

THE JULY-AUGUST 1976
VOL. 55 NO. 6
BELL SYSTEM
TECHNICAL JOURNAL

| | | |
|--|--|------------|
| J. G. Evans, F. W. Kerfoot, and R. L. Nichols | Automated Network Analyzers for the 0.9- to 12.4-GHz Range | 691 |
| A. E. Rosenberg | Evaluation of an Automatic Speaker-Verification System Over Telephone Lines | 723 |
| P. Balaban | Statistical Evaluation of the Error Rate of the Fiber- guide Repeater Using Importance Sampling | 745 |
| V. Ramaswamy and R. D. Standley | A Phased, Optical, Coupler- Pair Switch | 767 |
| D. Marcuse | Mode Mixing With Reduced Losses in Parabolic-Index Fibers | 777 |
| J. A. Lewis | The Boundary Layer in a Con- centrated, Multicomponent Electrolyte | 803 |
| | Contributors to This Issue | 825 |
| | Abstracts of Bell System Papers | 827 |

THE BELL SYSTEM TECHNICAL JOURNAL

ADVISORY BOARD

D. E. PROCKNOW, *President,*
Western Electric Company, Incorporated

W. O. BAKER, *President,*
Bell Telephone Laboratories, Incorporated

W. L. LINDHOLM, *President,*
American Telephone and Telegraph Company

EDITORIAL COMMITTEE

D. GILLETTE, *Chairman*

W. S. BOYLE

I. DORROS

A. G. CHYNOWETH

W. R. HARLIN

C. L. COYNE

C. B. SHARP

T. H. CROWLEY

B. E. STRASSER

W. A. DEPP

I. WELBER

EDITORIAL STAFF

L. A. HOWARD, JR., *Editor*

P. WHEELER, *Associate Editor*

J. M. SZNYTER, *Art and Production Editor*

F. J. SCHWETJE, *Circulation*

THE BELL SYSTEM TECHNICAL JOURNAL is published ten times a year by the American Telephone and Telegraph Company, J. D. deButts, Chairman and Chief Executive Officer; W. L. Lindholm, President; C. L. Brown, Executive Vice President and Chief Financial Officer; F. A. Hutson, Jr., Secretary. Editorial enquiries should be addressed to the Editor, The Bell System Technical Journal, Bell Laboratories, Whippany, N. J. 07981. Checks for subscriptions should be made payable to American Telephone and Telegraph Company and should be addressed to the Treasury Department, Room 1038, 195 Broadway, New York, N. Y. 10007. Subscriptions \$15.00 per year; single copies \$1.75 each. Foreign postage \$1.00 per year; 15 cents per copy. Printed in U.S.A.

THE BELL SYSTEM TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING
ASPECTS OF ELECTRICAL COMMUNICATION

Volume 55

July–August 1976

Number 6

Copyright © 1976, American Telephone and Telegraph Company. Printed in U.S.A.

Automated Network Analyzers for the 0.9- to 12.4-GHz Range

By J. G. EVANS, F. W. KERFOOT, and R. L. NICHOLS

(Manuscript received December 22, 1975)

Two highly accurate automated network analyzers have been developed to cover the frequency ranges of 0.9 to 4.2 GHz and 3.7 to 12.4 GHz. In addition to the directly measured quantities of insertion loss and phase, the analyzers provide insertion delay, impedance, and two-port parameters as derived quantities.

The functional architecture of the systems is described along with some of the novel techniques used to obtain the high signal-to-noise ratios and system stabilities needed for the precise measurements of loss, phase, and delay. General descriptions of these techniques are given along with some of the design considerations.

Experimental results that demonstrate measurement accuracy are presented for one-port and two-port devices using both coaxial and waveguide connectors.

I. INTRODUCTION

New automated measuring instrumentation has been developed for making network and component characterizations in the microwave frequency range from 0.9 to 12.4 GHz. The development was stimulated by characterization requirements growing out of the design of new communication systems and the redesign of older systems. These systems are achieving greater capacity due to advances in technology, both in the areas of improved analysis and improved components. These advances have led to requirements on measurements at micro-

wave frequencies that heretofore had been needed only at lower frequencies.

This new instrumentation was implemented as an extension to the Computer Operated Transmission Measuring Sets (cotms)¹ operating in the Transmission Technology Laboratories at Bell Laboratories. This implementation avoids the duplication of costly common equipment and allows the utilization of many of the proven cotms measurement techniques.

The microwave instrumentation, in concert with the cotms system, is intended for the characterization of linear two-port networks, either active or passive. Measurement terminals are provided for networks with either coaxial or waveguide terminals. Provision is made for complete four-parameter characterization of network elements such as transistors (e.g., H parameters, Y parameters).

The full 0.9- to 12.4-GHz frequency range is covered by two microwave systems of very similar design. The lower frequency system, completed in November 1972, covers the range 0.9 to 4.2 GHz.* Its first applications were in characterizing solid state components. The second instrument, completed in May 1971, covers the range 3.7 to 12.4 GHz,[†] spanning the most widely used common-carrier frequency bands. It found immediate use in characterizing filter elements, and later complete networks for the TN1 radio system.

This paper reviews the measurement capabilities and the overall design of both measuring systems. Several system-design considerations of special interest and error mechanisms that limit accuracy are described. This paper concludes by presenting experimental results that illustrate the applications of the measurement systems and demonstrate their high measurement accuracy.

II. MEASUREMENT SET CHARACTERISTICS

2.1 General

The two microwave measuring sets which cover the frequency ranges of 0.9 to 4.2 GHz and 3.7 to 12.4 GHz are fully automated network analyzers that were specifically designed for high-accuracy measurements. The sets are extensions of the 50-Hz to 1-GHz Computer Operated Transmission Measuring Set (cotms). The cotms system provides input frequencies for the microwave signal source, IF detection circuitry, and the control computer. Extensive use is made of the computation capability to enhance measurement convenience and accuracy. Measurements are made under computer direction and the

* For brevity this system will be referred to as the 1- to 4-GHz measuring set.

† For brevity this system will be referred to as the 4- to 12-GHz measuring set.

calculations for derived parameters are made in real time. Accuracy is enhanced by the correction of known systematic errors and the averaging of random errors. The resulting measurement speed and accuracy are significantly greater than those previously achieved with manual techniques.

2.2 Measurements

The basic quantities measured by the test sets are loss (or gain) and phase. Insertion loss and insertion phase of a network are determined by subtracting the loss and phase obtained with the test set connectors directly connected from the loss and phase measured with the network inserted between the connectors. Special care was taken to achieve an uncertainty in test-port impedance of less than 2 percent on both the coaxial and waveguide test ports in order to minimize impedance mismatch errors.

Envelope delay, τ , is a derived parameter approximated from insertion-phase measurements at two closely spaced frequencies.

$$\tau = \frac{\partial\theta}{360\partial F} \doteq \frac{\theta_2 - \theta_1}{360(F_2 - F_1)} \text{ for } \theta \text{ in degrees.} \quad (1)$$

The return loss or reflection coefficient of a single-port device under test (DUT) is also a derived parameter.² The measurement is made with the hardware switched (for example, see the S_{11} and S_{22} configurations in Fig. 4) to couplers on the coaxial test ports and magic tee hybrids on the waveguide test ports. Three measurements with independent reflection calibration standards are used to substantially reduce errors associated with impedance mismatches and poor coupler directivity.²

The four-parameter characterization of a DUT entails a complicated set of calibration measurements and switching of the hardware into four independent measurement configurations. The four measurements are closely related to the scattering (S) parameters² and are easily transformed to other parameter representations. Determination of the S parameters entails the measurement of three reflection coefficient standards on each test port and a transmission standard connecting the test ports. Automatic switching provides the proper interconnections for two transmission and two reflection measurements corresponding to S_{21} , S_{12} , S_{11} , and S_{22} . The measurements are followed by a calculation to convert S parameters that are referenced to test set impedances to S parameters referenced entirely to the calibration standards, thereby eliminating any impedance-mismatch errors associated with imperfect test set impedances.³ More details on the hardware used for the above measurements are presented in later sections.

2.3 Test set performance

The uncertainties of measurements made on both test sets are dependent upon random errors, uncorrected systematic test set errors, and uncertainties in the calibration standards. The random errors can be reduced by signal averaging. For a typical measurement, the user can specify an upper bound on the numerically computed variance of the mean of every directly measured quantity. Averaging slows down the measurement rate. The measurement rate is five frequency points per second taking a single observation. The rate decreases to one frequency point per second if 32 observations are averaged. The typical standard deviation of a single observation at low insertion loss is 0.0017 dB and 0.01 degree.

Table I gives a representative list of the quantities measured and their respective resolutions. The resolution, or differential accuracy, is important in the characterizations of the small loss and phase variations in the passband of a filter.

Systematic test set errors may give rise to uncertainties in absolute characterizations that are larger than the differential uncertainties. Some of these errors, such as detection nonlinearity, are functions of many parameters and are not practical to remove by computation.

Table I — Test set performance

| Characteristic | Range | Resolution |
|---|---|---------------------|
| Frequency | 0.9 to 4.2 GHz; 3.7 to 12.4 GHz in two test sets. | 1 Hz |
| Test rate | Five measurements per second with no averaging for insertion loss and phase | |
| Insertion measurements* | | |
| Loss | 0 to 80 dB | 0.002 dB |
| Phase | 0° to 360° | 0.01° |
| Delay | (0 to 1980)/ ΔF ns [†] | 0.06/ ΔF ns |
| Reflection measurements* | | |
| Reflection coefficient | 0 to 1 | 0.02% |
| Two-port measurements* | | |
| Reflection parameters (S_{11} , S_{22}) | 0 to 1 | 0.02% |
| Transmission parameters (S_{12} , S_{21}) | 0 to 80 dB 0° to 360° | 0.002 dB 0.02° |
| Incident power level | -15 to -55 dBm (0.9 to 4.2 GHz) -5 to -30 dBm (3.7 to 12.4 GHz) | 1.0 dB 2.0 dB |

* Loss and phase random uncertainties increase as loss increases, decreasing the signal-to-noise ratio. The variance of a single measurement increases as follows:

$$\sigma_{\text{loss}} = 0.0017 \times 10^{L/20} \text{ dB}$$

$$\sigma_{\text{phase}} = 0.01 \times 10^{L/20} \text{ deg.}$$

L is the loss in dB of the DUT. The random uncertainty can be reduced by averaging.

[†] ΔF is the frequency difference in eq. (1) expressed in MHz.

They have been reduced by careful system design to a magnitude typically less than the relationships given below when the best available calibration standards are used.*

$$|\Delta S_{jj}| \leq 0.0015(1 + F/12)|1 - S_{jj}^2 - S_{ij} \cdot S_{ji}| \quad (2)$$

$$\left| \frac{\Delta S_{ij}}{S_{ij}} \right| \leq 0.0001(1 + F) + 0.0001|S_{ij}|^{-1} + K(1 + F/12)(|S_{11}| + |S_{22}|) \quad (3)$$

$K = 0.01$ for a simple insertion measurement

$K = 0.002$ for mismatch errors after computer correction.

The parameter (F) is the measurement frequency expressed in gigahertz. The family of parameters (S_{ij}) are the scattering parameters of the device under test.

Equation (2) can be applied to a one-port reflection coefficient characterization by setting $|S_{12} \cdot S_{21}|$ equal to zero. Uncertainty in the characterization of the standards is the dominant contributor to this partially empirical relationship.

Equation (3) is composed of a crosstalk uncertainty, an impedance mismatch uncertainty, and a residual uncertainty from many other mechanisms. When a simple insertion measurement is made without compensation for the deviation of the test set impedances from nominal values, then the coefficient on the last term must be increased from 0.002 to 0.01. A more comprehensive discussion of the error mechanisms is given in Section V.

III. BLOCK DIAGRAM DESCRIPTION

3.1 Overall description

A simplified block diagram illustrating the relationship between cotms and the broadband microwave test sets is shown in Fig. 1. The test sets are quite similar in system design. They consist of source, comparison unit, heterodyne converter, and interface circuitry.

The cotms source supplies a computer-controlled low-frequency input signal to source circuitry in each test set. The microwave measurement frequency is developed in the source circuitry by heterodyning and frequency multiplication of the input signal. Signal purity of the output signal is enhanced by computer-controlled YIG filters.

The comparison circuitry interconnects the source and detector via a switching network that rapidly interchanges an unknown path, containing the DUT, and a standard path. Full two-port characteri-

* Equation (2) is valid if the uncertainty in the reflection coefficients of the impedance standards is less than 0.001. Equation (3) is valid if the uncertainty in S_{12} and S_{21} of the transmission standard is less than 0.0001.

zation of the unknown is provided by additional switching circuitry in the unknown path.

The heterodyne circuit provides a fixed intermediate frequency input signal to the computer-controlled cotms detector. This signal contains information on the differences in amplitude and phase between the unknown and standard paths. The heterodyne circuit also provides a phase-reference input signal to the cotms detector at the same fixed intermediate frequency.

The interface provides digital control of the test system circuitry by the cotms computer and informs the computer of circuit status. All digital lines between cotms and the test systems pass through the interface. Manual control of each line is also provided by a toggle switch for maintenance and other special purposes.

3.2 Source

The test frequency in each applique is developed by heterodyning a signal at frequency f_1 received from cotms with a signal at frequency f_2 (see Fig. 2). The signal f_2 is derived from reference signals coherent with the synthesizer-derived signal f_1 . The signal at frequency $f_1 + f_2$ is applied to the input of a comb generator which produces a harmonic

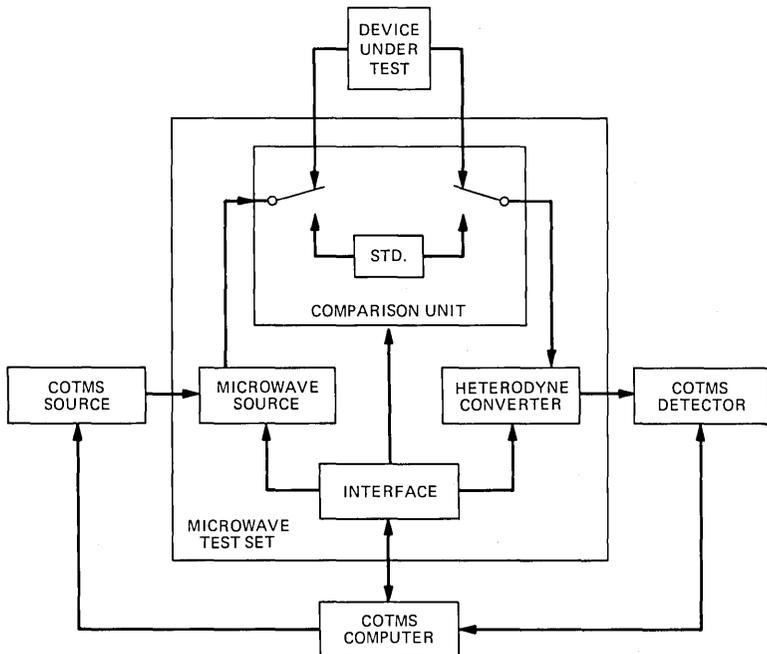
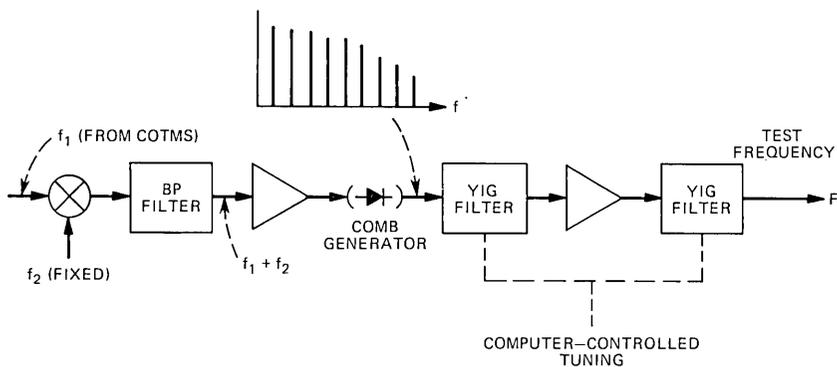


Fig. 1—Simplified block diagram of the cotms-microwave measuring systems.



1-4 GHz MEASURING SET : $f_2 = 198 \text{ MHz}$, $22 \text{ MHz} \leq f_1 \leq 77 \text{ MHz}$
 4-12 GHz MEASURING SET : $f_2 = 825 \text{ MHz}$, $75 \text{ MHz} \leq f_1 \leq 300 \text{ MHz}$

Fig. 2—Block diagram of the source circuitry.

output spectrum extending beyond 13 GHz. A single sinusoidal harmonic is selected by four stages of computer-controlled YIG filtering, producing a test signal at frequency $F = N(f_1 + f_2)$. This weak signal is amplified with solid state amplifiers in the 1- to 4-GHz test set and with a traveling-wave-tube amplifier in the 4- to 12-GHz test set.

3.3 Heterodyne section

In the heterodyne section (see Fig. 3), the microwave source signal (F) is split to excite the DUT, generate a reference-phase signal, and generate the local-oscillator signal. The local-oscillator (LO) signal and the signals at frequency F are combined to form fixed intermediate-frequency signals (IF) which are sent to the COTMS detection circuitry. These IF signals are processed by COTMS to extract phase and amplitude information.

The source signal applied to the upper path shown in Fig. 3 excites the DUT in the comparison circuit. The DUT modulates the amplitude and phase of this signal, which are ultimately detected in COTMS. To be compatible with the COTMS frequency range, it is necessary to heterodyne this signal with an LO signal to generate a lower-frequency IF signal.

A reference signal is necessary to make a phase measurement. This reference signal is generated in the lower path in Fig. 3 in an almost identical fashion as that described above.

The LO signal common to both IF mixers is offset from the measurement frequency by 0.495 GHz. This signal is generated by the mixer

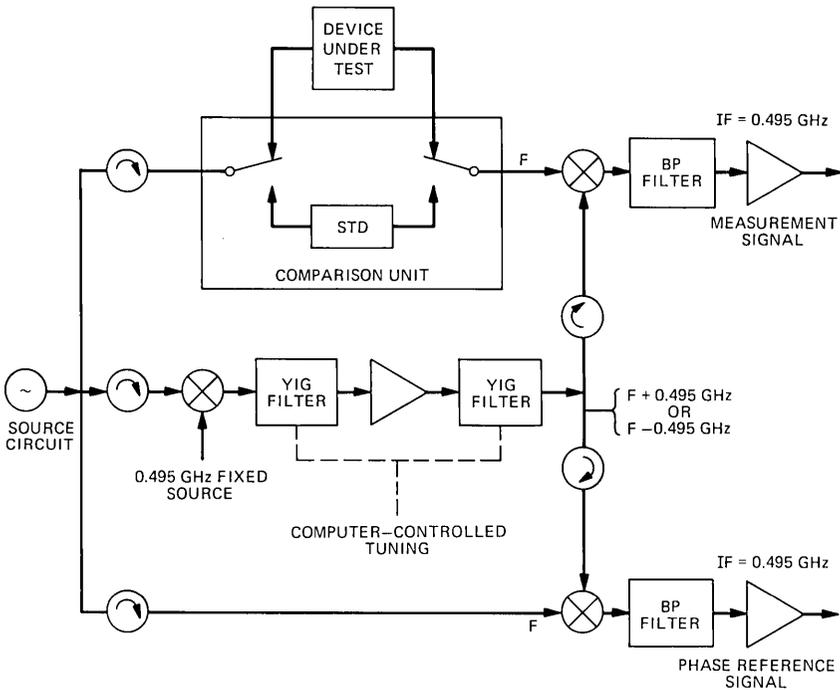


Fig. 3—Block diagram of the heterodyne circuitry.

circuit shown in the middle path of Fig. 3. The fixed 0.495-GHz signal is derived from reference signals coherent with the synthesizer-based test signal F . The overall LO path bandwidth is minimized by switching from the $F + 0.495$ GHz sideband to the $F - 0.495$ GHz sideband at each test set's midband point. The proper sideband is selected by four stages of YIG filtering. This filtering attenuates all unwanted tones, especially the F tone, by 140 dB with respect to the wanted sideband tone. This technique for the generation of the LO signal eliminates the need for a second microwave source and achieves a coherence between the F and the $F \pm 0.495$ -GHz signals that is important in reducing phase noise. This latter consideration is discussed further in Section 4.4.

3.4 Comparison circuit

The comparison circuit is a switching network which excites the DUT at various test ports and performs a rapid comparison switching to an internal transfer standard. In each test set, there are test ports with a reflection purity exceeding 40 dB used for insertion measurements and a four-configuration switching network for the measurement

of the four S -parameters of a two-port device. This latter network is illustrated in Fig. 4. A computer-directed calibration procedure provides the data for the automatic removal of errors arising from crosstalk and impedance deviations. Additional test ports are provided in the 4- to 12-GHz applique for the measurement of reflection coefficients in three waveguide sizes.

The rapid comparison switch allows measurements in a differential mode to increase accuracy. The DUT is compared to the internal standard and the difference in loss ($M_x - M_s$ dB) and phase ($\theta_x - \theta_s$) are recorded. Similarly, an external standard is measured, obtaining ($M_z - M_s$) and ($\theta_z - \theta_s$). The loss and phase of the DUT relative to the external standard are derived from the simple operations:

$$(M_x - M_z) = (M_x - M_s) - (M_z - M_s) \text{ dB} \quad (4)$$

$$(\theta_x - \theta_z) = (\theta_x - \theta_s) - (\theta_z - \theta_s) \text{ deg.} \quad (5)$$

This technique eliminates errors from drifts in the hardware external to the comparison switches and from the effects of asymmetry in the two switch positions. The comparison rate was chosen as 30 Hz to achieve an effective filtering of 60 Hz AC interference.

3.5 Interface circuitry

Digital control of the test sets by the COTMS computer and computer monitoring of their circuit status is provided by the interface circuitry.

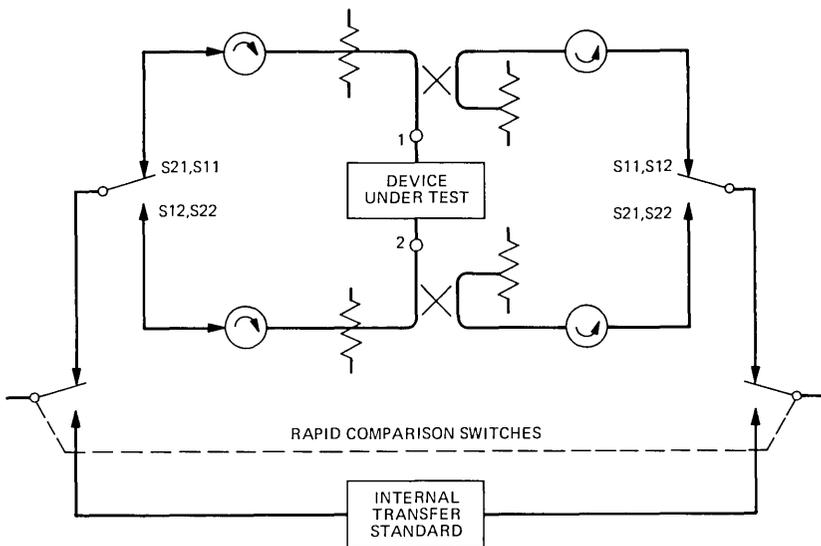


Fig. 4—Block diagram of the comparison circuit.

This circuitry processes all digital lines between the sets and cotms, with manual control of each line also provided by a toggle switch. Computer or cotms-generated logic signals provide control of the following major functions: s-x comparison switching, test-port configuration selection, dut rf level control, and yig filter tuning through a d/a converter.

The interface also monitors critical test system circuits for operation within established limits. Detection of an out-of-limit or fault condition halts the measurement process and energizes a panel lamp indicating the condition. Interface panel meters associated with each of the monitored circuits are also provided as troubleshooting aids.

IV. SYSTEM DESIGN CONSIDERATIONS

4.1 General

Previously, measurements of the required accuracy (see Table I) were made only with equipment of limited range and with painstaking care on a point-by-point basis. The need for measurements over extended ranges, with measurements performed rapidly and automatically, required some special system designs. Some of the techniques needed to achieve frequency accuracy, high signal-to-noise ratios, thermal stability, and impedance purity are discussed below.

4.2 Frequency accuracy

Frequency accuracy is necessary to minimize the uncertainties in the measurement of loss, phase, and delay of networks that have large variations of these parameters with frequency. The delay determination (see eq. 1) is directly sensitive to inaccuracies in frequency which lead to uncertainties in the quantity $(F_2 - F_1)$.

Exceptional frequency accuracy is achieved in each test system by the generation of all signals from the synthesizer and reference signals from cotms; these signals are derived from the same high-quality crystal source.

4.3 Amplitude noise

An amplitude signal-to-noise ratio exceeding 80 dB was a design objective to achieve a loss-measurement resolution of 0.002 dB for small values of loss. This was achieved through the use of a narrow detection bandwidth in cotms (500 Hz) and through noise-suppression circuitry.

The amplitude-noise suppression occurs in the level-control circuitry which appears in the source and local-oscillator sections. These level circuits serve the dual function of maintaining nominal signal

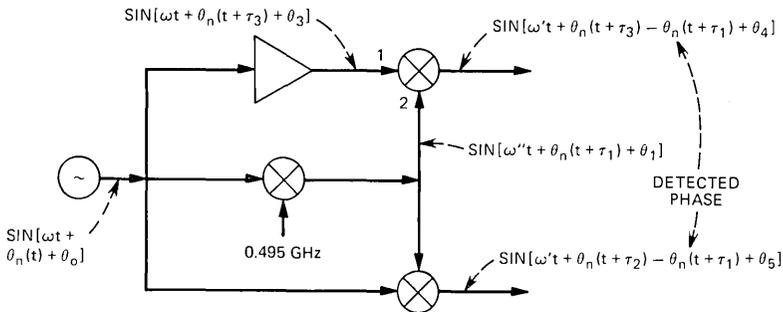


Fig. 5—Phase-noise cancellation mechanism.

levels and suppressing amplitude noise in a bandwidth exceeding 500 Hz centered about the leveled carrier signals.

4.4 Phase noise

A design consideration of considerable importance was the ability to measure delay with small uncertainty. The objective was to have an uncertainty not exceeding 0.1 ns (rms) when measuring narrow filters, where the frequency aperture, $F_2 - F_1$ in eq. (1), couldn't exceed 0.5 MHz. This objective is commensurate with a signal-to-phase noise ratio exceeding 80 dB.*

Phase noise was reduced to accomplish this objective in two ways: the first by reducing the phase noise present on the microwave signal and the second by noise cancellation in the frequency-conversion and phase-detection circuitry. Phase noise is increased in the microwave source generation process, for it is multiplied in direct proportion to frequency multiplication. The amount of multiplication has been reduced in each measuring system by the use of an intermediate stage of heterodyne conversion with a relatively phase-noise-free signal f_2 (see Fig. 2). The reduction is approximately four times or 12 dB. At 12 GHz, where the multiplier is largest, the phase signal-to-noise ratio is approximately 30 dB.

The second means of phase-noise reduction is by cancellation, of which there are two mechanisms. The first is demonstrated in Fig. 5.

Phase noise in the source $\theta_n(t)$ propagates via the measurement path and the LO generation path, appearing at ports 1 and 2 of the measurement-path IF mixer. The phase terms $[\theta_i; i = 0, 1, 2, \dots, 5]$ are constants that are not important to this evaluation. The delay terms $[\tau_i; i = 1, 2, 3]$ account for differences in the lengths of the various

* A simple calculation from eq. 1 reveals the following relationship between the variance in τ and the variance of the measured phase: $\sigma_\tau \doteq \sigma_\theta(\sqrt{2})/[360(F_2 - F_1)]$. An uncertainty, $\sigma_\tau = 0.1$ ns, is commensurate with $\sigma_\theta = 0.01$ degree.

signal paths. The IF signal results from the argument differences of the input signals, which causes phase-noise cancellation to the following degree:

$$\theta_n(t + \tau_3) - \theta_n(t + \tau_1). \quad (6)$$

The variance* of this IF phase signal is equal to

$$\sigma_{\theta_n}^2 = E[\theta_n(t + \tau_3) - \theta_n(t + \tau_1)]^2 = 2[R_{\theta_n}(0) - R_{\theta_n}(\tau)], \quad (7)$$

where $\tau = \tau_3 - \tau_1$. When $\tau = 0$, there is perfect cancellation. For a worst-case noise distribution,[†] assuming $\tau \leq 100$ ns, and an IF bandwidth ≤ 500 Hz the PM-S/N is improved in excess of 76 dB. Phase noise is similarly canceled in the phase-reference-path mixer.

The second phase-noise-cancellation mechanism arises from the use of a phase-reference signal. The difference in phase between the measurement signal and the phase-reference signal is detected. Phase noise common to both signals is rejected. This mechanism cancels phase noise originating in the source signal and the 0.495-GHz signals shown in Fig. 5. The cancellation exceeds 40 dB for the COMMS phase detector.

The combined effectiveness of the above mechanisms was measured by deliberately adding sinusoidal phase modulation exceeding 20 degrees peak to peak to the microwave source channel. No increase in the variance of a phase measurement (typically 0.01 degree) was detected.

4.5 Thermal stability

The rapid comparison technique of measurement described in Section 3.4 allows slowly varying changes in the test system lying external to the comparison switches to be canceled in the difference measurements. Changes in the transmission parameters of components lying inside the switches are not canceled. The latter component count is large due to the need to switch among the many test ports. The primary cause of parameter change is temperature fluctuation.

The temperature of the component assembly was stabilized in each test set by attaching it to a large aluminum plate, which in turn is surrounded by an insulating material. The composite structures have a thermal time constant exceeding 8 hours. This design has made thermal drift of these components a negligible source of error. The cables that connect the DUT to this circuitry are not enclosed in the stabilized environment and are a significant source of error (see Section 5.2).

* The stochastic variable θ_n is assumed to be a wide sense stationary process.⁴

[†] All energy concentrated at the extreme edges of the 500-Hz IF bandwidth.

4.6 Test port impedance invariance

By careful design, the impedance seen looking into the test ports used for simple insertion-loss measurements varies less than 2 percent from the nominal characteristic value of the coaxial line or waveguide. This is necessary to minimize reflection errors. With current hardware, it is not possible to achieve this impedance purity on the more complicated S -parameter switching networks. For measurements on these ports, the reflection errors are mathematically removed after the actual test set impedances are determined. This determination is accomplished through the measurement of calibration standards and a special hardware design.³

The S -parameter hardware shown in Fig. 4 is designed so that the impedances seen looking into each test port are constant, regardless of which parameter is measured. This invariance makes all S -parameter measurements referenced to the same impedances* and allows for the selfmeasurement of the physical impedances Z_1 and Z_2 seen by the DUT. This consistency allows the transformation of the measured data into an S -parameter representation free of reflection errors.

The invariance of the test port impedances in each test system is achieved by the use of isolators, attenuators, and the directivity of couplers to mask impedance changes caused by the switches. Variation of test port impedances of less than 0.05 percent has been achieved.

V. MEASUREMENT ERRORS

5.1 General

The nature of error mechanisms must be understood to be able to predict uncertainties that may arise under actual test conditions. This understanding was obtained during the development of the two test systems by directly characterizing the error-producing mechanisms and then by observing the errors in characterizing networks of predictable properties. Confidence was obtained from the consistency of the predicted errors and observed errors. The most important error-producing processes in the appliques are: thermal instabilities, nonlinearities, crosstalk, second-order-noise products, and impedance-reflection errors.

5.2 Thermal instabilities

Thermal instabilities, which cause drift of the apparatus in the comparison unit between calibration and device measurements, are the

* In the S_{11} switch configuration, the detected loss and phase is directly related to $S_{11}(Z_1, Z_2)$ of the device under test; Z_1 and Z_2 are the test port impedances. Once Z_1 , Z_2 and the four-parameter set $S(Z_1, Z_2)$ are determined, it is possible through known transformations to obtain another representation such as $S(50, 50)$. The transformed parameters are then free of reflection errors.

largest source of uncertainty in absolute measurements. This error mechanism was anticipated in the system designs and measures were taken to stabilize the temperature of critical components. These measures are discussed in Section 4.5.

The paths that connect the test set connectors to the DUT are not controlled in temperature. These paths typically contain several feet of coaxial cable to achieve a flexible connection to the DUT. In the case of waveguide measurements, they also contain coax-waveguide transducers.* Careful measurements have correlated almost all of the observed output data drifts to temperature changes in these external connecting paths. The following sensitivities in both test sets have been observed:

$$\Delta\theta/\Delta T \doteq 0.1^\circ/\text{F}^\circ \text{ per GHz} \quad (8)$$

and

$$\Delta L/\Delta T \doteq 0.01 \text{ dB}/\text{F}^\circ \text{ per GHz.} \quad (9)$$

External path temperature cycling of 1°F peak to peak has been observed over a 1-hour period in the air-conditioned laboratory environment. The slowly varying thermal drifts do not cause significant uncertainties in the more critical differential measurements of delay [see eq. (1)] and loss slope.

5.3 System nonlinearities

The onset of nonlinearities in the test systems determines the high-signal-level limit of operation. Signal levels and components have been chosen to assure a measurement uncertainty less than 0.001 dB and 0.01 degree due to deviations from linear detection. Devices with gain are measured at a reduced excitation level so that the output-signal level lies in the linear-detection region.

One technique used to confirm linear performance consists of measuring a 6-dB loss unbalance between the two comparison-switch portions at the highest level and at a lower level. Adequate isolation exists between the comparison switch and the level-adjusting network to avoid errors due to impedance changes. The differences between the nominal 6-dB measurements are then used as a measure of errors caused by nonlinear behavior. The data indicates deviations from linear behavior are within accuracy objectives for both test sets. The components that limit linear behavior are the modulators in the heterodyne circuit.

* The coax-waveguide transducers are reciprocal and low loss for the four *S*-parameter measurements. The transducers include isolators and matching attenuators for the simple insertion-loss measurements.

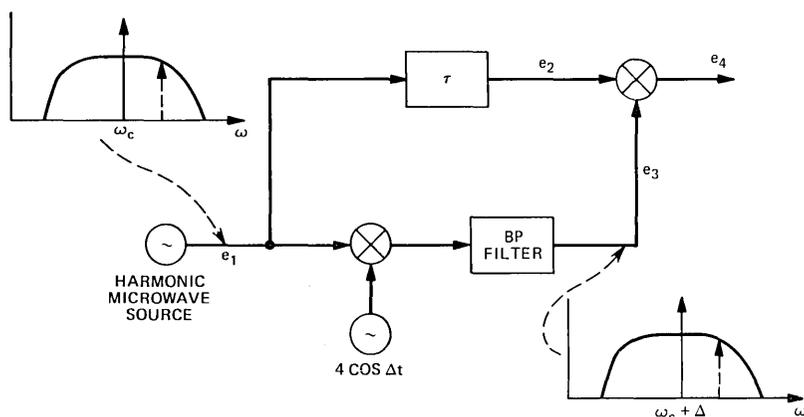


Fig. 6—Second-order noise product mechanism.

5.4 Crosstalk

Measurement uncertainties occur for DUTs with high loss when signals can reach the detector by other paths than through the DUT. By design, these leakage signals are more than 80 dB below the zero-loss-detected signal. For a DUT with loss $Mx(\text{dB})$, the uncertainty due to leakage or crosstalk is as follows:

$$|\text{Uncertainty}(\text{dB})| < 8.68 \times 10^{Mx-80/20} \quad (10)$$

$$|\text{Uncertainty}(\text{deg})| < 57.3 \times 10^{Mx-80/20}. \quad (11)$$

For example, the uncertainties for a device having 60 dB of loss could be 0.9 dB and 5.7 degrees. Crosstalk uncertainties are removed from reflection-coefficient and return-loss measurements but not from transmission measurements.* The second uncertainty term in eq. (3) accounts for this crosstalk.

5.5 Second-order noise products

The high degree of correlation between the RF signal and the RF local oscillator signal in both appliques leads to a small bias in the measured data. The mechanism can be understood by observing the simplified diagram of half of the comparison heterodyne section shown in Fig. 6. The harmonically generated microwave signal e_1 has associated with it a noise spectrum whose envelope is determined by the YIG filters and the spectral shape of the noise phenomena. The phase noise spectral shape typically follows a $1/\Delta f$ decay about the carrier frequency. The amplitude noise is typically flat and is significant only

* The removal of crosstalk uncertainties from the transmission measurement is possible with additional calibration measurements and data processing.

in the 4- to 12-GHz test set due to the use of TWT amplifiers with noise figures exceeding 40 dB.

A frequency-shifted replication of the noise spectrum of e_1 is present on the local oscillator signal e_3 . A noise component on e_1 , illustrated in Fig. 6 as a dashed line, interacts in the converter with a replica of the same component on e_3 to produce a coherent IF signal. The cumulative effect of all noise components produces a small signal at frequency $\Delta = 495$ MHz, leading to a bias in the measured quantities.

A detailed analysis* of the phenomena (see Appendix) indicates that the average value of the IF signal amplitude must be adjusted by the multiplicative factor,

$$[1 + Ram(\tau) + Rpm(\tau)]. \quad (12)$$

$Ram(\tau)$ and $Rpm(\tau)$ are the autocorrelation functions of the AM and PM noise spectrums, respectively, that are common to the signals e_1 and e_3 . Similarly, the phase is adjusted by the additive factor,

$$[2 Ram/pm(\tau)] \text{ radians.} \quad (13)$$

$Ram/pm(\tau)$ is the cross-correlation function of the AM and PM noise spectrums. The parameter (τ) is the delay difference between the two paths to the converter shown in Fig. 4. It is therefore a function of the delay through the DUT.

The uncertainty caused by this second-order-noise interaction is proportional to the changes in the above terms due to the differential nature of the measurements [see eqs. (4) and (5)]. These uncertainties are to first order:

$$\text{Uncertainty (dB)} = 8.68[R\acute{a}m(\tau) + R\acute{p}m(\tau)](\tau x - \tau z) \quad (14)$$

$$\text{Uncertainty (deg)} = 115[R\acute{a}m/pm(\tau)](\tau x - \tau z). \quad (15)$$

The delay difference ($\tau x - \tau z$) is the difference in delay between the DUT and the calibrating device. The prime denotes the first derivative with respect to (τ).

This analysis gave the qualitative understanding of this phenomena that led to the use of an intermediate stage of heterodyne conversion to reduce source noise (see Section 4.4). A quantitative understanding is prevented by a lack of knowledge of the noise-process statistics. Measurements made to reveal these uncertainties have shown that they are negligible for relatively broadband networks and for differential measurements such as delay and loss slope. Uncertainties of 0.01 dB and 0.03 degree were observed on the steep skirts of a 10-GHz

* Assuming the heterodyne converter is approximated as forming the product $e_2 \cdot e_3$.

cavity filter with a 40-MHz bandwidth. The uncertainties decrease directly with frequency since the phase noise dominates (see Section 4.4). These uncertainties are commensurate with the performance discussed in Section 2.3.

5.6 Errors from impedance mismatches

Errors are made in measuring a DUT when this device, calibration standards, and the test set terminals have impedances that differ from the assumed nominal values. In the microwave test systems, these errors are quite significant and are fundamentally different in the various measurement modes.

The simplest mode is the measurement of insertion transmission. The impedance mismatch uncertainty in measuring either the forward- or reverse-scattering parameters S_{21} and S_{12} has been discussed by many authors.^{1,2,5} A general relationship is:

$$\frac{\overline{S_{ij}}}{M_{ij}} = \frac{1 - \rho_o \overline{S_{jj}} - \rho_l \overline{S_{ii}} + \rho_o \rho_l \overline{\Delta S}}{1 - \rho_o \rho_l}, \quad (16)$$

where

M_{ij} is the voltage-scattering parameter actually measured between the test set generator impedance Z_g and the test set load impedance Z_l .

$$\rho_o = \frac{Z_g - R_g}{Z_g + R_g} \quad \rho_l = \frac{Z_l - R_l}{Z_l + R_l}. \quad (17)$$

The S -parameters are those that would be obtained if the test set had ideal impedances R_g and R_l .

In both microwave test systems, the insertion transmission measurements are made with relatively simple hardware. This simplicity allowed the impedances of the test ports to be kept within 2 percent of nominal. Examples of the impedance mismatch uncertainty that can occur in measuring unknowns (assuming $|\overline{S_{12}} \cdot \overline{S_{21}}| \leq 1$; $|\rho_o| \leq 0.01$, $|\rho_l| \leq 0.01$) of varying input and output return losses have been computed from eq. (16) and are given below:

| DUT Return Loss* | Max. Amplitude Error | Max. Phase Error |
|------------------|----------------------|------------------|
| 40 | 0.0035 dB | 0.023° |
| 30 | 0.0072 dB | 0.048° |
| 20 | 0.0191 dB | 0.126° |

The impedance or reflection coefficient of a DUT is determined from one measurement of the DUT and one measurement on each

* Return loss = $-20 \log |\overline{S_{ii}}|$ $i = 1, 2$.

of three calibration standards. The standards are typically two reactive circuits (nominally an open and a short) and a resistive circuit (Z_s) nominally matched to a transmission line characteristic impedance (R_s). The uncertainty that is incurred if $Z_s \neq R_s$ is (Ref. 2):

$$M - \Gamma = \frac{\Gamma - \rho_s}{1 - \rho_s \Gamma} - \Gamma \doteq \rho_s (\Gamma^2 - 1) \quad (18)$$

$$\rho_s = \frac{Z_s - R_s}{Z_s + R_s}. \quad (19)$$

M is the measured voltage-reflection coefficient and Γ is the reflection coefficient that would be measured if $Z_s = R_s$. ρ_s is smaller than 0.001 for the better resistive standards. The errors arising from uncertainties in the reactive standards are much smaller and are ignored. Notice that (18) vanishes for $\Gamma = \pm 1$. This occurs when the reflection measurement reduces to a differential comparison of the DUT with either the open or short standard. Notice that the use of calibration standards makes the mismatch errors dependent on the quality of the calibration standards and not upon the impedances of the test system's terminals.

Four-parameter characterizations are made with complex switching hardware in which the test-port impedances can vary with frequency up to 20 percent from the design values. The large impedance-mismatch uncertainties that could arise are reduced by processing measurements of the DUT with measurements on calibration standards. The characterizations are performed in three steps. First, all S parameters defined with respect to the actual test set impedances, Z_1 on port 1 and Z_2 on port 2, are determined (see Section 4.6). These parameters are, by definition, free of impedance-reflection errors. Next, the impedances Z_1 and Z_2 are measured using the same three reflection coefficient standards mentioned above. Finally, these $S(Z_1, Z_2)$ parameters are transformed to another representation, most commonly $S(R_s, R_s)$. The reflection errors that arise in the transformed representation are dependent upon the accuracy to which Z_1 and Z_2 or their voltage-reflection coefficients ρ_1 and ρ_2 are characterized. Thus, the impedance mismatch errors are related to uncertainties in the calibration standards and not the test set impedances.

The quantities ρ_1 and ρ_2 are,* from eq. (18), approximately equal to ρ_{s1} and ρ_{s2} , respectively. ρ_{s1} and ρ_{s2} are the reflection coefficients of the resistive calibration standards R_{s1} and R_{s2} used on ports 1 and 2, respectively. The relationships between the measured parameters

* $|\rho_i| \ll 1 \quad i = 1, 2.$

$[S(Z_{s1}, Z_{s2})]$ and the error-free parameters $[S(R_{s1}, R_{s2})]$ are found by known transformations.²

$$\frac{S_{ij}(R_{s1}, R_{s2})}{S_{ij}(Z_{s1}, Z_{s2})} = \frac{1 - \rho_{si}S_{ii} - \rho_{sj}S_{jj} + \rho_{si}\rho_{sj}\Delta S}{(1 + \rho_{si})(1 - \rho_{sj})} \Big|_{R_{s1}, R_{s2}} \\ \doteq 1 - \rho_{si}[S_{ii}(R_{s1}, R_{s2}) + 1] \\ - \rho_{sj}[S_{jj}(R_{s1}, R_{s2}) - 1]. \quad (20)$$

$$S_{ii}(Z_{s1}, Z_{s2}) = \frac{S_{ii} - \rho_{si} + \rho_{si}\rho_{sj}S_{jj} - \rho_{sj}\Delta S}{1 - \rho_{si}S_{ii} - \rho_{sj}S_{jj} + \rho_{si}\rho_{sj}\Delta S} \Big|_{R_{s1}, R_{s2}} \\ \doteq S_{ii}(R_{s1}, R_{s2}) - \rho_{si}[1 - S_{ii}^2(R_{s1}, R_{s2})] \\ + \rho_{sj}S_{ij}(R_{s1}, R_{s2})S_{ji}(R_{s1}, R_{s2}). \quad (21)$$

These relationships coupled with extensive experimental data were used to derive the uncertainty bounds given by eqs. (2) and (3).

VI. MEASUREMENT RESULTS

6.1 General

Ideally, measurement accuracy should be confirmed with network standards selected to accentuate various error mechanisms. At microwave frequencies, no complete set of such networks exists. The measurement results given below* examine accuracy with a few network standards whose properties are based upon geometric dimensions. Additional indications of accuracy are obtained from characterizations of a given network under different measurement conditions that are expected to give the same results, and by comparing characterizations on both test systems in the 3.7- to 4.2-GHz band of overlap.

6.2 Transmission measurements

6.2.1 Delay of an airline

The ability to accurately measure envelope delay is of critical importance to the development of transmission networks. The delay-measuring uncertainties in both measuring sets were evaluated by measuring an airline of known electrical length, one of the few available standards of delay. The data shown in Fig. 7 is a measurement of the delay of a 7-mm airline of 20.35 cm in electrical length. A frequency difference of 20 MHz was used in the on-line calculation of delay using eq. (1). The results show a measurement uncertainty of less than ± 0.022 ns about the nominal delay of 0.678 ns.

* Errors caused by the failure of connectors to make good contact are not discussed. These errors are not significant for waveguide connections. For this reason, many of the networks measured for accuracy verification have waveguide connectors.

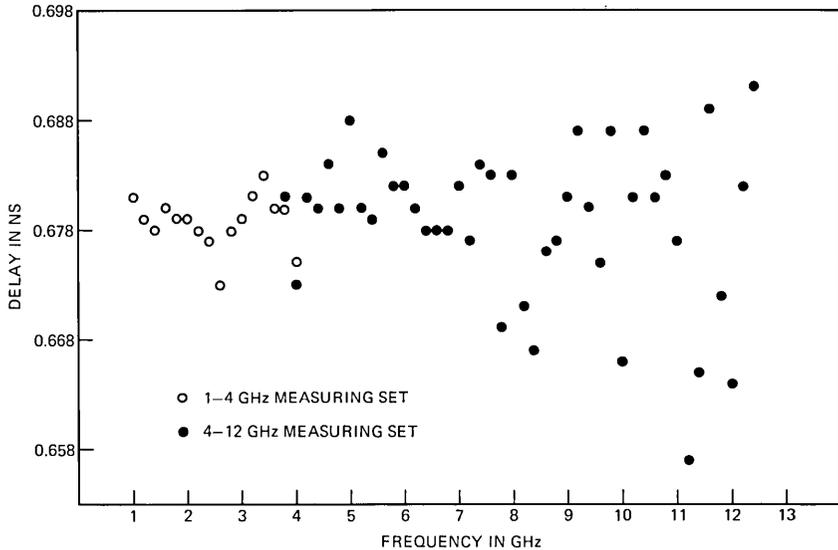


Fig. 7—Delay of an airline with a 20.35-cm electrical length.

6.2.2 Insertion parameters of a waveguide filter

The loss and delay of a narrow-bandwidth waveguide filter was measured on both test sets to compare their accuracy. Waveguide isolators were used to improve the impedances of the coax-waveguide transducers needed to make measurements on the 1- to 4-GHz test set. The same isolators were used in the measurement made on the 4- to 12-GHz test set so that the DUT would see essentially the same test set impedances in each measurement. The filter characteristic showing the passband and a portion of the stop band measured on the 1- to 4-GHz test set is shown in Fig. 8. Equivalent data taken on the 4- to 12-GHz test set (not shown) coincides point for point except at loss levels above 80 dB. At loss values above 80 dB, the measured loss is limited by crosstalk to about 88 dB.

The difference between characterizations on the two sets in a magnified portion of the passband is shown in Fig. 9. The worst-case loss and delay* differences are 0.004 dB and 0.08 ns. The average absolute differences are 0.0012 dB and 0.019 ns. These differences are less than the bound on the typical uncertainty computed from eq. (3), 0.005 dB and 0.1 ns.

* Delay was computed on-line using eq. (1) with a frequency difference of 1 MHz.

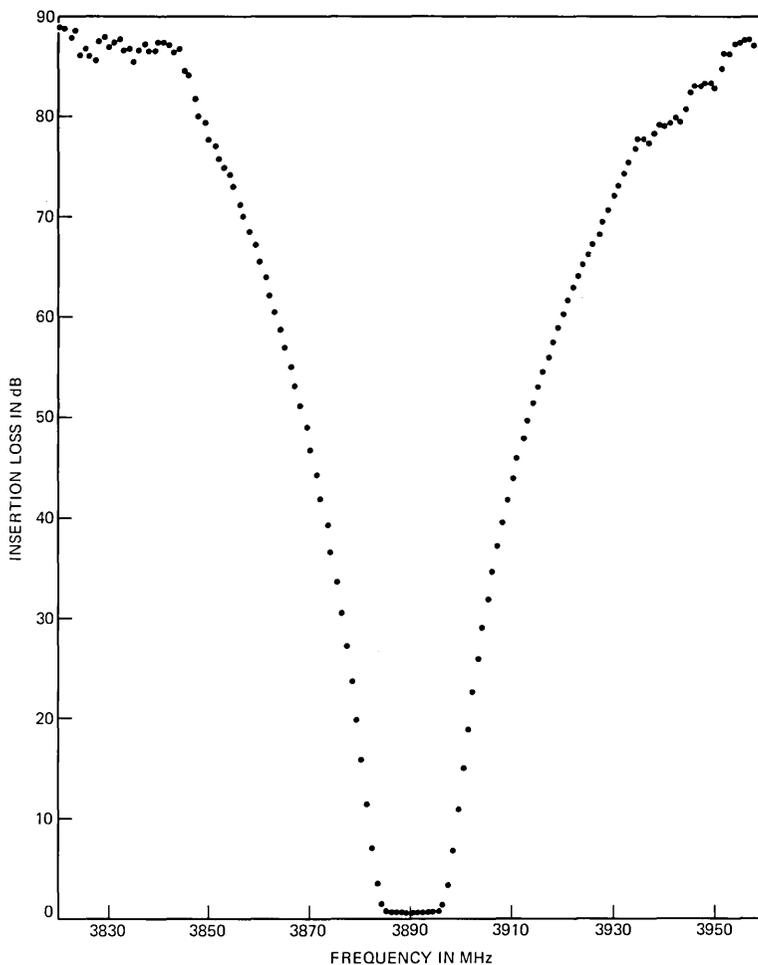


Fig. 8—Insertion loss of a waveguide filter measured on the 1- to 4-GHz measuring set.

6.2.3 Measurement stability

The insertion loss and phase of a 12-dB attenuator was measured at the beginning and end of a 20-minute time interval to evaluate the stability of insertion measurements. The attenuator was surrounded by thick insulation to increase its response time to temperature changes. Such precautions are sometimes necessary to make precise measurements at microwave frequencies depending on the temperature dependence of the DUT.

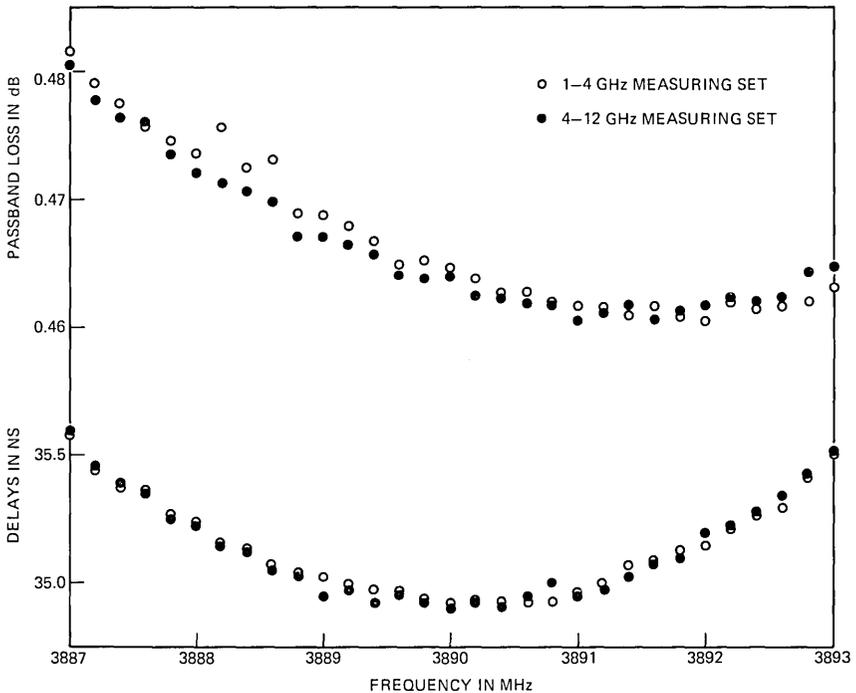


Fig. 9—Passband loss and delay of waveguide filter measurements.

The differences in the two measurements are shown in Fig. 10. The loss differences are typically smaller than 0.004 dB. The phase differences are approximately proportional to frequency as if there were a change in the electrical length of the attenuator or test-system circuitry. The data indicate an equivalent change in electrical length of less than 0.001 inch over the 20-minute interval. This corresponds to a change of less than four parts in one million in the more than 20 feet of equivalent circuit length associated with either measuring system (see Section 4.5).

6.3 Return-loss measurements

Return-loss or one-port characterizations are measured on the 4- to 12-GHz test set on both waveguide and coaxial test ports. These measurements are made on the 1- to 4-GHz test set using coaxial test ports. Low-loss transducers can be used on either set to make measurements in other connectors. The measurement of the DUT and calibration is under computer direction. Three calibration standards described by a constant magnitude and a phase that varies linearly with propagation coefficient are used. This description allows for the

correction of some perturbations from nominal values and for the mathematical translation of the reference plane of measurement. The calibration standards typically chosen are a nominal short circuit, open circuit, and a resistive termination. These features will be demonstrated below in the measurement of some precisely machined waveguide networks.

6.3.1 Inductive iris

A symmetric inductive iris 1.7 inches wide by 0.2 inch thick was constructed in WR229 waveguide and measured on both test systems. Measurements on the 4- to 12-GHz system were made with connection to a high-directivity magic-tee hybrid. Measurements on the 1- to 4-GHz system were made with connection to the internal coaxial hybrid through a lossless waveguide to coax transducer. Calibration was performed with three waveguide standards, a short, an offset short, and a match. A different set of standards was used on each test system. Figure 11 gives the measurement results on the 1- to 4-GHz system. Data obtained from the 4- to 12-GHz system (not shown) exhibit a mean absolute difference of 0.01 dB and 0.08 degree from

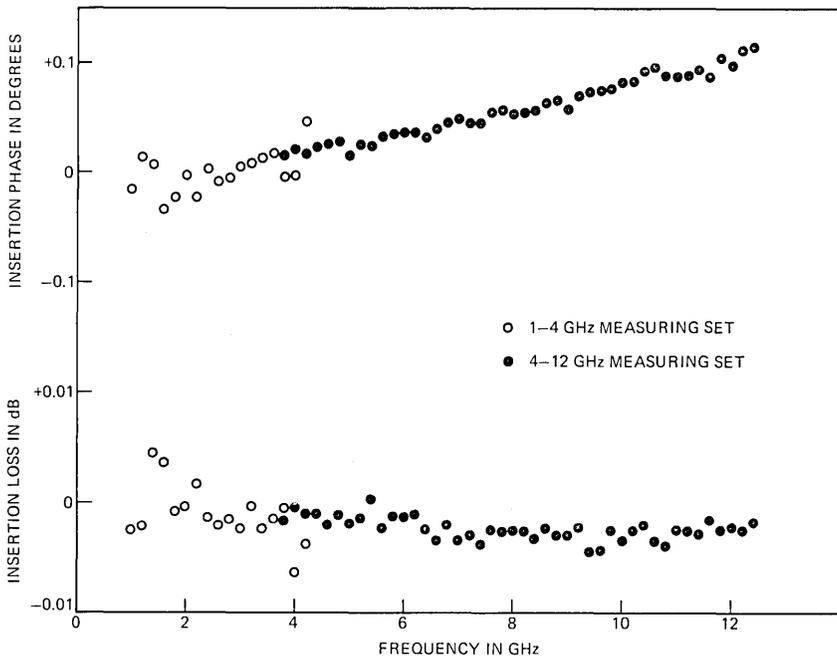


Fig. 10—Change in insertion loss and phase of a 12-dB attenuator after a 20-minute time interval.

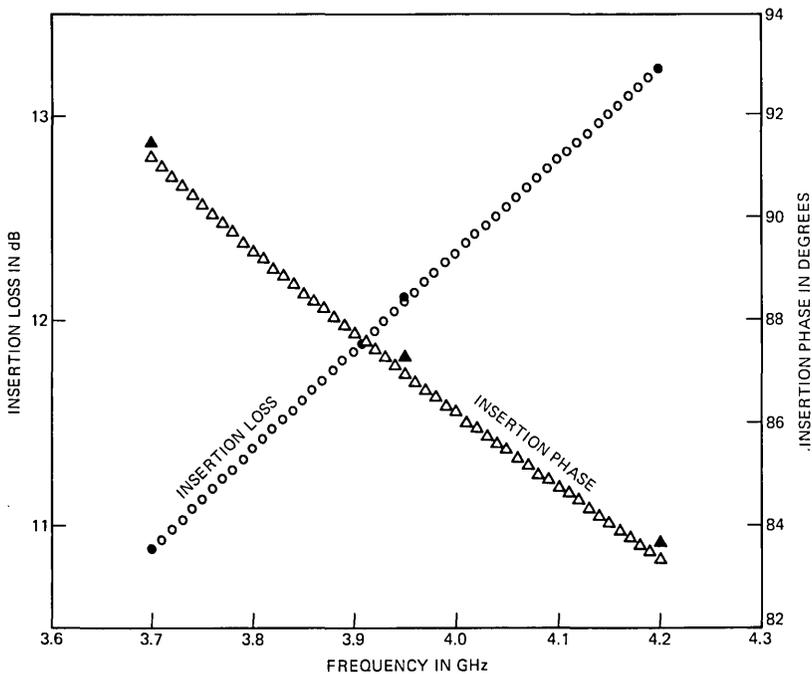


Fig. 11—Insertion loss and phase of a symmetrical inductive window 1.700 inches wide by 0.2 inch thick.

the 1- to 4-GHz system data, with peak differences of 0.03 dB and 0.15 degree. The solid data points are derived from theoretical computer computations based upon the mode-matching technique.⁶ These differences are less than the uncertainty of 0.037 dB and 0.24 degree computed from eq. (2). Experimental results in the next section reveal that a major contributor to the uncertainties in these measurements is the deviations of the match from a zero reflection coefficient.

6.3.2 Offset short

A short circuit displaced 0.4090 cm in WR90 waveguide was measured as a circuit standard. This circuit has a reflection coefficient of unit magnitude (0 dB) and an angle of -90 degrees in the middle of the 10.7- to 11.7-GHz microwave radio band. The calibration standards used for the measurement were a zero length short, a match, and a short offset by 0.8197 cm. The latter calibration standard has a reflection coefficient of unit magnitude and an angle of 180 degrees in midband. The match has a return loss exceeding 60 dB. The measurement of this circuit should produce the worst-case uncertainties since

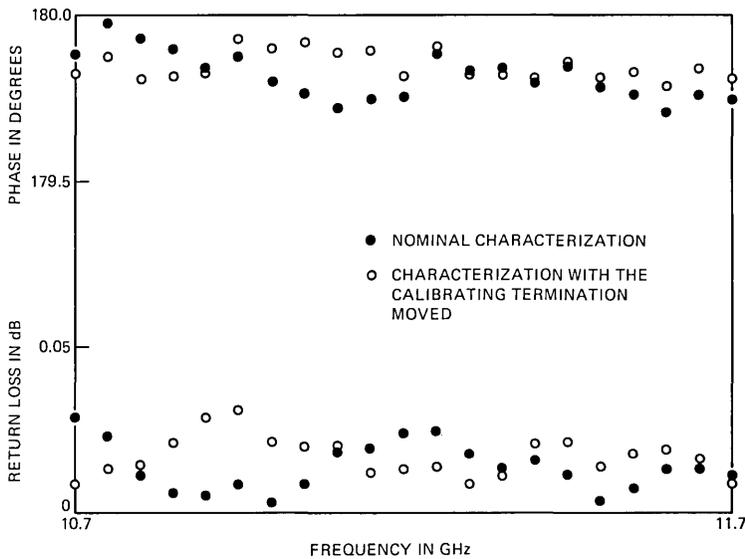


Fig. 12—Return loss of an offset short in WR90 waveguide.

its reflection coefficient has the greatest distance on the Smith chart from the reflection coefficients of the calibrating standards.

The measurement results are shown in Fig. 12. The measurement reference plane was mathematically moved to the plane of the circuit's short circuit by an on-line program. The nominal characterization at this displaced plane is a reflection coefficient of unit magnitude and constant 180-degree phase angle. Two characterizations are presented, one with the tapered absorber in the match in its nominal position and a second with the absorber moved 0.83 cm ($\sim\lambda/4$) within its electroformed waveguide. The changes in the characterizations reveal that deviations of match from a zero reflection coefficient are a significant source of measurement uncertainty. The worst-case deviations are less than the uncertainties of 0.052 dB and 0.34 degree computed from eq. (2).

6.4 Four-parameter characterizations

Four-parameter or two-port characterizations are measured on both microwave test sets using automatically switched hardware and a computer-directed calibration technique. These features were briefly discussed in Sections 2.2 and 5.6. Voltage-scattering parameters (S parameters) are computed in real time. The calibration-standard models used assume a constant magnitude and a phase that varies linearly with propagation coefficient. This description allows for the

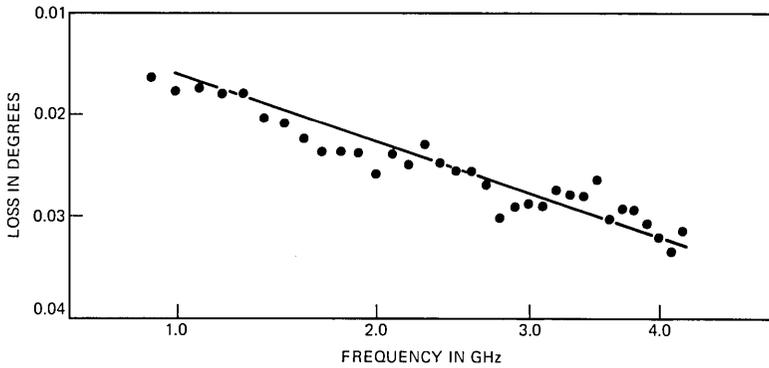


Fig. 13— S_{21} phase of a 50-ohm airline.

correction of some perturbations from nominal values and for the translation of the reference planes of measurement. Some of these features will be demonstrated in the measurements below. Parameter representations other than S parameters are available through processing following the measurements.

6.4.1 Airline characterizations

An airline is a two-port network whose characteristics are precisely known. Errors can arise in characterizing an airline if the test set impedances deviate significantly from nominal. Impedance deviations

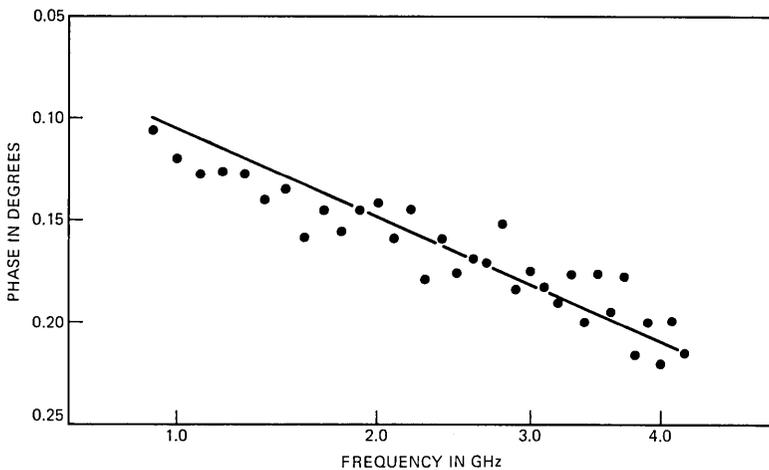


Fig. 14— S_{21} magnitude of a 50-ohm airline.

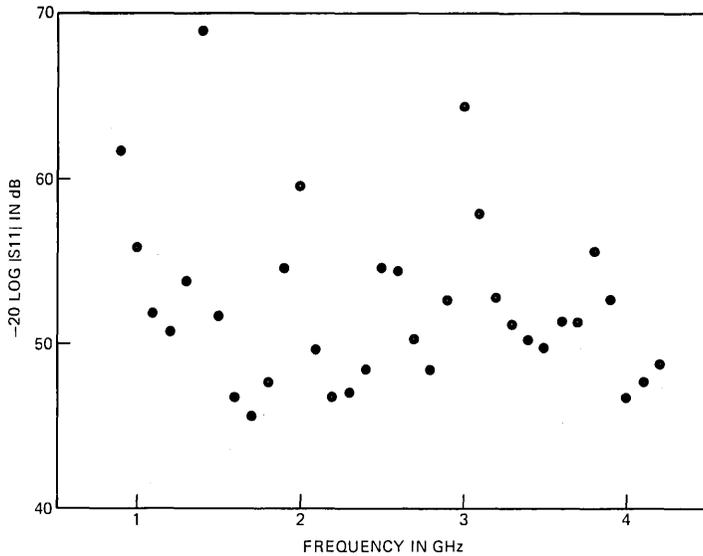


Fig. 15— S_{11} magnitude of a 50-ohm airline.

in excess of 20 percent are observed in the test sets. These errors are avoided by the calibration method discussed in Section 5.6.

A 30-cm-long 14-mm airline was characterized on the 1- to 4-GHz measuring system to illustrate these points and to demonstrate the accuracy obtained. The S_{21} measurements are plotted in Figs. 13 and 14 on a square-root-of-frequency scale. The reference planes of measurement were translated to the center of the airline so that the phase data is the difference between the measured airline and an ideal airline. Observe that the data indicates characteristics nearly linear* with respect to the square root of frequency as expected from theoretical models, taking skin losses into account.² The maximum deviation from linear is 0.004 dB and 0.05 degree. The typical deviation derived from eq. (3) is predicted to be less than 0.005 dB and 0.04 degree. An S_{21} measurement was also made without calibration to demonstrate the accuracy improvement provided by computationally removing the test set impedance-reflection errors. Deviations from a square-root-of-frequency characteristic in excess of 0.1 dB and 0.7 degree were observed.

The magnitude of S_{11} (dB) is shown in Fig. 15. The uncertainty of this measurement is dominated by deviation of the resistive calibration

* The solid line in Fig. 14 was extrapolated from resonant Q measurements made at 250 MHz.

standard from 50 ohms, corresponding to a reflection coefficient uncertainty of 0.005 (46 dB).

6.4.2 Transistor characterization

The 1- to 4-GHz measuring system is frequently used for the characterization of microwave transistors to be used in amplifier circuits. Their small size and impedances that vary considerably from 50 ohms make them particularly difficult to characterize accurately. Two characterization techniques have been used. The first technique uses simplified handling by calibrating with large coaxial standards at the interface of the transistor jig and test set. The defined reference planes are then translated to the jig-transistor interface making the approximation that the jig is an ideal 50-ohm transmission line. The second technique uses calibration standards of the same size scale as the transistor inserted at the jig-transistor interface. The parasitic circuit elements of these standards are significantly smaller than those associated with the transistor jig. A detailed characterization effort is under way so that a definitive accuracy evaluation can be made.

Measurements of S_{11} using the two techniques to characterize a beam-led transistor mounted on a 0.1-inch-square substrate are

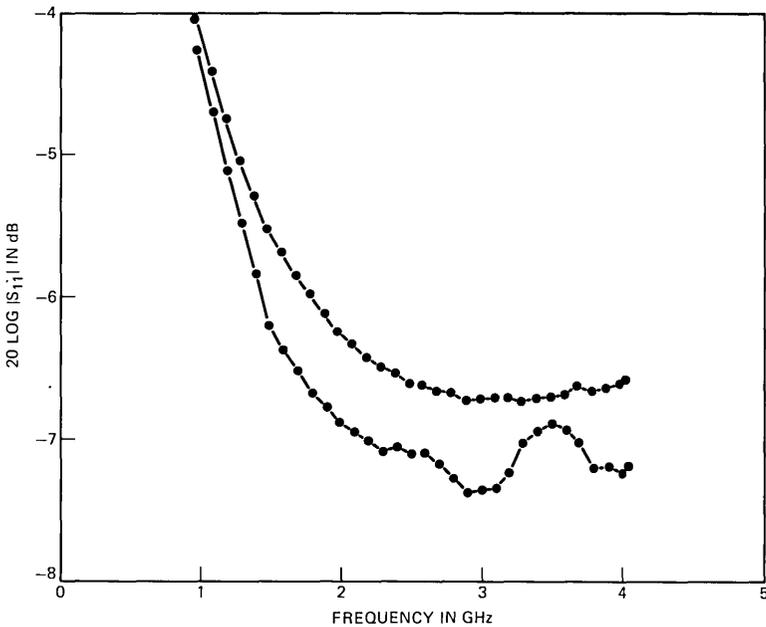


Fig. 16—Input reflection coefficient of a beam-led microwave transistor.

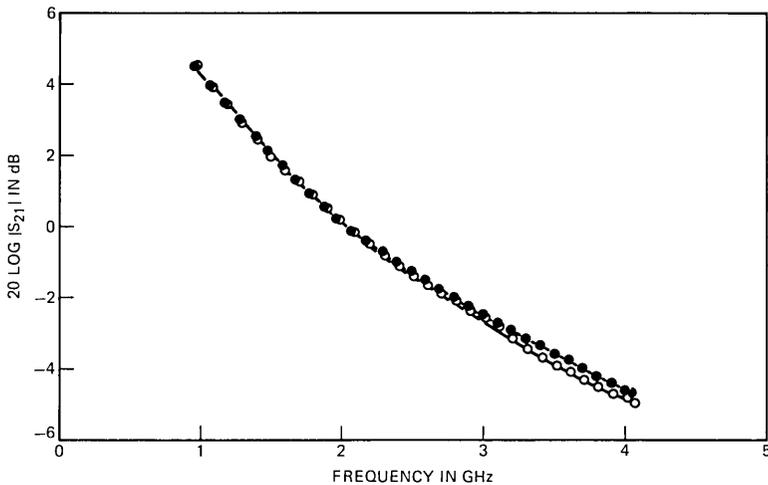


Fig. 17—Forward scattering parameter of a beam-led microwave transistor.

shown in Fig. 16. The top curve is the characterization using the transistor-sized standards. This curve is free of the impedance cycling with frequency that is evident in the second curve measured with the coaxial standards. This cycling cannot be attributed to the transistor and is found in measurements on the jig with the transistor replaced with a through connection.

The measurements of S_{21} using the two calibration techniques is shown in Fig. 17. The measurements using the coaxial standards cycle about the measurements using the transistor-sized standards, ending up 0.5 dB lower at 4 GHz. These results reveal that measurement accuracy can be improved by calibration directly at the interface of the DUT and the test set.

VII. SUMMARY

The development of two general-purpose network analyzers covering the frequency range of 0.9 to 12.4 GHz has been discussed. These network analyzers were developed to provide flexible, fast, and accurate circuit characterizations, which are necessary feedback in the circuit design process.

Automatic control and highly accurate baseband detection is provided by the Computer Operated Transmission Measuring Set. The techniques used to extend baseband-measurement capabilities to microwave frequencies while retaining accuracy were discussed. The most prominent sources of measurement errors were quantitatively examined. Measurements results that confirm accuracy were presented.

APPENDIX

Analysis of Second Order Noise Products

Rice's representation (Ref. 4, p. 374) of a stochastic process is convenient for the analysis of the heterodyne process modeled by Fig. 6.

$$e_1 = \sin(\omega_c t) + n(t) \sin(\omega_c t) + m(t) \cos(\omega_c t). \quad (22)$$

The first term is the normalized carrier signal. The second term is in phase noise which can be thought of as AM noise for small $n(t)$. The last term is in quadrature noise which can similarly be thought of as phase noise. The stochastic variables are assumed to be stationary and to have the following properties.*

$$E[n(t)] = E[m(t)] = 0. \quad (23)$$

$$E[n(t + \tau)n(t)] = Ram(\tau) = Ram(-\tau). \quad (24)$$

$$E[m(t + \tau)m(t)] = Rpm(\tau) = Rpm(-\tau). \quad (25)$$

$$E[m(t + \tau)n(t)] = Ram/pm(\tau) = -Ram/pm(-\tau). \quad (26)$$

The delayed signal e_2 and the up-converted signal e_3 are given by

$$e_2 = \sin(\omega_c t + \omega_c \tau) + n(t + \tau) \sin(\omega_c t + \omega_c \tau) + m(t + \tau) \cos(\omega_c t + \omega_c \tau). \quad (27)$$

$$e_3 = 2 \sin(\omega_c + \Delta)t + 2n(t) \sin(\omega_c + \Delta)t + 2m(t) \cos(\omega_c + \Delta)t. \quad (28)$$

The output converter is approximated as forming the product $e_4 = e_2 \cdot e_3$.

$$e_4 = \cos(\Delta t - \omega_c \tau) + [n(t + \tau)n(t) + m(t + \tau)m(t)] \cos(\Delta t - \omega_c \tau) + [m(t + \tau)n(t) - n(t + \tau)m(t)] \sin(\Delta t - \omega_c \tau) + [n(t + \tau) + n(t)] \cos(\Delta t - \omega_c \tau) + [m(t + \tau) - m(t)] \sin(\Delta t - \omega_c \tau). \quad (29)$$

Filterable tones falling at $2\omega_c$ have been omitted from eq. (29).

Taking the expected value of e_4 and substituting the properties of eqs. (23) through (26) yields,

$$E(e_4) = \cos(\Delta t - \omega_c \tau) + [Ram(\tau) + Rpm(\tau)] \cos(\Delta t - \omega_c \tau) + 2Ram/pm(\tau) \sin(\Delta t - \omega_c \tau). \quad (30)$$

$$E(e_4) \doteq [1 + Ram(\tau) + Rpm(\tau)] \cos[\Delta t - \omega_c \tau - 2Ram/pm(\tau)]. \quad (31)$$

* $E[\cdot]$ denotes expected value. If $Ram(\tau) = Rpm(\tau)$, then the noise process also models an additive as well as a modulating stationary process.⁴

The amplitude and phase of the mean IF signal e_4 are seen in eq. (31) to be modified by the noise present on the microwave signals.

REFERENCES

1. W. J. Geldart, G. D. Haynie, and R. G. Schleich, "A 50 Hz-250 MHz Computer-Operated Transmission Measuring Set," *B.S.T.J.*, 48, No. 5 (May 1969), pp. 1339-1381.
2. J. G. Evans, "Measuring Frequency Characteristics of Linear Two-Port Networks Automatically," *B.S.T.J.*, 48, No. 5 (May 1969), pp. 1313-1338.
3. J. G. Evans, "Linear Two-Port Characterization Independent of Measuring Set Impedance Imperfections," *Proc. IEEE*, 59, No. 4 (April 1968), pp. 754-755.
4. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, New York: McGraw-Hill Book Company, 1965, pp. 300-340.
5. D. Leed, "An Insertion Loss, Phase and Delay Measuring Set for Characterizing Transistors and Two-Port Networks Between 0.25 and 4.2 gc," *B.S.T.J.*, 45, No. 3 (March 1966), pp. 397-440.
6. N. Marcuvitz, *Waveguide Handbook*, New York: McGraw-Hill, 1951, pp. 255-257.

Evaluation of an Automatic Speaker-Verification System Over Telephone Lines

By A. E. ROSENBERG

(Manuscript received September 9, 1975)

An automatic speaker-verification system accessed by test customers from their own telephones over dialed-up lines has been evaluated. The test population consisted of over 100 male and female speakers who called up nominally once each working day over a period of five months. The operation of the system is based on a set of functions of time obtained from acoustic analysis of a fixed, sentence-long utterance. These functions are compared with stored reference functions to determine whether to accept or reject an identity claim. The system is implemented on a NOVA 800 laboratory computer. Telephone line access to the computer is via a data set hookup. Identity claims are made by keying an identification number on a Touch-Tone® dial. Instructions and responses to the customer are made by means of a programmed voice-response system. Reference data was computed off-line and updated with the analysis data of accepted utterances. The evaluation indicated an error rate of approximately 10 percent for new customers and approximately 5 percent for adapted customers.

I. INTRODUCTION

Speaker verification is the authentication of an individual's claimed identity by analysis of his spoken utterances. Research on an automatic system for speaker verification at Bell Laboratories has been reported in previous papers.¹⁻³ The system is based on an acoustic analysis of a fixed, sentence-long utterance resulting in a function of time or contour for each feature analyzed. Features selected for analysis in previous evaluations have included pitch, intensity, the first three formants, and selected predictor coefficients. The system compares the set of sample contours obtained from an unknown individual with the set of reference contours corresponding to the identity claimed by that individual. If the comparison results in an overall measure of dissimilarity which is smaller than a predetermined threshold, the identity claim is accepted. Otherwise, it is rejected.

Previous evaluations of the system have concentrated on investigating features to be analyzed and developing comparison procedures to make the system as effective as possible in terms of reducing overall error rate. The speech samples used in these evaluations were collected by wideband recording of male speakers in a sound booth. The recordings were selected and edited to eliminate botched utterances and non-speech acoustic events. From the outset, however, the intent has been to provide a completely automatic system which could operate via dialed-up lines from telephones on the user's own premises and to include both male and female speakers in the user population. The purpose of the evaluation described in this paper was to determine how well the system would operate under these broadened, "real-world" conditions.

There are several "real-world" difficulties which are expected to be adverse to system performance. First, there are the uncontrolled and degraded environmental and transmission conditions encountered during the recording of sample utterances. Environmental conditions involve acoustic background noise and disturbances generated at the user's end and by telephone equipment. Transmission conditions involve signal modification over dialed-up telephone lines. Telephone transmission is nominally over a band from 300 to 3000 Hz. The roll-off at 300 Hz may be gradual due to the attenuation characteristics of the carbon-button transmitter, or quite sharp due to the attenuation characteristics of repeaters in some toll lines. Moreover spectral and phase distortions and variations are likely to be encountered.

The second class of problems is largely behavioral. For example, can a stable and adequate initial reference file be established based on a small number of sample utterances collected in one sitting? Also, can both day-to-day and long-term variations in speaking behavior be tolerated and tracked, and can reference files be updated to reflect these changes in behavior?

Since the principal goal of this evaluation was to study the effect of these "real-world" conditions rather than to achieve optimum performance, the system was made more tractable by using only pitch and intensity features for analysis.

II. SYSTEM OPERATION

Although the operation of the system has been described in previous papers, it will be outlined again here, with departures pointed out, to provide a basis for discussion of the present evaluation.

Figure 1 provides an outline of the system operation in the form of a block diagram. The entire system has been implemented in software on a Data General NOVA 800 laboratory computer. The two

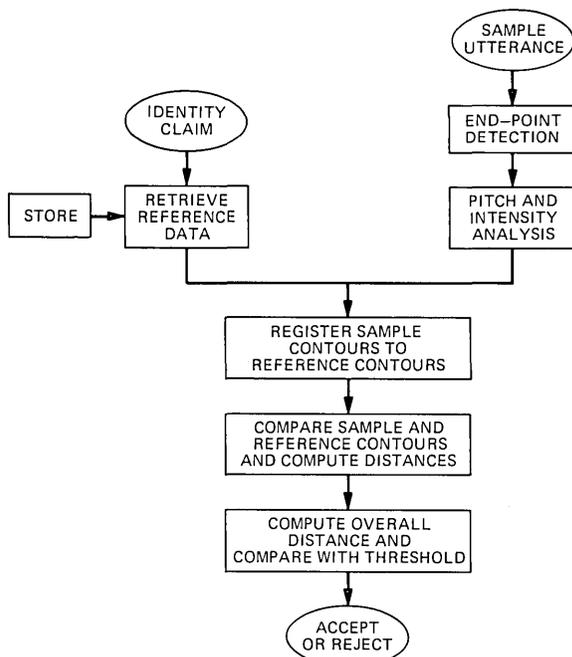


Fig. 1—System-operation block diagram.

inputs to the system are an identity claim and a sample utterance. The sample utterance used for purposes of evaluation is the all-voiced sentence "We were away a year ago." A marked interval is provided for input of the utterance. The input is subjected to 900-Hz, low-pass, analog filtering by two sections of a Rockland 1520 Dual Filter with a combined roll-off of 48 dB/octave. The filtered input is digitized at 10 kHz by means of a 12-bit analog-to-digital converter and stored on disk. The digitized input is then scanned forward from the beginning of the recording interval and backward from the end to determine the beginning and end of the actual sample utterance. The end-point detection is accomplished by means of an energy calculation. The delimited portion is subjected to feature analysis which in this implementation consists of a pitch-and-intensity analysis. Pitch analysis is accomplished by means of the time-domain parallel-processing technique of Gold and Rabiner,⁴ modified by Rabiner for application to telephone speech and extension to female talkers. A pitch period value is obtained every 10 ms through the course of the utterance with resolution to 100 μ s, the sampling period. The resulting pitch contour is smoothed nonlinearly to bridge across unvoiced gaps and to diminish

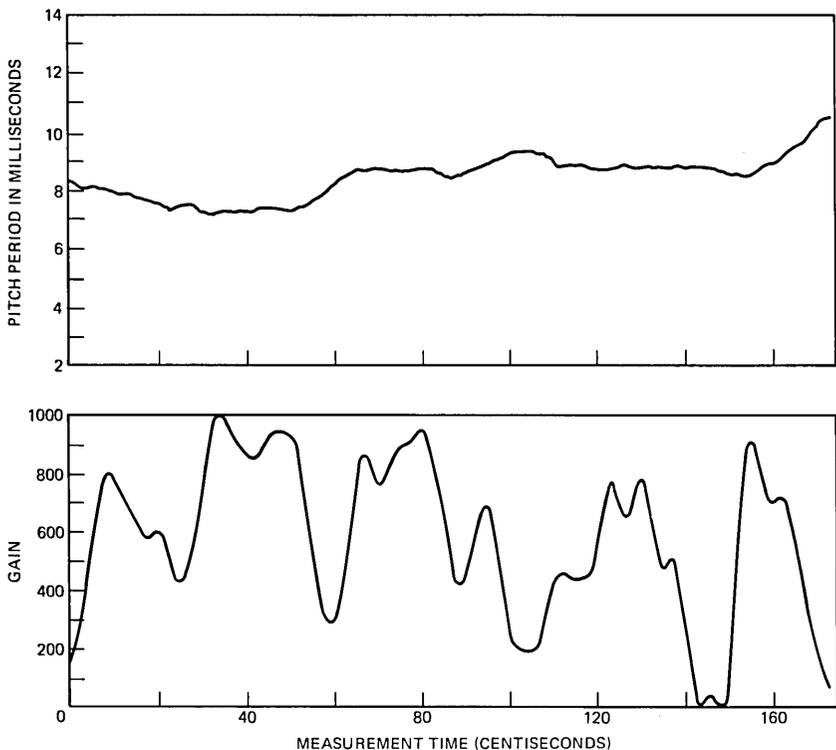


Fig. 2—Smoothed pitch period and intensity contours obtained from analysis of a sample of the test utterance.

the effect of singular values.* The contour is further subject to 16-Hz low-pass smoothing. In addition, an intensity or energy measurement is calculated every 10 ms through the course of the utterance to obtain an intensity contour. This contour is also subject to 16-Hz low-pass filtering and, in addition, is normalized to the peak intensity measurement resulting in a contour of relative intensity values. A typical set of pitch-and-intensity contours is shown in Fig. 2.

These contours comprise the basic patterns for verification. They are compared with a corresponding set of reference patterns associated with the claimed identity. The reference patterns are obtained by averaging and combining a set of patterns obtained from sample utterances of the person whose identity is claimed. (The referencing process has been described in Ref. 2.) Before comparing the sample

*The pitch detector modifications, especially the nonlinear smoothing, were largely motivated by a preliminary study of the effects of telephone transmission on automatic speaker verification by O. M. M. Mitchell.⁵

and reference contours, an additional operation, time registration, is carried out. In this operation, the events of the sample contour are brought into the best possible registration with corresponding events of the reference contour by replotting the sample contour versus a modified function of time. This step is necessary to account for the normal, expected variations in speaking behavior observed in the repetition of a sample utterance by the same speaker. In previous implementations, this operation was accomplished by means of a method of steepest ascent.¹ In the present implementation, a dynamic programming technique is used. The technique is similar to those described by Sakoe and Chiba, Itakura, and Ellis.⁶⁻⁸ The intensity contour is the guide contour for the procedure. The sample intensity contour is linearly stretched or compressed to the normalized length of the reference intensity contour. Then a distance is calculated between the i th point (or set of points) in the sample contour and the j th point (or set of points) in the reference contour for each i and j . The dynamic programming algorithm is used to find the path of least accumulated distances through the matrix of distances $\{d_{ij}\}$. The optimal path $i = I(j)$ $j = 1, \dots, N$ determines the warping function required to replot the sample contour registered to the reference contour. A number of constraints are imposed so that the resulting path does not deviate excessively from the path of no warping $i = j$. The warping function obtained for the intensity contour is also applied to the pitch contour. Time registration of the intensity contour is illustrated in Fig. 3.

Following registration, the pitch and intensity contours are divided into 20 equal-length segments, as shown for intensity in the bottom panel of Fig. 3. In each segment, a set of measurements is applied to both the sample and reference contours and a squared difference is calculated specifying the dissimilarity between these contour segments for each measurement. The squared difference for each measurement and segment is weighted inversely by a variance which is calculated from the set of sample contours used to construct the reference (see Section 2.2). The effect of the variances is to weight most heavily those segments in which a particular measurement was most consistent over the set of sample contours comprising the reference. A distance for each measurement is calculated by summing the weighted squared differences over the 20 segments of a contour. In addition to four distances for each contour, based on segment-by-segment measurements, there is also a distance based on the overall cross-correlation of sample and reference contours. There are also a distance based on the cross-correlation of the pitch and intensity contours and two distances based on the amount of warping required to register the sample

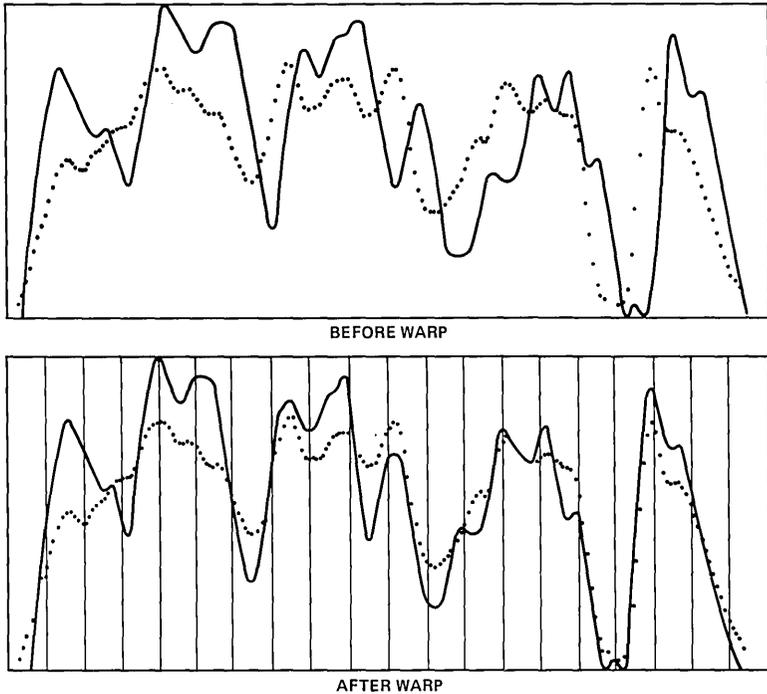


Fig. 3—Time registration of sample and reference intensity contours. The reference contour is plotted using a dotted line in both the top and bottom panel. The sample contour, solid line, is shown with its end points aligned to the reference contour, before internal registration in the top panel, and after internal registration in the bottom panel.

contours to the reference contours. The overall distance is obtained by a simple average over the entire set of 13 individual distances or the average over a subset of these distances selected *a priori* for each speaker. The speaker-dependent distance-selection technique is described in Ref. 3. Finally, the overall distance is compared with a pre-determined threshold to determine whether to accept or reject the identity claim.

2.1 Experimental setup and typical transaction

A block diagram of the experimental setup is shown in Fig. 4. The basic elements are the customer's phone with a *Touch-Tone* dial, a data set (Western Electric 407A), an analog-to-digital converter, and a Data General NOVA 800 computer, in which reside both the automatic speaker-verification programs and a programmed voice-response facility which is used to provide instructions and responses to the customer during each transaction. The voice-response system, which

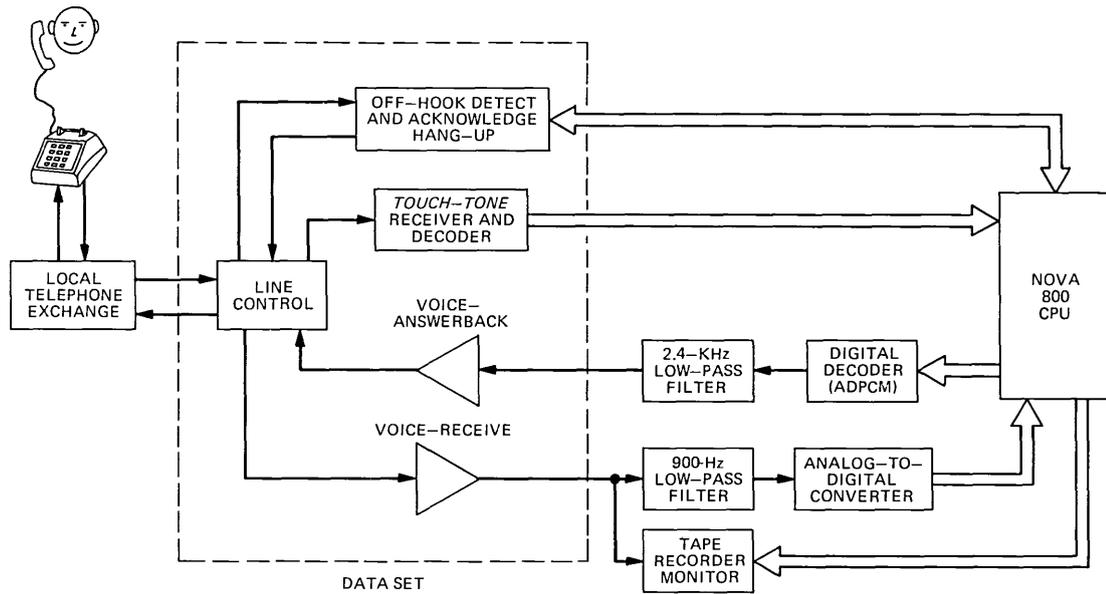


Fig. 4—Experimental setup block diagram showing the principal components of apparatus linking the customer's telephone and the computer.

uses ADPCM-coded speech message units, has been described in a paper by Rosenthal et al.⁹

A typical transaction begins with a customer dialing up the data set from his own premises via the local telephone exchange. The data set acknowledges the incoming call and provides an interrupt to the CPU, which initializes an identifying voice-response message and the instruction to the customer to dial in his identification number via the *Touch-Tone* dial on his phone. The decoded identification number constitutes the customer's identity claim. The customer is then requested to speak the test utterance within a tone-marked recording interval and the input utterance is then processed as described in the previous section. (There are several conditions that can occur during the end-point detection or pitch-and-intensity analysis indicating a bad recording that will cause the customer to be instructed to repeat the test utterance.) Following the analysis and comparison, the customer is advised of the decision to accept or reject his identity claim. A rejection, of course, constitutes a system error, a false alarm. The entire transaction from recording to verification response takes 20 to 30 seconds. A large fraction of this time is consumed by the software pitch detector. A breakdown of the computation times is shown in Table I. The major decision parameters for each transaction are appended to a log file set up for each customer.

One hundred four "customers," approximately evenly divided between adult males and females, participated actively in the evaluation by calling in nominally once each working day over a period of five months. These were all native American or Canadian speakers of English with no overt speech defects. The customers were instructed to speak the test utterance naturally and consistently from day to day.

2.2 Reference information

The establishment and updating of reference information is an important element of the system. On the customer's initial call to the system he is requested to provide five recordings of the test utterance with approximately 10 seconds between each recording. These utterances are analyzed and the analysis data are stored in disk files. These data are used to construct the customer's initial reference file. The actual reference construction is carried out off-line during nonoperating hours and requires about 10 minutes per customer.

The following operations are included in the construction of a reference file:

- (i) *Reference contours.* A set of reference pitch and intensity contours is constructed. The intensity contours obtained from up

Table I — Processing times for a single transaction

| Operation | Time (s) |
|---|------------------|
| Intensity computation & end-point detection | 4 |
| Pitch analysis (3.5 s/s) | 5.5 (Typical) |
| Time registration | 3.5 |
| Comparison & decision | 2 |
| | 15 |
| System overhead | 10 (Approximate) |
| Total | 25 |

to 10 customer sample files are mutually registered and averaged to provide a reference intensity contour. The corresponding sample pitch contours are mutually registered and averaged using the same warping parameters obtained for the reference intensity contour.

- (ii) *Distance weights.* For each kind of measurement made on the contours, a variance is calculated over the sample contours used to construct the reference contours in the form

$$\sigma_j^2 = \frac{1}{N} \sum_{n=1}^N (s_{jn} - r_j)^2,$$

where s_{jn} is the j th measurement on the n th sample contour, r_j is the j th measurement on the reference contour, and N is the number of sample contours.

- (iii) *Measurement selection.* The subset of the original measurement set which is most effective in separating the overall distance distribution of customer and impostor sample utterances is found. Each customer sample file and a sample file from each of 30 different customers of the same sex are used to provide the customer and impostor distributions.
- (iv) *Threshold computation.* Estimates of equal-error thresholds for both the overall distance based on all measurements and the overall distance based on selected measurements are computed. The same sample customer and impostor files used in measurement selection are used to estimate the thresholds.

For the initial reference file, the operating threshold is set at 1.5 times the estimated equal-error threshold. This is done to compensate for the fact that the sample files used to estimate the threshold coincide exactly with the sample files used to construct the reference contours. Moreover, these sample files are obtained from utterances collected

in one session. For these reasons, the measurements and distances extracted from this set of sample files are expected to be highly correlated and are not likely to adequately reflect the expected variation over a series of independent trials. The factor of 1.5 for augmenting the threshold is somewhat arbitrary and was arrived at by inspection of the data from a preliminary experiment. The accept-reject criterion following the initial reference is based on all measurements rather than a selected set of measurements. This, again, is done because the initial sample data files do not adequately reflect the expected range of variation to provide a stable selection of measurements.

Following establishment of the initial reference, the customer calls in nominally once a day. The data files for each trial in which the customer's claim is accepted are saved. When five of these files are accumulated, the reference file is updated. The reference file is updated a second time when five additional "accepted" customer sample files are accumulated. For the fourth reference and thereafter, 10 additional customer files must accumulate. For the fourth and successive references, there are a total of 25 customer files available of which 10 are used to construct the contours and all are used to select measurements and calculate the thresholds. Following reference construction, the oldest 10 customer files are deleted, so that the maximum number of sample files per customer allowed to accumulate is 25. Also, from the fourth reference and thereafter, the operational accept/reject criterion switches over to selected measurements and the threshold is allowed to adapt. The adaptation mechanism is as follows: at each trial for which the customer claim is accepted, if the overall distance is greater than 75 percent of the current threshold, the threshold is increased by 10 percent; if the overall distance is less than 40 percent of the current threshold, the threshold is decreased by 10 percent. The reference updating procedures are summarized in Table II.

Table II — Reference updating procedures

| Ref. No. | No. of Utterances Analyzed | | Threshold Setting |
|----------|----------------------------|-------------------------------------|--|
| | Total | No. Used to Construct Ref. Contours | |
| 1 | 5 | 5 | $1.5 \times \text{EET}^*$ (all measurements) |
| 2 | 10 | 10 | $1.5 \times \text{EET}^*$ (all measurements) |
| 3 | 15 | 10 | $1.3 \times \text{EET}^*$ (all measurements) |
| 4+ | 25 | 10 | EET^* (measurements selected and allowed to adapt to the next update) |

* EET = Equal-error threshold.

2.3 Results

The results are described in terms of the two types of error that can occur: rejecting a (legitimate) customer claim and accepting an (improper) impostor claim. The customer-rejection rate was calculated simply from a tabulation of the number of rejections experienced by the customer population over the five-month-long period of evaluation. For each trial, overall distances were calculated both for all measurements and for selected measurements. However, the operational accept/reject criterion is based on all measurements for the first three references and on selected measurements thereafter, as described in Section 2.2. Excluded from this tabulation were trials in which it was known that an unauthorized person used an identification number, trials in which the customer deliberately altered his utterance, trials during which the analog-to-digital recording system was operating defectively, and trials in which a faulty end-point-detection program produced misaligned analyses. (The latter two situations were quickly rectified.) Also excluded were the trials from customers whose first references failed. A first-reference failure is defined as one in which a new customer obtained three rejections before obtaining the five acceptances necessary for the first update. Approximately 20 or 25 percent of the initial references failed in this way. When this occurred, the customer was requested to provide a new sample of five utterances with which to construct the initial reference. Invariably this new initial reference posed no problems.

A typical customer history is plotted in Fig. 5. Each point represents the overall distance for a particular trial. The current threshold is plotted as a broken horizontal line. An error occurs for each trial in which a point lies above this line, that is, where the overall distance exceeds the threshold. Three such errors occurred over the 76-trial history of this customer. The reference update in effect for each series of trials is indicated by the numbers above the horizontal axis. The threshold was allowed to adapt trial-by-trial following the fifth reference. The general trend for both distance and threshold is an initial elevation followed by a levelling off after 20 or 25 trials. This general behavior is expected and relatively easy to track. The occasional large trial-to-trial variations are not easy to track even with adaptation.

The overall results of the evaluation expressed in terms of average error rates are shown in Table III. Three columns of figures are shown under "reject customer." The "operational" criterion, as already mentioned, uses all measurements for the first three references and selected measurements thereafter. With the operational criterion, the average rejection rate over all customers and approximately 4500 trials is

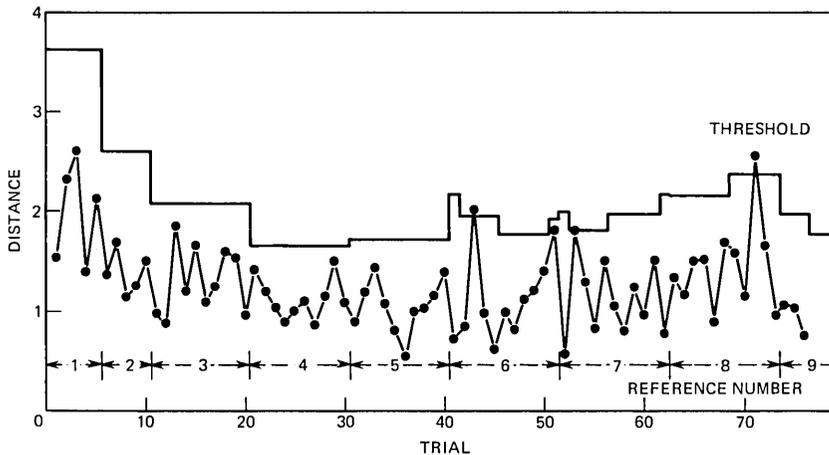


Fig. 5—Individual customer history showing the trial-by-trial overall distance throughout the period of the evaluation. Also shown is the operational threshold and reference updates through the same period.

approximately 9 percent. The median is somewhat less, approximately 7 percent. The error rates for female customers are consistently slightly higher than those for male customers. However, a statistical hypothesis test, derived by a likelihood ratio based on normal population distributions, indicated that the hypothesis of identical populations cannot be rejected at any reasonable level of significance. Under the assumption of identical populations, observed differences in error rates could be expected with a frequency as great as 20 percent. Note that the operational error rate is consistently lower than either the error rate for all measurements or for selected measurements. This is because the all-measurement criterion is preferable for early references and the selected-measurement criterion is preferable for later ones, as shown in Fig. 6. The left half of this figure shows the average reject-

Table III — Average error rates

| Customers | Reject Customer | | | Accept Impostor | | |
|--------------|-----------------|------------------|-----------------------|-----------------|------------------|-----------------------|
| | Operational | All Measurements | Selected Measurements | Operational | All Measurements | Selected Measurements |
| Males (56) | 8.20 | 9.46 | 10.65 | 8.78 | 8.91 | 6.06 |
| Females (48) | 9.92 | 10.37 | 12.2 | 10.92 | 10.94 | 7.54 |
| Total (104) | 8.99 | 9.88 | 11.33 | 9.72 | 9.79 | 6.70 |

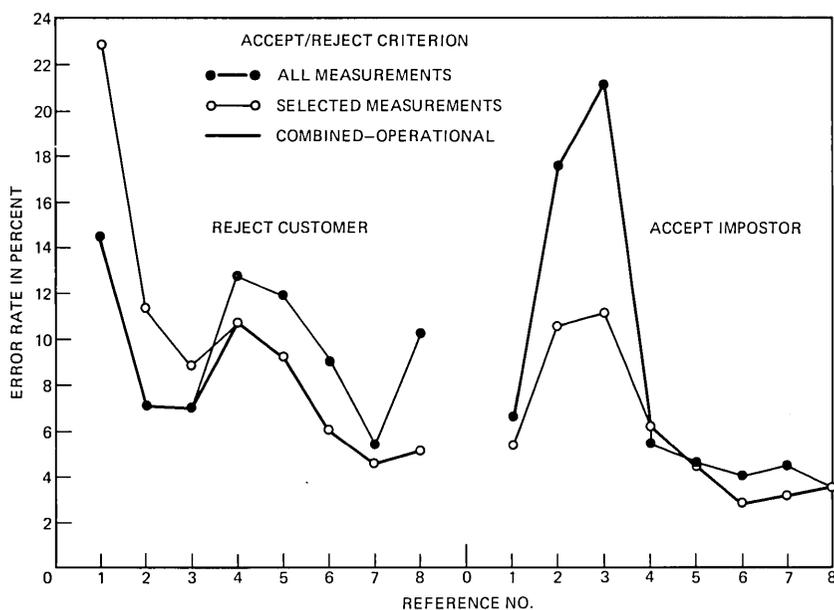


Fig. 6—Average error rates vs reference update. Reject-customer rates are plotted on the left and accept-impostor rates on the right. The thick line in each case shows the error rate under the operational criterion.

customer rate as a function of reference update. The rates for the all-measurement criterion and the selected-measurement criterion are plotted separately and the combined operational strategy is indicated by the thick line. The general trend for both criteria is the same as that observed for the history of a single customer shown in Fig. 5: a high initial error rate followed by a levelling off. It can also be seen that the operational strategy produces the minimum error rate. Overall, at early reference update stages, the reject-customer rate is of the order of 10 percent, while in later stages, the rate approaches 4 percent.

The second part of the error analysis, the tabulation of accept-impostor data, is accomplished differently. It is not practical to provide access to the system to a separate population of impostors for this purpose. Instead, a procedure was set up to systematically cross-compare selected sample utterances of each customer with the references of every other customer of the same sex. The sample file for every tenth accepted utterance of each customer is selected for this purpose, and the results tabulated. This is accomplished during the same off-line period used to update the reference files. The error rates shown in the right half of Table III are based on approximately 16,000 comparisons. The accept-impostor rates are approximately 10 percent

for the all-measurement criterion and 7 percent for selected measurements, indicating a clear advantage for the selected-measurement criterion. Unfortunately, the operational criterion produces no such improvement. The reason for this is seen in the right half of Fig. 6, where the accept-impostor rate is plotted versus reference number. Most of the advantage for selected measurements occurs for customer references in the early stages of updating. At these early stages, however, the operational criterion is the all-measurement one. In the latter stages of stable references there is no clear advantage for either all measurements or selected measurements. In these latter stages, as was the case for the reject-customer rate, the accept-impostor rate is approximately 4 percent.

In summary, error rates of the order of 4 or 5 percent, for both reject-customer and accept-impostor, are obtained for customer references in advanced stages of reference updating. The overall error rates are about twice as high because of the adverse effect of errors occurring at early stages of reference updating.

Another interesting question is the effect on the error rates of varying the threshold. To determine the effect of threshold variation, all the customer log files were scanned by varying the actual thresholds in steps of 10 percent and tabulating the number of errors at each step by comparing the actual overall distances with the varied threshold. The results are shown in Fig. 7. All the actual thresholds are normalized to 1. Thresholds are plotted for the operational accept/reject criterion. At the normalized threshold value of 1, the error rate is approximately 9 percent, the same value shown in Table III. The plot of accept-impostor error rate as a function of threshold was obtained by comparing selected customer samples left on file at the end of the experiment with customer references. (Since the comparison data is from the period at the end of the experiment, most references were in an advanced updating stage and the error rate at the normalized threshold value of 1 is approximately 5.5 percent, considerably less than the 9.7 percent rate obtained throughout the entire period of the experiment, as shown in Table III.) It is possible to get a feeling for the amount of tradeoff obtainable when the threshold is varied. For example, if the threshold is set for a reject-customer rate of 4 percent, the corresponding accept-impostor rate is approximately 15 percent. Conversely, if the threshold is set for an accept-impostor rate of 4 percent, the corresponding reject-customer rate is approximately 11 percent.

Finally, it is of interest to survey individual error rates to get a feeling for the range of performance over the customer population. Histograms of individual error rates have been plotted in Fig. 8. The

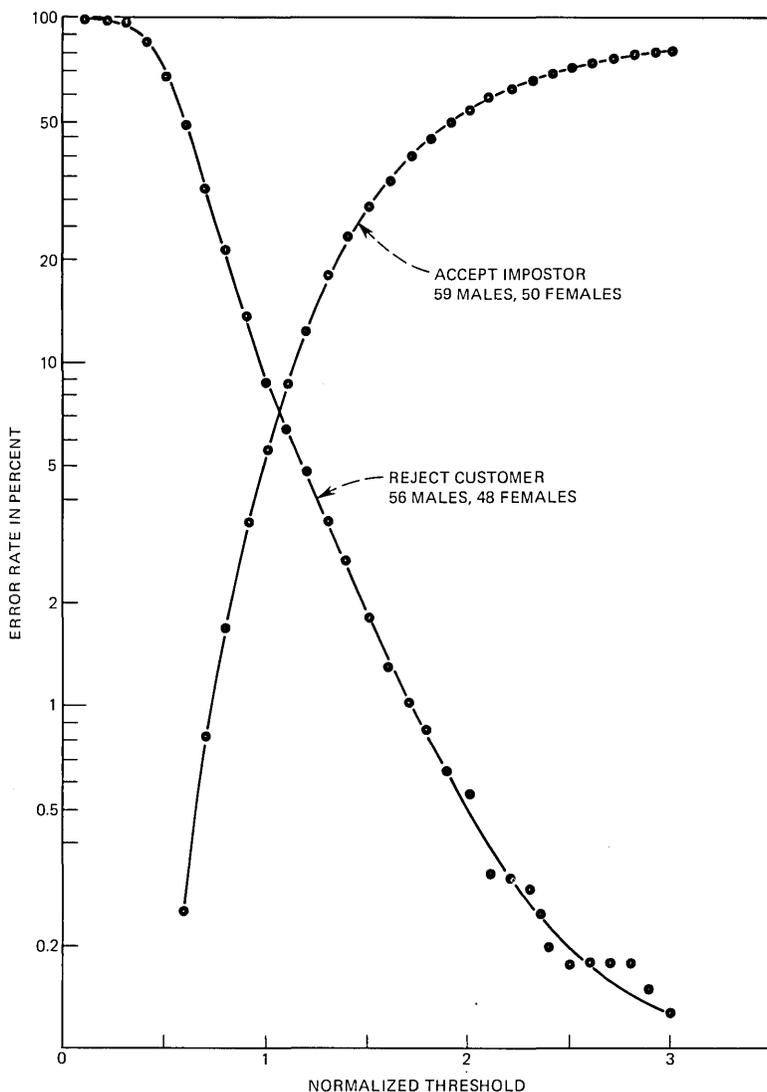


Fig. 7—Error rate vs normalized threshold. The effects are given of threshold variation on the reject-customer and accept-impostor rates using the operational accept/reject criterion.

top half shows the distribution of individual reject-customer rates under the operational criterion. About 80 percent of the customers have error rates less than 15 percent. There is, however, a small fraction of the population with excessively large rates of rejection. Some of these customers have been identified as special cases and will be

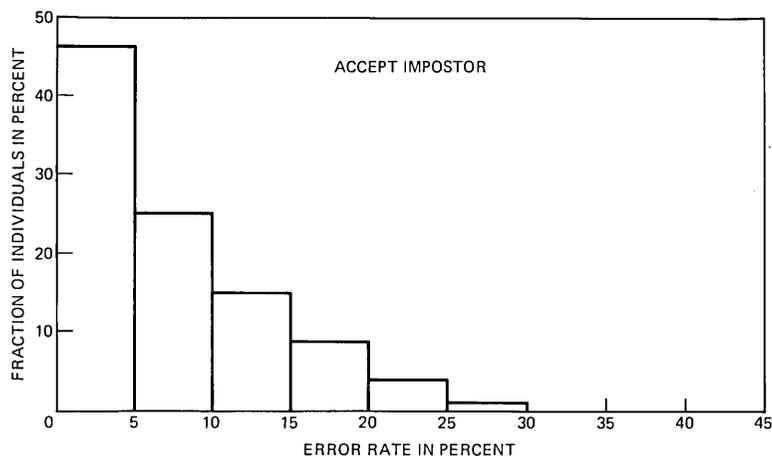
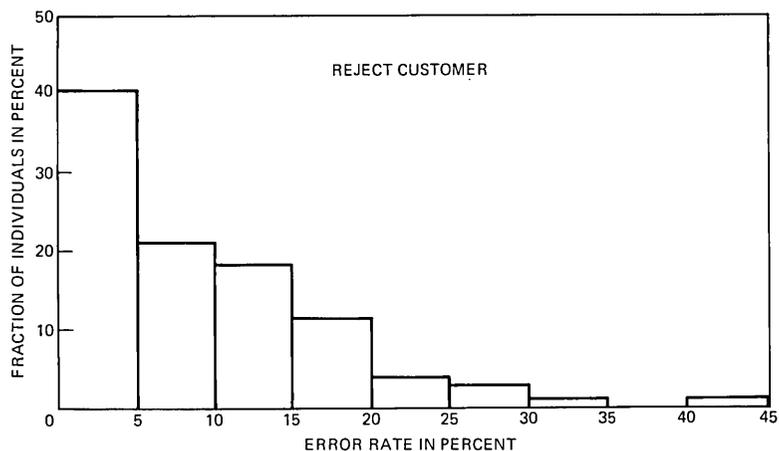


Fig. 8—Histograms showing the distribution of individual error rates over the 104-customer population. The top half shows the reject-customer distribution while the bottom half shows the accept-impostor distribution.

discussed in the next section. In the bottom half of the figure, a similar histogram is shown for the accept-impostor rates under the selected distance criterion. This distribution is somewhat tighter, with fewer outliers than the reject-customer distribution.

III. DISCUSSION

As stated in Section I, the purpose of this evaluation was to determine how well the system could perform under “real-world” conditions. First, “real-world” conditions make it difficult for reference files to adequately represent and keep pace with normal variations in speaking behavior. To avoid encumbering the customer, the initial

reference file is constructed from a small number of highly correlated utterances collected in one sitting. Two questions are (i) how adequate is this initial reference file, and (ii) can it be satisfactorily updated and adapted to track both trial-to-trial and long-term variations in speaking behavior? In contrast, in previous laboratory evaluations, reference files were constructed from independent samples generally spanning a relatively long period of time with a distinct set of test samples spanning the same period of time.

Second, the recordings are obtained over dialed-up lines from the customer's own telephone and are therefore uncontrolled and degraded in comparison with the carefully executed wideband recordings of previous laboratory evaluations.

In a previous laboratory evaluation, the equal-error rate using a pitch-and-intensity analysis was approximately 6 percent with the all-measurement criterion and 3 percent with the selected-measurement criterion.* An error rate of approximately 5 percent would be considered quite satisfactory for the "real-world" evaluation. In fact, an overall error rate of approximately 9 percent was obtained. As anticipated, the error rate varied considerably over the course of the customers' access and reference-update history. For well-established and adapted customer references, an error rate of approximately 4 percent was obtained, which is quite acceptable. However, the error rate of 15 or 20 percent obtained for initial customer references is unsatisfactory.

This "start-up" or initialization problem is in fact compound. In the first place, the number of sample utterances available during the first few reference updates is only marginally adequate to calculate reliable reference data. This is especially true for the initial reference in which the same five utterances are used to construct contours as well as to calculate weights and thresholds. It is preferable to have an independent set of utterances with which to calculate weights and thresholds. The second part of the "start-up" problem is the customer's talking behavior. More than likely there will be large variations in talking behavior from trial to trial through the "warm-up" period of early trials until a stable or habituated talking pattern is established.

As just mentioned, the initial reference is a special case because the five sample utterances used to calculate the reference data are collected in one session. Generally, the talking behavior from utterance to utterance in this session will be highly correlated. The reference data calculated from this set of utterances will therefore be quite "tight,"

* With a predictor coefficient analysis added to the pitch-and-intensity analysis, this same evaluation yielded a 3-percent and a 1-percent error rate for all and selected measurements, respectively.

encompassing only the limited range of behavior of that session. The utterances in subsequent sessions, however, can be expected to vary considerably from those in the initial session and from each other due to variations in behavior during the "warm-up" period. Without compensation, then, a large customer-reject rate can be expected for the trials immediately following the initial reference. The compensation that was attempted was to augment the calculated threshold by 50 percent. However, it is clear that this was not sufficient since 20 or 25 percent of the initial references "failed," as described in the previous section, and the elevated rejection rate was 14 percent or 15 percent at this stage, as shown in Fig. 6. Additional augmentation could well be tolerated since the accept-impostor rate at this stage is only 6 percent. Thus, the threshold could have been adjusted to yield approximately a 10-percent equal-error rate.

The situation is quite different at the first reference update, reference no. 2. At this stage, the sample utterances used to calculate the reference data consist of the five utterances from the initial session plus five more obtained in succeeding trials. The latter, as just mentioned, can be expected to vary considerably from the initial set and from each other. The result is a set of reference data which is likely to be considerably "looser" than the initial set. Consequently, there is a considerably reduced customer-reject rate together with a considerably increased impostor-accept rate. The same situation still holds true at the second reference update when an additional five utterances are added to the reference data. In both of these reference updates, the calculated equal-error threshold was augmented. This is perhaps not the correct strategy in light of the error-rate trend.

From the third reference update, reference no. 4 and thereafter, the error rates decrease generally monotonically. It is reasonable to believe that a better choice of thresholds for the first three references could provide a strictly monotonically decreasing equal-error rate through the course of reference updates, say from 12 percent to 4 or 5 percent.

To provide a reduced-error rate at initial stages, a larger set of initial samples, preferably recorded at different sessions, could be obtained. Another approach which may be more practical is to provide an additional decision category, that of "repeat" or "defer decision." Thus, if the distance on any trial is within a specified fraction of the threshold, additional utterances are requested until a decision can be made. This strategy is useful because there is a considerable probability of acceptance on a trial after rejection on a previous trial. In this experiment, a tabulation of the frequency of acceptance following rejection showed that 82 percent of the trials in which a customer was rejected were followed by trials in which the customer was accepted.

(For well-adapted customer references, reference no. 5 and thereafter, this figure increases to 91 percent.)

Another tabulation was carried out to assess the effect on error rate of withholding decision on trials for which $|D/T - 1| < \Delta$, where D and T are the overall distance and threshold, respectively. For $\Delta = 0.05$, the customer-reject rate fell from 8.8 percent to 7.4 percent, a 15-percent improvement, with decisions withheld on 4 percent of the trials. For the same value of Δ , a 22-percent improvement was noted for the accept-impostor rate. For $\Delta = 0.1$, the customer-reject rate fell by 24 percent and the accept-impostor rate by 39 percent, with decisions withheld on approximately 7 percent of the trials.

The system was shown to be capable of adapting to long-term variation in speaking behavior by means of periodic updating of reference data and the use of an adaptive threshold. Regular and continuous usage of the system is probably necessary for such adaptation. Several customers had absences of two or more weeks during the course of the experiment but were not rejected with greater frequency on their return. However, the sample was too small to be conclusive. It seems likely, in fact, that many individuals will have difficulty being accepted with respect to old reference patterns after prolonged absences.

The other principal "real-world" condition of interest was the effect of transmission and background noise over dialed-up telephone lines and the background noise at the calling location. Although this was not studied in any concentrated or organized manner, generally speaking, the system seemed quite tolerant of the transmission and background conditions encountered. Over a nine-day period the following levels were monitored. The standard deviation of peak speech levels was approximately 8 dB. The background noise during recording was approximately 35 dB below the average peak speech level. Occasionally, background noise reached levels 10 dB higher than this average level. (A background noise level greater than -18 dB with reference to average peak speech level, or a peak speech level less than -12 dB with reference to average peak speech level, would prompt a request for a new recording.) Peak speech levels for female speakers were generally 2 or 3 dB below peak levels for males, while levels for calls originating from outside the local exchange were about 3 dB below the levels for local calls. (Approximately 25 percent of the total population called from outside the local exchange, generally via toll lines.) The only condition observed to be definitely detrimental was recording in the presence of pulse-like background noise, such as the kind originating from a typewriter or teletype console in close proximity to the telephone transmitter. One customer who habitually called under this condition is one of the outliers in Fig. 8.

Abnormal voice conditions were also monitored in an informal way. Mild upper respiratory infections were not observed to have any effect on the customer-rejection rate. However, a fairly severe case of laryngitis was observed for which the speaker could not provide an acceptable utterance because his voice "broke" at each attempt. The most severe voice condition problem observed was diplophonia. Diplophonia is a condition associated with a husky or "raspy" voice quality. An inspection of the speech waveform for diplophonic individuals reveals voiced speech intervals in which alternate pitch periods are more strongly correlated than adjacent pitch periods. Pitch analysis for such speakers is quite difficult. Two customers, one male and one female, were observed to be diplophonic. The female speaker provided a rejection rate of 44 percent. It seems clear that for any system making use of pitch analysis, diplophonic speakers should be identified and, if possible, pitch analysis modified or eliminated.

It is useful to keep in mind the results of some previous supplementary experiments in the light of the present evaluation. One experiment assessed the ability of human listeners to perform a speaker-verification task.¹⁰ Using the same 8-customer, 32-casual-impostor speaker set used in the earliest evaluation of the automatic system, listeners performed at an average equal-error rate of 4 percent in a series of A-B comparison trials. That is, in 4 percent of the trials in which the speakers were the same, the listeners judged them to be different, and in about the same fraction of the trials in which the speakers were different, the listeners judged them the same. This performance level is somewhat better than what has been observed for this telephone evaluation, but not as good as previous laboratory evaluations.

Another study¹¹ indicated that intensively trained professional mimics have considerably better chances for acceptance than casual impostors. With pitch-and-intensity analysis alone, mimic acceptance was found to be of the order of 15 percent when thresholds were set for a customer-reject rate of 1 percent. With the inclusion of predictor coefficient analysis, mimic acceptance falls sharply to a tolerable 4-percent level. Thus, pitch-and-intensity analysis by itself in a speaker-verification system may be rather vulnerable to the efforts of determined mimics.

The practical acceptability of a speaker-verification system depends on its expected error rate and its intended application. There are many noncritical screening applications in which the present error rate of 9 percent is acceptable. This error rate can be improved considerably (perhaps by a factor of 2 or more) with the use of individual

test phrases and multiple-phrase sequential strategies. For applications with more stringent requirements, say error rates of the order of 1 percent and good protection against mimics, extended analysis techniques are necessary. As already mentioned, a laboratory evaluation which included predictor coefficient analysis satisfied these requirements.

IV. SUMMARY

The feasibility of operating the automatic speaker-verification system in the "real world" has been demonstrated. The system proved to be tolerant of many of the degraded and uncontrolled transmission and environment conditions which occur in the "real world" when customers access the system from their own premises via dialed-up telephone lines. The error rate obtained for stable and established reference patterns is approximately 5 percent, which is quite acceptable considering the abbreviated analysis used in the experiment. The greatest weakness seems to lie in the establishment of adequate initial reference patterns. It is felt that at least a partial remedy for this difficulty can result from a collection of initial samples at more than one recording session and the use of a "deferred-decision" category if the distance is within a specified tolerance of the threshold.

V. ACKNOWLEDGMENTS

A project of this complexity required the assistance of many people. The author wishes to express his appreciation to Don Bock who was responsible for providing and maintaining the system hardware, to Nancy Graham for writing the dynamic programming routines, to Carol McGonegal and Kathy Shipley for their programming assistance, to Judy Dudgeon for recruiting the "customers" and monitoring the daily experimental data, and to all the "customers" who cheerfully participated in the experiment.

REFERENCES

1. G. R. Doddington, "A Computer Method of Speaker Verification," Ph.D. Dissertation, Department of Electrical Engineering, University of Wisconsin, 1970.
2. R. C. Lummis, "Speaker Verification by Computer Using Speech Intensity for Temporal Registration," *IEEE Trans. Audio and Electroacoust.*, *AU-21* (April 1973), pp. 80-89.
3. A. E. Rosenberg and M. R. Sambur, "New Techniques for Automatic Speaker Verification," *IEEE Trans. Acoustics, Speech and Signal Processing*, *ASSP-23* (April 1975), pp. 169-176.
4. B. Gold and L. R. Rabiner, "Parallel Processing Techniques for Estimating Pitch Periods of Speech in the Time Domain," *J. Acoust. Soc. Am.*, *46* (August 1969), pp. 442-448.
5. O. M. M. Mitchell, "Speaker Verification via Telephone," private communication.

6. H. Sakoe and S. Chiba, "A Dynamic Programming Approach to Continuous Speech Recognition," Proc. Int. Congr. Acoust., Budapest, Hungary, 3 (1971), pp. 65-68.
7. F. Itakura, "Minimum Prediction Residual Principle Applied to Speech Recognition," IEEE Trans. Acoustics, Speech, and Signal Processing, *ASSP-23* (February 1975), pp. 67-72.
8. J. H. Ellis, "Algorithm for Matching 1-Dimensional Patterns in Multidimensional Space, with Application to Automatic Speech Recognition," Electron. Lett., 5 (July 24, 1969), pp. 335-336.
9. L. H. Rosenthal, L. R. Rabiner, R. W. Schafer, P. Cummiskey, and J. L. Flanagan, "A Multiline Computer Voice Response System Utilizing ADPCM Coded Speech," IEEE Trans. Acoustics, Speech, and Signal Processing, *ASSP-22* (October 1974), pp. 339-356.
10. A. E. Rosenberg, "Listener Performance in Speaker Verification Tasks," IEEE Trans. Audio and Electroacoustics, *AU-21* (June 1973), pp. 221-225.
11. R. C. Lummis and A. E. Rosenberg, "Test of an Automatic Speaker Verification Method with Intensively Trained Mimics," J. Acoust. Soc. Am., 51 (January 1972), p. 131(A).

Statistical Evaluation of the Error Rate of the Fiberguide Repeater Using Importance Sampling

By P. BALABAN

(Manuscript received December 12, 1975)

Simulation of repeaters for an optical-fiberguide digital-communication system requires the calculation of the statistical error rate of the signal. The calculations of these error rates are difficult because of the non-gaussian nature of the noise in the optical detector. In this paper, statistical techniques are described which are useful for simulating tails of distributions. In particular, the importance-sampling procedure is used to modify the probability densities of the input values in a way that makes simulation possible. Application of this procedure gives more accurate results in reasonable computer times. The method is applied to the calculation of the error rate of a fiberguide repeater. Realistic examples are simulated. Results compare favorably with experimental measurements. The number of samples needed for simulation was reduced by five to six orders of magnitude. An alternative numerical solution to the problem is also developed.

I. INTRODUCTION

Traditional computation of error rates of digital repeaters is often done by assuming that the noise is gaussian in nature. For a linear equalizer, the error rate at a given sampling time can then be easily computed if the signal and standard deviation of the noise are given.¹

The fiberguide repeater has essentially three kinds of noise sources. Two noise sources are located in the detector, while the third is a thermal gaussian noise generated by the input stage of the preamplifier (see Fig. 1).

The detector, which is an avalanche photodiode is governed by two intrinsic noise processes: the primary photo current, which is a Poisson process, and the avalanche multiplication process which produces the gain in the diode. Neither of these processes is gaussian.

The probability density functions (PDF) for the output current of the diode were derived by Personick^{2,3} and McIntyre.⁴ The expressions

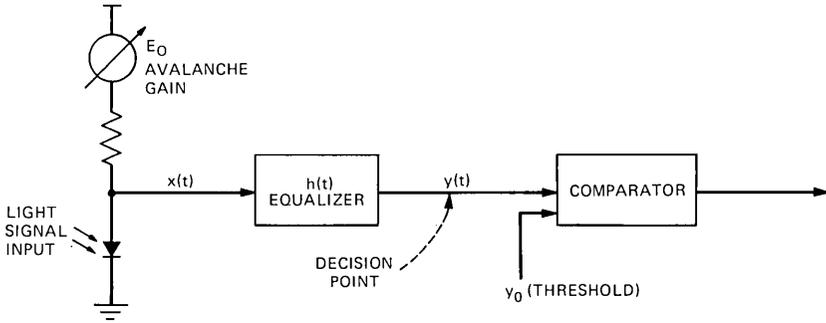


Fig. 1—Optical repeater.

describing these distributions are quite complex, and excessively time consuming for computer simulation of the repeater. An approximation to these probability densities was obtained by Webb, McIntyre, and Conradi (wmc).⁵ The approximate PDF (referred to below as the wmc PDF) is, according to the authors, fairly accurate in the practical operating range, and is easily implemented on the computer.

The wmc PDF for the current at the output of the diode is:

$$p(x) = \frac{1}{\sqrt{2\pi} \sigma \left(1 + \frac{x - M}{\sigma\lambda}\right)^{\frac{3}{2}}} \exp\left\{\frac{-(x - M)^2}{2\sigma^2 \left(1 + \frac{x - M}{\sigma\lambda}\right)}\right\}, \quad (1)$$

where

- $M = (n_e q / T) \mathbf{G} =$ mean output current
- $q =$ charge of an electron
- $T =$ time interval
- $\mathbf{G} =$ average avalanche gain
- $n_e = (\eta T / \hbar \Omega) \mathbf{P}_{\text{opt}} =$ number of primary electrons
- $\eta =$ optical conversion efficiency
- $\mathbf{P}_{\text{opt}} =$ average optical power in time T
- $\hbar \Omega =$ energy of a photon
- $\sigma^2 = n_e \mathbf{G}^2 F_e (q/T)^2 =$ variance of the diode current
- $\lambda = \sqrt{n_e F_e} / (F_e - 1)$
- $F_e = k_{\text{eff}} \mathbf{G} + [2 - (1/\mathbf{G})](1 - k_{\text{eff}}) =$ excess noise factor
- $k_{\text{eff}} =$ effective ionization ratio (assumed constant over range of \mathbf{G} considered)
- k_{eff} and η are photodiode parameters.

Equation (1) is, more precisely, the PDF of the mean diode current over the time interval T for a given average optical input power,

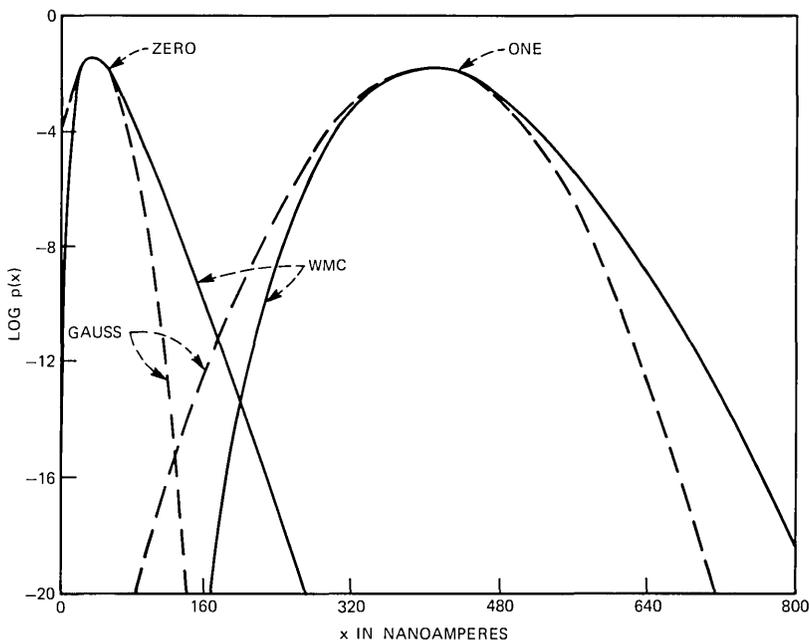


Fig. 2—Probability density function for ONE and ZERO signals using wmc and gaussian approximation.

P_{opt} , and a given average avalanche gain G . It corresponds to the sum over all primary electrons of the product of the primary electron PDF with the avalanche gain PDF. This equation, therefore, combines the effects of both detector noise sources.

In Fig. 2 the solid lines represent current PDFs for two signals: a ONE where the optical power is $P_{opt} = -51$ dBm and a ZERO where $P_{opt} = -61$ dBm. The other parameters are $G = 80$, $\eta = 0.667$, $k_{eff} = 0.04$, $T = 22$ ns. The broken lines represent error PDFs of the same signals and same diode parameters where a normal approximation to the distribution is assumed. The asymmetry of the wmc distributions is clearly seen.

The noise generated by the first stage of the amplifier has to be added to this detector-generated noise.

II. POSSIBLE SOLUTION TO THE PROBLEM

To compute the error rate of the repeater, one has to know the signal probability distribution at the decision-making point, i.e., at the output of the equalizer in Fig. 1.

The output of the equalizer with impulse response $h(t)$ is

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t - \tau)d\tau, \tag{2}$$

where

- $x(t) = x_s(t) + x_{TH}(t)$
- $x_s(t)$ = signal output of the avalanche diode
- $x_{TH}(t)$ = gaussian thermal noise generated by the amplifier.

The problem is to find the distribution of $y(t)$ for given distributions of $x_s(t)$ and $x_{TH}(t)$ and impulse response $h(t)$.

Figure 3 shows the shapes of two distributions: $P_1(y)$ is the cumulative distribution of y for a ONE signal transmitted, $P_0(y)$ is the probability that the output current exceeds y for a ZERO signal transmitted. y_0 is the decision threshold of the comparator (Fig. 1). These distributions are

$$P_1(y) = \int_{-\infty}^y p(y/\text{ONE})dy \tag{3}$$

$$P_0(y) = \int_y^{\infty} p(y/\text{ZERO})dy,$$

where $p(y/\text{ONE})$ and $p(y/\text{ZERO})$ are the respective conditional PDFs. It should be noted that, for practical laser sources, the ZERO input signal could be on the order of 10 percent of the ONE signal,

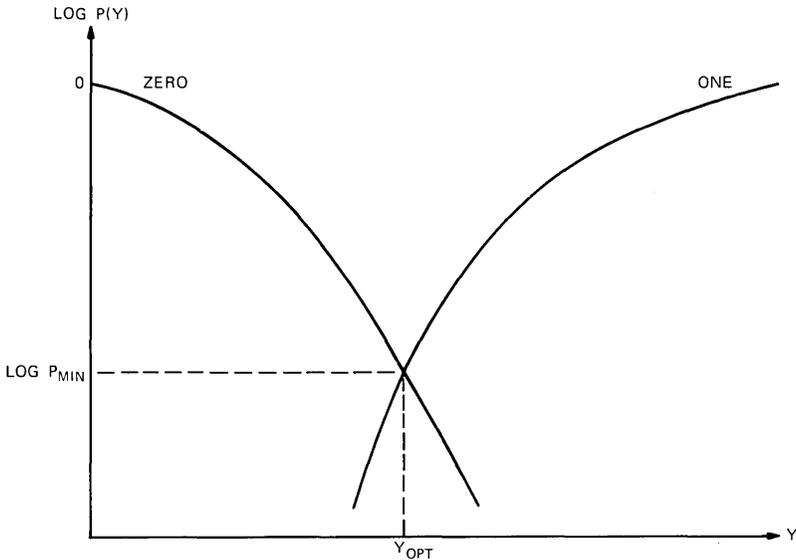


Fig. 3—Distributions for ZERO and ONE signals.

since the laser is not turned off completely. In addition, there is the dark current of the detector diode present. The error rate of a digital repeater is defined as

$$P_e = aP_1(y_0) + bP_0(y_0). \quad (4)$$

(The factors a and b are determined by the *a priori* distribution of the data. For unbiased data, $a = b = 0.5$.)

One of the features that has to be determined is the optimum placement of the decision threshold y_{opt} , that is, where P_e is minimum.

A general analytical solution to this problem is not possible; therefore, the problem has to be approximated numerically or simulated with the aid of a computer. Analysis of the system using a gaussian approximation to the detector noise sources was considered by S. D. Personick.⁶ This analysis, however, predicts optimum decision thresholds considerably different from actual measured values.⁷ Another solution to this problem is to replace the exact solution by a simpler solvable formulation that gives an upper-bound solution. This was done by Personick^{3,8} using the Chernoff bounds. Finding upper bounds is a very effective technique since it gives a degree of confidence. Nevertheless, a method that gives more accurate results is needed to estimate how close the bounds are to the real solutions for this class of problems.

The problem can be solved either by the numerical solution of eq. (3) or by statistical Monte Carlo simulation using importance sampling. Both solutions were attempted and are described below. Particular emphasis was placed on the statistical simulation, since numerical solutions require excessive amounts of computer time.

III. DISCRETE STATISTICAL ANALYSIS (DIRECT COMPUTATION OF OUTPUT PDF)

The output of equalizer $y(t)$ is given in eq. (2). Since we are interested in the output at the sampling time t_0 , it is not necessary to compute the convolution for every t . The value of t_0 is selected to maximize $y(t_0)$:

$$y(t_0) = \int_{-\infty}^{\infty} x(\tau)h(t_0 - \tau)d\tau \quad (5)$$

since

$$\begin{aligned} x(t) &= x_s(t) + x_{\text{TH}}(t) \\ y(t_0) &= \int_{-\infty}^{\infty} x_s(\tau)h(t_0 - \tau)d\tau + \int_{-\infty}^{\infty} x_{\text{TH}}(\tau)h(t_0 - \tau)d\tau \\ &= y_s(t_0) + y_{\text{TH}}(t_0). \end{aligned} \quad (6)$$

The problem is to find the probability distribution of $y(t_0)$.

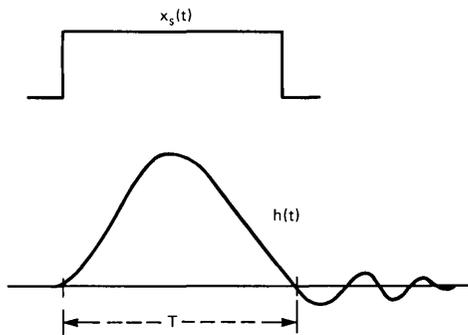


Fig. 4—Input signal and impulse response of the equalizer.

The probability distribution of $y_{TH}(t_0)$ is gaussian since x_{TH} is a thermally generated noise. The standard deviation of y_{TH} is

$$\sigma_{TH} = \sqrt{\frac{4kT B_w}{R_{TH}}}, \quad (7)$$

where B_w is the effective noise bandwidth of $h(t)$ and R_{TH} is the thermal resistance.

The effective noise bandwidth is defined as

$$B_w = \frac{1}{|H_0|^2} \int_0^\infty |H(f)|^2 df, \quad (8)$$

where $H(f)$ is the one-sided Fourier transform of $h(t)$ and $|H_0|$ is the maximum absolute value of $H(f)$. If the impulse response is approximated by a raised cosine of duration τ , $B_w = 3/4\tau$.

Typical waveforms for $x_s(t)$ and $h(t)$ are shown in Fig. 4. The discrete approximation of the signal $y_s(t_0)$ is:

$$y_s = \frac{T}{N} \sum_1^N x_n h_{N-n}, \quad (9)$$

where T is the duration of the signal and N is the number of samples used in the approximation. The samples x_n are independent and have a distribution of the form of eq. (1) for a time interval $\Delta T = T/N$.

3.1 Numerical computation of $p(y)$

If the random variables x_n are independent, then the density of their sum equals the convolution of their respective densities

$$p(y_s) = \frac{1}{\Delta T^N \cdot \prod_n |h_n|} \left[p\left(\frac{x_1}{\Delta T \cdot h_{N-1}}\right) * p\left(\frac{x_2}{\Delta T \cdot h_{N-2}}\right) * \dots * p\left(\frac{x_N}{\Delta T \cdot h_0}\right) \right]. \quad (10)$$

Since the thermal noise is also additive and independent, the total density of the signal is

$$p(y) = p(y_s) * p(y_{TH}). \quad (11)$$

3.2 Fourier transform approach

The simplest and fastest way to compute $p(y)$ is to transform all partial densities using the FFT, multiply them all, and take an inverse FFT:

$$p(y) = \frac{1}{|(\Delta T)^N \prod_n (h_n)|} F^{-1} \left(\prod_n \left\{ F \left[p \left(\frac{x_n}{\Delta T \cdot h_{N-n+1}} \right) \right] \right\} \cdot F[p(y_{TH})] \right). \quad (12)$$

Unfortunately, the Fourier transform is computed as a sum of positive and negative numbers and, therefore, suffers from large numerical errors. Since we are interested in the tails of the distribution (error rate of 10^{-9}), we conclude that the numerical errors will be excessive.

3.3 Direct convolution

Direct convolution produces less truncation error, because densities are positive functions. But direct convolution requires an excessive amount of arithmetic operations, e.g., if we have n distributions each having M points, the number of operations (multiplication and addition) will be

$$N = \frac{n!M^2}{2^n}. \quad (13)$$

For $M = 100$ and $n = 2$, $N = 10^4$ which is a manageable number; but for $M = 100$ and $n = 10$, $N = 3.5 \times 10^7$ and the number of operations becomes excessive. A program using direct convolution was written for $n = 2$. A density for the average power of the input pulses was convolved with the density of the thermal noise. This can be regarded as a response of an integrate-and-dump equalizer (capacitive input).

The results of this convolution are shown in Fig. 5. The solid lines are the distributions for a ONE and ZERO being transmitted. The average power of the ONE in this case is -51 dBm and of the ZERO -61 dBm, with $T = 22$ ns, $R_{TH} = 2.3$ kilohms, and $B_w = 31$ MHz, which corresponds to a $\sigma_{TH} = 14.68$ nA.

The decision threshold θ is usually expressed in percent of the mean range of the output current as:

$$\theta = \frac{Y_0 - \bar{Y}_0}{\bar{Y}_1 - \bar{Y}_0} 100 \text{ [percent]}, \quad (14)$$

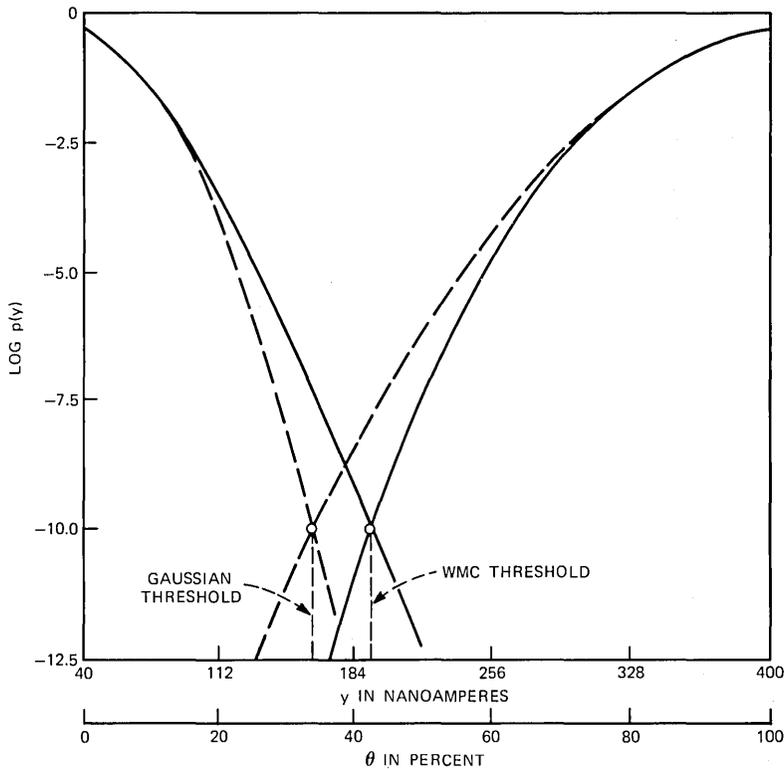


Fig. 5—Distributions for ONE and ZERO signals with thermal noise; wmc and gaussian approximations.

where

$$\begin{aligned} \bar{Y}_1 &= \text{mean output current for ONE} \\ \bar{Y}_0 &= \text{mean output current for ZERO} \\ Y_0 &= \text{threshold current.} \end{aligned}$$

As seen from the graph, the optimum decision threshold θ for this case is 42.3 percent and the minimum error rate is 9.86×10^{-11} . These results were also observed in an experimental setup.⁷ The broken lines in Fig. 5 represent the same input but with the assumption that the probability density of the detector current is normal. As seen from the graph, the resultant threshold is considerably different from the one obtained with the wmc distributions.

IV. STATISTICAL (MONTE CARLO) ANALYSIS

Monte Carlo (mc) analysis has proven over the years to be a very effective method for analysis of systems whose parameters or inputs have known statistical distributions.⁹

However, straightforward application of mc analysis for calculation of error rates is impractical. The reason is that the error rates of interest are in the order of 10^{-6} to 10^{-9} . To have some confidence in the results, 10 to 100 samples have to be in that region of low probability. This will require on the average 10^{11} samples, which is impractical.

To overcome this shortcoming, we introduced a modified sampling method called importance sampling.¹⁰ Conceptually, the idea is simple. If the regions which contribute to the result are known, modify the distributions in such a way that more samples are taken from the important regions.

Let us consider the situation where the problem is to estimate the value of

$$\xi = \int_{-\infty}^{\infty} g(x)p(x)dx, \quad (15)$$

where $p(x)$ is a PDF. The importance-sampling procedure consists in introducing another PDF $p^*(x)$, which is preferable for sampling purposes. Then

$$\begin{aligned} \xi &= \int_{-\infty}^{\infty} g(x) \frac{p(x)}{p^*(x)} p^*(x) dx \\ &= \int_{-\infty}^{\infty} g^*(x) p^*(x) dx. \end{aligned}$$

ξ is then estimated from

$$\xi = \frac{1}{N} \sum_1^N g^*(x_i),$$

where x_i is picked from the PDF $p^*(x_i)$.

4.1 Defining the important regions for error-rate computation

The problem is to define the regions where the errors come from. In Fig. 6a is shown the input waveform for a ONE, while Fig. 6b depicts the impulse response of the equalizer and 6c is the PDF of the diode current for one time interval ΔT .

Since the input signal is always positive, it is obvious that for positive h_n most of the errors, that is, outputs that are below the threshold θ , will come from region A in the PDF of Fig. 6c. Conversely, if h_n is negative, most of the errors will come from region C. Exactly the reverse is true for transmitted ZERO signals, since the signal is considered erroneous if it exceeds the threshold θ .

A very similar argument can be constructed for thermal noise, where the "important" regions are the positive tail for a ZERO transmitted and the negative tail for a ONE transmitted.

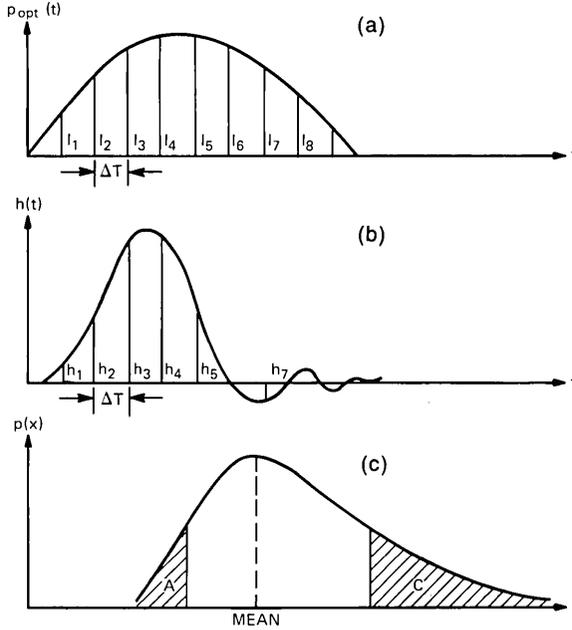


Fig. 6—(a) Optical input signal. (b) Impulse response of the equalizer. (c) PDF of average detector current for ΔT .

4.2 Biasing of the PDF

To implement the importance-sampling technique, the distributions have to be modified so that the important regions are emphasized. The PDFs in the random-number generator that we use are stored in tabular form, as shown in Fig. 7a. The region in question can be biased by a constant multiplier as in Fig. 7b, where $p^* = pB$ for $a \leq x \leq b$ and where the bias B is constant. In Fig. 7c, the PDF is biased such that it is quasi-uniform in the region prescribed. The bias B_i is constant in intervals between two adjacent points in the table. For each interval (x_{i-1}, x_i) , this bias is

$$B_i = \frac{\alpha p_{\max}}{p_i}, \quad (16)$$

where p_i is the value of the PDF at x_i and α is a coefficient specified by the user. The biasing and normalization is done automatically by the program. The output of the random-number generator consists of two values: a random number RV and its associated bias BV . BV is the bias of the interval that RV was picked from.

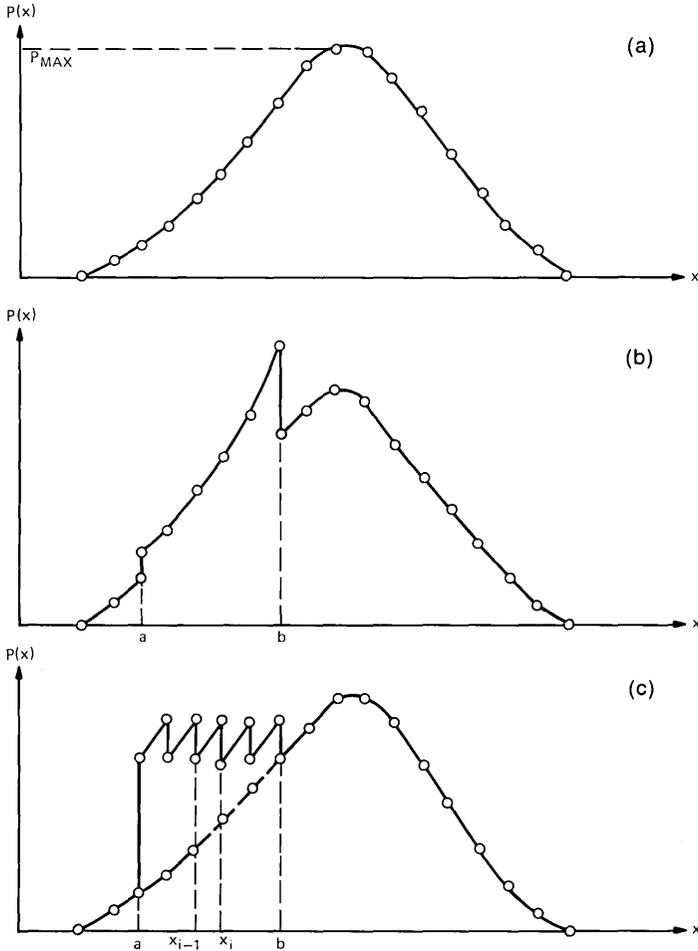


Fig. 7—Probability density functions. (a) Unbiased. (b) Constant bias. (c) Bias for quasi-uniform PDF.

4.3 Computing the histogram

4.3.1 The usual histogram

The standard output of a Monte Carlo analysis is the histogram. Usually the histogram is constructed by subdividing the output variable into M bins, Fig. 8.

The statistical analysis program counts the number of outputs which fall into each bin. For a large number of samples ($N \rightarrow \infty$), the histogram approximates the PDF. The probability p_i that the output y will

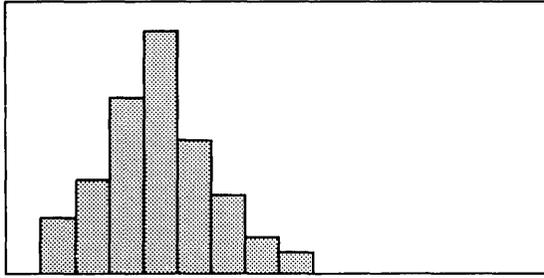


Fig. 8—Histogram.

be in the interval $y_{i-1} \leq y \leq y_i$ is

$$p_i = \lim_{N \rightarrow \infty} \frac{n_i}{N}, \quad (17)$$

where

$$\begin{aligned} n_i &= \text{number of samples in bin } i \\ N &= \text{total number of samples.} \end{aligned}$$

The accuracy of estimation of p_i can be established by¹¹

$$\epsilon = 2 \frac{\sigma_i}{p} = 2 \sqrt{\frac{(1-p)}{Np}}, \quad (18)$$

where $\pm \epsilon$ is the 95.4-percent confidence limit of p_i , p is the desired probability, and σ_i is the standard deviation of p_i .

4.3.2 Histogram for importance sampling

Let us assume we have M independent random variables $x_1, \dots, x_m, \dots, x_M$,

$$y = f(x_1, \dots, x_m, \dots, x_M),$$

$p^*(x_m)$ is the biased PDF of the random variable x_m . How do we compute the unbiased histogram for y .

If a sample x_{mi} is selected from $p^*(x_m)$, its probability to be selected is increased by the bias B_{mi} . Therefore, the weight of the sample x_{mi} is

$$w_{mi} = \frac{1}{B_{mi}}.$$

Since y_i is determined by M independent variables x_m the probability of y_i , to be selected from these regions, is increased by the product of all the biases of x_m . Therefore, the weight of y_i is $w_i = 1/(\prod_m B_{mi})$. The average weight of all the n_j samples that fell into bin j is given by

$$w_{avj} = \frac{1}{n_j} \sum_{i=1}^{n_j} w_{ij} = \frac{1}{n_j} \sum_{i=1}^{n_j} \left(\frac{1}{\prod_m B_{mi}} \right)_j. \quad (19)$$

The probability that the unbiased output will be $y_{j-1} \leq y \leq y_j$ is

$$p_j = w_{avj} \frac{n_j}{N} = w_{avj} p_j^* \quad (20)$$

Since this method increases the probability p in eq. (18), the error ϵ is reduced considerably.

$$\epsilon = 2\sqrt{\frac{(1-p^*)}{Np^*}} = 2\sqrt{\frac{w_{av} - p}{Np}} \quad (21)$$

For error-rate computation $p \ll 1$,

$$\epsilon \approx \frac{2}{\sqrt{Np^*}} = 2\sqrt{\frac{w_{av}}{Np}} \quad (22)$$

and since p_j^* is increased in the interesting areas by $1/w_{avj}$, the confidence limit is decreased on the average by

$$\gamma = \frac{\epsilon_{IS}}{\epsilon_C} = \sqrt{w_{avj}} \quad (23)$$

where

- ϵ_{IS} = Monte Carlo analysis with importance sampling
- ϵ_C = conventional Monte Carlo analysis.

4.4 Example

A simple example to illustrate the principle of importance sampling is the problem of calculating the probability of obtaining a total of three when one tosses two dice. Each die has six faces labeled from one to six and each face has the probability $p = 1/6$ of being on top.

The problem, of course, can be solved analytically. Any particular combination of the dice has the probability of $1/6 \cdot 1/6 = 1/36$ of occurring. Since there are two combinations which make three (one-two and two-one), the probability of getting a three in a random toss of the dice is $2/36 = 1/18$.

If the probabilities of a one and a two are biased so that they occur twice as often as usual, that is, $p = 1/3$ rather than $1/6$, then the probability of getting a three will be four times as great as in the unbiased case. That is $p = 2/9$ instead of $1/18$. The weight for this example is

$$w_3 = 1/(B_1B_3) = 1/4$$

and the probability of obtaining a 3 is being approximated by

$$\hat{p} = \frac{1}{4} \frac{n_3}{N}$$

The error in eq. (18) is reduced by slightly more than a factor of two, since for

$$p = 1/18, \quad \epsilon = \frac{8.24}{\sqrt{N}}$$

and for

$$p = 2/9, \quad \epsilon = \frac{3.74}{\sqrt{N}}.$$

V. ERROR-RATE ANALYSIS

The importance-sampling method described above was applied to the error-rate analysis of the fiberguide repeater with an avalanche diode detector.

To demonstrate the analysis procedure, we will analyze one specific case. Let us assume that the input to the photodiode is as shown in Fig. 9. The ONE is a square pulse of -51 dBm power of 22 ns duration. ZERO is 10 percent of the ONE power or -61 dBm. The avalanche diode data are

$$\begin{aligned} \text{average avalanche gain} &= 80.0 \\ \text{ionization ratio} \quad k &= 0.04 \\ \text{conversion efficiency} \quad \eta &= 0.667. \end{aligned}$$

The equalizer impulse response $h(t)$ is as in Fig. 10 and the thermal input resistance is $R_{TH} = 2.3$ kilohms.

The pulses (ONE and ZERO) are subdivided into nine intervals [eq. (9)], since this resolution is considered adequate for the impulse response in Fig. 10. Nine distributions* of the form in eq. (1) are computed for the ONE and ZERO pulses.

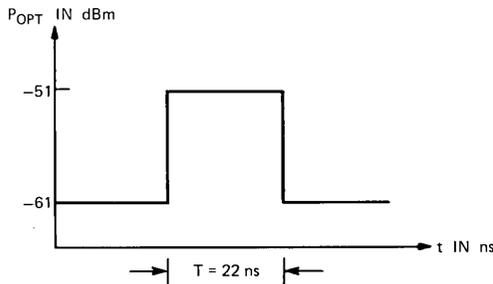


Fig. 9—Optical input.

* In this case the distributions are identical since the pulse is rectangular.

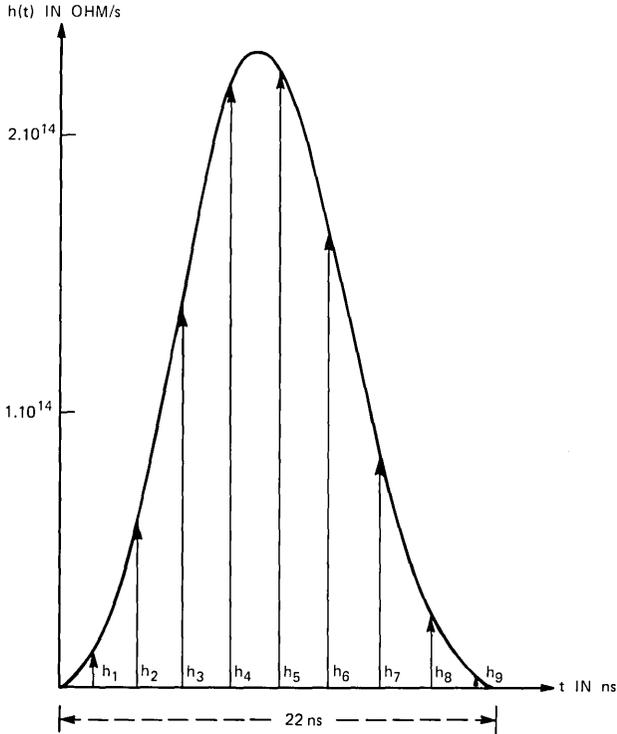


Fig. 10—Impulse response.

The mean for ONE is $M_1 = 400.3 \text{ nA}$

The standard deviation for ONE is $\sigma_1 = 102.66 \text{ nA}$

The mean for ZERO is $M_0 = 40.03 \text{ nA}$

The standard deviation for ZERO is $\sigma_0 = 32.46 \text{ nA}$

The effective noise bandwidth [eq. (8)] for the thermal noise is $B_w = 31 \text{ MHz}$. The standard deviation is computed from eq. (7):

$$\sigma_{TH} = 14.68 \text{ nA.}$$

From the above considerations, the output for N intervals is

$$y = \frac{T}{N} \sum_{n=1}^N x_n h_{N-n} + \frac{T}{N} \sum_{n=1}^N h_n x_{TH}, \quad (24)$$

where the x_n and x_{TH} are samples from the respective PDFs and h 's are discrete values obtained from $h(t)$. The impulse response is normal-

ized to produce unity gain

$$\begin{aligned}
 y^* &= \frac{T}{N} \sum_{n=1}^N h_n \cdot y \\
 h_n^* &= \frac{h_{N-n}}{\sum_{n=1}^N h_n} \\
 y^* &= \sum_{n=1}^N h_n^* x_n + x_{\text{TH}}.
 \end{aligned} \tag{25}$$

In addition, the above case has been analyzed for $N = 1$. This corresponds to the integrate-and-dump equalizer which was analyzed numerically in Section 3.3.

5.1 Computation of distributions and biasing

The calculations of distributions for the diode output are straightforward using the wmc distributions of eq. (1). The biasing is a more complex procedure and sometimes requires a few trials. The idea is to get a reasonable amount of samples in the "important" region of the output. The PDF for ONE was computed for $x_n = 0$ to 800 nA, and was biased according to eq. (16),

$$B_i = \alpha \frac{p_{\text{max}}}{p_i} \text{ for } 0 \text{ to } 400 \text{ nA with } \alpha = 1.$$

The PDF for ZERO was computed from 0 to 400 nA and was biased from 40 nA to 400 nA with an $\alpha = 0.1$. The gaussian distribution for the thermal noise was tabulated for $\pm 7\sigma$. The biased regions were -7σ to 0 for the ONE signal and 0 to 7σ for the ZERO signal with an $\alpha = 0.5$. The PDFs were stored in tabular form for 100 points. Since the ZERO signal has a PDF which requires more detail for small currents (Fig. 2), the independent variable was spaced logarithmically. All others were linearly spaced.

The distributions were computed by a discrete form of eq. (3):

$$\begin{aligned}
 \text{Cum1}(n) &= \sum_{i=1}^n \text{Hist1}(i) \\
 \text{Cum0}(n) &= \sum_{i=n}^M \text{Hist0}(i),
 \end{aligned} \tag{26}$$

where $\text{Hist1}(i)$ and $\text{Hist0}(i)$ are histograms of the outputs.

To calculate the minimum error rate of eq. (4), $\text{Cum1}(n)$ and $\text{Cum0}(n)$ were smoothed and approximated by asymptotic functions. The minimum of the sum of these functions is the minimum error

rate, and the current where this minimum occurs is the optimum threshold.

5.2 Smoothing of the cumulative distributions

It was shown by S. B. Weinstein¹² that for the class of "exponential-type" distributions the error-probability distribution has an asymptote in the tails of the form

$$P_e(x) \sim \exp\left[-\left(\frac{x}{\sigma}\right)^\nu\right]. \quad (27)$$

For $\nu = 2$ the distribution is gaussian.

To obtain better accuracy, this asymptotic function was expanded to

$$P_e(x) = \exp\{-\exp[a(\log x)^2 + b \log x + c]\} \quad (28)$$

or

$$\log(-\log P_e) = a(\log x)^2 + b \log x + c;$$

i.e., the log log of P_e is a second-order polynomial of $\log x$. This equation reduces to Weinstein's form, eq. (27), for $a = 0$.

The second-order polynomials were then least square fit to both Cum1 and Cum0.

5.3 Minimum of error and optimum threshold

The probability of error, eq. (4), is defined as

$$\begin{aligned} P_e(I) &= \frac{1}{2}[P_{e_1}(I) + P_{e_0}(I)] \\ &= \frac{1}{2}(\exp\{-\exp[a_1(\log I)^2 + b_1(\log I) + c_1]\} \\ &\quad + \exp\{-\exp[a_0(\log I)^2 + b_0(\log I) + c_0]\}), \end{aligned} \quad (29)$$

where I is the decision threshold current.

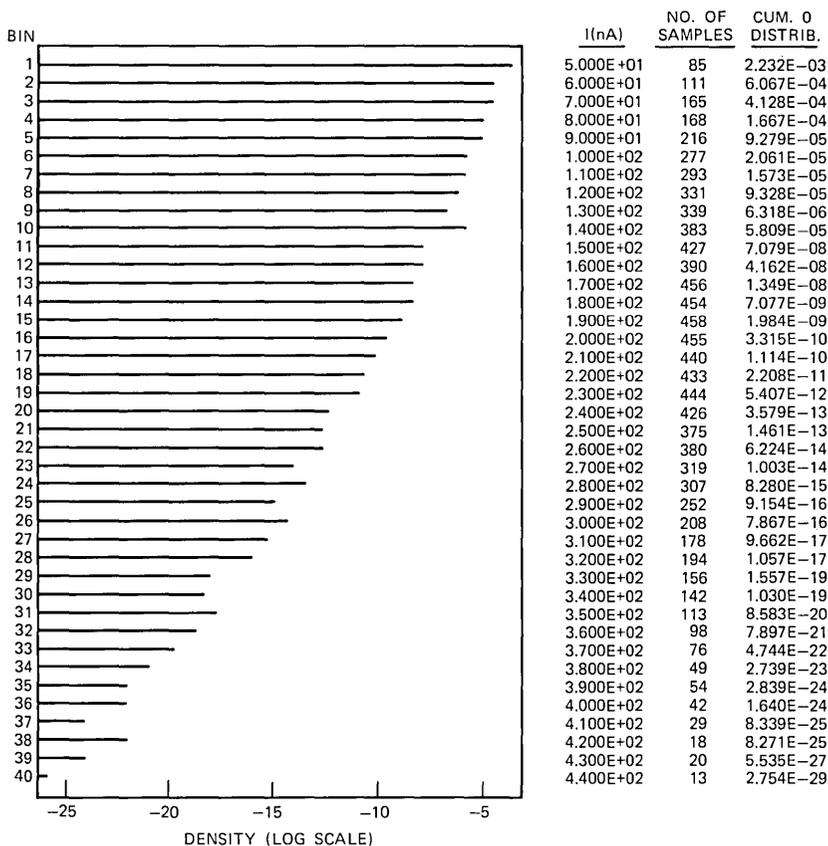
To find the minimum of $P_e(I)$, we compute I_{opt} from

$$\begin{aligned} \frac{dP_e(I)}{dI} \Big|_{I=I_{\text{opt}}} &= 0. \\ P_e(I)_{\text{min}} &= P_e(I_{\text{opt}}). \end{aligned} \quad (30)$$

Equation (30) is solved using the Newton-Raphson method. The threshold θ is usually expressed in percent as in eq. (14).

VI. RESULTS

Figure 11 shows a histogram for the ZERO signal for 10,000 samples. The abscissa is scaled logarithmically. In the tables on the right side are indicated the ordinate values for the bins, the number of samples in each bin, and the distribution for the bin. Below the histogram are the



COEF. OF LEAST SQ. FIT $A_0 = 0.9954$ $A_1 = -1.2324$ $A_2 = 0.53019$
 NUMBER OF SAMPLES = 10,000

Fig. 11—Histogram for ZERO signal.

coefficients of the fitted polynomial $R(I) = a_0 + a_1 \log I + a_2 (\log I)^2$, where $R(I) = \log_{10} [-\log_{10} P_{eo}(I)]$.

Figure 12 shows the smoothed distributions for the above example for one section ($N = 1$) and 1000 samples. The stars represent numerical solutions for the same input parameters, as described in Section III. As seen from the data, if biasing is chosen rightly, a comparatively small sample size is adequate to simulate the extreme tails of the PDF.

Figure 13 shows the smoothed distributions when $h(t)$ is represented by nine sections ($N = 9$). The broken lines represent the distributions for $N = 1$. The sample size was 10,000 in both cases. As seen from the graph in this example, an integrate-and-dump equalizer analysis can be used as a first-order approximation to an equalizer design with an impulse response similar to the $h(t)$ shown in Fig. 10.

The minimum error-rate and optimum threshold were

$$N = 9 \quad P_{e\min} = 3.03 \times 10^{-10} \quad \theta = 43.61 \text{ percent}$$

$$N = 1 \quad P_{e\min} = 6.6 \times 10^{-11} \quad \theta = 42.68 \text{ percent.}$$

Figure 14 again shows the distribution of the above example ($N = 9$). The circled points are results obtained using Chernoff bound approximations as obtained by J. H. Bobsin⁸ using Personick's³ method. As seen from the graph, the optimal threshold predicted by the Chernoff approximation is identical with the one obtained by the statistical analysis. The minimum error rate shown by the Chernoff bound is about two orders of magnitude larger from the statistically simulated value. However, the error rate obtained by the Chernoff approximation is an upper bound. Therefore, if the result falls within the acceptable design specifications, no further analysis is necessary. If, on the other hand, the predicted error rate is marginal, a more accurate analysis technique such as the one described in this paper is advisable.

Figure 15 depicts curves for the minimum error rate $P_{e\min}$ and optimum threshold as a function of the standard deviation of the

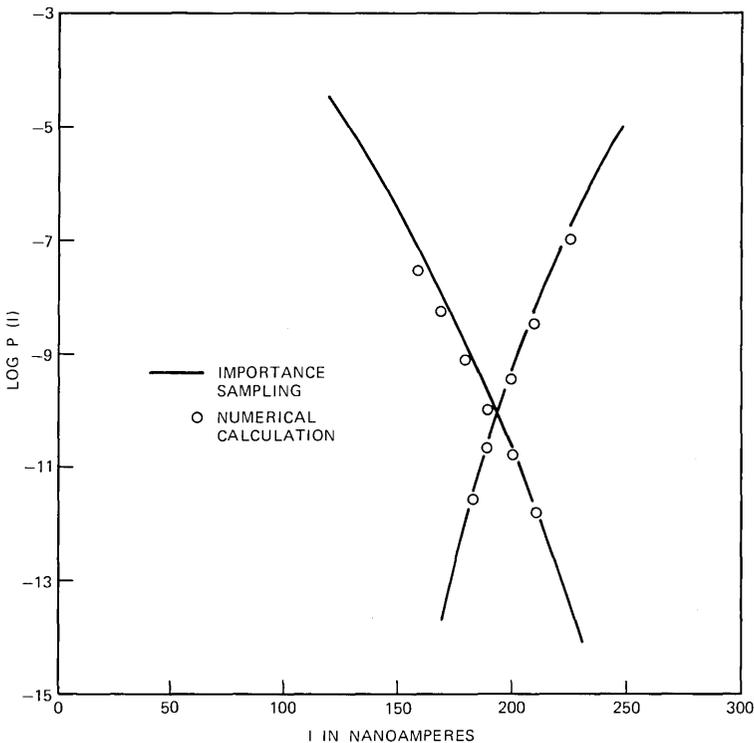


Fig. 12—Distributions for ONE and ZERO signals.

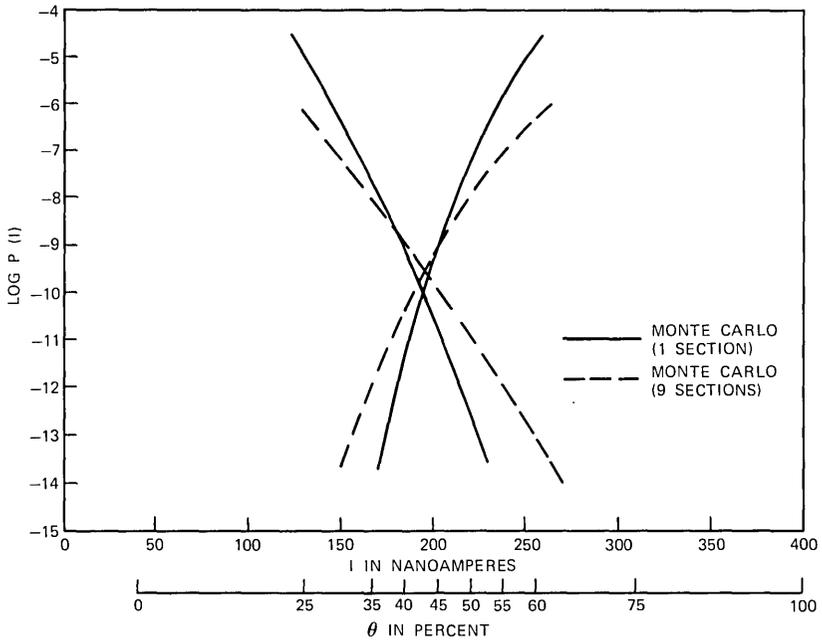


Fig. 13—Distributions for ONE and ZERO signals using Monte Carlo method for 1 and 9 sections.

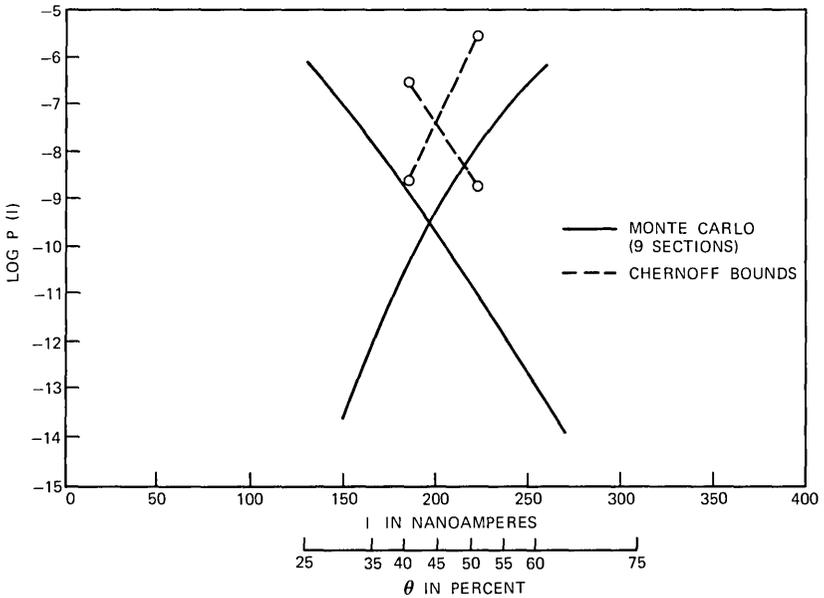


Fig. 14—Distributions for ONE and ZERO signals using Monte Carlo method and Chernoff bounds.

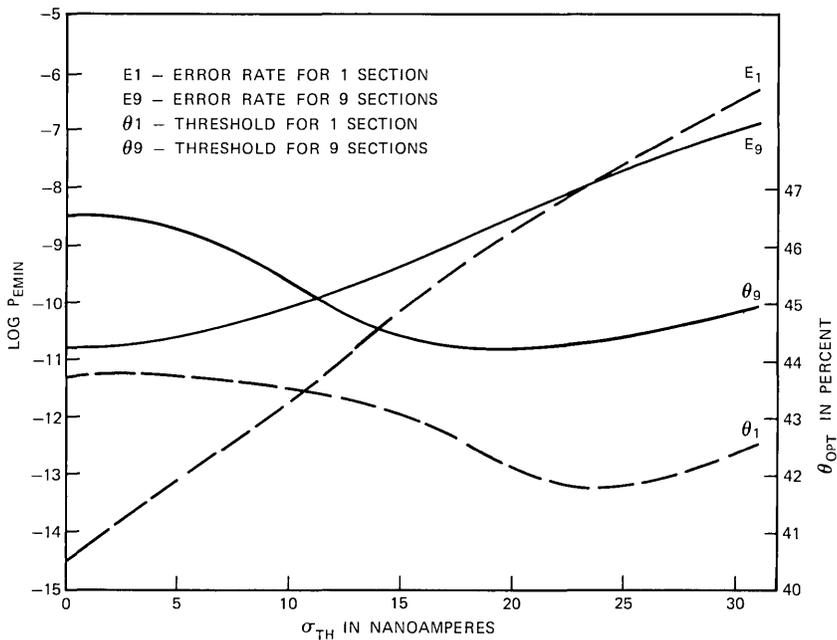


Fig. 15—Minimum error rate and optimum threshold for 1 and 9 sections as a function of thermal noise.

thermal noise σ_{TH} . For this example,

$$\sigma_{TH} = \frac{23.32}{\sqrt{R_{TH}}} \text{ [nA]},$$

where R_{TH} is the input impedance of the equalizer in kilohms.

The solid lines represent $P_{e\min}$ and θ_{opt} for $N = 9$. The broken lines depict the same case analyzed with $N = 1$.

As seen from the graphs, the optimum threshold for both types of receivers is between 42 and 46 percent. Interestingly, the threshold has a minimum at $\sigma = 25$ nA ($R_{TH} = 0.87$ kilohm).

The integrate-and-dump analysis has a lower error rate than the nine-section equalizer analysis in the normal operating range of the repeater.

Throughout this analysis, the avalanche gain was assumed to be constant. In a more realistic design, the avalanche gain can be adjusted to optimize the error rate. This will result in a higher avalanche gain for higher thermal noise.

In this example, the input signals and the equalizer impulse response were chosen so that there is no intersymbol interference. This was done to simplify the example. However, the analysis procedure is not limited to this case.

VII. CONCLUSIONS

The statistical importance-sampling technique was used to calculate error rates of optical repeaters. The application of this technique substantially reduced the number of samples needed for simulation of the repeater and thus made computer simulation practical.

The importance-sampling technique is very effective and has general applicability in areas where tails of distributions have to be computed or particular sensitive areas of the statistical distribution have to be emphasized.

This method cannot, however, be used automatically; a good prior knowledge of the behavior of the system is necessary. To get an accurate statistical simulation of a section of an output, all the important regions of input PDFs which may contribute to this section have to be emphasized.

VIII. ACKNOWLEDGMENTS

The author is grateful to D. D. Sell and S. D. Personick for introducing him to problems related to optical communication and for helpful discussion on the subject, C. L. Semmelman for his comments and suggestions in the area of statistical sampling techniques, and R. B. Hawkins for guidance and encouragement.

REFERENCES

1. R. W. Lucky, J. Salz, and E. J. Weldon, Jr., *Principles of Data Communication*, New York: McGraw-Hill, 1968.
2. S. D. Personick, "New Results on Avalanche Multiplication Statistics with Applications to Optical Detection," *B.S.T.J.*, *50*, No. 1 (January 1971), pp. 167-189.
3. S. D. Personick, "Statistics of a General Class of Avalanche Detectors With Applications to Optical Communications," *B.S.T.J.*, *50*, No. 10 (December 1971), pp. 3075-3095.
4. R. J. McIntyre, "The Distribution of Gains in Uniformly Multiplying Avalanche Photodiodes: Theory," *IEEE Trans. Electron Dev.*, *ED-19* (June 1972), pp. 703-713.
5. P. P. Webb, R. J. McIntyre, and J. Conradi, "Properties of Avalanche Photodiodes," *RCA Rev.*, *35* (June 1974), pp. 234-276.
6. S. D. Personick, "Receiver Design for Digital Fiber Optic Communication Systems, I and II," *B.S.T.J.*, *52*, No. 6 (July-August 1973), pp. 843-886.
7. D. D. Sell, unpublished material.
8. S. D. Personick, P. Balaban, and J. H. Bobsin, unpublished material.
9. Special issue on "Statistical Circuit Design," *B.S.T.J.*, *50*, No. 4 (April 1971).
10. H. Kahn, "Use of Different Monte Carlo Sampling Techniques," *Symposium on Monte Carlo Methods*, edited by H. A. Mayer, New York: John Wiley, 1956, pp. 146-190.
11. R. V. Hogg and A. T. Craig, *Introduction to Mathematical Statistics*, New York: MacMillan, 1970.
12. S. B. Weinstein, "Estimation of Small Probabilities by Linearization of the Tail of a Probability Distribution Function," *IEEE Trans. Commun. Technol.*, *19*, No. 6 (December 1971), pp. 1149-1155.

A Phased, Optical, Coupler-Pair Switch

By V. RAMASWAMY and R. D. STANDLEY

(Manuscript received December 15, 1975)

We analyze an optical switch consisting of two directional couplers with an intermediate phase-compensating network. The conditional requirements on coupler length, phase mismatch between guides in each coupler, and phase compensation between couplers are established such that two switching states exist. It is shown that, to switch between states,

- (i) *The couplers must be 3 dB or stronger.*
- (ii) *Each coupler must be "adjusted" to become a hybrid by controlling the phase mismatch between guides.*
- (iii) *The differential phase ϕ must be initially set to turn on one state.*
- (iv) *An additional increment of π in the phase shift causes switching to the second state.*

It is estimated that the controlling voltage required to activate the bipolar switch is of the order of 1 to 2 V for a 3- μ m guide with 3-mm-long electrodes in Ti-diffused LiNbO₃. We are currently in the process of fabricating such devices.

I. INTRODUCTION

A number of methods¹⁻⁴ have been suggested for realizing optical switches in integrated optical format; in many cases, the schemes are extensions of the microwave art.^{5,6} Recently, Papuchon et al.⁷ and Campbell et al.⁸ obtained switching by perturbing the phase synchronism in indiffused strip-guide directional couplers of about one coupling period in length. Kogelnik et al.⁹ have proposed a switch consisting of tandem couplers with adjustable phase mismatch between guides in each coupler. This approach results in relaxed dimensional tolerances over previously described coupler switches. However, the adjustment of phase mismatch in both couplers is required to switch between states.

In this paper, we propose and analyze an optical switch consisting of directional coupler pairs with an intermediate phase-compensating network. We show that the switching between states can be achieved by a π change in the intermediate phase shift providing that each

coupler has been initially adjusted for hybrid performance, i.e., for 3-dB coupling. Prior to adjustment, the coupling strengths are required to be 3 dB or stronger, implying coupler lengths of $\frac{1}{2}$ to $1\frac{1}{2}$ basic coupling periods.

II. PRELIMINARY ANALYSIS

We assume directional couplers with uniform coupling having transfer matrices of the form¹⁰

$$\begin{bmatrix} E_3 \\ E_4 \end{bmatrix} = \begin{bmatrix} A & B \\ -B^* & A^* \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \end{bmatrix}, \quad (1)$$

where $|A|^2 + |B|^2 = 1$ (see Fig. 1). For coupler 1, the matrix coefficients relating the field amplitudes E_k are given by¹¹

$$A_1 = \cos \alpha_1 - j \frac{\gamma_1}{\sqrt{\gamma_1^2 + 1}} \sin \alpha_1, \quad (2)$$

$$B_1 = j \frac{\sin \alpha_1}{\sqrt{\gamma_1^2 + 1}}, \quad (3)$$

and, for coupler 2,

$$A_2 = \cos \alpha_2 - j \frac{\gamma_2}{\sqrt{\gamma_2^2 + 1}} \sin \alpha_2, \quad (4)$$

$$B_2 = j \frac{\sin \alpha_2}{\sqrt{\gamma_2^2 + 1}}, \quad (5)$$

where

$$\alpha_{1,2} = \sqrt{\gamma_{1,2}^2 + 1} \frac{\pi l_{1,2}}{2L_0}, \quad (6)$$

$$\gamma_{1,2} = \frac{\Delta\beta_{1,2}L_0}{\pi}, \quad (7)$$

$\Delta\beta_{1,2}$ = propagation constant difference between the guides in couplers 1 and 2, respectively,

$l_{1,2}$ = the interaction length,

L_0 = transfer length for total power transfer for $\gamma = 0$.

If we connect the two couplers in series via a differential phase-shifting network (Fig. 1), then, for unit excitation of port 1 of the first coupler, we obtain for the output signals

$$E_5 = A_1A_2 - B_1^*B_2e^{j\phi}, \quad (8)$$

$$E_6 = -A_1B_2^* - A_2^*B_1^*e^{j\phi}, \quad (9)$$

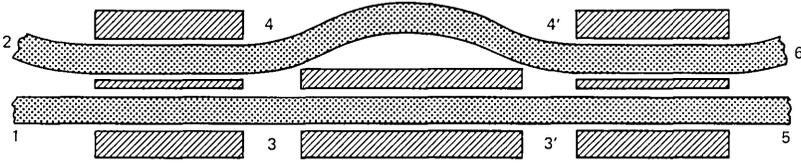


Fig. 1—Two couplers with adjustable $\Delta\beta$, connected in series with an intermediate phase shifter.

where ϕ is the differential phase. The two states of the switch are

$$\begin{aligned} \text{State I: } & E_5 = 0, \quad |E_6| = 1 \\ \text{State II: } & |E_5| = 1, \quad E_6 = 0. \end{aligned}$$

Thus, for State I, the condition is

$$e^{j\phi_1} = \frac{A_1 A_2}{B_1^* B_2}, \quad (10)$$

i.e.,

$$\left| \frac{A_1 A_2}{B_1^* B_2} \right| = 1. \quad (11)$$

This simply implies that, if the couplers have coupling coefficients that are the inverse of each other in magnitude, then a unique $\phi = \phi_1$ exists that permits the realization of State I.

For State II, the condition is

$$e^{j\phi_2} = -\frac{A_1 B_2^*}{A_2^* B_1}, \quad (12)$$

i.e.,

$$\left| \frac{A_1 B_2^*}{A_2^* B_1} \right| = 1. \quad (13)$$

This implies that, if the magnitudes of the coupling coefficients are identical, then there exists a $\phi = \phi_2$ that permits the realization of State II.

It follows, then, that with the parameters of the one coupler specified, there exist values of ϕ that permit the realization of States I or II for each allowed value of γ_2 and α_2 . To realize both states by adjustment of ϕ alone, hybrid couplers are necessary, i.e., $|A_1| = |A_2| = |B_1| = |B_2| = 1/\sqrt{2}$.

Low-loss strip waveguides made by diffusion of metals in electro-optic materials¹²⁻¹⁴ are available to facilitate the fabrication of these switches. However, the difficulty in forming coupler pairs is the maintenance of uniform guide width, separation, and refractive in-

dex differential.* Thus, realization of hybrid pairs is rather difficult. The above analysis suggests that, given couplers with a power-coupling coefficient greater than 3 dB, States I or II can be achieved by appropriately altering the phase mismatch in the couplers (electro-optically or otherwise) and introducing a specific differential phase shift ϕ between couplers.

Thus, our requirements are

- (i) Each coupler must initially have coupling coefficient $|B_1|$, $|B_2| \geq 1/\sqrt{2}$.
- (ii) Direct-current biasing can then be provided in each coupler to alter the phase mismatch in order to achieve $|B_1| = |B_2| = 1/\sqrt{2}$.
- (iii) A differential phase shift ϕ_1 is then introduced between couplers for State I to exist, and then additional phase shift of π , viz., $\phi_2 = \phi_1 + (2m + 1)\pi$, causes the switch to go to State II. Thus, activation of either state can be achieved by the addition or removal of the control voltage that causes a phase shift of π between couplers.

Section III provides a detailed analysis and establishes the initial conditions to realize the switch.

III. DETAILED ANALYSIS

In this section, we derive the equations governing the states of the switch in terms of the coupler parameters and the differential phase shift.

Substituting eqs. (2) through (5) in (10), we obtain

$$e^{j\phi_1} = \left[\frac{\gamma_1^2 + \cos^2 \alpha_1}{\sin^2 \alpha_1} \cdot \frac{\gamma_2^2 + \cos^2 \alpha_2}{\sin^2 \alpha_2} \right] e^j \left\{ \tan^{-1} \left[\frac{-\gamma_1 \tan \alpha_1}{(1 + \gamma_1^2)^{\frac{1}{2}}} \right] + \tan^{-1} \left[\frac{-\gamma_2 \tan \alpha_2}{(1 + \gamma_2^2)^{\frac{1}{2}}} \right] \right\}. \quad (14)$$

The phase shift ϕ_1 required to null port 5 is

$$\phi_1 = \tan^{-1} \left[\frac{-\gamma_1 \tan \alpha_1}{(1 + \gamma_1^2)^{\frac{1}{2}}} \right] + \tan^{-1} \left[\frac{-\gamma_2 \tan \alpha_2}{(1 + \gamma_2^2)^{\frac{1}{2}}} \right], \quad (15)$$

and the additional condition established in eq. (14) is

$$\frac{\gamma_1^2 + \cos^2 \alpha_1}{\sin^2 \alpha_1} = \frac{\sin^2 \alpha_1}{\gamma_2^2 + \cos^2 \alpha_2}. \quad (16)$$

Equation (16) relates the coupler parameters such that we have

* The effect of dimensional tolerances on individual coupler performance has been treated by several authors (Refs. 15-18). The effect of such tolerances on interference-switch performance has been considered by Shelton (Ref. 19).

couplers whose ratio of the power-coupling coefficients are the inverse of each other in *magnitude*, i.e., $|A_1|/|B_1| = |B_2|/|A_2|$. Inspection of (15) and (16) reveals that, for a given value of γ_1 , α_1 , there are many combinations of γ_2 , α_2 , and ϕ_1 that will meet the inverse condition requirement and therefore realize State I.

Equations (2) through (5) can be substituted in eq. (12) to yield the following conditions for the phase shift ϕ_2 required to null port 6:

$$\phi_2 = \pi + \tan^{-1} \left\{ \frac{-\gamma_1 \tan \alpha_1}{(1 + \gamma_1^2)^{\frac{1}{2}}} \right\} + \tan^{-1} \left\{ \frac{-\gamma_2 \tan \alpha_2}{(1 + \gamma_2^2)^{\frac{1}{2}}} \right\}. \quad (17)$$

Using eq. (15), (17) can be written as

$$\phi_2 = \phi_1 + \pi, \quad (18)$$

and the additional condition is

$$\frac{\gamma_1^2 + \cos^2 \alpha_1}{\sin^2 \alpha_1} = \frac{\gamma_2^2 + \cos^2 \alpha_2}{\sin^2 \alpha_2}. \quad (19)$$

Equation (19) illustrates the need for the ratio of the coupling coefficients to be identical, i.e., $|A_1|/|B_1| = |A_2|/|B_2|$. Again, we have many combinations γ_2 , α_2 , and ϕ_2 that will meet this requirement to realize State II for a given value of γ_1 , α_1 . Although (18) is written in terms of ϕ_1 , note that the values of γ_1 , α_1 and the associated values of γ_2 , α_2 , and ϕ_2 are distinctly different from those needed to realize State I.

Obviously, if we set $|B_1| = |B_2| = 1/\sqrt{2}$, i.e., if the couplers become hybrids, then eqs. (16) and (19) are simultaneously satisfied and both states can be simultaneously realized by simply changing ϕ by adding π . In other words, the dc biasing in each coupler is so chosen that (γ_1, α_1) and (γ_2, α_2) satisfy

$$\frac{|\sin \alpha_{1,2}|}{(\gamma_{1,2}^2 + 1)^{\frac{1}{2}}} = \frac{1}{\sqrt{2}}. \quad (20)$$

For illustrative purposes, we consider only the fundamental solution of eq. (20) for positive γ , which is plotted in Fig. 2. The plot shows the normalized length of the coupler as a function of the normalized phase mismatch γ to achieve hybrid performance. In practice, the couplers as fabricated are not usually exact hybrids. For any phase mismatch between guides, the coupler lengths have to be within $\frac{1}{2}$ to $1\frac{1}{2}$ transfer lengths (L_0). In practice, this can be easily achieved.

To illustrate the adjustment for hybrid performance, let us consider two couplers with the following initial specifications:

$$\begin{aligned} \text{Coupler } x: \quad & \gamma_x = 0.24, \quad l/L_0 = 1.2, \\ \text{Coupler } y: \quad & \gamma_y = 1.1, \quad l/L_0 = 1.1. \end{aligned}$$

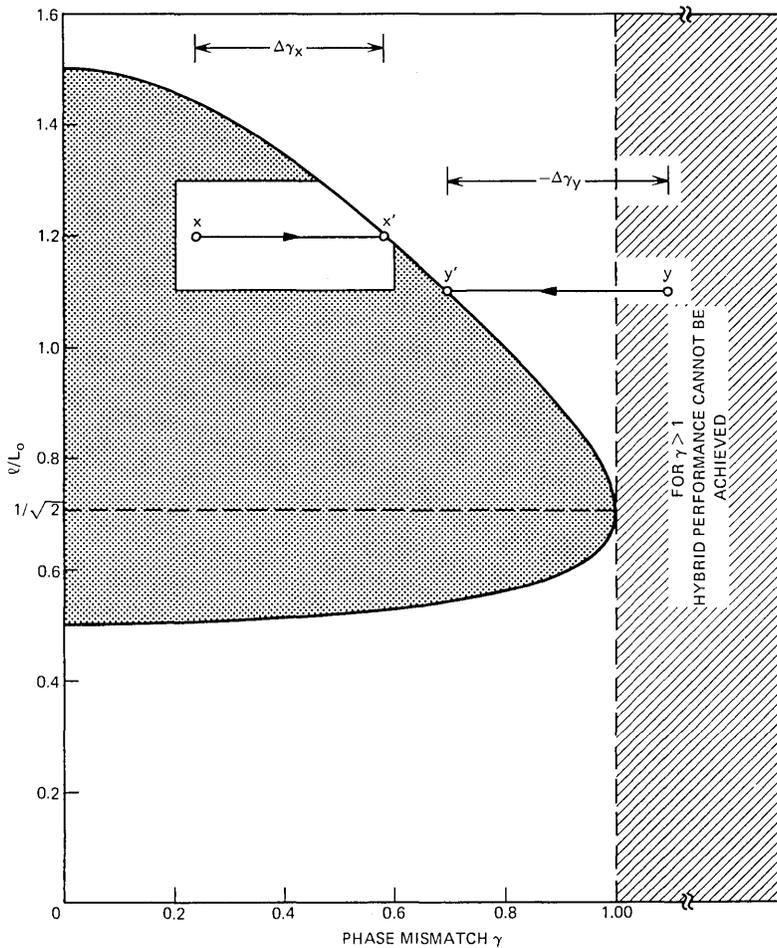


Fig. 2—Required coupler length (l/L_0) to obtain hybrid operation as a function of the phase mismatch parameter γ of the coupler.

As seen from Fig. 2, biasing the two couplers to change γ 's by $\Delta\gamma_x$ and $-\Delta\gamma_y$, respectively, results in hybrid performance, with $\gamma_{x'} = 0.58$, $\gamma_{y'} = 0.7$. Then, initializing the phase shifter to a value ϕ_1 given by eq. (15), the switch can then be operated by changing the value of ϕ_1 by adding π .

To obtain the phase change ϕ_1 , let us further assume the two couplers are identical in performance, i.e., $\gamma_1 = \gamma_2$, $\alpha_1 = \alpha_2$. (This assumption is easily satisfied, especially with fabrication processes using electron-beam lithography with analog scan or step-and-repeat techniques.)

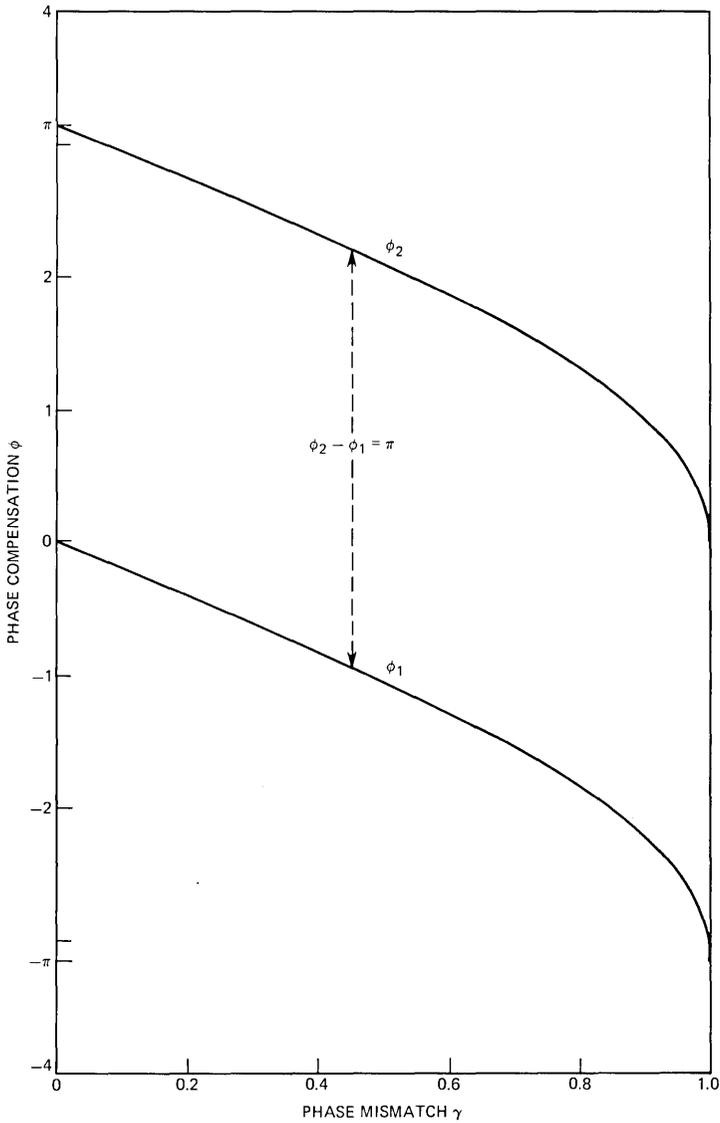


Fig. 3— ϕ_1 and ϕ_2 required to switch on State I and II, respectively, assuming identical coupler pairs and both being adjusted to provide hybrid operation.

For this case, then,

$$\phi_1 = 2 \tan^{-1} \left[\frac{-\gamma_1}{(1 - \gamma_1^2)^{\frac{1}{2}}} \right] \quad (21)$$

and

$$\phi_2 - \phi_1 = \pi. \quad (22)$$

Equations (21) and (22) are plotted in Fig. 3. Assuming Ti-diffused LiNbO₃ strip waveguides,²⁰ the control voltage required to produce a phase shift of π is estimated to be 1 to 2 V for a 3- μ m guide width and $L = 3$ mm. The special case with $\phi = 0$ is treated in the appendix.

IV. CONCLUSIONS

Equations have been established governing the states of a switching network consisting of two dephased, directional couplers interconnected via a differential phase network. The results imply that, in practice, one would strive to fabricate coupler pairs where power-coupling coefficients are greater in magnitude than 3 dB and, by appropriate bias, adjust each for 3-dB coupling. Two unique values of differential phase will then exist such that both switching states can be realized. The voltage required to switch between states is estimated to be of the order of 1 or 2 V.

APPENDIX

If we set $\phi_1 = 0$, then we have the equations governing the switch described in detail by Kogelnik et al.²¹ It can be shown that, in addition to requirements of eq. (16) (State I), we must have

$$-\frac{\gamma_2}{\gamma_1} = \frac{\sin^2 \alpha_1}{\gamma_1^2 + \cos^2 \alpha_1} \quad (23)$$

and

$$\tan \alpha_2 = 2 \frac{\sqrt{\gamma_1^2 + \cos^4 \alpha_1}}{\sin^2 \alpha_1}. \quad (24)$$

These simply imply that, given γ_1 and α_1 , γ_2 , and α_2 are uniquely fixed. This is because $\phi = 0$ is a unique state and the second coupler must satisfy both *magnitude* and *phase* conditions.

The graphical data presented by Kogelnik are quite useful in that the magnitude requirement of eqs. (16) and (19) is implied in those results.²¹

REFERENCES

1. E. A. J. Marcatili, "Dielectric Rectangular Waveguide and Directional Coupler for Integrated Optics," B.S.T.J., 48, No. 7 (September 1969), pp. 2071-2102; and E. A. J. Marcatili, "Laser Beam Circuits," November 1968, unpublished work.
2. S. Kurazono, K. Iwasaki, and N. Kumagai, "A New Optical Modulator Consisting of Coupled Optical Waveguides," Electron. Commun. Jap., 5b-C(1) (1972), p. 103.
3. H. F. Taylor, "Optical Switching and Modulation in Parallel Dielectric Waveguides," J. Appl. Phys., 44, No. 7 (1973), p. 3257.

4. F. Zernike, "Integrated Optics Switch," WA5 Topical Meeting on Integrated Optics, New Orleans, 1974.
5. W. L. Teeter and K. R. Bushore, "A Variable Ratio Microwave Power Divider and Multiplexer," IRE Trans. Microw. Theory Tech., *MTT-5* (October 1957), pp. 227-229.
6. R. M. Vaillancourt, "Analysis of the Variable-Ratio Microwave Power Divider," IRE Trans. Microw. Theory Tech., *MTT-6* (April 1958), pp. 238-239.
7. M. Papuchon et al., "Electrically Switched Optical Directional Coupler: Cobra," Appl. Phys. Lett., *27* (September 1975), p. 289.
8. J. C. Campbell, F. A. Blum, D. W. Shaw, and K. L. Lawley, "GaAs Electro-optic Directional-Coupler Switch," Appl. Phys. Lett., *27* (August 1975), p. 202.
9. H. Kogelnik and R. V. Schmidt, "Switched Directional Couplers with Alternating $\Delta\beta$," IEEE J. Quantum Electron. (July 1976).
10. D. A. Watkins, *Topics in Electromagnetic Theory*, New York: John Wiley, 1958, p. 69.
11. S. E. Miller, "Coupled Wave Theory and Waveguide Applications," B.S.T.J., *33*, No. 3 (May 1954), pp. 661-719.
12. J. M. Hammer and W. Phillips, "Low-Loss Single-Mode Optical Waveguides and Efficient High-Speed Modulators of $\text{LiNb}_2\text{Ta}_{1-z}\text{O}_3$ on LiTaO_3 ," Appl. Phys. Lett., *24* (1974), p. 545.
13. R. D. Standley and V. Ramaswamy, "Nb-Diffused LiTaO_3 Optical Waveguides: Planar and Embedded Strip Guides," Appl. Phys. Lett., *25*, (1974), p. 711.
14. R. V. Schmidt and I. P. Kaminow, "Metal-Diffused Optical Waveguides LiNbO_3 ," Appl. Phys. Lett., *25* (1974), p. 458.
15. A. Ihaya, H. Furuta, and H. Noda, "Directional Coupling Between Thin Film Optical Guides," Fujitsu Scientific & Technical Journal, *101* (June 1973).
16. H. F. Taylor, "Frequency Selective Coupling in Parallel Dielectric Waveguides," Opt. Commun., *8* (August 1973), p. 421.
17. M. G. F. Wilson and G. A. Teh, "Tapered Velocity Coupling Improves Tolerance in Optical Directional Couplers," WBI Digest of Topical Meeting on Integrated Optics, New Orleans, January 1974.
18. D. B. Ostrowsky, "Fabrication, Duplication, and Examination of Passive Integrated Optical Components," Digest of Topical Meeting on Integrated Optics, New Orleans, January 1974.
19. J. C. Shelton, unpublished work.
20. I. P. Kaminow, L. W. Stulz, and E. H. Turner, "Efficient Strip-Waveguide Modulator," Appl. Phys. Lett., *27* (November 1975), p. 555.
21. H. Kogelnik, R. V. Schmidt, and W. L. Mammel, private communication.

Mode Mixing With Reduced Losses in Parabolic-Index Fibers

By D. MARCUSE

(Manuscript received January 13, 1976)

We present design criteria for the construction of a modified parabolic-index fiber with intentional mode coupling. Mode coupling serves the purpose of reducing multimode pulse dispersion and is accomplished by introducing carefully designed index fluctuations into the fiber core or by controlled "random" bends of the fiber axis. Radiation losses due to mode coupling can be minimized by terminating the parabolic-index fiber core in an abrupt index discontinuity. The additional modes introduced by this step must be filtered out by periodic mode filters that consist of parabolic-index fiber sections without the refractive-index step.

I. INTRODUCTION

In multimode optical fibers, light pulses are carried by many modes. Such multimode operation limits the information-carrying capacity of an optical fiber communications system because of pulse spreading. At the fiber input, all modes receive part of the energy of the light pulse simultaneously. However, at the fiber output, the light pulse is stretched out in time because each mode (or mode group) travels at a different group velocity. The length of the stretched pulse is proportional to the fiber length. The amount of pulse spreading in a multimode fiber depends on its construction. The dependence of the group velocity of the modes on the mode labels is influenced by the distribution of the refractive-index profile of the fiber in radial direction.¹ Step-index fibers* typically exhibit more pulse spreading than fibers with graded refractive-index distributions. Fibers whose index profiles show a parabolic (or square law) dependence on the radial coordinate have the property that all modes have nearly the same group velocity so that pulse spreading on parabolic-index fibers is nearly minimized.¹

Our remarks so far apply to multimode transmission in perfect fibers where each mode propagates independently of all the other modes.

* Step-index fibers have a core with constant refractive index and a lower index cladding.

Inhomogeneities of the axial distribution of the refractive index or geometrical imperfections of the fiber geometry tend to couple the modes among each other. Mode coupling has the undesirable consequence that some of the power is coupled to unguided radiation modes resulting in radiation losses.² On the other hand, coupling of the guided modes results in a continuous interchange of power between fast and slow modes so that a light pulse that is distributed over all the modes assumes a new shape. Instead of spreading according to the different group velocities of the independent modes, the light power carried by coupled modes is forced to travel at an average velocity and exhibits a narrower width, which spreads only proportionally to the square root of the fiber length.³ Thus, mode coupling can be intentionally introduced to improve the pulse performance of multimode fibers.⁴ However, this technique for improving the pulse performance must be used with great care to avoid an unacceptable increase in the power loss of the fiber.

To understand how mode coupling can be tailored to minimize losses, it is necessary to consider the coupling process in more detail. Each mode has a characteristic propagation constant β_M . The label M is used to identify the mode. A unique mode designation requires that M consists actually of two symbols representing the radial and azimuthal mode numbers. For simplicity, we combine the double label in the single symbol M . As mentioned before, mode coupling is provided by some deviation of the fiber from its perfect geometry and composition.⁵ We use a function $f(z)$ to describe the axial dependence of the deviation of the refractive-index distribution or of the core-radius deviation from their nominal, perfect values. In addition to this function of the length coordinate z , we need its Fourier transform which we define as

$$F(\theta) = \lim_{L \rightarrow \infty} \left\{ \frac{1}{\sqrt{L}} \int_0^L f(z) e^{i\theta z} dz \right\}. \quad (1)$$

Coupling between two modes labeled M and N is mediated by a particular Fourier component of $f(z)$ according to the law²

$$\beta_M - \beta_N = \theta. \quad (2)$$

It is thus clear that two modes remain uncoupled if $F(\theta) = 0$ for the particular θ value required by (2). Furthermore, if the differences between the propagation constants of neighboring guided modes depend on the mode number, $F(\theta) \neq 0$ is required over a certain range of θ values if all guided modes are to be coupled. Using these rules and certain "selection rules," it has been demonstrated⁴ that most guided modes of a step-index fiber can be coupled with very little radiation

loss if the Fourier spectrum (1) is limited to a carefully selected range, such that Fourier components exist for coupling the guided modes according to (2), but that coupling between guided and radiation modes [also obeying the law (2)] is prevented.

Mode coupling in parabolic-index fibers requires only a narrow spectrum of spatial frequencies θ because the differences between propagation constants of neighboring modes are almost independent of the mode label. This feature of the modes of the parabolic-index fiber causes a problem since it makes it harder to discriminate between coupling among guided modes and coupling from guided to radiation modes. With an appropriate selection rule, coupling among guided modes of the step-index fiber has the property that the differences (2) increase with increasing mode number.⁴ Cutting the Fourier spectrum (1) off at a maximum spatial frequency $\theta = \theta_{\max}$ thus stops mode coupling at a given mode number, so that coupling between guided modes of lower order and modes with the highest mode numbers is prevented. Since only the highest-order modes are near (in mode number space) radiation modes, coupling between guided and radiation modes is thus avoided. Because of the almost constant differences between propagation constants of neighboring modes, this strategy fails in the parabolic-index fiber.

We show in this paper that we may couple the guided modes of a parabolic-index fiber and still avoid radiation losses by using a modified parabolic-index profile. An ideal parabolic-refractive-index profile has the form

$$n = n_0 \left[1 - \left(\frac{r}{a} \right)^2 \Delta \right] \quad |r| < \infty. \quad (3)$$

r is the radial coordinate and Δ/a^2 determines the gradient of the index profile. The true parabolic-index profile cannot be realized because the refractive index of ordinary solid materials cannot be less than unity. Typical parabolic-index fibers have index profiles of the form

$$n = \begin{cases} n_0 \left[1 - \left(\frac{r}{a} \right)^2 \Delta \right] & |r| \leq a \\ n_2 = n_0(1 - \Delta) & a \leq |r| \leq b \\ 1 & b \leq |r| < \infty. \end{cases} \quad (4)$$

The cladding region $a \leq |r| \leq b$ is usually so thick that, mathematically, we may assume $b \rightarrow \infty$. The guided modes do not carry significant amounts of power inside of the cladding region so that they behave almost as if they were guided by the ideal index distribution (3). Modes with significant amount of power in the cladding region are no longer guided but belong to the continuous spectrum of radia-

tion modes. By providing a narrow Fourier spectrum of spatial frequencies for the purpose of coupling the guided modes, we necessarily couple the highest-order guided modes to radiation and lose power.

The situation is changed if we modify the index profile to the following form:

$$n = \begin{cases} n_0 \left[1 - \left(\frac{r}{a} \right)^2 \Delta \right] & |r| \leq a \\ n_2 < n_0(1 - \Delta) & a < |r| \leq \infty \end{cases} \quad (5)$$

($b = \infty$ was assumed for simplicity). The index profiles of (3), (4), and (5) are shown in Figs. 1a, 1b, and 1c. We divide the guided modes of the index profile (5), shown in Fig. 1c, into two classes. There are modes whose field distributions are essentially limited to the region $0 \leq |r| \leq a$. These modes have negligibly small field intensities in the region $|r| \geq a$ and behave as though they belonged to the idealized medium defined by (3). Because modes of this kind are essentially the modes of the parabolic-index medium, we call them P-modes. There are, in addition, modes of order higher than the P-modes whose field distributions reach strongly into the region near $r = a$. These modes are guided by the index discontinuity at $r = a$ and behave similar to the modes in a step-index fiber. For this reason we call them S-modes. The differences (2) between neighboring P-modes are nearly identical, while the differences between S-modes are much larger and, if an additional selection rule is introduced, increase with increasing mode number. If a narrow band of spatial Fourier components is provided to couple the P-modes among each other, S-modes will remain uncoupled from P-modes and also remain uncoupled among each other. Thus, we have achieved coupling among the P-modes and have improved their pulse performance. However, if S-modes were allowed to reach the detector, the pulse performance of the fiber would be degraded very seriously because of the different group velocities of P-modes and S-modes and the large group velocity spread among the S-modes. It is thus necessary to suppress S-modes before they reach the detector. This can be done easily by adding at the end of our fiber with index distribution (5) (or Fig. 1c) another fiber section that does not allow S-modes to propagate. A fiber with an index profile according to (3), or Fig. 1b, has this property.

The strategy just outlined would work if there were truly no coupling between P- and S-modes. Some residual coupling is, however, unavoidable because of imperfect fiber tolerances. A small amount of power will always be coupled from P-modes to S-modes causing a "noise" background to reach the detector. This unwanted noise can be reduced by installing mode filters periodically along the fiber. As mentioned above, mode filters for S-modes consist simply of fiber sections with an index profile according to (4), or Fig. 1b. If we construct a fiber whose

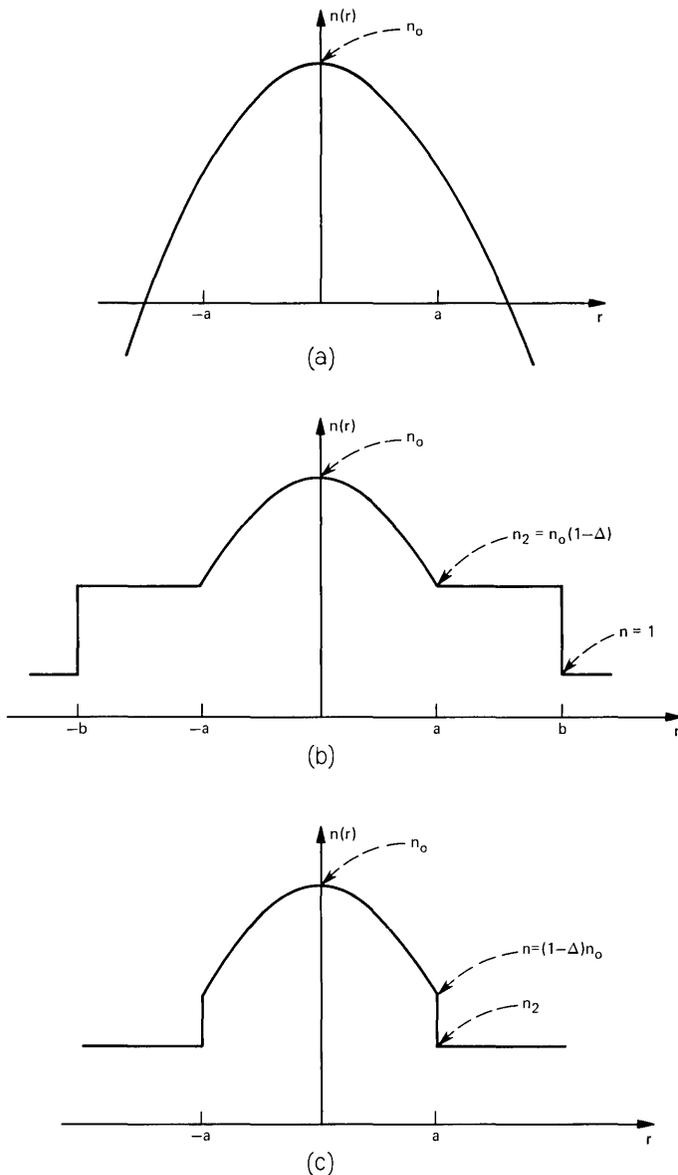


Fig. 1—(a) Ideal parabolic-index profile of eq. (3). (b) Truncated parabolic-index profile of eq. (4). This profile is used for mode filters. (c) Modified parabolic-index profile of eq. (5). This profile is used for the fiber with intentional mode coupling.

index profile is given by (5), or Fig. 1c, for most of its length, but which is changed to assume the form of (4) shown in Fig. 1b for relatively short sections periodically interspersed with the rest of the fiber, we obtain a fiber guide with mode filters for S-modes. Mode filters must

not be spaced too closely in order to avoid excessive losses. The additional loss occurs because we cannot avoid coupling between P-modes and a small group of S-modes that lie immediately adjacent (in mode number space) to P-modes. Along the boundary between P- and S-modes, mode spacing cannot be controlled so that we must assume that the intentionally introduced strong coupling mechanism will couple P-modes to their immediate S-mode neighbors along their common mode boundary. The mode filters strip away all S-modes, thus causing a small amount of loss of power that has been coupled to the S-mode group near the mode boundary. Design criteria for optimum mode filter spacings will be given in this paper.

If intentional mode coupling is achieved by introducing index fluctuations into the fiber core, the mode filters can be incorporated into the fiber by the same manufacturing process that was used to produce the intentional fiber "imperfections." On the mode filter sections, no strong coupling will be provided to avoid additional losses. We shall show later that coupling can also be provided by small bends of the fiber axis.

In the following chapters, we provide the necessary information to explain the mechanism and give design criteria for a modified parabolic-index fiber with mode coupling.

II. MODE SPECTRUM OF THE MODIFIED PARABOLIC-INDEX FIBER

We have defined P-modes and S-modes in the introduction. P-modes have field distributions that have (negligibly) small values at the core boundary $r = a$. Their properties are nearly identical with the modes of the ideal, infinitely extended square-law medium. We may associate rays with these modes. The rays corresponding to P-modes spiral around the fiber axis in helical paths. Axial rays cross the fiber axis and move out to a turning point where the tangent to their path is parallel to the fiber axis. Spiraling rays encounter two turning points, one near the fiber axis and the other one at larger radii.^{1,6} An approximate field description utilizes the fact that modes in weakly guiding fibers are nearly linearly polarized⁷ (at least suitable superpositions of exact fiber modes are nearly linearly polarized). Their dominant transverse electric field component can be approximately expressed by the WKB approximation in the form^{1,6}

$$E = A \frac{\cos\left(\psi - \frac{\pi}{4}\right) e^{-i\nu\phi} e^{-i\beta z}}{\left(\{[n(k)r]^2 - \beta^2\} r^2 - \nu^2\right)^{\frac{1}{2}}} \quad (6)$$

with

$$\psi = \int_{r_1}^r \frac{1}{r} \left(\{[n(r)k]^2 - \beta^2\} r^2 - \nu^2\right)^{\frac{1}{2}} dr. \quad (7)$$

A is an amplitude coefficient, the integer ν is the azimuthal mode number, β the propagation constant, and

$$k = \omega(\epsilon_0\mu_0)^{\frac{1}{2}}, \quad (8)$$

the propagation constant of plane waves in free space. The inner turning point r_1 is the smaller root of the equation

$$\{[n(r)k]^2 - \beta^2\}r^2 - \nu^2 = 0. \quad (9)$$

Since the denominator of the expression in (6) vanishes, the wkb approximation seems to fail at the turning points. A more careful analysis (not used here) is capable of bridging the gap and extending the validity of the wkb solution across the turning points. The requirement of continuity of the field solutions across the turning points leads to the condition^{6,8}

$$\int_{r_1}^{r_2} \frac{1}{r} (\{[n(r)k]^2 - \beta^2\}r^2 - \nu^2)^{\frac{1}{2}} dr = (2p + 1) \frac{\pi}{2}. \quad (10)$$

The upper turning point r_2 is the larger root of eq. (9), the integer p is the radial-mode number. Equation (10) defines the propagation constant β and is called the eigenvalue equation. For P-modes we have $r_2 < a$ and obtain by substitution of (5) or (3) into (10) the expression

$$\beta = \left\{ n_0^2 k^2 - 2n_0 \frac{k}{a} (2\Delta)^{\frac{1}{2}} (2p + \nu + 1) \right\}^{\frac{1}{2}}. \quad (11)$$

S-modes reach the core boundary. It is well known that an abrupt index change at the core boundary forces the electric field intensity to assume very low values at $r = a$. For all modes, with the exception of modes very close to cutoff, it is permissible to approximate the actual value of the electric field by $E = 0$ at $r = a$. This condition in conjunction with (6) and (7) leads to the following eigenvalue equation for S-modes,

$$\int_{r_1}^a \frac{1}{r} (\{[n(r)k]^2 - \beta^2\}r^2 - \nu^2)^{\frac{1}{2}} dr = (2p + \frac{3}{2}) \frac{\pi}{2}. \quad (12)$$

Substitution of (5) and integration leads to

$$\begin{aligned} & [(\kappa_0 a)^2 - 2(n_0 k a)^2 \Delta - \nu^2]^{\frac{1}{2}} \\ & - \nu \left[\arctan \left(\frac{\frac{1}{2}(\kappa_0 a)^2 - \nu^2}{\nu [(\kappa_0 a)^2 - 2(n_0 k a)^2 \Delta - \nu^2]^{\frac{1}{2}}} \right) + \frac{\pi}{2} \right] \\ & + \frac{(\kappa_0 a)^2}{2n_0 k a (2\Delta)^{\frac{1}{2}}} \left[\arcsin \left(\frac{(2n_0 k a)^2 \Delta - (\kappa_0 a)^2}{[(\kappa_0 a)^4 - (2n_0 k a \nu)^2 2\Delta]^{\frac{1}{2}}} \right) + \frac{\pi}{2} \right] \\ & = (2p + \frac{3}{2})\pi \quad (13) \end{aligned}$$

with

$$\kappa_0^2 = n_0^2 k^2 - \beta^2. \quad (14)$$

The guided and radiation modes of the modified parabolic-index fiber can be displayed in mode-number space. Each mode is characterized by two integers, the azimuthal mode number ν and the radial mode number p . Mode-number space displays the values of ν and p in the plane shown in Fig. 2. For P-modes we can introduce a compound mode number

$$M = 2p + \nu + 1 \quad (15)$$

and express the propagation constant (11) in the form,

$$\beta = \left\{ (n_0 k)^2 - 2 \frac{n_0 k}{a} (2\Delta)^{\frac{1}{2}} M \right\}^{\frac{1}{2}}. \quad (16)$$

Modes with constant values of M have the same propagation constants and lie on diagonal dotted lines, two of which are shown in Fig. 2. Cutoff for the S-modes is defined by the condition

$$\beta = n_2 k \quad (17)$$

or

$$\kappa_0 = (n_0^2 - n_2^2)^{\frac{1}{2}} k. \quad (18)$$

When we substitute (18) into (13), we obtain an implicit equation for the mode boundary between S-modes and radiation modes shown in Fig. 2 as the solid line. The boundary between P-modes and S-modes is defined by the condition in (16) written in the form

$$(\kappa_0 a)^2 = 2 n_0 k a (2\Delta)^{\frac{1}{2}} M \quad (19)$$

and from

$$(\kappa_0 a)^2 = 2(n_0 k a)^2 \Delta + \nu^2. \quad (20)$$

Equation (20) expresses the requirement that the transition from P-modes to S-modes occurs when the turning point falls on the core boundary and is obtained by combining (5), (9), (14), and the condition $r = r_2 = a$. By eliminating $\kappa_0 a$ from (19) and (20), we obtain a functional relation between ν and p which was used to calculate the dash dotted mode boundary shown in Fig. 2. This figure was computed with the help of the following parameters,

$$\left. \begin{aligned} n_0 &= 1.53 \\ n_2 &= 1.5 \\ ka &= 150 \\ \Delta &= 0.0098 \end{aligned} \right\}. \quad (21)$$

These values lead to the following value for the refractive index at the core boundary: $n_0(1 - \Delta) = 1.515$.

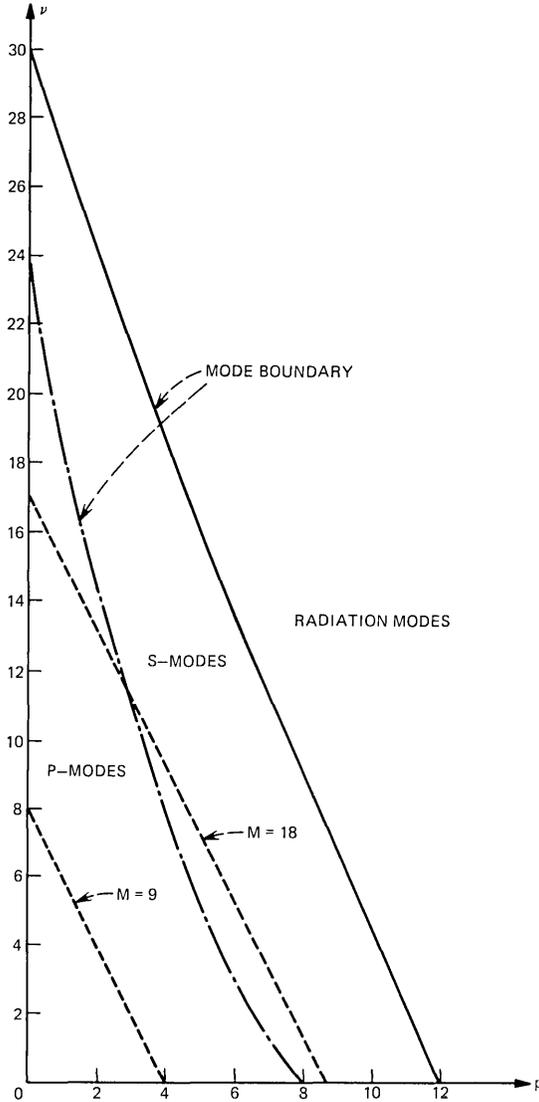


Fig. 2—Mode number space. The solid line is the boundary between S-modes and radiation modes. The dash-dotted line is the boundary between S-modes and P-modes. The dotted lines labeled $M = 9$ and $M = 18$ are curves of constant compound mode number M . $M = 18$ is used as approximate mode boundary.

For the purpose of designing a coupling mechanism for the P-modes, we need to know the differences between the propagation constants of the modes. If we let ν change by $\delta\nu$ and p by δp , we compute from

the eigenvalue equation for the P-modes in the form (19),

$$\frac{\delta\kappa_0}{\kappa_0} = \frac{\delta M}{2M} \quad (22)$$

with

$$\delta M = 2\delta p + \delta\nu. \quad (23)$$

For the S-modes, we find from (13)

$$\frac{\delta\kappa_0}{\kappa_0} = \frac{\pi(\delta p) + \frac{1}{2}(\delta\nu) \left[\arctan\left(\frac{\frac{1}{2}(\kappa_0 a)^2 - \nu^2}{\nu S}\right) + \frac{\pi}{2} \right]}{(2p + \frac{3}{2})\pi + \nu \left[\arctan\left(\frac{\frac{1}{2}(\kappa_0 a)^2 - \nu^2}{\nu S}\right) + \frac{\pi}{2} \right] - S} \quad (24)$$

with the abbreviation

$$S = [(\kappa_0 a)^2 - 2(n_0 k a)^2 \Delta - \nu^2]^{\frac{1}{2}}. \quad (25)$$

The difference of the propagation constants follows from (14):

$$\delta\beta = -\frac{\kappa_0^2}{\beta} \frac{\delta\kappa_0}{\kappa_0}. \quad (26)$$

We have plotted values for $|\delta\beta a|$ in Fig. 3. These values were computed from (22) and (26) for P-modes and from (13), (24), and (26) for S-modes using the numbers in (21). We also assumed that $\delta M = \pm 1$. For P-modes, $\delta M = \pm 1$ clearly yields the least separation between guided modes. ($\delta M = 0$ would lead to $\delta\beta = 0$ and is excluded.) Just as in the case of the step-index fiber, we introduce a selection rule by properly designing the coupling mechanism (see Section III below). The selection rule is

$$\delta\nu = \pm 1. \quad (27)$$

We may now achieve a transition between neighboring modes with $|\delta M| = 1$ in two different ways. We can either use

$$\delta\nu = \pm 1, \quad \delta p = 0 \quad (28)$$

or

$$\delta\nu = \pm 1, \quad \delta p = \mp 1, \quad (29)$$

where the upper or lower set of signs belong together respectively. All other combinations lead to bigger values of $|\delta M|$ and larger differences for the separation between the S-modes.

Figure 3 shows that the values of $|\delta\beta a|$ for P-modes are very nearly independent of the mode number M . By using a Fourier spectrum (1) with spatial frequencies in the range $\theta a = 0.14$ to 0.1415 , we couple all neighboring P-modes with $|\delta M| = \pm 1$. The hatched area in Fig. 3 labeled "S-modes" indicates the range of $|\delta\beta a|$ values that occurs for

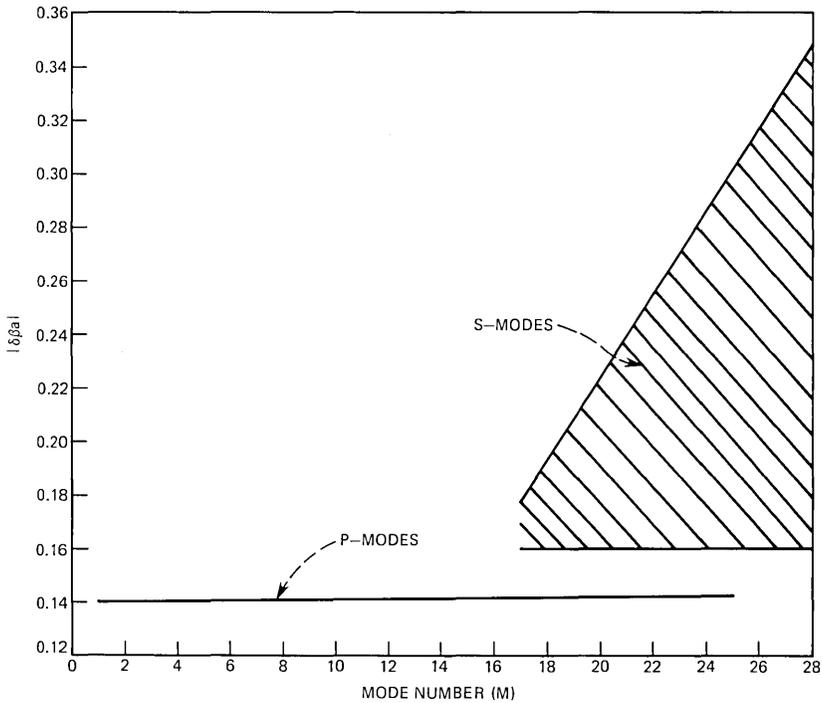


Fig. 3—The differences of propagation constants of adjacent modes with $|\delta M| = 1$ as functions of the compound mode number M .

a given value of M . It is apparent how widely the differences between S-modes with $|\delta M| = \pm 1$ vary. Typically, the combination (28) leads to smaller $\delta\beta$ differences. However, even the smallest spacing between adjacent S-modes is so much larger than the corresponding spacing between P-modes that it should be relatively easy to design the coupling mechanism so that S-modes are not intentionally coupled among each other. The differences between adjacent P-modes and S-modes along the dash-dotted mode boundary in Fig. 2 cannot be calculated from our simplified theory, so that we must assume that P-modes, lying along the mode boundary, may be coupled to their S-mode neighbors.

We have now proved that P-modes can be coupled among each other by a spatial Fourier spectrum of very narrow width and that it is relatively easy to discriminate between P-mode and S-mode coupling. It remains to discuss mode coupling, study the degradation of pulse performance that results from coupling a few S-modes along the mode boundary to P-modes, and to discuss the performance of the mode filters.

III. MODE COUPLING AND PULSE SPREADING

Mode coupling is provided by implanting a refractive-index perturbation into the fiber core.² Instead of the index distribution n given in (5), we now use a perturbed refractive-index distribution \mathbf{n} so that we have

$$\mathbf{n}^2 - n^2 = \frac{r}{a} f(z) \cos \phi \quad |r| \leq a. \quad (30)$$

The dependence of (30) on $\cos \phi$ imposes the selection rule (27).² Without such a selection rule we could not uncouple P-modes and S-modes. The linear r dependence indicated in (30) is quite arbitrary. It is necessary that (30) vanishes at $r = 0$ in order to have a well defined value of the function at this point. The linear r dependence is not only the simplest function that vanishes at $r = 0$ but also one for which the coupling coefficient can be evaluated. Other r dependent functions could be used as factors in (30). The function $f(z)$ is assumed to be a random function with the narrow Fourier spectrum (1) that assures coupling among the P-modes but prevents coupling between P-modes and S-modes.

The system of coupled P-modes can be described by coupled-power equations.² The coupling coefficients are obtained by solving overlap integrals including the function (30) multiplied with the electric field functions of the two modes whose coupling is to be evaluated. The P-modes are described very well by the Laguerre-Gaussian field solutions of the ideal parabolic-index distribution (3).^{9,10,11} We obtain for the power-coupling coefficients²:

$$h_{\nu p, \nu-1, p} = K(p + \nu) \quad (31)$$

$$h_{\nu p, \nu-1, p+1} = K(p + 1) \quad (32)$$

$$h_{\nu p, \nu+1, p} = K(p + \nu + 1) \quad (33)$$

$$h_{\nu p, \nu+1, p-1} = Kp \quad (34)$$

with

$$K = \frac{k}{16n_0^3 a (2\Delta)^{\frac{1}{2}}} \langle F^2(\theta) \rangle. \quad (35)$$

The symbol $\langle \rangle$ indicates that an ensemble average has been taken. The argument of the Fourier spectrum function (1) is the appropriate difference of the propagation constants of the coupled modes according to (2). [See eq. (75).]

It is our intention to solve the coupled-power equations for the P-modes with the coupling coefficients given by (31) through (35). The coupling mechanism postulated by the refractive-index perturbation (30) couples a given mode (ν, p) to its neighbors $(\nu + 1, p)$,

$(\nu - 1, p + 1)$, $(\nu - 1, p)$, and $(\nu + 1, p - 1)$. In the first two cases, the M number is increased by one while it is decreased by one in the third and fourth case. The coupled-power equations thus assume the form²

$$\begin{aligned} \frac{\partial P_{\nu,p}}{\partial z} + \frac{1}{v_M} \frac{\partial P_{\nu p}}{\partial t} + \alpha P_{\nu p} = & h_{\nu p, \nu+1, p} (P_{\nu+1, p} - P_{\nu p}) \\ & + h_{\nu p, \nu-1, p+1} (P_{\nu-1, p+1} - P_{\nu p}) + h_{\nu p, \nu-1, p} (P_{\nu-1, p} - P_{\nu p}) \\ & + h_{\nu p, \nu+1, p-1} (P_{\nu+1, p-1} - P_{\nu p}). \end{aligned} \quad (36)$$

We have indicated by our notation that the group velocity v_M depends only on the compound mode number M defined by (15) and not on the individual values of ν and p . In addition, all modes are assumed to suffer the same loss α caused by absorption and random scattering processes.

To be able to solve equation system (36), we assume that modes with equal values of the compound mode number M carry equal amounts of power. It is not obvious that this must be the case, but we may argue that modes with equal mode number M have identical propagation constants and, thus, are coupled among each other by the zero-spatial-frequency component of $F(\theta)$. The zero-frequency components of random distortions tend to be very large. Consider, for example, the departures from perfect straightness of the fiber axis. No fiber in actual use is ever perfectly straight. In fact, its slow variations tend to be particularly large so that its power spectrum peaks at zero spatial frequencies. This fact has been observed whenever power spectra of fiber distortions have been measured.^{12,13} Consequently, we assume that modes with equal values of M are strongly coupled and, hence, equally populated by a coupling process that is not explicitly taken into account. It is only incorporated into the analysis by assuming that modes with the same values of M carry equal amounts of power. Making this assumption and substituting (31) through (35) into (36) allows us to write,

$$\left. \begin{aligned} P_{\nu p} &= P_M \\ P_{\nu+1, p} &= P_{\nu-1, p+1} = P_{M+1} \\ P_{\nu-1, p} &= P_{\nu+1, p-1} = P_{M-1} \end{aligned} \right\} \quad (37)$$

and

$$\begin{aligned} \frac{\partial P_M}{\partial z} + \frac{1}{v_M} \frac{\partial P_M}{\partial t} + \alpha P_M = & K[(M + 1)(P_{M+1} - P_M) \\ & - M(P_M - P_{M-1})]. \end{aligned} \quad (38)$$

If the number of guided modes is very large, we may regard them as a quasi-continuum and treat M approximately as a continuous vari-

able.¹⁴ This assumption allows us to write the equation system (38) as a partial differential equation,²

$$\frac{\partial P_M}{\partial z} + \frac{1}{v_M} \frac{\partial P_M}{\partial t} = -\alpha P_M + K \frac{\partial}{\partial M} \left[M \frac{\partial P_M}{\partial M} \right]. \quad (39)$$

Using (16), we can approximate the inverse group velocity in the following way

$$\frac{1}{v_M} = \frac{1}{c} \frac{\partial \beta}{\partial k} = \frac{n_0}{c} + WM^2 \quad (40)$$

with the abbreviation

$$W = \frac{\Delta}{n_0 c (ka)^2}. \quad (41)$$

c is the velocity of light in vacuum. We have indicated in Section I that a group of S-modes along the mode boundary between P- and S-modes are coupled to the P-modes. These modes have different group velocities. We incorporate this mode group into our theory by assuming that the modes with the highest value of M , $M = N$, have a group velocity that differs from the law given by (40) for P-modes. Thus, we use

$$\frac{1}{v_M} = \frac{n_0}{c} + W[M^2 + \rho N^2 \delta(M - N)]. \quad (42)$$

The parameter ρ is the relative amount by which the inverse velocity difference $1/v_M - n_0/c$ of the mode group with $M = N$ differs from the normal value. If the S-modes followed the law for the P-mode group velocity, we would have for the mode group with $M = N$

$$\frac{1}{v_N} = \frac{n_0}{c} + WN^2. \quad (43)$$

Equation (42) states instead that the highest-mode group has inverse group velocity

$$\frac{1}{v_N} = \frac{n_0}{c} + (1 + \rho)WN^2. \quad (44)$$

We solve the time-dependent coupled-power equations in the usual way.² First, the time-independent problem is solved. The only difference from the normal procedure is that, in our present case, we do not assume that the highest-mode group is depleted by a loss process. Thus, instead of requiring $P_N = 0$, we assume that power outflow stops at $M = N$ and require correspondingly

$$\frac{\partial P_M}{\partial M} = 0 \quad \text{at} \quad M = N. \quad (45)$$

The time-independent problem, thus, has the solution

$$P_M = e^{-\alpha z} \sum_{i=0}^{\infty} c_i \frac{J_0(u_i \sqrt{M/N})}{\sqrt{N} J_0(u_i)} e^{-\sigma_i z}. \quad (46)$$

The parameters u_i are the roots of the Bessel function $J_1(x) = 0$ with u_0 defined as zero; thus,

$$u_i = 0, 3.832, 7.016, 10.123, \dots \quad (47)$$

The eigenvalues are

$$\sigma_i = K \frac{u_i^2}{N} \quad (48)$$

with K of (35). The coefficients c_i determine the initial power distribution at $z = 0$. For large values of z , P_m reaches its steady-state solution

$$P_M = \frac{c_0}{\sqrt{N}} e^{-\alpha z}, \quad (49)$$

which indicates that all modes carry equal amounts of power.

The time-dependent, steady-state solution is known to represent a pulse with gaussian shape.^{2,3} Its full width (in time) between $1/e$ points is given by second-order perturbation theory² and can be calculated to assume the following form:

$$T = T_0 \left(1 + \frac{5\rho}{N} \right). \quad (50)$$

For $\rho = 0$, we obtain the pulse width of P-modes that are completely independent of the S-modes,

$$T_0 = \sqrt{L} \frac{2^{1/4} n_0^{1/2} \Delta^{5/4} N^{5/2}}{kc(ka)^{3/2} [\langle F^2(\theta) \rangle]^{1/2}}. \quad (51)$$

L is the length of the fiber that has been traversed by the pulse. The pulse width spreads only proportional to the square root of the distance traveled. Equation (50) shows that the group of S-modes that is coupled to the P-modes widens the pulse more if its group velocity departs more from that of the P-modes—that is, with increasing values of ρ . Its influence is reduced with increasing values of the maximum mode number N .

A comment needs to be made regarding our analysis. We have calculated the impulse-response width (50) by treating the mode boundary between P- and S-modes as being parallel to the lines $M = \text{const}$. Figure 2 shows that this is not strictly true. In addition, it is not true that all S-modes along the mode boundary have equal group velocities. Our result must thus be regarded as an estimate and

we must use an average group velocity for the S-modes. For $M = N$, we use the value that is obtained by setting $\nu = N/1.5$ in (19) and (20) and obtain,

$$N = [0.573n_0ka\sqrt{2\Delta}]_{\text{int}}. \quad (52)$$

The subscript "int" attached to the bracket indicates that the nearest integer to the number in brackets must be taken. The approximate mode boundary (52) is shown in Fig. 2 as the dotted line labeled $M = 18$.

IV. THE INFLUENCE OF RESIDUAL COUPLING BETWEEN P- AND S-MODES

If the intentionally introduced coupling were the only mechanism by which the guided modes interact, we would have no further problem. However, residual coupling between all the modes is unavoidable so that we must consider the problem of power coupling from P-modes to S-modes and vice versa.

If the P-modes carry light pulses, some of their power will couple to S-modes. Since S-modes travel with different group velocities, the power they receive from P-modes spreads out and forms an almost continuous background of noise power. Some of this power is coupled back into P-modes so that this noise background reaches the detector even if we filter out the S-modes before they reach the fiber end. Periodically spaced mode filters will reduce this noise problem, but filters increase the overall losses of the system so that it is necessary to reach a compromise between excess loss due to mode filters and undesirable noise caused by unintentional S-mode coupling.

We treat the noise problem in two stages. First, we consider the power-coupling process from P-modes to S-modes, and in the next step we allow this power to couple back to the P-modes.

Consider the equation system²

$$\frac{\partial Q_n}{\partial z} + \frac{1}{w_n} \frac{\partial Q_n}{\partial t} = \sum_{\mu} H_{n\mu} P_{\mu}. \quad (53)$$

These equations describe only coupling from P-modes to S-modes. Coupling in the reverse direction is ignored as are losses. For weak coupling in short, low-loss fiber sections, this equation system is a reasonable approximation to the complete coupled power equation. $H_{n\mu}$ represents the residual, undesired coupling mechanism; the summation extends over all P-modes and the indexing system is simplified by using one symbol for the complete set of mode labels. By using a double Fourier integral transformation, it is possible to derive the following solution of (53),

$$Q_n = \frac{1}{2} \left(\sum_{\mu} H_{n\mu} E_{\mu} \right) \frac{w_n v}{v - w_n} \left[\operatorname{erf} \left(\frac{t - z/v}{T/2} \right) - \operatorname{erf} \left(\frac{t - z/w_n}{T/2} \right) \right], \quad (54)$$

which is based on the assumptions that the pulses carried by the P-modes travel with a uniform velocity v (these pulses are strongly coupled by the intentional coupling mechanism) and that their shape is gaussian,^{2,3}

$$P_\mu = A_\mu \exp \left[- \frac{(t - z/v)^2}{(T/2)^2} \right] \quad (55)$$

with energy content

$$E_\mu = \int_{-\infty}^{\infty} P_\mu dt, \quad (56)$$

and that the S-modes are slower than the P-modes,

$$w_n < v \quad \text{for all } n. \quad (57)$$

The shape of the pulse carried by one S-mode is shown in Fig. 4 as the solid line. The shape of the P-mode pulses is indicated by the dotted line; the relative height of the two pulses is of no significance. The width of the S-pulse is

$$\delta z = t(v - w_n) = (v - w_n) \frac{z}{w_n}. \quad (58)$$

Because the P-pulses are spaced as closely as possible to maximize the information rate, we may safely assume that the S-pulses overlap after a distance z that is short compared to the total length of the fiber. The number of S-pulses that overlap at any given point is, on average,

$$M_s = \frac{\delta z}{D} = \frac{v - w_n}{w_n} \frac{z}{D}, \quad (59)$$

where D indicates the spatial separation between the original P-pulses.

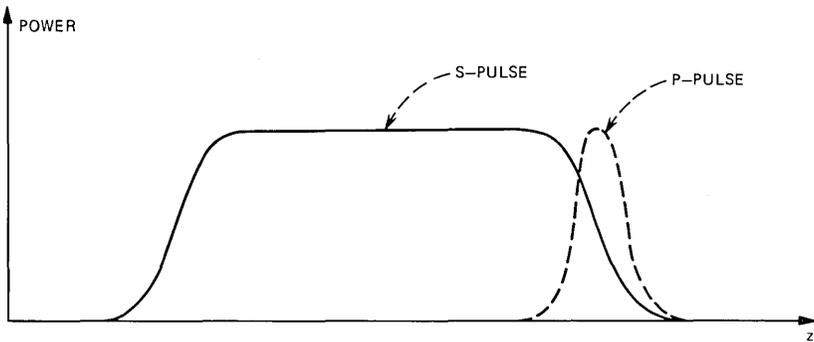


Fig. 4—The solid line is the shape of the S-pulse that results from coupling of P-mode power whose shape is indicated by the dotted line. The relative height of the two pulses is of no significance.

The average power carried by a given S-mode is, thus,

$$Q_n = M_s Q_n = v \frac{z}{D} \sum_{\mu} H_{n\mu} E_{\mu}. \quad (60)$$

[The difference of the two error functions in the bracket in (54) assumes the value 2 over most of the region where its value does not vanish. Equation (60) can also be obtained directly from (53).]

We have now determined the average power in a given S-mode due to unintentional power coupling from the P-modes. We have also convinced ourselves that the S-modes do not carry pulses, because of extensive pulse overlap, but carry a quasi-continuous background noise signal.

The amount of noise power that is coupled back from S-modes into P-modes is obtained from the time-independent coupled-power equations²

$$\frac{\partial \bar{P}_{\nu}}{\partial z} = \sum_n H_{\nu n} Q_n. \quad (61)$$

This equation is solved by simple integration. Summing over all modes and substituting (60) yields the following expression for the total noise power that is carried by all the P-modes,

$$N_p = \sum_{\nu} \bar{P}_{\nu} = v \frac{z^2}{2D} \sum_{\nu} \sum_n \sum_{\mu} H_{\nu n} H_{n\mu} E_{\mu}. \quad (62)$$

Greek summation symbols are used to indicate summation over P-modes while Latin symbols express summation over S-modes.

We assume that the fiber length between mode filters is $z = d$. At the end of the fiber of length L , we collect the noise contribution from L/d fiber sections between mode filters. The total noise is, thus,

$$N_t = \frac{vLd}{2D} \sum_{\nu} \sum_{\mu} \sum_n H_{\nu n} H_{n\mu} E_{\mu}. \quad (63)$$

So far we have ignored absorption and scattering losses. If all modes suffer identical losses, we only need to multiply (63) by the total loss that a signal suffers in traveling through the fiber. However, since our objective is to derive an expression for the signal-to-noise ratio, the absorption loss drops out in the end because signal and noise suffer identical losses. The average signal power at the end of the fiber is

$$S = \frac{1}{T} \sum_{\mu} E_{\mu}. \quad (64)$$

The signal-to-noise ratio at the fiber end is, therefore,

$$\frac{S}{N_t} = \frac{2D}{vTLd} \frac{\sum_{\mu} E_{\mu}}{\sum_{\nu, \mu} \sum_n H_{\nu n} H_{n\mu} E_{\mu}}. \quad (65)$$

If the unintentional coupling coefficients H_{nm} were known, we could calculate the signal-to-noise ratio from (65). Equation (65) tells us that the signal-to-noise ratio improves as the spacing D between pulses is increased or the pulse width T is decreased. The signal-to-noise ratio deteriorates with increasing unintentional coupling strength, in fact this decrease is inversely proportional to the square of the coupling strength. The signal-to-noise ratio decreases linearly with total fiber length L and with increasing fiber length d between mode filters. It is thus advantageous to space the mode filters as closely as the additional loss that filters introduce will allow.

For practical purposes, it will be necessary to obtain information about the unintentional mode coupling by making signal-to-noise measurements on a representative fiber sample. The signal-to-noise ratio of a fiber sample of length L_c follows from (65) if we set $L = d = L_c$. We may use pulses of length T_c for this measurement (T_c or T are measured at the end of each fiber) and use a pulse spacing D_c for the calibration measurement. This allows us to express the signal-to-noise ratio of a fiber with mode filters in terms of the measured signal-to-noise ratio $(s/n)_c$ of a representative fiber sample,

$$\frac{S}{N_t} = \frac{DT_c L_c^2}{D_c T L d} (s/n)_c. \quad (66)$$

V. LOSSES CAUSED BY MODE FILTERS

We explained in the introduction that mode filters are required to suppress the buildup of power in the S-modes. A mode filter is simply a section of fiber—without the mechanism for intentional mode coupling for P-modes—whose refractive index profile is modified from the shape of Fig. 1c (used for most of the fiber) to the shape of Fig. 1b.

Mode filters introduce additional losses because the strong coupling mechanism provided for the P-modes couples these modes to a group of S-modes immediately adjacent to the common mode boundary. Let us first consider how many modes there are for a given compound mode number. If we count the number of combinations of ν and p that lead to a fixed value of M defined by (15), we find the following expression,

$$N_M = \left[\frac{M-1}{2} \right]_{\text{int}} + 1. \quad (67)$$

The subscript “int” indicates in this case that the integer smaller than the number in brackets must be taken. For simplicity we use the approximation

$$N_M \approx \frac{M}{2}. \quad (68)$$

As a further approximation, we replace the mode boundary by the straight line labeled $M = 18$ in Fig. 2. The total number of P-modes is now approximately

$$N_{T_p} = \sum_{M=1}^N \frac{M}{2} = \frac{1}{4}N(N+1). \quad (69)$$

Because of the strong intentional coupling, each P-mode and each mode in a group of N_N S-modes along the common mode boundary carries the same amount of energy. The energy reaching the σ th fiber section is E_{T_σ} . The total energy loss on the σ th fiber section just after the mode filter is, thus,

$$E_{T,\sigma+1} - E_{T_\sigma} = - \left[\frac{N_N}{N_{T_p} + N_N} e^{-2\alpha d} + (1 - e^{-2\alpha d}) \right] E_{T_\sigma}. \quad (70)$$

The symbol α indicates the losses caused by absorption and random scattering. We solve the difference equation and use the relation

$$\sigma d = z \quad (71)$$

to obtain with the help of (68) (with $M = N$) and (69),

$$E_{T_\sigma} = E_{T_0} \left(\frac{N+1}{N+3} \right)^\sigma e^{-2\alpha z}. \quad (72)$$

At the end of the fiber, we have a total of L/d mode filter sections. The excess loss caused by the mode filters is, thus,

$$\frac{E_{TL}}{E_{T_0}} e^{2\alpha L} = \left(\frac{N+1}{N+3} \right)^{L/d} = e^{-2\alpha_f L}. \quad (73)$$

This allows us to define the approximate power-loss coefficient per unit length for the excess filter-loss penalty,

$$2\alpha_f = \frac{2}{Nd} = \frac{3.5}{n_0 k a \sqrt{2\Delta} d}. \quad (74)$$

Equation (52) was used to obtain the right-hand side of this expression. Short fiber sections of length d between mode filters (in other words, more mode filters) thus increase the excess loss.

VI. DISCUSSION AND NUMERICAL EXAMPLES

In the introduction, we have outlined our strategy for reducing pulse spreading in parabolic-index fibers. We are now in a position to offer specific design criteria for our approach. The most important aspect of the fiber design consists in incorporating index fluctuations into the fiber core that obey the relation (30). As pointed out earlier, the r dependence of the index perturbations is not particularly important.

The linear dependence was chosen for convenience. The function $f(z)$ is a nearly periodic, random function that must contain spatial frequencies in the range

$$\left. \begin{aligned} \theta &= \frac{2\pi}{\Lambda} = \frac{\sqrt{2\Delta}}{a} \rightarrow \frac{\sqrt{2\Delta}}{a} + \delta\theta \\ \text{with } \delta\theta &= 1.15 \theta\Delta \end{aligned} \right\}. \quad (75)$$

[This formula follows from (2) and (26) with the help of (16), (18), (22), with $\delta M = 1$, and (52)]. The parameter Δ is defined by (3). For the numerical example given by (21), we find that θa (a is the fiber-core radius) covers the range from 0.14 to 0.1415. To gain insight into the meaning of these numbers, we assume that the vacuum wavelength of the light transmitted through our fiber is $\lambda = 1 \mu\text{m}$. This means that $ka = 150$ leads to a core radius of $23.87 \mu\text{m}$. The spatial period length of the almost periodic random function must thus be $\Lambda = 1.07 \text{ mm}$.

Next, we must decide what amount of index fluctuation is required to achieve a desired reduction in the width of the uncoupled pulse (pulse carried by uncoupled modes). In the absence of coupling, the pulse width is given by

$$T_{uc} = \frac{n_0 L}{2c} \Delta^2. \quad (76)$$

The pulse width in the presence of mode coupling must be smaller than this value, otherwise steady state has not been reached and (50) and (51) are not applicable. First, let us consider how much harm results from coupling between P-modes and the unwanted group of S-modes along the common mode boundary. Using the values in (21), we find from (52) that $N = 18$. Let us (arbitrarily) assume that $\rho = 1$; we then find from (50) that the pulse width of the coupled modes is 28 percent larger than the width of coupled P-modes in the absence of S-mode coupling. Ignoring the slight pulse-broadening effect of residual S-mode coupling, we define an improvement factor R as

$$R = \frac{T_0}{T_{uc}}. \quad (77)$$

It is desirable to make R as small as possible and only values with $R < 1$ are meaningful. From (51), (52), and (76) we get

$$R = \frac{1.41n_0^2\Delta^{\frac{1}{2}}}{\left(\frac{L}{a^2}\langle F^2(\theta)\rangle\right)^{\frac{1}{2}}} = \frac{0.33}{\left(\frac{L}{a^2}\langle F^2(\theta)\rangle\right)^{\frac{1}{2}}}. \quad (78)$$

The numerical value on the right-hand side was obtained with the

help of (21). To achieve, for example, $R = 0.1$ would require

$$\left[\frac{L}{a^2} \langle F^2(\theta) \rangle \right]^{\frac{1}{2}} = 3.3.$$

To understand what this result means in terms of the amplitude of the refractive-index fluctuations, let us consider that the function $f(z)$ appearing in (30) has the form

$$f(z) = A \sin [\Omega z + \psi(z)], \quad (79)$$

where $\psi(z)$ is a random phase function with correlation length D_{corr} . The power spectrum of this function defined by (1) is⁴

$$\langle F^2(\theta) \rangle = A^2 \frac{\sin^2 \left[(\theta - \Omega) \frac{D_{\text{corr}}}{2} \right]}{(\theta - \Omega)^2 D_{\text{corr}}}. \quad (80)$$

Instead of the correlation length D_{corr} , we may introduce the width $\delta\theta$ of the spectral band by the relation

$$\delta\theta = \frac{4\pi}{D_{\text{corr}}}. \quad (81)$$

With $\Omega = \theta$, we thus obtain from (80) and (81)

$$\left[\frac{L}{a^2} \langle F^2(\theta) \rangle \right]^{\frac{1}{2}} = A \left(\frac{\pi}{a\delta\theta} \frac{L}{a} \right)^{\frac{1}{2}} = 3.3. \quad (82)$$

(The number on the right-hand side pertains to our example.) This important relation shows how the amplitude A of the refractive-index fluctuations is related to the spectral bandwidth $\delta\theta$, the core radius, and the length of fiber over which a certain improvement factor R (in our case $R = 0.1$) is to be achieved. For our example, θa ranges from 0.14 to 0.1415 so that we have $a\delta\theta = 0.0015$. Assuming $L = 1$ km, and using $a = 23.87 \mu\text{m}$, we have from (82)

$$A = 1.1 \times 10^{-5}. \quad (83)$$

The refractive-index fluctuations follow from (30) and (79) if we assume that $\mathbf{n} - n \ll 1$ [n is the perfect index distribution (5)]:

$$\mathbf{n} - n = \frac{1}{2n_0} \frac{r}{a} A \sin [\theta z + \psi(z)] \cos \phi. \quad (84)$$

It is apparent from (83) that very slight index fluctuations are very effective for intentional mode coupling. Much more substantial pulse-length shortening than $R = 0.1$ should, thus, be easily achievable. The random phase fluctuations with correlation length D_{corr} can be produced by keeping the phase $\psi(z)$ of (79) constant over a fixed

distance and introducing a random phase jump periodically in length intervals D_{corr} . The relation between D_{corr} and the desired bandwidth of the random function is given by (81) and (75).

Not much has been said about the problem of actually implementing the refractive-index fluctuations prescribed by (30) or (84). In principle, it is possible to introduce the refractive-index fluctuation during the process of preform fabrication, since it is necessary to take special care to produce the parabolic-index profile in any case. It may thus be possible to produce the intentional deviations from the perfect parabolic-index profile in the core by programming the process of chemical vapor deposition, or whatever process is used to control the refractive-index profile. However, there is a much simpler way of realizing a refractive-index fluctuation of the kind required by (30). Let us write the index profile inside the fiber core in cartesian coordinates:

$$n = n_0 \left[1 - \frac{x^2 + y^2}{a^2} \Delta \right]. \quad (85)$$

If we displace the index profile in x direction from its symmetric position, we may make the substitution

$$x \rightarrow x - g.$$

Assuming the displacement g to be small, we obtain instead of (85)

$$\mathbf{n} = n + 2n_0 \frac{xg}{a^2} \Delta \quad (86)$$

with n once more given by (85). By transforming the cartesian coordinates to cylindrical coordinates using

$$x = r \cos \phi, \quad (87)$$

we obtain

$$\mathbf{n}^2 - n^2 \approx 2n_0(\mathbf{n} - n) = \frac{r}{a} \left(4n_0^2g \frac{\Delta}{a} \right) \cos \phi. \quad (88)$$

Comparison of (30) and (88) allows us to make the following association:

$$f(z) = 4n_0^2g \frac{\Delta}{a}. \quad (89)$$

The actual displacement of the fiber axis from perfect straightness can be expressed as

$$g = B \sin [\Omega z + \psi(z)], \quad (90)$$

with B being the displacement amplitude. Comparison of (79), (89),

and (90) allows us to use the relation

$$B = \frac{aA}{4n_0^2\Delta}. \quad (91)$$

For our example values (21), we thus obtain from (83)

$$B = 2.9 \times 10^{-3} \mu\text{m}. \quad (92)$$

This discussion shows that we may introduce the desired perturbation by actually bending the fiber axis in the form indicated by (90) with the very small amplitude B . Such bending could be accomplished by surrounding the fiber with a suitably strained plastic jacket.

The abrupt refractive-index discontinuity of magnitude $n_0(1 - \Delta) - n_2$ is necessary to insulate the P-modes from the radiation modes by creating a buffer region of S-modes. The amount of this index step is actually quite arbitrary. However, it is clear from the general formula (65) for the signal-to-noise ratio that the sum over n in the denominator of this expression is larger when it consists of more terms, that is, if there are more S-modes. To keep the signal-to-noise ratio as small as possible requires keeping the number of S-modes small, which is achieved by keeping the abrupt index discontinuity small. What is the minimum index step that is needed for effective isolation of the P-modes? We can estimate a minimum index step by the following consideration. The lowest possible value of the propagation constant of P-modes is given by

$$\beta_m = n_0(1 - \Delta)k. \quad (93)$$

The largest propagation constant of the radiation modes (actually, what we call radiation modes become cladding modes in a fiber with finite cladding thickness) is given by

$$\beta_r = n_2k. \quad (94)$$

Adjacent groups of P-modes with the same value M of the compound mode number are spaced a distance θ (in β -space) apart. A sufficiently wide buffer zone of S-modes is required to isolate the P-modes from radiation modes. Thus, we require that the P-mode group with propagation constant (93) is separated (in β -space) by 3θ from the radiation modes. This requirement leads us to the desired condition for the minimum height of the index step

$$n_0(1 - \Delta) - n_2 = \frac{3\theta}{k} = 3 \frac{\sqrt{2\Delta}}{ka}. \quad (95)$$

For the values listed in (21), we find $n_0(1 - \Delta) - n_2 = 0.003$.

It remains to consider the design of the mode filters. Mode filters consist of fiber sections with the refractive index distribution of Fig.

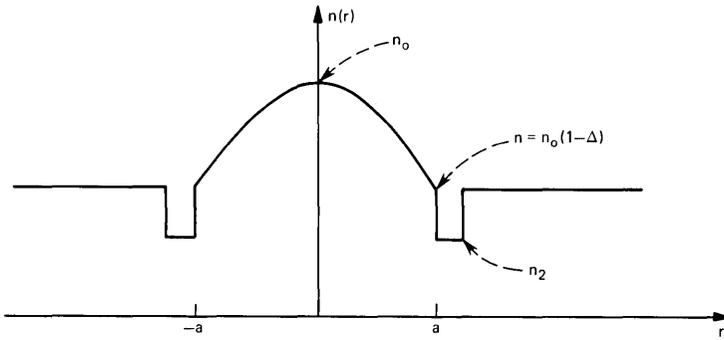


Fig. 5—Refractive-index profile of a fiber providing continuous mode filtering for S-modes by electromagnetic tunneling.

1b and serve the purpose of stripping off S-modes. The length of the filter sections depends on the S-mode losses of the filters. Since S-modes are leaky waves on the filter sections, their losses are very high, so that filters 10 cm to 1 m long should be sufficient. We have seen that the signal-to-noise ratio (66) increases as the distance d between mode filters decreases. The only limitation on the length of d or the number of mode filters L/d is the additional loss caused by the filters. If we decide to tolerate an additional filter loss of 1 dB/km ($2\alpha_f = 0.23 \text{ km}^{-1}$) for the entire fiber, we find from (74) with the numbers of (21)

$$d = 0.47 \text{ km.} \quad (96)$$

If the filter spacing is reduced below this distance, the filter-loss contributions increase above 1 dB/km.

The signal-to-noise ratio that is caused by unintentional coupling between P-modes and S-modes can be computed from (66) if a calibration measurement has been made. In principle, it would be possible to compute this value from (65); however, such a calculation would require a detailed knowledge of the unintentional power-coupling coefficients $H_{\nu n}$. In the absence of accurate information of unintentional fiber imperfections, such a calculation is not possible.

Instead of using discrete mode filters spaced at certain intervals it is also possible to design the fiber so that mode filtering is achieved continuously over its entire length. The index profile of such a fiber is shown in Fig. 5.

VII. ACKNOWLEDGMENT

E. A. J. Marcatili suggested the possibility of implementing the intentional fiber perturbations by bending the fiber axis and of using a fiber designed with continuous mode filters. Illuminating discussions with him are also gratefully acknowledged.

REFERENCES

1. D. Gloge and E. A. J. Marcatili, "Multimode Theory of Graded-Core Fibers," *B.S.T.J.*, *52*, No. 9 (November 1973), pp. 1563-1578.
2. D. Marcuse, *Theory of Dielectric Optical Waveguides*, New York: Academic Press, 1974.
3. D. Personick, "Time Dispersion in Dielectric Waveguides," *B.S.T.J.*, *50*, No. 3 (March 1971), pp. 843-859.
4. D. Marcuse, "Reduction of Pulse Dispersion by Intentional Mode Coupling," *B.S.T.J.*, *53*, No. 9 (November 1974), pp. 1795-1815.
5. D. Marcuse, "Coupled Mode Theory of Round Optical Fibers," *B.S.T.J.*, *53*, No. 6 (July-August 1973), pp. 817-842.
6. D. Gloge, "Propagation Effects in Optical Fibers," *IEEE Trans. Microw. Theory Tech.*, *MTT-23*, No. 1 (January 1975), pp. 106-120.
7. D. Gloge, "Weakly Guiding Fibers," *Appl. Opt.*, *10*, No. 10 (October 1971), pp. 2252-2258.
8. K. Petermann, "The Mode Attenuation in General Graded Core Multimode Fibers," *Arch. Elec. Uebertr.*, *29*, No. 7/8 (July-August 1975), pp. 345-348.
9. G. Goubau and F. Schwering, "On the Guided Propagation of Electromagnetic Beam Waves," *IRE Trans. Ant. Propag.*, *AP-9*, No. 3 (May 1961), pp. 248-256.
10. H. Kogelnik and T. Li, "Laser Beams and Resonators," *Appl. Opt.*, *5*, No. 10 (October 1966), pp. 1550-1567.
11. D. Marcuse, "Excitation of Parabolic-Index Fibers With Incoherent Sources," *B.S.T.J.*, *54*, No. 9 (November 1975), pp. 1507-1530.
12. D. Marcuse and H. M. Presby, "Mode Coupling in an Optical Fiber with Core Distortions," *B.S.T.J.*, *54*, No. 1 (January 1975), pp. 3-15.
13. L. S. Watkins, "Instrument for Continuously Monitoring Fiber Core and Outer Diameters," *Optical Fiber Transmission*, A digest of technical papers presented at the Topical Meeting on Optical Fiber Transmission, January 7-9, 1975, Williamsburg, Virginia; Washington, D. C.: Optical Society of America, 1975, pp. TuA4-1 to TuA4-4.
14. D. Gloge, "Optical Power Flow in Multimode Fibers," *B.S.T.J.*, *51*, No. 8 (October 1972), pp. 1767-1783.

The Boundary Layer in a Concentrated, Multicomponent Electrolyte

By J. A. LEWIS

(Manuscript received January 8, 1976)

We show formally how the application of boundary layer techniques to a well-stirred, concentrated, multicomponent electrolyte yields essentially the same results as those for a dilute, binary electrolyte. In particular, the coupled integral equations for the reactant concentration and the electric potential, derived by J. L. Blue for a dilute, binary electrolyte, remain valid in the general case with modified values of certain combinations of physical constants appearing in the equations.

I. INTRODUCTION

In this paper, we use the method of matched asymptotic expansions to derive boundary layer equations for a well-stirred, concentrated, multicomponent electrolyte containing a single ion reacting at the electrodes. We have two main objectives. First, we wish to show how the boundary layer approximation, up to now applied mainly to dilute, binary electrolytes or to the case of excess supporting electrolyte,¹ may be systematically generalized. Second, and more important, we show how the generalization leads to equations almost identical with those valid for a dilute, binary electrolyte, for which J. L. Blue² has derived an elegant and efficient method of solution.

In the general case, the singular perturbation in the boundary layer leads to a set of coupled convective diffusion (CD) equations for the concentrations or, in other words, a *vector* CD-equation in place of the single, scalar CD-equation for a dilute, binary electrolyte. The vector CD-equation is diagonalized by introducing the eigenvalues and eigenvectors of a reduced diffusivity matrix. The resulting uncoupled, scalar CD-equations can then be solved by Blue's method.

The perturbation yields linear CD-equations in the electrolyte, but the strongly nonlinear dependence of the electrode current density on the electrolyte potential and the reactant ion concentration is retained. In fact, the first approximation is simply the so-called "secondary current distribution," the potential satisfying Laplace's equation in

the electrolyte, but satisfying nonlinear boundary conditions at the electrodes. Furthermore, if, as is often the case, the current-carrying capacity of the electrode surface reaction is large compared with that due to ohmic conduction, the potential is approximately constant on the electrodes, giving the so-called "primary current distribution."

We begin in Sections II and III with a semiquantitative discussion of some of the approximations involved. Sections IV through IX contain the detailed analysis. The reader not interested in the details may skip to Sections X and XI, which contain very brief summaries of the formal perturbation solution and the equivalent boundary integral solution developed by J. L. Blue.

II. THE FIRST APPROXIMATION

A first approximation to the cathode current density serves as a convenient introduction to the general concentrated, n -component electrolyte. In a well-stirred electroplating bath, the ion concentrations are almost constant, except in thin diffusion boundary layers surrounding the electrodes. If the electrode current densities are bounded, so are the concentration gradients in the boundary layers. Thus, the changes in concentrations across the boundary layers are small when the boundary layer thicknesses are small enough. To a first approximation then, the concentrations are constant throughout the electrolyte, current being carried entirely by ohmic conduction. In the electrolyte, the dimensionless potential φ (equal to $F\Phi/RT$, for Faraday constant F , gas constant R , and absolute temperature T) satisfies Laplace's equation

$$\nabla^2 \varphi = 0.$$

On the electrodes, the current densities are determined by surface reactions producing cations (e.g., Cu^{++}) at the anode and depositing them at the cathode, the reaction rates being functions of the local concentrations and potentials. The potential φ then satisfies the boundary condition

$$k^{(0)} \partial \varphi / \partial n = k^{(0)} \mathbf{n} \cdot \nabla \varphi = \begin{cases} -j_a, & \text{on the anode,} \\ j_c, & \text{on the cathode,} \\ 0, & \text{otherwise,} \end{cases}$$

where \mathbf{n} is the outward normal, $k^{(0)}$ is the dimensionless bulk conductivity, and, for $e = a, c$,

$$j_e = j_e^{*(0)} [\exp(\alpha_e |\varphi - \varphi_e|) - \exp(-\beta_e |\varphi - \varphi_e|)].$$

The dimensionless ion exchange current density $j_e^{*(0)}$, here constant, but in general a function of the concentrations, and the exponents α_e , β_e characterize the surface reaction rates. The quantity $\varphi - \varphi_e$ is the

so-called "electrode surface overpotential" for electrolyte potential φ and electrode potential φ_e .

When $j_e^{*(0)}$ is large, $|\varphi - \varphi_e|$ must be small or, to a first approximation, $\varphi = \varphi_e$ on the electrodes, giving the so-called "primary current distribution." Note, however, that the primary current distribution leads to unbounded current density at electrode edges or reentrant corners, while the secondary current distribution does not. The exponential form of the electrode current relation, however, suggests that the primary and secondary current distributions differ only in a very small neighborhood of such a boundary singularity.

Once φ has been determined, the electrode current densities can be calculated. Furthermore, the approximation does not depend on the specific electrolyte, but only on the existence of thin boundary layers across which the concentration changes are small. It thus applies to dilute, binary and concentrated, multicomponent electrolytes alike, although, of course, the various constants, e.g., the conductivities, may be vastly different in the two cases.

To obtain a better approximation, we must include the effect of concentration variation across the boundary layers. Before doing so, however, let us consider a method of practical solution for φ , the method of boundary integrals widely used for the numerical solution of potential problems. Since φ is harmonic, it satisfies Green's third identity, which in the plane has the form

$$2\pi\varphi(P) = \int_{\Gamma} \left[\frac{\partial\varphi}{\partial n}(Q)G(P, Q) - \varphi(Q)\frac{\partial G}{\partial n}(P, Q) \right] ds(Q),$$

where Γ is the curve bounding the electrolyte and $G(P, Q) \sim \ln(1/|PQ|)$, as the point $P \rightarrow Q$, and is harmonic otherwise. When P lies on Γ , the identity becomes an integral equation for φ , nonlinear because $\partial\varphi/\partial n$ is a nonlinear function of φ . Very efficient numerical methods for solving this equation have been developed.

III. THE BINARY ELECTROLYTE

For the dilute, binary electrolyte, J. L. Blue² has extended the above boundary integral method to include the effect of concentration changes across the boundary layers. He derived a relation giving the concentration changes across the boundary layers in terms of integrals of the product of electrode current density and a kernel function whose form depends upon the local velocity distribution. These concentration changes affect the electrode current density in two ways. First, as we have already noted, the ion exchange current density is a function of concentration. Second, the potentials at the electrodes differ from those at the boundary layer edge, where the previous φ is actually

evaluated, by the so-called "concentration overpotentials" $\Delta\varphi_e$, $e = a, c$. These overpotentials are due to the net current carried by ions diffusing at different rates through the boundary layers. To the present degree of approximation, they are linear functions of the concentration changes across the boundary layers.

In a binary electrolyte, in which the two ion concentrations must be proportional, since the electrolyte is electrically neutral, $\Delta\varphi_e$ is proportional to the integral giving the concentration change. This constant of proportionality, the constant giving the concentration to be inserted into the ion exchange current density, and the bulk conductivity are the only constants containing material properties in the final relations. They thus characterize the electrolyte completely in this case and may thus be used in place of the individual diffusivities and mobilities.

In the above, we did not qualify the term "binary electrolyte" with the term "dilute." The discussion of the previous paragraph applies as well to concentrated binary electrolytes, provided that the boundary layer concentration changes are small. In this case, in the boundary layers, the equations are linearized around a state of constant bulk concentrations, so that the diffusivities, in general forming a 2×2 matrix, and the mobilities, both functions of the concentrations, are constants, evaluated at the constant values of the bulk concentrations. Using the charge neutrality condition, the potential and one concentration are eliminated from the two boundary layer equations governing the ion fluxes. The resulting equation has exactly the same form as that for a dilute, binary electrolyte, so that the solution has the same form, except that the proportionality constants have different values.

In the following, we extend this result to concentrated, multicomponent electrolytes, the case of the ternary ($n = 3$) electrolyte being of particular interest. As above, we linearize the equations around a state of constant bulk concentrations, using the boundary layer approximation (*not* the assumption of excess supporting electrolyte). We eliminate the potential and one concentration to obtain a set of $(n - 1)$ boundary layer equations for the concentrations, coupled by a reduced $(n - 1) \times (n - 1)$ diffusivity matrix. This system is solved by diagonalization, using the principal values and vectors of the reduced matrix. Finally, these solutions are combined to calculate the electrode concentrations and concentration overpotentials. We find that both of these quantities are proportional to a single integral as before, so that the electrolyte can be characterized by specifying the two constants of proportionality, instead of n^2 diffusivities and n mobilities.

IV. THE COMPLETE BOUNDARY VALUE PROBLEM

The steady flow of ions in a nonreacting, concentrated, n -component electrolyte is governed by the n conservation laws*

$$\nabla \cdot \mathbf{N}_i = 0, \quad (1)$$

for $i = 1, \dots, n$, where the ion flux \mathbf{N}_i is given in terms of the ion concentrations C_1, \dots, C_n and the potential Φ by

$$\mathbf{N}_i = C_i \mathbf{V} - \sum_{j=1}^n D_{ij} \nabla C_j - F z_i U_i C_i \nabla \Phi. \quad (2)$$

In these n relations, \mathbf{V} is the circulation velocity of the electrolyte due to stirring, F is the Faraday constant, z_i the charge number of the i th ion, and the diffusivities D_{ij} and mobilities U_i are functions of C_1, \dots, C_n .† We shall not impose any restrictions on the form of the D_{ij} 's, although it seems reasonable to require, for example, that for $i \neq j$, $D_{ij} \rightarrow 0$, as $C_i \rightarrow 0$. This makes the i th ion flux \mathbf{N}_i vanish, when the i th ion is absent, and gives the correct form for a dilute electrolyte. With

$$\nabla \cdot \mathbf{V} = 0, \quad (3)$$

as is appropriate for a liquid electrolyte, eqs. (1) and (2) yield the set of n nonlinear, second-order, partial differential equations

$$\nabla \cdot \left(\sum_{j=1}^n D_{ij} \nabla C_j + F z_i U_i C_i \nabla \Phi \right) = \mathbf{V} \cdot \nabla C_i. \quad (4)$$

The set of equations is completed by the charge neutrality condition

$$\sum_{i=1}^n z_i C_i = 0, \quad (5)$$

which implies that the current density

$$\mathbf{J} = \sum_{i=1}^n F z_i \mathbf{N}_i = -K \nabla \Phi - \sum_{i=1}^n \sum_{j=1}^n F z_i D_{ij} \nabla C_j, \quad (6)$$

with the conductivity

$$K = \sum_{i=1}^n F^2 z_i^2 U_i C_i, \quad (7)$$

satisfies

$$\nabla \cdot \mathbf{J} = 0. \quad (8)$$

* Capital letters denote dimensional quantities; lower case letters denote dimensionless quantities.

† In Appendix C, we show how D_{ij} and U_i can be expressed in terms of the electrochemical potential.

We assume that there is no flux through the surface bounding the electrolyte, except that flux due to the production of cation 1 (e.g., Cu^{++}) at the anode and deposition of the same cation at the cathode. Thus, for outward normal \mathbf{n} ,

$$\mathbf{V} \cdot \mathbf{n} = \mathbf{N}_i \cdot \mathbf{n} = 0 \quad (9)$$

for $i = 2, \dots, n$, while

$$\mathbf{N}_1 \cdot \mathbf{n} = \mathbf{J} \cdot \mathbf{n} / Fz_1 \quad (10)$$

on the boundary, where

$$\mathbf{J} \cdot \mathbf{n} = \begin{cases} -J_a, & \text{on the anode,} \\ J_c, & \text{on the cathode,} \\ 0, & \text{otherwise.} \end{cases} \quad (11)$$

We shall assume that the electrode current densities determined by the surface reaction rates have the form

$$J_e = J_e^*(C_1/C_e^*)^{\gamma_e} [\exp(\alpha_e F |\Phi - \Phi_e| / RT) - \exp(-\beta_e F |\Phi - \Phi_e| / RT)], \quad (12)$$

for $e = a, c$, where the ion exchange current density $J_e^*(C_1/C_e^*)^{\gamma_e}$ is assumed to be a function only of the concentration of the cation participating in the surface reactions, R is the gas constant, T the absolute temperature, Φ the electrolyte potential, Φ_e the electrode potential, and $|\Phi - \Phi_e|$ the surface overpotential. These conditions complete the boundary value problem.

Now let $L, V, C, RT/F, D, D/RT$, and FDC/L be some typical length, velocity, concentration, potential, diffusivity, mobility, and current density and set

$$\begin{aligned} (x, y, z) &= (X, Y, Z)/L, \\ c_i &= C_i/C, \\ \varphi &= F\Phi/RT, \\ \mathbf{v} &= \mathbf{V}/V, \\ \mathbf{j} &= L\mathbf{J}/FDC, \\ d_{ij} &= D_{ij}/D, \\ u_i &= RTU_i/D. \end{aligned}$$

The equations and boundary conditions then take the dimensionless form

$$\nabla \cdot \left(\sum_{j=1}^n d_{ij} \nabla c_j + z_i u_i c_i \nabla \varphi \right) = \text{Pe} \mathbf{v} \cdot \nabla c_i, \quad (13)$$

for $i = 1, \dots, n$, where the Peclet number $\text{Pe} = VL/D$,

$$\sum_{i=1}^n z_i c_i = 0, \quad (14)$$

$$\nabla \cdot \left(k \nabla \varphi + \sum_{i=1}^n \sum_{j=1}^n z_i d_{ij} \nabla c_j \right) = 0, \quad (15)$$

where

$$k = \sum_{i=1}^n z_i^2 u_i c_i. \quad (16)$$

With $\mathbf{v} \cdot \mathbf{n} = 0$ on the boundary,

$$\sum_{j=1}^n d_{ij} \partial c_j / \partial n + z_i u_i c_i \partial \varphi / \partial n = \begin{cases} -\mathbf{j} \cdot \mathbf{n} / z_1, & \text{for } i = 1, \\ 0, & \text{for } i = 2, \dots, n, \end{cases} \quad (17)$$

where

$$\begin{aligned} -\mathbf{j} \cdot \mathbf{n} &= k \partial \varphi / \partial n + \sum_{i=1}^n \sum_{j=1}^n z_i d_{ij} \partial c_j / \partial n \\ &= \begin{cases} j_a, & \text{on the anode,} \\ -j_c, & \text{on the cathode,} \\ 0, & \text{otherwise,} \end{cases} \end{aligned} \quad (18)$$

and, for $e = a, c$,

$$j_e = j_e^* c_i^{\gamma_e} [\exp(\alpha_e |\varphi - \varphi_e|) - \exp(-\beta_e |\varphi - \varphi_e|)], \quad (19)$$

$$j_e^* = (L J_e^* / FDC) (C / C^*)^{\gamma_e}. \quad (20)$$

V. THE BOUNDARY LAYER APPROXIMATION

To obtain explicit results, we consider a plane, rectangular region with a cathode at $Y = y = 0$ and an anode at $Y = L, y = 1$. We assume that the electrolyte is well stirred, with $Pe \gg 1$, so that the concentrations are nearly constant, except in thin boundary layers, of thickness of order ϵL , $\epsilon = (Pe)^{-\frac{1}{2}}$, balancing diffusion and convection near the electrodes. Far from the electrodes, we introduce the outer expansions

$$\begin{aligned} c_i &= \bar{c}_i(x, y, \epsilon) = c_i^{(0)} + O(\epsilon^3), \\ \varphi &= \bar{\varphi}(x, y, \epsilon) = \bar{\varphi}^{(0)}(x, y) + \epsilon \bar{\varphi}^{(1)}(x, y) + \dots, \end{aligned} \quad (21)$$

since eq. (13) implies that $\mathbf{v} \cdot \nabla \bar{c}_i = O(\epsilon^3)$, where the $c_i^{(0)}$'s are the constant bulk concentrations and $\bar{\varphi}^{(0)}, \bar{\varphi}^{(1)}$ are harmonic functions; i.e.,

$$\bar{\varphi}_{xx}^{(0)} + \bar{\varphi}_{yy}^{(0)} = \bar{\varphi}_{xx}^{(1)} + \bar{\varphi}_{yy}^{(1)} = 0. \quad (22)$$

In the boundary layers, we set

$$\begin{aligned} c_i &= \tilde{c}_i(x, \tilde{y}, \epsilon) = c_i^{(0)} + \epsilon \tilde{c}_i^{(1)}(x, \tilde{y}) + \dots, \\ \varphi &= \tilde{\varphi}(x, \tilde{y}, \epsilon) = \tilde{\varphi}^{(0)}(x, \tilde{y}) + \epsilon \tilde{\varphi}^{(1)}(x, \tilde{y}) + \dots, \end{aligned}$$

where the stretched variable $\tilde{y} = y/\epsilon$ in the cathode boundary layer and $\bar{y} = (1 - y)/\epsilon$ in the anode boundary layer. The assumption that the electrode current densities are bounded has already been introduced by choosing the leading term in the expansion of \tilde{c}_i to be constant. In the boundary layer eqs. (13) and (15) become

$$\left(\sum_{j=1}^n \tilde{d}_{ij} \tilde{c}_{j\tilde{y}} + z_i \tilde{u}_i \tilde{c}_i \tilde{\varphi}_{\tilde{y}} \right)_{\tilde{y}} - (\tilde{\psi}_{\tilde{y}} \tilde{c}_{ix} - \tilde{\psi}_x \tilde{c}_{i\tilde{y}}) = -\epsilon^2 \left(\sum_{j=1}^n \tilde{d}_{ij} \tilde{c}_{jx} + z_i \tilde{u}_i \tilde{c}_i \tilde{\varphi}_x \right)_x, \quad (23)$$

$$\left(\tilde{k} \tilde{\varphi}_{\tilde{y}} + \sum_{i=1}^n \sum_{j=1}^n z_i \tilde{d}_{ij} \tilde{c}_{j\tilde{y}} \right)_{\tilde{y}} = -\epsilon^2 \left(\tilde{k} \tilde{\varphi}_x + \sum_{i=1}^n \sum_{j=1}^n z_i \tilde{d}_{ij} \tilde{c}_{jx} \right)_x, \quad (24)$$

where we have introduced the reduced stream function

$$\tilde{\psi}(x, \tilde{y}, \epsilon) = \epsilon^{-2} \psi(x, \epsilon \tilde{y}),$$

with the x and y components of \mathbf{v} given by

$$u = \psi_y, \quad v = -\psi_x,$$

vanishing on the electrodes. The tildes on d_{ij} , etc., mean $\tilde{d}_{ij} = d_{ij}(\tilde{c}_1, \dots, \tilde{c}_n)$, etc. Note that the \tilde{d} 's, \tilde{u} 's, \tilde{k} , and $\tilde{\psi}$ must all be expanded around $\epsilon = 0$ to obtain the equations satisfied by $\tilde{c}_i^{(1)}$, $\tilde{\varphi}^{(0)}$, $\tilde{\varphi}^{(1)}$, \dots . The boundary conditions, for example at the cathode $y = \tilde{y} = 0$, have the form

$$\sum_{j=1}^n \tilde{d}_{ij} \tilde{c}_{j\tilde{y}} + z_i \tilde{u}_i \tilde{c}_i \tilde{\varphi}_{\tilde{y}} = \begin{cases} \epsilon \tilde{j}_c / z_1, & \text{for } i = 1, \\ 0, & \text{for } i = 2, \dots, n, \end{cases} \quad (25)$$

where

$$\tilde{j}_c = j_c^*(\tilde{c}_1)^{\nu_c} [\exp(\alpha_c \tilde{\varphi}) - \exp(-\beta_c \tilde{\varphi})], \quad (26)$$

\tilde{c}_1 and $\tilde{\varphi} > 0$ being evaluated at $\tilde{y} = 0$ and the cathode assumed to be grounded, so that $\varphi_c = 0$. Obviously \tilde{j}_c must also be expanded in powers of ϵ . Finally, multiplying eq. (25) by z_i and summing yields

$$\tilde{k} \tilde{\varphi}_{\tilde{y}} + \sum_{i=1}^n \sum_{j=1}^n z_i \tilde{d}_{ij} \tilde{c}_{j\tilde{y}} = \epsilon \tilde{j}_c. \quad (27)$$

The set of relations from which the boundary layer equations are obtained is completed by matching $\tilde{c}_i(x, \tilde{y}, \epsilon)$ with $\bar{c}_i(x, y, \epsilon)$ and $\tilde{\varphi}(x, \tilde{y}, \epsilon)$ with $\bar{\varphi}(x, y, \epsilon)$ for small y and large \tilde{y} . Specifically, we use intermediate matching, setting

$$y = \epsilon^\alpha y^*, \quad \tilde{y} = \epsilon^{\alpha-1} y^*,$$

for fixed y^* and $0 < \alpha < 1$, and considering the limit $\epsilon \rightarrow 0$. Matching

$$\tilde{c}_i(x, \epsilon^{\alpha-1} y^*, \epsilon) = c_i^{(0)} + \epsilon \tilde{c}_i^{(1)}(x, \epsilon^{\alpha-1} y^*)$$

with

$$\bar{c}_i(x, \epsilon^\alpha y^*, \epsilon) = c_i^{(0)} + O(\epsilon^3)$$

trivially yields the boundary conditions

$$\bar{c}_i^{(1)}(x, \infty) = 0,$$

for $i = 1, \dots, n$. Similarly, matching $\bar{\varphi}^{(0)}$ and $\bar{\varphi}^{(0)}$ gives

$$\bar{\varphi}^{(0)}(x, \infty) = \bar{\varphi}^{(0)}(x, 0).$$

The differential equation (24) and the boundary condition (27) imply that $\bar{\varphi}_{\bar{y}\bar{y}}^{(0)} = \bar{\varphi}_{\bar{y}}^{(0)} = 0$, so that

$$\bar{\varphi}^{(0)}(x, \bar{y}) \equiv \bar{\varphi}^{(0)}(x, 0).$$

To a first approximation, the outer potential $\bar{\varphi}^{(0)}$, as yet undetermined, simply penetrates the boundary layer.

VI. FIRST-ORDER MATCHING: THE SECONDARY CURRENT DISTRIBUTION

Boundary conditions on $\bar{\varphi}^{(0)}$ are found by matching

$$\bar{\varphi} = \bar{\varphi}^{(0)} + \epsilon \bar{\varphi}^{(1)}$$

with

$$\bar{\varphi} = \bar{\varphi}^{(0)} + \epsilon \bar{\varphi}^{(1)}.$$

The zero-order inner potential $\bar{\varphi}^{(0)}$ has already been determined. The first-order potential $\bar{\varphi}^{(1)}$ satisfies the differential equation

$$\left[k^{(0)} \bar{\varphi}_{\bar{y}}^{(1)} + \sum_{i=1}^n \sum_{j=1}^n z_i d_{ij}^{(0)} \bar{c}_{j\bar{y}}^{(1)} \right]_{\bar{y}} = 0,$$

and the boundary condition

$$k^{(0)} \bar{\varphi}_{\bar{y}}^{(1)} + \sum_{i=1}^n \sum_{j=1}^n z_i d_{ij}^{(0)} \bar{c}_{j\bar{y}}^{(1)} = j_c^{(0)},$$

at $\bar{y} = 0$, where $d_{ij}^{(0)}$, $k^{(0)}$ are evaluated at $c_1^{(0)}, \dots, c_n^{(0)}$ and $j_c^{(0)}$ at $c_1^{(0)}$, $\bar{\varphi}^{(0)}(x, 0) = \bar{\varphi}^{(0)}(x, 0)$, so that

$$j_c^{(0)}(x) = j_c^* [c_1^{(0)}]^{r_c} \{ \exp[\alpha_\epsilon \bar{\varphi}^{(0)}(x, 0)] - \exp[-\beta_\epsilon \bar{\varphi}^{(0)}(x, 0)] \}.$$

Two integrations yield

$$\bar{\varphi}^{(1)}(x, \bar{y}) = \hat{\varphi}^{(1)}(x) + \left[\bar{y} j_c^{(0)}(x) - \sum_{i=1}^n \sum_{j=1}^n z_i d_{ij}^{(0)} \bar{c}_j^{(1)}(x, \bar{y}) \right] / k^{(0)}, \quad (28)$$

where $\hat{\varphi}^{(1)}(x)$ is an arbitrary function of x . The last term can be identified as the concentration overpotential. Since, for large \bar{y} , $\bar{c}_i^{(1)} \rightarrow 0$,

$$\bar{\varphi} \sim \bar{\varphi}^{(0)}(x, 0) + \epsilon [\hat{\varphi}^{(1)}(x) + \epsilon^{\alpha-1} y^* j_c^{(0)}(x) / k^{(0)}],$$

for $\bar{y} = \epsilon^{\alpha-1}y^*$ and $\epsilon \rightarrow 0$. On the other hand, expanding in Taylor series gives

$$\bar{\varphi} \sim \bar{\varphi}^{(0)}(x, 0) + \epsilon^{\alpha}y^* \bar{\varphi}_y^{(0)}(x, 0) + \epsilon \bar{\varphi}^{(1)}(x, 0) + \dots,$$

for $y = \epsilon^{\alpha}y^*$. First-order matching then yields the zero-order boundary condition

$$\bar{\varphi}_y^{(0)}(x, 0) = j_c^{(0)}(x)/k^{(0)}$$

and

$$\hat{\varphi}^{(1)}(x) = \bar{\varphi}^{(1)}(x, 0),$$

another penetration condition. A similar analysis produces the boundary condition

$$\bar{\varphi}_y^{(0)}(x, 1) = j_a^{(0)}(x)/k^{(0)}$$

at the anode $y = 1$. Since $j_c^{(0)}$ and $j_a^{(0)}$ depend only on $\bar{\varphi}^{(0)}$, we now have a complete boundary value problem for $\bar{\varphi}^{(0)}$. To a first approximation then, the potential φ satisfies Laplace's equation

$$\varphi_{xx} + \varphi_{yy} = 0, \quad (29)$$

in the electrolyte, and the nonlinear boundary condition

$$k^{(0)} \partial \varphi / \partial n = \begin{cases} j_a, & \text{on the anode,} \\ -j_c, & \text{on the cathode,} \\ 0, & \text{otherwise,} \end{cases} \quad (30)$$

where, for $e = a, c$,

$$j_e = j_e^* [c_1^{(0)}]^{\gamma_e} [\exp(\alpha_e |\varphi - \varphi_e|) - \exp(-\beta_e |\varphi - \varphi_e|)], \quad (31)$$

yielding the so-called "secondary current distribution."

Although we have derived this secondary current distribution specifically for small boundary layer thickness and moderate electrode current density, it probably is valid over a much wider range. The basic assumption is simply that the concentration variations are small. Thus, it is probably still valid when the flow is turbulent and the present simple description of the flow field does not suffice or at loci of boundary singularities (sharp edges, etc.) where the gradients are not necessarily normal to the boundary, as is assumed tacitly when one introduces the stretched variable \bar{y} in the boundary layer approximation. In any event, by calculating the first-order terms $\bar{c}_i^{(1)}$, $\bar{\varphi}^{(1)}$, we can obtain an error estimate or an improved approximation.

VII. CONCENTRATION VARIATIONS

Although we have already derived boundary conditions on the $\bar{c}_i^{(1)}$'s, we have not yet displayed the equations which they satisfy. With

constant leading terms $c_i^{(0)}$ and $\tilde{\psi}$ expanded in the form

$$\tilde{\psi} = \tilde{\psi}_0 + \epsilon \tilde{\psi}_1 + \dots = \tilde{y}^2 f(x) + \epsilon \tilde{y}^3 g(x) + \dots,$$

where $f(x) = \psi_{y y}(x, 0)/2$, $g(x) = \psi_{y y y}(x, 0)/6$, \dots , eq. (23) becomes, to first order,

$$\sum_{j=1}^n d_{ij}^{(0)} \tilde{c}_{j\tilde{y}\tilde{y}}^{(1)} + (k_i^{(0)}/z_i) \tilde{\varphi}_{\tilde{y}\tilde{y}}^{(1)} = \tilde{\psi}_y^{(0)} \tilde{c}_{ix}^{(1)} - \tilde{\psi}_x^{(0)} \tilde{c}_{i\tilde{y}}^{(1)}, \quad (32)$$

for $i = 1, \dots, n$, where the i th partial conductivity

$$k_i^{(0)} = z_i^2 u_i^{(0)} c_i^{(0)},$$

while the boundary condition (25) at $\tilde{y} = 0$ yields

$$\sum_{j=1}^i d_{ij}^{(0)} \tilde{c}_{j\tilde{y}}^{(1)} + (k_i^{(0)}/z_i) \tilde{\varphi}_{\tilde{y}}^{(1)} = \begin{cases} j_c^{(0)}/z_1, & \text{for } i = 1, \\ 0, & \text{for } i = 2, \dots, n. \end{cases} \quad (33)$$

We can use the charge neutrality condition

$$\tilde{c}_n^{(1)} = - \sum_{j=1}^{n-1} (z_j/z_n) \tilde{c}_j^{(1)}$$

and the relation

$$k^{(0)} \tilde{\varphi}_{\tilde{y}}^{(1)} = j_c^{(0)} - \sum_{p=1}^n \sum_{j=1}^n z_p d_{pj}^{(0)} \tilde{c}_{j\tilde{y}}^{(1)},$$

found from eq. (28), to eliminate $\tilde{c}_n^{(1)}$ and $\tilde{\varphi}^{(1)}$. We obtain the reduced boundary layer equations

$$\sum_{j=1}^{n-1} d'_{ij} \tilde{c}_{j\tilde{y}\tilde{y}}^{(1)} = \tilde{\psi}_{\tilde{y}}^{(0)} \tilde{c}_{ix}^{(1)} - \tilde{\psi}_x^{(0)} \tilde{c}_{i\tilde{y}}^{(1)}, \quad (34)$$

for $i = 1, \dots, n - 1$, and the boundary conditions at $\tilde{y} = 0$

$$\sum_{j=1}^{n-1} d'_{ij} \tilde{c}_{j\tilde{y}}^{(1)} = \begin{cases} (1 - t_1^{(0)}) j_c^{(0)}/z_1, & \text{for } i = 1, \\ -t_i^{(0)} j_c^{(0)}/z_i, & \text{for } i = 2, \dots, n - 1, \end{cases} \quad (35)$$

where $t_i^{(0)} = k_i^{(0)}/k^{(0)}$ is the i th transference number and the reduced diffusivity matrix $[d'_{ij}]$ has elements

$$d'_{ij} = d_{ij}^{(0)} - (z_j/z_n) d_{in}^{(0)} - t_i^{(0)} \sum_{p=1}^n (z_p/z_i) [d_{pj}^{(0)} - (z_j/z_n) d_{pn}^{(0)}], \quad (36)$$

for $i = 1, \dots, n - 1$, $j = 1, \dots, n - 1$. The boundary value problem is completed by the conditions at infinity

$$\tilde{c}_i^{(1)}(x, \infty) = 0, \quad (37)$$

and the "initial condition"

$$\tilde{c}_i^{(1)}(0, \bar{y}) = 0. \quad (38)$$

As usual, the boundary layer approximation changes the equations from elliptic to parabolic type, requiring a condition at $x = 0$ in the time-like variable x . Equation (38) is equivalent to the assumption that the concentrations have their free-stream, bulk values near the leading edge $x = \bar{y} = 0$. We now have a complete set of coupled equations and boundary conditions for the first-order boundary layer concentration variations $\tilde{c}_i^{(1)}$. We note in passing that the corresponding set for the second-order variations $\tilde{c}_i^{(2)}$ would be much more complicated, containing not only $\tilde{\psi}^{(1)}$, as in the case of the dilute binary electrolyte, but also first-order terms $\tilde{d}_{ij}^{(1)}$, etc., due to the concentration dependent diffusivities and mobilities.

VIII. THE VECTOR CD-EQUATION AND ITS SOLUTION

Equation (34) and boundary condition (35) can be written in the vector form

$$[d'][\tilde{c}^{(1)}]\bar{y}\bar{y} = \tilde{\psi}_{\bar{y}}^{(0)}[\tilde{c}^{(1)}]_x - \tilde{\psi}_x^{(0)}[\tilde{c}^{(1)}]_{\bar{y}} \quad (39)$$

in $\bar{y} > 0$, and

$$[d'][\tilde{c}^{(1)}]_{\bar{y}} = [T^{(0)}]j_c^{(0)} \quad (40)$$

on $\bar{y} = 0$, where $[d']$ is the $(n - 1) \times (n - 1)$ matrix

$$[d'] = \begin{bmatrix} d'_{11} & \cdots & d'_{1(n-1)} \\ \cdots & \cdots & \cdots \\ d'_{(n-1)1} & \cdots & d'_{(n-1)(n-1)} \end{bmatrix},$$

and $[\tilde{c}^{(1)}]$ and $[T^{(0)}]$ the column vectors

$$[\tilde{c}^{(1)}] = \begin{bmatrix} \tilde{c}_1^{(1)} \\ \vdots \\ \tilde{c}_{n-1}^{(1)} \end{bmatrix}, \quad [T^{(0)}] = \begin{bmatrix} (1 - t_1)/z_1 \\ \vdots \\ -t_{n-1}/z_{n-1} \end{bmatrix}.$$

By way of comparison, for the binary electrolyte ($n = 2$), one has

$$\tilde{c}_{\bar{y}\bar{y}}^{(1)} = \tilde{\psi}_{\bar{y}}^{(0)}\tilde{c}_x^{(1)} - \tilde{\psi}_x^{(0)}\tilde{c}_{\bar{y}}^{(1)}, \quad (41)$$

for $\bar{y} > 0$, and

$$\tilde{c}_{\bar{y}}^{(1)} = T^{(0)}j_c^{(0)}, \quad (42)$$

for $\bar{y} = 0$, where $d'_{11} = 1$, $\tilde{c}^{(1)} = \tilde{c}_1^{(1)}$, $T^{(0)} = T_1^{(0)}(1 - t_1)/z_1$. The solution of this problem is given in Appendix A, together with that of the corresponding axially symmetric problem, important for jet plating and for the rotating disk electrode.

We now introduce the eigenvalues $d^{(p)}$ and the eigenvectors $[u^{(p)}]$ of $[d']$, such that*

$$[d'] [u^{(p)}] = d^{(p)} [u^{(p)}],$$

for $p = 1, \dots, n - 1$. If we set

$$[\bar{c}^{(1)}] = \sum_{p=1}^{n-1} c^{(p)} [u^{(p)}],$$

$$[T^{(0)}] = \sum_{p=1}^{n-1} \tau^{(p)} [u^{(p)}],$$

eqs. (39) and (40) are satisfied if

$$d^{(p)} c_{\bar{y}\bar{y}}^{(p)} = \bar{j}_{\bar{y}}^{(0)} c_x^{(p)} - \bar{j}_x^{(0)} c_{\bar{y}}^{(p)}, \quad (43)$$

in $\bar{y} > 0$, and

$$d^{(p)} c_{\bar{y}}^{(p)} = \tau^{(p)} j_c^{(0)}, \quad (44)$$

on $\bar{y} = 0$, for $p = 1, \dots, n - 1$.

Now, in Appendix A we show that, if $c(x, y)$ satisfies

$$c_{yy} = \psi_y c_x - \psi_x c_y,$$

with $\psi = y^2 f(x)$, and

$$c(0, y) = c(x, \infty) = 0,$$

then

$$c(x, 0) = \int_0^x c_y(x', 0) K(x, x') dx', \quad (45)$$

where the kernel

$$K(x, x') = -\left(\frac{1}{6}\right)^{\frac{1}{2}} \left(\int_{x'}^x [f(z)]^{\frac{1}{2}} dz \right)^{-\frac{3}{2}} / \Gamma\left(\frac{2}{3}\right). \quad (46)$$

Thus,

$$c^{(p)}(x, 0) = [d^{(p)}]^{-\frac{1}{2}} \tau^{(p)} \int_0^x j_c^{(0)}(x') K(x, x') dx', \quad (47)$$

with a factor $(d^{(p)})^{\frac{1}{2}}$, found by replacing f by $f/d^{(p)}$ in K , and a factor $(d^{(p)})^{-1}$ from the boundary conditions. Finally, to first order in ϵ , the concentrations at the cathode are given by

$$\bar{c}_i(x, 0) = c_i^{(0)} + \epsilon \mu_i \int_0^x j_c^{(0)}(x') K_c(x, x') dx', \quad (48)$$

for $i = 1, \dots, n - 1$, where

$$\mu_i = \sum_{p=1}^{n-1} [d^{(p)}]^{-\frac{1}{2}} \tau^{(p)} u_i^{(p)}, \quad (49)$$

* We assume that such u 's and d 's exist. This must be verified in each particular case.

and we now write K_c for K to indicate that the kernel depends on the velocity distribution near the cathode. A similar expression

$$\tilde{c}_i(x, 0) = c_i^{(0)} + \epsilon \mu_i \int_0^x j_a^{(0)}(x') K_a(x, x') dx', \quad (50)$$

with x measured from the anode leading edge and $i = 1, \dots, n - 1$, is valid for the concentrations at the anode. The concentration of the n th ion at cathode and anode is found from the charge neutrality relation

$$\tilde{c}_n(x, 0) = - \sum_{i=1}^{n-1} (z_i/z_n) \tilde{c}_i(x, 0). \quad (51)$$

With the $\tilde{c}_i^{(1)}$'s determined, we are almost ready to calculate the first-order corrections to the electrode current densities. For example, at the cathode we have

$$j_c = j_c^* (\tilde{c}_1)^{\gamma_c} [\exp(\alpha_c \tilde{\varphi}) - \exp(-\beta \tilde{\varphi})],$$

where

$$\begin{aligned} \tilde{c}_1 &= c_1^{(0)} + \epsilon \tilde{c}_1^{(1)}, \\ \tilde{\varphi} &= \tilde{\varphi}^{(0)} + \epsilon \tilde{\varphi}^{(1)}, \end{aligned}$$

all quantities being evaluated at $\tilde{y} = 0$. With

$$j_c = j_c^{(0)} + \epsilon \tilde{j}_c^{(1)} + \dots,$$

we find

$$j_c^{(0)} = j_c^* [c_1^{(0)}]^{\gamma_c} [\exp(\alpha_c \tilde{\varphi}^{(0)}) - \exp(-\beta_c \tilde{\varphi}^{(0)})],$$

as before, and

$$\begin{aligned} \tilde{j}_c^{(1)} &= j_c^* [c_1^{(0)}]^{\gamma_c} \{ [\gamma_c \tilde{c}_1^{(1)} + \alpha_c \tilde{\varphi}^{(1)}] \exp(\alpha_c \tilde{\varphi}^{(0)}) \\ &\quad - [\gamma_c \tilde{c}_1^{(1)} - \beta_c \tilde{\varphi}^{(1)}] \exp(-\beta_c \tilde{\varphi}^{(0)}) \}, \end{aligned}$$

where

$$\tilde{\varphi}^{(1)} = \tilde{\varphi}^{(1)}(x, 0) - \sum_{i=1}^n \sum_{j=1}^n z_i d_{ij}^{(0)} \tilde{c}_j^{(1)}(x, 0) / k^{(0)}. \quad (52)$$

We still must determine the first-order outer potential $\tilde{\varphi}^{(1)}$. It is the solution of the linear potential problem

$$\tilde{\varphi}_{xx}^{(1)} + \tilde{\varphi}_{yy}^{(1)} = 0, \quad (53)$$

in the electrolyte,

$$k^{(0)} \partial \tilde{\varphi}^{(1)} / \partial n = \begin{cases} \tilde{j}_a^{(1)}, & \text{on the anode,} \\ -\tilde{j}_c^{(1)}, & \text{on the cathode,} \\ 0, & \text{otherwise,} \end{cases} \quad (54)$$

where the $\tilde{j}_e^{(1)}$'s are linear functions of $\tilde{\varphi}^{(1)}$, with coefficients which

depend on $c_1^{(0)}$, $\bar{\varphi}^{(0)}$, and $\bar{z}_i^{(1)}$. This boundary condition is found from second-order matching of $\bar{\varphi}$ and $\bar{\varphi}$, just as the boundary condition on $\bar{\varphi}^{(0)}$ was found from first-order matching. As is usual in boundary layer problems, the complexity of matching increases rapidly with order so this rather messy matching is relegated to Appendix B. Once $\bar{\varphi}^{(1)}$ has been determined, the electrode current densities can be calculated to first order in ϵ .

IX. THE CONCENTRATION OVERPOTENTIALS

The concentration variations $\bar{z}_i^{(1)}$ appear in the $\bar{j}_e^{(1)}$'s in two ways. First of all, there is the power law dependence of the ion exchange densities on the concentration \bar{z}_1 of the "active" cation (e.g., Cu^{++}). This concentration is given by

$$\bar{z}_1 = c_1^{(0)} + \epsilon\mu_1 \int_0^x j_e^{(0)} K_e dx' \quad (55)$$

in terms of the bulk concentration $c_1^{(0)}$ and the constant μ_1 .

The second concentration dependence is contained in the so-called "concentration overpotential," the difference between the potential $\bar{\varphi}(x, 0)$ at the electrode and the potential $\bar{\varphi}(x, 0)$ at the edge of the boundary layer. To first order in ϵ , this concentration overpotential is given by

$$\Delta\varphi_e = \epsilon(\bar{\varphi}^{(1)} - \bar{\varphi}^{(1)}) = \epsilon\delta \int_0^x j_e^{(0)} K_e dx', \quad (56)$$

where

$$\delta = - \sum_{i=1}^n \sum_{j=1}^{n-1} z_i [d_{ij}^{(0)} - (z_j/z_n) d_{in}^{(0)}] \mu_j / k^{(0)}. \quad (57)$$

X. FORMAL PERTURBATION SOLUTION

Although we shall eventually abandon this approach in favor of an equivalent scheme more convenient for numerical analysis, let us briefly summarize the successive steps by which the electrode current densities are calculated to first order in ϵ . There are three such steps:

- (i) The nonlinear potential problem. First, we calculate the harmonic function $\bar{\varphi}^{(0)}$, satisfying nonlinear, mixed-boundary conditions. As is usual in boundary layer theory, we begin with a nonlinear problem, the only nonlinear problem in the sequence. It yields first approximations to the electrode current densities, which may be adequate for many purposes. It also furnishes a boundary condition for the boundary layer concentration variations.

- (ii) Concentration variations. The concentration variations $\bar{c}_1^{(1)}$, or rather $\bar{c}_1^{(1)}$ and $\Delta\varphi_e$, which are the only quantities involved in $\bar{j}_e^{(1)}$, are calculated from eqs. (55) and (56), giving them in terms of a single integral of the product of electrode current density and given kernels.
- (iii) The linear potential problem. With $\bar{c}_1^{(1)}$, $\Delta\varphi_e$ known, the harmonic function $\bar{\varphi}^{(1)}$, satisfying linear, mixed-boundary conditions in $\bar{\varphi}^{(1)}$, $\bar{c}_1^{(1)}$, $\Delta\varphi_e$ is calculated and, from it, a second approximation to the electrode current densities, good to order ϵ .

XI. SOLUTION BY BOUNDARY INTEGRALS

Instead of following the above rather cumbersome procedure, we can adopt the approach suggested by J. L. Blue,² forming a nonlinear, coupled set of integral equations whose solution, e.g., by successive approximation, yields the electrode current densities directly to order ϵ . From this point of view, the formal perturbation scheme simply validates this method. For example, perturbation tells us that the integral solution of the boundary layer equation is good to order ϵ and no more.

We begin with Green's third identity for a plane region:

$$2\pi\varphi(P) = \int_{\Gamma_a} [j_a(Q)/k]G(P, Q)ds(Q) - \int_{\Gamma_c} [j_c(Q)/k] \cdot G(P, Q)ds(Q) - \int_{\Gamma} \varphi(Q) \frac{\partial G}{\partial n}(P, Q)d(Q),$$

where G has the fundamental singularity at $P = Q$, Γ_a is the anode, Γ_c the cathode, and Γ the whole boundary, and we now denote $\bar{\varphi}^{(0)} + \epsilon\bar{\varphi}^{(1)}$ by φ and the bulk conductivity $k^{(0)}$ by k . For each set of values of j_a , j_c , φ , and points P on the boundary, the solution of this integral equation gives φ . Similarly, from the integrals

$$c_1 = c_1^{(0)} + \epsilon\mu_1 \int_0^x j_e K_e dx',$$

$$\Delta\varphi_e = +\epsilon\delta \int_0^x j_e K_e dx',$$

with $e = a, c$, we can calculate c_1 and $\Delta\varphi_e$ from the given values of j_a and j_c . With φ , c_1 , and $\Delta\varphi_e$ now known on the electrodes, we can calculate new values of j_a and j_c from the relations

$$j_e = j_e^*(c_1)^{\gamma_e} [\exp(\alpha_e |\varphi + \Delta\varphi_e - \varphi_e|) - \exp(-\beta_e |\varphi + \Delta\varphi_e - \varphi_e|)],$$

for $e = a, c$, the anode potential $\varphi_a > 0$, and the cathode potential $\varphi_c = 0$ (for a grounded cathode). Using these new values of electrode

current densities, compared with the old values, a successive approximation scheme can be developed which converges rapidly.

Note that this scheme is formally identical with that for the dilute, binary electrolyte. Furthermore, within the framework of this calculation, the concentrated, n -component electrolyte can be characterized completely by giving its bulk conductivity k and the constants μ_1 and δ . The values of the individual diffusivities d_{ij} and mobilities u_i , from which μ_1 and δ can be calculated, are of no importance. Only k , μ_1 , and δ enter the final set of equations.

APPENDIX A

Integral Solution of the Boundary Layer Equation

Let $c(x, y)$ satisfy the boundary layer equation

$$c_{yy} = \psi_y c_x - \psi_x c_y \quad (58)$$

in $x > 0$, $y > 0$, with $\psi = y^2 f(x)$ and

$$c(0, y) = c(x, \infty) = 0. \quad (59)$$

For use in the method of boundary integrals, we wish to express the boundary values of c in terms of an integral over the boundary values of c_y , i.e., we seek the kernel function K in the relation

$$c(x, 0) = \int_0^x c_y(x', 0) K(x, x') dx'. \quad (60)$$

Following Levich³ and Lighthill,⁴ we make the change of variable

$$\xi = \int_0^x [f(z)]^{\frac{1}{2}} dz, \quad \eta = (\psi)^{\frac{1}{2}} = y[f(x)]^{\frac{1}{2}},$$

the form of η being motivated by the observation that the right-hand side of eq. (58) vanishes when c is a function of ψ alone. We find that $v(\xi, \eta) = c(x, y)$ satisfies the equation

$$v_{\eta\eta} = 2\eta v_{\xi}, \quad (61)$$

if $\xi > 0$, $\eta > 0$, and

$$v(0, \eta) = v(\xi, \infty) = 0. \quad (62)$$

Denote the boundary values of v_η by g , so that

$$v_\eta(\xi, 0) = g(\xi).$$

Then the Laplace transform \bar{v} , given by

$$\bar{v}(\eta, s) = \int_0^\infty e^{-s\xi} v(\xi, \eta) d\xi,$$

satisfies the ordinary differential equation

$$\bar{v}'' - 2s\eta\bar{v} = 0,$$

in $0 < \eta < \infty$, and the boundary conditions

$$\bar{v}'(0, s) = \bar{g}(s), \quad \bar{v}(\infty, s) = 0,$$

where \bar{g} is the transform of g . Now suppose $u(\xi, \eta)$ is any solution of eq. (61), satisfying eq. (62). Then we have the identity between transforms

$$\bar{v}(\eta, s) = \bar{g}(s)\bar{u}(\eta, s)/\bar{u}'(0, s),$$

giving

$$\bar{v}(0, s) = \bar{v}'(0, s)[\bar{u}(0, s)/\bar{u}'(0, s)]. \quad (63)$$

A convenient choice for u is the similar solution of eqs. (61) and (62), satisfying the condition

$$u(\xi, 0) = 1,$$

so that $\bar{u}(0, s) = 1/s$. In this case, u is a function of $\zeta = \eta/\xi^{\frac{1}{3}}$ alone, given by

$$u = \left[\int_{\zeta}^{\infty} \exp(-2z^3/9) dz \right] / \left(\frac{1}{6} \right)^{\frac{1}{3}} \Gamma\left(\frac{1}{3} \right),$$

so that

$$u_{\eta}(\xi, 0) = -\xi^{-\frac{1}{3}} / \left(\frac{1}{6} \right)^{\frac{1}{3}} \Gamma\left(\frac{1}{3} \right)$$

and, with $\Gamma(k)/s^k$ the transform of ξ^{1-k} ,

$$\bar{u}'(0, s) = -\Gamma\left(\frac{2}{3} \right) / \left(\frac{1}{6} \right)^{\frac{1}{3}} \Gamma\left(\frac{1}{3} \right) s^{\frac{1}{3}}.$$

Thus,

$$\bar{u}(0, s)/\bar{u}'(0, s) = -\left(\frac{1}{6} \right)^{\frac{1}{3}} \Gamma\left(\frac{1}{3} \right) / \Gamma\left(\frac{2}{3} \right) s^{\frac{1}{3}}.$$

Inversion of eq. (63), using the convolution theorem, then yields

$$v(\xi, 0) = \int_0^{\xi} v_{\eta}(\xi', 0) \kappa(\xi - \xi') d\xi',$$

where κ , the inverse of $\bar{u}(0, s)/\bar{u}'(0, s)$, is given by

$$\kappa(\xi) = -\left(\frac{1}{6} \right)^{\frac{1}{3}} \xi^{-\frac{1}{3}} / \Gamma\left(\frac{2}{3} \right).$$

Using the relations

$$v(\xi', 0) d\xi' = c_v(x', 0) dx',$$

$$\xi - \xi' = \int_{x'}^x [f(z)]^{\frac{1}{3}} dz,$$

we finally obtain eq. (60), with

$$K(x, x') = -\left(\frac{1}{6} \right)^{\frac{1}{3}} \left(\int_{x'}^x [f(z)]^{\frac{1}{3}} dz \right)^{-\frac{1}{3}} / \Gamma\left(\frac{2}{3} \right). \quad (64)$$

K is of convolution type, i.e., a function of $x - x'$ alone, only when f is constant, e.g., for Poiseuille flow near the wall of a channel or tube. For the laminar boundary layer flow on a semi-infinite flat plate,

$$f \sim x^{-\frac{1}{2}},$$

so that

$$K \sim (x^{\frac{1}{2}} - x'^{\frac{1}{2}})^{-\frac{3}{2}}.$$

A slight modification of the above yields the integral relation

$$c(r, 0) = \int_0^r c_z(r', 0)L(r, r')r'dr', \quad (65)$$

for the axially symmetric problem, where

$$c_{zz} = r^{-1}(\psi_z c_r - \psi_r c_z),$$

in $r > 0, z > 0$, with $\psi = z^2 f(r)$ and

$$c(0, z) = c(r, \infty) = 0,$$

where the kernel function

$$L(r, r') = - \left(\frac{1}{6} \right)^{\frac{1}{2}} \left(\int_{r'}^r [f(s)]^{\frac{1}{2}} ds \right)^{-\frac{3}{2}} / \Gamma\left(\frac{2}{3}\right). \quad (66)$$

We merely replace the previously defined ξ by

$$\xi = \int_0^r [f(s)]^{\frac{1}{2}} s ds.$$

Equations (65) and (66) apply, for example, to the rotating disk electrode and to "jet plating," where a jet of electrolyte is directed against an electrode.

APPENDIX B

Second-Order Potential Matching

To obtain a boundary condition for the first-order outer potential $\bar{\varphi}^{(1)}$, we must match

$$\bar{\varphi} = \bar{\varphi}^{(0)} + \epsilon \bar{\varphi}^{(1)} + \epsilon^2 \bar{\varphi}^{(2)}$$

with

$$\bar{\varphi} = \bar{\varphi}^{(0)} + \epsilon \bar{\varphi}^{(1)} + \epsilon^2 \bar{\varphi}^{(2)},$$

where

$$\bar{\varphi}^{(0)} = \bar{\varphi}^{(0)}(x, 0), \quad (67)$$

$$\bar{\varphi}^{(1)} = \bar{\varphi}^{(1)}(x, 0) + \bar{y} j_e^{(0)}/k^{(0)} - \sum \sum z_i \bar{d}_{ij}^{(0)} \bar{c}_j^{(1)}/k^{(0)}. \quad (68)$$

The second-order inner potential $\bar{\varphi}^{(2)}$ satisfies the equation

$$\left[k^{(0)} \bar{\varphi}_{\bar{y}\bar{y}}^{(2)} + \bar{k}^{(1)} \bar{\varphi}_{\bar{y}\bar{y}}^{(1)} + \sum_{i=1}^n \sum_{j=1}^n z_i (d_{ij}^{(0)} \bar{c}_{j\bar{y}}^{(2)} + \bar{d}_{ij}^{(1)} \bar{c}_{j\bar{y}}^{(1)}) \right]_{\bar{y}} = -k^{(0)} \bar{\varphi}_{xx}^{(0)}, \quad (69)$$

and the boundary condition at $\bar{y} = 0$

$$k^{(0)} \bar{\varphi}_{\bar{y}}^{(2)} + \bar{k}^{(1)} \bar{\varphi}_{\bar{y}}^{(1)} + \sum_{i=1}^n \sum_{j=1}^n z_i (d_{ij}^{(0)} \bar{c}_{j\bar{y}}^{(2)} + d_{ij}^{(1)} \bar{c}_{j\bar{y}}^{(1)}) = \bar{j}_c^{(1)}, \quad (70)$$

where k , d_{ij} , j_c have been expanded in the forms

$$\begin{aligned} k &= k^{(0)} + \epsilon \bar{k}^{(1)} + \dots, \\ d_{ij} &= d_{ij}^{(0)} + \epsilon \bar{d}_{ij}^{(1)} + \dots, \\ j_c &= j_c^{(0)} + \epsilon \bar{j}_c^{(1)} + \dots. \end{aligned}$$

Note that $\bar{k}^{(1)}$, $\bar{d}_{ij}^{(1)}$, $\bar{j}_c^{(1)}$ are linear functions of the $\bar{c}_i^{(1)}$'s; for example,

$$k = k^{(0)} + \epsilon \sum_{p=1}^n (\partial k^{(0)} / \partial c_p) \bar{c}_p^{(1)}.$$

Two integrations give

$$\begin{aligned} \bar{\varphi}^{(2)} &= \bar{\varphi}^{(2)} + \bar{y} \bar{j}_c^{(1)} / k^{(0)} - \bar{y}^2 \bar{\varphi}_{xx}^{(0)} / 2 - \sum_{i=1}^n \sum_{j=1}^n z_i d_{ij}^{(0)} \bar{c}_j^{(2)} \\ &+ \int_0^{\bar{y}} \left[\sum_{i=1}^n \sum_{j=1}^n z_i (\bar{k}^{(1)} d_{ij}^{(0)} / k^{(0)} - \bar{d}_{ij}^{(1)}) \bar{c}_{j\bar{y}}^{(1)} - \bar{k}^{(1)} j_c^{(0)} / k^{(0)} \right] d\bar{y}' / k^{(0)}, \end{aligned}$$

where $\bar{\varphi}^{(2)}$ is an arbitrary function of x and eq. (68) has been used to eliminate $\bar{\varphi}^{(1)}$. We wish to calculate $\bar{\varphi}^{(2)}$ in the limit $\bar{y} = \epsilon^{\alpha-1} y^*$, with $0 < \alpha < 1$, y^* fixed, $\epsilon \rightarrow 0$. The first three terms all contribute to the limit, while the fourth term vanishes, since $\bar{c}^{(2)} \rightarrow 0$, as $\bar{y} \rightarrow \infty$. Finally, we must evaluate the limit of the integral. Consider the last term first. Since $j_c^{(0)}$ is a function of x only and

$$\bar{k}^{(1)} = \sum_{p=1}^n (\partial k^{(0)} / \partial c_p) \bar{c}_p^{(1)},$$

we have n integrals like

$$\int_0^{\bar{y}} \bar{c}_p^{(1)}(x, \bar{y}') d\bar{y}'.$$

Changing variable gives

$$\int_0^{\epsilon^{\alpha-1} y^*} \bar{c}_p^{(1)}(x, \bar{y}) d\bar{y} = \epsilon^{\alpha-1} \int_0^{y^*} \bar{c}_p^{(1)}(x, \epsilon^{\alpha-1} z) dz \rightarrow \epsilon^{\alpha-1} I(\epsilon),$$

where $I(\epsilon) \rightarrow 0$, as $\epsilon \rightarrow 0$. Similar limits are obtained for the other integrals with integrands of the form $\bar{c}_p^{(1)} \bar{c}_{j\bar{y}}^{(1)}$. Thus,

$$\begin{aligned} \bar{\varphi} &\sim \bar{\varphi}^{(0)}(x, 0) + \epsilon [\bar{\varphi}^{(1)}(x, 0) + \epsilon^{\alpha-1} y^* j_c^{(0)} / k^{(0)}] + \epsilon^2 [\bar{\varphi}^{(2)}(x) \\ &+ \epsilon^{\alpha-1} y^* \bar{j}_c^{(1)} / k^{(0)} - (\epsilon^{\alpha-1} y^*)^2 \bar{\varphi}_{xx}^{(0)}(x, 0) / 2], \end{aligned}$$

since $\bar{\varphi}_{xx}^{(0)} = \bar{\varphi}_{xx}^{(0)}(x, 0)$, to be matched with

$$\bar{\varphi} \sim \bar{\varphi}^{(0)}(x, 0) + \epsilon^{\alpha} y^* \bar{\varphi}_y^{(0)}(x, 0) + (\epsilon^{\alpha} y^*)^2 \bar{\varphi}_{yy}^{(0)}(x, 0)/2 + \epsilon [\bar{\varphi}^{(1)}(x, 0) + \epsilon^{\alpha} y^* \bar{\varphi}_y^{(1)}(x, 0)] + \epsilon^2 \bar{\varphi}^{(2)}(x, 0).$$

Since $\bar{\varphi}^{(0)}$ is harmonic, $\bar{\varphi}_{xx}^{(0)}(x, 0) = -\bar{\varphi}_{yy}^{(0)}(x, 0)$, and matching term by term gives the previous boundary condition on $\bar{\varphi}^{(0)}$

$$\bar{\varphi}_y^{(0)}(x, 0) = j_c^{(0)}/k^{(0)},$$

the boundary condition on $\bar{\varphi}^{(1)}$

$$\bar{\varphi}_y^{(1)}(x, 0) = \tilde{j}_c^{(1)}/k^{(0)},$$

and the "penetration" condition

$$\hat{\varphi}^{(2)}(x) = \bar{\varphi}^{(2)}(x, 0).$$

APPENDIX C

The Electrochemical Potential

In Section IV, we expressed the fluxes in terms of the gradients in the form

$$\mathbf{N}_i = C_i \mathbf{V} - \sum_{j=1}^n D_{ij} \nabla C_j - F z_i C_i U_i \nabla \Phi, \quad (71)$$

in effect *defining* the diffusivities D_{ij} and mobilities U_i , functions of the concentrations, by this relation. In electrochemistry, it is more usual to give the gradients of the electrochemical potentials, i.e., well-defined thermodynamic functions, in terms of the fluxes. Thus, Newman¹ gives the relations (for $n + 1$ species in our notation)

$$c_i \nabla \mu_i = \sum_{j=0}^n K_{ij} (\mathbf{V}_j - \mathbf{V}_i), \quad (72)$$

for $i = 0, 1, \dots, n$, where the μ_i 's are the electrochemical potentials, the K_{ij} 's are the interaction coefficients, functions of the concentrations, and \mathbf{V}_i 's are subject to certain constraints. Thus,

$$\sum_{i=0}^n C_i \nabla \mu_i = 0$$

and

$$\sum_{i=0}^n \rho_i \mathbf{V}_i = \rho \mathbf{V},$$

where \mathbf{V} is the fluid velocity, $\rho_i = M_i C_i$ is the partial density for molecular weight M_i , and

$$\rho = \sum_{i=0}^n \rho_i.$$

These restrictions imply that

$$\sum_{i=0}^n (K_{ij} - K_{ji}) = 0,$$

for $j = 0, 1, \dots, n$. In particular, Newman assumes the Stefan-Maxwell relations $K_{ij} = K_{ji}$, i.e., pairwise interaction between species. With these restrictions, eq. (72) has a unique solution for $\mathbf{V}_i = \mathbf{V}$. Thus we find, following Newman, that

$$\mathbf{V}_i - \mathbf{V}_0 = - \sum_{p=1}^n L_{ip} C_p \nabla \mu_p,$$

for $i = 1, \dots, n$, where $-[L_{ij}]$ is the inverse of $[M_{ij}]$, with

$$M_{ij} = \begin{cases} - \sum_{k=1 \neq i}^n K_{ik}, & \text{for } i = j, \\ K_{ij}, & \text{for } i \neq j. \end{cases}$$

(Note that we may assume that $K_{ij} = 0$, for $i = j$.) Once $\mathbf{V}_i - \mathbf{V}_0$ is calculated, $\mathbf{V}_i - \mathbf{V}$ may be determined from the identities

$$\mathbf{V}_0 - \mathbf{V} = - \sum_{k=1}^n (\rho_k/\rho) (\mathbf{V}_k - \mathbf{V}_0),$$

$$\mathbf{V}_i - \mathbf{V} = \mathbf{V}_i - \mathbf{V}_0 + \mathbf{V}_0 - \mathbf{V}.$$

With

$$\nabla \mu_p = \sum_{j=0}^n (\partial \mu_p / \partial C_j) \nabla C_j + (\partial \mu_p / \partial \Phi) \nabla \Phi,$$

we finally find

$$D_{0j} = - \sum_{k=1}^n \sum_{p=1}^n (\rho_k/\rho) L_{kp} C_0 C_p (\partial \mu_p / \partial C_j),$$

$$Fz_0 U_0 = - \sum_{k=1}^n \sum_{p=1}^n (\rho_k/\rho) L_{kp} C_p (\partial \mu_p / \partial \Phi),$$

$$D_{ij} = \sum_{p=1}^n \left[L_{ip} - \sum_{k=1}^n (\rho_k/\rho) L_{kp} \right] C_i C_p (\partial \mu_p / \partial C_j),$$

$$Fz_i U_i = \sum_{p=1}^n \left[L_{ip} - \sum_{k=1}^n (\rho_k/\rho) L_{kp} \right] C_p (\partial \mu_p / \partial \Phi).$$

REFERENCES

1. J. S. Newman, *Electrochemical Systems*, 1st ed., New York: Prentice-Hall, 1973.
2. J. L. Blue, unpublished work.
3. V. G. Levich, *Physicochemical Hydrodynamics*, 1st ed., Prentice-Hall, 1962.
4. M. J. Lighthill, "Contributions to the Theory of Heat Transfer Through a Laminar Boundary Layer," Proc. Roy. Soc., *A202* (1950), pp. 359-377.

Contributors to This Issue

Philip Balaban, Dipl. Ing., 1950, Technical University, Munich; Ph.D., 1966, Polytechnic Institute of Brooklyn; Scientific Dept., Israel Ministry of Defense, 1950–1956; Contraves AG, Zurich, 1956–1958; Computer Systems, Inc., 1958–1963; Bell Laboratories, 1967—. Mr. Balaban has been engaged in research and development in the areas of analog and digital computers, circuit design, and communication systems. From 1966–1967 he was a member of the faculty of the Electrical Engineering Department at the Polytechnic Institute of Brooklyn, and he spent the year 1970 as a Visiting Professor at the Technion, Israel Institute of Technology. His current work involves the development and application of computer aided analysis techniques to the design of transmission systems. Senior Member, IEEE; Member, Sigma Xi.

James G. Evans, B.E.E., 1963, M.E., 1964, Cornell University; E.E., 1973, Ph.D. (E.E.), 1974, New York University; Bell Laboratories, 1963—. Mr. Evans has been engaged in the design and development of computer-controlled measuring systems. He is currently working in the areas of nonlinear circuit characterization and digital circuit fault location. He is a supervisor in the Measuring Systems Design Department. Member, IEEE, Eta Kappa Nu, Tau Beta Pi.

Franklin W. Kerfoot, III, B.S.E.E., 1967, Rensselaer Polytechnic Institute; M.S.E.E., 1969, Polytechnic Institute of Brooklyn; Bell Laboratories, 1967—. Mr. Kerfoot has been involved in the design and development of computer-controlled analog instrumentation to support transmission systems development. He is currently developing automatic procedures to locate faults in digital circuit boards. Member, Eta Kappa Nu.

John A. Lewis, B.S., 1944, Worcester Polytechnic Institute; M.S., 1948, and Ph.D., 1950, Brown University; Bell Laboratories, 1951—. Mr. Lewis has worked on problems in piezoelectricity, heat conduction, and electroplating. He is currently concerned with optical fiber drawing. Member, American Mathematical Society, Society for Industrial and Applied Mathematics, Mathematical Association of America.

Dietrich Marcuse, Diplom Vorpruefung, 1952, Dipl. Phys., 1954, Berlin Free University; D.E.E., 1962, Technische Hochschule, Karls-

ruhe, Germany; Siemens and Halske (Germany), 1954–1957; Bell Laboratories, 1957—. At Siemens and Halske, Mr. Marcuse was engaged in transmission research and studying coaxial cable and circular waveguide transmission. At Bell Laboratories, he has been engaged in studies of circular electric waveguides and work on gaseous masers. He spent one year (1966–1967) on leave of absence from Bell Laboratories at the University of Utah. He is presently working on the transmission aspect of a light communications system. Mr. Marcuse is the author of three books. Fellow, IEEE; member, Optical Society of America.

Robert L. Nichols, RCA Institutes, Inc., 1951; Bell Laboratories, 1952—. Mr. Nichols has been engaged in the design and development of automated analog instrumentation to support transmission systems development. He is currently developing computer-operated instrumentation to characterize the nonlinear properties of networks.

V. Ramaswamy, B.Sc. (Physics), 1957, Madras University, India; D.M.I.T. (Electronics), 1960, Madras Institute of Technology, Chromepet, Madras, India; M.S., 1962, and Ph.D., 1969, Northwestern University; Zenith Radio Corporation, 1962–1965; Bell Laboratories, 1969—. Mr. Ramaswamy's work has included microwave components, diode parametric amplifiers, and wave propagation in semiconductor plasmas. At Bell Laboratories, he is engaged in research related to thin film optical waveguides and devices. Member, Sigma Xi, IEEE.

Aaron E. Rosenberg, S.B. (E.E.), and S.M. (E.E.), 1960, Massachusetts Institute of Technology; Ph.D. (E.E.), 1964, University of Pennsylvania; Bell Laboratories, 1964—. Mr. Rosenberg is presently engaged in studies of systems for man-machine communication-by-voice in the Acoustics Research Department. Member, IEEE, Eta Kappa Nu, Tau Beta Pi, Sigma Xi, IEEE Acoustics, Speech and Signal Processing Group Technical Committee on Speech Processing; Fellow, Acoustical Society of America.

R. D. Standley, B.S., 1957, University of Illinois; M.S., 1960, Rutgers University; Ph.D., 1966, Illinois Institute of Technology; USASRD, Ft. Monmouth, N.J., 1957–1960; IIT Research Institute, Chicago, 1960–1966; Bell Laboratories, 1966—. Mr. Standley has been engaged in research projects involving microwave, millimeter wave, and optical components. He is presently concerned with electron beam lithography as applied to fabrication of integrated optic devices. Member, IEEE, Sigma Tau, Sigma Xi.

Abstracts of Papers by Bell System Authors Published in Other Journals

CHEMISTRY

Amorphous Ferromagnetic Ag-X (X = Ni, Co, Gd) Alloys. J. J. Hauser, *Phys. Rev. B*, *12* (December 1975), pp. 5160-5165. The critical concentration for the appearance of ferromagnetism (41-at. % Ni) and the increase of T_c with Ni concentration ($\approx 9^\circ\text{K/at.}\%$) for amorphous Ag-X (X = Ni, Co, Gd) films are quite close to the values reported for crystalline Ni-Cu alloys, suggesting again the validity of the virtual-bound-state model over the rigid-band model.

Field Evaporation Experiments With a Magnetic Sector Atom-Probe FIM. T. Sakurai and E. W. Müller*, *Surface Sci.*, *50* (May 1975), pp. 38-44. Field evaporation of Rh in the presence of ^3He and ^4He gases revealed that the formation of the helium compound $(\text{RhHe})^{2+}$ is quite sensitive to He gas pressure; no helium compounds were observed below 5×10^{-7} Torr and all field-evaporated as helium compounds above 5×10^{-6} Torr at 78 K. * Pennsylvania State University.

The Growth of LiNbO_3 Thin Films by Liquid Phase Epitaxial Techniques. A. A. Ballman, H. Brown, and P. K. Tien, *J. Cryst. Grow.*, *29* (May 1975), pp. 289-295. A liquid phase epitaxial dipping process has been developed for the growth of thin films of LiNbO_3 on isostructural LiTaO_3 substrates. The process, employing a Li_2WO_4 flux, produces single crystalline layers from 1 to 10 μm in thickness. Low-loss optical waveguides have been fabricated from these films which show a sharp step-change of index of refraction with the Li-TaO₃ substrate.

MOS (Si-Gate) Compatibility of RF Diode and Triode Sputtering Processes. A. K. Sinha, *J. Electrochem. Soc.*, *123* (January 1976), pp. 65-71. The radiation damage induced in the gate regions of CMOS capacitors by various thin-film sputtering processes has been evaluated. It is shown that trapped hydrogen can be used for annealing interface states in sealed MOS structures, which are impervious to hydrogen during post-damage annealing.

Photoreactions of Ketones With Amines. CIDNP Criteria for the Intermediacy of Aminoalkyl Radicals and Aminium Radical Ions. H. D. Roth and M. L. Manion, *J. Amer. Chem. Soc.*, *97* (November 1975), pp. 6886-6888. The photoreactions of aromatic ketones with tertiary amines result in nuclear spin polarization and nmr line broadening. The intensity pattern of the polarization and the degree of line broadening depend upon the hyperfine coupling patterns of the paramagnetic intermediates and may distinguish between neutral aminoalkyl radicals and aminium radical ions resulting from net hydrogen abstraction and electron transfer, respectively.

Preparation of $\text{BaTi}_5\text{O}_{11}$ by Solid-State Reaction. H. M. O'Bryan, Jr. and J. Thomson, Jr., *J. Amer. Ceram. Soc.*, *58* (September-October 1975), p. 454. $\text{BaTi}_5\text{O}_{11}$ forms as a reaction intermediate when a mixture of $\text{BaCO}_3 + 4.5 \text{TiO}_2$ is heated at 1000 to 1100°C. Powder diffraction data are presented for $\text{BaTi}_5\text{O}_{11}$. The conditions for forming this compound show similarities with those required to form the unstable BaTi_2O_5 compound.

Thermal Stability of PtSi Films on Polysilicon Layers. A. K. Sinha, S. E. Haszko, and T. T. Sheng, *J. Electrochem. Soc.*, *122* (December 1975), pp. 1714-1718. Films containing PtSi (2000 Å) on poly-Si (1000 Å) have been considered for application as first-level metallization in bipolar LSI devices. The microstructure, resistivity, and stresses of these films were evaluated for various thermal treatments. No agglomeration was observed at up to 900°C; at 1000°C, a liquid-eutectic phase formed.

ELECTRICAL AND ELECTRONIC ENGINEERING

Aberrations and Tolerances in a Double-Deflection Electron Beam Scanning System. M. G. R. Thomson, *J. Vac. Sci. Technol.*, *12* (November–December 1975), pp. 1156–1159. A magnetic double-deflection system intended for use in an electron-beam lithography instrument has been theoretically analyzed. The first-order chromatic and third-order geometric aberrations are calculated so that the maximum field size can be found. The misalignment tolerance is estimated for different levels of required performance.

An Annotated Bibliography of Microwave Circulators and Isolators 1968–1975. R. H. Knerr, *IEEE Trans. Microw. Theory Tech.*, *MTT-23* (October 1975), pp. 818–825. A bibliography of microwave circulators and isolators from 1968 to 1975 is presented (151 references). Some topics, which the author considered the most significant, include the eigenvalue approach to 3-port Y-junction circulator design, the peripheral mode isolator, electronic isolators, and wideband UHF circulators. The bibliography is subdivided into waveguide circulators, stripline circulators, lumped element circulators, other types of circulators, general papers applicable to all circulators, isolators, and books.

Deep Level Spectroscopy, Low Temperature Defect Motion and Nonradiative Recombination in GaAs and GaP. C. H. Henry, *J. Electron. Mater.*, *4* (October 1975), pp. 1037–1052). Recent development in junction capacitance measurements allow deep levels in semiconductors to be conveniently studied for the first time. The levels are often strongly coupled to the lattice which causes rapid nonradiative recombination. The released electronic energy causes violent vibrations near the defect which can promote low-temperature defect motion.

Hopping Conductivity in Amorphous Carbon Films. J. J. Hauser, *Solid State Commun.*, *17* (December 1975), pp. 1577–1580. The electrical resistivity of amorphous carbon films getter-sputtered at 95°K is well fitted between 300 and 20°K by the relation $\rho = \rho_0 \exp[(T_0/T)^{1/2}]$ with $T_0 \approx 7 \times 10^7$ K. This behavior suggests a hopping conductivity very similar to that found in other amorphous semiconductors.

Mass-Spectrometric Sampling of the Ionic and Neutral Species Present in Different Regions of an RF Discharge in Methane. M. J. Vasile and G. Smolinsky, *Int. J. Mass. Spectrom. Ion Phys.*, *18* (October 1975), pp. 179–192. Ions originating in the sheath between the RF electrode and the plasma are characteristic of higher energy ionization processes than those observed from the plasma itself. Ion-molecule reactions observed in the RF sheath indicate a substantial proportion of encounters which occur at higher than thermal kinetic energy.

A Per-Channel A/D Converter Having 15-Segment μ -255 Companding. J. C. Candy, W. H. Ninke, and B. A. Wooley, *IEEE Trans. Commun.*, *COM-24* (January 1976), pp. 33–42. This encoder provides 8-bit compressed PCM and 13-bit uniform PCM at 8 kilowords per second. Also available is a 1-bit code that is useful for loop and switching applications. When suitably decoded the processed signal exhibits idle channel noise below 15 dBmCo, harmonic distortion less than -40 dB, and gain tracking within ± 0.5 dB for circuit tolerances of 2%. The encoder is being integrated on two chips, and a decoder on a single chip.

PHYSICS

Adiabatic Following Model for Two-Photon Transitions: Nonlinear Mixing and Pulse Propagation. D. Grischkowsky,* M. M. T. Loy,* and P. F. Liao, *Phys. Rev. A*, *12* (December 1975), pp. 2514–2533. The adiabatic following approximation is used to obtain solutions to a vector model describing a two-photon resonance. Explicit expressions are obtained for power-dependent nonlinear susceptibilities which are responsible for two-photon, resonantly enhanced, parametric processes. The results are also used to study the propagation of pulses nearly resonant with a two-photon transition. * IBM Thomas J. Watson Research Center.

Impact Ionization Rates for Electrons and Holes in GaAs_{1-x}Sb_x Alloys. T. P. Pearsall, R. E. Nahory, and M. A. Pollack, *Appl. Phys. Lett.*, *28* (April 1, 1976), pp. 403-405. We have measured the impact ionization rates for electrons and holes in GaAs_{1-x}Sb_x for alloys with $x = 0.05, 0.10, \text{ and } 0.12$. The ionization rates for electrons (α) and holes (β) are different and a function of Sb content. For GaAs_{0.88}Sb_{0.12}, $\alpha/\beta = 2.5$, while for GaAs, $\alpha/\beta = 0.25$.

Ion-Ion Correlations and Diffusion in Beta Alumina. D. B. McWhan et al., *Phys. Rev. Lett.*, *35* (October 6, 1975), pp. 953-956 and *36* (February 9, 1976), p. 344(e). A comparison of the correlation length for short-range order (5, 6, 11, 18 Å), the attempt frequency for diffusion (56, 28, 80, 85 cm⁻¹), and the previously reported values of tracer diffusion coefficient for Na, Ag, K, and Rb, respectively, suggests that ion-ion correlations inhibit diffusion in the β -aluminas.

Millimeter and Submillimeter Wave Laser Action in Symmetric Top Molecules Optically Pumped Via Perpendicular Absorption Bands. T. Y. Chang and J. D. McGee, *IEEE J. Quantum Electron.*, *QE-12* (January 1976), pp. 62-65. Ninety-nine new far infrared laser lines from 227 μm to 1.965 mm have been observed in CH₃CN, CH₃CCH, CH₃Cl, CH₃Br, and CH₃I by optically pumping these gases with CO₂ laser pulses of 150- μs duration.

Motion of pn Junctions in CuInSe₂. B. Tell, Sigurd Wagner, and P. M. Bridenbaugh, *Appl. Phys. Lett.*, *28* (April 15, 1976), pp. 454-455. We report the first pn junction delineation and diffusion study in a ternary chalcopyrite-type semiconductor. Pn junctions were formed in Zn or Cd plated p-CuInSe₂ by five-minute anneals at 200 to 450°C. By angle lapping and staining techniques, junction depths x_j varying from $\sim 1 \mu\text{m}$ to $\sim 130 \mu\text{m}$ were determined, from which the interdiffusion coefficient derived is $D(\text{cm}^2/\text{s}) = 164 \exp[-1.19(\text{eV})/kT]$. The large preexponential term indicates concentrations or mobilities of point defects substantially above those of the related II-VI compounds.

Novel Technique for Measuring Nitrogen Profiles in GaP:N. J. Shah, R. F. Leheny, and P. D. Dapkus, *J. Appl. Phys.*, *46* (December 1975), pp. 5244-5247. A high-resolution technique for determining the nitrogen-doping profiles in GaP:N using a laser-pumped dye cell as a broad-band fluorescence source is described. Results obtained suggest that, under some LPE growth conditions, vapor-liquid equilibration may be slow and, as a result, affect the nitrogen incorporated into the solid. The implications of these results for high-efficiency LED's are also discussed.

Observation of a New Nonlinear Photoelectric Effect Using Optical Levitation. A. Ashkin and J. M. Dziedzic, *Phys. Rev. Lett.*, *36* (February 2, 1976), pp. 267-270. We observe a new three-photon nonlinear photoelectric effect in glass with visible cw laser light using optical levitation of transparent particles. Electrons are pumped into the conduction band of glass by the absorption of two photons and subsequently ejected in a single photon step.

Polarization Rotation Induced by Resonant Two-Photon Dispersion. P. F. Liao and G. C. Bjorklund, *Phys. Rev. Lett.*, *36* (March 15, 1976), pp. 584-587. A polarization rotator is demonstrated which utilizes the dispersion associated with 3S to 5S two-photon transitions in sodium vapor. A linearly polarized beam at ν_1 is rotated by a circularly polarized beam at ν_2 , and $\nu_1 + \nu_2$ is near the two-photon transition frequency. When combined with a polarizer, a rapid optical shutter is obtained.

Preparation and Properties of InP/CdS and CuInSe₂/CdS Solar Cells. J. L. Shay et al., *Proc. 11th IEEE Photovoltaic Specialists Conf.*, 1975 (Scottsdale, Arizona, May 6-8), pp. 503-507. Described are the preparation and properties of the recently reported InP/CdS and CuInSe₂/CdS single-crystal solar cells. Solar power conversion efficiencies of 12.5 and 12% have been observed. The ultimate efficiency achievable with InP/CdS using state-of-the-art liquid encapsulated Czochralski p-type InP is calculated to be 17.2% for AM2 conditions and 14.0% for AMO conditions.

Push-rod AC Susceptibility Apparatus. J. J. Hauser and C. M. Antosh, *Rev. Sci. Instrum.*, *47* (January 1976), pp. 156-157. A push-rod ac susceptibility apparatus which permits the measurement of ultrathin films of superconductors and ferromagnets as a function of temperature and magnetic field is described. Films prepared at 77° K can be mounted under liquid nitrogen and measured without warmup.

Water and Its Relation to Broken Bond Defects in Fused Silica. R. H. Stolen and G. E. Walrafen, **J. Chem. Phys.* *64* (March 15, 1976), pp. 2623-2631. Vibrational modes associated with OH in fused silica were studied by Raman scattering. Intensities of lines at 604 and 490 cm^{-1} decrease with increasing OH and increase with fictive temperature, suggesting that OH is trapped at defect sites. A weak 970 cm^{-1} line was seen in fibers and assigned to the Si-(OH) stretch. * Dept. of Chem., Howard University.

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*, *Computer & Control Abstracts*, *Current Contents/Engineering, Technology & Applied Sciences*, *Current Contents/Physical & Chemical Sciences*, *Current Index to Statistics*, *Current Papers in Electrical & Electronic Engineering*, *Current Papers on Computers & Control*, *Electrical & Electronic Abstracts*, *Electronics & Communications Abstracts Journal*, *The Engineering Index*, *International Aerospace Abstracts*, *Journal of Current Laser Abstracts*, *Language and Language Behavior Abstracts*, *Mathematical Reviews*, *Metals Abstracts*, *Science Abstracts*, *Science Citation Index*, 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.



Bell System