Beam-Lead Technology

By M. P. LEPSELTER

A process has been developed to batch-fabricate beam-lead transistors, integrated circuits, and other components, where the leads serve a structural and protective as well as electrical function. Platinum silicide ohmic contacts, titanium and platinum sputtered layers, and electroformed gold beam leads constitute the metallurgical structure of the devices described. Test transistors have survived 350°C aging for hundreds of hours in corrosive ambients, and centrifuging at 135,000 g's.

I. INTRODUCTION

This paper describes a process developed to batch-fabricate semiconductor devices and integrated circuits with electroformed electrodes cantilevered beyond the edges of the wafer — hence, the name beam leads.[1] This type of structure simplifies the assembly and interconnection of individual units and integrated circuits, provides its own protective seal, and leads to a new class of integrated circuits[2] where the isolation is accomplished by etched trenches under the metal bridging connections.

Fig. 1 is a drawing of a silicon high-frequency beam-lead transistor. The 0.5-mil thick leads are cantilevered beyond the edge of the silicon chip, and are used for structural support of the chip as well as electrical contact. The beam leads may be bonded to a metal-filmed substrate with a matching pattern, obviating the need for eutectic brazing to the substrate. This structure imposes no electrical penalty on the
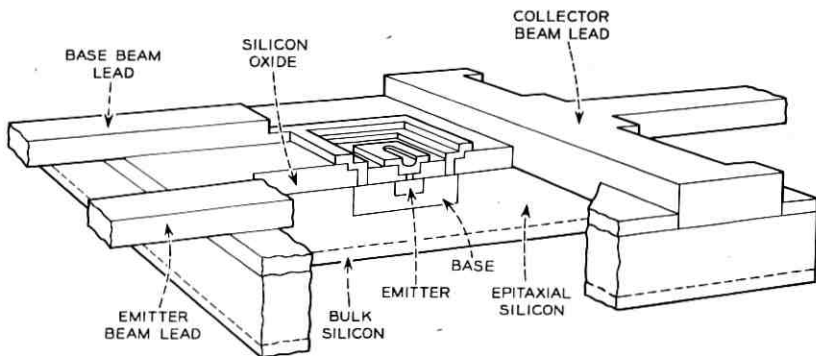
Fig. 1 — Beam-lead transistor structure.

device, the parasitic capacitance being at most equivalent to a conventional planar device.

Fig. 2 is a photo of an actual transistor bonded onto a metallized ceramic substrate. The diffused regions are facing the ceramic.

It can be seen that there are no external wires or brazed areas, as previously mentioned. This device structure has been centrifuged, without failure, to an acceleration level of over 135,000 g's on a test substrate.

Fig. 3 is a photograph of an array of transistors on a slice. There are three different electrode geometries which are used in conjunction with aging experiments. The spacing is approximately 15 mils center-to-center, with every other row missing to allow for test patterns. This spacing leads to a significantly higher packing density than is practical with standard scribe-and-break techniques.

Fig. 4 is a face-view photograph of a beam-lead transistor. The silicon chip is approximately 7-mils square, and the beam leads 1-mil wide. There is an approximate 0.2-mil clearance between the electrode fingers, which are thinner than the beam leads. Thus, the fingers can have the close spacing required for high-frequency operation without carrying the restriction imposed by the required beam thickness.

Basically, the process to be described starts with a standard slice of planar devices after the contact holes have been etched through the oxide. Platinum silicide ohmic contacts are formed in the contact holes; titanium and platinum layers are sputtered onto the silicon oxide; and the gold beam leads are electroformed, using the platinum layer as a base. The excess platinum outside the gold patterns is removed by glow-discharge etching, and the titanium is etched away. At this point, the
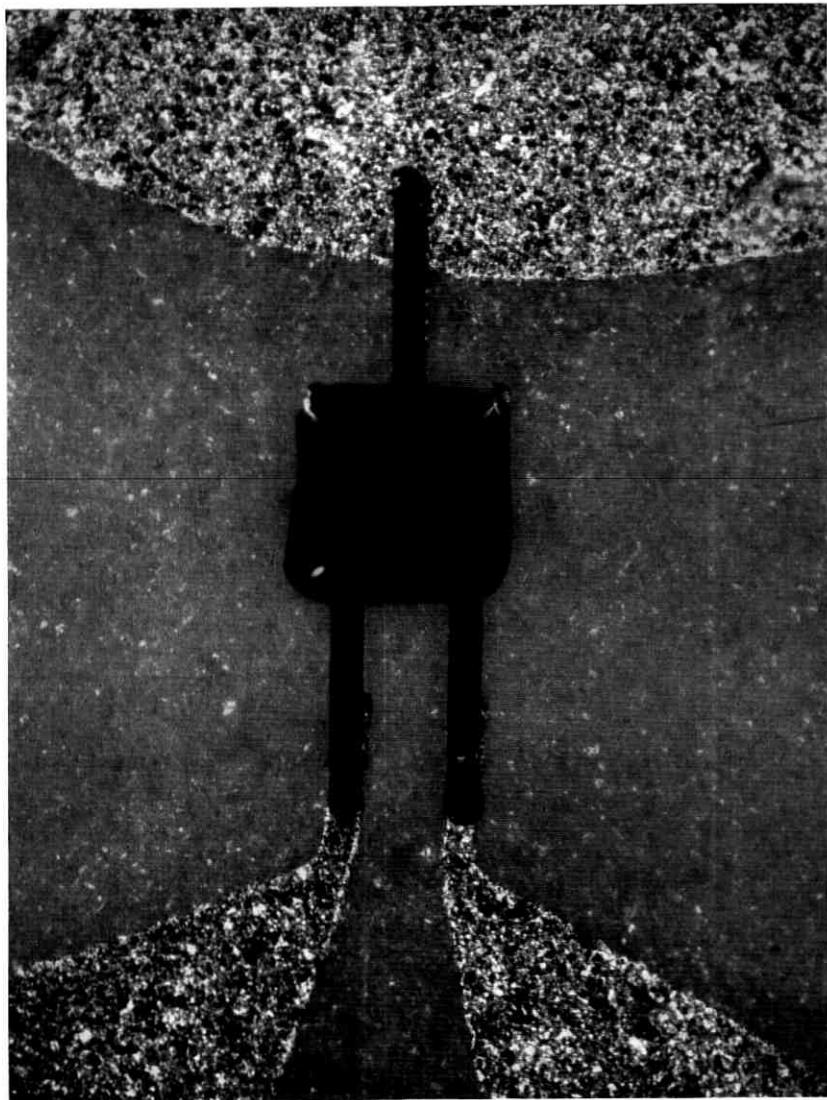
Fig. 2 — Bonded beam-lead transistor.

slice is turned over and etch-masking patterns developed in registry with the metallized patterns on the other side. The unmasked areas are etched away, leaving the individual devices with the beam leads cantilevered beyond the edges of the slice.
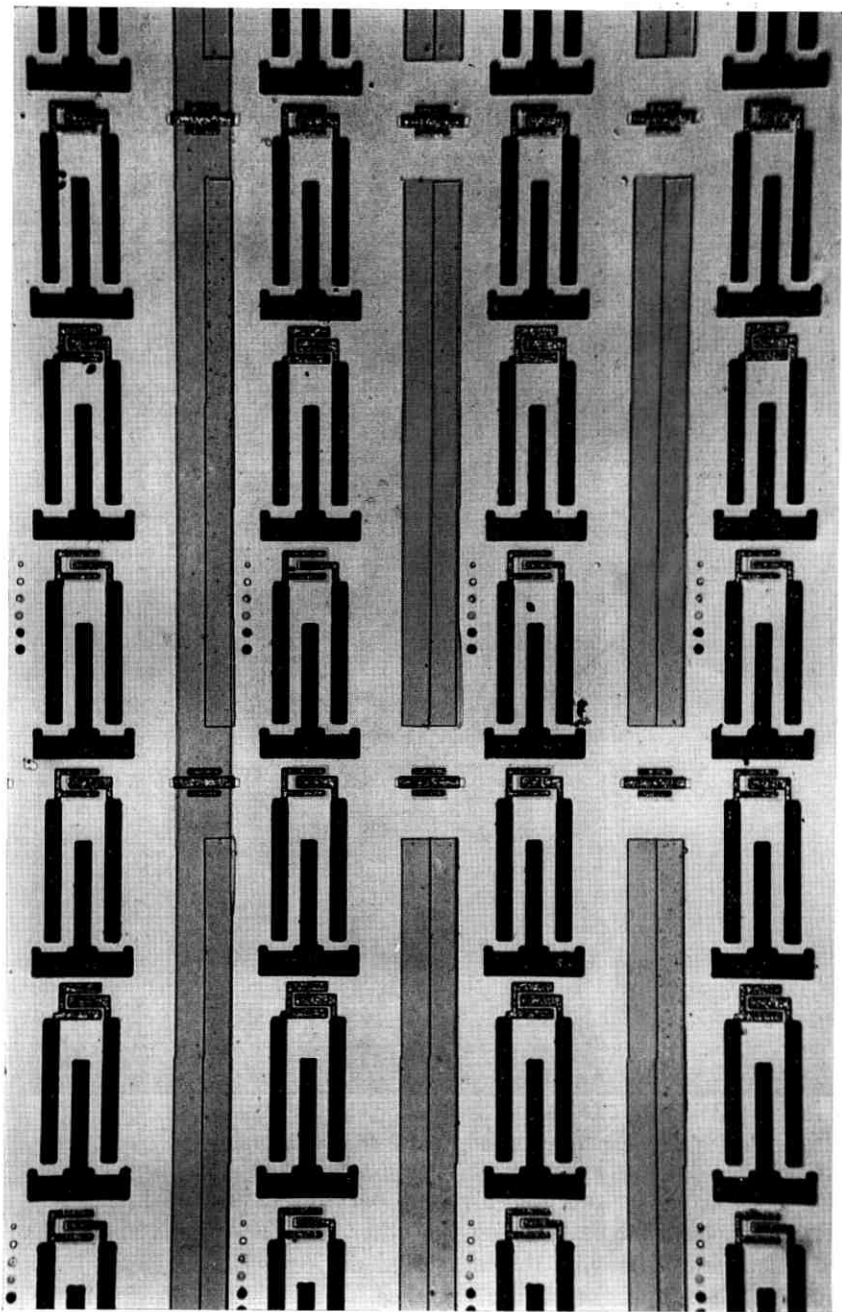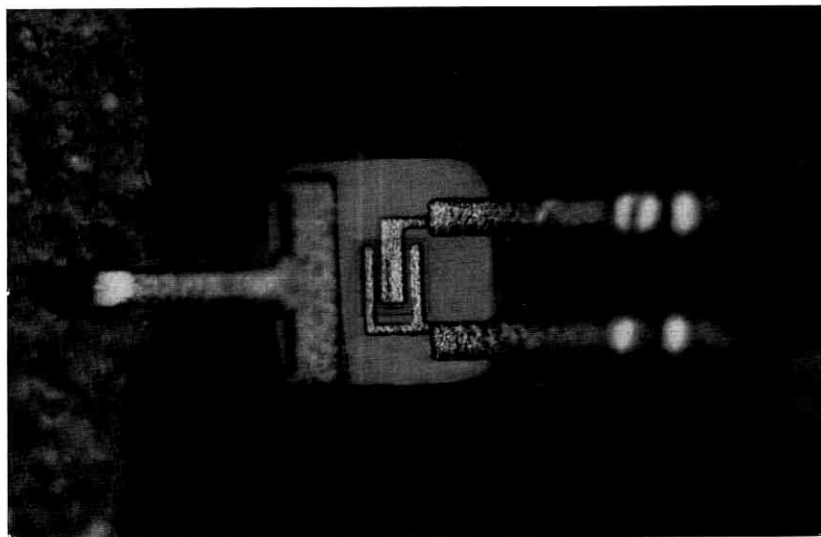
Fig. 3 — Array of transistors.

Fig. 4 — Face-view of transistor.

With no additional processing, the previous etching operation may be used to cut isolation trenches in integrated circuits, replacing isolation diffusions or solid-dielectric isolation.

The objective of this project was to produce a semiconductor device structure that would improve the present methods of device assembly, decrease fabrication complexity, yield electrical parameter stability without requiring a separate vacuum-tight enclosure, and decrease integrated circuit parasitic capacitance.

In Section II the beam-lead structure and its applications are discussed; Section III describes a process sequence developed to fabricate devices and integrated circuits; and Section IV is devoted to test results on devices fabricated with the techniques described in Section III.

II. BEAM-LEAD STRUCTURE AND APPLICATIONS

2.1 *Structure*

The beam-lead structure eliminates the need for chip brazing and external wire bonding. With the beam-lead device facing the substrate, the cantilevered leads extending beyond the edges of the chip are

readily aligned with matching patterns on the substrate, and the beam leads bonded to these patterns. The beam leads are used as the alignment guide: with multiple leads extending beyond the chip locating one lead automatically registers the rest, since both the beam-lead array and the substrate pattern are precisely oriented with respect to each other during the photolithographic operations. The surfaces to be joined to each other are both annealed gold, and thermocompression bonding or split-tip welding, among other techniques, may be used. A gold-gold bond is a very reliable metallurgical system; there are no oxide films to hinder the bonding, gold is one of the most ductile metals, thermocompression bonding occurs at low temperatures (300°C), and the joint formed is free of attack by oxidation, galvanic corrosion, or other corrosive media.

An extension of the beam-lead concept to multiple devices is shown by Fig. 5, a photograph of a string of beam-lead transistors fabricated with common rails connecting all the collectors in a row. This method allows devices to be simply handled in large numbers for feeding, electrical testing, and bonding. As required, one device is bonded to a substrate, and broken from the string. Since the string is initially fabricated as an array on the silicon slice, all dimensions are held to close tolerances, and any lead is precisely oriented with respect to any other point in the string, typically within 0.1 mil. In addition, the device need never be handled by the silicon chip. The gold leads and supporting ribs absorb the handling stresses and possible contamination.
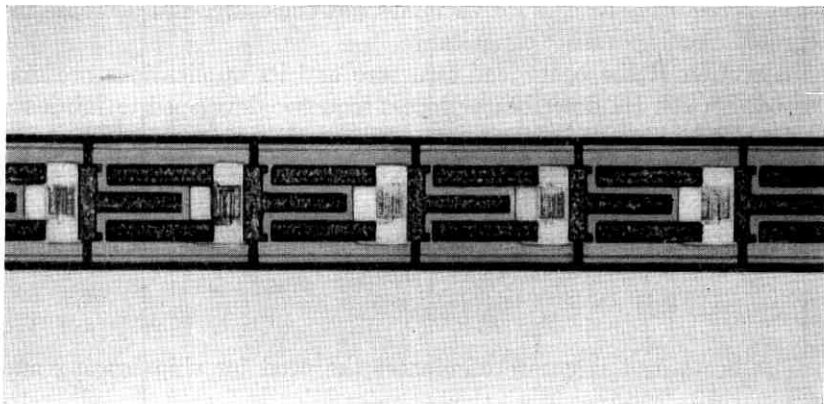
Fig. 5 — String of transistors.

2.2 *Applications*

An early application of beam-lead transistors to discrete-chip integrated circuits is seen in Fig. 6, which is a photograph of two DCTL gates with three and four inputs. There are seven high-frequency silicon transistors (3 leads), and nine beam-lead boron-diffused silicon resistors. Several circuits wired as a 3-stage ring oscillator, with a fan-in of 3-4, and a fan-out of 1, yielded a measured propagation delay of 4.2

Fig. 6 — Discrete-chip integrated circuit.

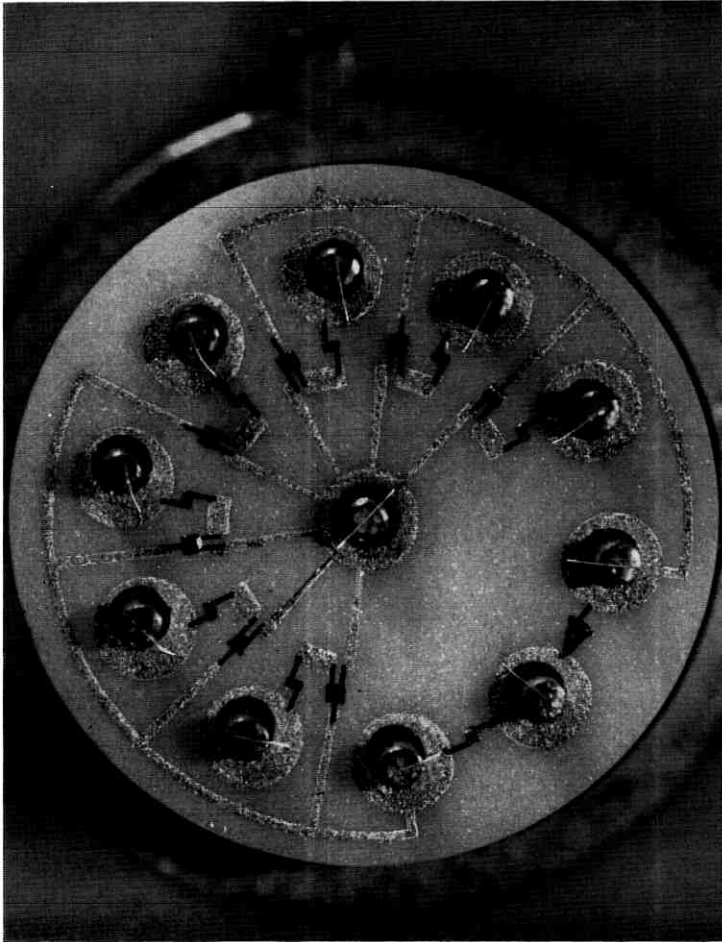ns. Measurements were made with a collector supply voltage of 5 volts, corresponding to an average power per gate of 23 mw. This is equivalent in performance to circuits fabricated with standard chip and wire construction.

Fig. 7 is an example of beam-lead transistors appliqued to a tantalum thin-film circuit, a 3-2 input gate DCTL switching circuit. With this structure, it is not necessary to braze the silicon chips to the substrate, subjecting precise resistors to the bonding temperature required for eutectic brazing. Ultrasonic or split-tip resistance welding may be employed to attach the beam-lead devices to the metallized
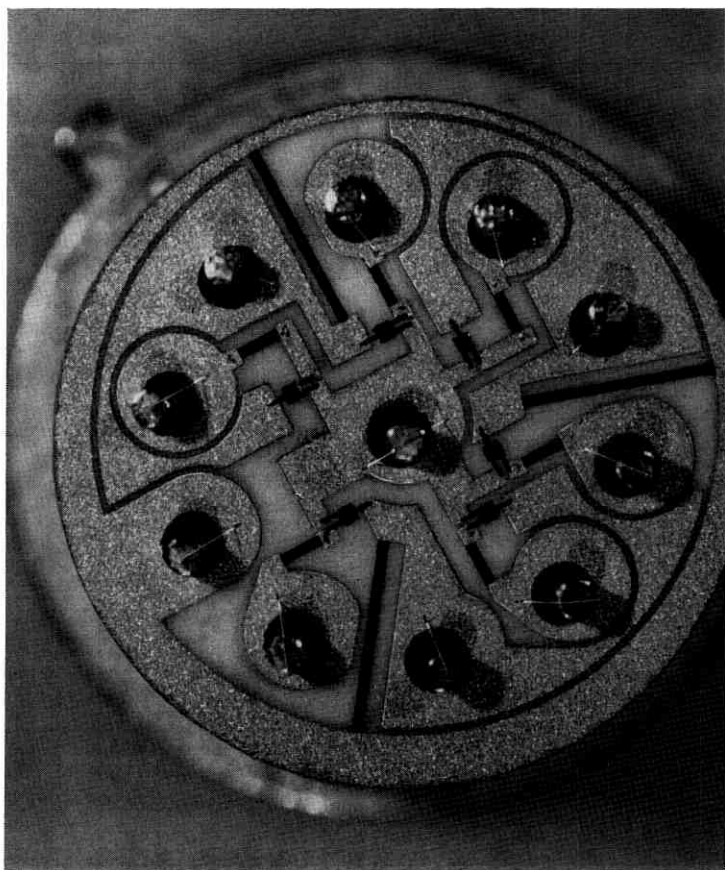

Fig. 7 — Hybrid thin-film digital circuit.

substrate. A TO-5 header and pin-connecting wires are used in this circuit for convenience only.

A specific application of beam-lead techniques to monolithic integrated circuits was investigated by H. A. Waggener.[2] As a consequence of etching away the unwanted silicon from under the beam leads, isolated pads of silicon may be attained, interconnected by the beam leads. The only capacitive coupling is then through the small metal-over-oxide overlay (0.05 pf typical). This is many times lower than the junction capacitance incurred with p-n junction isolation monolithic circuits. Ultrahigh-speed switching circuits are possible with this technique, with no increase in processing complexity. The isolation trenches are automatically formed during the silicon etching.

Fig. 8 is a physical-electrical schematic of a 4-input DCTL logic gate. This circuit was chosen because of its common occurrence, and sufficient complexity to illustrate "beam-lead" (isolation) techniques.

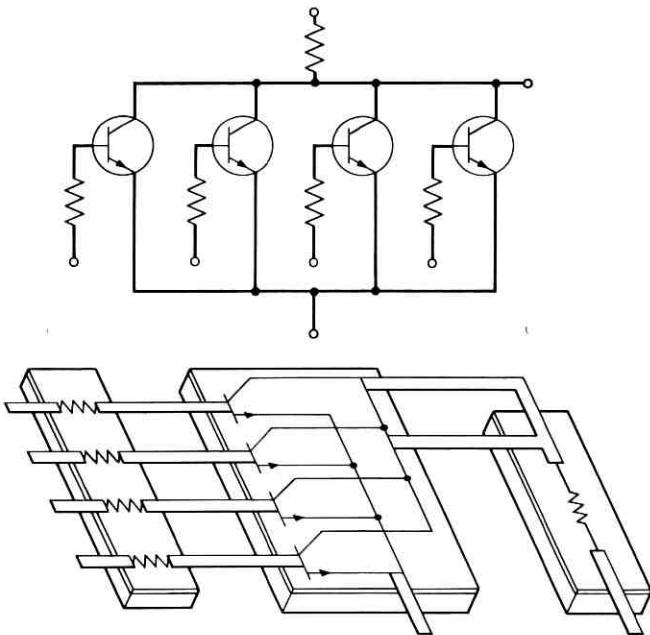Electrical measurements of switching performance were performed

Fig. 8 — Circuit schematic of DCTL gate and physical-electrical schematic showing component placement.

on 3-stage ring oscillators. Propagation delays of 4 ns were obtained, equivalent to the switching speed of comparable discrete device circuits. More complex circuits are currently under development.[3]

Figs. 9 through 11 are examples of beam-lead integrated circuits.

III. BEAM-LEAD PROCESSING

3.1 *Contact System*

Beam leads can be fabricated using many different metallurgical combinations, depending upon the environmental and thermal conditions the device must withstand. The objective of this project was to produce a contact structure for silicon devices that would survive 350°C oxidizing ambients, steam or air, continuously for 1000 hours without degradation. The metallurgical system Pt_5Si_2-Ti-Pt-Au was chosen, and the techniques developed to produce this structure are described in the following paragraphs.

The starting material is a slice of standard planar-oxidized silicon devices with contact holes etched in the SiO_2. At this stage the device
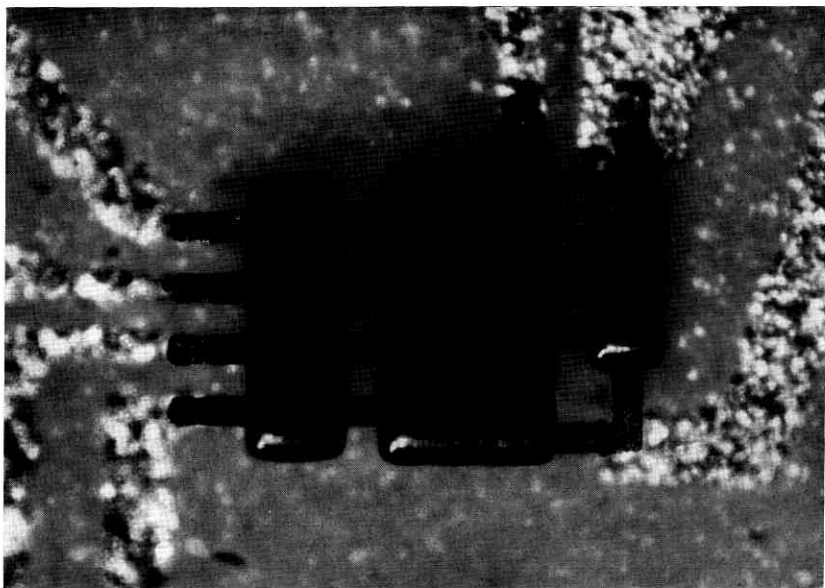

Fig. 9 — Beam-lead integrated DCTL gate mounted on gold-metallized ceramic, device side down.
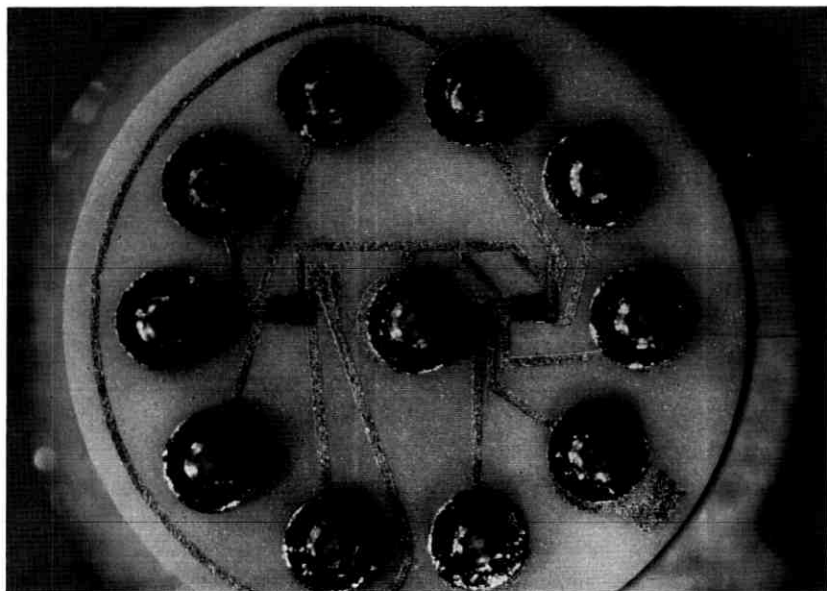
Fig. 10 — Two beam-lead integrated DCTL gates mounted on TO-5 header.

is ready for ohmic contacts. Platinum silicide was chosen as the ohmic contact material. It is one of the most stable compounds of silicon, has extreme corrosion resistance, is a solid phase up to 980°C,[4] has an optical reflectivity different from either silicon or platinum, and forms an ohmic contact to heavily-doped silicon. To form the compound in the contact holes, platinum is sputtered onto the whole slice after the contact holes are opened in the oxide, and heated to 700°C while still in the inert atmosphere. The platinum in the holes will react with the silicon to form Pt_5Si_2, which is a solid phase and will not ball up or creep beyond the edges of the contact holes as a liquid eutectic would. In addition, the silicide is a different color from platinum and can readily be distinguished from it under a microscope. Fig. 12 is a photograph of a device with the silicide formed in the contact hole. With improperly etched contact holes, spotty silicide formation or even no silicide at all is observed. An example of improper reaction is shown in Fig. 13. This is a photo of an integrated circuit. The square area with 3 contact stripes is a diffused base area; the two outside stripes are base contact stripes, and the center stripe is an emitter contact stripe. It can be seen that one of the base fingers in two different tran-
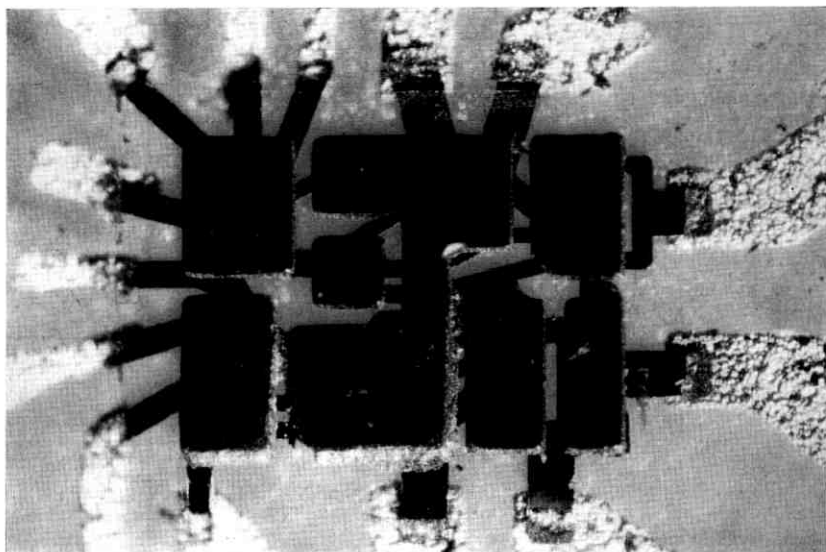
Fig. 11 — Dual 4-input DTL gate mounted on gold-metallized substrate, 0.030 inch over-all.

sistors has not formed the silicide; these two devices will not show proper electrical performance due to emission from only one edge of the emitter. This reason for substandard electrical characteristics would not have been detected if standard contacts had been used.

After the visual inspection, the platinum that is over the oxide is removed, leaving the silicide in the contact holes.

The next operation is to apply a material that will bond to both the SiO_2 and the Pt_5Si_2 and serve an an electrical connection to the external circuitry. No single element will satisfy all these conditions; however, a composite layered structure consisting of Ti-Pt-Au has been successfully employed. The first layer, titanium, is chosen because of its high oxygen activity, refractory nature, ability to absorb almost half its weight in reaction products interstitially[5] (commonly called "gettering"), and the natural oxide that it does form is completely self-passivating at temperatures up to 400°C. Although other metals, such as zirconium and hafnium, would probably be adequate, successful prior experience with titanium active metal semiconductor contacts[6] on such devices as shallow-junction solar cells for communications satellites and high-frequency silicon transistors makes titanium a logical choice for the oxide-bonding layer.
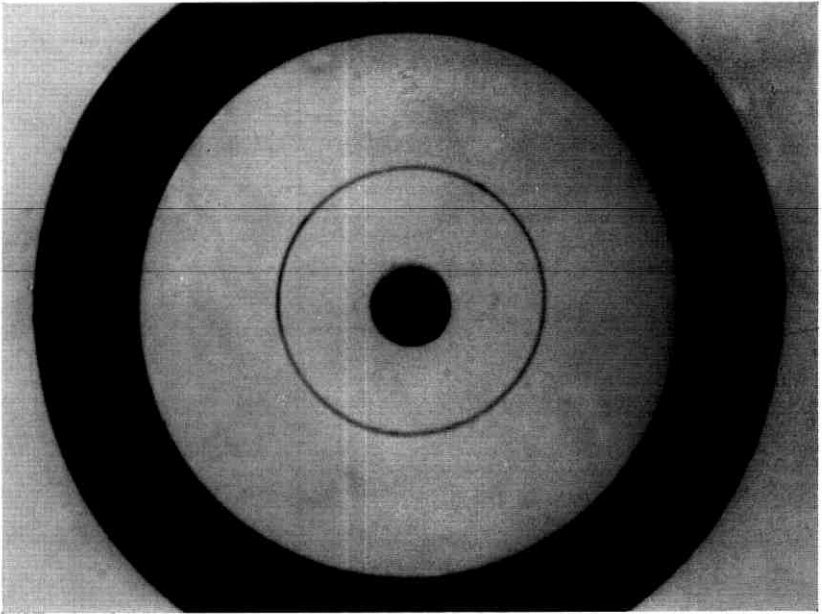
Fig. 12 — Platinum silicide ring-dot contact.

Table I is a list of some common metals, and the free energies of formation of their oxides,[7] in order of activity. This list may be used as a guide to the surface-bonding strengths. Titanium is the first active metal after silicon, and in addition to its great surface bonding strength may be used to penetrate thin SiO_2 layers. This property presumably allows titanium films to getter impurities from the passivating SiO_2 layer.

Gold is chosen as the outer layer because of its extreme resistance to corrosion, ease of bonding, low yield point, high elongation (allowing thermal expansion mismatch with the silicon substrate), and suitability for high-resolution electroforming. However, gold is a metallurgically reactive material, and reacts with titanium chemically at relatively low temperatures to form compounds which have none of the desirable characteristics of the individual metals. This leads to the third metal in the sandwich; platinum is chosen as the filler because of its inertness, ease of bonding the gold outer layer, and the low diffusion coefficient of gold into platinum (D less than 10^{-10} cm^2/sec at 900°C).[8]
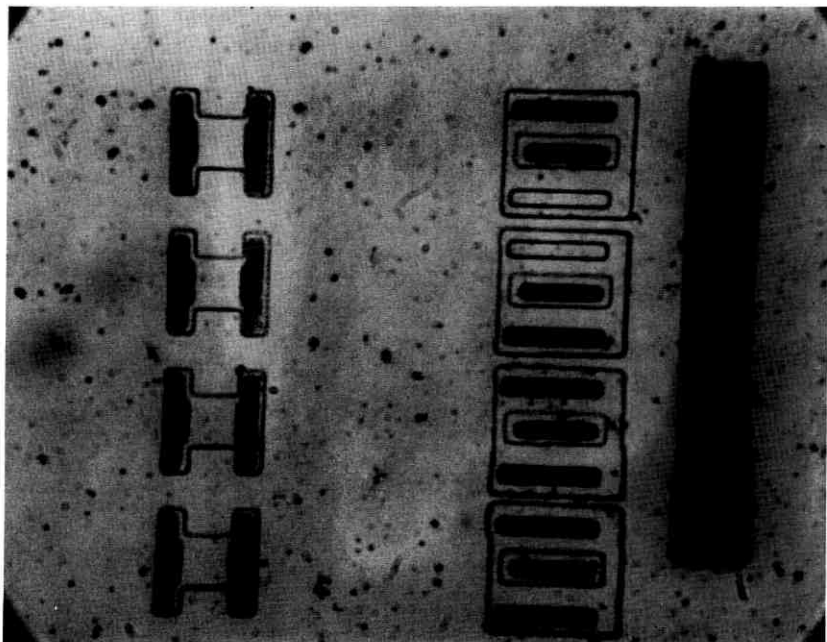
Fig. 13 — Improperly reacted contact holes.

3.2 *Sputtering*

Metal films are normally deposited onto semiconductors by vacuum evaporation. The use of the electrodes in beam-lead devices as structural supports, however, requires extreme adherence of the titanium to the SiO_2 layer. With the many processing operations and handling steps in producing a planar device, the surface cleanliness is variable from slice to slice. Since sputtered metals have many times the energy of thermally evaporated metals, and are capable of dislodging impurities, it was decided to use this method for the application of the titanium and platinum layers. The sputtering apparatus is shown in Fig. 14.

3.3 *Electroforming*

The method chosen to deposit the 0.5-mil thick gold beam leads is electroforming, the build-up of material by electroplating in selective areas. This technique has been developed to the point where extremely

TABLE I — FREE ENERGY OF FORMATION OF METAL OXIDES AT
500°K IN ORDER OF ACTIVITY

Oxide	$-F°$ (Kcal/mole)	$-F°$ (Kcal/gram atom oxygen)
NiO	46.1	46.1
MoO_2	114.5	57.2
Cr_2O_3	240.2	80.1
Na_2O	83.0	83.0
Ta_2O_5	434.9	87.0
SiO_2	187.9	94.0
TiO	112.2	112.2
ZrO_2	238.4	119.2
Al_2O_3	362.1	120.7

small geometries are feasible. Using high resolution KPR, dimensions under 0.2 mil are possible. Fig. 4 is a photograph of a high-frequency transistor with electroformed beam leads. The electrode fingers are plated to a thickness of 0.2 mil, thinner than the 0.5-mil thick beam leads, allowing the close spacing necessary for high-frequency operation. The aspect ratio of the fingers (height/width) can approach one for thin fingers.

3.4 *Back-Sputtering*

At this stage the electrodes and beams are defined by the gold. The remainder of the slice is coated with continuous layers of titanium and platinum, and patterns of gold beam leads and electrode fingers. Removing the platinum and titanium, in this order, from the areas outside the gold patterns remains to complete the metallizing process. Titanium is easily etched using the gold patterns as etch masks; however, platinum is much more difficult to etch than gold, especially without undercutting the narrow gold fingers. Therefore, a technique has been developed whereby the platinum is etched by glow discharge bombardment, now called "back-sputtering".

Fig. 15 is a drawing of the apparatus used for back-sputtering. A slice of oxidized and metallized silicon, up to one inch in diameter, is placed upon a ceramic spacer which is placed upon an aluminum cathode, and a glow discharge is excited with the cathode at −5 kv. The oxide thickness on the silicon slice is under 1 micron thick, and if this voltage appeared across the oxide, the dielectric stress would be 5×10^7 volts/cm, which is greater than the breakdown strength of the oxide. For thinner oxides, the stress would be correspondingly higher.
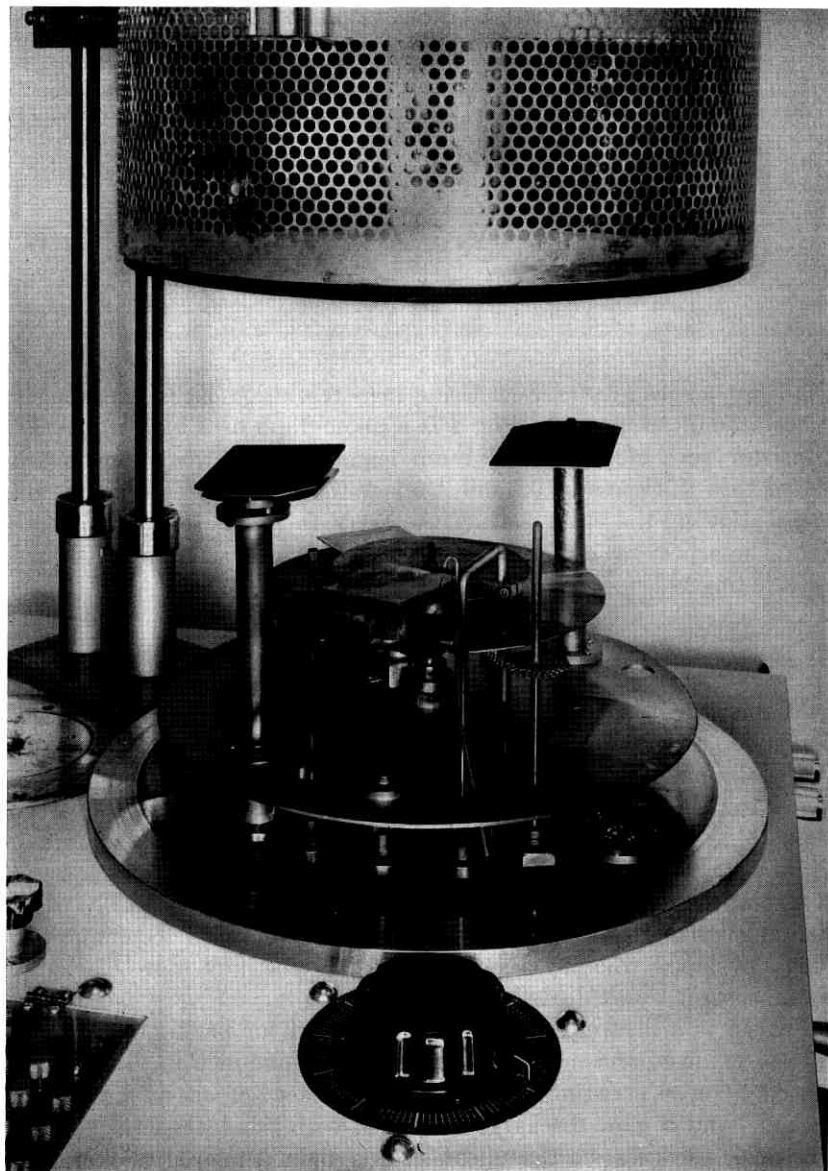
Fig. 14 — Sputtering apparatus.
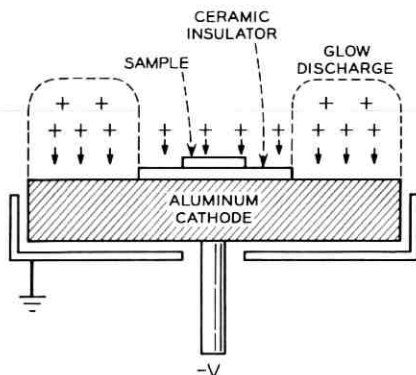
Fig. 15 — Drawing of back-sputtering apparatus.

The purpose of the ceramic spacer is therefore to prevent the dc field from appearing across the passivating oxide on the slice of devices should the metal film, shorting the edge of the slice, erode away.

Patterns of platinum, gold, rhodium, iridium, chrome-platinum, etc. have been delineated using this technique. In addition, holes have been etched into SiO_2 films using aluminum patterns as back-sputtering masks. The holes formed in this manner have very straight-sided edges due to the directional velocity of the bombarding ions. The limit of resolution of this technique is that of the masking procedure employed.

IV. TEST RESULTS

4.1 *Mechanical*

Special fixtures have been built to mechanically test beam-lead devices, and centrifuging has been performed to 135,000 g's without failure. Bending tests performed on the leads by bending over a 2-mil radius have yielded twenty 90° bends before breakage; 1.5-mil gold wire used as a control failed at 17 bends. Devices have also been successfully removed from substrates and reused. This attests to the strength and ductility of the electroformed leads.

Corrosion testing for periods up to 1000 hours in 350°C steam, 350°C air, and 350°C wet NaCl has revealed no physical or electrical deterioration of beam leads. In addition, temperature cycling and galvanic corrosion aging has been equally harmless.

4.2 *Junction Aging*

As a by-product of the metallization procedure described in the preceding section, there is considerable enhancement of surface stability and junction quality. This is attributed to the use of the sputtered Ti, a highly energetic getter-metal. pnp transistors, extremely sensitive to collector channel formation,[10] were fabricated and aged. One group was processed by normal procedure, using evaporated aluminum contacts; these devices acted as controls. Another group was processed with Pt-Ti-Pt-Au contacts, with half the group having junction-overlay contacts, and the other half non-overlay. An overlay contact is defined as a contact where the metal is delineated so that it remains on the oxide surface and extends beyond the junction. The two groups were then aged in 350°C steam. Fig. 16 is a graph of the collector leakage current during aging. The collectors of the control group (aluminum contacts) channeled within two hours, the collector leakage currents reaching hundreds of microamperes. The Ti-Pt contacted devices had leakage currents in the subnanoampere range after 400 hours, at which point the experiment was inadvertently terminated.

The curve of Fig. 17 illustrates the collector reverse current as a function of aging of a typical overlay pnp transistor under 21 volts reverse bias at 220°C. The reverse current, I_{cbo}, rose to 1 na after 50
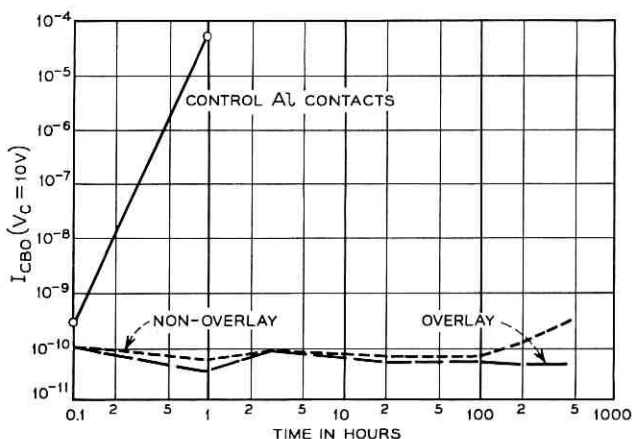
Fig. 16 — Graph of collector reverse current as a function of aging time in 350°C steam.
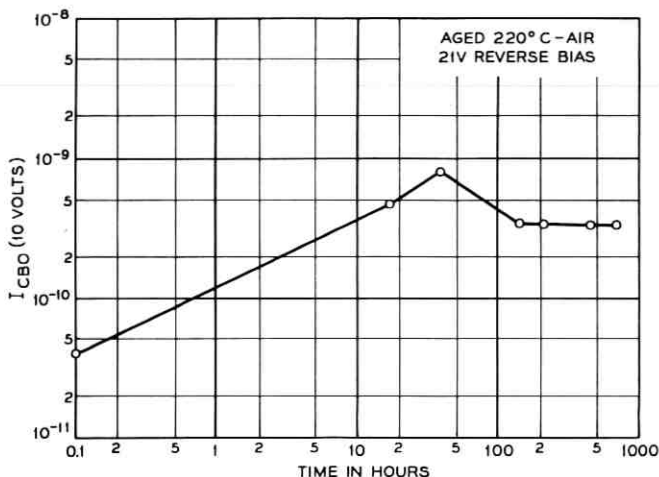
Fig. 17 — Graph of collector reverse current as a function of aging time under 21-volt bias and 220°C air.

hours, and then decayed to an equilibrium value in the 10^{-10} ampere range.

npn transistors, not prone to channel formation, have also been similarly aged, with favorable results for overlay, non-overlay, and completed beam-lead structures.

4.3 *Contact Resistance*

Forward junction characteristics have been monitored during aging in these hot, corrosive atmospheres. Typical resistance changes on contact areas of 0.2×1.6 mils are under 0.1 ohm after 1000 hours in 350°C air or steam; this is the measurement resolution. Obviously no "plaguing" problem has arisen with this metallurgical system since films formed at any of the metal interfaces would yield an increase in contact resistance, or require burning-in.

V. CONCLUSIONS

A semiconductor device structure and its application to high-frequency silicon switching transistors and ultrahigh-speed integrated circuits have been described. Beam leads have been shown to be rugged;

units have survived corrosive atmospheres (350°C steam), severe bending, and 135,000 g's centrifuging.

This structure imposes no electrical penalty; parasitic capacitance (under 0.05 pf per lead) is equivalent to wire-bonded and brazed-chip assembly. The beam-lead concept is applicable to many other electronic devices as well. In addition, the electroformed leads are not required to be of constant cross-section. They can taper out as they leave the silicon surface for lower inductance, stripline impedance matching, and even better heat conductance.

Assembly of these devices is considerably simplified. Chip-brazing and wire-bonding are eliminated; devices need not be handled by the silicon chip itself. The feasibility of this structure has been shown in the variety of digital, linear, and thin-film applique circuits which have been fabricated. Multiple arrays are the result of batch-fabrication of devices leading to the possibility of more automatic handling and assembly.

Isolation for integrated circuits is accomplished as a by-product of the structure. Parasitic capacitance has been shown to be negligible, and switching times are comparable to equivalent chip-and-wire circuitry. This has been achieved with no increase in process complexity over discrete devices.

The metallizing procedures developed to fabricate beam-lead devices yield an increase in device reliability; contact resistance is maintained at a low value during stringent accelerated aging tests; and passivation of junctions and surfaces has been enhanced. In addition, glow discharge etching has been developed, allowing the design of photo-reproduced geometries (down to 1 micron) of inert materials such as rhodium and iridium, compatible with normal device processing. Electroformed gold patterns of high resolution are employed as standard procedure.

REFERENCES

1. Lepselter, M. P. and MacDonald, R. W., Beam-Lead Devices, 1964 Electron Devices Meeting, Washington, D.C.
2. Lepselter, M. P., Waggener, H. A., and Davis, R. E., Beam-Leaded and Intraconnected Integrated Circuits, 1964 Electron Devices Meeting, Washington, D.C.
3. Murphy, B. T., Waggener, H. A., and Iwersen, J. E., 1965 IEEE Solid-State Circuits Conference, Philadelphia.
4. Hansen, M., Constitution of Binary Alloys, McGraw-Hill Book Company, Inc., New York, 1958.
5. McQuillan, A. D. and McQuillan, M. K., Titanium, Butterworths Publications Ltd., London, 1956.

6. Lepselter, M. P., U. S. Patent No. 3,106,489, Oct. 8, 1963.
7. Coughlin, J., Contribution to the Data on Theoretical Metallurgy, XII, Heats and Free Energies of Formation of Inorganic Oxides, U. S. Government Printing Office, Washington, D. C., 1954.
8. Matano, C., Japan J. Phys. (Trans.), *8*, 109, 1933.
9. Kuper, A. B. and Nicollian, E. H., J. Electrochemical Soc., *III*, 192C, abstract 131, August, 1964.

Errata

Measurements of Electromagnetic Backscattering from Known, Rough Surfaces, by J. Renau and J. A. Collinson, B.S.T.J., *44*, December, 1965, pp. 2203–2226.

On page 2204, replace the last two lines of the abstract with *an rms slope of 8° ± 4°, and a dielectric constant ε of 1.9 ± 0.3 at microwave frequences.* On page 2221, replace

$$m \equiv \frac{4\pi h}{\lambda} \cos \lambda \gtrsim \frac{1}{10}$$

with

$$m \equiv \frac{4\pi h}{\lambda} \cos \psi \gtrsim \frac{1}{10}.$$

From the experimental evidence presented, the condition on small m, p. 2221, for a surface to appear rough, in a radar sense, should have read $m \gtrsim 1$ instead of $m \gtrsim \frac{1}{10}$.

Techniques for Adaptive Equalization of Digital Communication Systems

By R. W. LUCKY

A previous paper described a simple adjustment algorithm which could be employed to set the tap gains of a transversal filter for the equalization of data transmission systems. An automatic equalizer was shown which used this algorithm during a training period of test pulse transmission prior to actual data transmission. The present paper extends the utility of this automatic equalization system by permitting it to change settings during the data transmission period in response to changes in transmission channel characteristics. Three schemes for accomplishing this adaptive equalization without the use of test signals are described and evaluated analytically. The first such scheme uses periodic estimates of channel response based on the received data signal to adjust or update the transversal filter settings. The second system is entirely digital and employs a sequential testing procedure to make adjustments aperiodically as they are required by changing conditions. The third system uses information obtained from a forward-acting error correction system for the purposes of adaptive equalization. Of the three systems described, the second is not only theoretically superior, but is practically the simplest. Experimental results for this second system are described.

I. INTRODUCTION

A previous paper[1] has dealt with the problem of automatic equalization for data transmission systems. In that paper, it was assumed that a finite-length transversal filter was to be used to correct the pulse response of a baseband (VSB) system at the sampling instants. A simple control system was shown which could be used to adjust the tap gains of the transversal filter to optimum positions using a series of test pulses transmitted prior to actual data transmission. After this training period the control system is disconnected, the tap gains remain fixed,

and normal data transmission ensues. This automatic equalization system has been used in conjunction with multilevel vestigial sideband modulation to achieve a rate of 9600 bits-per-second on private line voice facilities.[2,3,4]

Two limitations of this automatic equalization system are immediately apparent — it requires that test pulses be transmitted and it must be reset in another training period whenever the channel characteristics change. Other disadvantages of the present mode of operation include the long training period required to establish accurate final settings and the possibility of a nonlinear channel causing the transmission characteristics for data transmission to be slightly different from those for isolated pulse transmission.

For these reasons, it has been found advantageous to develop an equalizer capable of deriving its control signals directly from the transmitted data signal itself. Such an equalizer would be capable of tracking a time varying channel and would also circumvent the other difficulties associated with preset equalizer operation. To distinguish this equalizer from the previously described preset automatic equalizer, we shall call the tracking equalizer an *adaptive* equalizer. The purpose of the adaptive equalizer is to continually monitor channel conditions and to readjust itself when required so as to provide optimum equalization. To conserve signal power and bandwidth, the channel monitoring (or system identification) of the adaptive equalizer must be done using only the normal received data signal and without the benefit of added test information.

In this paper, we will present a few techniques which may be used to achieve adaptive equalization. Since these techniques are based on the use of a particular tap gain adjustment algorithm used in the preset equalizer, we will begin with a brief description of this equalizer and the algorithm.

II. PRINCIPLES OF TRANSVERSAL FILTER EQUALIZATION

In the preset automatic equalizer, a sequence of isolated test pulses is transmitted through the channel and demodulator. At the output of the demodulator, the pulse waveform is designated $x(t)$ as shown in Fig. 1. The pulse then passes through a $(2N + 1)$-tap transversal filter whose tap gain settings are $c_{-N}, \cdots, c_N$. The output pulse of the equalizer is

$$\hat{h}(t) = \sum_{j=-N}^{N} c_j x(t - jT). \tag{1}$$
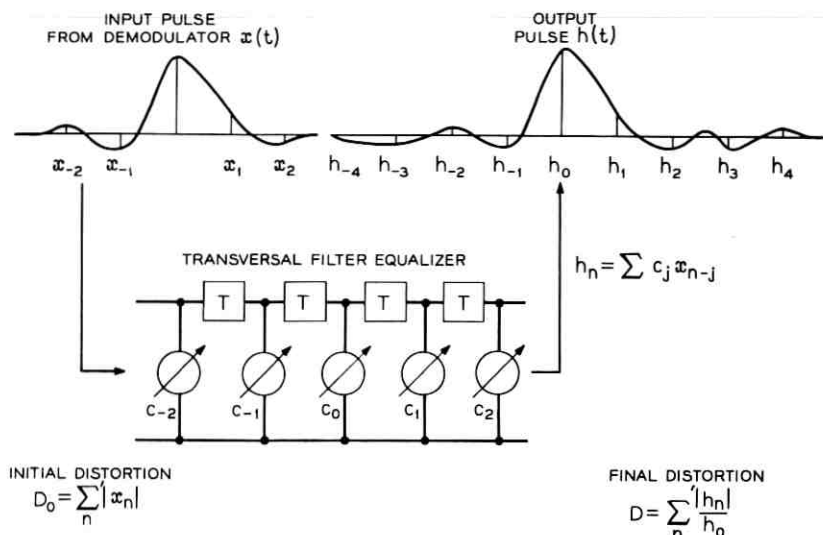
Fig. 1 — Transversal filter equalizer.

Since the output of the equalizer will be sampled at T second intervals during data transmission, we are only interested in the samples

$$\hat{h}_n = \hat{h}(nT)$$

of the output pulse. In terms of the input samples x_n we can write

$$\hat{h}_n = \sum_{j=-N}^{N} c_j x_{n-j} . \qquad (2)$$

The objective of the equalizer control circuitry is to set the tap gains such that the pulse distortion D is minimized, where we define[*]

$$D = \frac{1}{\hat{h}_0} \sum_{n=-\infty}^{\infty}{}' \mid \hat{h}_n \mid. \qquad (3)$$

This criterion is equivalent to requiring that the equalizer maximize the eye opening.

We assume that the input is normalized so that $x_0 = 1$ and that the center tap c_0 is used to satisfy the practical constraint $\hat{h}_0 = 1$. In Ref. 1 it was proved that if the initial distortion $D_0 < 1$, where

---

[*] Primes on summations indicate deletion of the zeroth term.

$$D_0 = \sum_{n=-\infty}^{\infty}{}' |x_n|, \tag{4}$$

then the output distortion D is at a minimum when the $2N$ tap gains c_j for $|j| \leq N, j \neq 0$ are adjusted so that $\hat{h}_n = 0$ for $|n| \leq N, n \neq 0$. In other words, if a binary eye is open before equalization, then using the tap gains to force zeros in the output response is optimum.

Also, if the initial distortion is less than unity, it was shown that a simple iterative procedure could be used to obtain optimum tap gain settings. In this procedure each tap gain c_j is adjusted an amount $-\Delta$ sgn $\hat{h}_j$ after each test pulse. (The center tap gain c_0 is adjusted by $-\Delta$ sgn $(\hat{h}_0 - 1)$.) Thus one simply inspects the polarities of the output pulse $\hat{h}(t)$ at the sampling times and uses this polarity information to advance or retard counter-controlled attenuators along the tapped delay line.

In constructing an adaptive equalizer we shall use this same simple adjustment algorithm. The problem now becomes one of finding the polarities of the channel pulse response $\hat{h}(t)$ without the benefit of test pulses. Once the decisions have been made as to the most likely polarities for the impulse response samples $\hat{h}_j$, we make discrete adjustments in the corresponding tap gains c_j to correct the equalizer. (This is shown diagrammatically in Fig. 2.) The adjustments in tap gains can be made periodically or only as required by changes in transmission characteristics. Since the normal telephone channel's characteristics change very slowly, the decisions regarding the channel response polarities can be
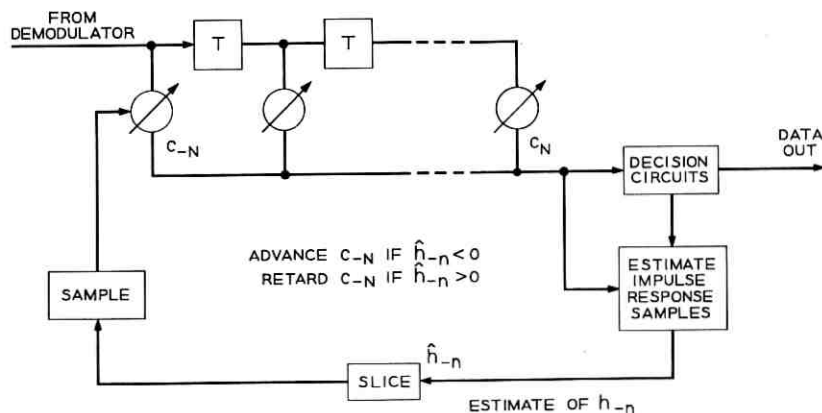

Fig. 2 — Adaptive equalizer.

made extremely accurate. In general the more accurate these decisions, the more precise the equalization which can be obtained, but with a corresponding increase in adaptive equalizer response, or settling time. We shall evaluate the accuracies and response times of the adaptive equalizers described in subsequent sections.

III. PERIODIC EQUALIZER ADJUSTMENT

3.1 *Maximum Likelihood Estimation of Response Values*

Let us assume that at the end of every KT seconds we wish to make a decision as to the current most probable impulse response polarities and to effect an incremental adjustment of the equalizer. The demodulated and equalized voltage at time t_k is

$$y_k = \sum_{n=-\infty}^{\infty} a_n \hat{h}_{k-n} + \eta_k, \tag{5}$$

where the a_n's are the input symbols chosen from an M-symbol alphabet, the h_n's are the samples of the overall (equalized) system impulse response, and the η_k's are noise samples.[*] We also make the following key assumptions.

(i)   The noise samples η_k are independent, identically distributed Gaussian variables with variance σ^2.

(ii) The input data symbols are uncorrelated.

(iii) The probability of error is relatively small, so that for practical purposes the sequence $\{a_n\}$ is available at the output of the detector.

(iv) The channel response samples $\hat{h}_n$ are essentially constant over the observation interval of KT seconds.

It is important to note that these assumptions are not made because they are true in a practical situation, but are made at this time to enable us to derive a reasonable system configuration and to assess its probable performance. Assumption (ii) is a particularly important one which we shall have to consider more carefully in predicting actual system performance.

Since we are using a decision-directed system — i.e., the detected symbols a_n are assumed correct — we can regard the signal samples y_k as being determined by noise and the set of parameters $\hat{h}_n$. By an appro-

---

[*] It is true that the noise sequence $\{\eta_k\}$ must pass through the equalizer. However, since in practice, all the tap gains c_j of the equilizer except for c_0 are very small we assume that the statistics of the noise sequence are unaffected by the equalizer.

priate statistical technique we can make estimates of the response
samples h_n from the signal samples y_k. Using the set of assumptions (i)
through (iv), the probability of receiving the sequence $\{y_k\}$, $k = 1$,
$\cdots$, K for a particular choice of the parameter set $\hat{h}_n$ is

$$p(\mathbf{y}\,|\,\mathbf{h}) = \prod_{k=1}^{K} \frac{\exp - \dfrac{1}{2\sigma^2}\left(y_k - \displaystyle\sum_{n=-\infty}^{\infty} a_n\hat{h}_{k-n}\right)^2}{\sigma\sqrt{2\pi}}. \tag{6}$$

The likelihood function $L(\mathbf{y}\,|\,\mathbf{h})$ is the logarithm of $p(\mathbf{y}\,|\,\mathbf{h})$. Apart from
a constant, this is

$$L(\mathbf{y}\,|\,\mathbf{h}) = \sum_{k=1}^{K} -\frac{1}{2\sigma^2}\left(y_k - \sum_{n=-\infty}^{\infty} a_n\hat{h}_{k-n}\right)^2. \tag{7}$$

The maximum likelihood estimates of the $(2N + 1)$ response values of
$\hat{h}_j$, $j = -N, \cdots, +N$, needed to adjust the transversal filter tap gains
are determined by the $(2N + 1)$ simultaneous equations $\partial L/\partial\hat{h}_j = 0$.
Thus

$$\sum_{k=1}^{K} a_{k-j}\left(y_k - \sum_{n=-\infty}^{\infty} a_n\hat{h}_{k-n}\right) = 0, \quad \text{for } j = -N, \cdots, +N. \tag{8}$$

These equations are more conveniently rewritten in the form

$$\frac{1}{K}\sum_{k=1}^{K} a_{k-j}y_k - \sum_{n=-\infty}^{\infty} \hat{h}_n A_{nj} = 0, \tag{9}$$

where

$$A_{nj} = \frac{1}{K}\sum_{k=1}^{K} a_{k-n}a_{k-j}. \tag{10}$$

Because of assumption (ii) that the input symbols are uncorrelated,
for a reasonably long averaging period K we can use $A_{nj} = S\delta_{nj}$ where
S is the average signal power. This simplification yields the estimators

$$\hat{h}_j = \frac{1}{KS}\sum_{k=1}^{K} a_{k-j}y_k. \tag{11}$$

We could easily now find the variance of the estimates (11) under
the assumption of random data, and we would discover that they are
poorly converging estimates with typically about 50,000 samples y_k re-
quired before we could move the equalizer taps with any degree of con-
fidence. This exorbitant settling time is caused by the presence of a
large parameter, $\hat{h}_0 \approx 1$, among a set of typically very small parameters.

Bear in mind that in normal operation the equalization will be very close to perfect and the samples $\hat{h}_j$ for $j \neq 0$ will be generally much less than 0.01 in magnitude.

This difficulty is circumvented by estimating $(\hat{h}_0 - 1)$ which is comparable in size to $\hat{h}_j$ for $j \neq 0$ instead of directly estimating $\hat{h}_0$. Therefore, we define

$$h_j = \begin{cases} \hat{h}_j & \text{for} \quad j \neq 0, \\ \hat{h}_0 - 1 & \text{for} \quad j = 0. \end{cases} \tag{12}$$

The samples h_j represent equalization error in the output pulse response. Following estimation of these values each of the taps c_j is advanced if h_j is negative and retarded if h_j is positive — the center tap being handled the same as any other tap.

Substitution of (12) in (7) gives the likelihood function

$$L(\mathbf{y} \mid \mathbf{h}) = \sum_{k=1}^{K} -\frac{1}{2\sigma^2} \left(y_k - a_k - \sum_{n=-\infty}^{\infty} a_n h_{k-n} \right)^2. \tag{13}$$

The quantities $(y_k - a_k)$ will be used frequently so we designate these e_k since they represent the error between the received sample y_k and the detected level a_k. The maximum likelihood estimates of the equalization error become

$$\hat{h}_j = \frac{1}{KS} \sum_{k=1}^{K} a_{k-j} e_k. \tag{14}$$

Fig. 3 shows a block diagram of an adaptive equalizer employing the estimates (14). In this system, the detected levels a_k are converted to analog form and subtracted from the received samples y_k to form the error samples e_k. The error samples e_k are then correlated simultaneously with each of the detected symbols a_{k-j} for $j = -N, \cdots, +N$. To accomplish this, the error samples e_k must be delayed NT seconds while the detected samples a_k are passed along a $(2N + 1)$-tap delay line. The delay line may consist of parallel shift registers since the samples a_k are in digital form. The outputs of the correlators (multipliers — low-pass filters) are sampled at K symbol intervals and appropriate actions are taken on the transversal filter attenuators.

3.2 Performance of the Adaptive Equalizer Under Ideal Conditions

In assessing the performance of an adaptive equalizer, we are primarily concerned with accuracy and settling time. The action of the equalizer
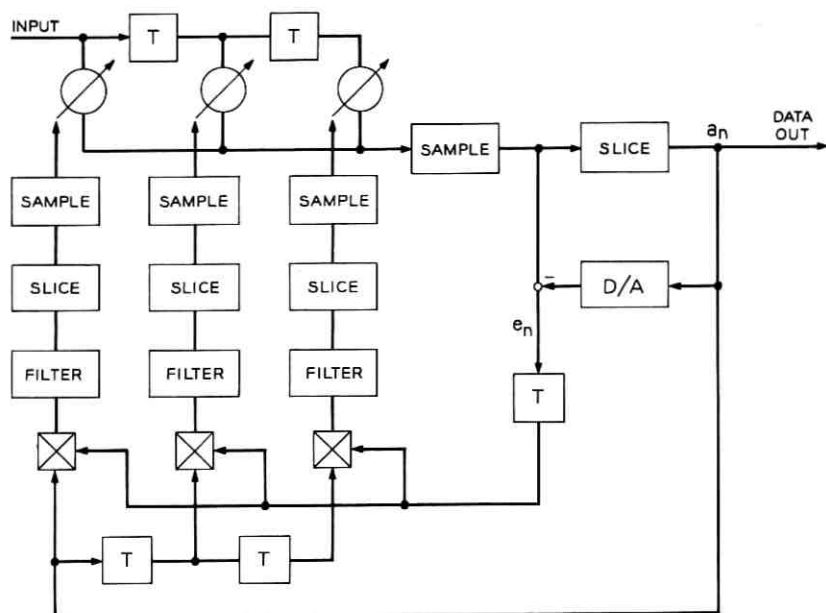
Fig. 3 — Periodically sampled adaptive equalizer.

is that of a multidimensional random walk with a bias toward correct equalization being supplied by the estimation circuitry. We shall make the assumption of small tap interaction and consider that each of the tap gains exhibits a one-dimensional walk independent of the action of other taps. The fundamental quantity involved in questions of accuracy and settling time is the probability $P(\text{sgn } \hat{h}_j = \text{sgn } h_j) = p_c$. For given noise and data statistics this probability is a function of the value of h_j and represents the probability of making a correct adjustment of tap c_j. When $|h_j|$ is large, i.e., equalization is poor, we expect that p_c will be close to unity, while for small $|h_j|$ the probability p_c approaches 0.5 and the tap gain tends to wander.

The adaptive equalizer must be designed to keep the inevitable wander of the tap gain within bounds imposed by accuracy requirements. Generally, we would wish to take full advantage of the inherent accuracy of the attenuator setting apparatus, i.e., each attenuator is adjusted in steps of Δ. Thus, when the equalizer is in perfect adjustment each tap gain will have an error of about 0.5Δ, which means that each of the samples h_j will be about 0.5Δ in magnitude. At the end of K symbol

durations, each of the gains will be increased or decreased by Δ. If $|h_j| = 0.5\Delta$ and a mistake is made in the polarity of $\hat{h}_j$, then the next value of h_j will be about 1.5Δ in magnitude and distortion will be considerably increased. We should generally design the system so that each tap gain spends a great majority of the time in the state where $|h_j| = 0.5\Delta$ and a small amount at 1.5Δ. Thus, for example, we may wish p_c to be 0.99 when $h_j = 0.5\Delta$.

We now evaluate $p_c(h_j)$ under the ideal conditions given in assumptions (i) through (iv). When the averaging period K is large the estimates $\hat{h}_j$ become normally distributed. The mean value of $\hat{h}_j$ is

$$\bar{\hat{h}}_j = \frac{1}{KS} \sum_{k=1}^{K} \overline{a_{k-j}e_k}. \tag{15}$$

The error sample e_k may be written

$$e_k = y_k - a_k = \sum_{n=-\infty}^{\infty} a_n h_{k-n} + \eta_k \tag{16}$$

so that (15) becomes

$$\bar{\hat{h}}_j = \frac{1}{KS} \sum_{k=1}^{K} \sum_{n=-\infty}^{\infty} \overline{a_{k-j}a_n}\, h_{k-n} + \frac{1}{KS} \sum_{k=1}^{K} \overline{a_{k-j}\eta_k}. \tag{17}$$

Since $\overline{a_k a_j} = S\delta_{kj}$ we have the necessary result

$$\bar{\hat{h}}_j = h_j. \tag{18}$$

The variance, σ_j^2, of $\hat{h}_j$ may be evaluated in straightforward fashion.

$$\sigma_j^2 = \frac{1}{K^2S^2} \sum_{k=1}^{K} \sum_{e=1}^{K} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \overline{a_{k-j}a_n a_{e-j}a_m}\, h_{k-n}h_{e-m}$$
$$+ \frac{1}{K^2S^2} \sum_{k=1}^{K} \sum_{e=1}^{K} \overline{a_{k-j}a_{e-j}\eta_k\eta_e} - h_j^2. \tag{19}$$

The four-fold sum in (19) may be partitioned into three sums involving pairwise equality in subscripts and an overlap term. A little manipulation yields

$$\sigma_j^2 = \frac{1}{K} \sum_{\substack{n=-\infty \\ n \neq j}}^{\infty} h_n^2 + \frac{2}{K^2} \sum_{n=1}^{K-1} (K-n)h_{j+n}h_{j-n} + \frac{\sigma^2}{KS}. \tag{20}$$

To get some sort of feel for the values involved, let's assume that $h_j = 0.5\Delta$ or -0.5Δ with equal probability when $|j| \leqq N$ and is zero otherwise. Then the expected value of σ_j^2 becomes

$$\overline{\sigma_j^2} = \frac{1}{K}\left(\frac{N\Delta^2}{2} + \frac{\sigma^2}{S}\right) \tag{21}$$

For p_c to be 0.99 when $h_j = 0.5\Delta$, we require that $0.5\Delta = 2.33\sigma_j$, so that for the zero noise condition

$$K \approx 10.8N. \tag{22}$$

For example, a 13-tap equalizer in the absence of noise requires about 140 symbols (0.058 sec. at a baud rate of 2400) to make a sufficiently accurate estimate. However, for a typical phone line application we require extremely accurate equalization so that Δ might typically be on the order of 0.0025 while the signal-to-noise ratio (S/σ^2) is about 30 db. In this case, the noise term in (21) completely swamps the "inter-symbol clutter" term $N\Delta^2/2$ and the 0.99 accuracy at 0.5Δ condition means that

$$K \approx \frac{21.7\sigma^2}{\Delta^2 S}. \tag{23}$$

For this example, 3470 samples (1.45 sec) are required per step of equalization.

3.3 *Performance of the Adaptive Equalizer Under Adverse Conditions*

Now that the equalization system configuration has been established under the assumption of ideal conditions, it becomes necessary to judge deterioration when these conditions are not met in practice. Throughout this paper, the specific application is assumed to be high-speed, voice telephone channel, data transmission.

Assumption (*i*) regarding Gaussian noise is usually justified in practice (the impulsive noise is not a determining factor for high-speed transmission) and in any event is not a very crucial assumption. Assumption (*iii*) that the received symbols are detected correctly is generally amply satisfied during normal data transmission with error rates of 0.01 or less. Remember that the information from a thousand or more symbols may be averaged to make one decision regarding equalization. Thus, we only require that a large majority of decisions are correct. Obviously the system performance deteriorates and finally "breaks" as the error rate becomes higher and higher, but the analytical evaluation of the effect appears difficult. Experimental results will be mentioned in a later section of this paper, but it should be said here that the system will work well with error rates on the order of 0.1.

Once the equalizer begins to work, the error rate quickly drops back to a more normal value.

The time variation rate of the channel must be matched with the accuracy requirement to arrive at an averaging time KT seconds during which the channel does not vary greatly compared to the size step Δ being taken on the equalizer taps. In the phone line application, this appears simple even for very high accuracies since the transmission characteristics are not usually observed to continually change at any great rate.

The most troublesome of the four ideal assumptions is the one involving uncorrelated input symbols. During normal data transmission one would expect that the sequence $\{a_n\}$ would appear random over an interval of a thousand symbols, but unfortunately this is frequently not the case. Long steady sequences of ones or of zeros may be used to hold the line, or the dotting pattern of alternate ones and zeros may be employed for some such purpose. In any event where a short, repetitive pattern is transmitted the spectrum of the transmitted signal consists of a number of discrete lines. Obviously it is impossible at the receiver to extract any information about the channel's transmission characteristics except at a few discrete points. Any adaptive equalizer must be prepared to weather this period and await new random data upon which meaningful decisions can be made. Fortunately, it can be shown that the adaptive equalizer of Fig. 3 acts on whatever information is available in the received data and retains its settings through periods of bad sequences.

We return now to (17) and remove the assumption concerning uncorrelated data. Then

$$\bar{h}_j = \sum_{n=-\infty}^{\infty} h_n r_{j-n} \qquad (24)$$

where r_j is the normalized autocorrelation function of the input data sequence (which is assumed to be stationary).

$$r_j = \lim_{K \to \infty} \frac{1}{2KS} \sum_{k=-K}^{K} a_k a_{k-j} . \qquad (25)$$

For ideal equalization, we require that the action of the equalizer cause $h_n = 0$ for $|n| \leq N$. The adaptive equalizer tries to accomplish this by forcing $\bar{h}_j = 0$ for $|j| \leq N$. As can be seen from (24), a slight error may be introduced if $r_j \neq 0$ for $j \neq 0$. The only error involved, however, is to cause some influence of the samples h_n for $|n| > N$ on

the tap settings. Generally, these samples which are outside the range of equalization are quite small (otherwise the number of equalizer taps needs to be increased) and their effect is only multiplied by the tails of the input autocorrelation function. Suppose, for example, that $h_n = 0$ for $|n| > N$. Then after equalization

$$\bar{h}_j = \sum_{n=-N}^{N} h_n r_{j-n} = 0 \quad \text{for} \quad |j| \leq N \tag{26}$$

and the only solution to this set of $(2N + 1)$ equations is the perfect state $h_n = 0$ for $|n| \leq N$ provided the matrix

$$R = \begin{matrix} r_0 & r_1 \cdots r_{2N} \\ r_{-1} & r_0 \\ \vdots & \\ r_{-2N} & \qquad r_0 \end{matrix} \tag{27}$$

is nonsingular.

If the data sequence consists of a repetitive pattern of period L symbols, then $r_j = r_{j+L}$ and the rank of R cannot be greater than L. Thus, if the period of pattern repetition L is less than the number of taps on the equalizer $(2N + 1)$, then the equalizer does not reach optimum settings. Consider then what settings the equalizer does reach. In view of $r_j = r_{j+L}$ we can rewrite (24) in the form

$$\bar{h}_j = \sum_{n=0}^{L-1} r_n \left\{ \sum_{m=-\infty}^{\infty} h_{j-n-mL} \right\} \tag{28}$$

and

$$\bar{h}_j = \bar{h}_{j+L} . \tag{29}$$

Thus, only L taps of the equalizer represent independent feedback loops, while the other $(2N + 1 - L)$ tap gains are slaves to the L gains considered independent. In fact, they receive the identical error signals and are thus incremented identical amounts. The equalizer solves the L simultaneous equations to arrive at

$$\sum_{m=-\infty}^{\infty} h_{j-mL} = 0 \tag{30}$$

for $j = 0, \cdots, L - 1$.

Actually, this solution minimizes the data distortion for the particular sequence being transmitted. The received samples y_k may be written in terms of the equalization error samples h_n using (5) and (12).

$$y_k = \sum_{n=-\infty}^{\infty} a_n h_{k-n} + a_k + \eta_k . \tag{31}$$

But, since $a_k = a_{k+L}$, we have

$$y_k = \sum_{n=0}^{L-1} a_n \left\{ \sum_{m=-\infty}^{\infty} h_{k-n-mL} \right\} + a_k + \eta_k \tag{32}$$

and we see that an equivalent channel response could be defined as $\{g_n\}$ where

$$g_n = \sum_{m=-\infty}^{\infty} h_{n-mL} \qquad 0 \le n \le L - 1$$

$$g_n = 0 \qquad\qquad \text{elsewhere.} \tag{33}$$

The distortion for the equivalent channel is minimized by zeroing the samples g_n within the range of the equalizer, but this is precisely what (30) indicates is done.

If the equalizer is at a nominally perfect setting ($h_n = 0; |n| \le N$) when the repetitive sequence is begun, then the equalizer holds its settings over the time of periodic transmission. There is no possibility of a drift in settings since taps c_k and c_{k+L} are "locked" together and yet must maintain the solutions $g_n = 0$. There are L free taps and L independent equations for which the taps are initially at a solution. Thus, the equalizer can hold its settings over an indefinite time while periodic sequences are transmitted so long as the channel characteristics remain fixed.

If the channel response changes while periodic sequences are being transmitted then the equalizer will move to a new solution of (30) to minimize the distortion for the particular periodic sequence. However, this solution will not infer that $h_n = 0$ for $|n| \le N$ and the equalization will not be perfect when random data starts again.

The action of the equalizer for unfavorable data sequences can be summarized by saying that the system always attempts to do the best equalization possible for the data being transmitted. In no case is the equalization deteriorated because of the adaptive loop over the performance of a fixed equalizer.

IV. AN IMPROVED DIGITALIZED ADAPTIVE EQUALIZER

4.1 *Description*

There is a great premium attached to the use of digital circuits where possible for reasons of equipment cost and size. It is possible to con-

siderably simplify the implementation of the adaptive equalizer of Fig. 3 by discarding linear concepts and using only polarity information throughout. Thus, instead of correlating the error signal e_k with the detected symbols a_{k-j} we add mod 2 the binary symbols corresponding to the polarities of e_k and a_{k-j}. The resulting simplification can be seen in Fig. 4. The symbol polarity sgn a_{k-j} is obtained by passing the most significant digit of the detected symbol (in binary format) through a shift register. In the Gray code commonly used for binary-to-multilevel conversion the first bit indicates polarity of the symbol.

The polarity of error sgn e_k can be produced by the simple expedient of adding an additional stage of slicing to the $\log_2 M$ slicers required for M-level transmission. Each stage of the $\log_2 M$ slicers "folds" the signal value about the last threshold, so that an extra stage simply produces automatically the polarity of the error sgn e_k. Fig. 4 illustrates a 16-level transmission system. Five stages of slicing are employed. The first four stages deliver the four detected bits indicating the received symbol a_k while the first and fifth bits are used for equalization purposes.

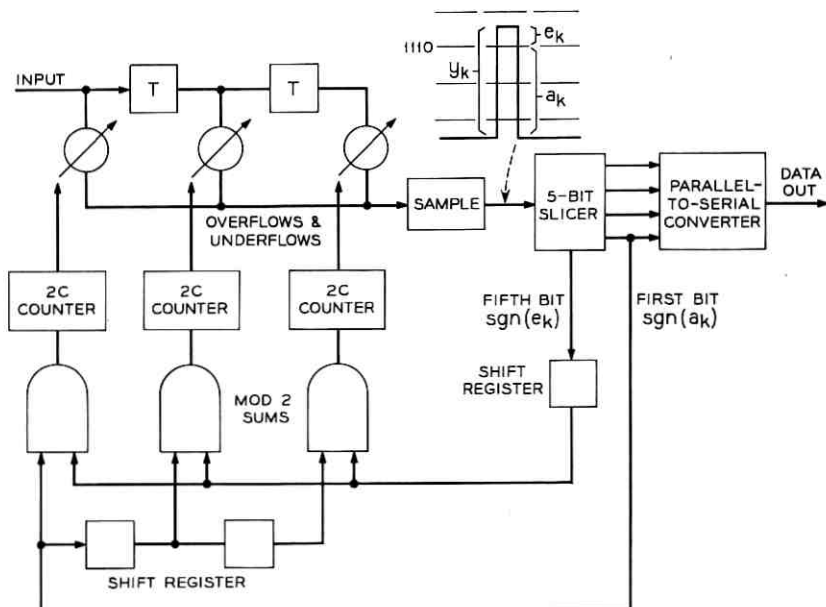After passing the error polarity through an N-stage shift register

Fig. 4 — 3-tap digital adaptive equalizer.

(in order to be able to correlate with N future and N past symbol polarities), the error polarities and symbol polarities are summed using exclusive-OR circuits. At this point we are able to easily introduce an improvement over the low-pass filter and periodic slicer used in the equalizer of Fig. 3. The problem with periodic slicing is that a compromise time interval KT must be chosen for averaging which is based on the most critical situation of near perfect equalization. The equalizer moves just as slowly when equalization is poor and the correlator output samples need much less averaging for a given accuracy.

A sequential testing procedure is clearly called for in this application. In a sequential test, the interval between decisions is determined by the input data itself. Instead of averaging data over a KT second interval and then sampling to determine polarity, running sums of the exclusive-OR outputs are kept. Positive and negative thresholds are set and the tests are terminated whenever these thresholds are crossed. This procedure is most easily implemented in digital form using up-down binary counters whose capacities of $2C$ counts determine the decision threshold value. Whenever a one is emitted from the exclusive-OR the counter advances one count, while a zero retards the counter one count. When the counter overflows we decide the polarity of h_k is positive and reduce the gain of on tap c_k. The counter is then reset to the center position of C counts. Similarly, an underflow adjusts c_k one step higher and resets the counters. The $2C$ storage counters are of course tied directly to the up-down counters which control the tap gain to accomplish this task in a most simple manner.

Thus, the equalizer of Fig. 4 is surprisingly simple. It requires only an N-stage shift register, a slicer, and $(2N + 1)$ binary counters of capacity $2C$ in order to convert a preset equalizer to the adaptive mode. Since the storage counters are used for averaging during stepup for the preset equalizer we finally arrive at an adaptive equalizer which costs almost nothing more than a preset equalizer.

The question arises as to why the preset mode (test pulses before transmission) is needed at all. In many cases it may not be needed provided a period of initial equalization is alloted during which data is transmitted, but not used due to its unreliability. As we shall find, the adaptive equalizer can accomplish a given degree of accuracy in equalization in less time than a preset equalizer providing the error rate is not too high. However, during initial setup the error rate is generally so high that the adaptive equalizer operates very slowly or not at all. Thus, a short period of test pulses can be profitably used to bring the

error rate down to manageable values before adaptive equalization is begun.

4.2 *Analytical Evaluation*

In this section we will evaluate the probability of correct adjustment of tap c_j and the average time required for an adjustment. The probability of correct adjustment p_c depends on the size ($2C$) of the storage counters and on the probability of an up-count p, and of a down-count q, on the jth storage counter.

The kth count of the jth counter is obtained by multiplying the polarities of e_k and a_{k-j}. Let us assume for convenience that h_j is positive (for negative h_j the situation is, of course, entirely similar). The probability of a correct adjustment is then the probability of an overflow occurring before an underflow. The probability of an up-count is

$$p = P(e_k > 0, a_{k-j} > 0) + P(e_k < 0, a_{k-j} < 0). \tag{34}$$

These probabilities are identical so we use

$$p = 2P(e_k > 0, a_{k-j} > 0). \tag{35}$$

Equation (35) can be rewritten in terms of the conditional probability

$$p = 2P(e_k > 0 \mid a_{k-j} > 0)P(a_{k-j} > 0). \tag{36}$$

The symbols a_j will be taken as independent and equally likely to assume any of the M values. Take $2d$ as the distance between adjacent levels, so that d is the distance from any level to the nearest slicing (decision) threshold. The amplitudes a_j can assume are then $d(2i - 1)$ for $i = -M/2 + 1, \cdots, M/2$. Since a_{k-j} can be positive or negative with equal likelihood, (36) becomes

$$p = P(e_k > 0 \mid a_{k-j} > 0) \tag{37}$$

$$p = \frac{2}{M} \sum_{i=1}^{M/2} P[e_k > 0 \mid a_{k-j} = d(2i - 1)]. \tag{38}$$

Now we need to evaluate the conditional probabilities in (38). With equalization error samples h_n at times nT and noise samples η_n, the received voltage at time kT is

$$y_k = \sum_n a_n h_{k-n} + a_k + \eta_k. \tag{39}$$

The error voltage e_k is

$$e_k = \sum_{n=-\infty}^{\infty} a_n h_{k-n} + \eta_k. \tag{40}$$

We remove the term involving a_{k-j} in (40) to obtain

$$e_k = a_{k-j} h_j + \left[\sum_{n \neq k-j} a_n h_{k-n} + \eta_k \right]. \tag{41}$$

The assumption is made that the sum of the intersymbol interference and noise (the terms in brackets) is Gaussian distributed, with mean zero and variance σ^2. The error e_k is then Gaussian with mean $a_{k-j} h_j$, and variance σ^2.

The probability density of e_k is sketched in Fig. 5. The conditional probabilities in (38) should be interpreted as the probability that, given given a_{k-j} was at level i, we *decide* that e_k is positive. If e_k crosses the decision threshold on the right-hand side of Fig. 5, it will appear to the receiver that e_k was negative since a_k will be incorrectly received as the next higher symbol. If e_k crosses the decision threshold on the left, e_k is interpreted as being positive. Thus, the conditional probability may be written (refer to Fig. 5)

$$P(e_k > 0 \mid a_{k-j}) = \tfrac{1}{2} + p_1 - (p_2 - p_3) \tag{42}$$

where

$$p_1 = \frac{1}{\sqrt{2\pi}\sigma} \int_{-a_{k-j}h_j}^{0} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx \tag{43}$$

$$(p_2 - p_3) = \frac{1}{\sqrt{2\pi}\sigma} \int_{d-a_{k-j}h_j}^{d+a_{k-j}h_j} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx. \tag{44}$$
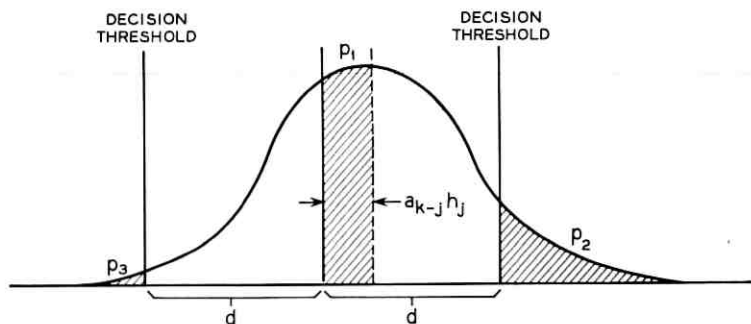

Fig. 5 — The probability density $p(e_k \mid a_{k-j})$.

With the equalization even fairly good (as it must be if meaningful data are being transmitted), $a_{k-j}h_j$ must be a small number so that the ranges of integration in (43) and (44) are small. Thus, we make the approximations

$$p_1 \cong \frac{a_{k-j}h_j}{\sqrt{2\pi}\sigma} \tag{45}$$

$$(p_2 - p_3) \cong \frac{2a_{k-j}h_j \exp\left(-\dfrac{d^2}{2\sigma^2}\right)}{\sqrt{2\pi}\sigma}. \tag{46}$$

Using these approximations we can write

$$P[e_k > 0 \mid a_{k-j} = d(2i - 1)] = \tfrac{1}{2}$$
$$+ \frac{d(2i - 1)h_j}{\sqrt{2\pi}\sigma}\left[1 - 2\exp\left(-\frac{d^2}{2\sigma^2}\right)\right]. \tag{47}$$

This expression is substituted into (38) and the summation is easily performed over i to give

$$p = \tfrac{1}{2} + \left[1 - 2\exp\left(-\frac{d^2}{2\sigma^2}\right)\right]\frac{h_j dM}{2\sigma\sqrt{2\pi}}. \tag{48}$$

For normal operation the exponential term in the brackets is small in comparison with unity and may be neglected. This term gives the contribution due to slightly more errors being made in the right-hand region (p_2) than in the left-hand region (p_3). An examination of the relative probabilities p and $(p_2 - p_3)$ without the approximations (45) and (46) shows that $(p_2 - p_3)$ is generally small in comparison with p, even for error rates if 0.01 and higher. Thus, the system is able to estimate the polarity of h_j even when the eye is completely closed.

The final approximation for the probability p of a correct step in the jth counter becomes

$$p = \tfrac{1}{2} + \frac{h_j dM}{2\sigma\sqrt{2\pi}}. \tag{49}$$

Generally, this probability is only a little larger than 0.5, hence many counts must be averaged before a decision can be made as to whether $p > \tfrac{1}{2}$, in which case $h_j > 0$, or whether $p < \tfrac{1}{2}$ and consequently $h_j < 0$. This averaging is best done using the 2C-count storage devices, since these devices effect a sequential test of the two hypotheses $p > \tfrac{1}{2}$ and $p < \tfrac{1}{2}$. This sequential test will require less time on the average

for a given accuracy than straightforward averaging with a fixed sample size as in the adaptive equalizer of Fig. 3. In addition, it is more easily implemented than the former technique.

The probability p_c of an overflow before an underflow when the probability of an up-count is p and the probability of a down-count is $q = 1 - p$ is taken from Feller's analysis of the problem of the gambler's ruin.[5] For a $2C$ counter initially set to its midpoint of C, Feller gives

$$p_c = 1 - \frac{(q/p)^{2C} - (q/p)^C}{(q/p)^{2C} - 1}. \tag{50}$$

If p_c is to be close to unity, $(q/p)^{2C}$ must be small compared with $(q/p)^C$, so we approximate (50) as

$$p_c \approx 1 - (q/p)^C. \tag{51}$$

We can further simplify this expression since q and p are both very close to 0.5. Writing

$$\left. \begin{array}{l} p = 0.5 + \varepsilon \\ q = 0.5 - \varepsilon \end{array} \right\} \tag{52}$$

we obtain

$$(q/p) \approx 1 - 4\varepsilon \tag{53}$$

and we use

$$(1 - 4\varepsilon)^C = \exp [C \log (1 - 4\varepsilon)] \tag{54}$$

to obtain the approximation

$$p_c = 1 - \exp (-4C\varepsilon). \tag{55}$$

Finally, we substitute the value of ε from (49).

$$p_c = 1 - \exp \left(- \frac{2Ch_j dM}{\sigma \sqrt{2\pi}} \right). \tag{56}$$

The average number of counts required for an overflow (or underflow) is also given by Feller.

$$\bar{n} = \frac{C}{(q - p)} - \frac{2C}{(q - p)} \left[\frac{1 - (q/p)^C}{1 - (q/p)^{2C}} \right]. \tag{57}$$

Using the same approximations in (57) as we used in (51) results in the approximation

$$\bar{n} \cong \frac{C\sigma \sqrt{2\pi}}{dMh_j} .$$ (58)

In order to be able to compare this system with the previous adaptive equalization system which uses linear techniques and averages over a fixed interval of K symbols, we need to find p_c in terms of h_j, $\bar{n}$ and S/σ^2. The signal power S for an M-level system with separation $2d$ is

$$S = \frac{2}{M} \sum_{i=1}^{M/2} d^2 (2i - 1)^2 = \frac{d^2}{3} (M^2 - 1).$$ (59)

Now combining (59), (58), and (56) we arrive at (for $M \gg 1$)

$$p_c = 1 - \exp\left(-\frac{3}{\pi} \bar{n} h_j^2 \frac{S}{\sigma^2}\right) \quad \text{(digitalized)}$$ (60)

whereas an equivalent approximation for the previous equalizer is

$$p_c = 1 - \frac{\sigma}{h_j \sqrt{2\pi KS}} \exp\left(-\tfrac{1}{2} K h_j^2 \frac{S}{\sigma^2}\right).$$ (61)

As expected, this comparison shows the previous system requires about twice as much time to achieve a given degree of accuracy as does the digitalized system. For the specific example used earlier 3090 symbols (1.29 seconds) are required to achieve an accuracy $p_c = 0.99$ when $h_j = 0.5\Delta$, $\Delta = 0.0025$, and $S/\sigma^2 = 1000$.

This does not mean that the digitalized equalizer operates twice as fast as the previous equalizer. Actually, it operates much faster than that. The average time of 3090 symbols is required only when $h_j = 0.5\Delta$ and since this represents perfect equalization we don't care how long the equalizer takes to move to the state -0.5Δ. When the gain c_j is out of equalization by a single step h_j is approximately 1.5Δ and according to (58) the time $\bar{n}$ required is only a third as much — or 1030 symbols. Similarly, when c_j is 2 steps out $\bar{n}$ becomes 618 symbols. If the equalizer is turned on when equalization is relatively poor the steps are taken in nearly the minimum time of C symbols. The counter capacity may be calculated from the accuracy requirement using (56). For our example we obtain $C = 85$. (Of course, either a 7-stage ($2C = 128$) or an 8-stage ($2C = 256$) would have to be used in practice.)

The longest average time before an equalizer change is required when $h_j = 0$ and, consequently, $p_c = 0.5$. Here a long average time is desirable since disturbing the equalizer is detrimental. The time in such case is C^2 symbols — in our example about 7200 symbols. Thus, the average time

between equalizer adjustments varies between C (85) and C^2 (7200) symbols with short times used when urgency of movement is greatest and longer times used when leisurely adjustment is possible and, in fact, necessary because of stringent accuracy requirements.

V. ADAPTIVE EQUALIZATION USING ERROR CONTROL INFORMATION

5.1 Description

In many applications for adaptive equalization a forward-acting error correction system will be associated with the data transmission system. When the objective of system design is high-speed transmission, the modem is generally operated at an unacceptably high error rate. A detection-retransmission system cannot be solely relied upon, since the high error rate would necessitate constant requests for retransmission.

In the exploratory VSB system described in Refs. 2, 3, and 4, a (200, 175) Bose-Chaudhuri code with a minimum distance of 8 was used for triple error correction. In the event of a detectable error pattern containing more than three errors, a retransmission request was made. Using triple error correction with a modem error rate of roughly 2×10^{-3}, the frequency of requests for retransmission was low enough to not have appreciably affected the throughput of the system.

When the equalization is imperfect the error rate is naturally increased, but moreover the data system becomes pattern sensitive. Some patterns of input data are more likely to result in errors than other patterns because of the memory of the system. Given that an error has occurred, it is quite likely that such a bad pattern was transmitted. Since the bad patterns are simply related to the system impulse response we have the interesting possibility of using the information available in the error correction system for the purpose of adjusting the equalizer.

Briefly, a scheme based on this principle works as follows. Whenever an error is corrected by the error control unit, the direction of the error and the polarities of the surrounding symbols are observed. By "direction" of the error is meant whether a symbol has been changed to a higher amplitude level (+) or to a lower level (−). If the direction of the error is positive the polarities of the surrounding symbols are taken directly to the equalizer. If the direction of the error is negative all symbol polarities are inverted. These polarities are used to either advance or retard counters attached to the variable attenuators of the equalizer. The attenuators are incremented positively or negatively

whenever the corresponding counters underflow or overflow. A block diagram of this system is shown in Fig. 6.

For a specific example, suppose that 16-level transmission is used. The incoming binary data train is converted 4 bits at a time into the 16 symbols using a Gray code. Suppose that at the receiver the following sequence is received

$$0011 \quad 1011 \quad 1\overset{X}{*}10 \quad 0110 \quad 0001.$$

The error control locates and corrects an error in the bit marked with an X. As part of the error correction procedure the entire 200-bit word has been stored, so the polarities of the symbols surrounding the error are readily available. In fact, with a Gray code the polarity of the amplitude level corresponding to a given 4-bit symbol is determined by the first bit. Therefore, the polarities of the two symbols either side of the error are, in order,

$$- \; + \; (\text{error}) \; - \; -.$$

The symbol which was in error was 1010 changed to 1110. From a table of the Gray code we can determine that this error carried amplitude level 13 into amplitude level 12. Thus, the direction of the error is nega-
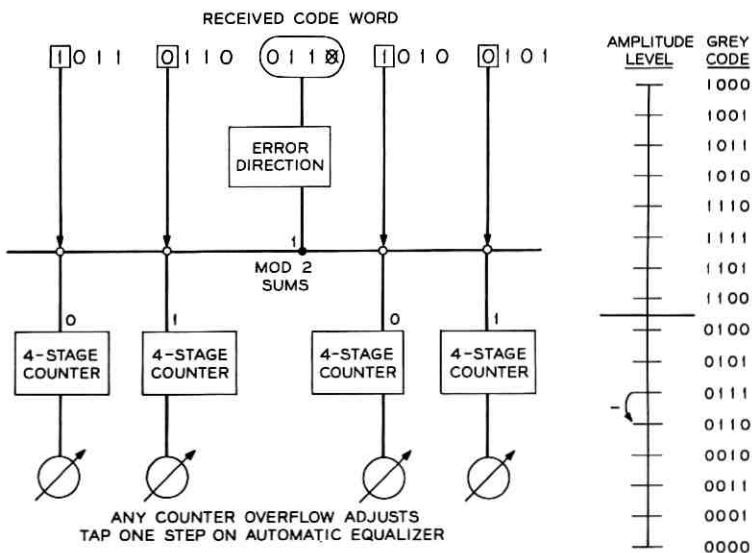

Fig. 6 — Example error control — adaptive equalizer coordination ($2C = 16$).

tive and the symbol polarities are reversed and used to increment counters.

Notice that in this system the storage counters are only incremented when errors occur. If no errors are being made, then the equalizer settings are not changed. Therefore, the counters are changed much less frequently in this system than in the previous adaptive equalizer. However, when the counters are changed, we shall find that the changes are more reliable and that smaller counters may be used to effect a comparably reliable statistical test.

5.2 Analytical Evaluation

Again we are going to evaluate the probability p of counting in the correct direction on the storage counter attached to attenuator c_j. We suppose that h_j is positive and with a spacing $2d$ between levels we have

$$p = P(a_{k-j} > 0 \mid e_k > d) + P(a_{k-j} < 0 \mid e_k < -d) \quad (62)$$

$$p = 2P(a_{k-j} > 0 \mid e_k > d). \quad (63)$$

With the M possible symbols equally likely we can write (63) in the form

$$p = \frac{\sum_{i=1}^{m/2} P[e_k > d \mid a_{k-j} = d(2i - 1)]}{MP(e_k > d)}. \quad (64)$$

As in Section 4.2, we write the error voltage e_k

$$e_k = a_{k-j}h_j + \left[\sum_{n \neq k-j} a_n h_{k-n} + \eta_k \right] \quad (65)$$

and assume the bracketed term is Gaussian, mean zero, variance σ^2.

Thus,

$$P[e_k > d \mid a_{k-j} = d(2i - 1)]$$
$$= \frac{1}{\sqrt{2\pi}\sigma} \int_{d-d(2i-1)h_j}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2} \right) dx \quad (66)$$

and for reasonably small error rates we make the approximation

$$P[e_k > d \mid a_{k-j} = d(2i - 1)]$$
$$\cong \frac{\sigma}{\sqrt{2\pi}\, d[1 - (2i - 1)h_j]} \exp\left(-\frac{d^2[1 - (2i - 1)h_j]^2}{2\sigma^2} \right). \quad (67)$$

If we also assume that when a_{k-j} is a random variable, e_k is Gaussian

distributed and make the same approximation involved in (67), we arrive at

$$P(e_k > d) \cong \frac{\sqrt{\sigma^2 + Sh_j}}{\sqrt{2\pi}\, d} \exp\left(-\frac{d^2}{2(\sigma^2 + Sh_j)}\right). \qquad (68)$$

Presumably $\sigma^2 \gg Sh_j$ when the system is near perfect equalization, so we drop the Sh_j terms. Equations (68) and (67) are then inserted into (64) to get

$$p = \frac{1}{M} \sum_{i=1}^{M/2} \frac{1}{[1 - (2i - 1)h_j]} \exp\left(-\frac{d^2[1 - (2i - 1)h_j]^2 + d^2}{2\sigma^2}\right). \qquad (69)$$

Our aim in evaluating (69) is an approximation which is accurate to terms linear in h_j, a small number. The denominator in (69) does not contribute terms on this order, so we are able to sum the geometric series giving

$$p = \frac{1}{M} \exp\left(\frac{d^2 h_j}{\sigma^2}\right) \left[\frac{1 - \exp\left(\frac{M d^2 h_j}{\sigma^2}\right)}{1 - \exp\left(\frac{2 d^2 h_j}{\sigma^2}\right)} \right]. \qquad (70)$$

Finally, we retain only terms linear in h_j to obtain the result

$$p = \tfrac{1}{2} + \frac{M d^2 h_j}{4\sigma^2}. \qquad (71)$$

Equation (71) is similar in form to (49) for the probability of a correct count in the digitalized equalizer. The principal difference is that (71) involves the threshold-to-noise ratio squared (d^2/σ^2) whereas (49) uses the ratio (d/σ). Thus, p in (71) is considerably more reliable than the probability for a correct count in the digitalized adaptive equalizer. However, the counts in the error control system occur at a much slower rate, namely $2P(e_k > d)$.

A counter of capacity $2C$ is used to store the counts from the error correction circuitry. The equations of Feller may again be used with suitable approximations to find the probability p_c of a correct equalizer adjustment and the average number of counts $\bar{n}$ required before a correction.

$$p_c = 1 - \exp\left(-\frac{C h_j d^2 M}{\sigma^2}\right) \qquad (72)$$

$$\bar{n} = \frac{2C\sigma^2}{M d^2 h_j}. \qquad (73)$$

In order to get the average number of symbols required before an equalizer adjustment, we must multiply $\bar{n}$ by the average number of symbols per error, which is approximately [from (68)]

$$n_0 = \frac{\sqrt{2\pi}\,d}{2\sigma e^{-d^2/2\sigma^2}} \cdot \tag{74}$$

The comparison of this system with the previous system becomes quite complicated because of the dependence of the error control system upon the number of levels M and on the error rate of the system (for which our approximation is only valid when equalization is exact). In general, it seems that for a given accuracy of equalization, the previous digitalized equalizer will require less time per adjustment. To follow through with our example we assume $h_j = 0.5\Delta$, $\Delta = 0.0025$, $S/\sigma^2 = 10^3$, and now $M = 16$. Since

$$S/\sigma^2 = \frac{d^2}{3\sigma^2}(M^2 - 1) \tag{75}$$

the threshold-to-noise ratio is

$$d^2/\sigma^2 = 11.76. \tag{76}$$

The probability of error for the system is about 6.5×10^{-4} and $n_0 = 1540$ symbols per error. From (72), we find that a counter capacity $2C = 40$ is required to ensure $p_c = 0.99$. This may be compared with $2C = 170$ for the previous system. The number of counts per adjustment $\bar{n} = 170$, but $n_0\bar{n} = 2.62 \times 10^5$ symbols per adjustment.

This comparison is somewhat unfair to the error control system since it must be pointed out that the speed of movement at $h_j = 0.5\Delta$ is immaterial. This condition merely determines the counter sizes necessary to meet accuracy requirements. We are much more concerned with the equalizer response when the equalization is imperfect. In this case, not only is $\bar{n}$ inversely proportional to h_j, but the error rate also increases so that counts are made more frequently. Even if one is willing to stretch the point quite a bit, it does seem that the error control equalizer coordination is unattractive in comparison with the previous adaptive system. Nevertheless the system implementation is quite simple and the concept sufficiently intriguing that perhaps a use can be found for such a coordination.

VI. EXPERIMENTAL RESULTS

Three systems for adaptive equalization of digital data systems have been described and analyzed. One of these three systems, the digitalized

adaptive equalizer described in Section IV, appears to be much more attractive than the other two, both from the standpoint of instrumentation ease and of performance. Therefore, although the error control coordinated system has also been constructed, only the digitalized adaptive system has been subjected to extensive testing.

The system constructed used a 13-tap delay line and a tap increment Δ of 0.0025, although this increment is tapered to considerably smaller values near the outside taps of the delay line. In line with the example values computed in Section IV, eight stages were used in the storage counters, resulting in a capacity of $2C = 256$ counts for each tap.

The system was tested in conjunction with the 9600 bit-per-second, 16-level VSB system described by F. K. Becker in Ref. 2. A good pictorial demonstration of the adaptive equalizer is shown in the sequence of photos in Fig. 7 where the VSB system is transmitting at 4800 bits-per-second, using only four levels. The reduced speed here is to enable us to easily discern the distinct levels in the eye pictures of Fig. 7. The first photo shows the normal, equalized eye pattern. The intensified dots on this picture indicate the position of the sampling time which in this first photo has been artificially moved to the right, completely out of the eye opening. This is equivalent to the sudden introduction of a constant delay into the transmission channel. It can be observed in this first photo that the error rate would be relatively high due to the mistiming.

In the subsequent pictures the timing position has been left fixed while the adaptive equalizer changes its settings to move the entire eye pattern to the right effecting a reequalization of the system. At first the pattern moves quite rapidly since decisions are made quickly by the testing counters. As the eye approaches the timing position again decisions are made at a slower rate and the movement slows. The entire process takes only about a few seconds in spite of the quite abnormally large disturbance of the transmission characteristic.

Fig. 8 shows a sequence of photographs of a 16-level folded eye picture following a sudden change in transmission characteristics. If the 16-level eye were shown in the same format as the 4-level eyes of Fig. 4, the 15 "holes" in the eye diagram would be too small to distinguish. Therefore, all 15 "holes" have been superimposed by folding the 16-level eye diagram over and over until the picture resembles a binary eye diagram. In the first second of operation, from 0 to 1 second, the eye is recaptured by the adaptive equalizer.

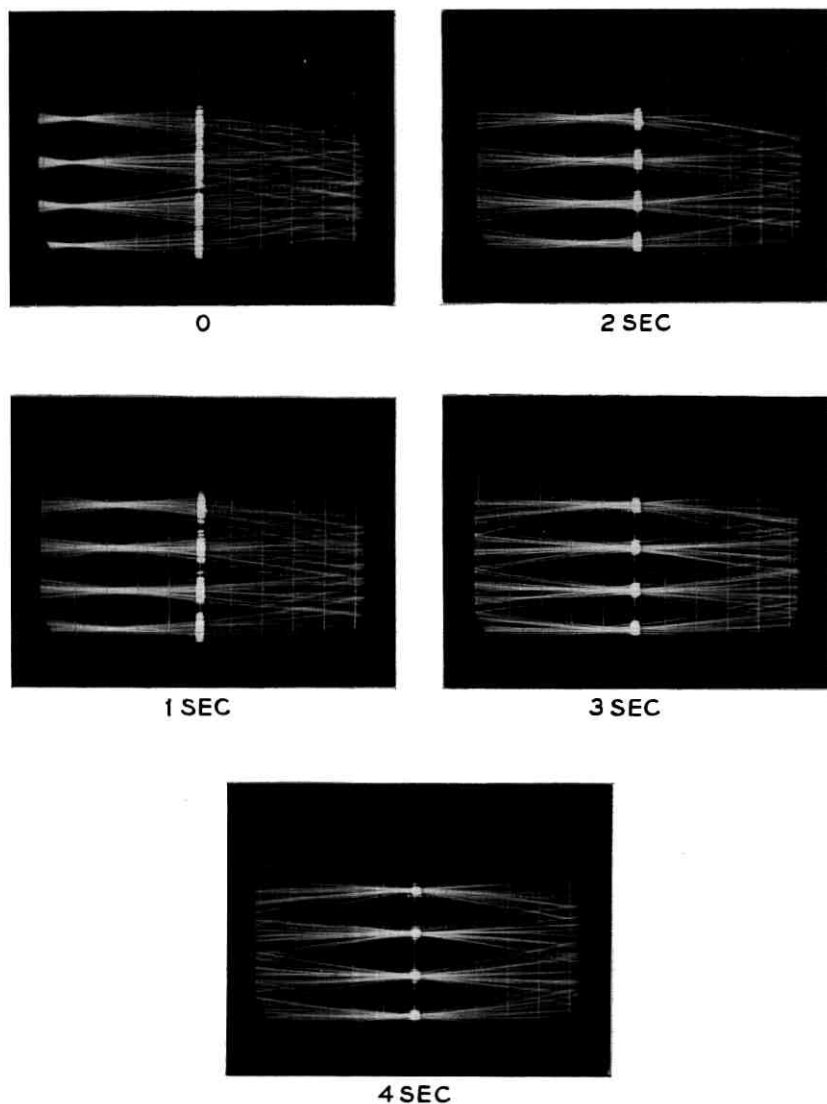At the positions shown in the first photos of Figs. 7 and 8 the equalizer

Fig. 7 — Four-level eye picture after timing is abruptly displaced.
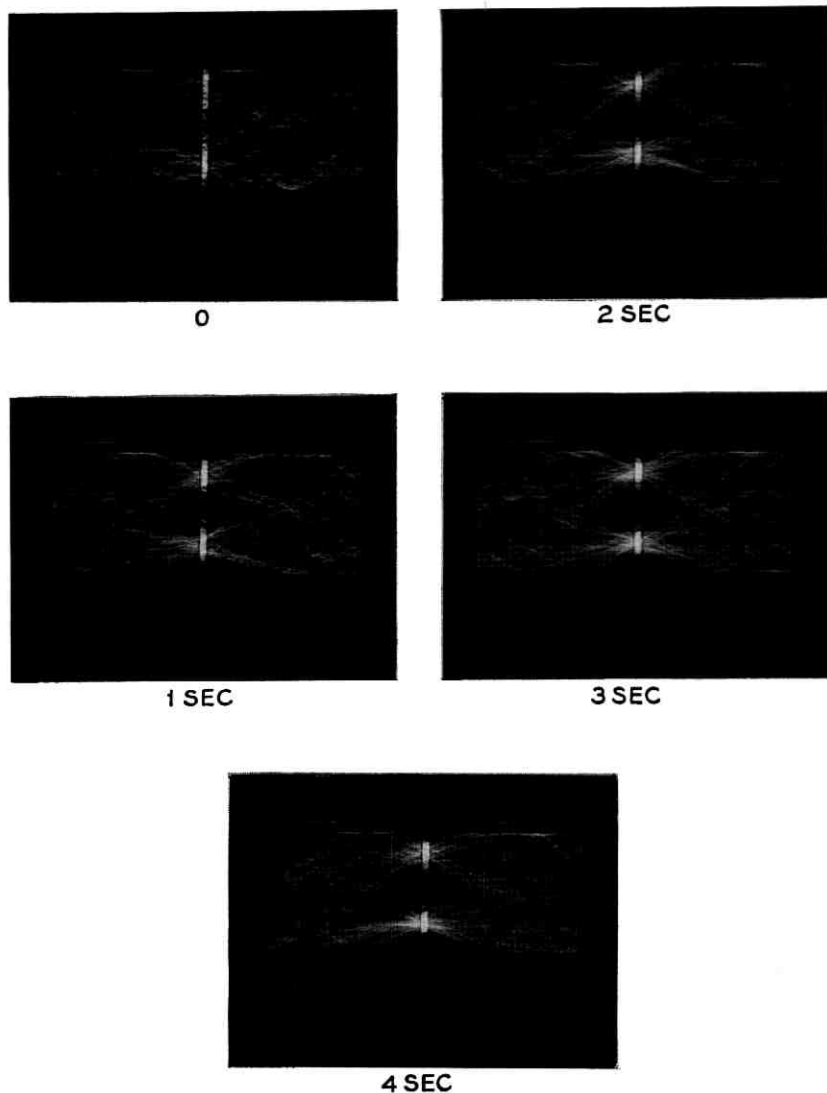
Fig. 8 — 16-level, folded eye picture after abrupt change in transmission characteristics.

acts quickly and decisively. It is possible to disturb the channel so badly that no semblance of an eye opening is left and the error probability is nearly 0.5. The mathematics of the equalizer operation here are quite complex and give little insight toward performance evaluation. It has

been observed experimentally that for binary (2-level) operation it has been impossible to find a setting or disturbance from which the equalizer will not eventually converge. Sometimes the eventuality takes as long as a minute as the equalizer makes slow (on the order of C^2) decisions — a sure indication of inaccuracy. After a period of what seems to be random hunting the equalizer reaches a position which it suddenly recognizes. The decisions come very quickly (on the order of C) and the eye appears seemingly from nowhere. Finally, perfect equalization is approached and the decisions once more come very slowly.

For higher level systems, 8- and 16-level, positions can be found from which the equalizer will not in all probability converge in a time that one is willing to wait around and watch. There seem to exist certain stable states where, for example, a 16-level eye pattern exists where 8-level transmission is being used — each level being split into two balanced levels. It should be emphasized that these conditions never occur during normal data transmission when the equalizer has an open eye to begin with and only has to track this eye through changing transmission characteristics. They are relevant, however, for the acquisition period if the equalizer is turned on without any setup period. Such procedure is not recommended for the higher level systems. (It will be obvious to the reader that it is possible to send a pattern of "outside" levels — i.e., a 2-level signal to start a higher level transmission. Also, it is possible to use quasi-random pattern generators, say maximal-length shift registers, at both transmitter and receiver to start transmission without worrying about transmission errors.)

A series of error performance runs was made on two test facilities — one looped via K-carrier to Boston from Holmdel, New Jersey and the other looped to Chicago via LMX-1 carrier from Holmdel. A number of 2 minute runs were made at 9600 bits-per-second at various times of day on each facility. In every case a control run of 2 minutes (1.152 × 10^6 bits) was made using preset equalization as described in Ref. 3. The results of these runs are plotted in Fig. 9 with preset error rate as the abscissa and adaptive error rate as the ordinate. In all cases, the error rate was diminished by the use of adaptive equalization. Even in the worst case the improvement factor was over three, while the best case represented an improvement factor of 50 and the average factor was approximately 10.

The importance of this improvement factor in the performance of this system cannot be overemphasized. While the order-of-magnitude improvement may not seem significant, it must be pointed out that this is raw error rate previous to error correction. If the curves relating customer error rate to raw error rate for the error control system used in
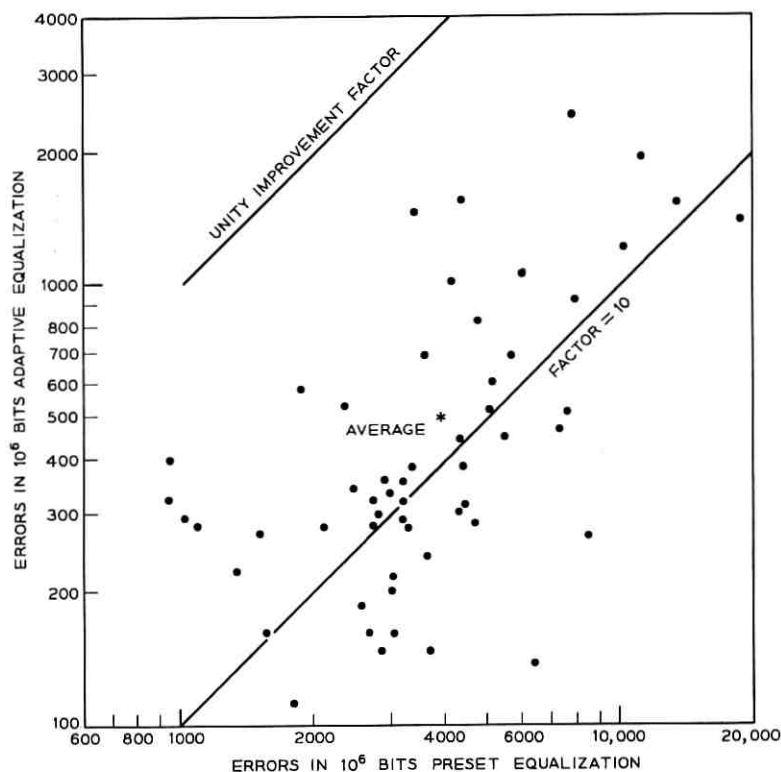
Fig. 9 — Error rates for adaptive and preset equalization on test runs.

conjunction with the VSB system which are shown in Ref. 4 are examined, it will be seen that the customer error rate is an extremely sensitive function of the raw error rate. An improvement of an order of magnitude in raw error rate results in an improvement of 2 or 3 orders of magnitude in the customer error rate. The customer error rates are not shown in Fig. 9 since in many cases the error rate for preset equalization was so high as to preclude synchronization, while frequently the adaptive error rate was too low to get any customer errors at all in the two-minute runs.

The question arises as to where the improvement comes from. First the improvement factor does *not* come from tracking the changing channel characteristics over the two-minute period. The error rate is virtually unchanged if the adaptive equalizer is turned off after a few

second's operation. The improvement comes from a number of other factors, chief of which is the improved setting accuracy. The preset equalizer operates for a period of 7 seconds during which 100 test pulses per second are transmitted. After this period an equilibrium distribution of equalizer positions has been reached in which the average tap error is about Δ, whereas with the adaptive equalizer 2400 symbols per second are used to extract information and the average tap error has been designed to be close to the minimum value of 0.5Δ.

The other factors, whose importance has not been quantitatively assessed, are the possibility of a nonlinearity in the transmission channel so that different characteristics are presented to the data signals than to test pulses, the possibility that the test pulses have some overlap, and the possibility of some bias in the examination of test pulse sample polarities. There seems to be no question that adaptive equalization using received data is superior to test pulse equalization.

Finally, an attempt was made to measure the time variation of some of the test lines. A critical measure of this variation can be achieved by averaging in low-pass filters the outputs of the exclusive-OR circuits in Fig. 4. The filter outputs are extremely sensitive indicators of the parameters h_j, although they are difficult to calibrate since the proportionality constant is a function of the noise variance [see (49)]. The recorder outputs from these filters are purposely not reproduced here lest the reader grant too much significance to the time variation records obtained. Generally, about a peak variation of one percent was found on the center 3 taps (h_{-1}, h_0, and h_1) and a negligible amount on other taps. The period of variation varied from about 10 seconds to a few minutes and then grew short again periodically. It was easily demonstrated that the variation was due to phase wander in the carrier system. Such slow phase wander was subsequently tracked using more sophisticated phase control apparatus.

A record of the exact setting reached by each tap gain control using adaptive equalizer setup was made periodically over a period of several weeks. Within statistical error the values reached were constant with the exception of the three central, phase-sensitive taps. (The reason these taps are phase sensitive is that the quadrature pulse ideally has nonzero values only at these points.)

The transmission facilities have been designed through the years to be relatively insensitive to temperature and humidity changes. It would seem that even for the most critical current data transmission usage, the telephone channel's phase and amplitude characteristics can be

assumed to be time invariant. Thus, it seemed that adaptive equalization was more useful as a refining device after test pulse equalization and as an insurance system rather than as a tracking equalizer for a significantly time varying channel. Nevertheless, the potentialities of the system considerably exceeded the demands of the environment and even so its benefits were quite strikingly apparent.

REFERENCES

1. Lucky, R. W., Automatic Equalization for Digital Communication, B.S.T.J., *44*, April, 1965, pp. 547–588.
2. Becker, F. K., An Exploratory, Multi-Level, Vestigial Sideband Data Terminal for Use on High Grade Voice Facilities, IEEE Annual Communications Convention, Boulder, Colorado, June 7–9, 1965.
3. Lucky, R. W., An Automatic Equalization System for Use with a High Speed, Voiceband Data Set, IEEE Annual Communications Convention, Boulder, Colorado, June 7–9, 1965.
4. Burton, H. O. and Weldon, E. J., An Error Control System for Use with a High Speed, Voiceband Data Set, IEEE Annual Communications Convention, Boulder, Colorado, June 7–9, 1965.
5. Feller, W., *An Introduction to Probability Theory and Its Applications*, John Wiley & Sons, New York, 1950.

Geometrical Representation of Gaussian Beam Propagation

By T. S. CHU

A single equation governing the various properties encountered in propagation of Gaussian beams is discussed. These characteristics may be graphically presented on an impedance chart such as a Smith chart or in the form of normalized curves. The geometrical representations highlight the relation between Gaussian mode theory and geometrical optics formulas.

The power coupling coefficient between two Gaussian modes suggests a complex mismatch coefficient whose geometrical representation is essentially the same as that of a complex reflection coefficient in transmission-line theory. Application of the Smith chart in determining a complex mismatch coefficient is illustrated by graphical solution of a beam-matching problem.

I. INTRODUCTION

The propagation of a Gaussian beam and its transformation through a lens has been well treated in previous literature.[1,2,3,4] This paper will show that a single, formally identical equation governs three properties of Gaussian beam propagation: (*i*) the phase front curvature and the beam radius in terms of the distance from the beam waist and the minimum beam radius; (*ii*) the propagation of a Gaussian beam in free space; (*iii*) the transformation of a Gaussian beam through a lens. Geometrical representations of these characteristics highlight the relationship between Gaussian beam propagation and geometrical optics.

Several recent papers[5,6,7,8] have been devoted to graphical solutions of Gaussian mode problems. One recalls that the Smith chart is a geometrical representation of complex reflection coefficient in transmission line theory. It seems logical, therefore, to look for the counterpart of a complex reflection coefficient in Gaussian mode theory in order that the full potential of the Smith chart may be realized in graphical solutions of Gaussian mode problems. In this paper, a complex mismatch coefficient will be defined such that the geometrical representation of this

coefficient is essentially the same as that of a complex reflection coefficient in transmission line theory. This observation immediately suggests the application of the Smith chart of complex mismatch coefficients to a graphical solution of beam matching problems.

II. THE ANALOGY AMONG THREE PROPERTIES OF A GAUSSIAN BEAM

In order to discuss the Gaussian beam transformation it is convenient to summarize the relationships among the parameters of a Gaussian beam propagating along the z-axis first.[4,9]

$$w(z) = \bar{w}\sqrt{1 + \left(\frac{\lambda z}{\pi \bar{w}^2}\right)^2} \qquad (1a)$$

$$r(z) = z\left[1 + \left(\frac{\pi \bar{w}^2}{\lambda z}\right)^2\right]. \qquad (1b)$$

In (1), w is the beam radius at which the field amplitude has fallen to $1/e$ of its maximum value on the z-axis, $\bar{w}$ is the minimum beam radius (called the beam waist) where one has a plane phase front at $z = 0$, and r is the radius of curvature of the phase front at z. It should be noted that the phase front is not exactly spherical; therefore, its radius of curvature is exactly equal to r only on the z-axis. The z-coordinate, measured from the beam waist, is taken to be positive to the right and negative to the left; the parameters are illustrated in Fig. 1. Equation (1) may be solved for $\bar{w}$ and z as follows:

$$\bar{w} = \frac{w}{\sqrt{1 + \left(\frac{\pi w^2}{\lambda r}\right)^2}} \qquad (2a)$$

$$z = \frac{r}{1 + \left(\frac{\lambda r}{\pi w^2}\right)^2}. \qquad (2b)$$

Equations (1) and (2) may be transformed into a single equation of complex variables:

$$\frac{1}{\pi w^2/\lambda} - i\frac{1}{r} = \frac{1}{(\pi \bar{w}^2/\lambda) + iz}. \qquad (3)$$

We are also interested in the transformation of a Gaussian beam going through a lens of focal length f as shown in Fig. 2. A beam with its minimum beam radius $\bar{w}_1$ located at d_1 will become another beam with its
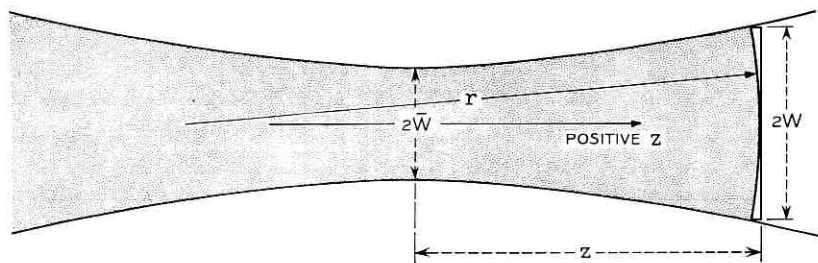
Fig. 1 — Parameters of a Gaussian beam.

minimum beam radius $\bar{w}_2$ located at d_2. Since the beam radius remains the same in passing through the lens we obtain from (1a)

$$\bar{w}_1 \sqrt{1 + (\lambda d_1/\pi \bar{w}_1{}^2)^2} = \bar{w}_2 \sqrt{1 + (\lambda d_2/\pi \bar{w}_2{}^2)^2}. \qquad (4)$$

The thin lens formula states that the change of the phase front curvature may be approximated by the reciprocal of the focal length. Thus, using (1b), one obtains
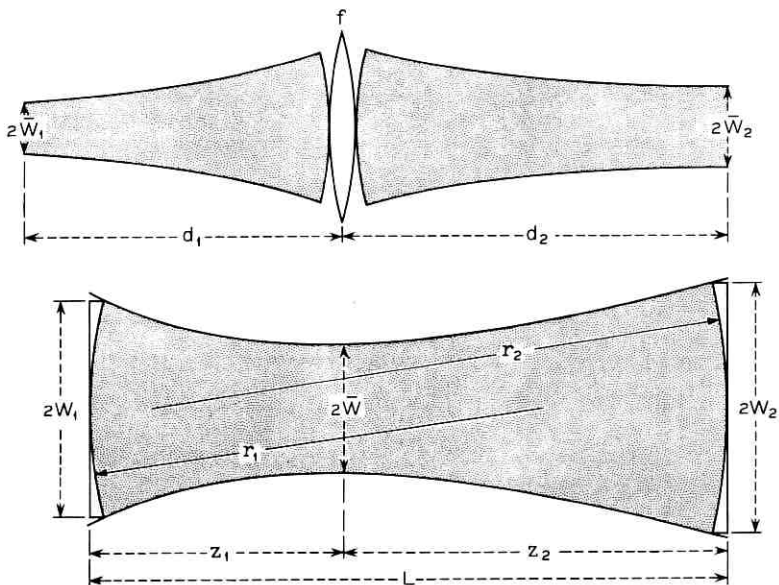

Fig. 2 — The analogy between the Gaussian beam transformation through a lens and the Gaussian beam propagation in free space.

$$\frac{1}{d_1[1 + (\pi\bar{w}_1^2/\lambda d_1)^2]} + \frac{1}{d_2[1 + (\pi\bar{w}_2^2/\lambda d_2)^2]} = \frac{1}{f} \tag{5}$$

where the sign of the second term on the left side is positive because d_2 has been taken to be positive to the left of the beam waist following the convention of Goubau[1] and Kogelnik.[2] Straightforward algebra will lead to the following solutions for (4) and (5):

$$\frac{d_2}{f} - 1 = \frac{\dfrac{d_1}{f} - 1}{\left(\dfrac{d_1}{f} - 1\right)^2 + \left(\dfrac{\pi\bar{w}_1^2}{\lambda f}\right)^2} \tag{6}$$

$$\frac{\bar{w}_2^2}{\bar{w}_1^2} = \frac{1}{\left(\dfrac{d_1}{f} - 1\right)^2 + \left(\dfrac{\pi\bar{w}_1^2}{\lambda f}\right)^2}. \tag{7}$$

The above two expressions are essentially rearrangements of formulas obtained by Goubau.[1] When $\pi\bar{w}_1^2/\lambda \ll |d_1 - f|$, (6) and (7) approach the thin lens formula and the magnification formula of geometrical optics.* When $\pi\bar{w}_1^2/\lambda f \to \infty$ the condition for geometrical optics focusing of parallel rays is obtained. These two equations have been plotted in Figs. 3 and 4 for various values of $p = \pi\bar{w}_1^2/\lambda f$. There it is seen that the singularity at $d_1 = f$ in geometrical optics is eliminated in Gaussian mode transformations. The maxima and minima in Fig. 3 may be easily found by differentiating (6) with respect to d_1/f, and they are given by

$$\frac{d_2}{f} = 1 \pm \frac{f}{2\pi\bar{w}_1^2/\lambda} \quad \text{when} \quad \frac{d_1}{f} = 1 \pm \frac{\pi\bar{w}_1^2}{\lambda f}. \tag{8}$$

The points of inflection are

$$\frac{d_1}{f} = 1 \quad \text{and} \quad 1 \pm \frac{\sqrt{3}\pi\bar{w}_1^2}{\lambda f}.$$

Equations (6) and (7) can be combined into one equation of complex variables:

$$\left(\frac{\pi\bar{w}_2^2}{\lambda f}\right) - i\left(\frac{d_2}{f} - 1\right) = \frac{1}{\left(\dfrac{\pi\bar{w}_1^2}{\lambda f}\right) + i\left(\dfrac{d_1}{f} - 1\right)}. \tag{9}$$

---

* Simple algebra may easily reduce them to the more familiar forms, $\dfrac{1}{d_1} + \dfrac{1}{d_2} = 1$ and $\dfrac{\bar{w}_2^2}{\bar{w}_1^2} = \dfrac{d_2^2}{d_1^2}$.
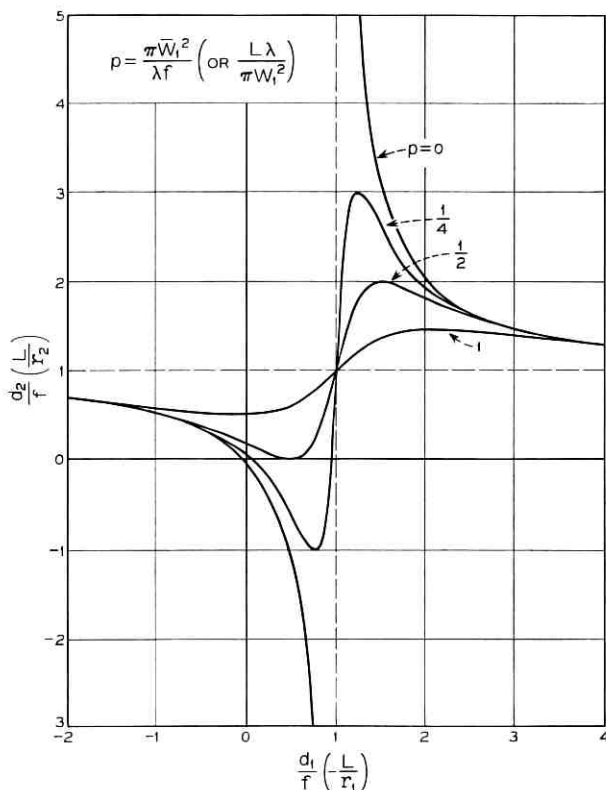
Fig. 3 — The graphical presentation of (5) and (14).

Next, let us consider the propagation of a Gaussian beam along the z-axis. One would like to express the beam radius w_2 and the curvature $1/r_2$ of the phase front at any given point in terms of the beam radius w_1 and the curvature $1/r_1$ at some other point along the axis. Since any Gaussian beam has a single beam waist at a definite location, one obtains from (2) the following conditions:

$$\frac{w_1}{\sqrt{1 + (\pi w_1^2/\lambda r_1)^2}} = \frac{w_2}{\sqrt{1 + (\pi w_2^2/\lambda r_2)^2}} \tag{10}$$

$$\frac{r_2}{1 + (\lambda r_2/\pi w_2^2)^2} - \frac{r_1}{1 + (\lambda r_1/\pi w_1^2)^2} = z_2 - z_1. \tag{11}$$

Let $L = z_2 - z_1$. Rearranging (10) and (11) yields

$$\frac{\dfrac{1}{\pi w_1^2/\lambda}}{\left(\dfrac{1}{\pi w_1^2/\lambda}\right)^2 + \left(\dfrac{1}{r_1}\right)^2} = \frac{\dfrac{1}{\pi w_2^2/\lambda}}{\left(\dfrac{1}{\pi w_2^2/\lambda}\right)^2 + \left(\dfrac{1}{r_2}\right)^2} \tag{12a}$$

and

$$-\frac{\dfrac{1}{r_1}}{\left(\dfrac{1}{\pi w_1^2/\lambda}\right)^2 + \left(\dfrac{1}{r_1}\right)^2} + \frac{\dfrac{1}{r_2}}{\left(\dfrac{1}{\pi w_2^2/\lambda}\right)^2 + \left(\dfrac{1}{r_2}\right)^2} = L. \tag{12b}$$

If we also rearrange (4) and (5) as follows:

$$\frac{\pi \bar{w}_1^2/\lambda}{(\pi \bar{w}_1^2/\lambda)^2 + d_1^2} = \frac{\pi \bar{w}_2^2/\lambda}{(\pi \bar{w}_2^2/\lambda)^2 + d_2^2} \tag{13a}$$

$$\frac{d_1}{(\pi \bar{w}_1^2/\lambda)^2 + d_1^2} + \frac{d_2}{(\pi \bar{w}_2^2/\lambda)^2 + d_2^2} = \frac{1}{f} \tag{13b}$$

the analogy* between (12) and (13) now emerges with the following one-to-one correspondences:

$$\frac{\pi \bar{w}_1^2}{\lambda} \leftrightarrow \frac{1}{\pi w_1^2/\lambda}$$

$$\frac{\pi \bar{w}_2^2}{\lambda} \leftrightarrow \frac{1}{\pi w_2^2/\lambda}$$

$$d_1 \leftrightarrow -\frac{1}{r_1}$$

$$d_2 \leftrightarrow \frac{1}{r_2}$$

$$\frac{1}{f} \leftrightarrow L.$$

Using the above analogy, one may immediately write down the solutions of (12) for w_2 and $1/r_2$

$$\frac{L}{r_2} - 1 = \frac{-\dfrac{L}{r_1} - 1}{\left(-\dfrac{L}{r_1} - 1\right)^2 + \left(\dfrac{L}{\pi w_1^2/\lambda}\right)^2} \tag{14}$$

---

* The analogy would look even better, if r_1 had been assumed positive to the left of the beam waist. This fact, of course, results from the sign convention of d_1, which is positive to the left of the lens.

$$\frac{L}{\pi w_2{}^2/\lambda} = \frac{L/\frac{\pi w_1{}^2}{\lambda}}{\left(-\frac{L}{r_1} - 1\right)^2 + \left(\frac{L}{\pi w_1{}^2/\lambda}\right)^2}. \tag{15}$$

Except for slightly different normalizations, these two expressions coincide with two formulas obtained by Rowe[10] using rather involved algebra. The curves in Figs. 3 and 4 represent (14) and (15), if the coordi-
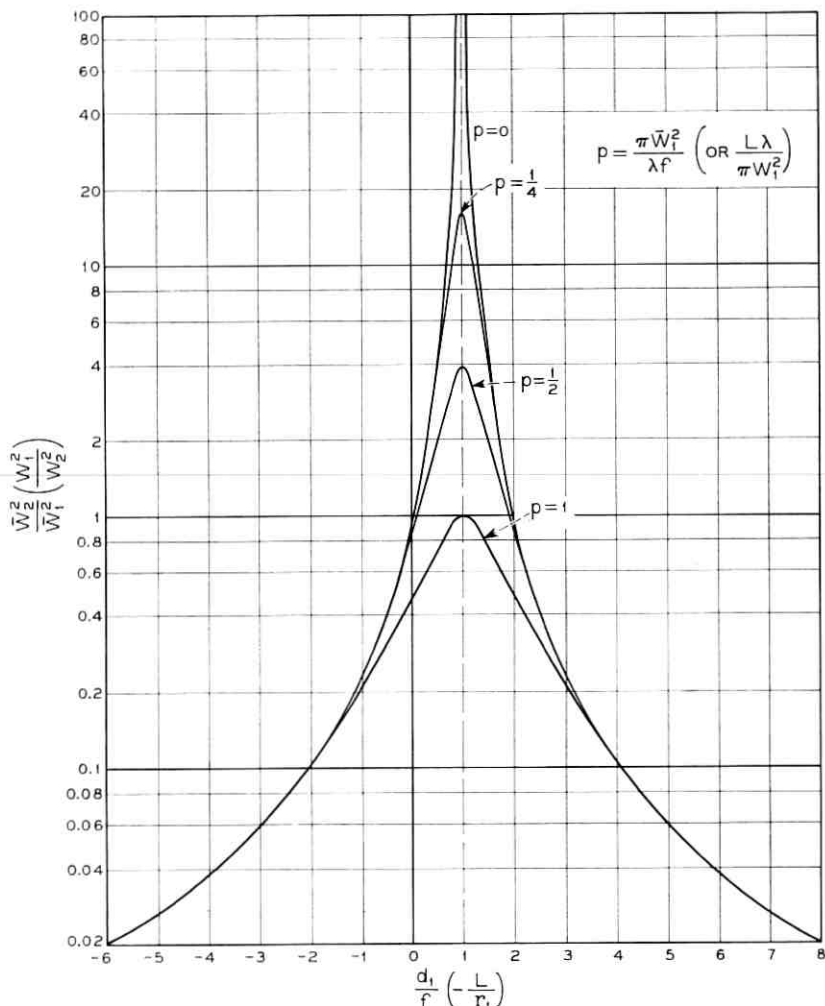

Fig. 4 — The graphical presentation of (6) and (15).

nates and the parameters are replaced by those shown in parentheses in the figures. When the wavelength approaches zero, the Gaussian beam becomes the spherical wave of geometrical optics. The singularity of the spherical wave is eliminated in the Gaussian beam representation. Equations (14) and (15) can also be combined into one equation of complex variables:

$$\frac{L}{\pi w_2{}^2/\lambda} - i\left(\frac{L}{r_2} - 1\right) = \frac{1}{\left(\dfrac{L}{\pi w_1{}^2/\lambda}\right) + i\left(-\dfrac{L}{r_1} - 1\right)}. \tag{16}$$

It is interesting to observe that (3), (9), and (16) are formally identical, and they are amenable to graphical representation on an impedance chart, e.g., a Smith chart. Thus, the circumference of the unit circle on a Smith chart corresponds to geometrical optics, while the interior describes the properties of a Gaussian beam. The transformation may be performed by simply taking diametrically opposite points on a Smith chart. It should be noted that while (3) has not been normalized, (9) and (16) are dimensionless. Representation of (3) by the Smith chart has also been proposed by Deschamps and Mast;[7] the analogy between (9) and (16) corresponds to the dual forms of the cartesian Gaussian beam chart.[5,6] Rearrangement of (14) and (15) will yield equations similar to those discussed in Ref. 8.

III. THE COMPLEX MISMATCH COEFFICIENT DIAGRAM

Using some of the results of the previous section, a graphical solution of mode matching problems using a properly defined complex mismatch coefficient will now be discussed. Equation (3) can be rewritten as

$$\frac{\pi \bar{w}^2}{\lambda} - iz = \frac{1}{\left(\dfrac{1}{\pi w^2/\lambda}\right) + i\dfrac{1}{r}} \tag{19}$$

and one may adopt the following notations:

$$\frac{1}{\pi w^2/\lambda} + i\frac{1}{r} = R + iX = Z \tag{20a}$$

$$\frac{\pi \bar{w}^2}{\lambda} - iz = G - iB = Y. \tag{20b}$$

Thus, a Gaussian beam may be characterized by either a beam impedance Z or a beam admittance Y. Now the following interesting identity can easily be verified.

$$\left| \frac{Z_1 - Z_0}{Z_1 + Z_0{}^*} \right|^2 = \left| \frac{Y_1 - Y_0}{Y_1 + Y_0{}^*} \right|^2 = 1 - \tau \qquad (21)$$

where the subscript 1 represents the incoming beam, 0 represents the fundamental mode of a receiving system,* denotes complex conjugate, and τ is the power coupling coefficient[11] between two Gaussian modes. The first two expressions in (21) certainly look like reflection coefficients in transmission lines. This observation suggests the designation of $(Z_1 - Z_0)/(Z_1 + Z_0{}^*)$ and $(Y_1 - Y_0)/(Y_1 + Y_0{}^*)$ as complex mismatch coefficients. They are indeed identical to reflection coefficients in form except for the complex conjugate in the denominator. If Z_1 (or Y_1) is normalized with respect to a real Z_0 (or a real Y_0), then the complex mismatch coefficient is formally identical to a complex reflection coefficient. Both the complex mismatch coefficient and the complex reflection coefficient are bilinear transformations of the impedance or admittance.

The Smith chart is a geometrical representation of the complex reflection coefficient. The same Smith chart may also represent the complex mismatch coefficient. With respect to a reference mode, any phase front of a Gaussian beam may be represented by two diametrically opposite points on a Smith chart, i.e., the normalized admittance and the normalized impedance. The propagation of a Gaussian beam corresponds to travel along a conductance circle, while imposing change of radius of curvature by a lens, say, corresponds to travel along a resistance circle. The pair of two diametrically opposite points may be reduced to one point by imposing a flip-over Smith chart on the original Smith chart, however, overcrowded coordinate lines are not desirable in practical graphical solutions. This latter version of geometrical representations corresponds to a bilinear transformation of the Gaussian beam chart discussed in Ref. 5, and is the same as that suggested in Ref. 7. It should be noticed that Collins' Gaussian beam chart[5] is a geometrical representation of the beam impedance, while Li's dual form[6] is that of the beam admittance. Here, however, the Smith chart is utilized as a complex mismatch coefficient diagram and its relationship to the power coupling coefficient between two Gaussian modes is identified.

IV. A NUMERICAL EXAMPLE OF BEAM MATCHING

In order to illustrate the application of the Smith chart, consider the problem of matching the output beam of an optical maser to an interferometer as shown in Fig. 5. The maser resonator consists of mirrors of 5 and 10-meter radius of curvature separated by 1 meter, and the interferometer consists of a pair of 10-meter mirrors separated by 10 cm.
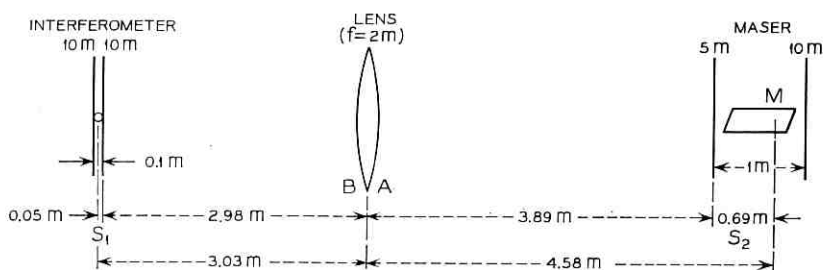
Fig. 5 — The geometry of a beam matching problem.

The maser beam exits from the 5-meter mirror and a lens with a focal length of 2 meters is available for matching. First, one calculates the minimum beam half-widths and their locations for both resonators by the following formulas:[4]

$$\frac{\pi \bar{w}^2}{\lambda} = \frac{\sqrt{d(R_1 - d)(R_2 - d)(R_1 + R_2 - 2d)}}{R_1 + R_2 - 2d} \qquad (22)$$

$$S = d \frac{R_2 - d}{R_1 + R_2 - 2d} \qquad (23)$$

where R_1 and R_2 are the radii of curvature of the mirrors, d is mirror separation, $\bar{w}$ is the minimum beam half-width, and S is the location of the beam waist from the mirror of radius R_1. Now we readily get $\pi \bar{w}_1^2/\lambda = 0.705$ m, $S_1 = 0.05$ m for the interferometer, and $\pi \bar{w}_2^2/\lambda = 1.727$ m, $S_2 = 0.692$ m for the maser resonator. If all the parameters are normalized with respect to $\pi \bar{w}_1^2/\lambda$, then the fundamental mode of the interferometer is represented by the unity conductance circle passing through the center of Smith chart in Fig. 6 where 0 corresponds to the beam wasit for the interferometer. The maser output beam is represented by the 2.45 $(= \bar{w}_2^2/\bar{w}_1^2)$ conductance circle passing through M which corresponds to the beam waist for the maser resonator. The inversions of these two circles with respect to the center 0 give the beam radius and the phase front curvature along the propagation path of the beam.

Let it be required that insertion of a 2-meter focal length lens changes the phase-front curvature of the maser output beam to that of the interferometer beam at the point where the two beams meet. Since the beam radius remains constant going through the lens, application of the thin lens formula demands that a segment of the beam resistance circle between the inverted maser and interferometer circles be equal to the length $\widehat{A'B'} = \pi \bar{w}_1^2/\lambda f = (\pi \bar{w}_1^2/\lambda)(1/r_1 - 1/r_2) = 0.3523$. After deter-
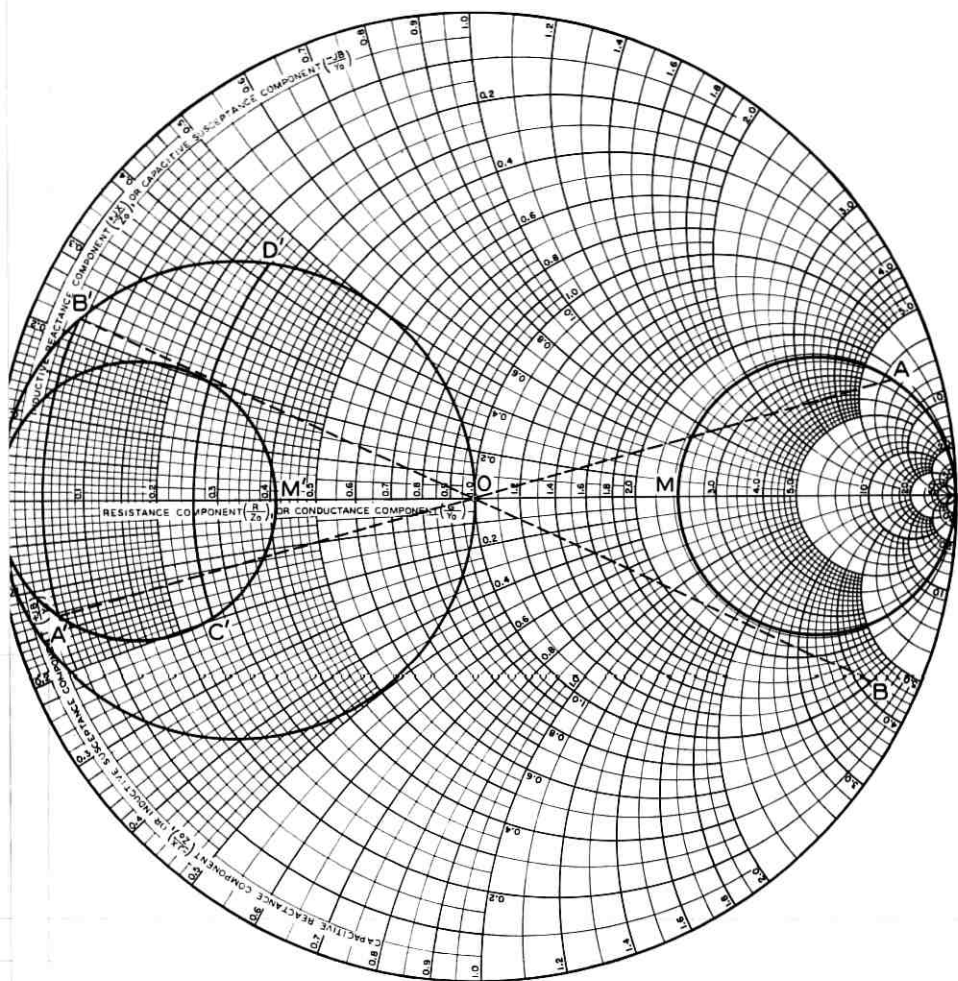
Fig. 6 — The graphical solution of a beam matching problem.

mining the two pairs of diametrically opposite points $A-A'$ and $B-B'$ (Fig. 6), the distances from the lens to the interferometer beam waist and to the maser beam waist can be read off the chart as $\widehat{OB} = 4.3 \times 0.705 = 3.03$ m and $\widehat{MA} = 6.5 \times 0.705 = 4.58$ m. The distances from the lens to the interferometer input mirror and to the maser output mirror are $3.03 - 0.05(S_1) = 2.98$ m and $4.58 - 0.69(S_2) = 3.89$ m, respectively.

V. DISCUSSION

The advantage of this complex mismatch coefficient diagram over the Gaussian beam charts discussed in Refs. 5 and 6, is similar to that of the Smith chart over cartesian impedance and admittance charts. Neither of the latter can represent an arbitrary transformation of a Gaussian beam, whereas the present mismatch coefficient chart, at least in principle, encompasses all possible transformations. Curvilinear coordinates are

Fig. 7 — An example of approximate beam matching.

not a problem here because a good network of circles is readily available in the form of Smith charts. Sometimes it may be impossible to achieve perfect beam matching due to limitations imposed by the inadequacy of a lens or of space. One is able to minimize the mismatch using the procedure discussed here.

In the above example, there exists a minimum focal length of the lens beyond which perfect matching is not possible. In Fig. 6, this minimum is represented by the maximum arc length $\overset{\frown}{C'D'}$ (0.64) of the beam resistance circle between the inverted maser and interferometer circles which corresponds to a focal length of $\pi \bar{w}_1^2 / \lambda \overset{\frown}{C'D'} = 0.705/0.64 = 1.102$ m. If one had only a one meter focal length lens available, $\overset{\frown}{C'D'}$ would be extended to E' as shown in Fig. 7 and the reciprocals of C' and E' are then found to be C and E. The distances from the lens to the maser output mirror and to the interferometer input mirror would be $\overset{\frown}{CM} - S_1 = 1.8 \times 0.705 - 0.05 = 1.22$ m and $\overset{\frown}{EF} - S_2 = 1.56 \times 0.705 - 0.69 = 0.41$ m for a resulting mismatch loss $|\overline{OF}|^2 = 0.015$.

VI. ACKNOWLEDGMENT

REFERENCES

1. Goubau, G., Optical Relations for Coherent Wave Beams, Proc. of 1962 Symposium on EM Theory and Antennas, Copenhagen, part 2, pp. 907–918.
2. Koglnik, H., Matching of Optical Modes, B.S.T.J., *43*, Jan., 1964, pp. 334–337.
3. Kogelnik, H., Imaging of Optical Modes and Resonators with Internal Lenses, B.S.T.J., *44*, March, 1965, pp. 455–494.
4. Kogelnik, H., Modes in Optical Resonators, in *Advances in Lasers* edited by A. K. Levine, Dekker Publishers, New York.
5. Collins, S. A., Jr., Analysis of Optical Resonators Involving Focusing Elements, Appl. Optics, *3*, Nov., 1964, pp. 1263–1275.
6. Li, T., Dual Forms of the Gaussian Beam Chart, Appl. Optics, *3*, Nov., 1964, pp. 1315–1317.
7. Deschamps, G. A. and Mast, P. E., Beam Tracing and Applications, Proc. of PIB Quasi-Optics Symposium, 1964.
8. Gordon, J. P., A Circle Diagram for Optical Resonators, B.S.T.J., *43*, Nov., 1964, pp. 1826–1827.
9. Yariv, A. and Gordon, J. P., The Laser, Proc. IEEE, *51*, Jan., 1963, pp. 4–29.
10. Rowe, H. E., unpublished work.
11. Kogelnik, H., Coupling and Conversion Coefficients for Optical Modes, Proc. of PIB Quasi-Optics Symposium, 1964.

The Attenuation of 3.392μ He-Ne Laser Radiation by Methane in the Atmosphere

By T. S. CHU and D. C. HOGG

Measurements of the propagation of infrared waves at 3.392μ and 3.508μ in clear weather have been carried out over a 2.6-km path of atmosphere at Holmdel, New Jersey. The measuring system employed antennas and detectors common to the two wavelengths. The excess attenuation at 3.392μ, after considering various corrections, is interpreted as absorption by methane in the atmosphere and has been found to be 5.5 (±0.5) db/km.

I. INTRODUCTION

The proximity of the 3.392μ He-Ne laser emission to one of the absorption lines of methane[1,2] predicts that this laser radiation may be somewhat attenuated by the earth's atmosphere. The characteristics of this attenuation are interesting not only in evaluating the potential use of the 3.392μ line for communications, but also because of possible application in measurement of the methane content in the atmosphere. An estimate of the absorption of 3.392μ laser emission by methane in the atmosphere has appeared in the literature.[1] The purpose here is to present a measurement which compares propagation of infrared waves at 3.392μ and 3.508μ over an atmospheric path length of 2.6 km. The differential transmittance at these two wavelengths, after accounting for the effect of water vapor, is used to determine the attenuation of the 3.392μ wave by methane in the atmosphere.

II. EXPERIMENTAL EQUIPMENT

The profile and terrain of the 2.6-km propagation path are illus-

trated in Fig. 1. The experimental site at Holmdel, N.J. is about 5 km south of the shore of Raritan Bay on the Atlantic Coast.

The transmitting and receiving equipment, located 150 feet and 370 feet above sea level, respectively, are sketched in the block diagram of Fig. 2. The transmitters consist of a 3.392μ He-Ne laser and a 3.508μ He-Xe laser; the laser outputs are chopped simultaneously at 330 cps. The two laser beams are aligned into a common path by means of a reflecting mirror and a beam splitter which in turn feeds a bisected near-field Cassegrainian telescope. This telescope consists of two bisected confocal paraboloids as shown in Fig. 2. The main reflector is one half of a 15-inch diameter spin-cast paraboloid of 12-inch focal length while the parabolic subreflector has a diameter of 0.5 inch and a focal length of 0.4 inch. In addition to desirable properties, such as good preservation of mode composition (a consequence of negligible aperture blocking), short focal length, and mechanical simplicity, this particular telescope design offers good impedance match which is especially important in receivers using high-gain laser amplifiers as described below. The concentrated beams from the subreflector of the receiving telescope pass through a He-Ne tube, which is a 3.392μ amplifier, in tandem with a He-Xe tube, which is a 3.508μ
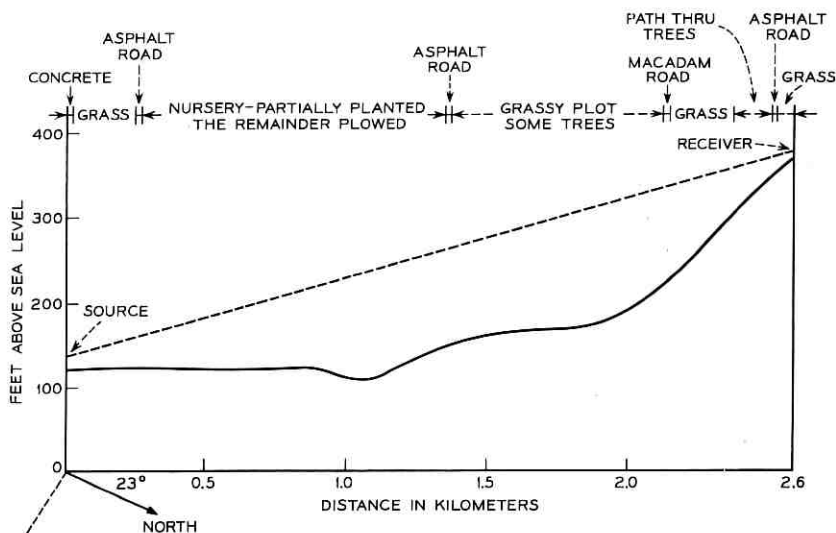

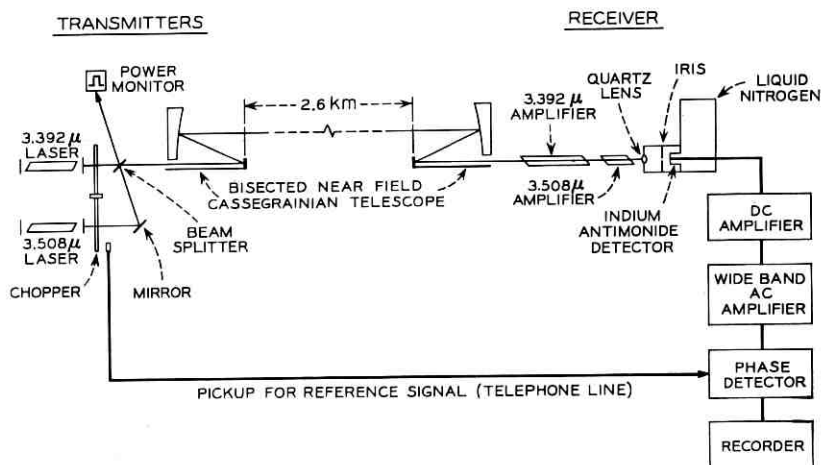Fig. 1 — Profile of transmission path.

Fig. 2 — Experimental equipment block diagram.

amplifier.* After these two amplifier tubes, the signal passes through a calcite polarizer, a quartz lens, and an iris, to a liquid nitrogen-cooled indium-antimonide photovoltaic infrared detector whose output is fed into a synchronous detection system. The comparison measurement is carried out by blocking one of the two laser outputs every other five minutes. The signal-to-noise ratios using a detector time constant of 2 sec are about 38 db for 3.508μ, at a laser amplifier gain of 15 db, and 23 db for 3.392μ, at a laser amplifier gain of 25 db.

III. MEASURED RESULTS

The measured data are summarized in Table I.

Having taken into account all conceivable correction factors, the unexplained excess attenuation of the 3.392μ signal over the 3.508μ signal is interpreted as the absorption by methane in the atmosphere. These factors, given in Table I, are determined as follows.

The difference in the laser outputs is obtained by a monitoring detector which is a biased lead-selenide photoconductor. The beam splitter (Fig. 2) transmits 20 per cent of the 3.392μ power while it reflects 60 per cent of the 3.508μ power into the telescope. The measured half-power widths of the transmitting beams are about 50 sec of arc. How-

* The He-Ne plasma introduces negligible loss to the 3.508μ wave; the same is true of the effect of the He-Xe plasma on the 3.392μ wave.

TABLE I — THE COMPARISON DATA

Wavelength	3.392μ	3.508μ Amplifier off	3.508μ Amplifier on
Laser output	+4 dbm	+7 dbm	+7dbm
Beam splitter loss	−7 db	−2.2 db	2.2 db
Diffraction loss	−(D + 1) db	−D db	−D db
Water vapor attenuation	−3 db	−1 db	−1 db
Amplifier gain used	25 db	0 db	15 db
Expected signal level	(18 − D) dbm	(3.8 − D) dbm	(18.8 − D) dbm
Expected signal ratio	$p_{3.39}/p_{3.5} =$	14.2 db	−0.8 db
Measured signal ratio	$p_{3.39}/p_{3.5} =$	0 db	−15 db

ever, the 3.508μ measured transmitting beamwidth is slightly narrower than at 3.392μ such that one expects about one db higher signal level at 3.508μ, this small difference in transmitting beamwidths being attributed to a slight difference in the laser beams and to imperfection in the laser-telescope alignment at 3.392μ. The diffraction loss, D, includes beam spread, telescope imperfections, and any insertion loss introduced by the amplifier tubes. The diffraction loss for the two wavelengths should differ only by the aforesaid one db; its absolute magnitude has not yet been determined with sufficient accuracy thus no numerical value is indicated for this item in the table. Of course, since the measurement is one of comparison, the absolute value is of no great significance here.

For the measurements under discussion, the water vapor density in the atmosphere was 10 gm/m³ (from the measured temperature and relative humidity, 68°F and 50°RH). Then the attenuation due to the water vapor content of the atmosphere, taken from standard tables[3], is 3 db and 1 db (total path) for 3.392μ and 3.508μ, respectively.

The measured gain of the 3.392μ amplifier tube is 25 db. However, it turned out, as shown by the second column in Table I, that the measured average signal level of 3.392μ with its amplifier tube operating was very closely equal to that of 3.508μ with its amplifier tube turned off. It follows that the over-all measured attenuation difference attributed to methane absorption is 14.2 db for the 2.6-km propagation path; this value, obtained on June 3 and 4, 1965, is typical of the measured data. The data distribution from several comparison tests and error estimates of the experimental equipment lead to the conclusion that the absorption of 3.392μ by methane in the atmosphere is 14.2/2.6 = 5.5 (±0.5) db/km.

IV. DISCUSSION

The above measured attenuation of 5.5 db/km agrees fairly well with the estimated transmittance $T = \exp(-1.1L)$ given by Edwards and Burch,[1] where L is the path length in kilometers. Their transmittance prediction was arrived at by laboratory measurements of the transmission of the 3.392μ He-Ne laser emission through different mixtures of methane and nitrogen in a cell, and by estimating the amount of methane (about 1 part in 10^6) in the atmosphere.[4] Edwards and Burch conclude from their measurements that the collision-broadened width of the absorption line of methane in the atmosphere is 3.9 (±1.2) kmc and the separation of the 3.392μ emission line from the center of the closest methane absorption line is 0.09 (±0.06) kmc. In Ref. 2, the methane absorption line closest to 3.392μ (2947.903 cm^{-1} or 88437.09 kmc) is given as 2947.888 (±0.015) cm^{-1} or 88436.64 (±0.45) kmc. It should be pointed out that Edwards and Burch only determined the separation between the 3.392μ emission line and the closest methane absorption line but not the absolute position of the 3.392μ emission line.

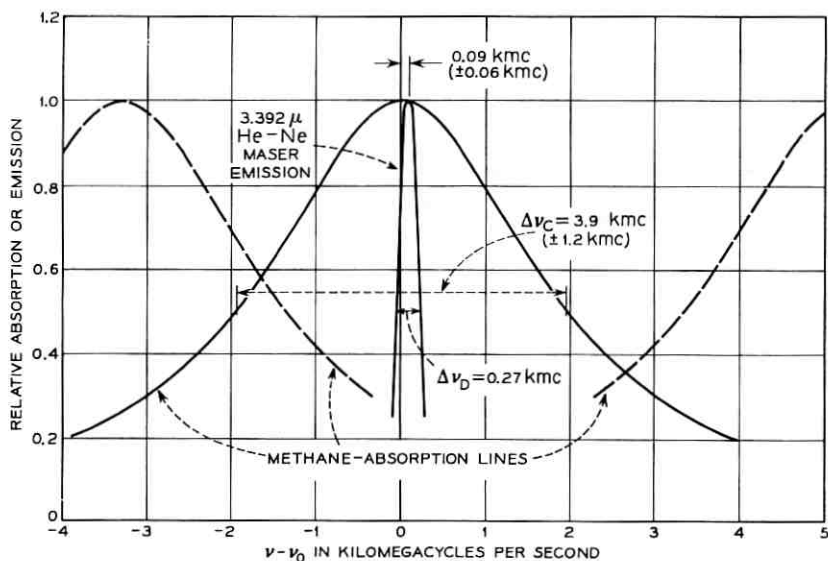The Doppler-broadened width (0.27 kmc) of the 3.392μ emission

Fig. 3 — Doppler broadening of the 3.392μ (88437.09 kmc) He-Ne Laser emission and collision broadening of the absorption lines of methane in the atmosphere.

line is small compared with the collision-broadened width of the absorption line of methane in the atmosphere and lies close to the center of the 2947.888 cm^{-1} methane absorption line as illustrated in Fig. 3. Therefore, the uncertainty of the cavity resonance within the Doppler broadening of the laser emission should not contribute significant error to the measurement of absorption of 3.392μ laser emission by methane in the atmosphere.

Consideration has also been given to magnetic field tuning of the laser oscillator as a possible scheme for investigating the profile of the methane absorption line. Fork and Patel[5] reported a maximum splitting of 11.84 kmc for 3.392μ emission using a magnetic field of 4 kilogauss. A frequency shift of 11.84 kmc should pull the 3.392μ laser line out of the collision-broadened width of the closest absorption line of methane in the atmosphere; however, one will run into the neighboring lines of the $P(7)$ branch of the methane absorption spectrum which are shown as dashed lines in Fig. 3.

V. ACKNOWLEDGMENT

REFERENCES

1. Edwards, B. N. and Burch, D. E., Absorption of 3.39 Micron Helium-Neon Laser Emission by Methane in the Atmosphere, J. Opt. Soc. Amer., 55, February, 1965, pp. 174–177.
2. Plyler, E. K., et al, Infrared Absorption Spectrum of Methane from 2470–3200 cm^{-1}, J. Res. NBS, 64A, May–June, 1960, pp. 201–212.
3. U. S. Air Force, Handbook of Geophysics, The MacMillan Co., 1960.
4. Fink, U., et al. Abundance of Methane in the Earth's Atmosphere, J. Opt. Soc. Amer., 54, April, 1964, pp. 472–474.
5. Fork, R. L. and Patel, C. K. N., Magnetic Field Tuning of Gaseous Laser Oscillators, Proc. IEEE, 52, February, 1964, pp. 208–209.

The Effect of a Finite-Width Decision Threshold on Binary Differentially Coherent PSK Systems

By W. M. HUBBARD

A binary PCM regenerator must decide which of two signal states was transmitted. This decision is normally based on whether some particular voltage is above or below a certain threshold at some particular time. In a real regenerator, the voltage in question must differ from this threshold value by some finite amount in order to cause the device to respond properly. The purpose of this calculation is to examine the effect of this "dead zone" on the probability of error in a differentially coherent PSK system both in the case where the signal is limited in amplitude (after the noise is added) and when it is not.

It is found that when the expected value of signal power, S, exceeds the threshold value, T, by more than 6 db and 10 db in the limited and unlimited cases, respectively, the effect on error probability is less than the effect of a 1-db degradation in signal-to-noise ratio. However, for smaller values of signal-to-threshold (S/T) the degradation becomes large. Numerical values for error probability as a function of signal-to-noise ratio are presented for $S/T = \infty$, 12, 9, 6, 4, and 3 db for both cases.

I. INTRODUCTION

A binary differentially coherent phase shift keyed (DCPSK) system is one in which the two (binary) states — "mark" and "space" — are transmitted as phase changes between adjacent time slots. In such a system, optimum results are obtained when the expected values of the two possible signal states in a given time slot differ in phase by π radians. For clarity it will be assumed in the following discussion that the expected value of phase change is 0 or π, respectively, according to whether a space or a mark was transmitted, even though phase changes of $\theta + 0$ and $\theta + \pi$ (θ an arbitrary but fixed angle) are equally suitable.

Such a system usually consists of a limiter, followed by a storage mechanism and a product demodulator, as shown in Fig. 1. If the signals in adjacent time slots have amplitudes A and B, respectively, and a phase difference ψ, the output of the product demodulator is proportional to

$$v = AB \cos \psi. \tag{1}$$

If the limiter is ideal, $A = B = 1$. The ideal regenerator samples the sign of v and regenerates a mark if $v < 0$ and a space if $v > 0$.

The probability that, due to additive Gaussian noise, $v > 0$ if a mark was sent or equivalently the probability that $v < 0$ if a space was sent has been calculated by many authors.[1-4] This probability is given by

$$\Pi = \tfrac{1}{2} \exp \left(-S/N\right) \tag{2}$$

where S/N is the signal-to-noise ratio.

In a practical system, the regenerator requires some finite value of v in order to regenerate a mark or a space reliably. While it is true that the effects of this finite-width decision level can be made arbitrarily small by providing sufficient gain ahead of the regenerator and sufficient dynamic range for the regenerator, it is often difficult — especially with some solid-state devices — to achieve either this gain or the necessary dynamic range. For this reason, it may be desirable to operate a regenerator under conditions where the decision threshold is important. In this paper we consider the following model:

(i) An error is made if $|v| > \varepsilon$ and, due to additive Gaussian noise, v has the wrong sign.

(ii) No error is made if $|v| > \varepsilon$ and v has the correct sign.

(iii) There is a 0.5 probability of error if $|v| < \varepsilon$, regardless of the sign of v.

The probability of error in such a system is clearly dependent on both the signal-to-noise ratio S/N and the signal-to-threshold ratio S/T. The latter is defined as the ratio of the expected value of signal power
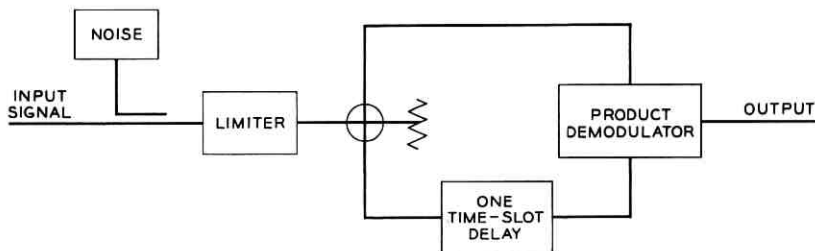

Fig. 1 — DCPSK receiver.

to the minimum value which causes the output of the product demodulator, v, to exceed ε in magnitude.

II. SYSTEMS WITHOUT A LIMITER

If the regenerator were ideal ($\varepsilon = 0$), the performance would be independent of whether or not the limiter in Fig. 1 is included, since the sign of v is independent of the product AB (both A and B are positive numbers). However, for the case where $\varepsilon > 0$, the limiter plays an important role under certain conditions. In this section, we consider the case where the limiter is omitted. The received signal can then be thought of as the vector sum of a unit vector representing the transmitted signal and a noise vector with Gaussian-distributed x and y components. The angle between the two vectors representing received signals one time slot apart together with their magnitudes A and B determine v. This is illustrated in Fig. 2 for a "mark". From Fig. 2, it is apparent that

$$v = AB \cos \psi = \mathbf{A} \cdot \mathbf{B} = A_x B_x + A_y B_y$$
$$= (A_o + a_x)(B_o + b_x) + a_y b_y \tag{3}$$

where $A_o = +1$ or -1 and $B_o = +1$ or -1 depending on the message. Cos ψ has the wrong sign if the sign of v differs from the sign of $A_o B_o$.

There will be two probability density functions $p_-(v)$ and $p_+(v)$, respectively, for the cases $A_o B_o = -1$ and $A_o B_o = +1$. From symmetry, it is apparent that the error probability is the same in these two cases. The probability of error can thus be determined from either

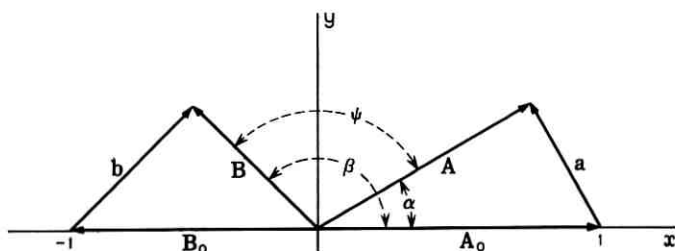$$\Pi = \frac{1}{2} \int_{-\varepsilon}^{\varepsilon} p_-(v) dv + \int_{\varepsilon}^{\infty} p_-(v) dv \tag{4}$$

or

$$\Pi = \frac{1}{2} \int_{-\varepsilon}^{\varepsilon} p_+(v) dv + \int_{-\infty}^{-\varepsilon} p_+(v) dv. \tag{5}$$

Bennett and Salz[3] derive the density functions $p_{\pm}(v)$ for the quadratic form of (3) where a_x, a_y, b_x, and b_y are independent Gaussian variables of zero mean and variance σ^2. They are given (in our notation) by

$$p_{\pm}(v) = \frac{1}{(2\pi)^{3/2}\sigma^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\exp\left[-(x^2 + y^2)/2\sigma^2\right]}{\sqrt{(1 + x)^2 + y^2}}$$
$$\cdot \exp\left(-\frac{[v \mp (1 + x)]^2}{2\sigma^2[(1 + x)^2 + y^2]}\right) dx \, dy. \tag{6}$$

σ^2 is the mean noise power.

A_0, B_0 = UNCORRUPTED SIGNALS

a, b = NOISE ON A, B RESPECTIVELY

A, B = RECEIVED SIGNALS

Fig. 2 — Phasor diagram of a pair of received signals.

In Appendix A it is shown that substituting (6) into (4) or (5) gives

$$\Pi = \frac{1}{8\pi\sigma^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left(-\frac{(x-1)^2 + y^2}{2\sigma^2}\right) \left\{ \operatorname{erfc} \frac{\varepsilon + x}{\sqrt{2}\,\sigma\,\sqrt{x^2 + y^2}} \right. \\ \left. + \operatorname{erfc} \frac{-\varepsilon + x}{\sqrt{2}\,\sigma\,\sqrt{x^2 + y^2}} \right\} dx\ dy. \tag{7}$$

This integral must be computed by numerical methods. The results of these computations are shown in Fig. 3.

III. SYSTEMS WITH AN IDEAL LIMITER

If an ideal limiter* is included in the system as shown in Fig. 1, (3) reduces to

$$v = \cos\psi. \tag{8}$$

In this case, it is simpler to perform the calculation of error probability from a consideration of the probability density of ψ itself. From Fig. 4 and the criteria for error listed in Section I, one obtains

$$\Pi = \frac{1}{2} \int_{\pi + \cos^{-1}\varepsilon}^{\pi + \cos^{-1}(-\varepsilon)} p(\psi)d\psi + \int_{\pi + \cos^{-1}(-\varepsilon)}^{\cos^{-1}\varepsilon} p(\psi)d\psi \\ + \frac{1}{2} \int_{\cos^{-1}\varepsilon}^{\cos^{-1}(-\varepsilon)} p(\psi)d\psi \tag{9}$$

---

* By an ideal limiter is meant a device which removes all amplitude variation from the signal without affecting the phase. That is, it transforms the signal $A(t)e^{j\varphi\,(t)}$ into the signal $A_0 e^{j\varphi\,(t)}$.
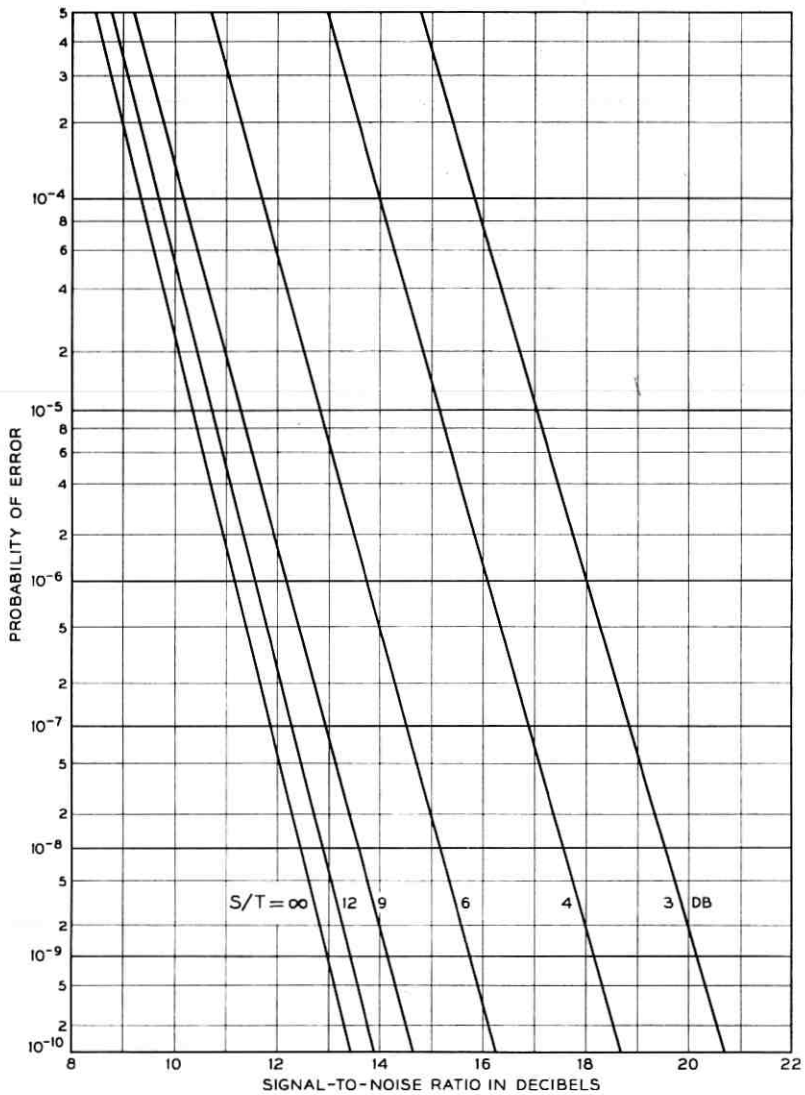
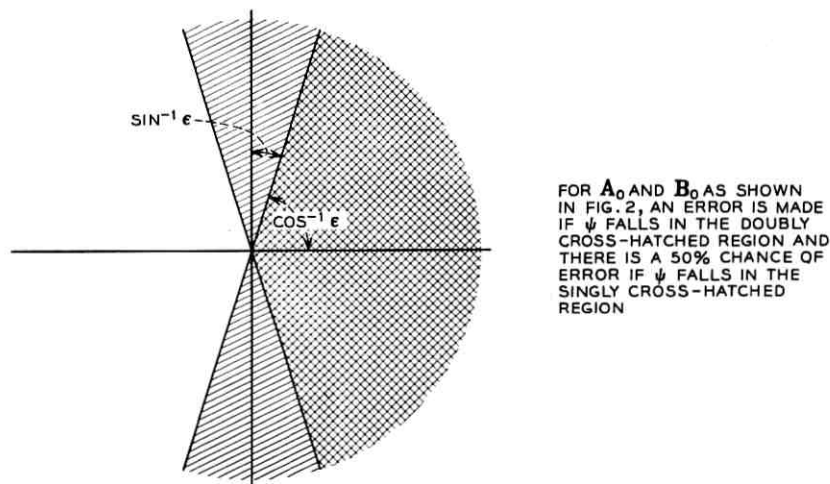Fig. 3 — Probability of error in an unlimited DCPSK system.

FOR $\mathbf{A_0}$ AND $\mathbf{B_0}$ AS SHOWN
IN FIG. 2, AN ERROR IS MADE
IF ψ FALLS IN THE DOUBLY
CROSS-HATCHED REGION AND
THERE IS A 50% CHANCE OF
ERROR IF ψ FALLS IN THE
SINGLY CROSS-HATCHED
REGION

Fig. 4 — Geometric interpretation of the limits of integration over $p(\psi)$.

where $p(\psi)$ is the probability density function of ψ. The function $p(\psi)$ is derived in Appendix B. It is given by:

$$
\begin{aligned}
p(\psi) = & -\frac{1}{2\pi} \exp\left(-\frac{1}{\sigma^2}\right) + \frac{1}{\pi} \exp\left(\frac{1}{2\sigma^2}\right) \\
& -\frac{1}{4\pi\sigma^2} \int_{-(\pi/2)}^{\pi/2} \cos\alpha \cos(\alpha + \psi) \\
& \cdot \exp\left(-\frac{\sin^2\alpha + \sin^2(\alpha + \psi)}{2\sigma^2}\right) d\alpha \\
& + \frac{1}{4\pi\sigma^2} \int_{-(\pi/2)}^{\pi/2} \cos\alpha \cos(\alpha + \psi) \\
& \cdot \exp\left(-\frac{\sin^2\alpha + \sin^2(\alpha + \psi)}{2\sigma^2}\right) \operatorname{erf} \frac{\cos\alpha}{\sqrt{2}\,\sigma} \\
& \cdot \operatorname{erf} \frac{\cos(\alpha + \psi)}{\sqrt{2}\,\sigma} \, d\alpha.
\end{aligned}
\tag{10}
$$

The probability of error is derived in Appendix C by substituting (10) into (9). The result is

$$
\Pi = \frac{1}{\pi} \int_0^{\pi/2} \exp\left[-\frac{(1 - \varepsilon^2)(1/2\sigma^2)}{1 - \varepsilon \sin 2\theta}\right] d\theta.
\tag{11}
$$

Again the integral must be computed numerically. The results are shown in Fig. 5.

IV. CONCLUSIONS

From the form of (2) one sees that for the ideal case (no dead zone) the probability of error, $\Pi(S/N)$, will plot as a straight line if the scale of the abscissa of the graph paper is proportional to the logarithm of the logarithm of the scale of the ordinate.* A grid of this type is used in Figs. 3 and 5. From these figures one sees that the graph of Π vs S/N is very nearly linear and very nearly parallel to the graph of the ideal case even when the dead zone is significant.

Since the family of curves of $\Pi(S/N)$ for various values of S/T is a set of approximately parallel lines, it is convenient to define a "threshold effect noise figure," N_T, as the change of S/N which would give a degradation in error-probability equivalent to that due to the dead zone. Since these curves are not exactly parallel lines, N_T will be a function of the value of Π at which it is determined, but this effect is quite small over the range of values plotted in Figs. 3 and 5.

The values of N_T as a function of S/T (evaluated at $\Pi = 10^{-9}$) are plotted in Fig. 6 for both the limited and the unlimited cases. From Fig. 6 it is observed that $N_T \leq 1$ db for $S/T \geq 9.4$ db and 5.8 db in the unlimited and limited cases, respectively. However, as S/T becomes smaller the effect of N_T on error-probability becomes quite important.

The difference in the value of N_T for the limited case and the unlimited case gives a measure of the improvement gained from the ideal limiting process. This improvement is very small for large S/T; however, for $S/T = 8$ db it already amounts to about 1-db improvement in N_T, for $S/T = 6$ db the improvement is about 2 db, and for $S/T = 3$ db this improvement exceeds 4 db.

V. ACKNOWLEDGMENTS

The author wishes to thank Mrs. C. L. Beattie for programing and performing the numerical calculations.

APPENDIX A

Derivation of the Error-Rate Equation for Unlimited Systems

Substituting (6) into (4) or (5) and making the change of variable

---

* It may be worth mentioning that for 1 db $\leq$ S/N $\leq$ 10 db one can calculate the error rate for the ideal case directly on a slide rule by reading 2Π on the LL3 scale directly opposite S/N (in bels) on the L scale.
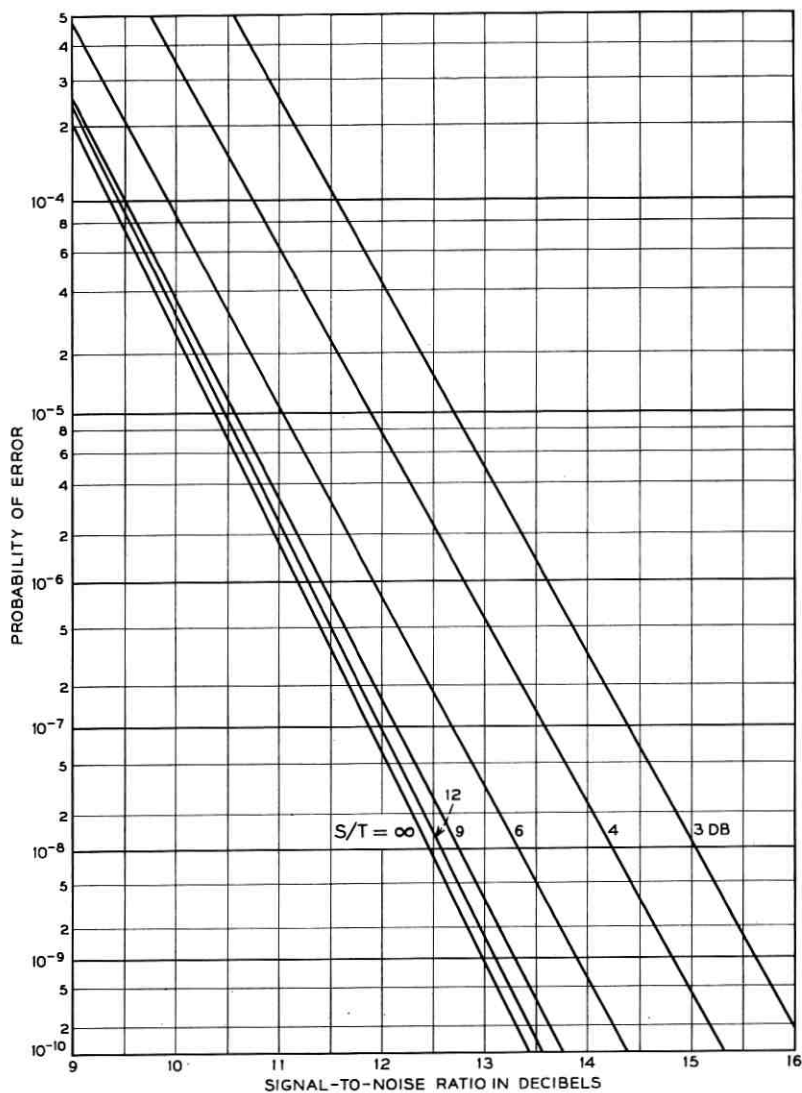
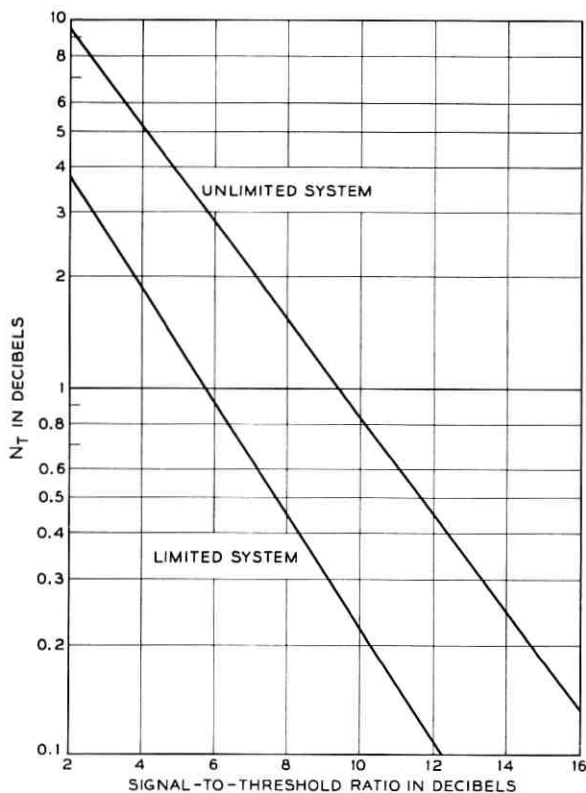Fig. 5 — Probability of error in a limited DCPSK system.

Fig. 6 — Threshold effect noise figure vs signal-to-threshold ratio.

$x \rightarrow x - 1$ gives

$$
\begin{aligned}
\Pi_{\pm} = \ & \frac{1}{(2\pi)^{3/2}\sigma^3} \frac{1}{2} \int_{-\varepsilon}^{\varepsilon}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\exp\left(-\dfrac{(x-1)^2 + y^2}{2\sigma^2}\right)}{\sqrt{x^2 + y^2}} \\
& \cdot \exp\left(-\frac{[v \mp x]^2}{2\sigma^2[x^2 + y^2]}\right) dx\, dy\, dv \\
& + \frac{1}{(2\pi)^{3/2}\,\sigma^3} \int_{\varepsilon}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\exp\left(-\dfrac{(x-1)^2 + y^2}{2\sigma^2}\right)}{\sqrt{x^2 + y^2}} \\
& \cdot \exp\left(-\frac{[\mp v + x]^2}{2\sigma^2[x^2 + y^2]}\right) dx\, dy\, dv.
\end{aligned}
\tag{12}
$$

Making the substitution

$$u = \frac{v + x}{\sqrt{2}\,\sigma\,\sqrt{x^2 + y^2}}$$

and changing the order of integration gives

$$\Pi_\pm = \frac{1}{4\pi\sigma^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left(-\frac{(x - 1)^2 + y^2}{2\sigma^2}\right) \left\{\frac{1}{2}\operatorname{erf}\frac{\varepsilon + x}{\sqrt{2}\,\sigma\,\sqrt{x^2 + y^2}}\right.$$
$$\left. -\frac{1}{2}\operatorname{erf}\frac{-\varepsilon + x}{\sqrt{2}\,\sigma\,\sqrt{x^2 + y^2}} + \operatorname{erfc}\frac{\varepsilon + x}{\sqrt{2}\,\sigma\,\sqrt{x^2 + y^2}}\right\} dx\,dy \tag{13}$$

where erf (z) and erfc (z) have the usual definitions

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \exp(-t^2)\,dt, \qquad \operatorname{erfc}(z) = 1 - \operatorname{erf}(z). \tag{14}$$

Equation (13) is readily simplified by means of (14) to give (7).

APPENDIX B

Derivation of the Probability Density $p(\psi)$

Bennett[5] and Davenport and Root[6] give the probability density of α (see Fig. 2) as

$$p(\alpha) = \frac{1}{2\pi} \exp\left(-\frac{1}{2\sigma^2}\right) + \frac{1}{2}\frac{1}{\sqrt{2\pi}\,\sigma} \cos\alpha \exp\left(-\frac{\sin^2\alpha}{2\sigma^2}\right)$$
$$\left[\operatorname{erf}\frac{\cos\alpha}{\sqrt{2}\sigma} + 1\right]. \tag{15}$$

A similar argument gives

$$p'(\beta) = \frac{1}{2\pi} \exp\left(-\frac{1}{2\sigma^2}\right) + \frac{1}{2}\frac{1}{\sqrt{2\pi}\,\sigma} \cos\beta \exp\left(-\frac{\sin^2\beta}{2\sigma^2}\right)$$
$$\left[\operatorname{erf}\frac{\cos\beta}{\sqrt{2}\sigma} - 1\right]. \tag{16}$$

From these results, one can evaluate $p(\psi)$.

$$p(\psi) = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \delta(\psi + \alpha - \beta)p(\alpha)p'(\beta)\,d\alpha\,d\beta$$
$$= \int_{-\pi}^{\pi} p(\alpha)p'(\alpha + \psi)\,d\alpha. \tag{17}$$

Cahn[7] gives numerical values for $p(\psi)$ for a wide range of values of S/N. However, an expression for $p(\psi)$ suitable for calculation of error-probabilities is obtained as follows. Substituting (15) and (16) into (17) and performing the indicated multiplication gives

$$
\begin{aligned}
p(\psi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} & \left\{ \frac{1}{2\pi} \exp\left(-\frac{1}{\sigma^2}\right) + \frac{1}{2} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{2\sigma^2}\right) \cos\alpha \right. \\
& \cdot \exp\left(-\frac{\sin^2\alpha}{2\sigma^2}\right) \mathrm{erf}\,\frac{\cos\alpha}{\sqrt{2}\sigma} + \frac{1}{2} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{2\sigma^2}\right) \\
& \cdot \cos(\alpha+\psi) \exp\left(-\frac{\sin^2(\alpha+\psi)}{2\sigma^2}\right) \mathrm{erf}\,\frac{\cos(\alpha+\psi)}{\sqrt{2}\,\sigma} \\
& + \frac{1}{4\sigma^2} \cos\alpha \cos(\alpha+\psi) \\
& \cdot \exp\left(-\frac{\sin^2\alpha + \sin^2(\alpha+\psi)}{2\sigma^2}\right) \\
& \cdot \mathrm{erf}\,\frac{\cos\alpha}{\sqrt{2\pi}\,\sigma} \mathrm{erf}\,\frac{\cos(\alpha+\psi)}{\sqrt{2}\,\sigma} - \frac{1}{4\sigma^2} \cos\alpha \cos(\alpha+\psi) \\
& \left. \cdot \exp\left(-\frac{\sin^2\alpha + \sin^2(\alpha+\psi)}{2\sigma^2}\right) \right\} d\alpha
\end{aligned}
\tag{18}
$$

where the other terms vanish because they are odd with respect to $\alpha \to \alpha + \pi$.

Consider

$$
I = \int_{-\pi}^{\pi} \cos(\alpha+\psi) \exp\left(-\frac{\sin^2(\alpha+\psi)}{2\sigma^2}\right) \mathrm{erf}\,\frac{\cos(\alpha+\psi)}{\sqrt{2}\,\sigma} \, d\alpha. \tag{19}
$$

We observe that the integrand is periodic in period π and the result is therefore independent of ψ. Thus, we can write

$$
I = 2 \int_{-(\pi/2)}^{\pi/2} \cos\alpha \exp\left(-\frac{\sin^2\alpha}{2\sigma^2}\right) \mathrm{erf}\,\frac{\cos\alpha}{\sqrt{2}\,\sigma} \, d\alpha
$$

which is easily integrated by means of the change of variable $x = \sin\alpha$ and the definition of the error function. The result is

$$
I = 2\sqrt{2\pi}\,\sigma\left[1 - \exp\left(-\frac{1}{2\sigma^2}\right)\right]. \tag{20}
$$

Substituting (20) into (18) and observing that all of the terms of the integrand of (18) are even with respect to $\alpha \to \alpha + \pi$, gives (10).

APPENDIX C

Derivation of the Error-Rate Equation for a Limited System

Designate the four terms on the right of (10) p_1, p_2, p_3, and p_4, respectively. We observe that

$$p_i(\psi) = p_i(\psi + \pi) \quad \text{for} \quad i = 1,2,4$$

$$p_3(\psi) = -p_3(\psi + \pi).$$

From this we have

$$\Pi = \int_{\pi+\cos^{-1}\varepsilon}^{\cos^{-1}\varepsilon} \{p_1(\psi) + p_2(\psi) + p_4(\psi)\}d\psi \qquad (21)$$

$$+ \int_{\pi+\cos^{-1}(-\varepsilon)}^{\cos^{-1}\varepsilon} p_3(\psi)d\psi.$$

The integration of p_1 and p_2 is trivial. p_4 can be integrated by changing the order of integration and applying (20) twice. Performing the first integral in (21) and making a change of variable in the second integral gives

$$\Pi = \frac{1}{2} - \frac{1}{4\pi\sigma^2} \int_{-1}^{1} \int_{-\sqrt{1-x^2}\cos\varphi + x\sin\varphi}^{\sqrt{1-x^2}\cos\varphi + x\sin\varphi} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) dxdy, \quad (22)$$

where

$$\varphi = \sin^{-1}\varepsilon.$$

Inspection of the limits of integration in (22) reveals that the integration is to be performed over an ellipse centered at the origin. Since the integrand is spherically symmetric we can integrate over any quarter of the ellipse — say the first quadrant. Then

$$\Pi = \frac{1}{2} - \frac{1}{\pi\sigma^2} \int_{0}^{\pi/2} \int_{0}^{\frac{\cos\varphi}{\sqrt{1-\sin\varphi\,\sin^2\theta}}} \exp\left(-\frac{r^2}{2\sigma^2}\right) rdrd\theta,$$

where we have made the change of variable $x = r\cos\theta$, $y = r\sin\theta$. This can be easily integrated over r to give (11).

REFERENCES

1. Lawton, J. G., Comparison of Binary Data Transmission, Proc. 1958 Conference on Military Electronics.
2. Cahn, Charles R., Performance of Digital Phase Modulation Communication Systems, IRE Trans. CS, May, 1959, pp. 3–6.

3. Bennett, W. R. and Salz, J., Binary Data Transmission by FM Over a Real Channel, B.S.T.J. *42*, September, 1963, pp. 2387–2426.
4. Bussgang J. J. and Leiter, M., Error Rate Approximation for Differential Phase-Shift Keying, IEEE Trans. CS 12, March, 1964, pp. 18–27.
5. Bennett, W. R., Methods of Solving Noise Problems, Proc. IRE, *44*, May, 1956, pp. 609–638.
6. Davenport, W. B. and Root, W. L., Introduction to the Theory of Random Signals and Noise, McGraw-Hill Book Company, Inc., New York, 1958.
7. Cahn, C. R., Performance of Digital Phase-Modulations Communications Systems, Thompson-Ramo-Wooldridge Report M-110-9U5.

Optics of General Guiding Media

By J. P. GORDON

Weakly focusing transparent media provide possible means for guided transmission of coherent light beams with relatively small loss. The scalar wave equation for the eigenmodes of propagation in such a medium is formally identical with Schroedinger's wave equation. Hence, the methods used in the solution of quantum-mechanical problems, such as the Wentzel-Kramers-Brillouin (WKB) approximation, are immediately applicable to this problem. Solutions for the eigenmodes and eigenvalues in the case of focusing in one dimension are given, and the Pöschl-Teller medium, whose index varies as

$$n = n_o[1 - (\alpha/2) \tan^2 \eta x]$$

is discussed in some detail. In addition, the relationship between the wave solutions and geometrical (ray) optics is examined.

I. INTRODUCTION

Weakly focusing transparent media, exemplified by the gas lens,[1] provide a possible means for guided transmission of coherent light beams with relatively small loss. The optics of a medium whose refractive index decreases quadratically away from some spatial axis (the z axis, say) has been the subject of much discussion in the literature.[2] In this paper, we would like to examine the properties of more general guiding media. Some work on this problem has been carried out recently by S. E. Miller.[3]

Consider a light beam traveling paraxially in the z direction in space, guided in a weakly focusing transparent medium. For simplicity, we shall consider focusing only in one of the two transverse dimensions. Thus, the refractive index of the medium and the electromagnetic field will be assumed to be dependent only on the x dimension, and to be independent of y.

Our concern is with transparent media whose index of refraction has the form

$$n = n_o(1 - \tfrac{1}{2}f(x)).$$

We make the following assumptions about $f(x)$

$$f(x) \ll 1 \quad \text{in the range of interest} \tag{1}$$

$$\frac{\partial^2 f}{\partial x^2} \geq 0 \tag{2}$$

$$f(0) = 0; \quad f(x) = f(-x). \tag{3}$$

We will investigate both geometrical and physical optics of such a medium in the realm of validity of the paraxial ray equation. Of the assumptions, (1) implies only gradual changes of index, (2) insures focusing properties, and is used in the approximation procedure of Section VII, while (3) is made only for mathematical convenience and can easily be relaxed.

II. GEOMETRICAL (RAY) OPTICS

The well-known paraxial ray equation

$$\frac{d^2 x}{dz^2} = \frac{1}{n}\frac{\partial n}{\partial x} = -\frac{1}{2}\frac{\partial f}{\partial x} \tag{4}$$

has the following general solution.

If we let

$$p = \frac{dx}{dz}, \tag{5}$$

i.e., p is the slope of the ray path; then (4) becomes

$$p\frac{dp}{dx} = -\frac{1}{2}\frac{\partial f}{\partial x}.$$

Hence,

$$p^2 + f = \text{const} \equiv \xi. \tag{6}$$

Inserting (6) in (5), and solving, we get

$$z = \int \frac{dx}{\sqrt{\xi - f(x)}} + \text{const.} \tag{7}$$

Note from (6), (5) and (3) that $\pm\xi^{\frac{1}{2}}$ is the slope of the ray path as it crosses the axis $x = 0$.

III. PHYSICAL OPTICS

In the paraxial ray approximation, the electromagnetic fields are always very nearly perpendicular to the "direction" of propagation,

which we take to be the z axis. Hence, we can use the scalar wave equation. For harmonic $[\exp(-i\omega t)]$ fields that are independent of y, we arrive at the wave equation

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial z^2} + k^2 n_o^2 (1 - f(x))\psi = 0 \qquad [k \equiv \omega/c] \qquad (8)$$

where we have used (1) to expand n^2. We look for solutions to (8) with a propagation constant β in the z direction, i.e.,

$$\psi \propto e^{i\beta z}; \qquad \frac{\partial^2 \psi}{\partial z^2} = -\beta^2 \psi \qquad (9)$$

yielding, for (8)

$$\frac{\partial^2 \psi}{\partial x^2} + [\beta_o^2 - \beta^2 - \beta_o^2 f(x)]\psi = 0 \qquad (10)$$

where we have substituted $n_o^2 k^2 = \beta_o^2 \equiv (2\pi/\lambda)^2$. With the further substitution

$$1 - \beta^2/\beta_o^2 = \xi \qquad (11)$$

(10) becomes

$$\frac{\partial^2 \psi}{\partial x^2} + \beta_o^2 (\xi - f(x))\psi = 0. \qquad (12)$$

Equation (12) is Schroedinger's wave equation for a particle in a one dimensional potential well.[4] It has a sequence of eigensolutions ψ_m and corresponding eigenvales ξ_m. The eigenfunctions ψ_m here represent the transverse distributions of the propagating field modes, while the eigenvalues ξ_m give the propagation constants. In accord with (1), the eigenvalues will be much smaller than unity, and so we can find the propagation constants β_m from (11) as

$$\beta_m = \left(1 - \frac{\xi_m}{2}\right)\beta_o. \qquad (13)$$

If we use the notation of the classical ray path equation (6), and set

$$p = p(x) = \sqrt{\xi - f(x)} \qquad (14)$$

then (12) becomes

$$\frac{\partial^2 \psi}{\partial x^2} + \beta_o^2 p^2 \psi = 0. \qquad (15)$$

We note from (2) and (3) that $f(x)$ increases monotonically with $|x|$. Consider the points $x = \pm A$, where

$$f(x) = f(A) = \xi. \tag{16}$$

For $|x| < A$, p^2 is positive, and ψ has oscillatory behavior. For $|x| > A$, p^2 is negative, and ψ has decreasing exponential behavior. The points $|x| = \pm A$ are inflection points of ψ. In the ray solution, the points $|x| = A$, where $p = 0$, are the turning points, where the ray has zero slope and, correspondingly, maximum excursion from the axis.

IV. THE WKB APPROXIMATION

A well-known solution to (15), valid in the range $|x| < A$ if p can be considered approximately constant over a few cycles of the oscillatory behavior of ψ, is

$$\psi \propto p^{-\frac{1}{2}} \cos \left\{ \beta_o \int_0^x p \, dx + \alpha \right\}. \tag{17}$$

With the symmetry of f assumed in (3), one can show directly from (12) that ψ must be either symmetrical or antisymmetrical in x. Thus, α must be a multiple of $\pi/2$. The eigenvalues of ξ are then determined by the condition that ψ must be matched through the turning point $x = A$ (where another approximate solution to (15) is necessary) into the *decreasing* exponential solution for $x > A$. Discussions of this problem may be found in most quantum mechanics texts.[5] If $f(x)$ can be assumed linear in x over a suitable region near A, then the asymptotic formulae for connection through $x = A$ are

$$\psi \propto p^{-\frac{1}{2}} \cos \left(\theta - \frac{\pi}{4} \right) \qquad\qquad x < A \tag{18a}$$

$$\rightarrow \left(\frac{2\pi\theta}{3p} \right)^{\frac{1}{2}} [J_{\frac{1}{3}}(\theta) + J_{-\frac{1}{3}}(\theta)] \qquad x \lesssim A \tag{18b}$$

$$\rightarrow \left(\frac{2\pi\theta}{3p} \right)^{\frac{1}{2}} [-I_{\frac{1}{3}}(\theta) + I_{-\frac{1}{3}}(\theta)] \qquad x \gtrsim A \tag{18c}$$

$$\rightarrow \tfrac{1}{2} p^{-\frac{1}{2}} \exp(-\theta) \qquad\qquad x > A \tag{18d}$$

where

$$\theta = \theta(x) = \left| \beta_o \int_A^x |p(x')| \, dx' \right|. \tag{18e}$$

The phase of ψ as it approaches the turning point must be as in (18a). Otherwise the connection through the turning point would give rise to the increasing exponential in the region $x > A$, and this is unallowable.

Expressions (18a) and (17) for ψ can be equated only when

$$\beta_o \int_0^A p \, dx = (m + \tfrac{1}{2}) \frac{\pi}{2} \tag{19}$$

where m is an integer. Even and odd symmetry solutions correspond respectively to even or odd values of m. Equation (19) is Bohr's quantization rule, and its solutions give the eigenvalues ξ_m. The eigenfunctions are then given by (18).

V. RELATION BETWEEN WKB AND GEOMETRICAL RAY PATHS

Referring to (19), let

$$I(\xi) = \frac{2\beta_o}{\pi} \int_0^A p \, dx. \tag{20}$$

For successive eigensolutions, I changes by unity. Now,

$$\frac{dI}{d\xi} = \frac{\beta_o}{\pi} \int_0^A \frac{dx}{p} \tag{21}$$

where we have used (14); i.e., $p = \sqrt{\xi - f(x)}$. For a change of I by unity, the propagation constant (13) changes approximately by

$$\delta\beta \approx \frac{d\beta}{dI} = \frac{d\beta}{d\xi} \bigg/ \frac{dI}{d\xi} = \frac{\pi/2}{\displaystyle\int_0^A dx/p}. \tag{22}$$

Now from the ray solution (7), the z distance a ray with the same value of ξ takes to go through one complete cycle of its transverse oscillatory motion is

$$z_\sim = 4 \int_0^A dx/p \tag{23}$$

hence, putting this in (22), we find

$$(\delta\beta)z_\sim = 2\pi. \tag{24}$$

Hence, adjacent modes undergo a relative phase shift of approximately 2π in the same distance the corresponding geometrical ray takes to complete one transverse cycle.

We can make the same argument more general by examining the behavior of solutions of (15) which are sums of component solutions (17) with nearby values of ξ. We look for trajectories $x(z)$ such that the phase differences between components remain constant. Then if we make up a wave packet, the packet will follow one of these trajectories.

Putting the z dependence back into (17), and keeping, for example,

only the positive imaginary part of the cosine term, we have

$$\psi \propto p^{-\frac{1}{2}} \left\{ \exp i \left[\beta z + \beta_o \int_0^z p \, dx + \alpha \right] \right\}. \tag{25}$$

The phase we have to maintain constant over a range of ξ is thus

$$\beta z + \beta_o \int_0^z p \, dx + \alpha. \tag{26}$$

Differentiating with respect to ξ, setting the result equal to zero, and using (13) and (14), we have

$$z = \int_0^z \frac{dx}{p} + \frac{2}{\beta_o} \frac{d\alpha}{d\xi}. \tag{27}$$

To complete the picture, we note that in order to make a wave packet initially concentrated at some point (x_o, z_o) should take all the phases initially equal at that point, so that

$$\alpha = -\beta z_o - \beta_o \int_0^{x_o} p \, dx + \text{const}$$

$$\frac{2}{\beta_o} \frac{d\alpha}{d\xi} = z_o - \int_0^{x_o} \frac{dx}{p}.$$

Hence, the wave packet trajectories of stationary phase are given by

$$z - z_o = \int_{x_o}^z \frac{dx}{p}. \tag{28}$$

This is identical with the ray path (7). In this last discussion we have neglected the discrete nature of the eigenvalues. Such an approximation should be valid for reasonably large mode numbers m.

VI. THE PÖSCHL-TELLER POTENTIAL

A number of functions $f(x)$ yield analytically integrable equations. The square law medium $f(x) = (x/b)^2$ and the square well medium $f(x) = 0$ for $|x| < A, f(x) \to \infty$ for $x > A$, suitably joined, are media with well-known solutions both for the geometrical and physical optics equations. An interesting function, which in a slightly more general form goes by the name of the Pöschl-Teller potential[6] in quantum mechanics, is the function

$$f(x) = \alpha \tan^2 (\eta x); \qquad -\pi/2\eta < x < \pi/2\eta. \tag{29}$$

Near $x = 0$

$$f(x) = \alpha\eta^2 x^2 \tag{30}$$

and there are impenetrable boundaries at $x = \pm\pi/2\eta$. For this function the ray equation (7), the wave equation (12), and the WKB approximation (17) can be directly integrated and compared. As the parameters α and η are changed, the function varies smoothly between a square well and a square law type.

Consider first the ray equation,

$$z = \int \frac{dx}{[\xi - \alpha \tan^2 \eta x]^{\frac{1}{2}}} + \text{const.} \tag{31}$$

This integrates to [Burrington ✳ 258]

$$z = \frac{1}{\eta[\xi + \alpha]^{\frac{1}{2}}} \sin^{-1}([1 + \alpha/\xi]^{\frac{1}{2}} \sin \eta x) + \text{const.} \tag{32}$$

If we set the constant equal to zero, then $z = 0$ when $x = 0$. This expression generates a ray path, with x taking on values between the turning points

$$x = \pm A = \pm\eta^{-1} \sin^{-1}[1 + \alpha/\xi]^{-\frac{1}{2}} \tag{33}$$

while the period of oscillation is

$$z_\sim = 2\pi/\eta[\xi + \alpha]^{\frac{1}{2}}. \tag{34}$$

Note that at the turning points, $f(x) = \xi$. We see that for $\alpha \gg \xi$, ηx is always small, and (32) reduces to

$$x = (\xi/\eta^2\alpha)^{\frac{1}{2}} \sin(\eta\alpha^{\frac{1}{2}}z)$$

which is appropriate to a square law medium, while for $\alpha \ll \xi$, the ray travels back and forth with constant slope $\xi^{\frac{1}{2}}$ between reflecting walls separated by π/η.

Pöschl and Teller have found exact solutions for the wave equation (12) with this potential. They have shown that the eigenvalues follow the simple law

$$\xi_m = (\eta^2/\beta_o^2)(m^2 + 2ma + a) \tag{35a}$$

where a is the positive root of the equation

$$a(a - 1) = \alpha\beta_o^2/\eta^2. \tag{35b}$$

The corresponding eigenfunctions are

$$\psi_m = \cos^a(\eta x) \sum_{\substack{k=0,\text{even} \\ k=1,\text{odd}}}^{m\,^{\text{even}}_{\text{odd}}} c_k \sin^k(\eta x) \tag{36}$$

where

$$\frac{c_{k+2}}{c_k} = \frac{(m - k)(m + k + 2a)}{(k + 1)(k + 2)}.$$

From (35) and (13), we find that the propagation constants of the modes follow the law

$$\beta_m = \beta_o - (\eta^2 a/2\beta_o)\left(1 + 2m + \frac{m^2}{a}\right). \tag{37}$$

To examine the square law limit, we set $\alpha \gg \xi$. Then [see (35)], $a \approx \alpha^{\frac{1}{2}}\beta_o/\eta \gg 1$,

and, from (37)

$$\beta_m \approx \beta_o - \eta\alpha^{\frac{1}{2}}(m + \tfrac{1}{2}) \tag{38}$$

while from (36)

$$\psi_m \approx \exp\left(-\tfrac{1}{2}\eta\alpha^{\frac{1}{2}}\beta_o x^2\right)H_m(\eta\alpha^{\frac{1}{2}}\beta_o x^2). \tag{39}$$

To examine the square well limit, we set $\alpha \ll \xi$. Then $a \approx 1$, and

$$\beta_m \to \beta_o - \tfrac{1}{2}(\eta^2/\beta_o)(m + 1)^2 \tag{40}$$

$$\psi_m \to \cos\left[(m + 1)\eta x\right] \quad \text{for} \quad m \text{ even}$$
$$\psi_m \to \sin\left[(m + 1)\eta x\right] \quad \text{for} \quad m \text{ odd.} \tag{41}$$

Returning now to the general solution for the propagation constants, note from (37) that the average of the propagation constant differences between the mth mode and its two neighbors is given by

$$\delta\beta = \tfrac{1}{2}(\beta_{m-1} - \beta_{m+1}) = (\eta^2/\beta_o)(m + a). \tag{42}$$

If we look at the ray equation, and insert the values of α and ξ_m (35) into (34), we find

$$z_\sim = 2\pi/\delta\beta. \tag{43}$$

This is a more precise version of (24).

Using the Pöschl-Teller potential, the WKB result (19) for the eigenvalues ξ_m can also be integrated. Referring to (21), (23), and (34), we have

$$\frac{dI}{d\xi} = (\beta_o/\pi)\int_0^A dx/p = \beta_o z_\sim/4\pi$$

$$= \beta_o/2\eta(\xi + \alpha)^{\frac{1}{2}}.$$

Hence,

$$I = \beta_o(\xi + \alpha)^{\frac{1}{2}}/\eta - \beta_o\alpha^{\frac{1}{2}}/\eta. \tag{44}$$

Here the constant of integration is determined by the condition that $I = 0$ at $\xi = 0$. Thus, (19) [see also (20)] is

$$\beta_o(\xi_m + \alpha)^{\frac{1}{2}}/\eta = \beta_o\alpha^{\frac{1}{2}}/\eta + (m + \tfrac{1}{2}) \tag{45}$$

or

$$\begin{aligned} \xi_m &= [(\eta/\beta_o)(m + \tfrac{1}{2}) + \alpha^{\frac{1}{2}}]^2 - \alpha \\ &= (\eta/\beta_0)^2[(m + \tfrac{1}{2})^2 + (2m + 1)\sqrt{a(a - 1)}] \end{aligned} \tag{46}$$

where we have used (35b). We see that the WKB result (46) agrees with the exact result (35a) in the limit of the square law medium, $a \gg 1$. In the square well limit, $a \to 1$, the WKB method gives eigenvalues proportional to $(m + \tfrac{1}{2})^2$ rather than $(m + 1)^2$ as does the exact result (35). This last discrepancy can be traced to incorrect matching of the boundary conditions by the WKB method, since here the WKB wave function (17) is the correct one. The factor $\tfrac{1}{2}$ in $m + \tfrac{1}{2}$ arises from matching the boundary condition, and in fact, we can get the WKB answer to equal the exact one for all a and m here by the artifice of replacing $m + \tfrac{1}{2}$ in (46) by

$$m + a - \sqrt{a(a - 1)} = m + a - \alpha^{\frac{1}{2}}\beta_o/\eta. \tag{47}$$

The number added to m varies between the value $\tfrac{1}{2}$ for large a to one as a approaches unity.

VII. FURTHER APPLICATION OF THE WKB METHOD

The results of Section VI give us reasonable faith in the WKB method of obtaining eigenvalues even at small values of m. The number we add to m may however, take on values between $\tfrac{1}{2}$ and 1 depending in general on m as well as on the form of the potential.

Let us consider the WKB result further. Equation (19), which we repeat here for convenience, may be written

$$\beta_o \int_0^{A_m} \sqrt{\xi_m - f(x)}\, dx = (m + \tfrac{1}{2})\frac{\pi}{2} \tag{48}$$

where

$$f(A_m) = \xi_m.$$

We can see two simple bounds. For functions that satisfy (2) and (3),

we have that in the range $0 \leq x \leq A_m$,

$$\xi_m \geq \xi_m - f(x) \geq \xi_m \left(1 - \frac{x}{A_m}\right) \qquad (49)$$

hence, performing the integral for the two bounds, (48) gives

$$\beta_o \sqrt{\xi_m} A_m \geq (m + \tfrac{1}{2}) \frac{\pi}{2} \geq \tfrac{2}{3}\beta_o \sqrt{\xi_m} A_m.$$

We can write this as

$$\sqrt{\xi_m} A_m = r(m + \tfrac{1}{2})\pi/2\beta_o \qquad (50)$$

where

$$1 \leq r \leq 1.5.$$

Note that

$$r^{-1} = \frac{1}{\xi^{\frac{1}{2}} A} \int_0^A \sqrt{\xi - f(x)} \, dx. \qquad (51)$$

If we further remember that we should replace the $\frac{1}{2}$ in $m + \frac{1}{2}$ by some number between $\frac{1}{2}$ and 1 which we can denote by s, we have

$$\sqrt{\xi_m} A_m = r(m + s)\pi/2\beta_o \qquad (52)$$

with

$$1 \leq r \leq 1.5 \quad \text{and} \quad 0.5 \leq s \leq 1.$$

Let us examine the implications of this formula in a definite example. We take the potential to be a pure power law[3]

$$f(x) = \left(\frac{x}{b}\right)^{2n}. \qquad (53)$$

We need the value of A_m;

$$f(A_m) = \xi_m \quad \text{therefore,} \quad A_m = b\xi_m^{1/2n}.$$

Then (52) is easily solved, yielding

$$\xi_m = \left\{\frac{\pi r(m + s)}{2\beta_o b}\right\}^{2n/n+1}. \qquad (54)$$

In this case, examination of (51) reveals that r is given by

$$r^{-1} = \int_0^1 \sqrt{1 - y^{2n}} \, dy \qquad (55)$$

so that r is independent of m, and varies between

$$\frac{4}{\pi} \quad \text{for} \quad n = 1 \quad \text{and} \quad 1 \quad \text{for} \quad n \to \infty.$$

Further, we know that $s = \frac{1}{2}$ for $n = 1$ (square law), and $s = 1$ for $n \to \infty$ (square well). As a guess, we would be tempted to try

$$s = \frac{n}{n + 1}$$

as a suitable interpolation. Then we have as our final result

$$\xi_m = \left\{ \left(m + \frac{n}{n + 1} \right) \frac{\pi r}{2\beta_o b} \right\}^{2n/n+1} \tag{56}$$

hence for the phase constant [repeating (13)]

$$\beta_m = \beta_o (1 - \xi_m/2). \tag{57}$$

VIII. ACKNOWLEDGMENT

REFERENCES

1. Berreman, D. W., A Gas Lens Using Unlike Counterflowing Gases, B.S.T.J., *43*, July, 1964, pp. 1469–1475; Berreman, D. W., A Lens or Light Guide Using Convectively Distorted Thermal Gradient in Gases; ibid, pp. 1476–1479.
2. Tien, P. K., Gordon, J. P., and Whinnery, J. R., Focusing of a Light Beam of Gaussian Field Distribution in Continuous and Periodic Lens-Like Media, Proc. IEEE, *53*, February, 1965, pp. 129–136; see also H. Kogelnik, On the Propagation of Gaussian Beams of Light through Lens-Like Media Including those with a Loss or Gain Variation, Applied Optics, December, 1965. Many references to other work may be found in these articles.
3. Miller, S. E., Light Propagation in Generalized Lens-Like Media, B.S.T.J., *44*, November, 1965, pp. 2017–2064.
4. This correspondence has been pointed out previously; Landau, L. D. and Lifshitz, E. M., *Electrodynamics of Continuous Media*, Pergamon Press Ltd., 1960, footnote p. 285.
5. See for example, Schiff, L. I., *Quantum Mechanics*, McGraw-Hill Book Company, Inc., New York, 1949, Chapter VII.
6. Pöschl, G. and Teller, E., Bemerkungen zur Quantenmechanik des Anharmonischen Oszillators, Zeits. Phys., *83*, 1933, pp. 143–151. The authors consider the functions

$$\frac{a}{\sin^2 \alpha x} + \frac{b}{\cos^2 \alpha x}$$

and

$$\frac{a}{\sinh^2 \alpha x} - \frac{b}{\cosh^2 \alpha x}.$$

The parameters a, b, and α are not the same as those used in the text. In addition, treatment of the function

$$-(b/\cosh^2 \alpha x)$$

can be found in Landau, E. M. and Lifshitz, E. M., *Quantum Mechanics, Non-Relativistic Theory*, Pergamon Press Ltd., 1958, pp. 69–70.

Contributors to This Issue

TA-SHING CHU, B.S., 1955, National Taiwan University; M.S., 1957, Ph.D., 1960, Ohio State University; Bell Telephone Laboratories, 1963—. Mr. Chu has worked in the field of electromagnetics with emphasis on surface waves and microwave antennas. At present he is working on optical and infrared wave propagation through the atmosphere. Member, IEEE, American Physical Society, Sigma Xi, Pi Mu Epsilon.

JAMES P. GORDON, B. S., Massachusetts Institute of Technology, 1949; M.A., 1951, Ph.D., 1955, Columbia University; Bell Telephone Laboratories, 1955—. His research in quantum electronics has involved work on paramagentic resonance, masers, lasers, and the quantum mechanical aspects of communication theory. He has written several technical articles in the field of quantum electronics. Member, AAAA, American Physical Society, Sigma Xi; senior member, IEEE.

D. C. HOGG, B. Sc., 1949, University of Western Ontario; M. Sc., 1950, Ph.D., 1953, McGill University; Bell Telephone Laboratories, 1953—. His work has included studies of artificial dielectrics for microwaves, diffraction of microwaves, and over-the-horizon, millimeter wave and optical propagation. Fellow, IEEE; member, Commission 2 of URSI, Sigma Xi, AAAS.

W. M. HUBBARD, B.S., 1957, Georgia Institute of Technology; M.S., 1958, University of Illinois; Ph.D., 1963, Georgia Institute of Technology; Bell Telephone Laboratories, 1963—. Mr. Hubbard has been concerned with analyses related to the design of millimeter wave solid-state repeaters for use in wave-guide transmission systems. Member, Sigma Xi, Tau Beta Pi, Phi Kappa Phi, American Physical Society.

MARTIN P. LEPSELTER, B.M.E., 1951, College of the City of New York; 1951–53, Ordnance Engineer, U. S. Navy Bureau of Ordnance; 1953–55, Principal Engineer, Hudson Fixtures, Inc.; 1955–57, Lecturer, M. E., College of the City of New York; Bell Telephone laboratories,

1957—. Mr. Lepselter has worked on process development for surface passivation of silicon devices, and developed the shallow-junction contacts for the Telstar solar cells currently in use on many satellites. He is presently completing the development of the metallization processing used in beam-lead devices and integrated circuits. Member, IEEE.

ROBERT W. LUCKY, B.S.E.E., 1957, M.S.E.E., 1959, Ph.D., 1961, Purdue University; Bell Telephone Laboratories, 1961—. Mr. Lucky has been concerned with various analysis problems in the area of digital data communications. His studies into signal distortion during transmission of data through telephone networks has led to the invention of various techniques to automatically correct for this distortion. Member, Eta Kappa Nu, Tau Beta Pi, Sigma Xi, IEEE.

B.S.T.J. BRIEFS

Hologram Transmission via Television

By L. H. ENLOE, J. A. MURPHY and C. B. RUBINSTEIN

Holography, or wavefront reconstruction photography, was first demonstrated by Gabor[1, 2, 3] over fifteen years ago, and it has been the subject of increased investigation over the last five years since the advent of lasers. Possible applications of holography suggest themselves in the fields of three-dimensional and multicolor television.[4] Furthermore, the statistics and redundancy of the hologram of an image may be deliberately made quite different from those of the original image. This has obvious possibilities in encoding schemes for television transmission in general. In this communication we report a first experimental step in this direction, namely the successful transmission via television of a Fresnel type of hologram in which the original object was a transparency.

In conventional two-dimensional hologram construction a transparency is illuminated from behind by a monochromatic, spatially-coherent light wave which then impinges on a photographic plate. This interferes with a reference or carrier wave derived from the same source that strikes the plate at an angle. A record, or hologram, is thus made of both the amplitude and phase of the light transmitted by the transparency. Reconstruction may be achieved by illuminating the hologram with a monochromatic, spatially-coherent light wave. This results in the production of a real image, a virtual image, and a direct wave. Most experimenters in the field have used relatively large angles between the reference and object beams, and the recording has been done on Kodak Spectroscopic plates with 649F emulsion. In the television experiment about to be described, the limited resolution of the equipment required the use of angles of less than one degree between the object beam and reference beam, and would severely limit three-dimensional and multicolor capabilities. The use of these small angles raised the problem of reconstructing the real image without interference from either the direct wave or the virtual image. This was solved by the use of a Fourier transform method which is described in the discussion of the reconstruction process. All recordings were on Poloroid Type 55 P/N film.

The experiment was basically a simple one designed to explore the technique for transmitting a hologram on television and recovering the image. Essentially the experiment consisted of transmitting a hologram of a transparency on television, recovering the hologram at the receiving end, and reconstructing the original image. The recovery was accomplished by photographing the hologram at the receiving end of the system, and then performing the reconstruction separately.

The hologram formation system is shown in Fig. 1. The light source is a He-Ne laser operated at 6328 Å and internally apertured to insure that it operated in the lowest order transverse mode. The output of the laser passed through a microscope objective which imaged the beam to a spot where a pinhole was placed to remove any imperfections in the beam caused by the transmission through the end mirror of the laser and the microscope objective. The beam was then allowed to diverge and fill the aperture of a lens which was located a focal distance away from the pinhole. At this point the beam was collimated and entered what was essentially a Mach-Zehnder interferometer. This was used to obtain as small an angle as desired between the reference and object beam. The angle, and consequently the interference fringe spacing, was adjusted by means of the final beam splitter. The transparency was placed in one arm of the interferometer and the reference beam passed through the other arm. Both beams then impinged directly on the vidicon without a lens and formed the hologram. This
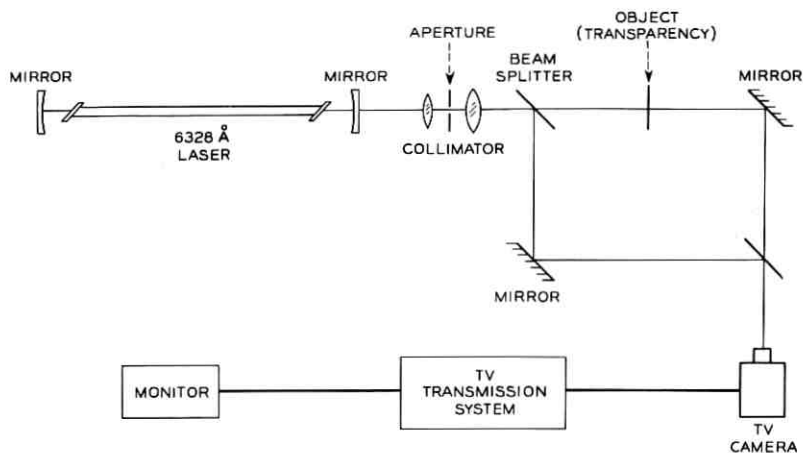

Fig. 1 — Block diagram of the hologram formation system.

signal was then transmitted to a television monitor and the resulting display was photographed.

The hologram reconstruction system is shown in Fig. 2. Everything up to and including the collimator is as shown in Fig. 1. The hologram, which is a Polaroid transparency, is placed in the collimated beam at the front focal plane of the lens and its Fourier transform appears in the back focal plane. At this point a filter, which is essentially an aperture, occults both the main beam and the virtual image beam. A second lens is used to perform another Fourier transform to give an image of the hologram minus the filtered information. The reconstruction of the object transparency, or real image, is then formed at the equivalent distance from the hologram image that the object transparency was from the vidicon.

Fig. 3 contains a photograph of the real image of a hologram transmitted by television. The original object was an opaque transparency containing the transparent word BELL underlined by three transparent lines. The angle between the reference and object beam was approximately 18 minutes of arc corresponding to a fringe spacing of approximately 122 microns on the front surface of the vidicon. This corresponds to an electrical frequency of 2.3 MHz for the standard (525 line) television system used. The photograph of the hologram pattern on the kinescope was made at an exposure of ⅕ second at f/16, and the photograph of the reconstructed image was made at an exposure of 30 seconds without a lens.

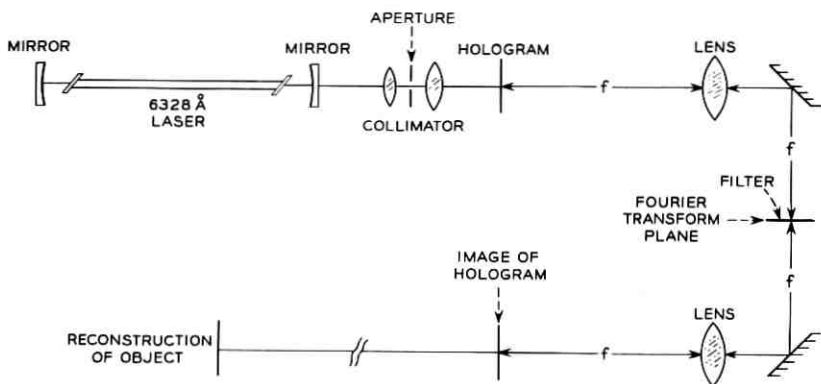In this experiment, we have not detected any noticeable degradation

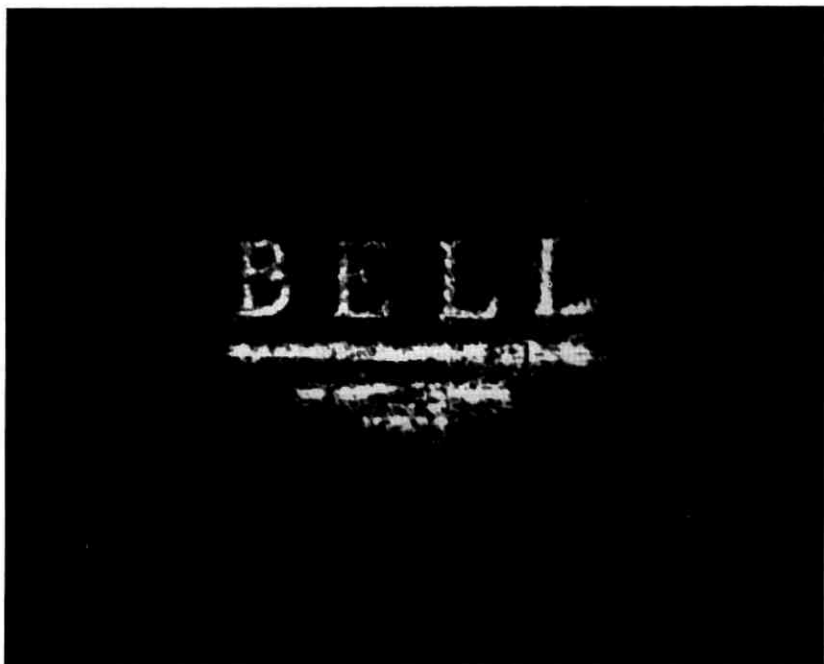Fig. 2 — Block diagram of the hologram reconstruction system.

Fig. 3 — Photograph of the reconstructed real image of a hologram transmitted via television.

of the image due to nonlinearity in the television system. On the other hand, there was a distinct degradation due to the "noise" introduced by the many optical components, and any nonlinearities in the television system were also overshadowed by the degradation caused by the nonrigidity of the Polaroid transparency. These problems are amenable to solution.

This experiment is a first step in the direction of a holographic television system. One important additional result of this work is that it pointed out the usefulness of viewing the laser output at any point in the optical system directly on a television monitor.

REFERENCES

1. Gabor, D., A New Microscope Principle, Nature, *161*, May, 1948, pp. 777–778.
2. Gabor, D., Microscopy by Reconstructed Wave-Fronts, Proc. Roy. Soc. (London), *A197*, July, 1949, pp. 454–487.
3. Gabor, D., Microscopy by Reconstructed Wave Fronts: II, Proc. Phys. Soc. (London), *B64*, June, 1951, pp. 449–469.
4. Leith, E. N., Upatnieks, J., Hildebrand, B. P., and Haines, K., Requirements for a Wavefront Reconstruction Television Facsimile System, J. SMPTE, *74*, October, 1965, pp. 893–896.

Laser Cavities with Increased Axial Mode Separation

By A. D. WHITE

Single axial mode operation of gas lasers may be achieved either by reducing the mirror separation, L, until the mode separation, $c/2L$, is comparable to the oscillation line width, Δv_a, or by using mode suppression or filtering techniques.[1, 2, 3] When the line width exceeds ≈ 3 gc (as in an argon ion laser for example), L must be less than 5 cm for an unstabilized laser and less than 10 cm for a laser stabilized at line center.* Lasers of this length have several disadvantages, the most serious of which (in the case of ion lasers) results from the close proximity of the Brewster windows to the discharge. In the case of the 6328 Å He-Ne laser, restricting the laser tube length to a few centimeters limits the available single axial mode output power to about 1 mw.

This brief describes two additional mode suppression techniques which permit single axial mode operation with mirror separations much larger than $c/2\Delta v_a$.

In common with other techniques,[3] mode suppression is achieved by splitting the beam and using a three-mirror cavity;* however, the specific configurations described in this note are believed to be new and may possibly be advantageous in certain circumstances.

The first technique is most simply understood by referring to Fig.

---

* The mode competition effects reported by Rigrod and Bridges (Electron Device Research Conference, University of Illinois, Urbana, June, 1965) for ion lasers may permit some increase of these lengths.

* It has been pointed out by E. I. Gordon and J. E. Geusic that all mode suppressors using a beam splitter and three mirrors are probably members of the same class of devices characterized by the microwave magic tee with two shorted arms.

REFERENCES

1. Gabor, D., A New Microscope Principle, Nature, *161*, May, 1948, pp. 777–778.
2. Gabor, D., Microscopy by Reconstructed Wave-Fronts, Proc. Roy. Soc. (London), *A197*, July, 1949, pp. 454–487.
3. Gabor, D., Microscopy by Reconstructed Wave Fronts: II, Proc. Phys. Soc. (London), *B64*, June, 1951, pp. 449–469.
4. Leith, E. N., Upatnieks, J., Hildebrand, B. P., and Haines, K., Requirements for a Wavefront Reconstruction Television Facsimile System, J. SMPTE, *74*, October, 1965, pp. 893–896.

Laser Cavities with Increased Axial Mode Separation

By A. D. WHITE

(Manuscript received October 28, 1965)

Single axial mode operation of gas lasers may be achieved either by reducing the mirror separation, L, until the mode separation, $c/2L$, is comparable to the oscillation line width, $\Delta \nu_a$, or by using mode suppression or filtering techniques.[1, 2, 3] When the line width exceeds ≈ 3 gc (as in an argon ion laser for example), L must be less than 5 cm for an unstabilized laser and less than 10 cm for a laser stabilized at line center.* Lasers of this length have several disadvantages, the most serious of which (in the case of ion lasers) results from the close proximity of the Brewster windows to the discharge. In the case of the 6328 Å He-Ne laser, restricting the laser tube length to a few centimeters limits the available single axial mode output power to about 1 mw.

This brief describes two additional mode suppression techniques which permit single axial mode operation with mirror separations much larger than $c/2\Delta \nu_a$.

In common with other techniques,[3] mode suppression is achieved by splitting the beam and using a three-mirror cavity;* however, the specific configurations described in this note are believed to be new and may possibly be advantageous in certain circumstances.

The first technique is most simply understood by referring to Fig.

* The mode competition effects reported by Rigrod and Bridges (Electron Device Research Conference, University of Illinois, Urbana, June, 1965) for ion lasers may permit some increase of these lengths.
* It has been pointed out by E. I. Gordon and J. E. Geusic that all mode suppressors using a beam splitter and three mirrors are probably members of the same class of devices characterized by the microwave magic tee with two shorted arms.

1(a) which illustrates one of several possible versions. In Fig. 1(a), G is a conventional Brewster window gas laser; M_1, M_2, and M_3 are spherical mirrors; and C is a birefringent crystal such as calcite with antireflection coated faces. The crystal is oriented so that plane polarized light defined by the Brewster window is split by the crystal into two parallel, orthogonally polarized beams (not necessarily of equal intensity) displaced from each other by a small distance, d, determined by the length of the crystal. The radius of M_3 is chosen to match the spot size and curvature of the beam formed by mirrors M_1 and M_2.

For light of the correct polarization and for a given transverse mode, the cavity has low loss only for those frequencies for which the optical path length L_{13} is an integral number of half wavelengths longer than L_{12}.* By proper adjustment of L_{23}, the low-loss modes for the "bicavity" may be spaced sufficiently far apart so that only one mode appears within the oscillation linewidth. This mode may be interpreted as a coincidence between axial modes of the cavities formed by mirrors $M_1 - M_2$ and $M_1 - M_3$ separately as indicated in Fig. 1(b). The difference in mode spacing

$$\Delta \nu_m = \frac{c}{2}\left(\frac{1}{L_{12}} - \frac{1}{L_{13}}\right)$$

must be large enough so that oscillation at modes adjacent to a coincidence does not occur† and small enough so that only one (or at most two) coincidences occur within the oscillation linewidth. The separation between coincidences, $\Delta \nu_c$, is given by $c/2(L_{13} - L_{12})$. If, for example, $L_{12} = 100$ cm, $L_{13} = 115$ cm, $\Delta \nu_m = 19.5$ mc and $\Delta \nu_c = 1000$ mc.

Numerous variations of the same general technique exist. Since calcite crystals capable of yielding beam displacements larger than 1–2 mm are difficult to obtain, the beam splitter may conveniently be a Wollaston or Rochon prism which provides a few degrees of angular separation, as illustrated in Fig. 2(a). In order to achieve higher single mode powers, the arrangement shown in Fig. 2(b) is advantageous since the tube length can be approximately doubled with no increase in over-all cavity length.

Because of the large difference in refractive index for e and o rays

---

* Light from M_2 and M_3 for which $(L_{13} - L_{12})$ is not an integral number of half wavelengths is elliptically polarized when recombined by the calcite prism and one component is partially reflected out of the cavity by the Brewster window.

† It can be shown that the cavity loss at modes adjacent to a coincidence increases with $(L_{23} - L_{12})$; consequently, the discrimination against adjacent modes increases with $\Delta \nu_m$.
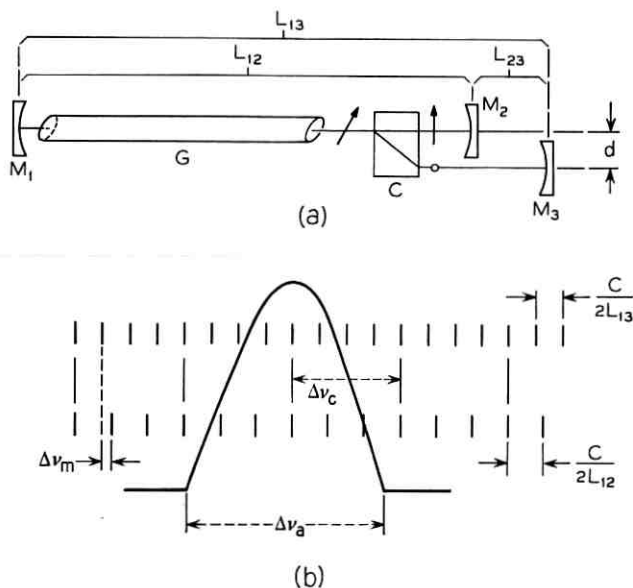
Fig. 1 — (a) Mode suppression technique using calcite rhomb beam splitter. (b) Coincidences between modes of the cavity shown in 1(a).
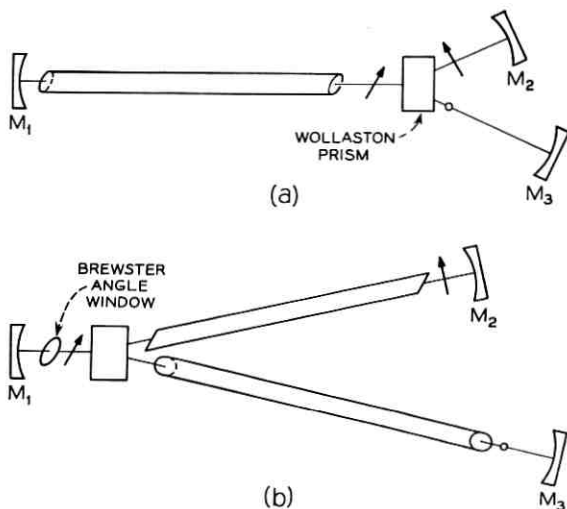

Fig. 2 — (a) Mode suppression technique using a Wollaston prism beam splitter. (b) Technique for increasing output power with no increase in over-all length.

in calcite, it is impossible to eliminate reflection for both rays simultaneously by antireflection coating unless birefringent coatings can be developed. This index difference is much smaller for crystal quartz, consequently the reflection coefficient can be made low for both rays. The deviation angle is reduced to about 2° which is still sufficiently larger than the diffraction angle of the beam in most gas lasers to allow complete separation of the beam in a few centimeters distance.

The feasibility of the bicavity approach has been demonstrated using a calcite Wollaston prism in conjunction with 60 cm long 6328 Å gas laser as illustrated in Fig. 2(a). In order to reduce reflection losses, glass windows having $\lambda/4$ cryolite coatings on one surface were oiled to each prism face. All three highly-reflecting (>99.6%) mirrors had a radius of curvature of 2 meters. L_{12} and L_{13} were approximately 122 cm and 142 cm, respectively, giving a difference in mode spacing $\Delta\nu_m$ of 17.3 mc and a separation between coincidences $\Delta\nu_c = 750$ mc. Since the 6328 Å oscillation linewidth is ≈ 1500 mc at most two axial modes can oscillate in this arrangement. The mode structure of the bicavity laser was examined with a plane parallel Fabry Perot interferometer having a free spectral range of 3 gc. Provided an aperture was used in the laser cavity to discriminate against all but the lowest order transverse mode, at most two axial modes with a separation estimated at ≈ 750 mc were observed to oscillate. Single-mode operation was obtained intermittently as one of the bicavity modes passed through line center.

Because of poor index matching, most of the output power appeared as reflections from the Wollaston prism with only a small amount (≈ 2 mw) being transmitted through the high-reflectivity mirrors. Reducing the reflection loss by using a quartz prism and low-loss coating should permit single-mode powers in excess of that available in a single axial mode from a multimode laser of the same length, as pointed out by Smith.[3]

A second technique, simpler than the first, but more restricted in application, makes use of the symmetry of the TEM_{01} mode. The line of symmetry of this mode (and other odd symmetric modes) is a region of zero field, consequently, splitting the mirror along this line should not appreciably increase the loss for this mode. One may form a bicavity configuration by placing each half of the split mirror at different distances from the common mirror as shown in Fig. 3(a). In order to minimize losses, the spot size at the mirror should be made as large as the tube bore permits by using a nearly concentric cavity.

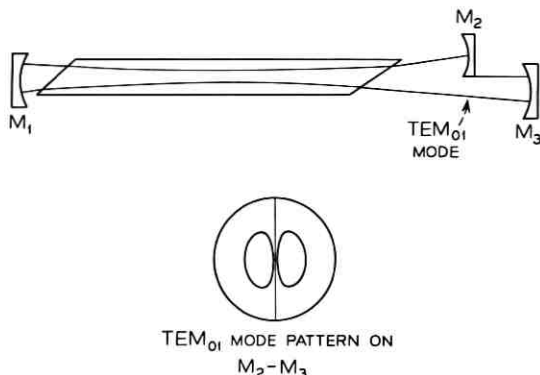The split mirror approach has been demonstrated experimentally,

TEM$_{01}$ MODE PATTERN ON
M$_2$–M$_3$

Fig. 3 — Mode suppression technique suitable for TEM$_{01}$ mode.

using for this purpose 1 meter radius mirrors spaced approximately 176 and 183 cm apart. Care was taken to produce a sharp edge at the front surface of the nearest split mirror (M$_2$ in Fig. 3(a)) to minimize scattering losses. The TEM$_{01}$ mode was selected by using an aperture at the common mirror. An interferometer was used to verify single axial mode operation as in the previous experiment.

Since each half of the split mirror has to be mounted separately, some difficulty was experienced in aligning the cavity to produce the desired mode. However, once the mirrors were oriented correctly, the mode was found to be quite stable.

Because of the increased diffraction angle, the TEM$_{01}$ mode is not the best choice for beam propagation. Unfortunately, there seems to be no way of applying the split mirror approach to the TEM$_{00}$ mode, but it may be possible to suppress axial modes of the "donut" TEM$_{01}$ circular mode.

The author is grateful to his colleagues for interesting and constructive discussions.

REFERENCES

1. Kogelnik, H. and Patel, C. K. N., Mode Suppression and Single Frequency Operation of Gaseous Optical Maser, Proc. IRE, 50, 1962, p. 2365; Kleinman, D. A. and Kisliuk, P. P., Discrimination Against Unwanted Orders in the Fabry-Perot Resonator, B.S.T.J., 41, 1962, p. 453.
2. Collins, S. A. and White, G. R., Interferometer Laser Mode Selector, Appl. Opt., 2, 1963, p. 448.
3. Fox, A. G., Patent Application; Smith, P. W., Stabilized Single Frequency Output from a Long Gas Laser, J. Quant. Electron. QEI, Nov., 1965; Di Domenico, M., to be published in Appl. Phys. Lett.