLUSIONS

We have described a magnetic bubble memory detector that will make correct decisions from memory sensor outputs even when large environmental or chip-to-chip signal transition variations occur in the detection interval (DI). The only constraints for successful detection are that for the selected functional B_k the maximum differences in $C_k = B_k - B_{k-1}$ for like symbols in contiguous data cycles are less than the minimum differences for unlike symbols. When the DI is properly selected to minimize interference, these constraints appear to be acceptable for typical outputs from a MBM design when the environmental changes in memory operation are slow relative to the field drive period.

At memory start up, the detector must have at least one cycle of output from the chip sensors in order to initialize. When input data $\{a_n\}$ is encoded to eliminate error propagation, then the recorded sequence is $\{b_n\}$, where $b_{-1} = 0$, and a single no-bubble sensor response is needed to initialize the detector. The comparison of sensor output to the uninitialized discharged reference condition in the detector circuit in the first read cycle can produce a single error if the first data bit in the $\{b_n\}$ data block is no-bubble. This is possible since a memory chip may have a no-bubble response B_k in excess of the optimum differential thresholds Γ_+ and Γ_- set in the difference detector. This detector-initializing no-bubble cycle is not a significant memory system constraint.

The detector circuit considered here incorporates peak sensing in a detection interval of the dc restored sensor response to generate the B_k each data cycle. Other forms of signal processing can be used to obtain B_k . In particular, if input noise is a problem in future MBM designs, the sensor outputs could be dc restored and integrated or matched filtered without dc restoration over the detection interval. But the resulting "1" and "0" responses for B_k must be sufficiently different to guarantee a detectable EYE. Also, if memory device design results in excessive intersymbol interference (ISI), signal equalization

can be considered to improve detection. At present, peak detection of bubble transitions with threshold comparison of $C_k = B_k - B_{k-1}$ seems to be adequate for differential detection of field-accessed MBM outputs. ISI equalization and special care in noise filtering have not been required to obtain a consistent differential detection probability of error below 10^{-8} for the memory types considered.

In light of the disadvantages frequently encountered with adaptive detectors and the advantages of a bubble memory detector that can accommodate a wide dynamic range for chip sensor response transitions, the differential detector described here has considerable merit. It is reasonable to integrate this detector as a monolithic bipolar IC such that the bubble memory detection step is accomplished by a single self-contained inexpensive DIP. Size and cost as well as the detector performance are very important memory system considerations, since even a modest bubble memory system can require many independent detector circuits.

REFERENCES

1. A. H. Bobeck and E. Della Torre, *Magnetic Bubbles*, Amsterdam: North-Holland, 1975, Section 4.5.
2. A. B. Smith, *Bubble-Domain Memory Devices*, Dedham, Ma.: Artech House, 1974, Chapter 8.
3. A. H. Eschenfelder, *Magnetic Bubble Technology*, New York: Springer-Verlag, 1980, Number 14 in Series in Solid State Sciences.
4. C. A. Belfiore and J. H. Park, Jr., "Decision Feedback Equalization," *Proc. IEEE*, 67, No. 8 (Aug. 1979), pp. 1143-56.
5. R. W. Lucky, "Techniques For Adaptive Equalization of Digital Communication," *B.S.T.J.*, 45 (Feb. 1966), pp. 255-86.
6. J. C. Hancock and P. A. Wintz, *Signal Detection Theory*, New York: McGraw-Hill, 1966, Section 6-6.
7. H. Kobayashi, "A Survey of Coding Schemes For Transmission Or Recording Of Digital Data," *IEEE Trans. Commun. Technol.*, COM-19, No. 6 (Dec. 1971), pp. 1087-1100.
8. A. Lender, "Correlative Level Coding For Binary Data Transmission," *IEEE Spectrum* (Feb. 1966), pp. 104-15.
9. J. Salz and B. R. Saltzberg, "Double Error Rates In Differentially Coherent Phase Systems," *IEEE Trans.*, CS-12 (June 1964), p. 202.

Subsampling of a DPCM Speech Channel to Provide Two “Self-Contained” Half-Rate Channels

By N. S. JAYANT

(Manuscript received November 12, 1980)

We consider the following channel-splitting problem: it is required to split a B -bits/s speech-code sequence into two “self-contained” $B/2$ -bits/s components, either of which can be used to reproduce acceptable speech; also, if both components are available at a receiver, it must be possible to reproduce speech with full B -bits/s quality. We propose a solution where for interesting values of B , the speech quality resulting from half-rate receptions approximately equals that from conventional full-rate receptions at $B/2$ bits/s. In the proposed solution, 3.2-kHz speech is sampled at 12 kHz and coded using PCM or differential PCM. The output sequence of code-words is split into odd- and even-word sequences. A full-rate receiver with access to both of the subchannels simply reconstitutes the output sequence prior to decoding, while a half-rate receiver with only the odd (or even) subchannel estimates the even (or odd) components by nearest-neighbor interpolation.

1. INTRODUCTION

The channel-splitting problem described in the abstract is redefined in Fig. 1. The receiving end of a speech communication system is supposed to operate in either a full-rate or half-rate mode depending on whether it has available to it both or only one of the speech subchannels. Respective qualities of speech reproduction are denoted by $Q_F(B)$ and $Q_H(B/2)$. The nonavailability of one subchannel is a good model for certain types of transmission failure, examples of which are signal fading in mobile radio and speech segment losses in packet switching. With appropriate forms of diversity reception, the second subchannel will be available with probability close to unity when the first subchannel is not. The channel-splitting problem has been re-

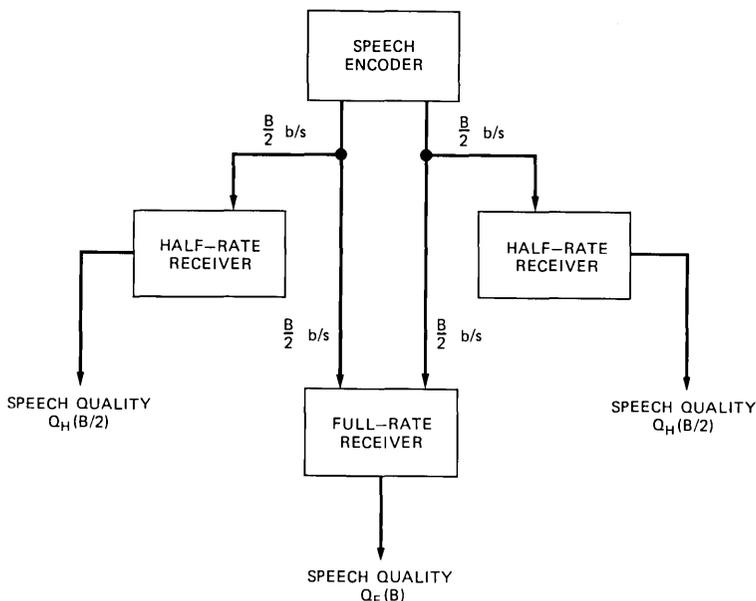


Fig. 1—Definition of the channel-splitting problem. When both of two half-rate components in the transmitted sequence are available at a receiver, conventional full-rate receptions result, with speech quality $Q_F(B)$. When only one of the components is present, a half-rate receiver recovers an approximation to the full-rate speech, with quality $Q_H(B/2)$.

cently analyzed for communication systems operating at rate-distortion limits with binary and Gaussian input signals.^{1,2}

The nontrivial nature of the channel-splitting problem can be appreciated through the simple example of a uniform quantizer. By combining two appropriately staggered R -bit quantizers (one of them a midrise, the other a midtread), one can realize an $(R + 1)$ -bit system, but not a $2R$ -bit system. For full-rate speech quality corresponding to 8-bit quantization, component quantizers would each need 7-bit (not $8/2 = 4$ bit) resolution for the combination to yield 8-bit quality. Thus, if the subchannels in Fig. 1 were simply uniform quantizers, and if speech were sampled at 8 kHz, one would need two 56-kbit/s quantizer systems so that a full-rate combination with 64-kbit/s quality can be realized. By contrast, in the differential pulse code modulation (DPCM) system proposed in this paper, the component receivers that combine to give 64-kbit/s quality are indeed half-bit rate, 32-kbit/s systems. Moreover, with an illustrative sentence-length speech input, the half-rate quality $Q_H(B/2)$ will be shown to exceed the full-rate quality $Q_F(B/2)$ of a conventional DPCM system operating at $B/2$ bit/s, for interesting values of B . The quality Q_F (32 kb/s) is quite acceptable

for speech communications, although somewhat short of toll quality. Note once again that two uniform quantizers operating at 32-kbit/s ($8 \text{ kHz} \times 4 \text{ bits}$) each can only give, in combination, 40-kbit/s ($8 \text{ kHz} \times 5 \text{ bits}$) quality, and not the desired 64-kbit/s quality. Similar arguments apply to lower bit rates as well.

The system of Fig. 1 is a special case of a communication scenario that can be generalized to include more than two subchannels, and/or subchannels that are non-equal-rate. The system of Fig. 1 can also be regarded as a special, symmetrical case of embedded coding,^{3,4} with a hierarchy consisting of two equally significant subcodes, viz., the half-rate sequences.

II. SUBSAMPLING AND INTERPOLATION

The utility of subsampling and interpolation has been demonstrated recently in the context of speech packet losses;⁵ speech-encoder outputs are partitioned into odd-sample and even-sample systems which are transmitted as separate packets. In the event of a lost odd (or even) packet, the lost samples are estimated using nearest-neighbor interpolations involving available even (or odd) samples. With the usual assumption of 3.2-kHz speech and 8-kHz sampling, the 1:2 subsampling (at 4 kHz) implies serious aliasing effects, but these errors are mitigated by an adaptive interpolation procedure where nearest-neighbor-weighting coefficients are varied to follow speech statistics, as reflected by appropriate extra information in packet headers.⁵ The system realizes dramatic improvements with packet loss probabilities up to about 10 percent; but as the component message loss probability approaches 100 percent, as in the channel-splitting problem, residual aliasing effects are quite unacceptable, even with adaptive interpolation.

The above observation has led us to the notion of 12-kHz sampling for the problem of Fig. 1. With 1:2 subsampling, the half-sampling rate will now be 6-kHz, which turns out to be just adequate for the 3.2-kHz speech inputs in telephony. We also considered 16-kHz sampling, but this is less preferable from the point of view of quantization noise. In 64-kbit/s decoding for example, 4-bit quantization of 16-kHz speech produces more quantization noise than the 5-bit quantizer that is possible with 12-kHz speech.

A second advantage of 12-kHz sampling is that it permits nonadaptive interpolations; adaptive interpolations yielded near-zero gains with 12-kHz speech. If 8-kHz is subsampled, it cannot be adequately reconstituted by nonadaptive interpolation even if the interpolation is invoked with a probability much less than 100 percent.⁵

Waveform reconstructions in the half-bit-rate (half-sample rate)

systems of this paper are described by

$$\hat{u}(n) = A_1 \cdot u(n-1) + A_2 \cdot u(n+1),$$

$$A_1 = A_2 = 0.5. \quad (1)$$

The samples $u(r)$ will be quantized DPCM prediction error samples in general; and more simply, they will be quantized speech samples in the special case of nonpredictive, or nondifferential PCM.

III. THE DPCM CODEC

Figure 2 shows block diagrams of full-rate and half-rate DPCM systems with fixed first-order predictors. In each case, decoding is defined by

$$y(n) |_{\text{DPCM}} = h_1 \cdot y(n-1) + q(n), \quad (2)$$

where $q(n)$ is the quantized prediction error signal, and h_1 is a first-order predictor. In the special case of nondifferential PCM ($h_1 = 0$), $q(n)$ is simply the quantized speech output:

$$y(n) |_{\text{PCM}} = q(n). \quad (3)$$

Subscripts H and F in Fig. 2 distinguish half- and full-rate versions of $y(n)$ and $q(n)$.

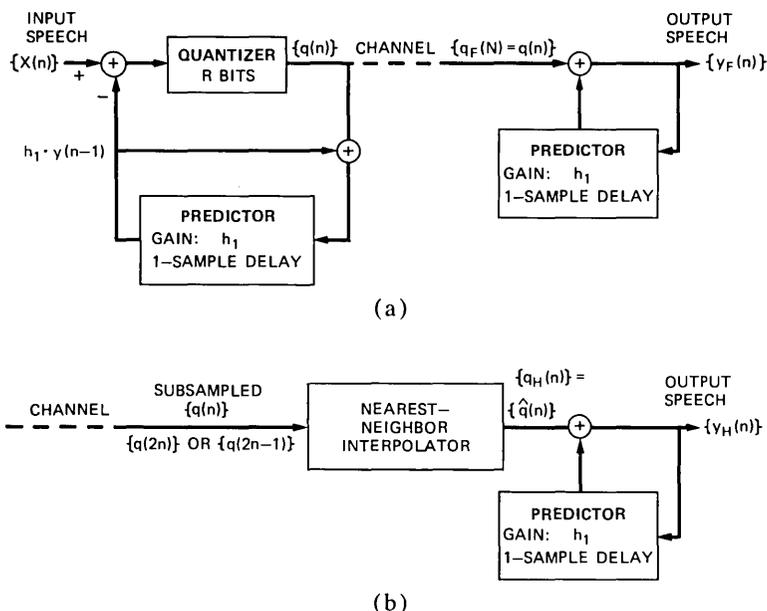


Fig. 2—DPCM block diagrams. (a) Conventional full-rate codec. (b) Decoder portion of half-rate receiver.

In contrast with well-known adaptive DPCM systems where the quantizer step size is adapted for every sample, the present paper assumes a system where the step size is adapted once for each block (of several ms duration), and held fixed for the duration of the block.⁶ A periodically updated, rather than instantaneously adaptive, quantizer is used in anticipation of interpolation procedures, which are known from recent experience to be unreliable when the step size exhibits sample-to-sample fluctuation.⁷

The periodically adaptive quantizers are defined by a block-specific step size Δ that is proportional to the root-mean-square value of a generalized first-difference in the form

$$\Delta |_{AQF} = K_1(R) \cdot [x(n) - h_1 \cdot x(n-1)]_{\text{rms}}, \quad (4a)$$

$$\Delta |_{AQB} = K_1(R) \cdot [y(n) - h_1 \cdot y(n-1)]_{\text{rms}}, \quad (4b)$$

with maximum and minimum constraints

$$\Delta_{\text{max}} = 128 \cdot \Delta_{\text{min}} = K_2(R) \cdot |x(n)|_{\text{max}}, \quad (4c)$$

where $[K_1(R), K_2(R)]$ are bit-rate-specific constraints with suggested values of $[0.25, 0.03]$, $[0.33, 0.06]$, $[0.58, 0.12]$ and $[1.0, 0.18]$ for 5, 4, 3 and 2-bit quantizers respectively. The subscripts *AQF* (4a) and *AQB* (4b) refer to forward-adaptive and backward-adaptive procedures; respective rms values in (4a) and (4b) are evaluated over the duration of a speech block to be coded in *AQF*, and over the duration of the most recent decoded speech block in *AQB*. The *AQB* procedure is less effective because of speech nonstationarity as well as the effect of quantization noise that is present in the $y(n)$ sequence used in (4b). However, step-size information in an *AQB* system need not be separately transmitted to a receiver; it is inherently available in the decoded $y(n)$ sequence. *AQF* procedures, by contrast, require the explicit transmission of step-size information (typically, about 5 bits worth, per block of 16 ms). In our experiments, quality losses in *AQB* were more noticeable in full-bit-rate speech than in half-bit-rate speech; and in each case the losses were of a second order of importance. With this in mind we have elected to cite only *AQF* results in section IV; these results can be regarded as upper bounds as far as quantizer performance is concerned.

In the context of 1:2 subsampling, the *AQB* procedure of (4b) cannot be implemented as such unless $h_1 = 0$ (PCM). However, step sizes obtained by setting h_1 to zero in (4) have been found to have fairly small effects on DPCM performance. Differences between PCM-optimal and DPCM-optimal step size are less significant than differences among step sizes of different speech blocks. Moreover, the suboptimality of PCM-matched step size becomes less significant as $h_1 \rightarrow 0$. The next

section shows that half-bit-rate DPCM favor values of h_1 that are indeed much smaller than those appropriate for conventional full-rate DPCM.

IV. RESULTS AND CONCLUSIONS

Figures 3 through 6 illustrate the performance of the interpolation procedure for the example of a 5-bit encoder. The waveform segments refer to two 20-ms blocks from a 3.2-kHz bandlimited, female utterance "The chairman cast three votes" sampled at 12 kHz.

Figure 3 shows full-rate and interpolated q waveforms ($h_1 = 0.5$) for the two segments; it demonstrates that the nonadaptive interpolator (1) is reasonably adequate even for the fast-varying unvoiced example. This is confirmed in Fig. 4 which shows corresponding full-rate and half-rate y waveforms. Notice that the half-bit-rate output provides a much better reproduction of the voiced segment, but is nevertheless reasonably effective in unvoiced speech reproduction. Perceptually,

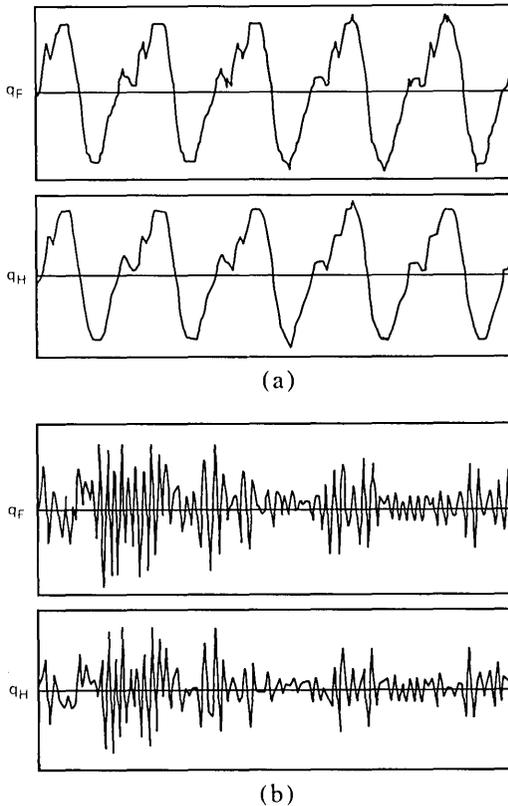


Fig. 3—Full-rate and half-rate (interpolated) sequences q_F and q_H ($R = 60$ kb/s, $h_1 = 0.5$) for (a) voiced speech segment and (b) unvoiced speech segment. The speech segments are 16 ms long.

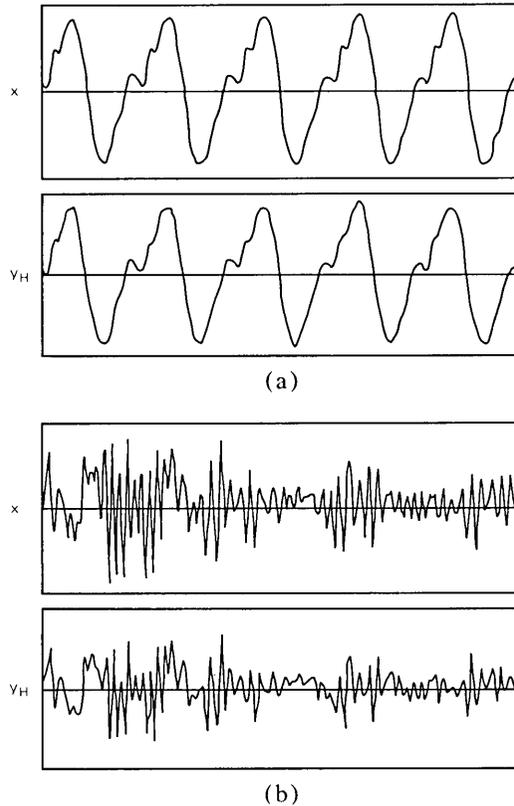


Fig. 4—Original and half-rate-decoded speech output $y(n)$ ($R = 60$ kb/s, $h_1 = 0.5$) for (a) voiced speech segment and (b) unvoiced speech segment. The speech segments are 16 ms long.

the waveform degradations in the half-bit coder are indeed fairly subtle for $R = 5$.

The significance of $h_1 = 0.5$ is demonstrated in Fig. 5, which shows DPCM performance as a function of predictor coefficient value. The objective quality measure used is the segmental signal-to-noise ratio SEGSNR defined as the average of $10 \log s/n$ (dB)-values measured over the totality of 20-ms blocks in the input. Notice that maximization of full-rate and half-rate quality call for $h_1 = 0.9$ and $h_1 = 0.5$, respectively; and notice also that these are not very sharp maxima, suggesting flexibility for practical implementations. The special, simple case of $h_1 = 0$ (PCM) leads to a noticeable quality degradation only for the full-rate system.

Figure 6 depicts the performance of full-rate and half-rate DPCM receivers as a function of encoder bit rate. The performance curves (i) and (ii) are for h_1 values that maximize half-rate speech quality (h_1

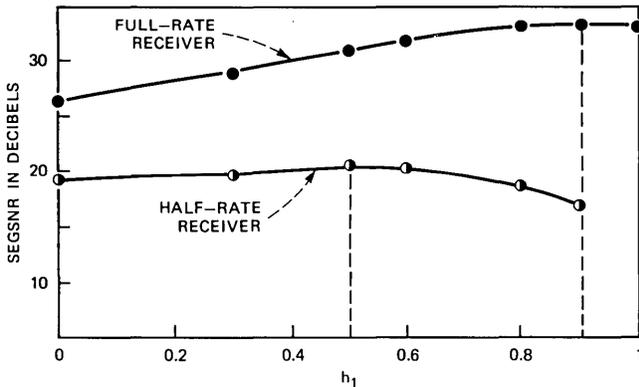


Fig. 5—Segmental SNR versus predictor coefficient h_1 ($R = 60$ kb/s). Values that maximize full-rate and half-rate quality are respectively 0.9 and 0.5.

$= 0.5$ for 5.4 bits/sample, $h_1 = 0.3$ for 3.2 bits/sample). The full-rate characteristic shows the expected 6-dB-per-bit behavior, while the half-bit-rate characteristic falls more gradually with decreasing R . Both characteristics tend to the expected 0-dB limit for no transmissions ($R \rightarrow 0$). The square dots in Fig. 6 represent the performance of a full-rate receiver in a system designed to maximize full-rate speech quality ($h_1 = 0.8$ to 0.9).

An important observation from Fig. 6 is that for encoder bit rates in the important range of 2 to 5 bits/sample,

$$Q_H(R/2) \approx Q_F(R/2); R < 5 \text{ bits/sample.} \quad (5)$$

This suggests that the half-bit-rate qualities in the subsample-inter-

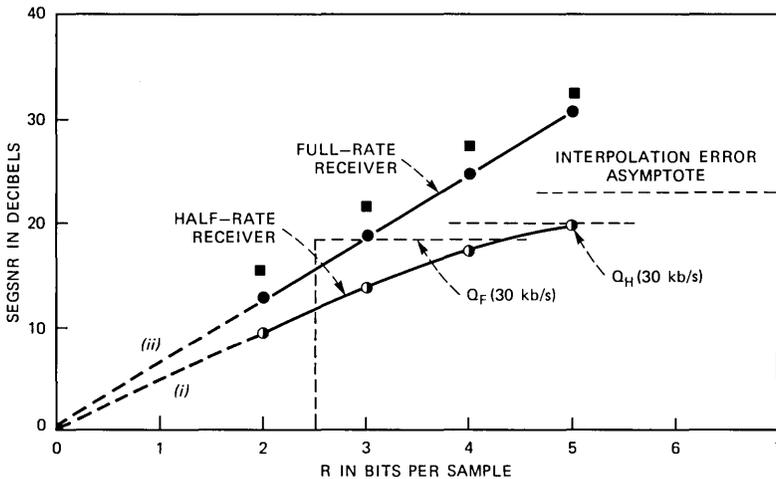


Fig. 6—Segmental SNR versus total transmitted bits/sample R . Curves (i) and (ii) refer to half-rate and full-rate codecs designed to maximize half-rate quality. The square dots represent a full-rate codec designed to maximize full-rate quality.

polate system are extremely good results, considering the crucial constraint that the half-bit-rate systems combine trivially to yield the full-rate performance $Q_F(R)$. The approximate equality in (5) is borne out very well in perceptual assessments of Q_H and Q_F . In contrast to (5), analytical results in Refs. 1 and 2 are quite pessimistic. This difference in conclusions is related to the fact that these analytical results apply at the rate-distortion limit, while the bit rates in this paper are nowhere close to the rate-distortion limit for speech. In fact, our bit rates are high enough that there is sufficient redundancy left in the coder output to permit subsampling and high-quality interpolations.

The relative performance of the half-bit-rate receiver diminishes with increasing bit rate. Clearly, as $R \rightarrow \infty$, the quantization noise contributions to Q vanish, $Q_F \rightarrow \infty$, and Q_H tends to a finite asymptotic value that shows the effect of nearest-neighbor interpolation noise. Results elsewhere⁵ can be used to show that this asymptotic value, for a first-order Markov signal example, is approximately given by the expected value of $10 \log [(1 + R_{xx}^2(1))/(1 - R_{xx}^2(1))] dB$, where $R_{xx}(1)$ is a block-specific adjacent sample correlation in the speech input $x(n)$.

Finally it would be appropriate to calibrate the Q_F and Q_H values in Fig. 6 with well-known definitions of toll-quality and communications quality (near perfect intelligibility with noticeable but not obtrusive degradation).⁸ Although simulations and conclusions have centered on a single earlier-cited test input, it appears that full-bit-rate DPCM realizes toll-quality for $R = 5$ and 4 bits/sample; and communications quality for $R = 3$ and 2 bits/sample. The half rate DPCM receptions in the proposed system approach toll quality with $R = 5$ bits/sample and maintain good communications quality at $R = 4$ and 3 bits/sample.

REFERENCES

1. J. K. Wolf, A. D. Wyner, and J. Ziv, "Source Coding for Multiple Descriptions," B.S.T.J., 59 (October 1980), pp. 1417-26.
2. L. Ozarow, "On a Source Coding Problem With Two Channels and Three Receivers," B.S.T.J., 59 (December 1980), pp. 1909-21.
3. D. J. Goodman, "Embedded DPCM for Variable Bit Rate Transmission," IEEE Trans. Commun., COM-28, No. 7 (July 1980), pp. 1040-46.
4. T. Bially, B. Gold, and S. Seneff, "A Technique for Adaptive Voice Flow Control in Integrated Packet Networks," IEEE Trans. Communications, COM-28, No. 3 (March 1980) pp. 325-33.
5. N. S. Jayant and S. W. Christensen, "Effects of Packet Losses in Waveform-Coded Speech and Improvements due to an Odd-Even Sample Interpolation Procedure," IEEE Trans. Commun., COM-29 (February 1981).
6. N. S. Jayant, "Step-Size Transmitting Differential Coders for Mobile Telephony," B.S.T.J., 54 (November 1975), pp. 1557-81.
7. N. S. Jayant et al., "On Soft Decision Demodulation for PCM and DPCM Encoded Speech," IEEE Trans. Commun., COM-28, No. 3 (March 1980), pp. 334-44.
8. J. L. Flanagan et al., "Speech Coding," IEEE Trans. Commun., COM-27 (April 1979), pp. 710-37.

Blocking States in Connecting Networks Made of Square Switches Arranged in Stages

By V. E. BENEŠ

(Manuscript received October 16, 1980)

Since the probability of blocking is a principal measure of the performance of a network, this paper examines the blocking states of a network. For two-sided connecting networks made of square switches arranged in stages, a parallel pair (PP) is a pair of paths through the network that meet at no switch. The blocking states of the network are closely related to those that have a busy PP. In particular, a routing algorithm can avoid all the blocking states if and only if it avoids all the states with busy PPs.

I. INTRODUCTION

Connecting networks made of rectangular switches have long been an important feature of telephone central offices, and their combinatorial properties still command substantial practical as well as theoretical interest. The probability of blocking, calculated or measured in a suitable setting, is a principal measure of the performance of a network, and it is of particular interest to identify and study those combinatorial features which give rise to blocked calls.

To be sure, blocking is known to be an unfortunate mismatch of available resources. But precisely, what structural features lead to it? What presages blocking? What are some tell-tale signs and necessary concomitants of blocking? When and how can blocking be avoided? To try to answer these questions, we focus on the blocking states of the network, those in which some call is blocked.

Although in a century of telephony much work has gone into calculations and simulations for blocking probability, there are (except for a few examples) virtually no results on the basic causes and nature of blocking. For networks made of square switches arranged in stages, this paper will partly remedy this lack by introducing some simple concepts suggested by examples; theorems based on these bear on the position and structure of the blocking states within the set (semilattice)

of states, and also on the possibility of routing rules that avoid blocking entirely.

II. PRELIMINARIES

In a network made of square switches each rectangular switch has the same number of inlets as outlets, so there is neither expansion nor concentration. The network is said to be arranged in stages if the switches are partitioned into sets called stages, two of which carry the inlets and outlets, while the others are internal stages, such that every path from an inlet to an outlet passes through each stage only once, always in the same order. The network is full access if each switch in a given stage has a link to every switch in the adjacent stages.

It has been pointed out repeatedly^{1,2,6} that the set S of permitted states of a connecting network forms a partially ordered system, in fact a semilattice, in which the order relation $x \leq y$ between states x , y means that x can be reached from y by zero or more hangups: In this case we say that x is below y . The usual Hasse diagram for the partial ordering is then a convenient way of visualizing the set S of states, marred only by the prodigious number of states that arise. This "curse of dimensionality" can be alleviated by considering only equivalence classes of states under a suitable group of symmetries,² and we do so when convenient without further discussion.

A state $x \in S$ is maximal if there are no states above x . It can be seen that in a network made of square switches all maximal states carry the same number of calls. A state x is a blocking state (or simply, is blocking) if an inlet u and an outlet v exist, both idle in x , such that there is no available allowed path from u to v , whose use would connect u to v and give rise to a state y covering (immediately above) x in the partial ordering; the call (u, v) is then said to be blocked in x . We use the convention that, in speaking of a call (u, v) , u is always an inlet, and v always an outlet. Figure 1 illustrates all the preceding concepts.

A routing algorithm or rule is a method for choosing routes for unblocked calls during the operation of the network to carry traffic. It specifies, for each state x and each call (u, v) not blocked in x , what route should be used to put up (u, v) if these terminals desired connection when the system was in state x . Needless to say, it is of interest to find and describe rules which make for good network performance as evaluated by the probability that an arriving call is blocked, i.e., the fraction of requests for service which cannot be completed because of blocking. Since such denials of service occur only when the network is in a blocking state, it is natural that a good routing strategy will seek to avoid the blocking states. Examples are known in which the blocking states can be avoided entirely; such networks are called nonblocking in the wide sense.^{1,3}

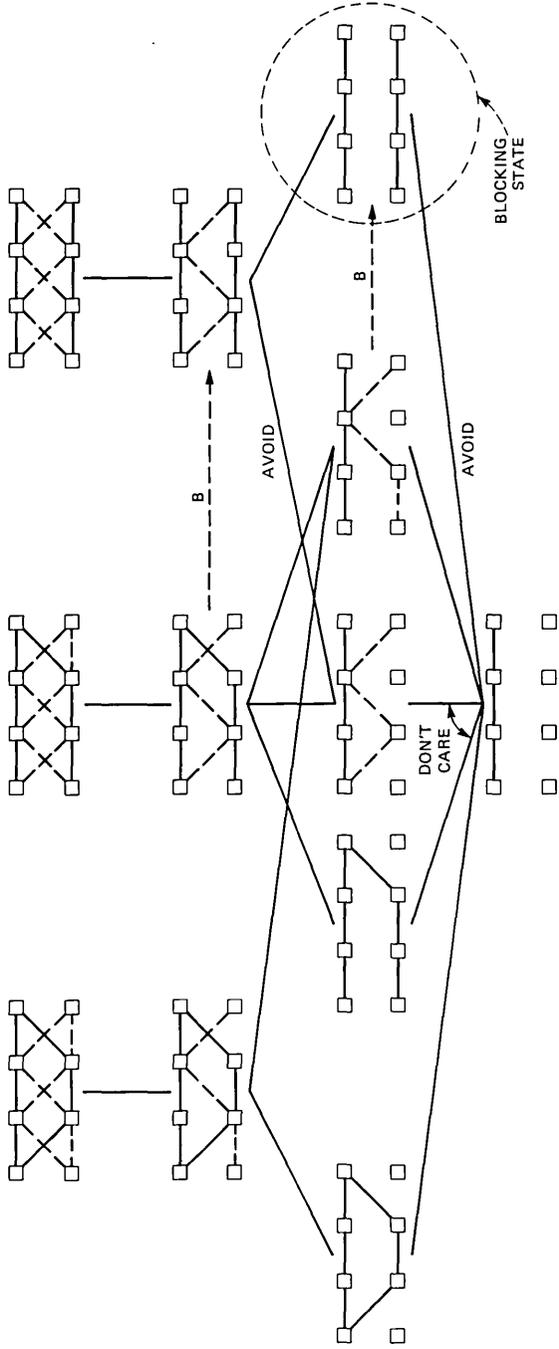


Fig. 1—Reduced state diagram.

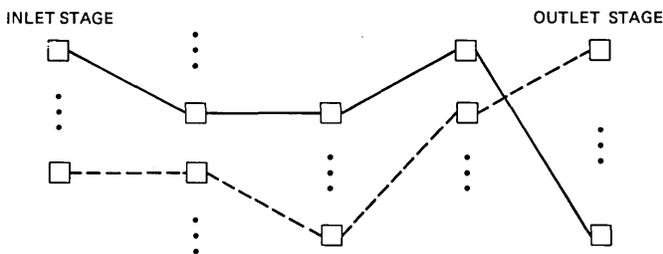


Fig. 2—A parallel pair (PP).

III. SUMMARY OF RESULTS

We assume as given a two-sided connecting network made of square switches arranged in stages, and we introduce the idea of a parallel pair (PP). This is a pair of paths through the network that meet at no switch (Figs. 2, 3). The basic message of this paper is that visiting states with busy PPs is bad for performance, because the property of being a blocking state is closely tied up with that of having a busy PP. Now having a blocked call and having a busy PP are usually not equivalent properties of states; the characterization of blocking is not that simple. There are nonblocking states which have a busy PP, and there are blocking states without a busy PP. Nevertheless, our intuitive conclusion is strongly supported by small examples,^{3,4} and here by the following results:

- (i) Every maximal state that is above a blocking state has a busy PP.
- (ii) Every blocking state is below a maximal state with a busy PP.
- (iii) A routing algorithm makes the network nonblocking in the wide sense if and only if it makes inaccessible all states with busy PPs.

It can be seen that in a state without any busy PPs, every path in use must meet every other at a switch. To maintain this property in an operating system, one must keep from adding paths that pass only through empty switches. Thus, the advice to avoid busy PPs almost coincides with (or at least, is very consistent with) the “packing” principle that to avoid blocking one should route calls in the most

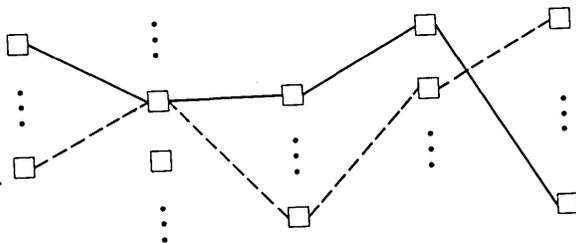


Fig. 3—Paths that meet at the second stage.

heavily loaded part of the network that will accept them: Clearly, in putting up a new call you best avoid busy PPs by having the route you choose meet as many calls in progress as possible. Our result says that completely avoiding blocking and never having any busy PPs are both possible, or both impossible, to achieve by a choice of routing rule.

IV. SOME EASY RESULTS

To fix ideas and record a few facts, we first list various results with straightforward proofs. The reader interested in warming up to the concepts in use here might go through these in some detail using Figs. 1 to 5.

(1) *Remark:* No state immediately below a maximal state is blocking.

(2) *Remark:* A blocking state two levels down from a maximal state (i.e., obtainable from a maximal state by two hangups) consists of an idle PP with all other terminals and links busy.

(3) *Lemma:* If x is a blocking state, then some PP is idle in x .

Proof: Let (u, v) be blocked in x . Since the switches are square, there is an idle path from u to the other side, and one from v to u 's side of the network. These paths must avoid each other if (u, v) is blocked. Thus, they constitute a PP idle in x .

(4) *Remark:* If x has a busy PP, then states y and z exist such that y is maximal, z is blocking, $y \geq x$, and $y \geq z$.

Proof: Starting in state x , put calls in until a maximal y is reached, then hang up the PP to form z , and reverse its assignment, i.e., if the busy PP provided the two calls (u_1, v_1) and (u_2, v_2) , then the calls (u_1, v_2) and (u_2, v_1) are both blocked in z . Compare Remark (2).

(5) *Remark:* If v is full access and x is blocking, then all inner switches are nonempty in x .

Proof: An empty switch is accessible from any switch in the adjacent stages, hence from any idle inlet, and also from any idle outlet. Thus, no call is blocked. See Fig. 4.

(6) *Remark:* x nonblocking $\not\Rightarrow x$ has no busy PPs.

Proof: Figure 4 provides a counterexample.

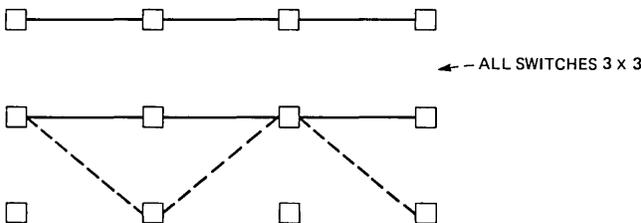


Fig. 4—State with two busy PPs but no blocking.

(7) *Remark:* x blocking $\not\Rightarrow x$ has a busy PP.

Proof: Figure 5 gives a counterexample.

V. BLOCKING STATES ARE BELOW STATES WITH BUSY PPS

(8) *Theorem:* A maximal state with no busy PPs has no blocking states under it. Every maximal state that is above a blocking state has a busy PP.

Proof: We embroider the argument in (3). Suppose that x is a blocking state, and $y \geq x$ is maximal. Let (u, v) be an inlet-outlet pair blocked in x ; u and v are on disjoint paths of a PP idle in x . Moreover, since $y \geq x$, it is possible to start with x and put in all the additional calls of y , by the same routes as they use in y , except those involving the terminal pair (u, v) . (This is *not* to say that a given routing algorithm would prescribe those routes in operation!) Since (u, v) is blocked in x , u and v are not talking to each other in y . Indeed, in y , u is connected to some switch other than v 's, and v to some switch other than u 's. If the paths carrying these calls in y met at some switch, (u, v) would not have been blocked in x . Hence, in the state formed from x by putting in all the calls of y except those involving u and v , these paths again form an idle PP with u on one path of the PP, and v on the other. Thus, y necessarily has a busy PP.

(9) *Remark:* We have already seen in (4) that a maximal state with a busy PP is above some blocking state. Thus, for maximal states y ,

y has a busy PP iff y lies above a blocking state.

(10) *Corollary:* Any blocking state lies below a maximal state which has a busy PP.

Proof: Every state x lies below some maximal y . If y had no busy PPs, no blocking state could be below it, by (8).

(11) *Corollary:* If x is below some maximal state which has no busy PPs, then x is nonblocking.

Proof: (8) or (9); the property is sufficient, but not necessary.

VI. THE CLOSURE TOPOLOGIES INDUCED BY \leq

A partial ordering \leq induces a topology in a familiar way. If $X \subseteq S$ is a subset of states, its lower closure, denoted X_{lc} , is the set of states

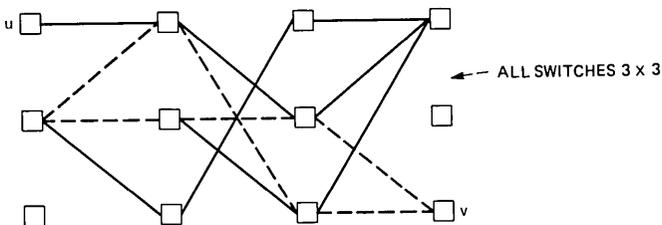


Fig. 5—Blocking state with no busy PPs: (u, v) is blocked.

that are below some member of X in the sense of \leq , including, of course, the elements of X as well. Lower closure satisfies the Kuratowski axioms for a topological closure operation. There is a dual topology of upper closure, wherein the dual \geq replaces \leq , denoted by X^{uc} .

By introducing notation for some of the important subsets of states we have encountered so far, we can use lower closure to express various preceding results in a concise way. Let us define

- Max = set of states that are maximal in \leq ,
- B = set of blocking states,
- $S - B$ = set of nonblocking states,
- bPP = set of states that have a busy parallel pair.

(12) *Remark:* $(bPP)_{lc} = (\text{Max} \cap bPP)_{lc}$.

Proof: If $x \in (bPP)_{lc}$, then there exists y with a busy PP such that $x \leq y$. Pick a maximal $z \geq y$. Then $z \in bPP$, since bPP is closed above. Thus, we have $z \geq x$ and $z \in \text{Max} \cap bPP$, and so also $x \in (\text{Max} \cap bPP)_{lc}$. The converse is trivial.

In these notations, Theorem (8) may be cast as saying that

(13) $B \cap (\text{Max} \cap (S - bPP))_{lc} = \emptyset = \text{empty set}$,

while Corollary (10) claims that

$$B \subseteq (\text{Max} \cap bPP)_{lc},$$

which by (12) simplifies to $B \subseteq (bPP)_{lc}$. Similarly Corollary (11) can be rendered as

$$[\text{Max} \cap (S - bPP)]_{lc} \subseteq S - B.$$

VII. WIDE-SENSE NONBLOCKING NETWORKS

A telephone network is called nonblocking in the wide sense (NBWS) if a rule exists for assigning routes to incoming calls that precludes all blocking, that is, under the rule no blocking state is accessible. Run according to the rule, the system can satisfy all demands for connection as they arise without ever rejecting any, or rerouting existing calls. The point is to try to reduce network complexity (from the extreme of strict nonblocking) by using a network that has blocking states, but to avoid them by clever routing.

Two necessary and sufficient conditions are known for a general connecting network to be NBWS: One is phrased in terms of denseness in a simple metric on the states,⁵ the other in terms of a semilattice homomorphism of S onto the set of assignments realized by the states in S .⁶ For networks made of square switches arranged in stages, we give a new necessary and sufficient condition, based, perhaps not surprisingly now, on avoiding states with busy PPs.

If α is a routing algorithm, we use $Ac(\alpha)$ to denote the set of states that are accessible under α , i.e., that can be reached by the operating

network if it is run using α to make routing choices. A good routing algorithm derives its worth from making bad states unlikely or inaccessible, so it is important to know what states can be reached. Further, it is clear that an NBWS network is one for which an α can be found such that $Ac(\alpha) \cap B$ is empty, i.e., no blocking state can be reached by operating under α . We note these properties of $Ac(\alpha)$:

(14) $Ac(\alpha)$ is closed below: $[Ac(\alpha)]_c = Ac(\alpha)$,

(15) $x \in Ac(\alpha) \Rightarrow Ac(\alpha) \cap \text{Max} \cap \{x\}^{uc} \neq \emptyset$.

Property (15) says that if x is accessible, then some maximal state above x is accessible. For networks made of square switches, however, a specialized sharp form of (15) holds, in which both hypothesis and conclusion are stronger, and which relates blocking to having a busy PP. To be specific, we shall use a precise version of the following heuristic idea: An algorithm α must complete an unblocked call by *some* route; so starting in a state x that has a blocked call (u, v) , it should be possible to keep picking successive unblocked new calls not involving u or v , and routing them by α , until as many terminals are busy as possible, and the system has arrived at a state z accessible from x under α by new calls, with (u, v) idle and blocked in z ; we claim in particular that the calls can be picked so that z consists of an idle PP which has u and v on switches that terminate its paths, and all other terminals are busy. Compare Remark (2).

(16) *Lemma: If (u, v) is a call blocked in state x , then an inlet s and an outlet t exist, with $s \neq u$ and $t \neq v$ such that EITHER*

(i) s, t, u , and v are the only terminals idle in x , and there is a PP idle in x one of whose paths joins s to v , the other u to t , OR

(ii) an inlet q and an outlet r exist, neither of which coincides with any of s, t, u , or v , such that q and r are both idle in x and (q, r) is not blocked.

Proof: We recall our convention that in a call (u, v) , u is always an inlet and v an outlet. Since (u, v) is blocked, there is a PP idle in x one of whose paths connects u 's switch to some switch other than v 's, which, since the switches are square, has an idle outlet t on it. Similarly, the other path of the PP connects v 's switch to some switch other than u 's on which there is an idle inlet s . Suppose that alternative (i) fails, so that some inlet q , distinct from each of s and u , is idle in x . Now since switches are square, q can reach an idle outlet r ; we shall show that r can be chosen to be distinct from both t and v . There are three cases.

Case (i): u and q are on the same switch. Since u can reach t , q must be able to reach some r other than t . This is because if we put up (u, t) in x , q could still reach some idle outlet, so *a fortiori* it can do so in x . Since (u, v) is blocked, so is (q, v) . But (q, r) is not blocked, so $r \neq v$. Thus, r is distinct from both t and v , though it may be on the same switch as t . So alternative (ii) holds in Case (i).

Case (ii): s and q are on the same switch. Then q can reach v 's switch and possibly t 's. If it can reach v 's switch *only*, then an idle outlet r must exist on v 's switch, $r \neq v$. This is because if an (s, v) call were put up, q could still reach some idle outlet, which must be on v 's switch, since adding a call does not make additional outlet switches reachable by q . Thus, $r \neq t$ and $r \neq v$, (q, r) is not blocked, and alternative (ii) holds. Suppose then that q can reach t 's switch as well as v 's. There are two subcases: (a) t 's switch has another idle outlet r distinct from t , or (b) t is the only idle terminal on t 's switch. In case (a), $r \neq v$, $r \neq t$, and (q, r) is not blocked, so we are done. In case (b), since both u and s can reach t , the final arc in the paths from u to t and s to t must be the same. Thus, these paths meet at an inner switch, and indeed on the way through the stages from inlets to outlets they meet for the last time at an inner switch σ . Thus, σ has at least two idle terminals on each side, and there is an idle path from σ to an outlet switch other than t 's, which is also reachable by u , because the paths meet. This outlet switch is therefore distinct from v 's because (u, v) is blocked; it has an idle outlet r on it distinct from each of t and v , and (q, r) is not blocked. Thus, again alternative (ii) holds.

Case (iii): q is on neither u 's switch nor s 's. If q can reach any switch other than t 's or v 's, we are done; for then there is an idle outlet r , distinct from each of t and v , and (q, r) is not blocked. The same is true if q can reach only t 's or v 's switch provided the one it can reach has more than one idle outlet. Suppose then that q can reach only t 's switch or v 's, and that whichever ones of these it can reach have exactly one idle outlet, viz, t or v as the case may be.

If q can reach t but not v , then (q, v) is blocked and moreover there is an inner switch at which q 's idle path to t meets one from u to t for the last time (indeed they coincide henceforth), and there is a path, subsequently disjoint from the remaining common path to t , from that inner switch to some outlet switch other than t 's or v 's, because (q, v) is blocked. On that outlet switch is the r we seek for alternative (ii). If q can reach v but not t , the same conclusion follows, with s playing the role of u , *mutatis mutandis*.

If q can reach both t and v , we argue similarly, as follows: Since t 's switch has exactly one idle link, an idle path from u to t and one from q to t must have that last idle link in common, there being no other way to reach t 's switch. Hence, these paths must meet at an inner switch at least once. As before there is then a last, in the ordering of stages from the inlet side to the outlet side, switch at which these paths meet, and they coincide thereafter. Since the paths are meeting, the switch has at least two idle links incoming from the inlet side, and at least two idle links outgoing on the outlet side. One of the latter is a part of the now common part of the paths on their remaining legs to t 's switch. The other idle link on the outlet side is part of an idle path

to the outlets; it does not meet the common part again, and so cannot end at t 's switch. If it reached v 's, (u, v) would not be blocked. The outlet switch it reaches then has an idle outlet r on it, distinct from each of t and v , and (q, r) is not blocked; hence alternative (ii).

(17) *Lemma: If x is a blocking state accessible under a routing algorithm α , then some maximal state $y \geq x$ is accessible under α , and y has a busy PP. In symbols,*

$$x \in B \cap Ac(\alpha) \rightarrow \exists y y \geq x, \quad y \in \text{Max} \cap Ac(\alpha) \cap bPP.$$

Proof: Let (u, v) be a call blocked in an accessible state x . If possible, use alternative (ii) of Lemma (16) to find a call (q, r) , unblocked in x , neither of whose terminals is on an idle PP with u and v . Putting up (q, r) according to α results in a state in which (u, v) is blocked. Add calls in this way repeatedly until no such (q, r) can be found, i.e., alternative (ii) fails. Then alternative (i) holds: The state z we have reached consists of an idle PP with one of u, v on each of its distinct paths. It is obvious that there is one maximal y above z , reachable under α by having the right two calls request connection, and that y has a busy PP.

(18) *Theorem: Let α be a routing algorithm. Some blocking state is accessible under α iff some state with a busy PP is accessible under α ; in symbols,*

$$Ac(\alpha) \cap B \neq \emptyset \quad \text{iff} \quad Ac(\alpha) \cap bPP \neq \emptyset.$$

Proof: Suppose $x \in Ac(\alpha) \cap bPP$. By (15) there exists $y \geq x$ such that y is maximal and $y \in Ac(\alpha)$. Since $y \geq x$ and bPP is upper closed, we see that y has a busy PP. By (9) there is a $z \in B$ such that $z \leq y$ and $z \in Ac(\alpha)$, since $Ac(\alpha)$ is closed below. Hence, $Ac(\alpha) \cap B \neq \emptyset$. The converse follows from (17).

We illustrate two aspects of Theorem (18) by reference to Figs. 4 and 5. According to the converse part of (18), if an algorithm α made the state in Fig. 4 reachable, then there must be a continuation (of the events that led to this state) leading to a blocking state. In Fig. 4 this happens if there is, first, a call between the two highest outer switches, and then, a call between the two second-highest outer switches. There are two ways of satisfying this extended assignment, and both lead to blocking. The direct part of (18) is exemplified in Fig. 5. If an algorithm α makes the state in Fig. 5 reachable, then according to (18) a state with a busy PP is also reachable. It can be verified that in Fig. 5 putting up any new call leads at once to a state with a busy PP.

(19) *Corollary: A routing algorithm α makes the network nonblocking in the wide sense if and only if no state with a busy PP is accessible under α .*

An example in which it actually happens that a natural routing rule makes all blocking states unreachable was shown in Fig. 1. The relation

B shown in Fig. 1 by arrows is to be interpreted as "better than." Any algorithm for routing which uses the "better" alternatives avoids all blocking and all busy PPs. Note that there are actually two such algorithms, as suggested by the "don't care" label.

REFERENCES

1. V. E. Beneš, *Mathematical Theory of Connecting Networks and Telephone Traffic*, New York: Academic Press, 1965, p. 62 and in references therein.
2. V. E. Beneš, "Reduction of Network States Under Symmetries," *B.S.T.J.*, 57, No. 1 (January 1978), pp. 111-49.
3. V. E. Beneš and R. P. Kurshan, "Wide-Sense Nonblocking Network Made of Square Switches," unpublished manuscript.
4. V. E. Beneš, "Programming and Control Problems Arising From Optimal Routing in Telephone Network," *B.S.T.J.*, 45, No. 9 (November 1966), pp. 1373-438.
5. V. E. Beneš, "Algebraic and Topological Properties of Connecting Networks," *B.S.T.J.*, 41, No. 4 (July 1962), pp. 1249-74.
6. V. E. Beneš, "Semilattice Characterization of Nonblocking Networks," *B.S.T.J.*, 52, No. 5 (May-June 1973), pp. 697-706.



Optical Fibers for Scanning Digitizers

By W. PFERD and G. C. STOCKER

(Manuscript received August 11, 1980)

Automated graphic activities such as computer cartography and computer-aided design require document-scanning systems that are stable, fast, and capable of handling large drawings and maps. In such a system the scanning assembly should accept documents readily and be compact and inexpensive enough to encourage wide usage. An attractive method for such automatic scanning and digitizing of engineering drawings combines optical fibers with a solid-state imaging array and a stepped friction feed. In this paper we describe an assembly of these elements employing a light-guide illuminator and receiver that are in direct contact with the drawing. Light signals, modulated by the dark areas on the drawing, are transmitted from the linear end of the lightguide to a grid configuration at the opposite end. The signals are optically coupled to and detected by a solid-state array sensor that scans the grid of optical fibers. The dimensional control necessary for fabricating the fiber grid to match the configuration of the solid-state array is critically important for the operation of this type of scanner. We therefore emphasize in this paper the dimensional requirements for the glass fiber assembly, the signals detection, and the accuracy of scanning.

I. INTRODUCTION

Document scanning systems for automated activities such as electronic mail, computer cartography, and computer graphics require an image-guiding system to capture visual details for conversion to electronic signals. The image arrangement is combined closely with, and influences the design of, the document-positioning mechanism. The document can be illuminated by diffused light in the general scanning area, the bounced light optically projected through a tiny aperture onto an appropriate sensor (flying aperture scanners), or by spot lighting where a tiny spot of light is projected onto the copy and the bounced spot is picked up directly by the sensor (flying spot scanners).

The surface condition alters the light that is reflected to the photoelectric detector during the scanning. Documents are scanned electronically, mechanically, or in combination, with the portion of the document being scanned mounted flat or in a circular arc.

The physical dimensions of scanning assemblies are controlled by the size of the document and the light paths required to scan the surface. In most assemblies, lenses and mirrors are combined with rotating or translating document holders.¹ In more recent assemblies, stationary copy is scanned by a flying spot from a cathode-ray tube,¹ a charge-coupled photodiode array,² or an acousto-optically modulated laser.^{3,4} Optical fibers are also used to pipe light from the document to a rotary scan head or solid state image sensor.⁵ In this paper, we describe the performance characteristics of a flying-spot type of optical fiber scanner that has inherent advantages for digitizing large documents. This paper gives special attention to glass fiber characteristics required to construct this new potentially low-cost scanning digitizer.

II. OPTICAL FIBER SCANNER

A digitizer must scan the document to resolve the line and character information into small picture cells or pixels. The optical response at each pixel is then used to control detection circuitry to generate positionally defined digital output. In general, the drawing can be scanned in one of three ways: individual pixels can be addressed in a series of parallel sweeps across the drawing surface; groups of pixels can be addressed by a sensor array that is stepped over the drawing; pixels can be addressed sequentially in parallel sweeps to locate lines and marks on the drawing that are subsequently followed or tracked. In all methods, some form of parallel scanning must first be made to locate the line information to a prescribed resolution. As will be shown for scanning large documents, optical fibers, suitably arranged for a raster output (method 1), offer high resolution, geometric fidelity, stability, and mechanical simplicity—all in a very compact assembly.

Figure 1 shows the schematic representation of an optical fiber scanning digitizer. Light is transmitted from the source via a multiplicity of lightguides to uniformly illuminate a narrow strip across the drawing. The fibers at the light source are randomized in a compact circular group and are illuminated with a small lamp. The receiver consists of a similar number of optical fibers as the illuminator. These are oriented in a line at one end and as a matrix or grid at the other end. The linear end of the illuminator and receiver are geometrically similar and located adjacent to each other in the fixed scanner head. Drawings are positioned directly against the fixed head by a spring-loaded pressure plate. The drawing is moved past the read head in small increments by rubber-coated friction rollers driven by a preci-

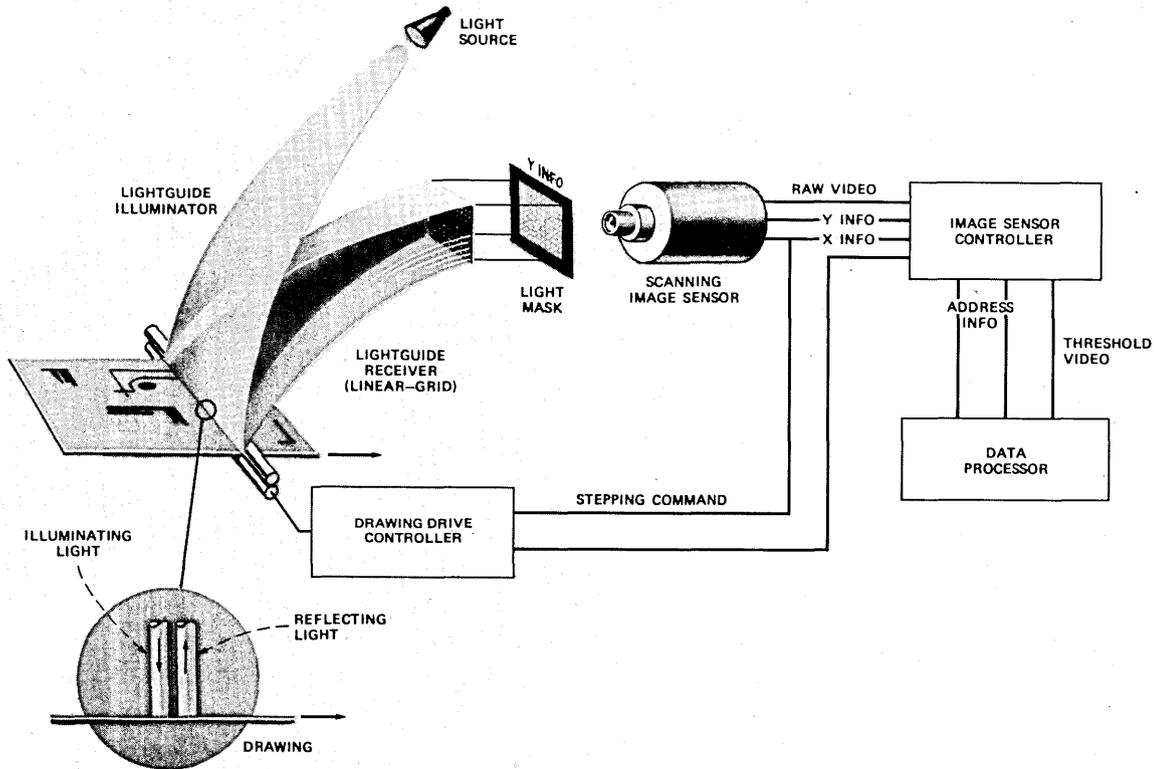


Fig. 1—Schematic representation of an optical fiber scanning digitizer.

sion-stepping motor. The stepping motor, operating at a few milliseconds per incremental movement, is driven by a sensor controller. The paper is moved when the controller signals that it has completed a linear scan of the document. The diffused and reflected light signals, attenuated by dark lines, characters, etc., on the drawing, are transmitted through the lightguide receiver and detected by an array of photosensitive elements that scan the lightguides. A one-to-one correspondence exists between the specific transverse location on the drawing, a receiver optical fiber, and an element in the solid state array.

A light mask and lenses are interposed between the receiver grid and the image sensor array. The light mask, prepared by a photolithographic process, is a metalized glass plate with an equal number of 1-mil-diameter holes as receiver fibers. The light-mask holes are located on the nominal centers of the fibers. The mask corrects for slight misalignments in the fiber grid and presents a dimensionally precise pattern to the array sensor. The lenses focus the light from each fiber (through the light-mask holes) onto the individual sensor elements (pixels). Gray scale or thresholded video output from each pixel is used to represent the digitized signals. These signals can be processed further⁶ to provide information that is suitable for reproduction, computer editing,⁷ or graphics teleconferencing.⁸

Table I gives a list of the important characteristics of an optical fiber scanning digitizer. Figure 2 shows the assembly used to scan digitize by optical fibers, which provided the experimental data described in this paper.

III. FIBER ASSEMBLY DESIGN

In scanning digitizers utilizing optical fibers, fabricating the fiber grid so that its tolerances match the location of the active pixels in solid state array sensors is critically important. While in communication applications fiber diameter is an important consideration regarding transmission parameters, in scanning digitizers by optical fibers it is crucial that the fiber grid, hole mask, and solid state array are correctly aligned. The degree of precision that can be achieved in

Table I—Characteristics of an optical fiber scanning digitizer

Type	Optical fiber/ccd array/roll feed
Copy Width	Up to 40 in.
Copy Length	Unlimited
Copy Feed	Variable, 1 ms/0.004 in. maximum
Scan Resolution	Variable, dependent on fiber diameter and spacing
Maximum Pixel Rate	10 mH (present maximum)
Output	Gray scale or threshold video
Size	Desk top (10 in. high, 20 in. deep, and 50 in. wide)

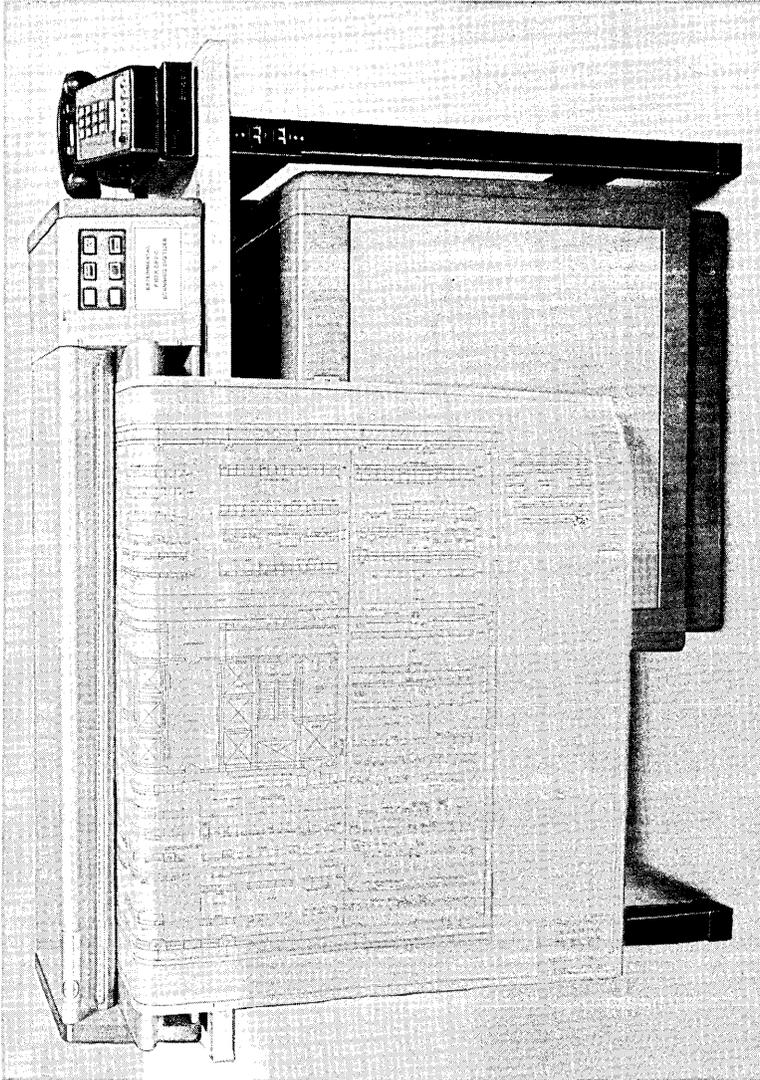


Fig. 2—Feasibility model of an optical fiber scanning digitizer.

positioning the fibers in a grid depends on the dimensional control of the fiber elements and the mode of fabrication of the assembly.

Problems arise in the construction of scanner grid assemblies that are similar to problems with fiber optic array connectors. In two-dimensional array connectors, the alignment mechanism is either a hole mask through which the fibers are threaded or thin chips are used that are grooved on both sides and stacked to hold the fibers.⁹ These methods work well and are economical for the relatively small number of fibers in cable. However, for aligning fibers for scanners that use many thousands of fibers, construction processes should be automated to avoid handling individual fibers. A method of forming plastic optical waveguide by selective photopolymerization¹⁰ offers one means of mass fabricating large linear arrays.¹¹ Unfortunately, the plastic sheets would cause fabrication problems. For example, dimensional variability of the sheet thickness would cause an imprecise alignment of the lightguides in the grid. A method that can be used successfully involves wrapping continuous fiber on a drum with a transverse groove for the subsequent injection of adhesive to fuse the assembly to form precise linear arrays. Further, the linear array can be split into bands of equal amounts of fibers for subsequent stacking into grids of the required configuration.

Figure 3 shows dimensional requirements for individual fibers and for the cross section of bands of fibers that are assembled by the drum method. In this example, there are 100 fibers in each band (250 fibers/in.). The 0.004 in. diameter (101.6 μm) is typical for optical fibers currently fabricated for communication systems, but the required tolerance of ± 0.00001 in. (± 0.25 μm) standard deviation is well beyond normal practice. To achieve this precision requires using advanced fiber drawing control¹² and instrumentation systems that use noncontacting on-line diameter measuring.^{13,14} As shown in Fig. 3, fibers in the band are uncoated and are positioned in tangential contact with each other. Epoxy fills the interstitial spaces on the bottom of the band. The adhesive joins the fibers together and forms a flat surface that is used to support the stack of other bands in making the grid.

Figure 4 shows an enlarged portion of a grid assembly. In this assembly, the vertical dimension is also controlled by the fiber diameter. A second application of adhesive, this time to fill the remaining interstitial spaces between fibers, is used to secure the bands to form a compact and precise grid. As noted previously, a light mask placed over the grid can compensate for minor misalignment induced by slight variations in fiber diameter and processing. The light mask consists of an array of 10,000 0.001-in. holes, located on 0.004-in. centers. The mask is fabricated by the same technology used for solid state sensor arrays. The final assembly therefore provides a grid of tiny emitting

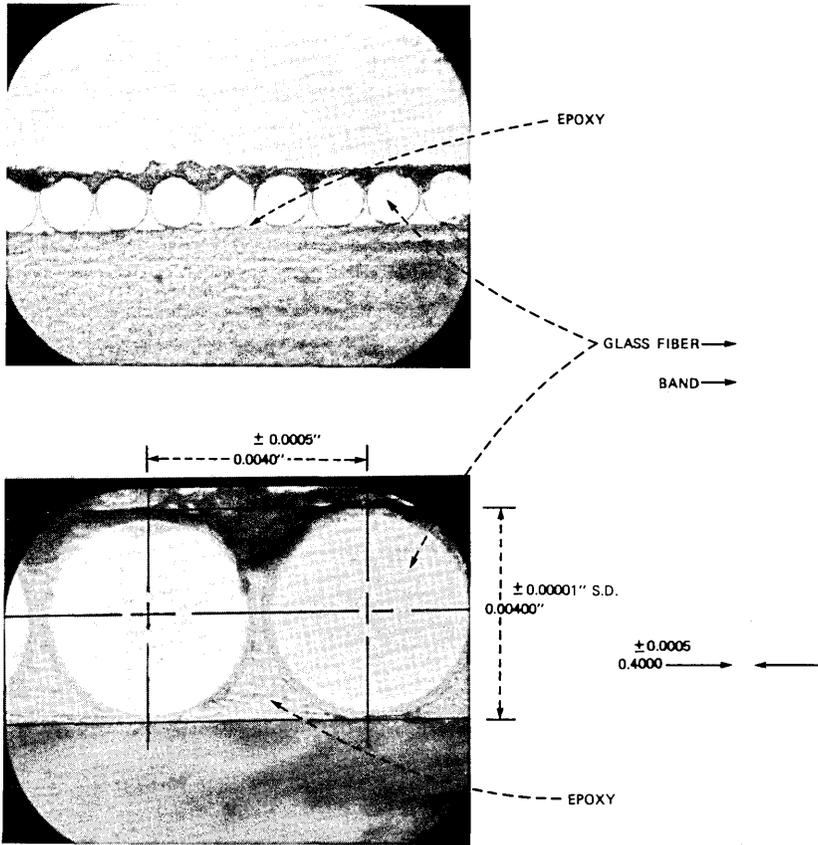


Fig. 3—Dimensional requirements for optical fiber linearly aligned lightguides at 250 bits/inch resolution.

holes with dimensional control similar to that obtained in solid state array fabrication.

IV. SCANNER PERFORMANCE OF FIBER OPTICS

Figure 5 shows the results of scanning a test pattern to evaluate the characteristics of a fiber assembly made by the drum and band method. Lines on the test pattern, varying in width from 0.0040 in. to 0.0145 in. produce detectable signals via the 0.004-in. diameter fibers. Fibers used in this test are of stepped-index type with a 0.7 core to cladding diameter ratio. The fibers are spaced nominally on 0.004-in. centers. As shown in the figure, this dimensional arrangement of fibers can resolve into discrete signal/line widths that are well below the 0.015 in. minimum that is the ANSI Y14.2 recommendation for engineering drawings.

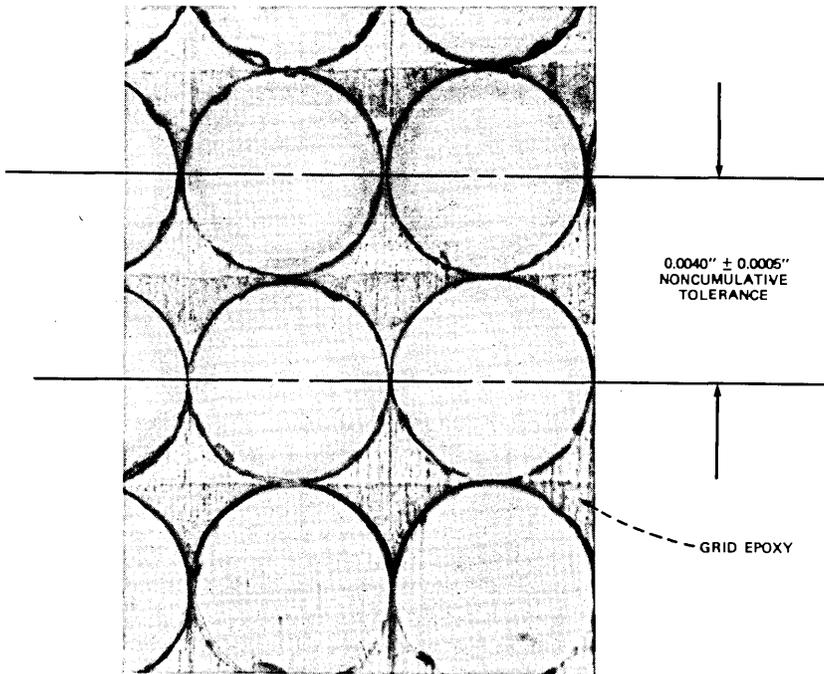


Fig. 4—Portion of optical fiber grid receiver.

In Figure 5 the digitized signals represent a partial scan across the drawing in the X direction. The absolute accuracy of scanned data in the X direction (perpendicular to the document motion), relative to the graphics on the drawing, is controlled by the positional alignment of the optical fibers as shown in Fig. 3. In the experimental unit each fiber is positioned within ± 0.0005 in. to adjacent fibers, and each band of 100 fibers varies in width by at most ± 0.0005 in. However, the placement of bands adjacent to each other to form the linear portion of the lightguide receiver introduces additional variations that cause apparent stretching or shrinking in the width of the engineering drawing. This effect can be compensated by an appropriate step in subsequent data processing to define the physical location of each lightguide rather than to assume the nominal spacing.

Accuracy in the Y-direction scan is controlled by the rotational characteristics of the stepping motor and dimensional variation in the friction feed rollers. Commercially available gear-driven stepping motors operate with an accuracy of ± 4 arc min. This results in a maximum single-step error of 0.00026 in. per step, but because of gearing, the error is noncumulative. Dimensional variations measured in experimental friction-feed rollers can contribute less than 0.2 percent varia-

tion to the individual step distance. This can produce a cumulative error and results in an apparent stretching or shrinking of the drawing in the Y direction. To obtain a more accurate digital rendition of the drawing, this condition can also be corrected by data processing.

Figure 6 shows the output from an experimental scanning digitizer for a portion of an engineering drawing. This display of thresholded video, as recorded in sections on a cathode-ray tube, shows the digiti-

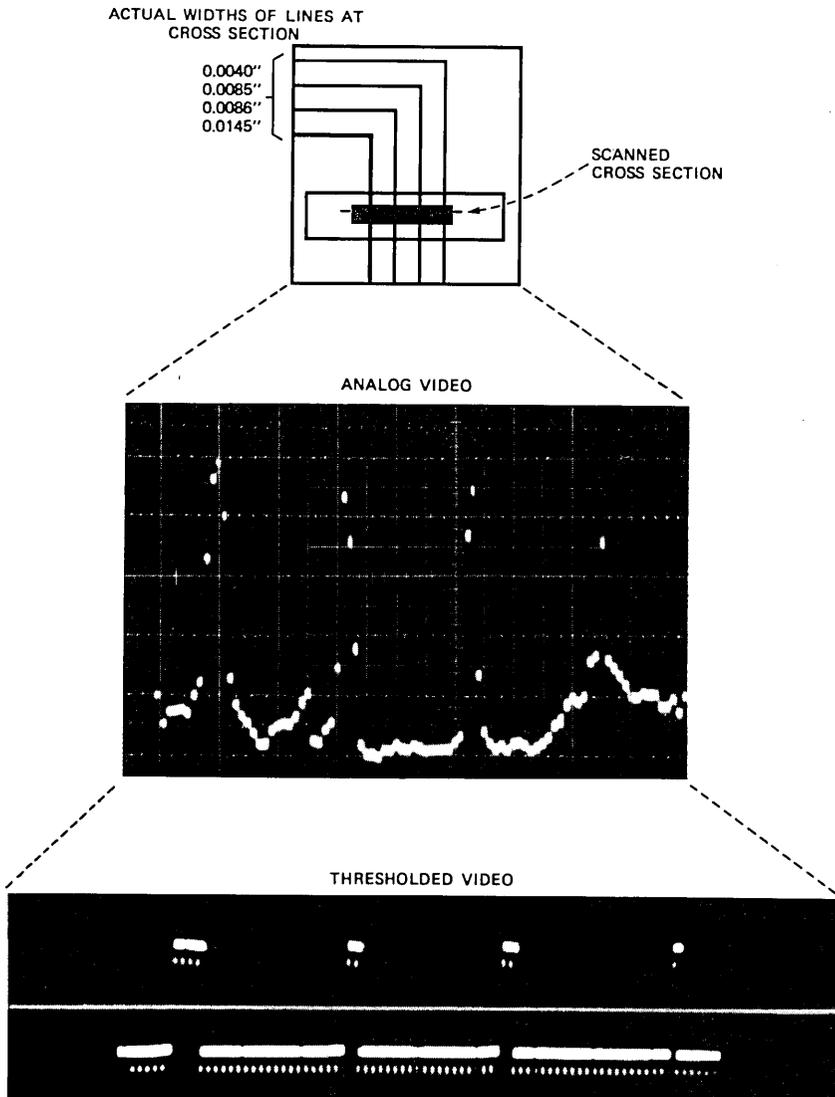


Fig. 5—Signals from a 250 fiber/inch scanner.

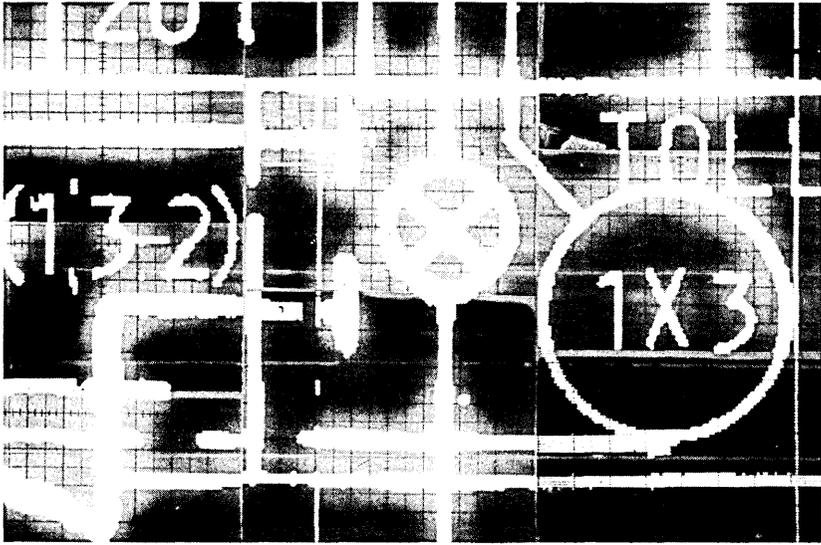


Fig. 6—Digitized output from an experimental optical fiber scanner as recorded in sections on a cathode-ray tube.

zation in the X and Y direction. Irregularities in the boundaries of the lines of the drawing appear as steps because of the digitization process. Files of this type can be processed further by computer for a variety of automated graphics activities. One use of this type of output is to create graphic files for computer-aided design systems.¹⁵

V. CONCLUSION

This paper has presented an overview of the considerations and requirements in the design of optical fibers for the scanning and digitizing of large documents. It shows that when combined with a solid state array and a roller feed, this type of assembly has distinct advantages. High resolution (250 bits per inch) can be obtained with fiber that has a diameter that is typical for communication systems. Geometric fidelity is inherent in the digitized output of this type of scanner because of the fixed-dimensional arrangement of fibers and the stepping motor-driven rollers. These aspects also contribute to the stability and mechanical simplicity of the assembly. Because the fibers are flexible, they can be folded into a compact arrangement to produce a desktop-size assembly for scanning the widest of engineering drawings. This type of assembly also permits convenient handling of the drawing with entry and exit from the front of the scanner.

Of critical importance to optical fiber scanning digitizers is the fabrication of the fiber grid to tolerances that match those of solid

state array sensors. For very large grids construction processes must be automated to avoid handling of individual fibers. We have found that wrapping continuous fiber on a drum with a transverse groove for the subsequent injection of adhesive forms precise and stable linear arrays. Further, the linear array can be split into bands of equal amounts of fibers for stacking and fusing into a grid. To achieve the required precision for the fibers, instruments that use noncontacting on-line diameter measuring, as well as advanced drawing and feedback controls, are necessary.

Using optical fibers for scanning digitizers can open new possibilities for computer-aided design and engineering. The availability of a technology that permits compact, high-resolution, and stable scanners should encourage wider use than now exists for automatic digitizing. Coupled with advanced data compaction software, convenient-to-use low-cost scanners can serve as a new input means for graphics processing systems and can permit the conversion to data of the largest of existing and newly prepared engineering drawings.

VI. ACKNOWLEDGMENTS

We sincerely thank the individuals in M. I. Cohen's Department in Laboratory 152 who provided special optical fiber for the experiments. John C. Williams and Frank V. DiMarcello graciously found time in their demanding schedules to draw the precision fiber. A special thanks also to Dave Smithgall of the Western Electric Engineering Research Center for his many insights and comments about the on-line measuring and feedback system, and for making available their precision measurement facilities. J. P. White, currently supervisor of the TOPES Development and Applications Group, is also recognized for his efforts in supervising the early development of the experimental optical fiber scanning digitizer and the data conversion software.

REFERENCES

1. D. M. Costigan, *Electronic Delivery of Documents and Graphics*, New York: Van Nostrand, 1978, Chapter 3.
2. E. H. Snow and G. P. Weckler, "Self-Scanned Charge Coupled Photodiode Sensor Arrays," *Proc. S.P.I.E.*, 116 (1977).
3. N. G. Altman, "Automatic Digitizing of Engineering Drawings," *Electro 78*, Session 23/3, Boston, May 1978.
4. M. H. Coden, "Optical Information Systems Applications," *S.P.I.E.*, 139 (1978).
5. Y. Katayama et al., "Polymer Optical Circuits for Multimode Optical Fiber Systems," *Appl. Opt.* (1979), p. 646.
6. W. Pferd and K. Ramachandran, "Computer Aided Automatic Digitizing of Engineering Drawings," *Compsac 78*, Chicago, Oct. 1978.
7. W. Pferd et al., "TOPES: A Time-Shared Computer Aided System for Office Planning and Engineering," *Comput. Aided Design*, 10, No. 6 (Nov. 1978), pp. 363-70.
8. W. Pferd, L. A. Peralta, and F. X. Prendergast, "Teleconferencing With Computer Aided Design Systems," *Eurographics 79*, Bologna, Italy, October 1979.

9. C. M. Miller, "Fiber Optic Array Splicing with Etched Silicon Chips," *B.S.T.J.*, *57*, No. 1 (January 1978), pp. 75-90.
10. W. J. Tomlinson et al., *Appl. Phys. Lett.*, *16* (1970), p. 486.
11. T. Kurokawa et al., "Fiber Optic Sheet Formation by Selective Photopolymerization," *Appl. Opt.*, *17* (1978), p. 4.
12. M. I. Cohen and R. J. Klaiber, "Drawing of Smooth Optical Fibers," *Tech. Dig. Top. Meet. Opt. Fiber Transm.*, 2nd, 1977, Paper TWB4.
13. L. S. Watkins, "Instrument for Continuously Monitoring Fiber Core and Outer Diameters," *Top. Meet. Opt. Fiber Transm.*, 1st, 1975, Paper TWA4-1.
14. D. H. Smithgall, L. S. Watkins, and R. E. Frazee, Jr., "High-Speed Noncontact Fiber-Diameter Measurement Using Forward Light Scattering," *Appl. Opt.*, *16* (1977), p. 2395.
15. K. Ramachandran, "Coding Method For Vector Representation of Engineering Drawings," *Proc. IEEE*, *68*, No. 7 (July 1980), pp. 813-7.

Magnetic Localization of Buried Cable by the SCARAB Submersible

By S. H. FRANCIS, H. R. LUNDE, T. E. TALPEY, and G. A. REINOLD

(Manuscript received June 19, 1980)

A magnetometer array has been installed aboard the SCARAB submersible to enable it to detect and track buried undersea telephone cable. Four three-axis magnetometers sense the magnetic field of a 25-Hz current applied to the cable. The magnetometer signals are filtered, amplified, digitized, and multiplexed onto SCARAB's umbilical cable for transmission to the control ship. A shipborne minicomputer processes these signals in real time to determine the cable location, which is continuously displayed to SCARAB's operator on a graphics terminal. This paper describes the development and capabilities of the system. Among the topics discussed are a history of the cable-locating problem, an analysis of the factors governing the achievable 25-Hz signal level, a description of the magnetic noise spectrum and its sources, and a discussion of the signal-processing techniques. We also examine the dependence of system performance on signal and noise levels.

I. INTRODUCTION

SCARAB (Submersible Craft Assisting Repair and Burial) is a new class of unmanned cable-controlled submersible craft whose mission is to facilitate the maintenance of undersea telephone cable systems. The first SCARAB, now under development, is owned and will be operated by a consortium including AT&T Long Lines and British, Canadian, and French telecommunications companies.* A second SCARAB belongs to Transpacific Communications, Inc., a subsidiary of AT&T Long Lines. The SCARAB vehicle, shown in Fig. 1, is typical of other recently designed unmanned submersibles. The basic structural element is a

* Transpacific Communications, Inc. (a subsidiary of AT&T Long Lines); Cable and Wireless, Ltd.; French Postes Télégraphe et Téléphonie; British Post Office; Teleglobe Canada.

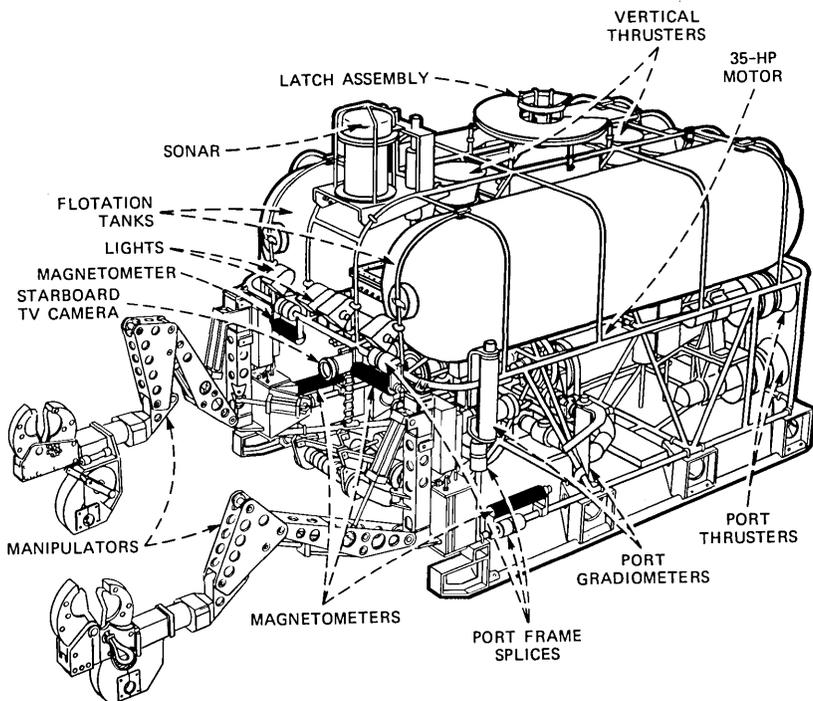


Fig. 1—SCARAB's major subsystems.

frame of welded aluminum rod. Flotation tanks are mounted on the frame to provide buoyancy, and seven thrusters are used for propulsion. Two of the thrusters are oriented vertically to allow the vehicle to dive and ascend. Two manipulators are provided for cutting and gripping cable. A water pump is used in conjunction with dredger and jetter attachments for unburying and burying cable. Additional equipment includes TV systems, acoustic devices, and a sensor system that includes the cable detection apparatus.

Operationally, SCARAB is designed to perform two specific functions: to assist during the repair of faulted cable, and to perform surveys of operational cables. During a repair operation, SCARAB magnetically locates the cable track and, subsequently, the point of the fault. It then unburies and cuts the cable as required, and attaches grippers so the cable ends can be lifted to the cable ship. Once both ends are aboard the ship, the cable is spliced, tested, and then lowered to the sea bed. The SCARAB then cuts away the lowering lines and reburies the exposed length of cable.

Basically, the survey operation is to follow the cable and measure the burial depth with the use of the magnetometer detection system.

If a point of shallow burial or cable exposure were detected, the vehicle would rebury using the jetter.

The purpose of the present paper is to describe in some detail SCARAB's cable-locating subsystem, designed and constructed by Bell Laboratories. More precisely, SCARAB has two distinct cable-locating systems, for reasons discussed below. The ac system described herein is considered the primary system; the dc system (not described) is intended to serve as a backup under certain operational circumstances.

II. HISTORY OF CABLE-LOCATING PROBLEM

Telephone cable systems use single coaxial cables whose outer diameters are between 2.5 and 5 cm, with repeaters inserted every 5 to 40 nautical miles, depending on the system. (The newer systems have the larger cable and closer repeater spacing.) Since the first transatlantic telephone cable was laid in 1956, the major threat to undersea cables has been the trawling and dredging operations of fishermen on the Atlantic continental shelves. In fact, about 95 percent of undersea telephone cable failures have been attributed to fishing activities. The principal response to this threat, aside from more heavily armoring the shelf sections of cables, was the development in 1966 of the first of a series of plows (Sea Plows I to IV) for burial of ocean cable.¹ Early versions of the sea plow buried the cable to a depth of 70 cm, leaving the repeaters exposed; the latest plow² buries the repeaters as well.

Burial reduces the likelihood of cable faults but by no means forestalls them. Even a buried cable system is vulnerable near unburied repeaters, or where cable burial was impeded by adverse bottom conditions, or where bottom erosion occurs after plowing, or where ships' anchors may drag.

This vulnerability, though minimized, nevertheless poses a problem: if a buried cable is damaged, how can it be found and its fault localized? An approximation to the fault location can be obtained from shore by combining knowledge of the cable's route with an estimate of the offshore distance to the fault, the latter obtained by any of several electronic fault-locating techniques applied at the cable's shore terminal. This procedure can identify the faulty repeater section, thereby isolating the fault to within a few miles, but the problem remains for the cable-repair ship to pinpoint the fault location and recover the cable. This problem is considerably exacerbated by burial of the cable, a fact recognized at the time the first sea plow became operational. To provide repair ships with the capability of finding buried cables, Bell Laboratories developed three varieties of electromagnetic probe: an electrode pair to be streamed behind the repair ship with a horizontal spacing of about 50 m; a 76-cm long magnetic induction coil to be

towed at low speeds in a vertical orientation near the sea floor; and a smaller induction coil for hand-held use by a diver.

To use the probes, a low-frequency signal (20 to 40 Hz) is applied between the cable's center conductor and ground at the shore cable terminal nearer the fault. This signal propagates along the cable, and its electric field can be detected at sea by the electrode system streamed behind the repair ship, allowing the ship to track the cable at relatively high speeds by zigzagging across it. Unless signal attenuation is excessive, the most seaward location at which the signal is detected provides a crude fault-location estimate, which can be refined by towing the induction coil at slow speeds so that it hangs nearly vertically, indicating by a null when it is above the cable. It was intended at this stage of the repair operation that a diver, using the hand-held probe to pinpoint the fault location, attach lifting lines and, if necessary, cut the cable. However, because of the depth limitations and inherent danger of diving operations, the usual procedure is to pick up the cable by grappling. Although grappling is quite straightforward for unburied cable, grappling has obvious disadvantages when applied to buried cable. The grapnel may not penetrate the bottom consistently to the cable's burial depth, thereby running the risk of missing the cable completely; or the cable may not have adequate tensile strength to withstand being pulled obliquely (often nearly horizontally) out of the bottom through more than a meter of mud or sand. These problems and similar ones motivated SCARAB's development.

III. CABLE LOCATING BY SCARAB: OPERATIONAL REQUIREMENTS AND SYSTEM CONCEPTS

3.1 SCARAB's cable-locating requirements

During a typical repair operation, SCARAB is called upon to find a buried cable and track it for possibly a mile or more until the cable fault is located. Thereupon it must position itself on the bottom, close to the cable, so its dredger and manipulators can be used to unbury the cable, cut it if necessary, and attach a lifting line. Cable reburial follows cable repair. SCARAB may also be required to estimate the cable's burial depth, both during repair operations and during survey of operational cable, to determine whether the cable should be buried more deeply.

In view of the complexity of these tasks and the difficulty of remotely maneuvering a swimming vehicle within a few meters of the bottom, the cable-locating system aboard SCARAB must be considerably more sophisticated than the ship-towed system described above. A principal limitation of the ship-towed system, from the viewpoint of its possible application aboard SCARAB (e.g., by mounting a single induction coil

on the vehicle), is its reliance on a single sensor, which measures only one component of the electric or magnetic field vector, and at only one point in space. Clearly, one instantaneous measurement of this kind is insufficient to localize the cable, which therefore can only be found by maneuvering the probe to a variety of positions near the cable (e.g., zigzagging), noting the field component as a function of time (and therefore of space), and using this time series to estimate the cable's location.

This procedure serves well where the information required concerning the cable's location is minimal and qualitative, such as during a grappling operation, where one usually needs only to know which of two opposite courses will cross the cable. However, for quantitative cable localization (e.g., to determine cable burial depth), this procedure would of course require real-time computer processing; but furthermore the processing algorithm would require knowledge of the vehicle's track over the bottom so that past history could be used to determine present cable location. Hence the computer would need accurate data concerning vehicle speed and ocean current; of these, the former is difficult to measure accurately at slow speeds and the latter is virtually impossible to measure from a swimming vehicle. (The problem is analogous to ship navigation using running fixes, a procedure whereby past and present lines of position are conjoined to determine the ship's position. The past line of position is updated using a dead-reckoned ship's track, which is subject to errors in the estimated course and speed of the ship and the set and drift of the current. The procedure is viable only because the ship's speed is usually large compared with the errors in its assumed speed over the bottom, and because the requirements for positional accuracy are not stringent. Neither of these helpful conditions pertains to SCARAB's cable-localization problem.)

So that estimates of track over the bottom do not have to be relied upon, a precision cable-locating system must have enough sensors so that their outputs, taken at a single instant of time, are sufficient to determine cable location without using any past data. For this purpose, a minimum of five sensors is required, since four parameters are needed to specify a given line in space (the cable location) and one parameter is needed to specify the cable current. In fact, for reasons discussed below, it proves convenient to use 12 sensors, packaged in four three-axis modules.

3.2 Choice of sensor

An attractive possibility for a cable-locating sensor is a bottom-penetrating sonar. In fact, a parametric sonar has succeeded in detecting telephone cable buried under 30 cm of sand in a laboratory tank experiment.³ Although the feasibility of a sonar cable detector has

been thereby strongly suggested, the time and funding required to develop an operational sonar system for SCARAB were judged to exceed the available resources. Accordingly, attention was focussed on passive electromagnetic sensors.

The next choice to be made was between electric and magnetic sensors. Because at least five sensors would have to be mounted on a small vehicle crowded with other equipment, compactness was an overriding concern. From this viewpoint, electric sensors (i.e., electrode pairs) are a disadvantage because their size, determined by their baseline length, is normally considerably larger than a comparably sensitive magnetic detector. Furthermore, the technology of commercially available magnetic sensors is well advanced, having benefited from long-standing geophysical and military interest in compact and reliable magnetic detectors (an interest with no counterpart in electric-field detection). For these reasons, and because the geometry of the cable's magnetic field offers some processing advantages over its electric-field geometry, magnetic cable detection was chosen.

Magnetic detectors fall into three categories, according to the magnetic quantity measured: total-field magnetometers (measuring the field magnitude), component magnetometers (measuring individual vector field components), and induction coils (measuring the time derivative of individual field components).

Of these, induction coils can be readily dismissed from further consideration. Their sole advantage over magnetometers is their lower self-noise level, which is irrelevant in this application because other noise sources dominate self noise. Their principal disadvantages are the size and the weight of their cores (important considerations for a small equipment-laden vehicle that must be positively buoyant).

The choice between total-field and component magnetometers is not as straightforward. To clarify the functional difference between them, consider the case of a total-field magnetometer in the presence of the earth's field \mathbf{B}_e and a cable field \mathbf{B}_c , where $|\mathbf{B}_c| \ll |\mathbf{B}_e|$ (as is always the case in practice). If $\hat{\mathbf{b}}$ is a unit vector parallel to \mathbf{B}_e , then the magnetometer output is $|\mathbf{B}_e + \mathbf{B}_c| \cong |\mathbf{B}_e| + \hat{\mathbf{b}} \cdot \mathbf{B}_c$, which is simply the component of the total field parallel to the earth's field (i.e., the perpendicular components of the cable field do not contribute). A total-field magnetometer is therefore completely equivalent to a component magnetometer oriented along the earth's field.

Suppose now that \mathbf{B}_c is time-varying, so that its contribution to the magnetometer output can be separated from that of the earth's static field by filtering, leaving a signal of $\hat{\mathbf{b}} \cdot \mathbf{B}_c$ from the total-field device and $\hat{\mathbf{a}} \cdot \mathbf{B}_c$ from the component device (where $\hat{\mathbf{a}}$ is a unit vector parallel to the magnetometer axis). The sole distinction between these outputs is that $\hat{\mathbf{b}}$ is fixed while $\hat{\mathbf{a}}$ is changeable (i.e., a component magnetometer can be oriented to measure any desired component).

This flexibility of the component magnetometer is both an advantage and a disadvantage. The advantage is that three component magnetometers with mutually orthogonal axes can be combined in a single package to create a vector sensor (i.e., a sensor which detects all three components of the vector field). Because the electromagnetic field is a vector field, there are crucial signal-processing advantages to be gained by employing a vector sensor, as will be seen. The disadvantage of the component magnetometer is that it is subject to a potentially severe source of noise if mounted on a moving vehicle. As the vehicle maneuvers, pitches, and rolls, it changes the orientation of the magnetometer's axis relative to the earth's field, so that in the earth coordinate system \hat{a} is time-varying. The magnetometer output is therefore $[\mathbf{B}_c(t) + \mathbf{B}_e] \cdot \hat{a}(t)$, and the contribution from the earth's field can no longer be removed by filtering unless $\hat{a}(t)$ has negligible frequency components in the frequency band of the signal $\mathbf{B}_c(t)$. Because the earth's field normally exceeds the cable field by nearly four orders of magnitude, even slight vehicle motions can produce intolerable levels of motional noise if these motions are in the signal frequency band.

The choice between total-field and component magnetometers therefore hinges on whether a usable signal frequency exists, at which motional noise is acceptably small. As will be shown below, SCARAB has sufficient inertia and stability, so when it is submerged, motional noise is negligible at frequencies above a few Hertz. This fact was not known when the choice of sensor had to be made, because SCARAB had not yet been built; accordingly, an experiment was conducted using the Navy's CURV III, a cable-controlled underwater recovery vehicle quite similar in design to SCARAB. A component magnetometer mounted aboard CURV III revealed not only that motional noise was acceptably low at all frequencies of possible interest, but also that magnetic noise from all sources (notably the thrusters) was low enough to encourage the development of the cable-locating system described below.

3.3 Choice between ac and dc signals

A further preliminary question concerned whether the system should operate at ac or dc. An ac system would detect the ac magnetic field of a low-frequency current on the cable. A dc system would detect the superposition of dc fields from two sources: dc current on the cable, and permanent magnetization of the cable's steel components (central strength member or external armor wires).

The overriding difference between ac and dc cable detection derives from the existence of a large dc component of the earth's field. To detect the small dc cable field in the presence of the earth's field, one must use the fact that the earth's field is spatially homogeneous,

whereas the cable field is a localized anomaly. Therefore, to detect the cable one must detect differences between field values at different locations in space; if these values differ, the cable is nearby. This procedure can be implemented with one sensor by moving it in space and watching for changes in signal. Or it can be implemented with multiple sensors by looking at the differences among their signals. The first of these techniques must employ a total-field sensor, since a vector sensor would be subject to changes in orientation as it was moved (the dc equivalent of motional noise). The second of these techniques could employ either total-field or component sensors, though usually component sensors are used, with two being packaged in the same housing to maintain their parallelism. Such a configuration is termed a gradiometer, since its output is proportional to the field gradient if the sensor separation is small.

One contrast between ac and dc cable detection, therefore, is in the type of sensor that can be used: ac detection can use a vector magnetometer, while dc detection requires a total-field magnetometer or gradiometer. Put another way, the ac sensor measures a vector quantity while the dc sensor measures either a scalar or a tensor quantity. Because the magnetic field is a vector, one might expect the vector sensor to lend itself more readily to data processing for cable localization, thereby favoring ac over dc. In fact, this expectation is borne out: the ac cable-locating algorithm described below is extremely simple, whereas a comparable dc algorithm has not yet been devised.

Nonetheless, a dc cable-locating system possesses certain advantages over ac. Since undersea telephone cables normally carry a dc current (with sea return) to power the repeaters, a dc detection system does not require a special signal to be placed on the cable, and the system works as well for an operational cable as for a broken cable. Furthermore, unlike an ac signal, the dc current does not diminish exponentially with distance from shore, so both ends of a broken cable can be powered with comparable current levels. In fact, when the cable is armored it does not need to be powered at all for dc detection, which can take advantage of the relatively large fields of the steel armor wires. In contrast, an ac signal cannot propagate across the ocean, and therefore, at best, it can reach only one side of a break in a transoceanic cable. Under certain circumstances, therefore, the unique capabilities of a dc system may be necessary.

In sum, dc sensors provide magnetic data of a type that is not suitable for cable localization, with the result that an ac system is normally preferable. However, operational circumstances exist for which an ac system would be ineffective and a backup system is desirable. For these reasons, a dc system is under development for SCARAB to supplement the ac system.

The remainder of this paper describes the implementation and operational performance of the ac system.

3.4 Cable-localization algorithm and array design

In the preceding sections we proposed that the optimum cable-localizing system should employ a vector sensor (i.e., a sensor detecting all three field components), because in some sense a vector sensor is better suited to characterizing a vector field. A corollary was that the optimum system was an ac system, to avoid the earth's dc field. A further proposition was that the optimum system should have enough sensors to determine cable location instantaneously, without having to rely on past history.

To particularize these general considerations, this section describes the method implemented aboard SCARAB for instantaneous cable localization by a number of ac vector sensors. The method relies on the fact that the magnetic field lines caused by the cable's current are circles (assuming the cable lies in a straight line). This geometry permits cable localization by a simple triangulation procedure most easily visualized in two dimensions (Fig. 2). Vector sensors are used to determine the direction of the field at two locations. At both locations a line perpendicular to the field is constructed. Because these two lines lie along radii of concentric circles, their point of intersection is the common center: the cable location. The generalization of this procedure to three dimensions is straightforward: at two locations, construct planes perpendicular to the field direction, and determine the line at which these planes intersect. This line of intersection is the cable's location.

In principle, this procedure allows cable localization with only two

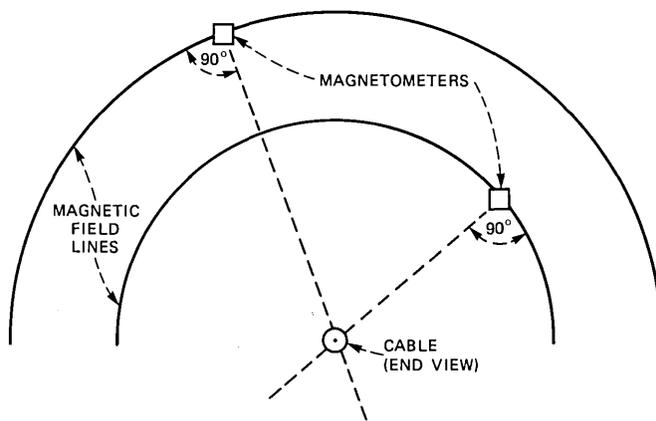


Fig. 2—Cable location in two dimensions. SCARAB's computer employs a three-dimensional generalization of this principle.

vector magnetometers. Moreover, even this basic system overdetermines the problem: a line in space can be described with four parameters, whereas two vector sensors provide six scalar measurements. These six measurements can be manipulated to provide information on field direction (four angles) and size (two magnitudes). Cable localization requires only the four angles; the two magnitudes can provide two cable-current estimates, which would of course coincide in the absence of noise and other sources of field distortion.

However, this simple system of two vector sensors must be embellished, for it fails when the two sensors and the cable all lie in the same plane (or, in the presence of noise, nearly in the same plane). This degenerate case cannot be ignored, for it occurs quite commonly for any two-sensor array. For example, to optimize cable tracking it would be logical to place the two sensors symmetrically to port and starboard, so that the sensors straddle the cable when SCARAB does. For such a placement, however, SCARAB would be unable to do cable localization while heading perpendicular to the cable, such as during initial cable acquisition.

The solution is to provide more than two sensors, and to implement a procedure for deciding which sensors shall be used and which ignored. To ensure that no cable location shall be coplanar with all pairs of sensors, a nonplanar array of at least four sensors is required.

The optimum position of these four sensors is governed by the desire that the planes whose intersection determines cable location should intersect nearly at right angles, since then the cable-location errors caused by noise-induced fluctuations of these planes will be minimized. For right-angle intersection, the baseline length between two sensors must be twice the distance from the baseline to the cable. However, SCARAB's maximum baseline length is constrained to be less than two meters (since booms extending beyond the vehicle's perimeter are undesirable), while SCARAB's height above the sea floor will generally be at least one meter. Consequently, right-angle triangulation is impossible to achieve, and the effects of noise are minimized simply by spacing the sensors as far apart as possible, maximizing the distance between each sensor and the plane defined by the other three. In fact, even this modest goal is impractical, because most potential sensor sites in the center and after parts of the vehicle are contaminated by thruster noise and frame-current distortion of the magnetic field (see below). Consequently, the best available array configuration comprises four coplanar sensors in a nearly vertical plane in the forepart of the vehicle (Fig. 1). This array is not well suited to localizing a cable which lies athwartships and underfoot (i.e., a cable in the plane of the sensors), but for all other cable locations it serves well.

Such an array of four sensors provides six possible sensor pairs and therefore six cable-location estimates, if the outputs from all pairs are

used. Since some pairs will be better situated than others to triangulate the cable, a choice must be made among them. This choice is made by calculating a weighted average of the six cable-location estimates, which de-emphasizes sensor pairs whose measured magnetic fields are nearly parallel. The weighting function used is $\sin \theta$,⁴ where θ is the angle between the magnetic fields and therefore between the planes whose line of intersection is sought. This weighting function satisfies the requirement that it increase monotonically from 0 at $\theta = 0^\circ$ (the most undesirable case) to 1 at $\theta = 90^\circ$ (the most desirable case), but beyond this requirement its choice is rather arbitrary.

This section has described how SCARAB's ac cable-location system works, in principle. Outputs from an array of four vector magnetometers are combined in a shipboard computer, using a rather simple cable-locating algorithm and weighting scheme, and the resulting estimate of cable location is displayed to SCARAB's operator so that he may maneuver the vehicle accordingly. How well this system works in practice depends on the relative levels of signals and noise. This is addressed in Section IV.

IV. SIGNAL AND NOISE

4.1 Signal

The magnetic field \mathbf{H} of a line current I in an infinite homogeneous nonconducting medium satisfies

$$|\mathbf{H}| = \frac{I}{2\pi\rho}, \quad (1)$$

where \mathbf{H} is in amperes/meter, I is in amperes, and ρ is the distance from the line current in meters. In a conducting medium such as sea water, this expression is valid only if $\rho \ll \delta$, where δ is the skin depth. In sea water, however, $\delta \cong 250/f^{1/2}$ meters, where f is the frequency in hertz, so that $\delta > 40$ m for all frequencies less than 40 Hz. Equation (1) is therefore well satisfied within about 20 m of the cable. This conclusion is unaffected by the presence of the sea bottom, whose lower conductivity causes no appreciable field distortions except at distances greater than a sea-water skin depth. An important implication is that the circular geometry of the field lines, on which SCARAB's cable-location method relies, is not appreciably distorted by the presence of the sea bottom except at ranges in excess of 20 m.

To characterize the signal, therefore, it remains only to specify I , the net current on the cable. When a signal current is fed to the center conductor at the shore terminal, return currents flow partially in the sea and partially in the outer conductor, metallic shielding tapes, and armor wires, if any. The net current I , which determines the field according to eq. (1), is the algebraic sum of all currents on the cable's

various conductors. This net current attenuates as the signal propagates seaward, because of losses in the cable, the sea water, and the repeaters (which introduce loss rather than gain because the low-frequency signal follows the dc power path through the repeaters rather than the path through the carrier-frequency amplifiers). For example, Fig. 3 shows a theoretical prediction of propagation at several frequencies along a 100-mile section of armored cable with 5-mile repeater spacing and with a cable break at the offshore end. The appendix describes the theory underlying Fig. 3. The net current is lower at higher frequencies, not only because the losses are greater, but also because the outer conductor carries a greater share of the return current (so that the net current, being the algebraic sum of center- and outer-conductor currents, is less). Clearly, from the viewpoint of signal propagation, we prefer the lower frequencies.

The choice of operating frequency, however, is ideally based on maximizing the signal-to-noise ratio. This ideal requires knowledge of the vehicle's noise spectrum, however, and SCARAB's development schedule did not permit delaying the choice of frequency until the vehicle was completed and its noise spectrum measured. Therefore, we provisionally chose 25 Hz for convenience, since 25 Hz is a frequency commonly used with ship-towed probes (though its choice in that context was influenced by considerations entirely irrelevant to SCARAB, such as the low-frequency falloff of induction-coil sensitivity and the expected low-frequency noise caused by turbulence-induced motion of

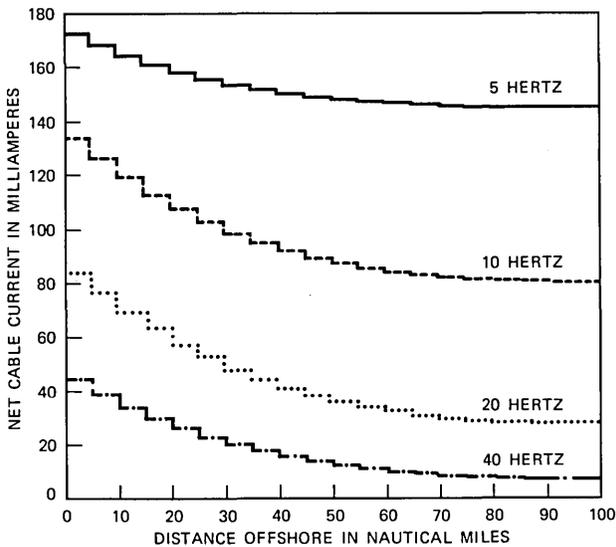


Fig. 3—Propagation of low-frequency signals on armored sc cable with a cable break 100 nmi offshore. The center-conductor current is 200 mA at the shore point, for all frequencies. The step discontinuities occur at the repeaters, which are unpowered.

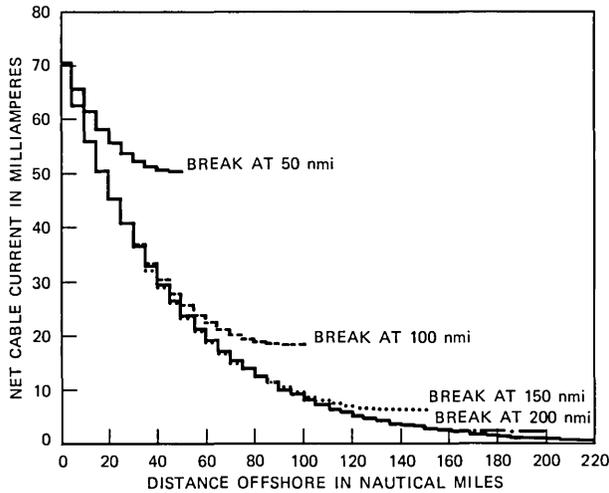


Fig. 4—Propagation at 25 Hz on armored SC cable with cable breaks at various locations. The unlabeled curve extending to 220 nmi is for an unbroken cable. The center-conductor current at the shore point is 200 mA (but note that the vertical scale extends only to 80 mA). The step discontinuities occur at the repeaters, which are unpowered.

a ship-towed coil). Although the choice of 25 Hz was therefore arbitrary to some degree, subsequent determination of SCARAB's noise spectrum (Section 4.2) showed that 25 Hz lies in a rather quiet region of the spectrum, and so we have retained it as SCARAB's operating frequency.

Figure 4 illustrates propagation at 25 Hz on the same type of cable system as Fig. 3, but with cable breaks at various distances offshore. Notice that a cable break enhances the signal by total reflection of the propagating mode, thereby roughly doubling the cable current at the fault relative to the current on an intact cable. A broken cable is therefore twice as easy to detect as an intact one (if both carry the same current at their shore points).

Aside from the question of signal attenuation, the maximum current that can be propagated to the desired detection point is restricted by the maximum center-conductor current tolerated by the repeaters. This value is 100 or 200 mA (rms), depending on the cable system, except when the system is operational, in which case no more than 20 mA (rms) can be tolerated without affecting service. Broken cables therefore can carry at least ten times the signal of operational cables, with a factor of at least five due to the higher allowed current and a factor of two due to the signal reflected from the cable break.

4.2 Noise

Knowledge of the noise field and its sources is a prerequisite to predictions of operational performance, and also may suggest methods

of performance improvement. SCARAB's magnetic noise originates from a variety of sources, described below in rough order of decreasing importance.

4.2.1 Thruster noise

SCARAB's propulsion is by thrusters comprising induction motors (2300 volt, four pole, three phase, 60 Hz) driving shrouded screws. Six 5-hp thrusters are used, four for fore-and-aft propulsion and two for vertical propulsion (Fig. 1). A seventh thruster, horizontally oriented athwartships, is powered by the hydraulic system, whose pump is driven by a 35-hp motor. The noise spectra of CURV's and SCARAB's thrusters were measured in tank tests and, as would be expected for any rotating machinery, consisted of spectral lines with no measurable broadband component.

The basic frequencies of an induction motor's operation are the shaft-rotation frequency and the synchronous frequency (i.e., the apparent rotation rate of the stator's magnetic field). For a thruster with no external load or internal friction, these frequencies are the same, equalling the power frequency (60 Hz) divided by the number of pole pairs (two for SCARAB, three for CURV). Under load the thruster slows down, reducing the shaft-rotation frequency by an amount which, when multiplied by the number of pole pairs, is called the slip frequency. For example, the synchronous frequency of the SCARAB's thrusters is 30 Hz, while the shaft and slip frequencies at one-third speed are 10 Hz and 40 Hz, respectively.

Thruster noise is principally at four frequencies: the slip and shaft frequencies, and their sum and difference frequencies. Noise at higher harmonics, although measurable at high speeds, is small. The geometry of the noise field at the slip and shaft frequencies is consistent with an oscillating dipole source aligned parallel to the thruster axis. In contrast, the geometry of the noise field at the sum and difference frequencies is consistent with an oscillating and rotating dipole source aligned perpendicular to the thruster axis. By far the strongest of these lines is the slip-frequency line. As the speed of SCARAB's thrusters varies between zero and 90 percent speed (the maximum attainable under load), the frequency of this line varies between 60 Hz and 6 Hz, and its root-mean-square amplitude (expressed as an effective source dipole moment) varies between 0 and $19.6 \gamma\text{-m}^3$ (where $1 \gamma = 1/400 \pi \text{ amp/m}$). The amplitude is roughly a linear function of the voltage, and therefore varies quite nonlinearly with motor speed (at half speed, the noise amplitude is about one quarter of its full-speed value).

The precise origin of these noise components is obscure, related in some way to secondary leakage fields and therefore to subtle details of motor design and manufacture. An illustration of the subtlety is that

a 5-hp SCARAB thruster has about six times the noise of a 10-hp CURV thruster when both are operated at maximum speed.

Thruster noise poses a potentially awkward problem, for it is composed of several spectral lines whose frequencies can attain any value between 0 and more than 60 Hz, depending on the operating speed. If the signal frequency were chosen higher than any attainable by thruster noise, a heavy penalty would be paid in the attenuation of the signal in its propagation from shore (Fig. 3). The alternative is to base the choice of frequency on other considerations and rely on being able to avoid thruster speeds producing excessive noise at the chosen frequency. The strongest constraint is on the use of the forward vertical thruster, which is closest to the magnetometer array. The operator must therefore be particularly careful in the use of that thruster while the cable-localization system is being used. In fact, during normal search and tracking operations the operator customarily relinquishes control of the vertical thrusters to SCARAB's automatic altitude control system, which servo-controls these thrusters to maintain the desired height above the bottom. Fortunately, this automatic altitude control system uses the vertical thrusters in a manner quite favorable to the cable-localization system.

To illustrate the effects of thruster noise, Fig. 5 shows a noise spectrum recorded using a three-axis magnetometer mounted on SCARAB at a point midway between the two upper magnetometers shown in Fig. 1. This point is well within a meter of the forward vertical thruster, is about 1.5 m from the 35-hp hydraulic-pump motor, and is more than 2 m from the thrusters at the stern. The spectrum was recorded during dockside tests in San Diego harbor, while SCARAB was swimming submerged at nearly full speed with the vertical thrusters controlled by the automatic altitude system. We feel that this spectrum is reasonably characteristic of the typical noise environment of all SCARAB's magnetometers.

In Fig. 5, the intense spectral line near 40 Hz comes from the forward vertical thruster, whose proximity to the magnetometer causes it to dominate the other noise sources. At the time the spectrum was taken, this thruster was operating at one-third speed, generating noise at its slip frequency near 40 Hz and at its sum frequency (i.e., shaft frequency plus slip frequency) near 50 Hz. These values of the frequency reflect typical operating speeds of the vertical thrusters, whose average speed in level flight is, of course, a direct function of the vehicle's buoyancy. The automatic altitude system maintains level flight using rather small speed excursions, in contrast with manual altitude control in which the natural tendency of the operator is to overreact by using large changes in vertical thrust.

In comparing the three components of the 40-Hz line, note that the

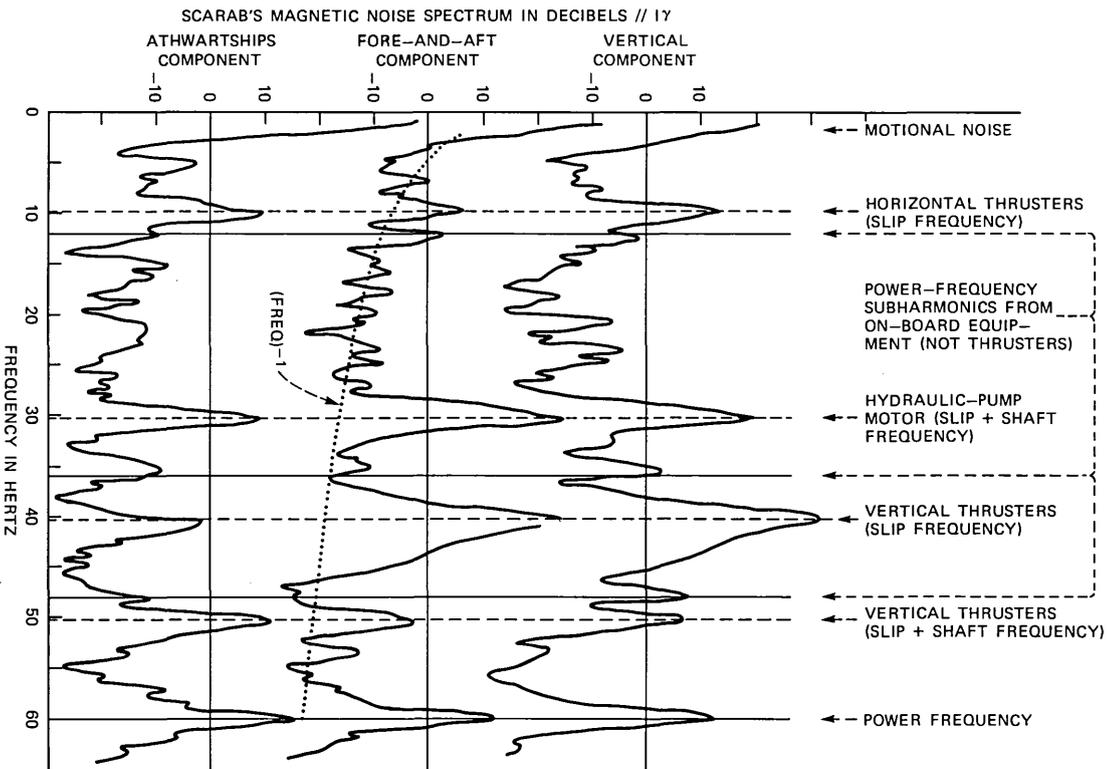


Fig. 5—Noise spectrum of vector magnetic field measured aboard SCARAB.

smallness of the athwartships component (relative to the other two components) is consistent with the tank-test result that the slip-frequency noise source is equivalent to a dipole parallel to the thruster axis.

Other spectral lines in Fig. 5 come from the horizontal thrusters and the hydraulic-pump motor. The horizontal thrusters, operating at nearly full speed, contribute noise at their slip frequency of about 10 Hz. Although they are generating more noise than the vertical thrusters (because of their higher speed and their greater numbers), their impact on the noise spectrum is substantially smaller because of their greater distance from the magnetometer.

The 35-hp hydraulic-pump motor operates at a constant speed just below its synchronous speed of 30 Hz, and its most important contribution to the noise spectrum is a strong line at its sum frequency (slip plus shaft frequency) just above 30 Hz. Its slip frequency is very low (near 1 Hz), and so its slip-frequency noise is of no practical importance.

From the spectrum in Fig. 5 it is clear that 25 Hz is a fortuitous choice of operating frequency, since it avoids the frequencies at which thruster noise is most severe. The frequencies of the horizontal and vertical thrusters are variable, of course, but it is quite practical to avoid the critical speed (58 percent) at which their slip frequency is 25 Hz. The vertical-thruster speed necessary to maintain level flight is substantially less than 50 percent at the vehicle's buoyancy, and the automatic altitude system maintains vehicle altitude without approaching the critical speed of 58 percent. On the other hand, we can expect the speed of the horizontal thrusters to pass through 58 percent in normal operation, but their noise level at this speed is substantially lower (by 8 dB) than their level in Fig. 5. As long as SCARAB does not swim continuously at this speed, the effects of the horizontal thrusters should be confined to small, short transients seen by the magnetometers each time the critical speed is passed.

In addition to the spectral lines discussed above, the spectrum of Fig. 5 contains noise at the power frequency and several other fixed frequencies. The 60-Hz line contains contributions from all the vehicle's equipment, but the predominant sources are the thrusters and the pump motor, since they draw the most current. Subharmonics of the power frequency at multiples of 12 Hz are also present, caused by unidentified equipment on board the vehicle other than the thrusters and pump motor. The subharmonic at 24 Hz is the weakest of these lines (not even showing on Fig. 5), and so its proximity to the 25-Hz operating frequency has no ill effects.

4.2.2 Frame currents

To continue the discussion of noise sources, this section describes the problem of distortion of the cable's field by currents induced in

SCARAB's structure. This type of noise is particularly insidious because of its perfect correlation with the signal. Although eddy currents in SCARAB's massive metal objects (such as the manipulator arms and flotation tanks) were found to have negligible effect, currents circulating in the supporting framework created a severe problem which had to be faced.

SCARAB is constructed as a welded framework of solid aluminum rods which support its flotation tanks, thrusters, sensors, and other equipment (see Fig. 1). The framework itself can be considered a passive electrical circuit which, because of the large cross section and high conductivity of the aluminum rods, has an extremely low impedance. This low impedance has an unfortunate consequence. Time-varying magnetic fields, inducing electromotive forces in the framework in accordance with Faraday's law, drive sizable currents in the frame members. These frame currents in turn cause secondary magnetic fields which, superposed on the primary fields, have the effect of distorting the field geometry on which the cable-locating system relies.

To appreciate the severity of this effect, consider a simplified framework in the form of a 1-m square composed of solid aluminum rods of 4-cm diameter (the size of SCARAB's main frame members). The calculated resistance of such a frame is $100 \mu\Omega$ and its inductive reactance is $17f \mu\Omega$ (where f is the frequency in Hz). These calculations neglect the resistance of the welds and assume a frequency sufficiently low so that the skin effect is negligible (not quite true at 25 Hz, but nearly so). At frequencies above 6 Hz the framework is primarily inductive, at lower frequencies it is primarily resistive.

Now suppose that an infinite straight wire carries a sinusoidal current, situated in the same plane as the square frame but 1 m away and parallel to one edge. A straightforward application of Faraday's and Ohm's laws shows that for frequencies at which the framework's impedance is inductive (i.e., above 6 Hz), the induced frame current is about 5 percent of the primary current. Since a magnetometer mounted on the framework could easily be 20 times closer to the frame than to the wire, the secondary magnetic field seen by the magnetometer could easily rival the primary field in magnitude.

Although SCARAB's framework is topologically more complicated, the same effect is present. To estimate its impact, we ran tests in which a current-carrying wire was placed at various locations within a meter of the vehicle, and measurements were made of the magnetic fields (primary plus secondary) and of the currents flowing in the frame members. The amplitudes of the frame currents, measured with a specially modified clamp-on ammeter, were of the same order of magnitude as predicted in the preceding paragraph. Consequently, the field anomaly was large enough to defeat the cable-locating algorithm.

We contemplated three possible solutions to the problem. One thought was to reduce the operating frequency, thereby reducing the induced EMF proportionately. This solution would require lowering the frequency to 1 Hz or below, however, because at frequencies above a few hertz the impedance is inductive and therefore decreases along with the EMF as the frequency is decreased, so that the currents remain the same. Since frequencies at or below 1 Hz are objectionable for many reasons (e.g., motional noise), we discarded this alternative. Another thought was to mount the magnetometers far away from any frame members, thereby reducing the adverse effects of the frame currents. In practice, this alternative requires four booms that hold the magnetometers one or two meters outside the vehicle's perimeter and retract to prevent damage during vehicle launch or recovery. We did not attempt to implement such a system, although it is probably feasible, because a simpler alternative was available.

The final solution was to prevent the frame currents from flowing in critical frame members by cutting these members and splicing them back together with insulating splices. Obvious choices for splicing were the four frame members on which the four magnetometers are mounted. Also spliced were the two vertical frame members at the front of the vehicle (see Fig. 1). Frame currents flow in the remaining frame members, of course, but their distance from the magnetometers reduces the adverse effects to acceptable levels.

The combination of frame currents and thruster noise constrains the placement of magnetometers to the front of the vehicle, where there are no thrusters and where the density of frame members has been minimized to improve TV visibility. Initial plans included the placement of one magnetometer further aft, but the high density of frame members at the intended location (each of which would have had to be cut and spliced) forced the relocation of this magnetometer to the front of the vehicle. The consequence, as mentioned in a preceding section, is that the cable-localization system is blind to a cable which lies under the vehicle and perpendicular to its center line, but this blind spot does not significantly reduce the usefulness of the system.

4.2.3 Motional noise

The strong source of noise evident below 5 Hz in Fig. 5 is motional noise, attributable to changes in magnetometer orientation relative to the earth's magnetic field as SCARAB pitches, rolls, and yaws. The identification of this noise is verified by its absence when SCARAB is on a stable platform (such as the sea bottom). When SCARAB is surfaced in a seaway, a spectral peak in the motional noise at the frequency of the surface water waves is sometimes identifiable. Above 5 Hz, however, motional noise is negligible relative to other noise sources.

4.2.4 Broadband noise

In addition to narrowband noise from the sources described above, SCARAB's noise spectrum includes broadband noise which falls off with increasing frequency roughly as $1/f$ (see Fig. 5). This frequency dependence matches the Fourier transform of a step function, and indeed if one looks at the noise in the time domain (restricting attention to a quiet frequency band such as the neighborhood of 25 Hz), one sees a series of impulsive transients. This behavior suggests that the broadband noise is due to the combined effects of a number of weak, random, impulsive sources.

One example of such an impulsive source is the actuation of a solenoid, of which many are used in SCARAB's hydraulic system. In fact, under quiet circumstances the actuation of SCARAB's solenoids is readily observable by a magnetometer more than 2 m away. Other potential impulsive noise sources include any movable equipment with residual magnetization, such as SCARAB's manipulator arms and 35-mm camera (whose casing is highly magnetized, causing large broadband transients during panning and tilting). Slight random shifts in the position of such equipment would cause broadband noise of the kind observed.

Although the identification of the principal broadband sources remains speculative, there are a number of sources which can definitely be excluded from consideration, since their levels are demonstrably insufficient to account for the observed broadband levels. These sources include detector noise (generated within the magnetometer core), atmospheric noise (from nearby man-made sources or from worldwide thunderstorm activity, ducted by the ionosphere), and tape noise (on the magnetic tape used for recording test data).

4.3 Signal-to-noise ratio and cable-localization performance

With the signal and noise levels known, the performance of the cable-locating system can be predicted. Consider a typical scenario in which SCARAB is looking for a cable break 100 miles offshore. Figure 4 shows that the net current at the break is about 40 mA (rms) and Fig. 5 implies that the noise level in a 1-Hz band at 25 Hz is roughly $\frac{1}{4}\gamma$ (rms). The range at which the cable's signal, given by eq. (1), equals the noise level is 32 m, which may be termed the "cable-detection range."

At this range, however, and even well inside it, the cable-localization algorithm still does not have adequate signal-to-noise ratio to localize the cable because it is the field differences (rather than the fields themselves) that contain the significant information for localization. Therefore, to calculate a "cable-localization range" by equating signal and noise, one should use as the signal the field difference $\Delta H =$

$(\Delta\rho)(\partial H/\partial\rho)$, where $\Delta\rho$ is a measure of the mutual separation of SCARAB's sensors. To allow for the noise contributions (assumed to be correlated) from two magnetometers, the noise level should be increased by a factor of two. With $\Delta\rho = 1$ m, this calculation gives a localization range of 4 m.

As SCARAB approaches the cable, therefore, it passes through three regions in which it can infer increasing levels of information concerning the cable. Outside the detection range, it does not sense the cable. Inside the detection range but outside the localization range, it detects the proximity of the cable but has insufficient information to determine the cable's location. Inside the localization range it can localize the cable, with the precision of this localization increasing as the distance to the cable decreases. For a net cable current $I = 40$ mA (rms), the detection and localization ranges are 32 m and 4 m, and for other values of I they scale as I and $I^{1/2}$, respectively.

Localization of the cable is reasonably precise only at quite modest ranges (a few meters or less). The fundamental obstacle to localization at greater ranges is SCARAB's compactness, which imposes severe constraints on the size of the magnetometer array. However, the description in Section VI of SCARAB's operational use shows that the present localization range is quite adequate.

V. CABLE-LOCATING SYSTEM: DETAILS OF DATA COLLECTION, PROCESSING, AND DISPLAY

Preceding sections have discussed the cable-locating system in rather general terms, omitting details of the actual hardware and software which carry out the functions of data collection, processing, and display. Some of these specifics will now be discussed, to promote a better understanding of how the system works in practice.

5.1 Data collection

The sensors are four commercial three-axis ring-core fluxgate magnetometers. Each magnetometer is mounted in a cylindrical aluminum pressure housing with three preamplifiers whose function is to remove from the magnetometer outputs the dc components (which are proportional to the three components of the earth's dc field) and to amplify the residuals for transmission to an electronics housing at the stern of the vehicle. The relative gain among the three preamplifiers is adjusted to compensate for the differential attenuation of the magnetic field components by the cylindrical aluminum container (whose eddy currents reduce the 25-Hz field inside to about 82 percent of its outside value, with slight attenuation differences between the axial component and the other two).

Each magnetometer signal is then passed through a bandpass filter

at 25 Hz and band-elimination filters at 30 Hz and 60 Hz (to eliminate the pump-motor and power-frequency noise). The signals are then fed to rms detectors whose outputs are sampled five times per second, digitized by a 12-bit analog-to-digital converter, and multiplexed onto SCARAB's umbilical cable for transmission to a minicomputer on the control ship.

In addition to the rms amplitudes, the cable-localization algorithm requires the relative phases of the three field components from each magnetometer to allow determination of the orientation of the vector magnetic field. These relative phases (which are either 0° or 180°) are continuously determined by phase comparators whose outputs, sampled five times per second, are digitally coded and packed with the digitized amplitudes for transmission to the surface minicomputer.

The process of digitizing the amplitudes introduces a certain amount of error because of the uncertainty in the digitizer's least-significant bit and because of bias and fluctuations in its reference voltage. This error normally affects only the two lowest-order bits, and is equivalent to magnetic noise of about 0.1 to 0.2 γ (rms), somewhat less than (but not negligible compared with) the vehicle noise at 25 Hz. This digital noise level could be reduced by increasing the preamplifier gain, with the penalty that the digitizer, which presently saturates at fields of about 120 γ (rms), would saturate at proportionately lower levels. The present gain setting was predicated on keeping the effects of digital noise below those of vehicle noise while maximizing the field at saturation. It follows that any future improvement or degradation in vehicle noise may require corresponding gain changes to maintain the desired relationship between digital noise and vehicle noise.

5.2 Data processing

The data collection system just described provides the shipboard minicomputer with three field-component amplitudes and their relative phases from each of four magnetometers. Sensors on the vehicle also provide its digitized heading, pitch, roll, and height above the bottom. New data are provided five times per second.

The first step in the real-time processing of these data, after the relative phases are used to assign signs to the rms component amplitudes, is to calculate a time average of the magnetic fields using an exponentially decaying weight function,

$$\langle \mathbf{H} \rangle = \int_0^\infty \frac{dt}{\tau} e^{-t/\tau} \mathbf{H}(-t), \quad (2)$$

where τ is an averaging time and $\mathbf{H}(-t)$ denotes the magnetic field at a time t in the past. This running average, whose purpose is to reduce the effects of the noise, is updated with each new input of data (five

times per second). The value of τ is chosen by SCARAB's operator, and may be changed at will as dictated by the operational situation. The rationale for particular choices of τ will be discussed below.

Note that this averaging is not the only averaging undergone by the data. The 25-Hz bandpass filter (whose $Q = 20$) has the side effect of averaging the analog signals according to eq. (2) with $\tau = 0.25$ s.

From these averaged data the computer determines cable location twice per second according to the algorithm described geometrically in Section 3.4. To express the algorithm mathematically, let the cable location be specified by the vectors \mathbf{r}_0 and $\hat{\mathbf{a}}$, where \mathbf{r}_0 is the vector from the coordinate origin to the closest point on the cable and $\hat{\mathbf{a}}$ is a unit vector parallel to the cable. If the cable carries a net current I , then the magnetic field observed at any point \mathbf{r} is

$$\mathbf{H}(\mathbf{r}) = \frac{I}{2\pi} \frac{\hat{\mathbf{a}} \times (\mathbf{r} - \mathbf{r}_0)}{|\mathbf{r} - \mathbf{r}_0 - \hat{\mathbf{a}}\hat{\mathbf{a}} \cdot (\mathbf{r} - \mathbf{r}_0)|^2}. \quad (3)$$

This equation is the vector generalization of the scalar eq. (1).

The core of the data-processing problem is to calculate the cable's location (expressed by \mathbf{r}_0 and $\hat{\mathbf{a}}$) from two measurements of the magnetic field (\mathbf{H}_m and \mathbf{H}_n) made by two magnetometers located at coordinates \mathbf{r}_m and \mathbf{r}_n . The dependence of \mathbf{r}_0 and $\hat{\mathbf{a}}$ on the measured parameters \mathbf{H}_m , \mathbf{H}_n , \mathbf{r}_m , and \mathbf{r}_n is given by the two equations below, whose derivation follows:

$$\mathbf{r}_0 = \frac{\mathbf{r}_m \cdot \mathbf{H}_m H_n^2 - \mathbf{r}_n \cdot \mathbf{H}_n \mathbf{H}_m \cdot \mathbf{H}_n}{H_m^2 H_n^2 - (\mathbf{H}_m \cdot \mathbf{H}_n)^2} \mathbf{H}_m + \frac{\mathbf{r}_n \cdot \mathbf{H}_n H_m^2 - \mathbf{r}_m \cdot \mathbf{H}_m \mathbf{H}_m \cdot \mathbf{H}_n}{H_m^2 H_n^2 - (\mathbf{H}_m \cdot \mathbf{H}_n)^2} \mathbf{H}_n, \quad (4)$$

$$\hat{\mathbf{a}} = \frac{\mathbf{H}_m \times \mathbf{H}_n}{|\mathbf{H}_m \times \mathbf{H}_n|}. \quad (5)$$

These equations can be derived using mathematical arguments that parallel the geometrical discussion of Section 3.4. For variety, however, we present a more formal derivation.

Suppose for the moment that \mathbf{H}_m and \mathbf{H}_n are not parallel, in which case it is clear that the three vectors \mathbf{H}_m , \mathbf{H}_n , and $\mathbf{H}_m \times \mathbf{H}_n$ comprise a basis spanning three-dimensional vector space. The vectors \mathbf{r}_0 and $\hat{\mathbf{a}}$ must therefore be expressible as

$$\mathbf{r}_0 = \alpha \mathbf{H}_m + \beta \mathbf{H}_n + \gamma \mathbf{H}_m \times \mathbf{H}_n, \quad (6)$$

$$\hat{\mathbf{a}} = \lambda \mathbf{H}_m + \mu \mathbf{H}_n + \nu \mathbf{H}_m \times \mathbf{H}_n, \quad (7)$$

for some choice of the scalars α , β , γ , λ , μ , ν . Since eq. (3) shows that $\hat{\mathbf{a}}$ must be perpendicular to both \mathbf{H}_m and \mathbf{H}_n , it follows that $\lambda = \mu = 0$. Choosing ν so that $\hat{\mathbf{a}}$ is a unit vector results in eq. (5) for $\hat{\mathbf{a}}$.

Turning now to eq. (6) for \mathbf{r}_0 , note that by definition \mathbf{r}_0 is perpendicular to $\hat{\mathbf{a}}$ (and therefore to $\mathbf{H}_m \times \mathbf{H}_n$), so $\gamma = 0$. To find α and β , take the inner product of eq. (6) with \mathbf{H}_m and \mathbf{H}_n in turn. The resulting scalar equations may be solved for α and β ,

$$\alpha = \frac{\mathbf{r}_0 \cdot \mathbf{H}_m H_n^2 - \mathbf{r}_0 \cdot \mathbf{H}_n \mathbf{H}_m \cdot \mathbf{H}_n}{H_m^2 H_n^2 - (\mathbf{H}_m \cdot \mathbf{H}_n)^2}, \quad (8)$$

$$\beta = \frac{\mathbf{r}_0 \cdot \mathbf{H}_n H_m^2 - \mathbf{r}_0 \cdot \mathbf{H}_m \mathbf{H}_n \cdot \mathbf{H}_m}{H_m^2 H_n^2 - (\mathbf{H}_m \cdot \mathbf{H}_n)^2} \quad (9)$$

But eq. (3) shows that $(\mathbf{r}_m - \mathbf{r}_0) \cdot \mathbf{H}_m = 0$, so that $\mathbf{r}_0 \cdot \mathbf{H}_m = \mathbf{r}_m \cdot \mathbf{H}_m$ (and similarly for n). When this identity is substituted into eqs. (8) and (9), which in turn (along with $\gamma = 0$) are substituted into eq. (6), one obtains eq. (4) for \mathbf{r}_0 .

SCARAB's computer uses eqs. (4) and (5) as follows. For each pair of magnetometers, the computer calculates a value of \mathbf{r}_0 and $\hat{\mathbf{a}}$, and furthermore calculates two estimates of the cable current by twice solving eq. (3) for the current necessary to account for each of the two observed field magnitudes. Because four magnetometers can be paired in six ways, these computations provide six estimates of \mathbf{r}_0 and $\hat{\mathbf{a}}$ which are then averaged with the weight function,

$$w_{mn} = \frac{|\mathbf{H}_m \times \mathbf{H}_n|^4}{H_m^4 H_n^4}, \quad (10)$$

which expresses the fourth power of the sine of the angle between \mathbf{H}_m and \mathbf{H}_n (a choice of w_{mn} whose motivation was discussed in Section 3.4). For example,

$$\langle \mathbf{r}_0 \rangle = \frac{\sum_{i>j=1}^4 w_{ij} \mathbf{r}_0(\mathbf{H}_i, \mathbf{H}_j, \mathbf{r}_i, \mathbf{r}_j)}{\sum_{i>j=1}^4 w_{ij}}, \quad (11)$$

and similarly for $\hat{\mathbf{a}}$ (which, in addition, is normalized to unit magnitude). An estimated cable current is also obtained, using a similar weighted average. In all these averages, magnetometer pairs for which \mathbf{H}_m and \mathbf{H}_n are nearly parallel (to within one degree) are omitted to avoid the computational singularity when $\mathbf{H}_m \parallel \mathbf{H}_n$ [which would cause the denominators of eqs. (4) and (5) to vanish]. Also omitted are any magnetometers whose contribution SCARAB's operator wishes to exclude (e.g., a malfunctioning magnetometer).

Having thus calculated an estimated cable location, the final computation uses the heading, pitch, and roll data to transform this cable location from a coordinate system fixed in the vehicle to a coordinate system independent of vehicle orientation. The cable location is then ready to be displayed to SCARAB's operator.

5.3 Data displays

The computer calculates a new cable-location estimate about twice per second, in the manner described above, and communicates this estimate to SCARAB's operator by means of a plasma display panel. Both computer and display are located in SCARAB's operator control hut, pictured in Fig. 6. The display format shown in Fig. 7 is one of several which the operator can invoke to aid in the various phases of SCARAB's operations. The particular format of Fig. 7 shows the estimated cable location in plan view relative to SCARAB's position, with SCARAB's heading indicated both numerically and by the orientation of the vehicle symbol (whose rectangular outline is drawn to scale to represent the actual location of the vehicle's perimeter). The estimated cable current and the operator-selected averaging time are also displayed.

When the vehicle is far from the cable, the magnetic data are dominated by noise and the cable-location estimates are random and meaningless. To suppress such estimates, the computer examines the cable-location estimates to establish their credibility before displaying them. For example, a cable-location estimate is not displayed if it is blatantly inconsistent with the preceding estimate. The criteria for credibility are sufficient to identify most estimates that are dominated by noise, so that in the absence of a cable the operator is not distracted by a random succession of meaningless lines.

A variation of this display format, shown in Fig. 8, includes a history of the estimated burial depth of the cable. This estimate is determined by subtracting SCARAB's height above the bottom (as measured by its altimeter) from its estimated height above the cable (given by the vertical component of r_0 , appropriately adjusted if a nonhorizontal \hat{a} indicates a sloping sea bottom). The estimated burial depth is more severely affected by noise than the estimated cable location, because the former is the difference between two larger quantities, each subject to noise. Accordingly, provision has been made for the burial depth to be time-averaged, with its averaging time chosen by the operator. This averaging time can be chosen substantially longer than the averaging time for the raw data, because at slow vehicle speeds the actual burial depth changes quite slowly with time.

A third display format (Fig. 9) shows a history of the three magnetic-field components measured by a single magnetometer (selected by the operator). This display is useful for troubleshooting, since the effects of a malfunctioning magnetometer are immediately apparent. It is also useful for cable detection when SCARAB is more than a few meters away from the cable (i.e., outside the cable-localization range but inside the cable-detection range). Furthermore, in cases when the signal is so weak that the cable-localization algorithm is ineffective at

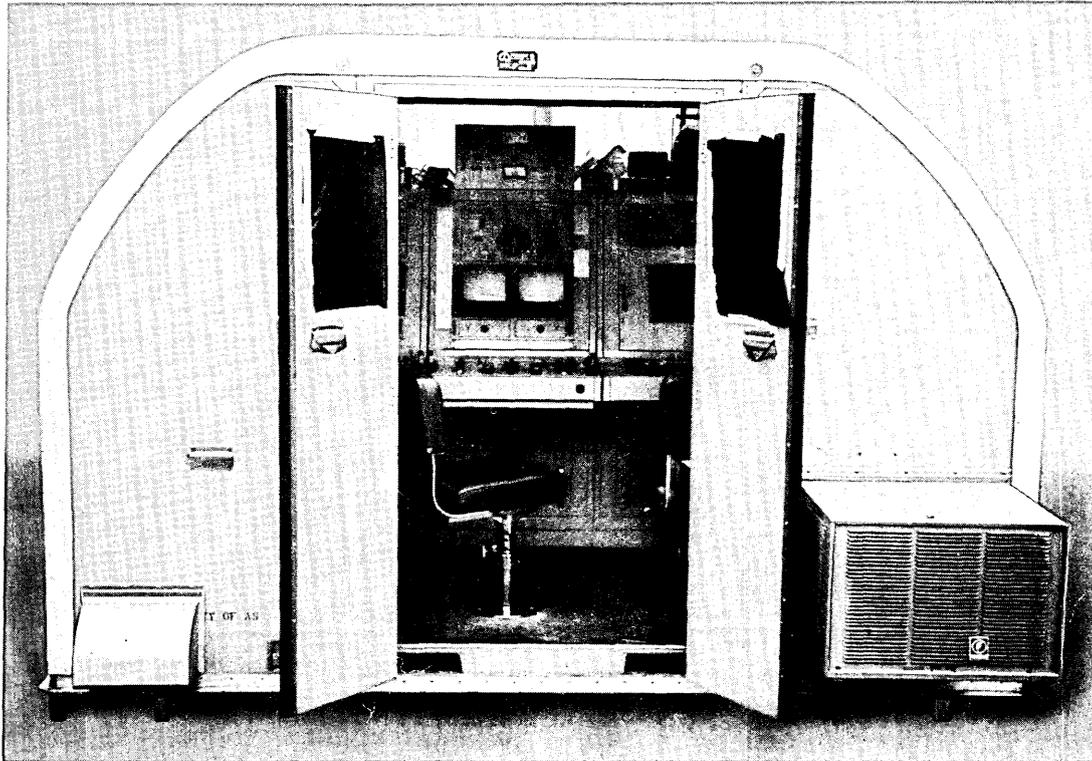


Fig. 6—SCARAB operator control hut. Computer keyboard and display are partially obscured but visible at the right-hand seat.

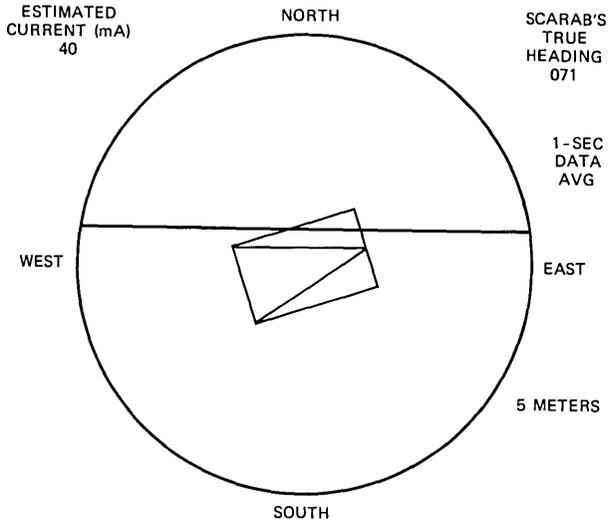


Fig. 7—Real-time computer display showing estimated cable location relative to SCARAB, in plan view.

any range, cable tracking using this display may still be possible, as described in the next section.

VI. OPERATIONAL USE OF THE SYSTEM

At the time of this writing, sea trials of the cable-locating system have been completed. The first tests were done by emplacing a length

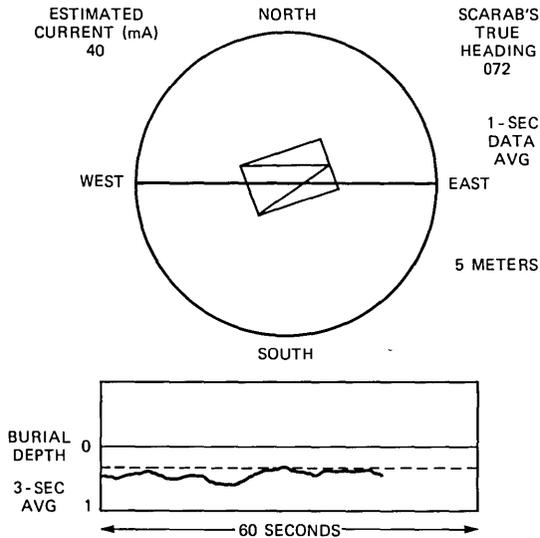


Fig. 8—Real-time computer display showing estimated cable location and a short history of estimated burial depth.

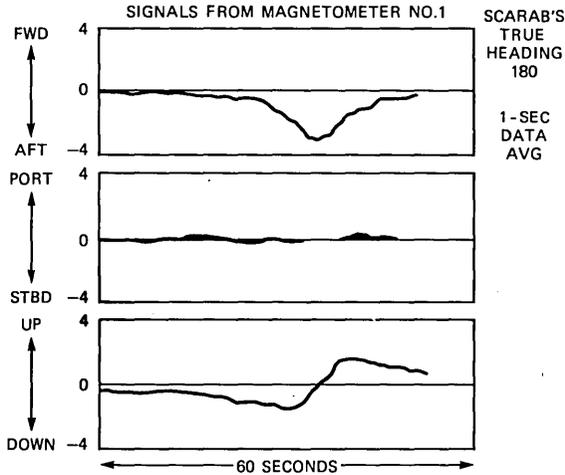


Fig. 9—Real-time computer display showing the three components measured by a magnetometer as a function of time.

of trial cable, applying a 25-Hz tone, and then acquiring and tracking the cable using the system. Later, tests were done on two operational cable systems. Finally, the system was used during the burial of a section of an operational cable.

Background noise levels were noted under various conditions throughout the trials. During vehicle pre-dive checkouts, noise levels were noted with the vehicle on deck. In this situation, magnetometer outputs were noted to vary randomly in the range of $\pm 0.3\gamma$. After the vehicle was launched it maneuvered on the surface for some time without using vertical thrusters and the outputs would decrease to $\pm 0.2\gamma$. At the beginning of a dive, the outputs increased to as much as several gammas. These levels result from the 25-Hz field components emitted from the vertical thrusters as they pass through the speed at which their slip frequency is 25 Hz. These noise levels were evident many times during descent as the vertical thruster levels were varied by the operators.

When the vehicle reached the bottom, vertical thrust was increased to 100 percent and magnetometer levels subsided to less than $\pm 0.5\gamma$. Transients of about 2γ were observed on all channels whenever vehicle lights were turned on (or off).

The location of the dive was generally far enough from the trial cable so that acquisition would be initiated by dead reckoning. As the vehicle approached the estimated cable position, a numerical display of magnetometer signals was observed to monitor the vertical fields measured by the magnetometers. (As an alternative, the vertical field of a single magnetometer could be observed in a time series using the

display of Fig. 8.) When an appreciable vertical field was observed, the display of Fig. 6 was used to allow the vehicle operator to track the cable.

When a specially laid trial cable was used an initial tone current of 200 mA (rms) was used. Field levels as large as 22γ were indicated and the cable was successfully tracked at an altitude of 4 meters.

After some confidence was established, tone current was reduced to 100 mA and then to 50 mA and indications were not noticeably degraded. Further reduction of the current to 10 mA (rms) produced an objectionable disturbance in the display, but it was still useful. There is some evidence that the actual cable currents may have been somewhat smaller than these values, because of injection losses where the current was applied to the cable; this effect was not quantified.

The burial-depth feature gave somewhat inaccurate results when the vehicle was maneuvering off the bottom, because of fluctuating altimeter readouts; however, stable and repeatable results were obtained when the vehicle was held on the bottom by its thrusters. In this mode of operation, the operator notifies the computer that SCARAB is being held on the bottom, and the computer therefore ignores the altimeter in determining burial depth. Thus, any inaccuracy in the altimeter is avoided. In addition, maneuver noise is reduced and vertical thruster levels tend to remain above the range in which 25 Hz emission occurs.

During the trials on operational cables, SCARAB was operated about 100 nmi from the cable terminal where the tone was applied to the cable. Tone currents of 20 and 30 mA (rms) were used. As a result, the actual 25-Hz current at the point of operation was only about 5 mA (rms). This proved marginal. Cable acquisition with the vehicle at an altitude of one meter could not be accomplished using the magmeter system; however, after the cable was acquired visually, tracking was possible. The system indicated an estimated tone of 3-6 mA during these tests. It should be remembered that during acquisition and tracking of faulted cables (SCARAB's principal mission), current levels can be used which are an order of magnitude higher than these.

The system was later used during the burial of an operational cable. Here, the indicated tone current was 10 mA and both acquisition and tracking were possible.

In view of the brevity of these trials, a full evaluation of the capabilities and limitations of the system must await actual operations. However, since the sea trials tended to confirm performance predictions, it seems likely that the system's performance is well enough known to support the following speculations as to the likely operational use of the system during a cable-repair operation.

With SCARAB on board, the cable-repair ship will steam to the estimated fault location and, perhaps with the aid of towed electrodes

or induction coils to sense the 25-Hz tone, will position itself shoreward of the fault and relatively near the known cable track (though not so near that the direction to the cable is ambiguous). After launch, SCARAB will proceed on or close to the seabed at a speed of about 1 knot in a direction that will cross the cable at right angles, with SCARAB'S crew monitoring the computer display of Fig. 9 to watch for the cable's presence.

At a few tens of meters from the cable, its signal will start to rise out of the noise. As the cable is crossed, the characteristic signature shown in Fig. 9 will be observed, indicating that the cable has been crossed. SCARAB'S operator will then turn the vehicle back toward the cable, intercepting it obliquely and then swimming directly above it toward the fault, using the display of Fig. 7 to facilitate cable tracking. Except in severe ocean currents, SCARAB should be able to stay within a few meters of the cable, so the restricted range of the localization algorithm should pose no handicap. Gentle bends in the cable should not affect the tracking algorithm adversely, as long as their radius of curvature is large compared with the dimensions of the magnetometer array. Sharper bends normally will be unburied, therefore causing no confusion if visibility is adequate. If the cable current is less than about 10 mA (rms), it may be insufficient for tracking using the tracking algorithm. However, SCARAB may still be able to track the cable using the display of Fig. 9, either by steering so as to maintain a zero vertical field component or by zigzagging across the cable, ascertaining from the display each time the cable is crossed.

When the cable fault is reached, it may very well be recognizable visually, either because the cable was pulled out of the bottom when it was broken or because the source of the damage (e.g., a trawler's otter boards) left visible tracks in the bottom. If there is no visual indication, the fault location may be recognized by the sudden decrease in the magnetic field, causing a deterioration of the cable-location estimate as the fault location is passed, accompanied by a sudden decrease in the estimated cable current. This effect results from the nature of the field near the end of the cable, where the field magnitude falls to zero in a horizontal distance comparable to the vertical distance of the vehicle above the cable. SCARAB'S cable-location estimate will therefore deteriorate completely in a distance of about a meter (corresponding to an elapsed time of about 2 s at a speed of 1 knot), showing that the fault has been passed. With this realization, SCARAB'S operator can circle back, reacquire the cable, and position SCARAB on the bottom near the fault, perhaps recognizing the precise fault location by the fact that when SCARAB is directly over the fault the magnetic field (and hence the estimated cable current) will have fallen to half the value it had before the area of the fault was reached. Once SCARAB

is on the bottom at the fault, it will use its dredger to unbury the cable, and the job of the cable-localization system is complete.

VII. SUMMARY

Increasing reliance on burial to protect undersea cable has led to the construction of the SCARAB submersible, whose ability to find and track such cable is clearly crucial to its success. SCARAB's primary cable-finding system is similar to previous ship-based cable-finding systems in that it requires a 25-Hz tone on the cable, but it is considerably more sophisticated than these previous systems so that the cable-tracking job is made as easy as possible. A ship-based minicomputer processes signals from four three-axis magnetometers aboard the vehicle, estimating the relative cable location, and displaying it pictorially to SCARAB's operator. The principal difficulties encountered in implementing the system were caused by two types of noise source on the vehicle: the thruster motors, and currents induced in SCARAB's framework by the cable's 25-Hz field. The thruster noise was circumvented by placing the magnetometers as far as possible from the thrusters and avoiding the thruster speed which would result in noise concentrated at 25 Hz. The frame-current problem was reduced to acceptable levels by inserting insulating splices in the frame members nearest the magnetometers, after mounting the magnetometers at locations where frame members were relatively sparse. Once these problems were overcome, the system was found to function satisfactorily with cable currents down to about 10 mA (rms). So far, operational use has been rather limited, and the outer limits of the system's performance have not yet been defined. Nevertheless, it seems clear that the system does provide SCARAB with the necessary capability for finding and tracking buried cable.

VIII. ACKNOWLEDGMENTS

The authors wish to thank J. C. Wyman for his valuable technical contributions to the project, and N. J. Zabusky for his guidance and support during the project's formative phases. We are also indebted to the crew of CURV III and the AT&T cables ship C.S. Long Lines and to the personnel of the Franklin Electric Company for their cooperation in the use of their facilities.

APPENDIX

Low-Frequency Propagation on Undersea Coaxial-Cable Systems

This appendix concerns the propagation on undersea coaxial-cable systems of signals whose frequency is sufficiently low that the electromagnetic fields are not confined to the cable but penetrate substan-

tially into the sea water. The purpose of the analysis is to calculate the signal levels in the neighborhood of the cable as a function of distance offshore, so that the performance of SCARAB's cable-locating system can be predicted. The analysis includes the effects of the repeaters, which have an important influence on signal attenuation.

In practice, an upper limit on the achievable signal level is given by the maximum low-frequency current tolerated by the repeaters. This maximum is 200 mA (rms) in the most recently installed system (the SG system), but is 100 mA (rms) on earlier systems. When in service, all systems require that the low-frequency signal be kept below 20 mA (rms). These maxima refer to the center-conductor current inserted at the shore termination; signal losses in the cable and repeaters govern how much of this signal actually reaches the point of interest (i.e., SCARAB's location), and shielding by the cable's outer conductor and armor wires determines how much of this center-conductor current can be seen by SCARAB's sensors.

The analysis of this problem divides naturally into two stages: a waveguide analysis using Maxwell's equations to determine the properties of the propagating modes, and a transmission-line analysis using these mode properties to study signal propagation on a repeatered cable system. The principal idealizations of the analysis, necessary to preserve the cylindrical symmetry of the waveguide problem, are twofold. The electromagnetic properties of the sea bottom are taken to be identical with those of sea water (insofar as they affect propagation on the cable), and the effects of armor wires (if present) are assumed the same as the effects of a cylindrical shell of equal conducting cross section. Both assumptions are plausible, but no detailed justification of either is attempted. (The theoretical analysis of coaxial cables with helically wound wire shields has recently begun to receive attention,⁴ but only in the limit of infinitely thin wires, which is an inappropriate limit for the problem at hand because the diameter of the armor wires exceeds the skin depth at the frequencies of interest.)

A.1 WAVEGUIDE ANALYSIS: THE PROPAGATING MODES

The electromagnetic fields propagating along a cable are governed by Maxwell's equations which, when expressed in cylindrical coordinates (ρ, ϕ, z) , become

$$\frac{1}{\rho} \frac{\partial}{\partial \rho} \rho H_{\phi} = \epsilon' j \omega E_z, \quad (12)$$

$$-j k_z E_{\rho} - \frac{\partial E_z}{\partial \rho} = -j \omega \mu H_{\phi}, \quad (13)$$

$$j k_z H_{\phi} = \epsilon' j \omega E_{\rho}, \quad (14)$$

where the field dependence on time t and longitudinal distance z is assumed to be $\exp(j\omega t - jk_z z)$, all fields are taken to be independent of the azimuthal angle ϕ , and

$$\epsilon' = \epsilon + \sigma/j\omega. \quad (15)$$

The constitutive parameters are the conductivity σ , the permeability μ , and the permittivity ϵ .

Equations (12) through (14) relate the three field components H_ϕ , E_ρ , and E_z which comprise cylindrically symmetric TM modes. A similar set of equations (not shown) relates the components E_ϕ , H_ρ , and H_z for cylindrically symmetric TE modes. TM and TE modes without cylindrical symmetry also exist. However, the cylindrically symmetric TM modes are the only propagating modes at the low frequencies of interest here, and the other modes are neglected.

Eliminating the electric fields from eqs. (12) through (14) yields the following equation for H_ϕ :

$$\left(\frac{\partial}{\partial \rho} \frac{1}{\rho} \frac{\partial}{\partial \rho} \rho + k_\rho^2 \right) H_\phi = 0, \quad (16)$$

where

$$k_\rho = (\omega^2 \epsilon' \mu - k_z^2)^{1/2} \operatorname{Im}(k_\rho) > 0. \quad (17)$$

The general solution of eq. (16),

$$H_\phi = AJ_1(k_\rho \rho) + BY_1(k_\rho \rho), \quad (18)$$

involves the Bessel functions J_1 and Y_1 and the arbitrary constants A and B . E_z can be derived from eqs. (12) and (18),

$$E_z = \frac{k_\rho}{\epsilon' j \omega} [AJ_0(k_\rho \rho) + BY_0(k_\rho \rho)]. \quad (19)$$

Coaxial cable consists of a number (N) of homogeneous cylindrical layers in which the electromagnetic fields take the form given by eqs. (18), (19), and (14) with A , B , k_ρ , and ϵ' assuming different values in each layer. The constants A and B in each of the N layers are constrained to satisfy the $2(N - 1)$ boundary conditions of continuous H_ϕ and E_z at the $N - 1$ cylindrical interfaces. Furthermore, the requirement that the solution be finite at $\rho = 0$ and $\rho = \infty$ implies that $B = 0$ in the central region and $B = jA$ in the outermost layer. These $2N$ constraints imply $2N$ linear equations that the $2N$ amplitudes A and B must satisfy. To permit a solution, k_z must be chosen so that the determinant of the $2N \times 2N$ matrix of coefficients vanishes.

It is easy to show that the number of such eigenvalues k_z corresponding to propagating modes equals the number of insulating layers. If the conductors (including the sea water) are perfect, this equality is

clear, for electromagnetic waves can propagate in each insulating layer independently, each such layer therefore supporting its own single propagating TEM mode. In this limit, the coaxial cable comprises a number of parallel and independent waveguides isolated from one another by perfectly conducting walls. If the conducting layers have finite conductivity, however, the waveguides are no longer independent, since electromagnetic waves can penetrate the conductors, but the number of modes remains the same. In this case the modes overlap spatially but differ from one another in their relative distribution of energy among the various insulating layers as well as in their longitudinal wave number k_z .

To illustrate the nature of these modes, we modeled a particular cable type (SG cable) and solved the eigenvalue problem. Unarmored SG cable comprises a central steel strength member (1.06 cm o.d.), a copper inner conductor (1.21 cm o.d.), a dielectric (4.32 cm o.d.), a copper outer conductor (4.37 cm o.d.), and a protective jacket (5.28 cm o.d.). For these calculations, the conductivities of copper, steel, and sea water are taken to be 5.9×10^7 , 4.8×10^6 , and 4 mhos/m, respectively, the permittivities of the insulators are $2.285\epsilon_0$ (where ϵ_0 is the permittivity of free space), the permittivity of sea water is $81 \epsilon_0$, and the permeability of steel is $180 \mu_0$ (with μ_0 the permeability of free space).

Figure 10 shows the field distributions corresponding to the two propagating modes of unarmored SG cable at 25 Hz. Mode 1 is the generalization of the TEM mode which, if the conductors were perfect, would be confined between the inner and outer conductors, while mode 2 is the generalization of the TEM mode which would be confined between the outer conductor and the sea. The attenuation lengths (defined as $-1/\text{Im}[k_z]$) of these modes are 330 nautical miles and 115 nautical miles at 25 Hz. At neighboring frequencies, the attenuation lengths scale roughly as the inverse square root of the frequency.

A.2 TRANSMISSION-LINE ANALYSIS

To predict signal propagation on a cable with N conductors (where $N = 2$ for unarmored cable and $N = 3$ for single-armored cable), we must calculate the appropriate amplitudes of the $2N$ modes (N types of modes, two directions of propagation) on each repeater section. The superposition of these modes must satisfy the relevant boundary conditions (expressed as constraints on voltages and currents) at each end of the section. This analysis requires an equivalent circuit for the repeaters, which is shown in Fig. 11.⁵ The repeater's equivalent impedance depends strongly on whether dc power is applied (as indicated in Fig. 11). The cable's outer conductor and armor wires (if present) are grounded to sea at each repeater, as shown.

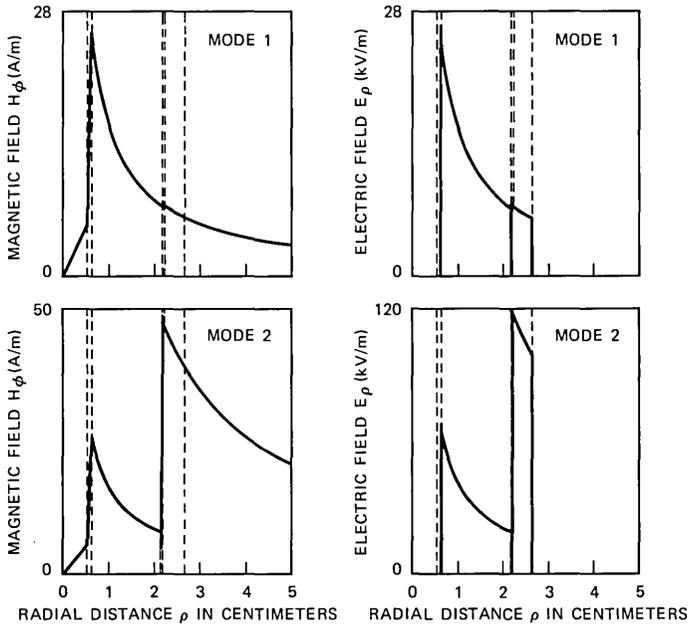


Fig. 10—Field amplitudes of the two propagation modes of unarmored sg cable at 25 Hz. The broken lines indicate the boundaries of the cable's layers: central strength member, inner conductor, dielectric, outer conductor, outer jacket, and sea water.

The transmission-line analysis is largely an exercise in matrix manipulation, and its details (which involve considerable notational complexity) will not be given in the interests of brevity. The procedure starts at the offshore cable break, where all N conductors are grounded to sea. These N constraints on the mode amplitudes, along with the $N - 1$ constraints that all conductors except the center conductor are

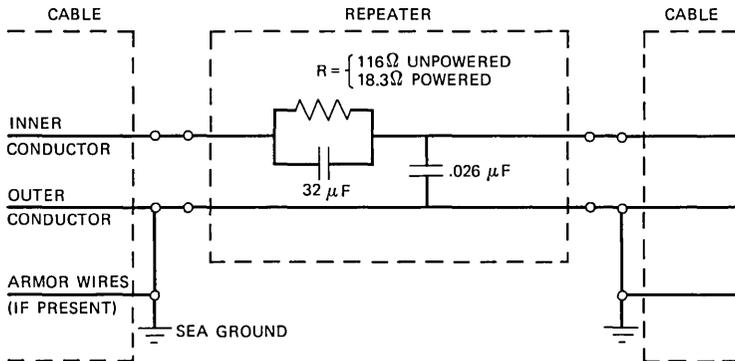


Fig. 11—Equivalent circuit of sg repeater at low frequencies. The equivalent resistance R is 18.3Ω if dc power is applied to the repeater, 116Ω otherwise.

grounded to sea at the offshore end of the most offshore repeater, provide $2N - 1$ constraints on the $2N$ modes propagating in the last cable section (i.e., the one between the last repeater and the cable break). The relative amplitudes of these modes can therefore be determined, leaving only an arbitrary multiplicative constant (common to all the modes) to be adjusted later. These mode amplitudes can be manipulated to provide the voltage and current at the offshore port of the last repeater, and conventional circuit analysis then provides the voltage and current at the shoreward port. These values of the voltage and current provide two constraints on the $2N$ modes in the next shoreward cable section, with the other $2N - 2$ constraints coming from the fact that all conductors but one (the center conductor) are grounded to sea at both ends of the cable section. The $2N$ mode amplitudes in the next-to-last cable section can therefore be found.

In this way, the solution is propagated from the cable break back toward the shore terminal, with the mode amplitudes being determined in each cable section so that the appropriate constraints on voltages and currents are satisfied at the cable break and at both ports of every repeater. Once the solution is complete, it is adjusted by a multiplicative constant so as to satisfy the driving-point condition (e.g., that the center-conductor current be 200 mA at the shore terminal).

Figure 12 shows the results of such a calculation, and assumes a

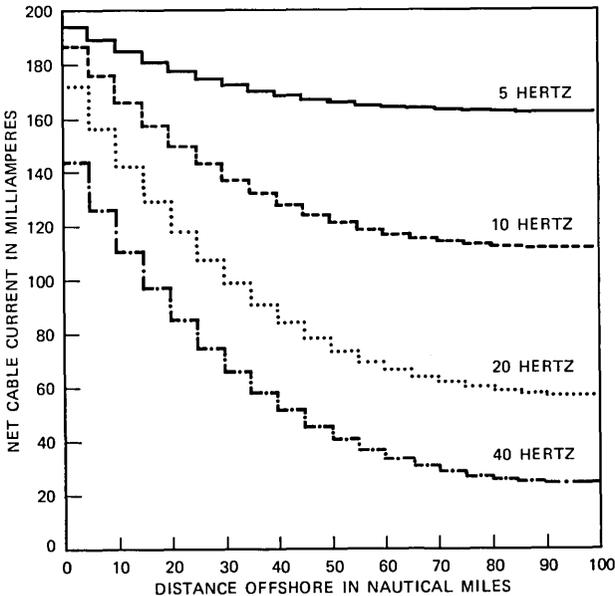


Fig. 12—Propagation of low-frequency signals on unarmored SC cable with a cable break 100 nmi offshore. The center-conductor current is 200 mA at the shore point, for all frequencies. The step discontinuities occur at the repeaters, which are unpowered.

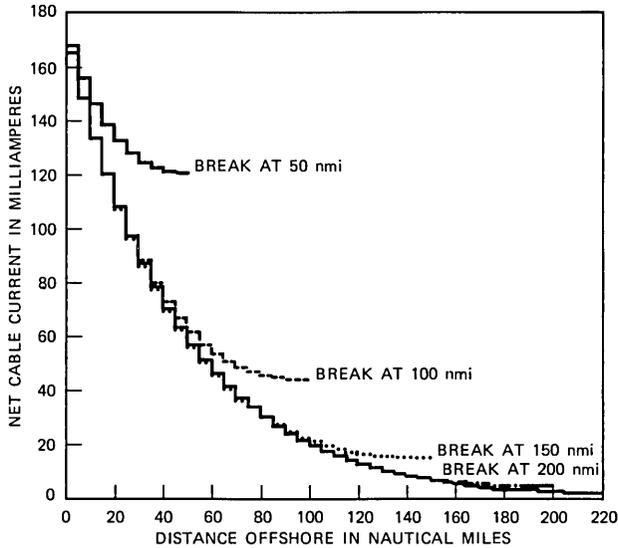


Fig. 13—Propagation of 25-Hz signals on unarmored SG cable, with cable breaks at various locations. The unlabeled curve extending to 220 nmi is for a cable with no break. The center-conductor current is 200 mA at the shore point. The step discontinuities occur at the repeaters, which are unpowered.

cable break to be 100 miles offshore. The repeaters are spaced at 5-mile intervals, and their presence is evident by the step discontinuities in the net cable current, which is defined as the algebraic sum of all currents on the cable's conductors. The step discontinuities are caused by the exchange of current between the outer conductor and the sea at each repeater. The center-conductor current itself (not shown) decays smoothly with distance, without appreciable step discontinuities, because the repeaters provide no low-impedance path from the center conductor to sea. The center-conductor current is 200 mA at the shore point in Fig. 12, but the net current is somewhat less because of the return current carried by the outer conductor. This outer-conductor current increases with frequency, so that the net current decreases with frequency.

Just as Fig. 12 illustrates the frequency dependence of the signal propagation, Fig. 13 illustrates the dependence on the distance to the cable break at SCARAB's chosen frequency of 25 Hz. One predominant feature of Fig. 13 is that the cable break reflects the propagating modes, doubling the signal amplitude relative to its value in the absence of a break. Another important conclusion supported by Fig. 13 is that the signal losses are primarily introduced by the repeaters, rather than by losses in the cable. This conclusion is borne out by the fact that the decay from 165 mA at the shore point to 1.7 mA at 220

nmi corresponds to an attenuation length of about 50 nmi, much less than the 330-nmi attenuation length of the cable's primary mode. Since the repeaters account for most of the signal attenuation (at least at 25 Hz), it follows that an improvement in propagation would result from a reduction in the repeater's impedance at 25 Hz. Because powering the repeaters reduces their impedance considerably (as Fig. 11 points out), 25-Hz propagation on a powered system (illustrated in Fig. 14) is much less attenuated than on an unpowered system. For example, the signal at 200 nmi is increased by more than a factor of 10 when the system is powered, as comparison of Figs. 13 and 14 reveals. Substantial further improvement in 25-Hz propagation on future cable systems would result from the inclusion of a bypass network allowing the 25-Hz signal to pass through the repeaters unattenuated. Such an improvement would increase the usefulness of SCARAB's cable-localization system by allowing useful signal levels to propagate farther offshore.

The calculations in this appendix so far have dealt with unarmored cable. Since many existing cable systems (and probably all future cable systems) employ armored cable on the continental shelves, it is of interest to see how armored and unarmored cable differ in their propagation characteristics. To this end, calculations involving armored cable have been carried out, modeling single-armored SG cable as identical to unarmored SG cable with the addition of a steel armor layer (6.3 cm o.d.) sandwiched between two jacket layers (5.75 and 6.76

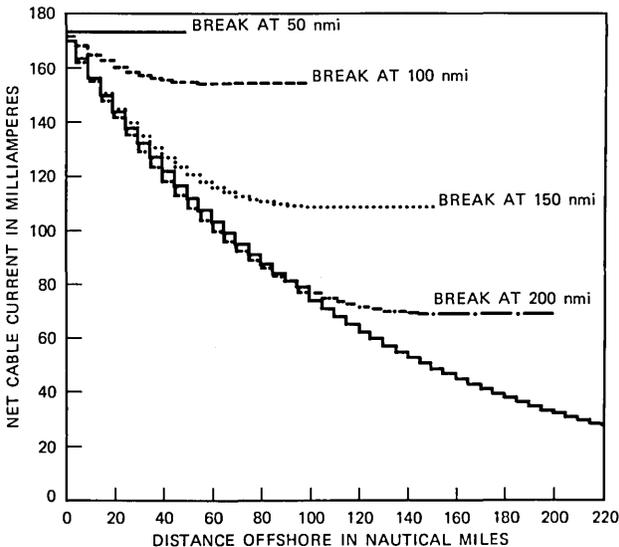


Fig. 14—Same as Fig. 13 except that the repeaters are powered, which reduces their impedance at 25 Hz.

cm o.d.). The three propagation modes of armored cable have attenuation distances at 25 Hz of 260, 140, and 86 nmi (compared with 330 and 115 nmi for unarmored cable). Propagation on armored cable at various frequencies is illustrated in Fig. 3, and propagation at 25 Hz with cable breaks at various locations is illustrated in Fig. 4. Comparison of Figs. 3 and 4 with their counterparts for unarmored cable (Figs. 12 and 13) shows one outstanding difference: the net current on the armored cable is reduced by the shielding effect of the armor layer (which carries a substantial amount of return current, particularly at the higher frequencies). At 25 Hz, for example, the net current on armored cable is only about 40 percent of that on unarmored cable. This ratio is about equal to the quantity $\exp(-\Delta\rho/\delta)$, where $\Delta\rho$ is the thickness of the armor layer and $\delta = (2/\omega\mu\sigma)^{1/2}$ is the skin depth in steel (which is a better shield even than copper because of its high permeability μ , which more than compensates for its lower conductivity σ). The attenuation of the signal with distance offshore, being governed primarily by the repeaters rather than the cable, is the same for both armored and unarmored cable.

A.3 SUMMARY

This appendix has presented the results of theoretical calculations of propagation on SG repeatered cable systems (both armored and unarmored). Propagation at various frequencies, with cable breaks at various locations, and with both powered and unpowered repeaters, is displayed in Figs. 3, 4, 12, 13, and 14. These figures show that, if 10 mA (rms) is taken as the necessary current level for cable localization by SCARAB, and if 200 mA (rms) of 25-Hz center-conductor current is applied at the shore terminal of an unpowered unarmored SG system, then breaks as far offshore as 170 nmi can be found (with this value reduced to 125 nmi for armored cable, because of the shielding effect of the armor wires). These ranges can be increased if dc power can be applied to the repeaters, thereby reducing their impedance. If less current is applied at the shore point (as on earlier cable systems, which can tolerate only 100 mA, or on operational systems, which can tolerate only 20 mA), the maximum effective range of SCARAB's cable-locating system is correspondingly reduced.

These operational ranges could be increased for future cable systems by designing the repeaters with a separate low-impedance path for the 25-Hz signal to avoid the constraints and losses imposed on this signal as it traverses the dc power path.

REFERENCES

1. H. A. Baxter and R. E. Mueser, "The Development of Ocean-Cable Plows," *IEEE Trans. Commun. Technol., COM-19*, No. 6 (December 1971), pp. 1233-41.

2. G. S. Cobb, "Sea Plow IV—An Underseas Vehicle for Burying Ocean Cable," in *OCEANS '76*, New York: Institute of Electrical and Electronics Engineers, 1976, pp. 19B-1 to 19B-6.
3. G. R. Barnard, Applied Research Laboratory, University of Texas, Austin, Texas, personal communication, 1975.
4. J. R. Wait, "Electromagnetic Theory of the Loosely Braided Coaxial Cable: Part I," *IEEE Trans. Microwave Theory Technol.*, *MTT-24*, No. 9 (September 1976), pp. 547-53.
5. C. D. Anderson, Bell Laboratories, Holmdel, New Jersey, personal communication, 1977.

CONTRIBUTORS TO THIS ISSUE

Václav E. Beneš, A.B., 1950, Harvard College; M.A. and Ph.D., 1953, Princeton University; Bell Laboratories, 1953—. Mr. Beneš has pursued mathematical research in traffic theory, stochastic processes, frequency modulation, combinatorics, servomechanisms, stochastic control, and filtering. In 1959–60 he was visiting lecturer in mathematics at Dartmouth College. In 1971 he taught stochastic processes at SUNY Buffalo, and in 1971–72 he was Visiting MacKay Lecturer in electrical engineering at the University of California in Berkeley. He is the author of *General Stochastic Processes in the Theory of Queues* (Addison-Wesley, 1963), and of *Mathematical Theory of Connecting Networks and Telephone Traffic* (Academic Press, 1965). Member, American Mathematical Society, Association for Symbolic Logic, Institute of Mathematical Statistics, SIAM, Mathematical Association of America, IEEE, Phi Beta Kappa.

A. Carnevale, B.S. (Physics), 1960, Fairleigh Dickinson University; Bell Laboratories, 1969—. At Bell Laboratories, Mr. Carnevale has been engaged in work on nuclear magnetic resonance, electron paramagnetic resonance, and computer software. For the last two years, his work has been devoted to fiber optics.

James W. Fleming, B.S. (Ceramic Engineering), 1970, M.S. (Ceramic Engineering), 1971, University of Missouri-Rolla; University of Missouri-Rolla, 1971–1972; Bell Laboratories, 1972—. At the University of Missouri-Rolla, Mr. Fleming continued work on internal friction in mixed alkali glasses, the subject of his master's thesis. At Bell Laboratories, he initially worked on preparation and properties of positive temperature coefficient barium titanate. By 1974 he was actively involved in lightguide glass preparation and analysis of optical dispersion in lightguides. He is currently interested in the preparation of high rate optical waveguides and novel glasses for optical communication systems. Member, American Ceramic Society, Optical Society of America. He was also a subeditor and session chairman for the American Ceramic Society.

S. H. Francis, B.A. (Physics), 1964, Yale University; A.M. (Physics), 1968, Ph.D. (Physics), 1970, Harvard University; Bell Laboratories, 1970—. Mr. Francis worked initially in research on atmospheric and

ionospheric dynamics, in conjunction with the Safeguard project. This research was followed by work on a cable-finding system for use aboard a remotely piloted submersible, for which Mr. Francis' responsibilities included system design verification, and software design and testing. Subsequent research in turbulent flow noise in towed sonar arrays was followed in 1978 by his present job as supervisor of the sonar systems analysis group, engaged in the development of towed sonar systems for Navy applications.

Nuggehally S. Jayant, B.Sc. (Physics and Mathematics), 1962, Mysore University; B.E., 1965, and Ph.D. (Electrical Communication Engineering), 1970, Indian Institute of Science, Bangalore; Research Associate at Stanford University, 1967-1968; Bell Laboratories, 1968—. Mr. Jayant was a Visiting Scientist at the Indian Institute of Science in 1972 and 1975. He has worked in the field of digital coding and transmission of waveforms, with special reference to robust speech communications. He is also editor of the IEEE Reprint Book, *Waveform Quantization and Coding*.

H. R. Lunde, B.S.M.E., 1969, Drexel University; M.S.M.E., Columbia University, 1971; Bell Laboratories, 1969—. Mr. Lunde has worked on the physical, thermal, and vibrational design of undersea repeaters. Since 1975 he has also been involved in the design of underwater work systems used to detect, repair, and maintain undersea cables.

Un-Chul Paek, B.S. (Engineering), 1957, Korea Merchant Marine Academy, Korea; M.S., 1965, Ph.D., 1969, University of California, Berkeley; Western Electric, 1969—. At the Western Electric Engineering Research Center, Princeton, N. J., Mr. Paek has been engaged primarily in research on laser material interaction phenomena and fiber optics. Member, Optical Society of America, American Ceramic Society, Sigma Xi.

George E. Peterson, B.S. (Physics), 1956, Ph.D. (Solid State Physics), 1961, University of Pittsburgh; Bell Laboratories, 1961—. At Bell Laboratories, Mr. Peterson initially was engaged in studies on low-noise amplifiers. He then studied laser crystals, nonlinear optic materials, and glass structure, and recently propagation of light in optical fibers. Member, American Physical Society, American Ceramic Society, Society for Glass Technology, and the American Crystallographic Association.

William Pferd, B.S.M.E., Rutgers University; M.S.M.E., New Jersey Institute of Technology; U.S. Army Air Force Intelligence, 1942-1945; Bell Laboratories, 1947—. During his Bell Laboratories career, Mr. Pferd has worked in the areas of customer products, systems engineering, and computer application. He participated in the development of ringers, rotary dials, card dialers, connectors, automatic sets, and public telephones. While at the Indianapolis Laboratory, his major responsibility was the development of the "single slot" coin telephone. On returning to the Whippany Laboratory, he led engineering studies in criteria for hardened facilities, equipment standards, central office planning, and building investment tax credit. He is currently Head of the Office Planning Department within the Building and Energy Systems Laboratory, responsible for the development of computer systems to facilitate the planning and engineering of Bell Systems buildings. Senior Member, IEEE; Member, ACM, SPIE.

G. A. Reinold, B.S.M.E., 1962, Union College; M.S.M.E., 1964, New York University; Bell Laboratories, 1962—. Mr. Reinold was first involved in the development of undersea cable installation and repair techniques. He later worked on methods for installing cables underground with minimal soil disturbance. He then participated in ocean cable design for the military. In 1969, he became supervisor of the undersea apparatus mechanical design group, where he was responsible for the physical design of the SG repeater, and the cable location system on the SCARAB submersible. He is currently responsible for the physical design of the undersea lightguide repeater (SL), and the complete refurbishment of the SCARAB submersible.

Gerhard C. Stocker, A.A.S. (Electrical Engineering) 1964, Mohawk Valley Community College; Bell Laboratories, 1970—. Mr. Stocker has worked on nuclear shock and vibration protection of military equipment. He also designed earthquake resistant cable rack systems for New Equipment-Building Systems (NEBS) buildings. Currently, as a member of the Office Standards and Data Conversion Group, he is responsible for the fabrication methods for optical fiber arrays and for the development of an experimental automatic scanning digitizer that uses optical fibers.

T. E. Talpey, B.E.E., 1946, Cornell University; M.S.E.E., 1948, Ph.D., 1954, University of Michigan; Doctorat d'Université, 1952, University of Grenoble, France; Bell Laboratories, 1953—. Mr. Talpey first worked on studies of noise in electron tubes and later became a

supervisor (1955–1969) in the electron tube development department with responsibilities that included high figure-of-merit amplifier tubes, long-life submarine cable tubes, and traveling wave tubes. On a leave of absence from 1960 to 1963, he worked as a research associate for Cornell University at the Arecibo Ionospheric Observatory in Puerto Rico. In 1963 he worked at Bell Laboratories on the design of a large phased-array radar receiver for the Nike system and in 1966 assumed his present position in ocean systems studies, where his work currently involves underwater acoustic signal propagation studies.

Uzi Timor, B.S. (Electrical Engineering), 1957, M.S. (Electrical Engineering), 1963, Technion, Israel Institute of Technology, Haifa, Israel; Ph.D., 1969, University of California, Berkeley; Armament Development Authority/Israel Ministry of Defence, 1959–1966, 1972—; Member of the Technical Staff, Jet Propulsion Laboratory, Pasadena, California, 1969–1972; Adjunct Faculty (Electrical Engineering), Technion, Israel Institute of Technology, Haifa, Israel, 1972–1980; Bell Laboratories, 1979–1980. At Jet Propulsion Laboratory, Mr. Timor was engaged in research in digital communications and coding. As a consultant at Bell Laboratories, he worked on digital communications and mobile radio. Member, IEEE, Eta Kappa Nu.

Woodson D. Wynn, M.S.E.E., 1961, Ph.D., 1965, Georgia Institute of Technology; Assistant Professor, Georgia Institute of Technology, 1964–1966; Bellcomm, Inc., 1966–1971; Bell Laboratories, 1971—. Mr. Wynn has been concerned with the theoretical analysis and practical solution of problems in signal design, transmission, processing and recording. His current interest is in the development and application of statistical communication and adaptive control theory. Member, IEEE, Tau Beta Pi, Pi Mu Epsilon, Eta Kappa Nu, and Sigma Xi.

PAPERS BY BELL LABORATORIES AUTHORS

PHYSICAL SCIENCES

- Amorphous Concentrated Spin Glasses: MnX(X=Ge, C, Si-Te).** J. J. Hauser, *Phys Rev*, 22 (September 1, 1980), pp 2554-9.
- Avalanche-Induced Hot Electron Injection and Trapping in Gate Oxides on p-Si.** A. K. Sinha, W. S. Lindenberger, W. D. Powell, and E. I. Povolonis, *J Electrochem Soc*, 127 (1980), pp 2046-9.
- Block Copolymer Theory. 6. Cylindrical Domains.** E. Helfand and Z. R. Wasserman, *Macromolecules*, 13 (July-August 1980), pp 994-8.
- Convection in Rectangular Cavities with Differentially Heated End Walls.** P. G. Simpkins and T. D. Dudderar, *Europhys Conf Abstr Physico Chemical Hydrodynamics* 3, 3F (April 1980), pp 268-73.
- Crystalline Transformations in Spherulites of Poly(vinylidene Fluoride).** A. J. Lovinger, *Polymer*, 21 (1980), pp 1317-22.
- Effects of Atomic Motion on Degenerate 4-Wave Mixing.** P. F. Liao and L. M. Humphrey, *Proceedings of the International Conference on Laser '79*, edited by V. J. Corcoran, McLean, Virginia: STS Press, 1980, pp 473-4.
- The Effects of Velocity-Changing Collisions on Two-Photon and Stepwise Absorption Spectroscopic Lineshapes.** P. F. Liao, J. E. Bjorkholm, and P. R. Berman, *Phys Rev A*, 21 (June 1980), pp 1927-38.
- Exploratory Data Analysis in the Geophysical Sciences.** B. Kleiner and T. E. Graedel, *Rev Geophys Space Phys*, 18, No. 3 (August 1980), pp 699-717.
- Hydrogen in Semi-Insulating Polycrystalline Silicon Films.** W. R. Knolte, H. R. Maxwell, Jr., and R. E. Benenson, *J Appl Phys*, 51 (August 1980), pp 4385-90.
- Hyperfine-Induced Singlet-Triplet Mixing in ^3He .** P. F. Liao, R. R. Freeman, R. Panock, and L. M. Humphrey, *Optics Commun*, 34 (August 1980), pp 195-8.
- The Influence of Structural Disorder on the Exciton Spectra of $\text{K}^+\text{CN}_x\text{-Cl}_{1-x}$.** M. A. Bosch, *J Phys C*, 25 (1980), pp L671-5.
- A Model of SIPOS Deposition Based on Infrared Spectroscopic Analysis.** W. R. Knolte and H. R. Maxwell, Jr., *J Electrochem Soc*, 127 (October 1980), pp 2254-9.
- Observation by NMR of the Formation of Localized Electronic States in an Ionic Alloy.** R. Dupree, D. J. Kirby, W. Freyland, and W. W. Warren, *Phys Rev Lett*, 45 (July 14, 1980), pp 130-3.
- Oxidation of Tantalum Silicide on Polycrystalline Silicon.** S. P. Jurarka, D. B. Fraser, W. J. Lindenberger, and A. K. Sinha, *J Appl Phys*, 51 (June 1980), pp 3241-5.
- Piezoelectric $\text{Cs}_2\text{S}_2\text{O}_6$: Room Temperature Crystal Structure.** R. Liminga, S. C. Abrahams, and J. L. Bernstein, *J Chem Phys*, 73 (August 1, 1980), pp 1432-8.
- Surface Enhanced Raman Scattering in Multilayer Film Structures (Abstract).** C. A. Murray, D. L. Allarn, and M. R. Rhinewine, *Proceedings of the VIIth International Conference on Raman Spectroscopy*, August 4-8, 1980, edited by W. F. Murphy, Amsterdam: North-Holland, 1980, p 406.

MATHEMATICS

- Baxter Permutations Rise Again.** C. L. Mallows, *J Combinatorial Theor, Ser A*, 27 (November 1979), pp 394-6.
- Indirect Rate Distortion Problems.** H. S. Witsenhausen, *IEEE Trans Inform Theor*, IT-26 (September 1980), pp 518-21.
- A Weighted Overlap-Add Method of Short-Time Fourier Analysis/Synthesis.** R. E. Crochiere, *IEEE Trans Acoustics Speech and Signal Process*, ASSP-28, No. 1 (February 1980), pp 99-102.

COMPUTING

- C on the Tandem 16.** A. M. Usas, *Proc Tandem Users Group Convention*, San Diego, California, September 15-17, 1980.

Information Theory in Digital Testing—A New Approach to Functional Test Pattern Generation. V. D. Agrawal, IEEE Int Conf on Circuits and Computers (October 1980), pp 928-31.

Logic and Fault Simulations of the DSP, A VLSI Digital Signal Processor. I. I. Edumiati and R. N. Gadenz, Proc 1980 Int Conf on Circuits and Computers, ICCO 80 (October 1980), pp 948-52.

A Model for a Tandem Software Development System. A. M. Usas, Proc Tandem Users Group Convention, San Diego, California, September 15-17, 1980.

A Modified Adaptive Transform Coding Scheme With Post-Processing Enhancement. J. M. Tribolet and R. E. Crochiere, Proc IEEE Int Conf on Acoustics Speech and Signal Process (April 1980), pp 336-9.

Network Management: Surveillance and Control of a Modern Telecommunications Network. D. A. Kettler, IEEE Int Conf on Circuits and Computers, 2 (October 1980), pp 834-7.

On Retrieval From a Small Version of a Large Data Base. A. Walker, Proc Sixth Int Conf Very Large Data Bases, IEEE Conf Proc (October 1980), pp 47-54.

Problems in Physical Property Data Retrieval. D. T. Hawkins, J Chem Inf Comput Sci, 20 (August 1980), pp 143-5.

ENGINEERING

Administration of Pair Gain Systems. G. R. Boyer and M. A. Schwartz, Natl Telecommun Conf 1980, 1 (December 1, 1980), pp 12.2.2-12.2.6.

Efficient Waveguide Electro-Optic TE \leftrightarrow TM Mode Converter/Wavelength Filter. R. C. Alferness, Appl Phys Lett, 36 (April 1979), pp 513-5.

ESD Damage from Triboelectrically Charged IC Pins. P. R. Bossard, R. G. Chemelli, and B. A. Unger, *Electrically Overstress/Electrostatic Discharge Symposium Proceedings, 1980*, New York: Reliability Analysis Center RADC/RBRAC Griffiss AFB.

A Fast-Switching Low-Loss, 12-GHz Microstrip 4-PSK Path Length Modulator. B. Glance and N. Amitay, IEEE Trans Commun, COM-28, No. 10 (October 1980), pp 1824-8.

A Fast Low-Loss Low-Drive 14-GHz Microstrip PIN Phase Shifter. B. Glance, IEEE Trans Microwave Theor Tech, MTT-28, No. 6 (June 1980), pp 669-71.

Integrated Services Digital Network. T. L. Powers, Signal, 34, No. 10 (August 1980), pp 45-6.

Intermag-80 Session Chairman's Summary. R. D. Pierce, IEEE Magnetics Society Newsletter, 16 (Summer 1980), p 1.

MOS Compatibility of High-Conductivity TaSi₂/n⁺ Poly-Si Gates. A. K. Sinha, W. S. Lindenberger, D. B. Fraser, S. P. Murarka, and E. N. Fuls, IEEE Trans Electron Devices, ED-27 (1980), pp 1425-30.

Numerical Analysis of the Resistance of Interference Fit Pin Connections. E. Guancial, IEEE Trans Components, Hybrids, and Manufact Tech, CHMT-3 (September 1980), pp 402-7.

Optimization of Multimode Fiber Bandwidth Via Differential Group Delay Analysis. M. J. Buckler, J. W. Shiever, and F. P. Partus, Conf Proc—Sixth European Conf on Optical Commun (September 15, 1980), pp 33-6.

Picosecond Optical Sampling Technique for Measuring the Speed of Fast Electro-Optic Switch/Modulators. R. C. Alferness, N. P. Economou, and L. L. Buhl, Appl Phys Lett, 37 (October 1, 1980), pp 597-9.

Refractory Silicides for Integrated Circuits. S. P. Murarka, J Vac Sci Technol, 17 (July-August 1980), pp 775-92.

Solid State Microwave Power Amplifiers Used in the Bell System. J. W. Gewartowski, Microwave J, 23 (September 1980), pp 68 and 72.

Stress Distributions Around an Interference-Fit Pin Convection in a Plated-Thru-Hole. R. P. Goel and E. Guancial, IEEE Trans Components, Hybrids, and Manufact Tech, CHMT-3 (September 1980), pp 392-402.

SOCIAL AND LIFE SCIENCES

An Interpretation of the Log Likelihood Ratio as a Measure of Waveform Coder Performance. R. E. Crochiere, J. M. Tribolet, and L. R. Rabiner, IEEE Trans Acoustics, Speech and Signal Process, ASSP-28, No. 3 (June 1980), pp 318-23.

Modern Methods of Investigation in Speech Production. O. Fujimura, *PHONETICA*, 37, Nos. 1, 2 (1980), pp 38-54.

On the Measurement of Waveform Coder Distortion Using the Log Likelihood Ratio. R. E. Crochiere, J. M. Tribolet, and L. R. Rabiner, *Proc 1980 Int Conf on Acoustics, Speech and Signal Process* (April 1980), pp 340-3.

Two Acoustical Inverse Problems in Speech and Hearing. M. M. Sondhi, a chapter in *Mathematical Methods and Application of Scattering Theory*, New York: Springer, 1981, pp 290-300.

CONTENTS, MAY-JUNE 1981

Dispersionless Single-Mode Lightguides With α Index Profiles

U. C. Paek, G. E. Peterson, and A. Carnevale

A Class of Closed Markovian Queuing Networks: Integral Representations, Asymptotic Expansion, and Generalizations

J. McKenna, D. Mitra, and K. G. Ramakrishnan

Time-Frequency Multiplexing (TFM) of Two NTSC Color TV Signals—Simulation Results

B. G. Haskell

Overflow Models for Dimension[®] PBX Feature Packages

L. Kaufman, J. B. Seery, and J. A. Morrison

Impact of Forecast Uncertainty on Feeder-Cable Sizing

N. H. Noe

The Effects of Misclassification Error on the Estimation of Several Population Proportions

J. D. Healy

Adaptive Post-Filtering of ADPCM Speech

N. S. Jayant

Variable Bandwidth Adaptive Delta Modulation

J. O. Smith and J. B. Allen

A Two-Pass Pattern-Recognition Approach to Isolated Word Recognition

L. R. Rabiner and J. G. Wilpon

Fast Decryption Algorithm for the Knapsack Cryptographic System

P. S. Henry

CONTENTS, MAY-JUNE 1981

Dispersionless Single-Mode Lightguides With α Index Profiles

U. C. Paek, G. E. Peterson, and A. Carnevale

A Class of Closed Markovian Queuing Networks: Integral Representations, Asymptotic Expansion, and Generalizations

J. McKenna, D. Mitra, and K. G. Ramakrishnan

Time-Frequency Multiplexing (TFM) of Two NTSC Color TV Signals—Simulation Results

B. G. Haskell

Overflow Models for Dimension[®] PBX Feature Packages

L. Kaufman, J. B. Seery, and J. A. Morrison

Impact of Forecast Uncertainty on Feeder-Cable Sizing

N. H. Noe

The Effects of Misclassification Error on the Estimation of Several Population Proportions

J. D. Healy

Adaptive Post-Filtering of ADPCM Speech

N. S. Jayant

Variable Bandwidth Adaptive Delta Modulation

J. O. Smith and J. B. Allen

A Two-Pass Pattern-Recognition Approach to Isolated Word Recognition

L. R. Rabiner and J. G. Wilpon

Fast Decryption Algorithm for the Knapsack Cryptographic System

P. S. Henry

THE BELL SYSTEM 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.

