THE BELL SYSTEM TECHNICAL JOURNAL

THE BELL SYSTEM TECHNICAL JOURNAL is published monthly, except for the May–June and July–August combined issues, by the American Telephone and Telegraph Company, C. L. Brown, Chairman and Chief Executive Officer; W. M. Ellinghaus, President; V. A. Dwyer, Vice President and Treasurer; F. A. Hutson, Jr., Secretary. Editorial inquiries should be addressed to the Editor, The Bell System Technical Journal, Bell Laboratories, Room 1J-319, 101 J. F. Kennedy Parkway, Short Hills, N. J. 07078. Checks for subscriptions should be made payable to The Bell System Technical Journal and should be addressed to Bell Laboratories, Circulation Group, 101 J. F. Kennedy Parkway, Short Hills, N.J. 07078. Subscriptions $20.00 per year; single copies $2.00 each. Foreign postage $1.00 per year; 15 cents per copy. Printed in U.S.A. Second-class postage paid at New Providence, New Jersey 07974 and additional mailing offices.

Comments on the technical content of any article or brief are welcome. These and other editorial inquiries should be addressed to the Editor, The Bell System Technical Journal, Bell Laboratories, Room 1J-319, 101 J. F. Kennedy Parkway, Short Hills, N. J. 07078. Comments and inquiries, whether or not published, shall not be regarded as confidential or otherwise restricted in use and will become the property of the American Telephone and Telegraph Company. Comments selected for publication may be edited for brevity, subject to author approval.

Video Colorization Diagnostics in Optical Telecommunications

By H. M. PRESBY and R. CHANG *

(Manuscript received October, 1981)

A new diagnostic tool is used to study the outputs of light-emitting diodes, lasers, and optical fibers, and to provide quality control of optical fiber preforms. The object is observed with a video camera, and the monochrome signal is processed to synthesize a color display from the different intensity levels. The color representation allows features not previously observable to be readily recognized and characterized. Correlation between separated, but equally bright, regions of a video image is also easily achieved in that they are now rendered in the same color. A calibration procedure allowing quantitative information to be quickly obtained from the display is described.

I. INTRODUCTION

Video diagnostics is capable of playing an important role in the study and characterization of the components used in optical telecommunications. The term itself refers to obtaining information from a video signal derived by viewing the object, or a property of the object of interest, with a video camera. Prime examples are observing the light output of light-emitting diodes (LEDs) or of fibers. The video image will then consist of a two-dimensional display of the intensity at the plane upon which the camera is focussed. For LEDs, this plane could correspond to the output face of the device so that the near-field

---

radiation pattern can be characterized and, in the case of fibers, it could be a screen upon which the fiber output is allowed to fall.

The display contains a very large amount of information in the form of point-to-point variations in intensity. This information could of course be computer-processed to yield the desired data and, indeed, a host of such systems exist in both specialized[1,2] and commercially available form.

However, in many applications, this sophistication and associated expense is not required as the desired information could be obtained by simply observing the monitor display if some means existed of distinguishing the various intensity levels. Unfortunately, it is difficult for the monochrome monitor screen and the eye to discern small changes in grey scale and to correlate them over an extended two-dimensional image. On the other hand, the human visual system is very sensitive to small color changes, which can then be used as a means of image enhancement.

This paper describes a method based on the latter effect to achieve distinction and subsequent correlation in a video image by converting different shades of grey into synthetic color signals. We apply the term video colorization to this procedure in analogy with video digitization, since the signal is represented by a discrete number of synthetic colors. By just observing the color display, all of the required information for many applications can then be obtained. Examples are given of studies of LED, laser, and fiber output light distributions and also of the internal structure of optical fiber preforms.

II. VIDEO COLORIZATION

The advantages of converting monochrome displays into color is receiving considerable attention[3] and is making a major impact in computer graphics.[4] Not only does color separate and highlight information better but it also serves to direct attention to particular areas of interest. This latter point has been found to improve operator recognition and lower probability of operator error in situations where judgments are based on monitor displays. In addition to its laboratory research value, the system to be described here is well-suited to a manufacturing environment for use by personnel in product testing and quality control.

Video colorization diagnostics is achieved with the arrangement shown in Fig. 1. A video camera observes the object or intensity distribution of interest either with a camera lens, through a microscope or by projecting directly onto the face of the vidicon itself. The vidicon is a silicon target type adjusted for linearity and uniformity of response.

The output of the camera is sent to a video quantizer.[5] This instru-

ment is specifically designed to process the grey-scale characteristics of a monochrome video input signal to synthesize color signals from different shades of grey. It achieves this by dividing the input signal into a maximum of eight slices and assigning a color to each slice. The width of the slices is adjustable to any amplitude level between black and white, and the color assigned to each slice can be any combination of red, blue, and green.

The video quantizer generates a test pattern on the red-green-blue color monitor to facilitate the above adjustments. Three such patterns which appear as vertical bands of color are shown in Fig. 2. The width of each band corresponds to a voltage range in the video signal. Optimal resolution is obtained if the total width of the bands is equal to the total voltage range of the input signal.

Figure 2a shows the condition in which each band has an equal width and the intensity progression goes from black to white with the colors in the order, black, red, yellow, green, light blue, dark blue, violet, and white. Thus, the darker parts of the image will be rendered black and red and the brighter segments in white and violet, with the other colors falling in between.

Figure 2b shows the test pattern setup for a logarithmic response and Fig. 2a, for equal brightness contour generation. In the latter case, as will be shown later, the image on the monitor is composed of a series of equal brightness contours.

The logarithmic response is of particular value in calibrating the conversion of grey scale into color. Figure 3a shows the monochrome display of a neutral-density step wedge as observed by the video camera. Eight steps are included in the field of view varying in density from 0.22 to 1.36, giving a range of relative brightness from 100 to 7 percent. Note that only four of the steps are clearly resolved. With the test pattern set for a logarithmic response to match the density steps of the wedge, Fig. 3b is observed. Note that now all eight steps are seen and that a particular color can be associated with a given grey-scale value. We now present the results of applying this system to a variety of objects.

III. OPTICAL FIBER OUTPUT

The output radiation pattern of a multimode optical fiber consists of a complex configuration determined by the mutual interference effects of the propagating modes. A typical pattern observed for laser excitation of the fiber is shown in monochrome form in Fig. 4a. The main feature noted is patches of light of varying intensity. Generally, from a systems perspective it is only the total power in this distribution which is of interest; however, recent attention being paid to modal noise makes details of the pattern important.[6] The intensity distribu-

tion is also of interest in characterizing the fiber or mode distributions themselves, and properties of it can be used in multimode fiber interferometry.

Characterizing the light output distribution or focusing attention upon a specific feature of it is made possible with video colorization. A colorized intensity pattern is shown in Fig. 4b. Regions of equal intensity are now clearly delineated and the brightest patches of light stand out in white. For comparison, an equal brightness contour display of the pattern is shown in Fig. 4c.

The output light distribution is extremely unstable to small perturbations of the fiber. Touching the fiber at any point along its length dramatically changes the observed pattern. Even air currents are sufficient to perturb the distribution as displayed in Fig. 5, which shows the output as observed over a several-minute time interval under normal ambient room conditions. If the total power of this distribution is not detected (as may occur with certain fixed apertures or splices in a fiber system), these variations become an important source of modal noise and video colorization offers a way to study the effects. Video colorization also opens up other possibilities in that now a specific feature of the pattern can be easily monitored and its changes as a function of external stimuli recorded. This allows multimode fibers to be used in sensitive interferometric-like detection schemes to monitor quantities like, pressure, temperature, or even electric and magnetic fields.[7]

IV. LIGHT-EMITTING DIODES AND LASERS

Video colorization is ideally suited to characterize the light outputs of LED and laser sources. The output consists of a two-dimensional intensity distribution which can be immediately analyzed once it is colorized. The LED devices are observed with a microscope focussed on their emitting surface to obtain the near-field light distribution.

Figure 6 shows a poor quality LED observed in black and white and in colorized form. The most intense emission regions are clearly seen in the colorized photos which show the output of the device as a function of increasing current.

An LED with a relatively uniform output light distribution is shown in Fig. 7. Note that the colorized display shows some nonuniformity in the contact region (near the upper left) and a small bright area in violet just to the left of center.

Two views of the output of a GaAlAs laser are shown in Fig. 8. Figures 8a and 8b show the near-field pattern, and Figs. 8c and 8d show the far-field pattern; the latter is obtained by focussing on a ground glass upon which the radiation impinged.

V. PREFORM STRUCTURE

As is well known, the quality of the refractive index profile of an optical fiber determines, in the absence of mode coupling, the bandwidth of the medium. In the modified chemical vapor deposition (MCVD) process, the profile is built up in a layered fashion by varying the dopant concentration in the deposited layers. Layer structure, then, is an important indicator of profile quality and techniques have been developed to render it visible for inspection in the preform state.[8]

One way of investigating the layer structure is by immersing the preform in matching oil and observing the core with a video camera.[9] A preform observed in this manner is shown in Fig. 9a. The central dark band outlined with two bright lines is the axial index depression. The two bright lines further out arise from a deposition layer with an index much higher than that of the adjacent layers and represent a deposition flaw. While this perturbation is visible in the monochrome display, its presence can be enhanced by video colorization as shown in Fig. 9b. In this display, the test pattern was set so that all grey levels below a certain value would be rendered in red. Grey levels above this threshold, representing deposition flaws, are rendered in yellow. Thus, a difinitive indicator is provided to check preform quality.

Similar observations on a single-mode preform are shown in Fig. 10. Figure 10a shows in monochrome the GeO_2-doped core in the lower region and the phosphosilicate layers above it which make up the deposited cladding. Figure 10b is a colorized display with the test pattern set to show up the uniform deposition structure in the cladding. For Fig. 10c, the test pattern was adjusted to best show uniformity of the core deposition.

Video colorization is also an aid in viewing the layer structure as observed in preform slab samples. Figure 11 shows such a sample and how colorization helps in making visible the fine details within the layers.

As a last example, consider Fig. 12 which demonstrates a use of general utility. The monochromatic display in Fig. 12a comes from viewing the output of a collimated tungsten-halogen light source intended to supply a uniform intensity distribution. How well this uniformity is achieved is clearly seen in the colorized display of Fig. 12b. We, thus, have a very sensitive procedure to test background or incident light distributions for uniformity.

In conclusion, we have seen that video colorization can be used as a tool that can enhance and, in some cases, even make possible the rapid interpretation of video images. We have investigated images arising from a number of areas related to optical fibers and are confident that this technique can be used to advantage in many other studies as well.

REFERENCES

1. H. M. Presby et al., "Rapid Automatic Index Profiling of Whole-Fiber Samples: Part II," B.S.T.J. *58*, (1979) pp. 883–902.
2. D. Marcuse and H. M. Presby, "Automatic Geometric Measurements of Single-Mode and Multimode Optical Fibers," Appl. Opt. *18* (1979) pp. 402–8.
3. P. Snigier, "Graphic Display Devices," Digital Design (April 1980), pp. 39–44.
4. Staff Report, "The Color Impact in Graphic Display," Optical Spectra (April 1980), pp. 46–53.
5. The Video Quantizer used in this work is a Colorado Video, Model 606A.
6. E. G. Rawson, J. W. Goodman, and R. E. Morton, "Experimental and Analytical Study of Model Noise in Optical Fibers," Proc. Sixth Conf. Opt. Commun., York, England, September 1980, pp. 72–5.
7. H. Heaton, "Thermal Straining in a Magnetrostrictive Optical Fiber Interferometer," Appl. Opt. *19* (1980), pp. 3719–20.
8. H. M. Presby and D. Marcuse, "Optical Fiber Preform Diagnostics," Appl. Opt. *18* (1979), pp. 23–30.
9. H. M. Presby and D. Marcuse, "Preform Index Profiling (PIP)," Appl. Opt. *18* (1979), pp. 671–7.
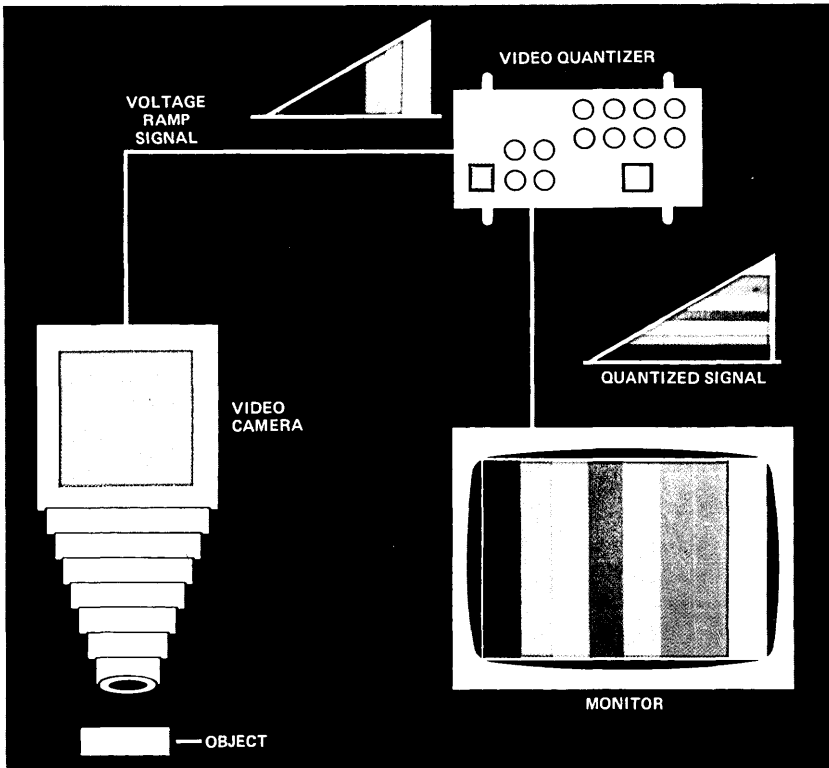
Fig. 1—Arrangement for video colorization diagnostics.
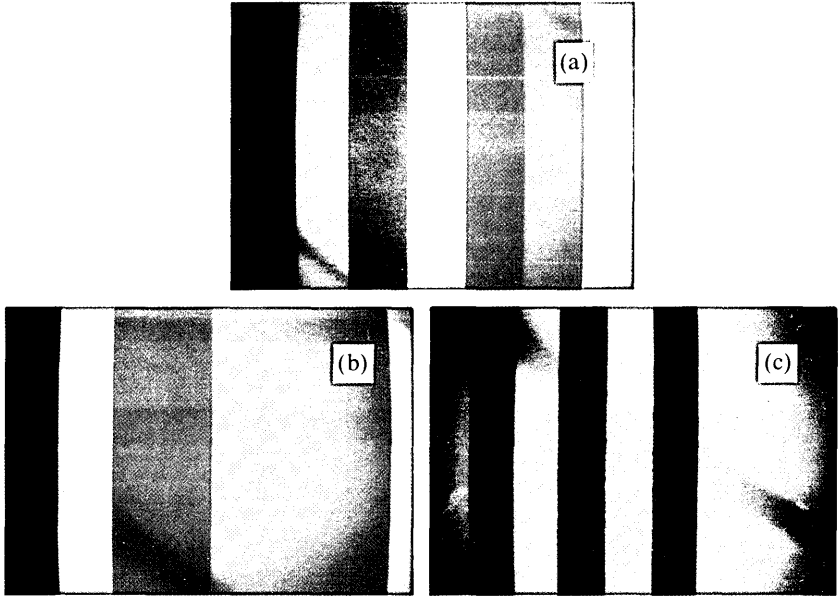
Fig. 2—Test pattern of video quantizer set for (a) bands of equal width; (b) logarithmic response, and (c) equal brightness contour generation.
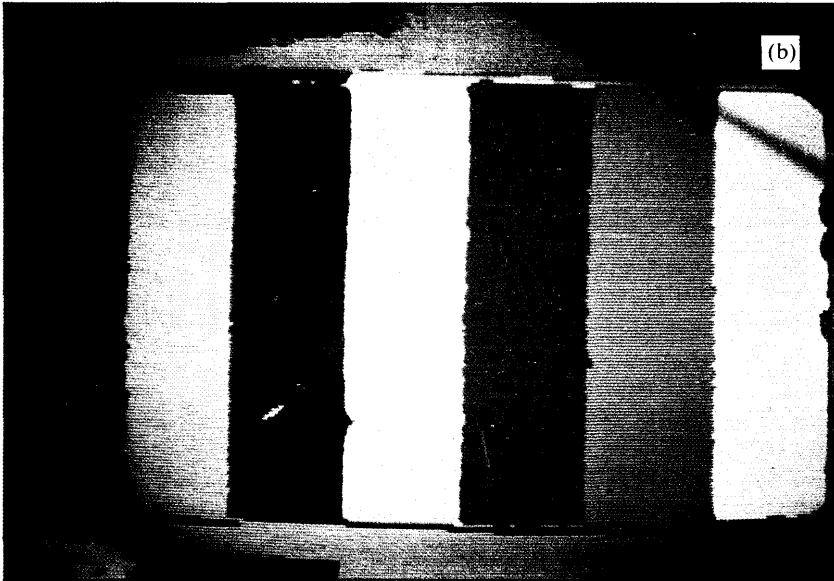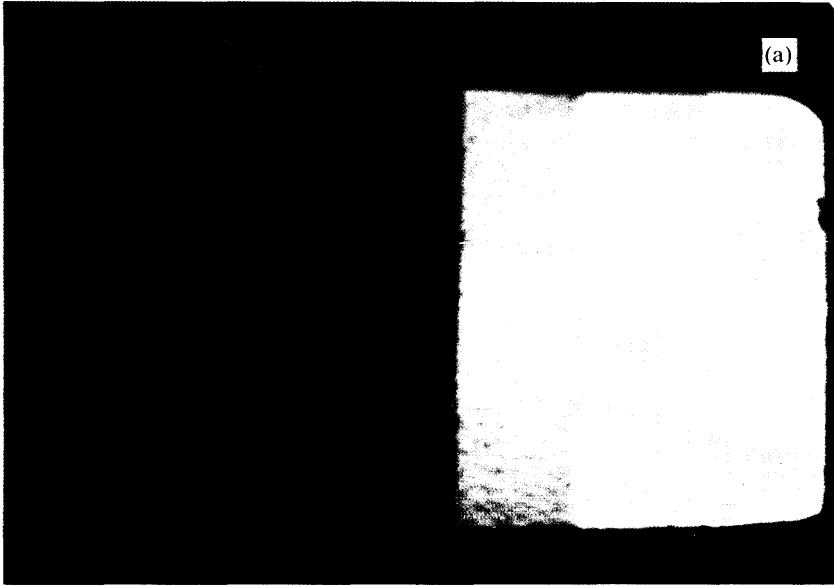
Fig. 3—Neutral density step wedge observed (a) monochromatically and (b) and (c) by colorization. Note that all eight steps are clearly seen in the latter case.
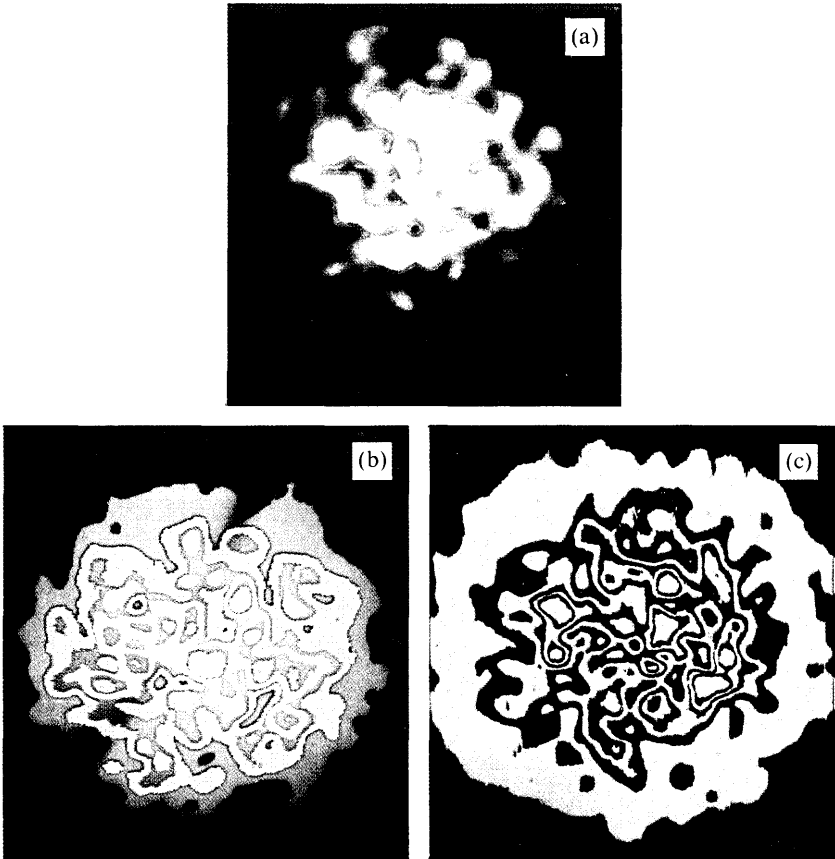
Fig. 4—Output light distribution of multimode optical fiber observed (a) monochromatically, (b) by colorization, and (c) by equal brightness contours.
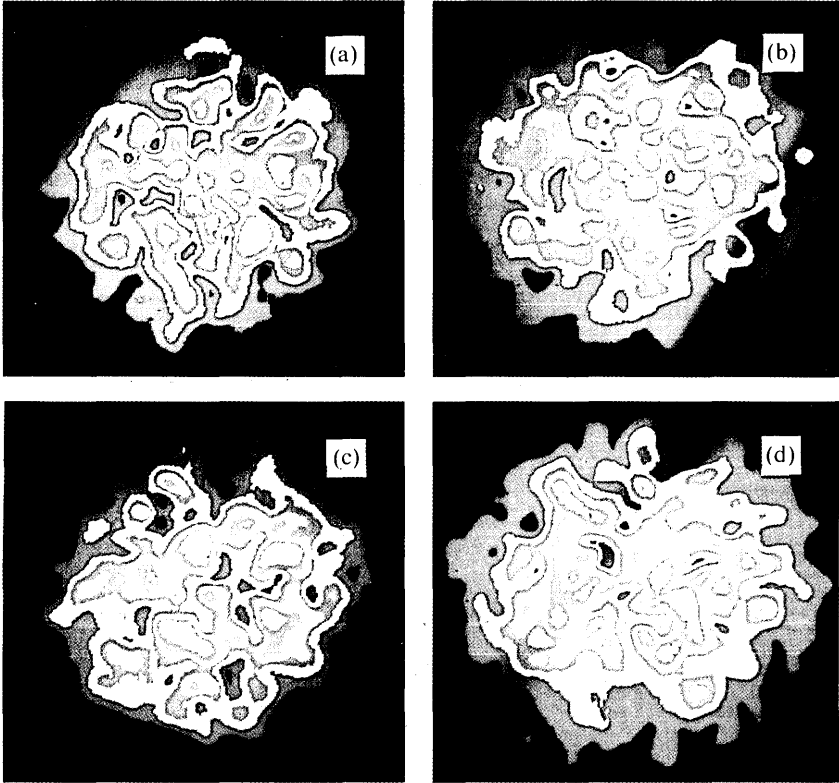
Fig. 5—Variations observed in output light distribution of multimode optical fiber as a function of several minutes' time under ambient room conditions.
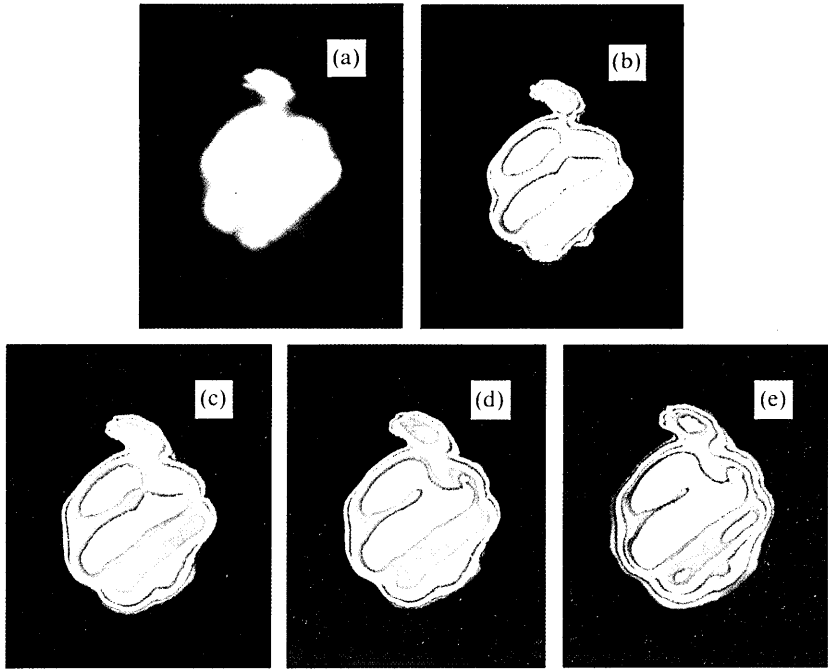
Fig. 6—Near-field light output of poor quality LED observed in black and white and in colorized form as a function of increasing current [(a) through (e)]. Note the changes in the brightest emitting regions shown in white.
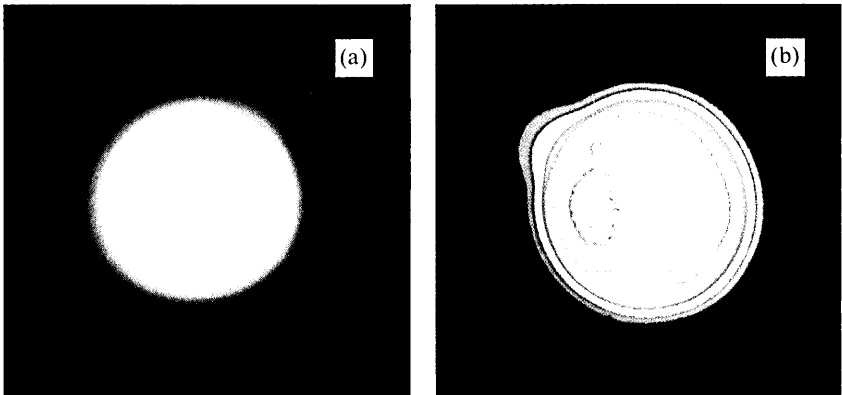
Fig. 7—Light-emitting diode with relatively uniform light output observed in non-colorized and colorized form.
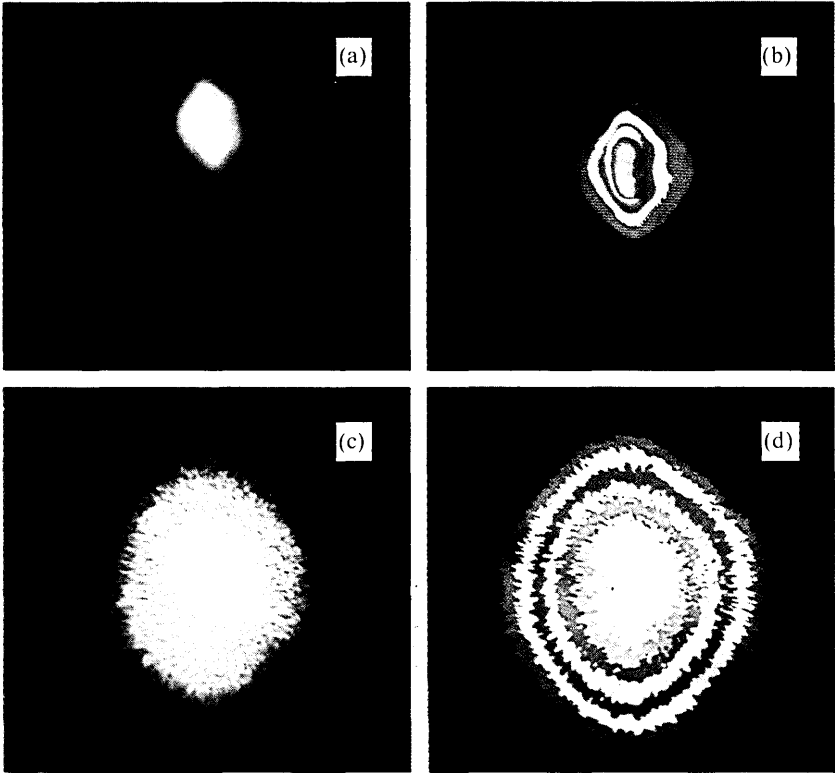
Fig. 8—Near-field [(a) and (b)] and far-field [(c) and (d)] light distribution of GaAlAs laser.
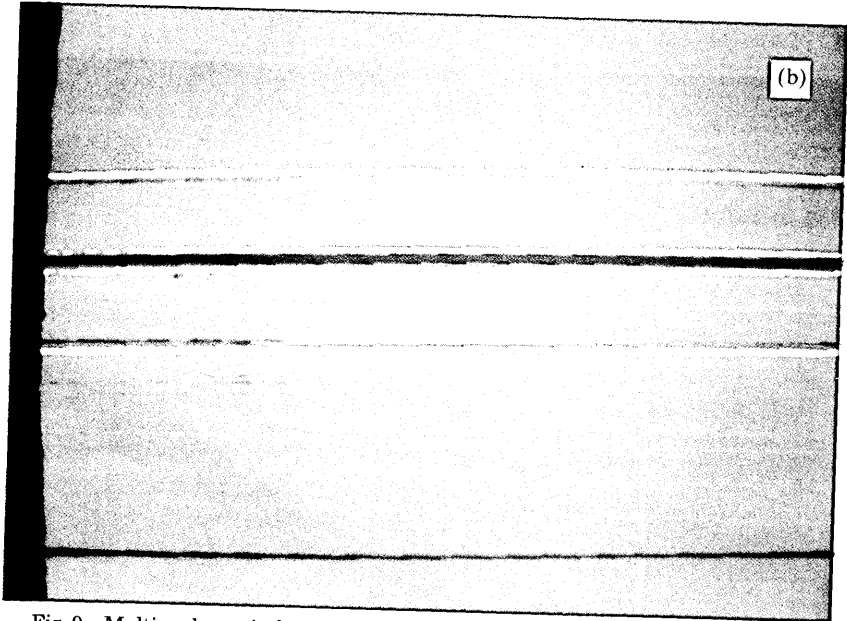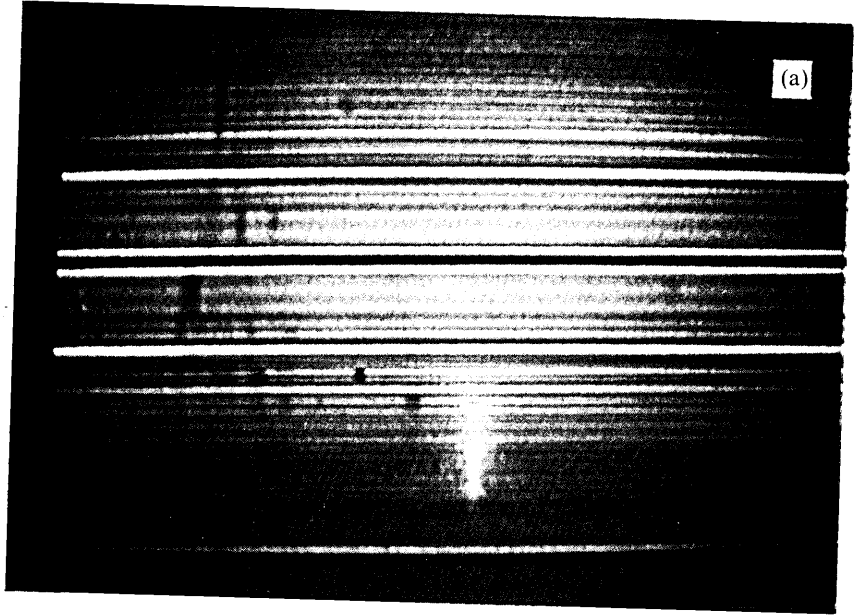
Fig. 9—Multimode optical fiber preform with layer perturbation, observed in (a) black and white and (b) colorization. The colorized form showing how significant process deviations can be readily detected.
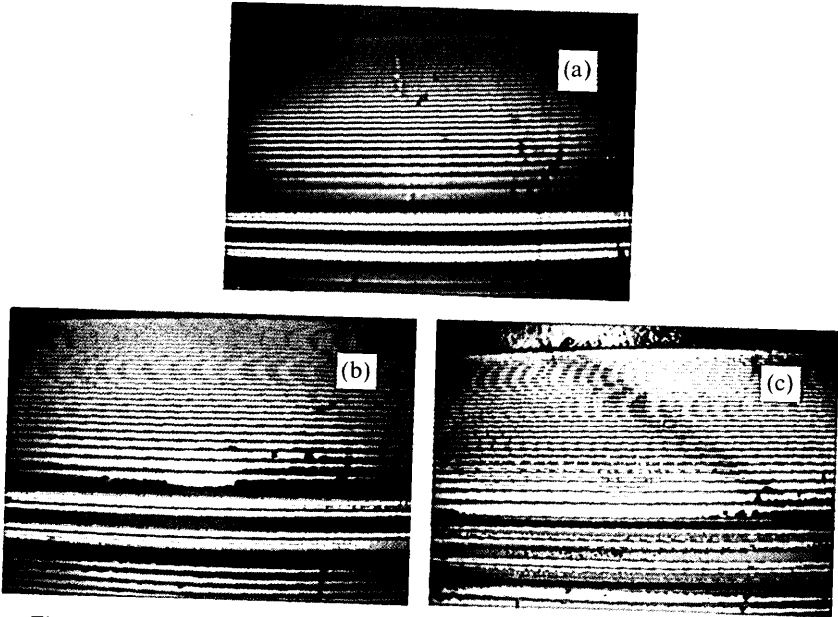
Fig. 10—Observations of single-mode optical fiber preform in (a) monochrome and (b) and (c), colorized form.
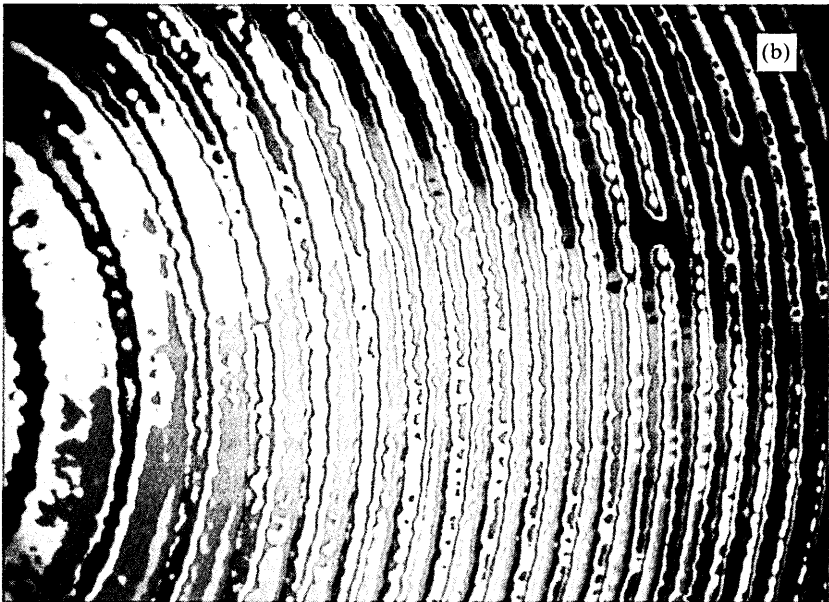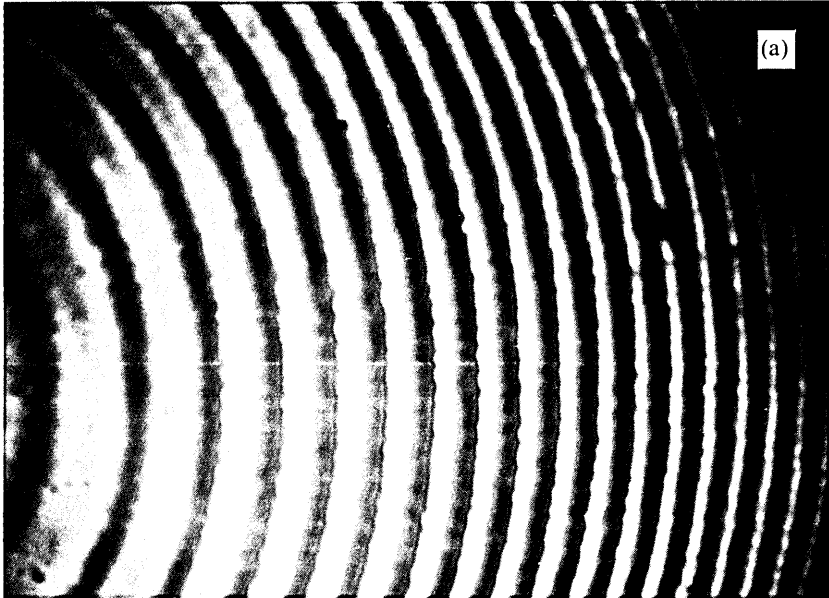
Fig. 11—Slab sample from tip of optical fiber preform showing improved rendering of layer structure achieved with video colorization.
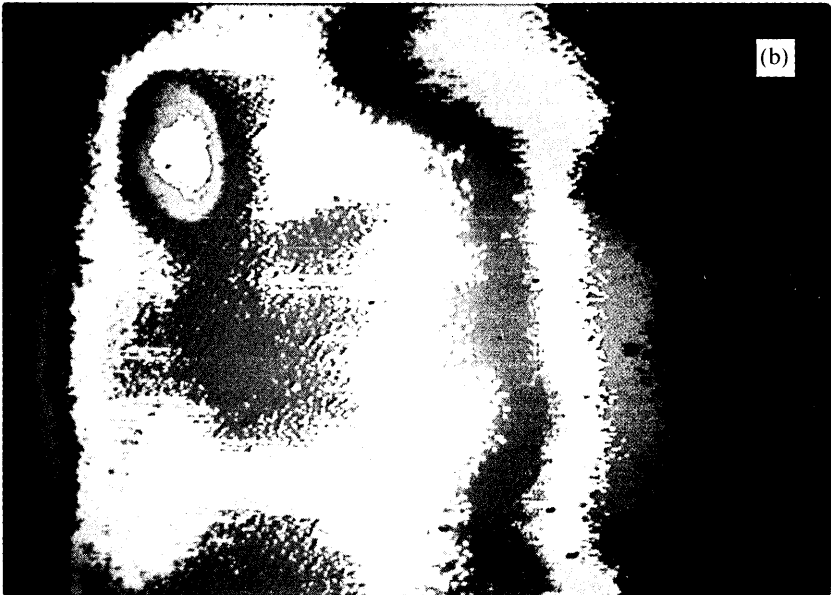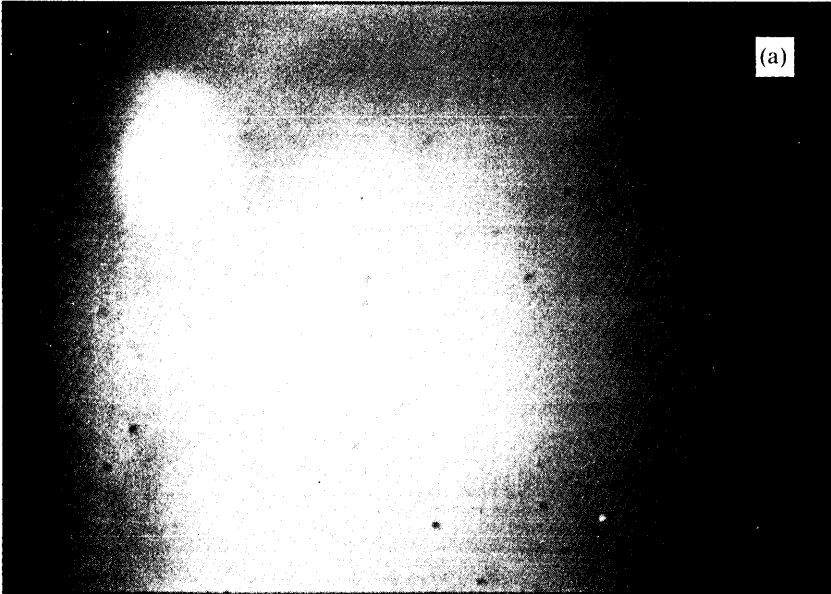
Fig. 12—Light output of tungsten-halogen source showing how nonuniformities are clearly seen with video colorization.

Measurements of OH Diffusion in Optical-Fiber Cores

By D. L. PHILEN

A color-center laser was used at 0.95 µm to determine the activation energy for OH diffusion in the cores of current multimode fibers. The temperature range of 600 to 800°C was investigated and the activation energy determined to be 19,700 ± 1600 cal/mole in agreement with previous values reported for bulk silica. Based on this value, some 23,000 years would be necessary at 90°C to produce a 3-percent change in the 0.95-µm OH absorption at the core center. Apparently, OH diffusion over a service life of optical fibers should not be a problem in presently envisioned lightwave applications.

I. INTRODUCTION

The diffusion of water (H_2O) into optical fiber cores can affect both fiber loss[1] and bandwidth.[2] Water contamination of the fiber core could occur during three different periods—preform manufacture, fiber drawing, and long-term environmental exposure. The first two have previously been studied;[3-5] however, environmental data on fused silica require longer exposure times[6] because of slow diffusion rates. It would be ideal to have an activation energy which would allow the calculation of concentration changes as a function of temperature. Thus, the effect of water diffusion in a fiber could be calculated based on the expected operating conditions for that fiber. Reported here is the first measurement of the activation energy for OH diffusion in a fiber core.

In the transmission wavelength region of current interest, 0.8 to 1.5 µm, two major OH absorption bands appear. These bands are at 1.38 and 0.95 µm and are, respectively, the first and second overtone bands of the fundamental stretching vibration at 2.7 µm. Two candidates for system wavelengths are 0.87 and 1.3 µm—very close to the OH absorption bands. A 10-ppm increase of OH would increase the loss by about 12 dB/km at 0.95 µm, 1.0 dB/km at 0.90 µm,[7] and 7 dB/km at 1.3 µm.[8] Thus, knowledge of the OH diffusion rate in optical fiber waveguides

becomes increasingly important as attention shifts to longer wave-lengths.

II. THEORY

Water probably does not exist as a molecule of H_2O in glass. On attacking a silica surface, water undergoes the following reaction:

$$H_2O + (\equiv Si—O—Si\equiv)_{glass} \rightleftarrows 2(\equiv Si—OH)_{glass}.$$
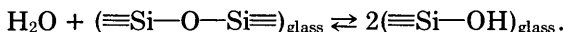
This process is interpreted as the conversion of one molecule of water into two OH groups by the addition to an Si—O—Si bridge to form two (Si—OH) groups. This mechanism is supported by data indicating the solubility of water in silica (concentration) to be proportional to the square root of the vapor pressure.[7]

Diffusion is then viewed as the reformation of the Si—O—Si bridge and the transfer of the OH group to an adjacent bridge.

$$(\equiv Si_A—OH) + (\equiv Si_B—O—Si_B\equiv) \rightleftarrows$$
$$(\equiv Si_B—O—Si_A\equiv) + (\equiv Si_B—OH).$$
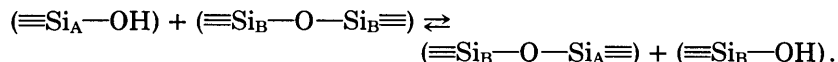
This process is unlike ordinary diffusion where preferred sites are involved, and it is probably incorrect to refer to OH in glass as OH ions or OH radicals since charged ions or free radicals are unlikely to exist in the above process. The measured loss attributed to OH is simply that arising from the O—H stretching vibration of the OH functional group.

The diffusion of a line source located at the origin in cylindrical coordinates is given by:

$$C(r, t) = \frac{\alpha}{4\pi Dt} e^{-r^2/4Dt} \qquad (1)$$

(see Appendix A for derivation),[9] where

α = concentration per unit length of the line source
C = concentration (mole fraction) at radius r and time t
D = diffusion coefficient
r = distance from the line source.

If r is in cm and t in seconds, then D has dimensions of cm^2/s. The diffusion coefficient is a function of temperature and is given by the following equation:

$$D = D_0\exp(-E/RT), \qquad (2)$$

where

D_0 = a constant
E = the activation energy
R = gas constant in cal mole^{-1} kelvin^{-1}
T = temperature in kelvin.

Equations (1) and (2) are used to obtain the activation energy E.

Two problems of interest are the diffusion of water from the cladding into the core of the fiber and from water surrounding the fiber into the cladding and the core. Since light does not propagate with low loss in the cladding, optical loss measurements in the cladding are difficult. Measurement of the diffusion of water out of the core region is easier since optical loss measurements in the fiber core are used for fiber diagnostics. Because of microscopic reversibility, the diffusion coefficient inside the core is the same for diffusion into or out of the core. This approach avoids surface recombination effects observed when H_2O diffuses onto and off a silica substrate.[10] The question we address is, Given that water is already inside the fiber, what is its diffusion rate into or out of the fiber core?

III. EXPERIMENTAL

The experimental setup illustrated in Fig. 1 incorporates a tunable, single-mode, color-center laser.[11] This laser is pumped by a continuous wave krypton-ion laser and is tuned to the OH absorption peak at 0.95 μm. The reel of fiber chosen for test was 1.9 km long and had abnormally high loss at 0.95 μm because of OH absorption. Fifty meters were reeled off and the OH spatial profile determined at the end of that 50-meter section by measuring the absorption component of the loss calorimetrically, while traversing the color-center laser beam across the fiber core. About 100 meters more were reeled off and the OH spatial profile measured again at the end of the 100-meter section. These two spatial profiles agree with each other within the
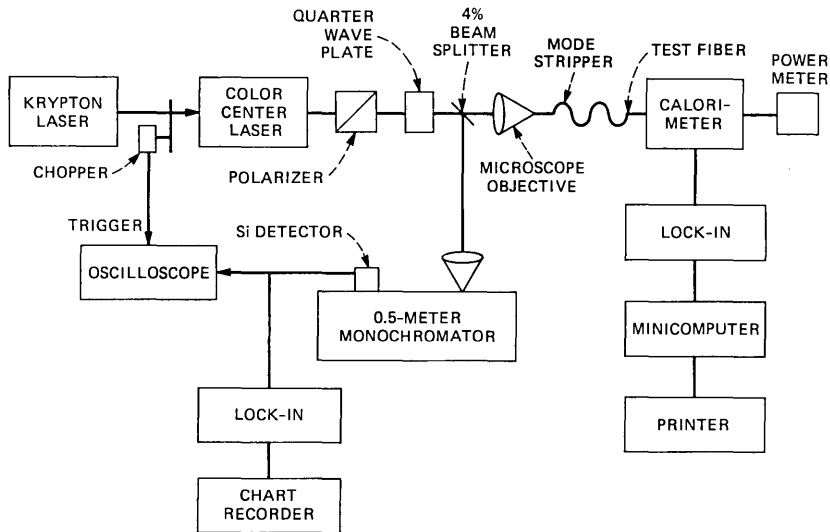
Fig. 1—Experimental setup for OH diffusion measurement.
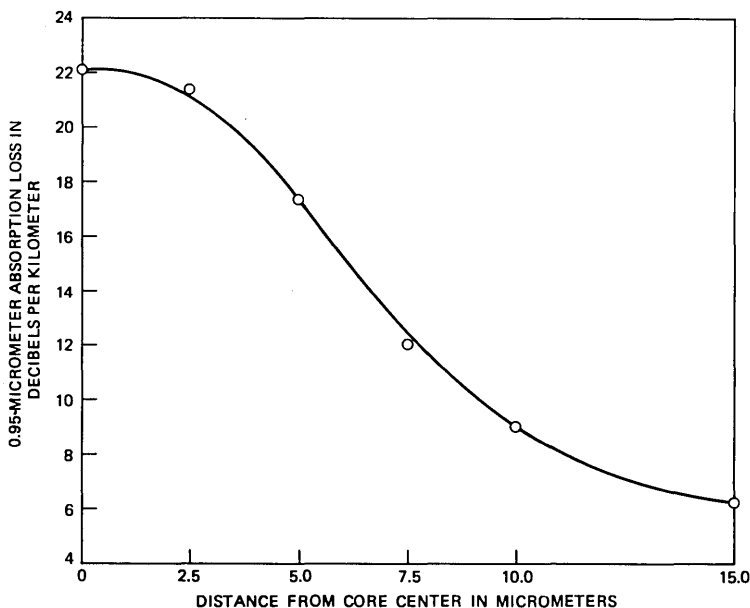
Fig. 2—The 0.95-μm loss because of OH in the control fiber.

error of the calorimeter measurement (±0.2 dB/km). This 100-meter section was used as the fiber for the diffusion experiments. Finally, the spatial profile was measured at the beginning of the 100-meter reel. This represents the control fiber and is illustrated in Fig. 2. It shows the OH absorption profile to be peaked at the fiber core center and decreasing towards the core-cladding boundary. These precautions ensured that the OH distribution was uniform over the length of the test fiber and ensured the same initial conditions for each test. This fiber had a germania-phosphosilicate core, with $\Delta = 1.3$ percent, and 1–2 percent phosphorus added for fining; its core was about 40 μm in diameter.

A number of two-meter samples were cut from the 100-meter fiber length. The coating was stripped from the samples, and each sample was heated in an oven for one day at a fixed temperature. Six temperatures were chosen to give measurable diffusion rates: 600, 655, 680, 700, 750, and 800°C. The oven temperature was monitored with a Chrome-alumel thermocouple and held constant within ±2°C. After removal from the oven, each sample was placed in the calorimeter and the OH spatial profile measured. The spatial resolution is limited by the focused spot size to about 5 μm.

Since this fiber has a core diameter of only 40 μm, the scattering loss due to the core-cladding boundary increased dramatically in the region near 20 μm. In Fig. 3 the spatial profiles for four of the temperatures

and for the unexposed control sample show how the absorption profile changes with increasing temperature. In the region of 15 μm, a rising loss versus radius may indicate in-diffusion from the cladding for the 800°C sample; however, this indication is not evident in the lower temperature samples. The rise is negligible at $r = 10$ μm, even for the 800°C sample; therefore, we neglected in-diffusion effects from the cladding in analyzing the loss measured at $r = 0$.

IV. RESULTS AND DISCUSSION

In evaluating the results, one must consider the following effects:

(*i*) The fiber as produced already contains diffused water as a result of preform manufacture and fiber drawing. Unfortunately, we do not know how long the preform or fiber was exposed to each temperature in this process. As a result, the OH in the starting fiber already has a Gaussian distribution rather than the line source distribution that would simplify the analysis.

(*ii*) In the measurement method used, the laser was focused to a spot on the fiber face and the loss at each radial position determined calorimetrically. Determining the concentration of OH by evaluating the loss at radius r in the fiber core is not straightforward since the exciting beam, localized at $r = r_1$, excites modes which cover a large range of r. To make the analysis tractable, only losses measured for
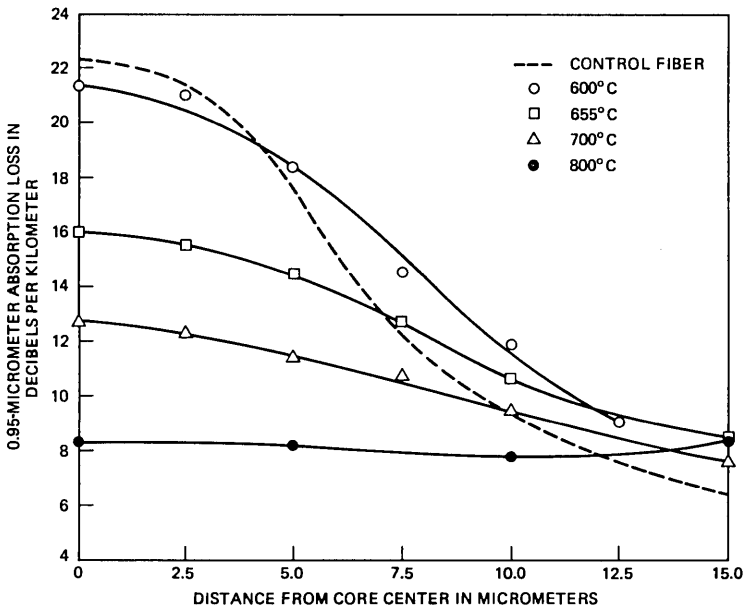
Fig. 3—Diffusion profiles for four of the exposure temperatures.

the $r = 0$ excitation condition will be used in deriving the activation energy.

Despite these complications, Appendix B shows that γ_{in} and $\gamma_{out}(T)$, the 0.95-μm absorption losses at $r = 0$ before and after exposing the fiber for one day at temperature T, are related by:

$$\ln[(\gamma_{in}/\gamma_{out}(T)) - 1] = \text{constant} - \frac{E}{RT}. \tag{3}$$

Figure 3 indicates that the $r = 0$ loss measured for the control fiber and the 600°C exposure are very nearly the same. Because of the uncertainty associated with each measurement, and their similar magnitude, the expression $\ln[(\gamma_{in}/\gamma_{out}(T) - 1]$ has a large uncertainty associated with it. Thus, the 600°C temperature was omitted from the data analysis.

Equation (3), plotted for the five remaining exposure temperatures in Fig. 4, should result in a straight line of slope $(-E/R)$. The value of the activation energy determined from the slope is 19,700 ± 1600 cal/mole or 0.854 ± 0.07 electron volts. While an uncertainty exists in the activation energy from thermodynamic considerations alone, the overall uncertainty in the measured value is the uncertainty in the slope.
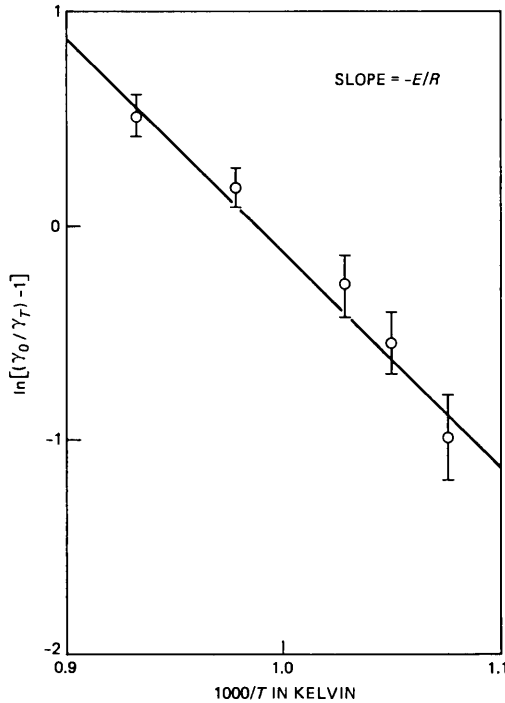
Fig. 4—Plot of eq. 3. The slope gives the activation energy.

It is this calculated uncertainty in the slope which is ±1600 cal/mole. This value for the activation energy compares quite favorably with the values 17,300 ± 2000 cal/mole and 18,300 ± 500 cal/mole determined between 600 and 1200°C in bulk silica by Moulson and Roberts.[10] They determined two different diffusion coefficients—one for absorbing of the OH centers,

$$D = (1.0 \pm 0.2) \times 10^{-6}\exp(-18{,}300 \pm 500)/RT \text{ cm}^2 \text{ s}^{-1}$$

and one for removal,

$$D = (2.7 \pm 1.0) \times 10^{-7}\exp(-17{,}300 \pm 2000)/RT \text{ cm}^2 \text{ s}^{-1}.$$

They point out, however, that the activation energy is the same within their experimental error and our value of 19,700 ± 1600 cal/mole overlaps theirs once the experimental uncertainty is taken into account.

Moulson and Roberts attribute the difference in the pre-exponential factor $(1.0 \pm 0.2) \times 10^{-6}$ for absorbing versus $(2.7 \pm 1.0) \times 10^{-7}$ for removal, to the process

$$H_2O + Si\text{—}O\text{—}Si \rightleftarrows 2(Si\text{—}OH)$$

being slightly exothermic by 6 kilocalories. Also, the recombination of hydroxyl groups at the glass surface may introduce an appreciable nonzero concentration of water just below the glass surface—an assumption made for their calculations, but unnecessary in this study.

From eq. (18), we can relate the different diffusion times for various temperatures. Suppose one fiber is exposed for time t at a temperature where the diffusion coefficient is D, and a second fiber for time t' at D'. If the fibers are initially identical ($\gamma_{in} = \gamma'_{in}$ in Appendix B) and the exposures are equivalent $[\gamma_{out}(T) = \gamma_{out}(T')]$, then eq. (18) gives $Dt = D't'$. From Fig. 3, apparently a temperature of about 600°C is required to produce diffusion in one day that results in a 3-percent change in the absorption at 0.95 μm. Using our measured value of E, the time required for this to occur at the highest acceptable exposure temperature in the field ($T' = 90$°C) is t', where

$$\frac{D_{600°C}}{D_{90°C}} = \frac{\exp(-E/RT)}{\exp(-E/RT')} = \frac{t'}{1 \text{ day}} = \frac{1.169 \times 10^{-5}}{1.375 \times 10^{-12}}$$

$$t' = 23{,}000 \text{ years}$$

which greatly exceeds the projected 40-year service life for an installed cable. Since the OH absorption loss at 1.3 μm is of the same order of magnitude as at 0.95 μm, this confirms that OH diffusion is not likely to be a problem for presently envisioned applications.[6]

We have made many assumptions in analyzing these data. We have assumed the diffusion coefficient to be independent of concentration,

and over the small concentration changes observed here, this is probably a valid assumption. We have assumed the activation energy to be independent of temperature. In addition, the equation used to model the data does not account for changes in viscosity or molal volume of the glass. Work is progressing to extend these measurements to lower temperatures to check for a temperature dependence of the activation energy. At a temperature of 325°C, this extension will require 6.2 months of exposure in the oven if the present calculations are correct.

Data on viscosities and molal volumes for ternary glasses like GeO_2—P_2O_5—SiO_2 are sparse or nonexistent. Moreover, little is known about the prediction of liquid diffusivities, to say nothing of diffusivities of very high viscosity materials such as glass. Refer to Refs. 12 and 13 for excellent texts on the subject of liquid and high-density transport phenomena.

V. SUMMARY

The OH profile and activation energy for diffusion have been measured in a fiber core using a color-center laser and calorimeter. The 600 to 800°C results presented here agree with the bulk silica results of Moulson and Roberts. Thus, the compositional gradients, internal stresses and geometry peculiar to the fiber apparently have no significant effect on the activation energy. This procedure is a novel approach to measuring diffusion and adds confidence that long-term OH diffusion will be no problem in currently proposed lightwave systems. Plans are being made to extend these measurements to even lower temperatures.

VI. ACKNOWLEDGMENTS

APPENDIX A
Derivation of the Diffusion Equation

In 1855, Adolph Fick empirically proposed that diffusion in one dimension was described by:

$$J_x = -D \frac{\partial C}{\partial x}, \tag{4}$$

where

J_x = the flux in the x direction per unit time
C = concentration of the diffusing species
D = diffusion coefficient.

Consider a volume element of material between x and $x + dx$ having

unit cross-sectional area as shown in Fig. 5. The flux into the element minus the flux out equals the rate of accumulation of a species within the element.

$$J_x - J_{x+dx} = \frac{\partial \bar{C}}{\partial t}\, dx, \tag{5}$$

where

$$\bar{C} = \text{Average concentration in the element}$$
$$\bar{C}dx = \text{Total amount in the element at time } t.$$

Expand J_{x+dx} about x in a Taylor series

$$J_{x+dx} = J_x + \frac{\partial J_x}{\partial x}\, dx + \frac{\partial^2 J_x}{\partial x^2}\frac{(dx)^2}{2} + \cdots. \tag{6}$$

In the limit as $dx \to 0$, eq. (5) has the form

$$-\frac{\partial J_x}{\partial x} = \frac{\partial C}{\partial t}. \tag{7}$$

This is simply a treatment of the conservation of matter. Combining eqs. (7) with (4) yields

$$\frac{\partial}{\partial x}\left(D\,\frac{\partial C}{\partial x}\right) = \frac{\partial C}{\partial t}. \tag{8}$$

If D does not vary with x, we have

$$D\,\frac{\partial^2 C}{\partial x^2} = \frac{\partial C}{\partial t}. \tag{9}$$

Also, if D does not vary with t, this is a linear partial differential equation known as Fick's Second Equation.

In the case of a fiber with cylindrical geometry, (9) becomes,

$$D\,\vec{\nabla}^2\, C = \frac{\partial C}{\partial t}. \tag{10}$$

This has the solution

$$C(r,\, t) = \frac{\alpha}{4\pi Dt}\,\exp[-(r^2/4Dt)] \qquad r > 0,\, t > 0, \tag{11}$$

where α is a constant relating to the initial concentration.
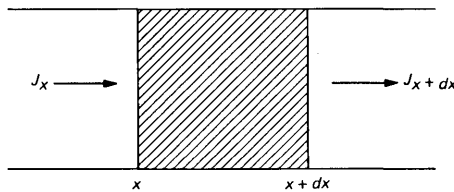
Fig. 5—Diffusion through a unit volume element.

APPENDIX B

Derivation of the Activation Energy

The distribution of power in a fiber core when the beam is focused at $r = 0$ is approximated by:

$$P(r) = P_0 \exp[-(r/r_0)^2].\tag{12}$$

The OH concentration for diffusion of a line source into a cylinder was derived previously [eq. (11)]:

$$C(r, t) = \frac{\alpha}{4\pi Dt} \exp[-(r^2/4Dt)],\tag{13}$$

when the laser beam is focused at $r = 0$, the parameter that is measured is the 0.95-μm absorption loss $\gamma(t)$:

$$\gamma(t) = K \int_0^a P(r)C(r, t)(2\pi r dr)\tag{14}$$

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

$$K = \text{a constant}.$$

Define a new constant:

$$r_1 = \left(\frac{1}{r_0^2} + \frac{1}{4Dt}\right)^{-1/2}.\tag{15}$$

Now substitute (12) and (13) into (14):

$$\gamma(t) = \frac{KMP_0}{2Dt} \int_0^a r \exp[-(r/r_1)]^2 dr$$

$$= \frac{KMP_0 r_0^2}{4Dt + r_0^2} \{1 - \exp[-(a/r_1)^2]\}.\tag{16}$$

If $a \gg r_1$, then $\exp[-(a/r_1)^2] \ll 1$; therefore, (16) becomes

$$\gamma(t) \cong \frac{KMP_0 r_0^2}{4Dt + r_0^2}.\tag{17}$$

The $a \gg r_1$ assumption is equivalent to the common practice of replacing

$$\int_0^a \quad \text{with} \quad \int_0^\infty.$$

If $r_0 < a/2$, then

$$\exp[-(a/r_1)^2] < 0.02$$

(regardless of $a^2/4Dt$), and the assumption is valid.

If γ_{in} is the loss when the fiber enters the oven and γ_{out} is the loss when the fiber is removed from the oven one day later, then eq. (17) gives

$$\gamma_{out}^1 - \gamma_{in}^{-1} = 4D \times (1 \text{ day})/KMP_0 r_0^2. \qquad (18)$$

The diffusion coefficient at the oven temperature T is $D = D_0 \exp(-E/RT)$, where E is the activation energy and R is the molar gas content. Multiplying (18) by γ_{in} and taking the ln then gives

$$\ln[(\gamma_{in}/\gamma_{out}) - 1] = \text{constant} - E/RT. \qquad (19)$$

Plotting

$$\ln[(\gamma_{in}/\gamma_{out}) - 1]$$

versus $1/T$ gives a straight line of slope $-E/R$.

REFERENCES

1. D. B. Keck, K. D. Mauer, and P. C. Schultz, "On the Ultimate Lower Limit of Attenuation in Glass Optical Waveguides," Appl. Phys. Lett., 22 (April 1973), pp. 307-9.
2. L. G. Cohen and S. J. Jang, "Interrelationship Between Water Absorption Loss and Dispersion in Multimode Fiber," Appl. Opt. 20, No. 9 (May 1981), pp. 1635-9.
3. M. Horiguchi, and M. Kawachi, "Measurement Technique of OH-Ion Distribution Profile in Rod Preform of Silica-Based Optical Fiber Waveguides," Appl. Optics, 17, No. 16 (August 15, 1978), pp. 2570-4.
4. P. Kaiser, "Drawing-Induced Coloration in Vitreous Silica Fibers," J. Opt. Soc. Am., 64, (1974), pp. 475-81.
5. B. J. Ainslie, P. W. France and G. R. Newns, "Water Impurity in Low Loss Silica Fibre," Mat. Res. Bull., 12, (May 1977), pp. 481-8.
6. M. J. Buckler and W. B. Gardner, "Measured Changes with Time in Optical Fiber OH Ion Absorption," Proc. Topical Meeting on Optical Fiber Communication, Washington, D.C., March 6-8, 1979, p. 106.
7. K. Nassau, "The Diffusion of Water in Fiber Optics," Mat. Res. Bull., 13, (January 1977), pp. 67-76.
8. R. Olshansky, "Optical Properties of Waveguide Materials for 1.2-1.8 Micron," Physics of Fiber Optics, Advances in Ceramics, 2, American Ceramic Society, Columbus, Ohio, 1981, pp. 40-54.
9. A. L. Ruoff, Materials Science, Englewood Cliffs, N.J.: Prentice Hall, 1973, pp. 445-9.
10. A. J. Moulson and J. P. Roberts, "Water in Silica Glass," Trans. Faraday Soc., 57, (July 1961), pp. 1208-16.
11. L. F. Mollenauer, "Color Center Laser," Methods in Experimental Physics, 15, Part B, New York: Academic Press, 1979, pp. 1-54.
12. R. B. Byrd, W. E. Stewart, and E. N. Lightfoot, Transport Phenomena, New York: John Wiley, 1966.
13. J. H. Hilderbrand, Viscosity and Diffusivity—A Predictive Treatment, New York: John Wiley, 1977.

A Class of Approximations for the Waiting Time Distribution in a *GI* / *G* /1 Queueing System

By A. A. FREDERICKS

Single server queueing models are important in the study of a wide variety of stochastic systems. A particularly important example is the study of task schedules for computer systems with real-time applications. In this paper, we present a class of approximations for the waiting time distribution in single server queueing systems with general independent (renewal) input and general (independent) service time distributions. These approximations allow the analyst to use as much (or as little) of the structure of the input and service processes as desired. Moreover, they also allow him to concentrate on specific quantities associated with the delay distribution that may represent the most relevant performance criteria. Examples include probability of delay, mean delay, and tails of the delay distribution. The methods given in this paper have been used to analyze the performance of task schedules in a variety of processor-based systems.

I. INTRODUCTION

Single server queueing models arise quite naturally in the study of a wide variety of stochastic server systems. A particularly important class is single processor computer systems, such as stored program control switching systems or nodes in a data communication network. In many such applications, it is important to keep as much of the structure of the interarrival and service time processes as possible in order to obtain realistic results; that is, the simplifying assumptions of exponential distributions cannot be made. While the resulting $GI/G/1$ queue may be extremely difficult to analyze, one is often content with reasonable approximations that incorporate the main features of the problem. In addition, one often desires results that are reasonably simple analytically, since the behavior of the $GI/G/1$

queueing model may be the input to the analysis of a more complex system.

Our main purpose here is to present a class of approximations for the waiting time distribution, $W(x)$, for such $GI/G/1$ systems which allows the analyst to use as much (or little) of the structure of the relevant input and service processes as desired. The resulting approximations can be extremely simple in form (e.g., a single exponential) or, with additional effort, more complex. In particular, we give relatively simple expressions for constants C and a, such that $W_A(x) = 1 - Ce^{-ax}$ provides a good fit to the probability of delay, $P_D = [1 - W(0)]$ and the mean delay, $\bar{w}$. Similar approximations are developed which provide a good fit to the tails of the delay distribution. Thus, these results can be used to study systems with a wide variety of performance criteria. Our results also lead to some new heavy traffic approximations, as well as a simple approximation for light traffic.

While we feel that the approximations presented here are of use in themselves, we hope that the results obtained will stimulate more active research into the general method presented. Of particular interest is the problem of obtaining more quantitative error bounds to guide the user in the application of these techniques.

We note that the methods given here have been used to analyze some rather complex $GI/G/1$ queueing systems which have arisen in the study of a certain class of computer systems with real-time applications.[1,2]

Some key results are summarized in the next section to give a general idea of the nature of this work. Our basic approach is introduced in Section III where several (single) exponential approximations are derived. In Section IV, we look briefly at the implications of the approximations for heavy traffic. In Section V, we look at some special cases which lead to analytic statements about the accuracy of our various approximations, and in Section VI, their accuracy is assessed via several numerical examples. Extensions to more general functional forms are considered in Section VII and illustrated via another numerical example. Some final remarks are given in Section VIII. A summary of our notation, key formulas, and approximations are given in the Appendix for easy reference.

II. SUMMARY OF SOME KEY RESULTS

The main idea of this work is to assume a functional form for an approximation $W_A(x)$ to the true waiting time distribution $W(x)$ for a $GI/G/1$ queue. The Lindley integral equation for $W(x)$ is then used to obtain the unknown coefficients. We obtain several approximations by assuming $W(x) \approx W_A(x) = 1 - Ce^{-ax}$. A key approximation of this

type, referred to as approximation A^* has C and a determined by eqs. (7) and (9) with $a = a_E$ (also see pages 299 and 300):

$$C = C^* = \frac{1 - K(0)}{1 - \hat{K}_-(a_E)}$$

and

$$\frac{C}{a} = \frac{C^*}{a^*} = \frac{\bar{u}_+}{K(0) - \hat{K}_-(a_E)},$$

where $K(u)$ is the distribution function for the difference, $\tilde{u}$,[†] between the service time and interarrival time and

$$\hat{K}_-(s) = \int_{-\infty}^{0} e^{su} dK(u)$$

$$\bar{u}_+ = \int_{0}^{\infty} u \, dK(u)$$

and a_E satisfies the characteristic equation

$$\int_{-\infty}^{\infty} e^{a_E u} dK(u) = 1.$$

(The equation numbers of this section correspond to those of later sections.) Limiting properties and special cases for this approximation, as well as others are discussed. Several numerical examples are also given to illustrate the accuracy of these approximations. For an example $H_2/E_3/1$ system, the maximum relative error in the approximation to P_D given by A* is found to be 5.0 percent, while the maximum relative error in the approximation to $\bar{w}$ given by A* is found to be 1.5 percent. Perhaps the simplest approximation considered (using only the first two moments of $\tilde{u}$) is the one referred to as approximation $A_{G,0}$, which is also of the form $W_A(x) = 1 - Ce^{-ax}$, but where now the constants C and a are determined by

$$a = a_H = \frac{-2\bar{u}}{\sigma_u^2}, \tag{14}$$

i.e., the heavy traffic a, and

$$C = C_{G,0} = [1 - \mathrm{Erf}(|\bar{u}|/\sqrt{2}\sigma_u)]/[1 + \mathrm{Erf}(|\bar{u}|/\sqrt{2}\sigma_u)], \tag{15}$$

where $\mathrm{Erf}(x)$ is the error function. Note that (15) has the limiting behavior for small $(\bar{u})$ (assuming σ_u tends to a finite limit)

[†] The symbol ~ will be used to denote a random variable (see the Appendix for a summary of notations used).

$$C_{G,0} \underset{\substack{\bar{u} \to 0 \\ (\rho \to 1)}}{\sim} 1 - 2\sqrt{\frac{2}{\pi}} \frac{|\bar{u}|}{\sigma_u}, \tag{16}$$

which then results in the following limit for the approximate mean delay, $\bar{w}_{G,0}$

$$\bar{w}_{G,0} = \frac{C_{G,0}}{a_H} \underset{\substack{\bar{u} \to 0 \\ (\rho \to 1)}}{\sim} \frac{1}{a_H} - \sigma_u \sqrt{\frac{2}{\pi}} \tag{17}$$

the right-hand side of eq. (17) is, in fact, a well-known lower bound for $\bar{w}$ (as $\bar{u} \to 0$, e.g., see Ref. 4).

The accuracy of these approximations, as well as several others we develop, are studied in considerably more detail in what follows. We also show how the approach taken leads to a class of approximations which can be made increasingly more accurate (with additional effort).

III. DEVELOPMENT OF THE EXPONENTIAL APPROXIMATION

We wish to obtain an approximate solution for the waiting time distribution, $W(x)$, for a $GI/G/1$ queueing system. More specifically, we have a single server queue where the interarrival times, $\tilde{t}$, between customers are independent and identically distributed, taken from a general distribution, $A(t)$ and the service times, $\tilde{\tau}$, are independently drawn from an arbitrary distribution $B(\tau)$. A first in-first out discipline is assumed. If we let $\tilde{u} = \tilde{\tau} - \tilde{t}$ and denote the distribution of $\tilde{u}$ by $K(u)$, then if an (equilibrium) waiting time distribution exists,[†] it satisfies the well-known Lindley integral equation[3]

$$W(x) = \int_{-\infty}^{x} W(x - y)dK(y), \qquad x \geq 0 \tag{1}$$

($W(x)$ is equal to 0 for $x < 0$).

In this section, we consider approximations $W_A(x)$ to $W(x)$ of the form

$$W_A(x) = 1 - Ce^{-ax}. \tag{2}$$

Our objective is to determine suitable values for the constants C and a. For this purpose, we will use the relation (1).

3.1 Pointwise matching

As is well known (see Ref. 5, pages 376 and 410), if the equation

$$\int_{-\infty}^{\infty} e^{au}dK(u) = 1, \tag{3}$$

[†] Throughout we assume stability, i.e., that $(1/\alpha) = \bar{\tau} < \bar{t} = (1/\lambda); \rho = (\lambda/\alpha) < 1$.

possesses a real nonzero root, a_E, then

$$1 - W(x) \underset{x \to \infty}{\sim} C_\infty e^{-a_E x}. \qquad (4)$$

Throughout this paper, we will tacitly assume that for all systems considered, eq. (3) does, in fact, possess such a root. Thus, if an exponential form is to be used, a_E is a reasonable choice for the parameter a in (2). On the other hand, substitution of (2) into (1) yields the relation

$$1 - Ce^{-ax} = K(x) - Ce^{-ax} \int_{-\infty}^{x} e^{ay} dK(y). \qquad (5)$$

Hence, we see that if we choose $a = a_E$, (5) will be valid to order $0(e^{-a_E x})^\dagger$ as x tends to infinity. We can now determine C by requiring that (5) be valid at another value of x. For x such that $[1 - K(x)] > 0$ we can solve (5) for C to obtain

$$C(x) = \frac{[1 - K(x)]e^{ax}}{1 - \int_{-\infty}^{x} e^{ay} dK(y)}. \qquad (6)$$

Now if we are interested in approximating the behavior of $W(x)$ near $x = 0$, a reasonable choice might be to solve for C from (6) with $x = 0$; that is, choose

$$C = C_0(a) = \frac{1 - K(0)}{1 - \hat{K}_-(a)} = \frac{1 - \hat{K}_-(0)}{1 - \hat{K}_-(a)}, \qquad (7)$$

where

$$\hat{K}_-(s) = \int_{-\infty}^{0} e^{su} dK(u).$$

[Note that clearly $0 \le C_0(a) \le 1$.]

We will refer to the exponential approximation of $W(x)$ obtained by choosing a to satisfy (3) ["(5) at $x = \infty$"] and C to satisfy (7) [(5) at $x = 0$] as approximation $A_{E,0}$, i.e., $W_{A_{E,0}} = 1 - C_0(a_E)e^{-a_E x}$. (The subscript E refers to the use of the exact dominant root—we will consider other alternatives shortly.)

Note that if the true delay distribution were given by a simple exponential, i.e., $W(x) = 1 - Ce^{-ax}$, then the value of $C(x)$ determined from (6) would be identically a constant and $W_{A_{E,0}}(x) \equiv W(x)$. Thus, the *variation* of $C(x)$ with x, as determined by (6), gives some indica-

---

$\dagger$ That is, the difference between the left-hand side and right-hand side will tend to zero even when multiplied by $Ce^{a_E x}$.

tion of *how* exponential the waiting time delay distribution is. A more "sophisticated" approximation might be obtained by minimizing the variation of $C(x)$ with respect to some suitable norm, however, we will concentrate here mostly on simpler techniques.

Note also that $C(x)$ as given by (6) appears in various (exponential) bounds for $W(x)$. For example, Kingman[6] and Rossberg and Siegel[7] prove the following inequality for the complementary waiting time distribution, $W^c(x) = 1 - W(x)$.

$$C_L e^{-a_E x} \le W^c(x) \le e^{-a_E x},$$

where a_E is given by (3) and

$$C_L = \inf_{\substack{x \ge 0 \\ [1-K(x)] > 0}} C(x).$$

Ross[8] gives the bound

$$W^c(x) \le C_U e^{-a_E x},$$

where

$$C_U = \sup_{\substack{x \ge 0 \\ [1-K(x)] > 0}} C(x).$$

(See also Ref. 9.) Hence, it seems quite natural to choose a specific value for x in $C(x)$ to obtain an exponential approximation to $W(x)$, with $C_L < C < C_U$.

3.2 An alternate method for determining C

As we shall see shortly, $A_{E,0}$ does provide a reasonable approximation to P_D, but the resulting approximation for $\bar{w}$ is not as good. One might consider choosing C from (6) with another choice of x; however, here we consider another alternative. If we compute the mean waiting time, $\bar{w}$, using eq. (1) we obtain

$$\bar{w} = \int_0^\infty x \, dW(x) = \int_0^\infty x \left[W(0) dK(x) + \int_{-\infty}^x d_x W(x - y) dK(y) \right]. \quad (8)$$

Again, assuming the exponential form (2) for $W_A(x)$ [and, hence, that $\bar{w}_A = (C/a)$], (8) yields

$$\bar{w}_A = \frac{C}{a} = \frac{\bar{u}_+}{K(0) - \hat{K}_-(a)} = \frac{\hat{K}'_+(0)}{\hat{K}_-(0) - \hat{K}_-(a)}, \quad (9)$$

where $\bar{u}_+ = \int_0^\infty u \, dK(u)$, $\hat{K}_+(s) = \int_0^\infty e^{su} dK(u)$ and $\hat{K}_-(s)$ is as in (7). If a is chosen from (3) and C then from (9), we will refer to the resulting approximation as $A_{E,1}$, i.e., $W_{A_{E,1}}(x) = 1 - C_1(a_E) e^{-a_E x}$. As we shall see, this will result in a good approximation for the mean delay via (9);

however, it is not clear from this relation that the resulting C need be less than 1. (It is clearly ≥ 0.) A correct probabilistic exponential form for $W_A(x)$ which incorporates both (7) (and, hence, results in a reasonable approximation for the probability of delay) and (9) (and hence results in a good approximation for the mean delay) can be obtained in the following manner. First, find a_E from (3), then $C = C^* = C_0(a_E)$ from (7), and then the approximate mean delay, $\bar{w} = C_1(a_E)/a_E$ from (9). The constant a is then chosen to be

$$a = a^* = \frac{C_0(a_E)}{C_1(a_E)} \, a_E.$$

We will refer to this approximation as A^*. We now look at a more organized way of obtaining such approximations.

3.3 Moment matching

Equation (9) was obtained by matching the first moment of the Lindley integral equation, i.e., from (8). We note that if we match the 0th moment in a similar way we are, in fact, also led to (7). That is, if we seek a solution to (1) of the form $1 - Ce^{-ax}$, and match the first two moments (0th and 1st) of (1) to determine *both* unknown constants, we are led to eqs. (7) and (9), i.e.,

$$C = \frac{1 - K(0)}{1 - \hat{K}_-(a)}$$

$$\frac{C}{a} = \frac{\bar{u}_+}{K(0) - \hat{K}_-(a)}$$

as a pair of equations in the two unknowns C and a. These can be combined and rewritten as

$$a = \frac{[K(0) - \hat{K}_-(a)][1 - K(0)]}{\bar{u}_+[1 - \hat{K}_-(a)]} \tag{10}$$

$$C = \frac{1 - K(0)}{1 - \hat{K}_-(a)}. \tag{11}$$

Thus, one can first determine a from (10) and then C from (11). Clearly, if such an a exists, $a \in [0, \infty]$, $C \in [0, 1]$.

Denoting the right-hand side of (10) by $f(a)$, we note that

$$f(a) \underset{a \to 0}{\to} 0$$

$$f(a) \underset{a \to \infty}{\to} \frac{K(0)[1 - K(0)]}{\bar{u}_+} < \infty$$

and also

$$f'(a) = \frac{-\hat{K}'_-(a)[1 - K(0)]^2}{\bar{u}_+[1 - \hat{K}_-(a)]^2}$$

has the properties

$$f'(a) > 0, \qquad a \geq 0$$

$$f'(0) = \frac{\bar{u}_+ - \bar{u}}{\bar{u}_+} > 1$$

$$f'(a) \underset{a \to \infty}{\to} 0.$$

Moreover,

$$f''(a) = \frac{-[1 - K(0)]^2}{\bar{u}_+} \frac{(\hat{K}_-(a)[1 - \hat{K}_-(a)] + 2(\hat{K}_-(a))^2}{[1 - \hat{K}_-(a)]^3} < 0, \qquad a \geq 0.$$

Thus, we see that (10) possesses a unique positive root, which can be obtained via the iteration scheme

$$a_{n+1} = f(a_n)$$

starting from any positive a_0.

Thus, we can always determine the desired root $a_{0,1}$ from (10) and then the resulting $C = C_{0,1}$ from (11). This approximation

$$W_{A_{0,1}}(x) = 1 - C_{0,1}e^{-a_{0,1}x}$$

is referred to as $A_{0,1}$.

3.4 Light traffic

Under quite general conditions, the function $\hat{K}_-(a)$ defined in (7) and with a satisfying (3) has the limit

$$\hat{K}_-(a) \underset{\rho \to 0}{\to} 0,$$

e.g., fix $B(\tau)$ and the "shape" of $A(t)$ and let $1/\lambda \to \infty$. Hence, for ρ small, (7) and (9) lead to the following intuitively appealing "light" traffic approximations (A_{LT})

$$C_{LT} = 1 - K(0) \tag{12}$$

$$\bar{w}_{LT} = \frac{C_{LT}}{a_{LT}} = \frac{\bar{u}_+}{K(0)}. \tag{13}$$

IV. HEAVY TRAFFIC

In the case of heavy traffic, an obvious possibility for the exponent a is the heavy traffic limit (e.g., see Ref. 4):

$$a_H = \frac{-2\bar{u}}{\sigma_u^2}, \tag{14}$$

where $\bar{u}$, σ_u^2 are the mean and variance, respectively, of $\tilde{u} = \tilde{\tau} - \tilde{t}$ (service minus interarrival time). We can use this value for a in place of a_E in the approximations of Section III. In particular, if we use $a = a_H$ in (7) we will refer to the resulting approximation as $A_{H,0}$. Note that the resulting C is strictly less than one unless $a_H = 0$ ($\rho = 1$). If we use $a = a_H$ in (9), we refer to the resulting approximation as $A_{H,1}$. These two approximations (as well as those we discuss next) offer potential improvements to the more standard heavy traffic approximation that can extend its applicability to lower load levels.

In many cases, where heavy traffic approximations are applied, the structure of $K(u)$ is either not known or too complex to be used in further analysis. Standard heavy traffic approximations generally only make use of the first two moments of $K(u)$, i.e., $\bar{u}$ and σ_u^2. We can also obtain results which use only $\bar{u}$, σ_u^2 by making the standard heavy traffic assumption that $K(u)$ is approximately Gaussian. For example, if we not only assume that $a = a_H$ in (7), but further that $K(u)$ is Gaussian, we obtain the following approximation for C using (7):

$$C_{G,0} = \frac{1 - \mathrm{Erf}\!\left(\dfrac{-\bar{u}}{\sqrt{2}\sigma_u}\right)}{1 + \mathrm{Erf}\!\left(\dfrac{-\bar{u}}{\sqrt{2}\sigma_u}\right)}. \tag{15}$$

We refer to this approximation as $A_{G,0}$. Note that (15) implies the limiting behavior for $|\bar{u}|$ small

$$C_{G,0} \underset{\substack{\bar{u}\to 0\\(\rho\to 1)}}{\sim} \frac{1 - \sqrt{\dfrac{2}{\pi}}\dfrac{|\bar{u}|}{\sigma_u}}{1 + \sqrt{\dfrac{2}{\pi}}\dfrac{|\bar{u}|}{\sigma_u}} \underset{\substack{\bar{u}\to 0\\(\rho\to 1)}}{\sim} 1 - 2\sqrt{\dfrac{2}{\pi}}\dfrac{|\bar{u}|}{\sigma_u}. \tag{16}$$

Now combining (16) with (14), we obtain the following (limiting) approximation to the mean delay for $|\bar{u}|$ small

$$\bar{w}_{G,0} = \frac{C_{G,0}}{a_H} \underset{\substack{\bar{u}\to 0\\(\rho\to 1)}}{\sim} \frac{1}{a_H} - \sigma_u\sqrt{\frac{2}{\pi}}. \tag{17}$$

In a similar manner, (9) yields the approximation $A_{G,1}$

$$\bar{w}_{G,1} = \frac{C_{G,1}}{a_H} = \frac{\dfrac{\sigma_u}{\sqrt{2\pi}}\,e^{-\frac{1}{2}\left(\frac{\bar{u}}{\sigma_u}\right)^2} + \dfrac{\bar{u}}{2}\left[1 - \mathrm{Erf}\!\left(\dfrac{-\bar{u}}{\sqrt{2}\sigma_u}\right)\right]}{\mathrm{Erf}\!\left(\dfrac{-\bar{u}}{\sqrt{2}\sigma_u}\right)}, \tag{18}$$

which has the limit

$$\bar{w}_{G,1} = \frac{C_{G,1}}{a_H} \underset{\substack{\bar{u}\to 0 \\ (\rho\to 1)}}{\sim} \frac{\sigma_u^2}{2|\bar{u}|} - \frac{\sigma_u}{2}\sqrt{\frac{\pi}{2}} \underset{\substack{\bar{u}\to 0 \\ (\rho\to 1)}}{\sim} \frac{1}{a_H} - \frac{\sigma_u}{2}\sqrt{\frac{\pi}{2}}. \tag{19}$$

Now if $K(u)$ were, indeed, Gaussian, then the resulting (heavy traffic) mean delay, $\bar{w}_G$, is known to have the limiting bounds (e.g., see Ref. 4).

$$\frac{1}{a_H} - \sigma_u\sqrt{\frac{2}{\pi}} \le \bar{w}_G \le \frac{1}{a_H}. \tag{20}$$

Compare (17), (19), and (20).

V. SOME SPECIAL CASES

In this section, we look at the general behavior of these approximations for some special cases where exact results are readily obtainable. We begin by looking at the $GI/M/1$ case where the approximations of Section III are exact. This will serve as a simple illustrative example, as well as provide some additional insight into the behavior of the heavy traffic approximations of Section IV.

5.1 Case: GI/M/1

For a $GI/M/1$ system, we can write (for $u > 0$)

$$K(u) = \int_0^\infty [1 - e^{-\alpha(u+t)}]dA(t) = 1 - pe^{-\alpha u}, \qquad u > 0, \tag{21}$$

where $1/\alpha$ is the mean of the (exponential) service time and

$$p = \int_0^\infty e^{-\alpha t}dA(t) = \hat{A}(\alpha).$$

We, thus, have

$$\bar{u}_+ = \alpha p \int_0^\infty xe^{-\alpha x}dx = p/\alpha \tag{22}$$

$$K(0) = 1 - p. \tag{23}$$

Now if a satisfies (3),

$$\hat{K}_-(a) = \int_{-\infty}^0 e^{au}dK(u) = 1 - \int_0^\infty e^{au}dK(u).$$

Hence,

$$\hat{K}_-(a) = 1 - \alpha p \int_0^\infty e^{au}e^{-\alpha u}du = 1 + \frac{\alpha p}{a - \alpha}. \tag{24}$$

Using (23), (24) in (7) yields

$$C_0(a) = \frac{\alpha - a}{\alpha}, \tag{25}$$

while (22), (23), (24) in (9) yield

$$\frac{C_1(a)}{a} = \frac{\alpha - a}{a\alpha}. \tag{26}$$

If a is chosen to be the exact root of (3), then these both lead to the exact waiting time distribution. (Recall, for a $GI/M/1$ system

$$W(x) = 1 - \frac{\alpha - a_E}{\alpha} e^{-a_E x},$$

e.g., see Ref. 10.)

If we use the heavy traffic approximation for a, i.e., (14) then (25) or (26) can be used to obtain the following approximation for C

$$C_{H,0} = \frac{\alpha + \dfrac{2\bar{u}}{\sigma_u^2}}{\alpha} = 1 - \frac{2}{\alpha^2 \sigma_u^2} \frac{(1 - \rho)}{\rho}; \tag{27}$$

that is, a possible improvement to $C = 1$ for heavy traffic. Thus, the resulting mean delay is found to be

$$\bar{w}_{H,0} = \frac{C_{H,0}}{a_H} = \frac{1}{a_H} - \frac{1}{\alpha} = \frac{\sigma_\mu^2}{2|\bar{u}|} - \frac{1}{\alpha}. \tag{28}$$

Note that this correction term to $1/a_H$ as an approximation to the mean delay is typical. For example, for an $M/G/1$ system it is easy to show that the true mean delay, $\bar{w}$, satisfies

$$\bar{w} = \frac{1}{a_H} - \frac{1}{2}\left(\frac{1}{\lambda} + \frac{1}{\alpha}\right).$$

5.2 Case: M/G/1

For this case, we have

$$K(u) = \int_0^\infty e^{\lambda(u-\tau)} dB(\tau) = q e^{\lambda u}, \qquad u < 0, \tag{29}$$

where $1/\lambda$ is the mean of the (exponential) interarrival time and

$$q = \int_0^\infty e^{-\lambda\tau} dB(\tau) = \hat{B}(\lambda).$$

We, thus, have

$$\bar{u}_+ = \int_0^\infty u dK(u) = \int_{-\infty}^\infty u dK(u) - \int_{-\infty}^0 u dK(u)$$

$$= \bar{u} - q\lambda \int_{-\infty}^0 u e^{\lambda u} du = \left(\frac{1}{\alpha} - \frac{1}{\lambda} + \frac{q}{\lambda} \right) \tag{30}$$

$$K(0) = q \tag{31}$$

and

$$\hat{K}_-(a) = q\lambda \int_{-\infty}^0 e^{au} e^{\lambda u} du = \frac{q\lambda}{a + \lambda}. \tag{32}$$

Using (7), we obtain the following approximation $(A_{E,0})$ for the probability of delay, P_D

$$C_0(a_E) = \frac{1 - q}{1 - \dfrac{q\lambda}{a_E + \lambda}}, \tag{33}$$

while from (9) we obtain the following approximation $(A_{E,1})$ to the mean delay, $\bar{w}$

$$\bar{w}_{E,1} = \frac{C_1(a_E)}{a_E} = \frac{(\lambda - \alpha + q\alpha)(a_E + \lambda)}{q\alpha\lambda a_E}. \tag{34}$$

(Since we know that for an $M/G/1$ queue, the true value of P_D is ρ, we could use this fact, together with (33) to obtain a simple approximation for a_E; we discuss this possibility shortly.)

Using the known limits for the true $a = a_E$ as $\rho \to 1$, (33) readily yields the following:

$$C_0(a_E) \underset{\substack{\lambda \to 0 \\ (\rho \to 0)}}{\sim} \rho(1 - \rho \overline{\tau^2} \alpha^2) \underset{\substack{\lambda \to 0 \\ (\rho \to 0)}}{\sim} \rho \tag{35}$$

$$C_0(a_E) \underset{\substack{\lambda \to \alpha \\ (\rho \to 1)}}{\sim} 1 - \frac{q}{1 - q} \frac{a}{\alpha} \underset{\substack{\lambda \to \alpha \\ (\rho \to 1)}}{\to} 1, \tag{36}$$

where we have used the fact that for $\lambda \to 0$

$$q = \int_0^\infty e^{-\lambda\tau} dB(\tau) \underset{\lambda \to 0}{\sim} 1 - \frac{\lambda}{\alpha} + \frac{\lambda^2 \overline{\tau^2}}{2}. \tag{37}$$

We similarly obtain from (34)

$$\bar{w}_{E,1} = \frac{C_1(a_E)}{a_E} \underset{\substack{\lambda \to 0 \\ (\rho \to 0)}}{\sim} \frac{\lambda \overline{\tau^2}}{2(1 - \rho)} \left(1 + \frac{\lambda}{a_E} \right) \underset{\substack{\lambda \to 0 \\ (\rho \to 0)}}{\sim} \frac{\lambda \overline{\tau^2}}{2(1 - \rho)} \tag{38}$$

and for $\lambda \to \alpha$, clearly

$$\bar{w}_{E,1} = \frac{C_1(a_E)}{a_E} \underset{\substack{\lambda \to \alpha \\ (\rho \to 1)}}{\sim} \frac{1}{a_E}. \tag{39}$$

Equation (38) shows that as $\lambda \to 0$, we obtain the familiar $P - K$ (Pollecek-Khinchin) formula for $\bar{w}$ in an $M/G/1$ system, while (39) shows that with

$$a = a_E, \qquad \bar{w}_{E,1} \underset{\rho \to 1}{\sim} \frac{1}{a_E} \underset{\rho \to 1}{\sim} \frac{1}{a_H}.$$

Now recall for the approximation A^*, we take $C^* = C_0(a_E)$ and

$$a^* = \frac{C_0(a_E)a_E}{C_1(a_E)}.$$

Hence, for A^*, the resulting exponent for the exponential approximation is given by

$$a^* = \frac{q\alpha\lambda a_E(1 - q)}{(\lambda - \alpha + q\alpha)(a_E + \lambda - q\lambda)}. \tag{40}$$

From (40) we obtain

$$a^* \underset{\substack{\lambda \to 0 \\ (\rho \to 0)}}{\to} \frac{2}{\alpha\overline{\tau^2}} = \frac{1}{\bar{R}} \tag{41}$$

$$a^* \underset{\substack{\lambda \to \alpha \\ (\rho \to 1)}}{\sim} a_E\left\{1 - \left[\frac{a_E(1 - q)}{\alpha}\right]\right\} \underset{\substack{\lambda \to \alpha \\ (\rho \to 1)}}{\sim} a_E, \tag{42}$$

where $\bar{R}$ is mean forward recurrence time of the service time distribution. Note we have used the fact that

$$a_E \underset{\rho \to 1}{\sim} 0.$$

The first of these is a very interesting relation. It is easy to show that for an $M/G/1$ system, that as $\rho \to 0$ the true dominant root a_E satisfies

$$\hat{B}(-a_E) = \int_0^\infty e^{a_E\tau}dB(\tau) \underset{\rho \to 0}{\sim} \infty. \tag{43}$$

Hence, for example, if $\hat{B}(s)$ is rational, a_E tends to the smallest (in magnitude) pole of $\hat{B}(s)$, which is not consistent with (41). However, it is easy to show that for an $M/G/1$ system, the *mean* delay conditioned on being delayed, is just the mean forward recurrence time, $\bar{R}$— the dominant contribution is from the case where the arrival finds one

customer in the system. That is, for $\rho \to 0$, the dominant root really may not be the best quantity to use for an approximation, indeed, the limit (41) is most likely preferable. We, thus, have via A^* an exponential approximation to the true delay distribution for $\rho \to 0$ which can be decidedly *nonexponential* but with the *correct* limiting probability of delay *and* mean delay.

The approximation $A_{0,1}$ also has some very interesting properties for the $M/G/1$ case. Considering (7) and (9) as defining two equations for the two unknowns C and a, we obtain the two relations

$$(a + \lambda) = \frac{Cq\lambda}{C + q - 1} \tag{44}$$

and

$$(a + \lambda) = \frac{Cq\alpha\lambda}{\lambda - \alpha + q\alpha}. \tag{45}$$

Combining (44) and (45) we find that

$$C = \frac{\lambda}{\alpha} = \rho. \tag{46}$$

That is, for the $M/G/1$ case, $A_{0,1}$ yields an *exact* expression for the probability of delay, P_D. Using (44) or (45), we obtain the following approximation for the dominant root:

$$a = \frac{\alpha\rho^2 q}{(\rho - (1 - q))} - \alpha\rho = \frac{\rho\alpha(1 - q)(1 - \rho)}{\rho - (1 - q)}. \tag{47}$$

Note that

$$a \underset{\rho \to 1}{\sim} \frac{\alpha(1 - q)(1 - \rho)}{q}, \tag{48}$$

but

$$a \underset{\rho \to 0}{\to} \frac{2}{\alpha \overline{\tau^2}}$$

as with A^*.

Note that since for an $M/G/1$ system, we know that $P_D = \rho$, we could have used this fact directly in either of the two approximations [(33), (34)] to obtain an approximation for the true dominant root, a_E. This, of course, would result in

$$a_E \cong \frac{\rho\alpha(1 - q)(1 - \rho)}{\rho - (1 - q)},$$

i.e., just (47).

For the mean delay, $\bar{w}$, $A_{0,1}$ yields the approximation

$$\bar{w}_{0,1} = \frac{\rho - (1 - q)}{\alpha(1 - q)(1 - \rho)}, \tag{49}$$

which has the asymptotic properties

$$\bar{w}_{0,1} \underset{\substack{\lambda \to 0 \\ (\rho \to 0)}}{\sim} \frac{\lambda \overline{\tau^2}}{2\left(1 - \rho - \dfrac{\lambda \alpha \overline{\tau^2}}{2}\right)} \tag{50}$$

$$\bar{w}_{0,1} \underset{\substack{\lambda \to \alpha \\ (\rho \to 1)}}{\sim} \frac{q}{\alpha(1 - q)(1 - \rho)}. \tag{51}$$

Comparing the asymptotic behavior of $A_{0,1}$ and A^*, we see that while the former has a more desirable property for the approximation to P_D, the behavior of the approximations to $\bar{w}$ and a_E are more desirable for A^* than for $A_{0,1}$. This leads us to consider another approximation, one that combines the best properties of each. The simplest method is to define approximation A^{**} via $w_{A^{**}}(x) = 1 - C^{**}e^{-a^{**}x}$, where

$$C^{**} = C_{0,1}$$

and

$$a^{**} = \frac{C_{0,1}}{\bar{w}_{E,1}} = \frac{C_{0,1}\,a_E}{C_{E,1}}.$$

The use of the heavy traffic a_H, of course, yields similar results to those for $A_{E,0}$, $A_{E,1}$, and A^* as $\rho \to 1$. Again, there is no simplification for the Gaussian case.

VI. SOME NUMERICAL EXAMPLES

We consider several numerical examples of varying complexity to illustrate the accuracy of the various approximations. Because of the large number of possible combinations of approximations and quantities of interest, we will not cover all possibilities for every example.

6.1 Case: M/D/1

For the $M/D/1$ case,

$$K(u) = P(\tilde{u} = \bar{\tau} - \tilde{t} \leq u) = \begin{cases} 1, & u > \bar{\tau} \\ e^{-\lambda(\bar{\tau} - u)}, & u \leq \bar{\tau}. \end{cases}$$

Hence, we can readily obtain eq. (3) for the dominant root, a_E

$$\int_{-\infty}^{\infty} e^{au}\mathrm{d}K(u) = \frac{\lambda}{a + \lambda}\,e^{a\bar{\tau}} = 1. \tag{52}$$

Rewriting this equation as

$$\lambda\bar{\tau}(e^{a\bar{\tau}} - 1) = a\bar{\tau}, \tag{53}$$

it is clear that in addition to $a = 0$, there is always one positive real root, a_E, for any finite λ. Moreover,

$$a_E \underset{\lambda\bar{\tau}=\rho\to1}{\to} 0, \qquad a_E \underset{\substack{\lambda\to0 \\ (\rho\to0)}}{\to} \infty.$$

The other quantities needed for the various approximations are

$$q = \hat{B}(\lambda) = e^{-\rho} \tag{54}$$

$$K(0) = q = e^{-\rho} \tag{55}$$

$$\hat{K}_{-}(a_E) = \frac{q\lambda}{a_E + \lambda} = \frac{e^{-\rho}\lambda}{a_E + \lambda} \tag{56}$$

$$\bar{u}_{+} = \left(\frac{1}{\alpha} - \frac{1}{\lambda} + \frac{q}{\lambda}\right) = \left(\bar{\tau} - \frac{1}{\lambda} + \frac{e^{-\rho}}{\lambda}\right) \tag{57}$$

$$\bar{u} = \left(\frac{1}{\alpha} - \frac{1}{\lambda}\right) = \bar{\tau} - \frac{1}{\lambda} \tag{58}$$

$$\sigma_{u}^{2} = \frac{1}{\lambda^{2}} \tag{59}$$

and, hence,

$$a_H = \frac{-2\bar{u}}{\alpha_{u}^{2}} = \frac{2\rho(1 - \rho)}{\bar{\tau}}. \tag{60}$$

With a_E from (53) and (54) to (60), we can compute the various quantities predicted by our approximations. For simplicity, we will assume $\bar{\tau} = 1$ in the following so that $\lambda = \rho$.

6.1.1 Probability of delay, P_D

Table I shows the value of P_D as predicted by the various approximations. Recall that P_D from $A_{0,1}$ is exact for this case. We see, however, that $A_{E,0}(A^*)$ produces a good approximation to this quantity. The value of P_D predicted by $A_{E,1}$ is seen to be significantly poorer, particularly for ρ small. Actually, our heavy traffic approximation $A_{H,0}$ does significantly better than $A_{E,1}$. Note that $A_{G,0}$ provides a reasonable

Table I—Probability of Delay, P_D, case $M/D/1$

ρ	True	$A_{0,1}(A^{**})$	$A_{E,0}(A^{*})$	$A_{H,0}$	$A_{G,0}$	$A_{E,1}$	A_{LT}
0.01	0.0100	0.0100	0.0100	0.0149	0.1920	0.0327	0.0100
0.1	0.1000	0.1000	0.0975	0.1406	0.2256	0.1986	0.0952
0.5	0.5000	0.5000	0.4756	0.5647	0.4462	0.6170	0.3935
0.9	0.9000	0.9000	0.8864	0.8975	0.8524	0.9276	0.5934
0.99	0.9900	0.9900	0.9884	0.9885	0.9842	0.9928	0.6284

Table II—Mean delay, $\bar{w}$, case $M/D/1$

ρ	True	$A_{E,1}(A^*,$ $A^{**})$	$A_{0,1}$	$A_{E,0}$	$A_{H,1}$	$A_{G,1}$	A_{HT}	A_{LT}
0.01	0.00505	0.00505	0.00507	0.00154	0.00759	12.53	50.51	0.00503
0.1	0.0556	0.0550	0.0565	0.0270	0.0832	1.589	5.556	0.0536
0.5	0.5000	0.4911	0.5415	0.3785	0.7026	1.033	2.000	0.3513
0.9	4.500	4.478	5.166	4.279	5.027	4.895	5.556	0.8379
0.99	49.50	49.56	57.55	49.35	50.13	49.88	50.51	0.9829

approximation in the heavy traffic region and that A_{LT} provides a good approximation in the light traffic region.

6.1.2 Mean delays, $\bar{w}$

Table II shows the mean delays resulting from the various approximations. We see that $A_{E,1}(A^*)$ provides the best approximation to the true $\bar{w}$. The approximations $A_{H,1}$ and $A_{0,1}$ are somewhat similar, with $A_{H,1}$ being better at higher loads, while $A_{0,1}$ appears better in the midrange. The approximation $A_{E,0}$ appears quite poor for low loads but improves with increasing load. While not comparable with $A_{H,1}$, $A_{G,1}$ does provide some improvement over A_{HT}. In the light load region, A_{LT} is seen to provide a good approximation. It is interesting to note that while $A_{H,1}$, $A_{G,1}$, and A_{HT} all use the heavy traffic a_H from (60), which has the poor behavior

$$a_H \underset{\rho \to 0}{\to} 0,$$

this only results in the anomalous behavior

$$\bar{w}_A \underset{\rho \to 0}{\to} \infty$$

for $A_{G,1}$ and A_{HT}.

6.1.3 Dominant root, a_E

Table III shows the exponent that would be used in the (exponential) approximation for $W(x)$ corresponding to each of the approximations. Recall that for $A_{E,0}$ and $A_{E,1}$ the exponent is exact, i.e., $a = a_E$. We see that A^* is somewhat better than $A_{0,1}$ for heavier loads and both tend to 2 as $\rho \to 0$—the inverse of the mean forward recurrence time of the service time. As should be expected, the heavy traffic a, a_H, is quite poor for light traffic.

Table III—Dominant root, a, $M/D/1$

ρ	True	A^*	$A_{0,1}$	A_H
0.01	6.475	1.974	1.974	0.0198
0.1	3.615	1.775	1.770	0.1800
0.5	1.256	0.9685	0.9234	0.5000
0.9	0.2071	0.1979	0.1742	0.1800
0.99	0.0200	0.0199	0.0172	0.0198

Note that the differences between P_D and $\bar{w}$ as predicted by $A_{E,0}$ and $A_{E,1}$, particularly for light loads, indicates that the true delay distribution is essentially nonexponential. Yet, we see that assuming an exponential form can result in excellent approximations for P_D and $\bar{w}$. Moreover, the reasonable results predicted by $A_{H,0}$ and $A_{H,1}$ for light loads where a_H is far from any reasonable choice of a dominant root indicates that these techniques are somewhat robust with respect to the choice of the exponent, a.

6.2 Case: M/E₂/1

We now look at several of the resulting approximations for the waiting time distribution, $W_A(x)$, and compare these with the exact results for an $M/E_2/1$ system. Specifically, we assume $A(t) = 1 - e^{-\lambda t}$ and that $B(t)$ is the convolution of two exponentials with unit mean. Hence, we have for the respective Laplace Stieltjes transforms

$$\hat{A}(s) = \frac{\lambda}{\lambda + s}$$

$$\hat{B}(s) = \frac{1}{(1 + s)^2}. \tag{61}$$

For this system, the exact delay distribution, $W(x)$ can readily be obtained via standard technique (e.g., see Ref. 10). The result is

$$W(x) = 1 - C_1 e^{-a_1 x} - C_2 e^{-a_2 x}, \tag{62}$$

where

$$a_1 = \frac{(2 - \lambda) - [\lambda(4 + \lambda)]^{1/2}}{2}$$

$$a_2 = \frac{(2 - \lambda) + [\lambda(4 + \lambda)]^{1/2}}{2}$$

$$C_1 = \frac{(1 - \rho)(1 - a_1)^2}{a_1(a_2 - a_1)}$$

$$C_2 = \frac{-(1 - \rho)(1 - a_2)^2}{a_1(a_2 - a_1)}.$$

From (61) and (62), we can readily compute all of the quantities we need. We have

$$q = \hat{B}(\lambda) = \frac{1}{(1 + \lambda)^2}; \qquad \bar{\tau} = \frac{1}{\alpha} = 2; \qquad \bar{u} = 2 - \frac{1}{\lambda},$$

and

$$\sigma_u^2 = 2 + \frac{1}{\lambda^2},$$

yielding expressions analogous to (54) to (60).

Tables IV, V, and VI show comparisons of the exact delay distribution, as well as the mean delay, with some of the approximations we developed for $\rho = \lambda \bar{\tau} = 2\lambda = 0.1$, 0.5, and 0.9. The approximations for the probability of delay, P_D, and mean delay, $\bar{w}$, show similar behavior to that for the $M/D/1$ case. In addition, we see that the approximation $A_{E,1}$ seems to have the best tail behavior (followed closely by A^{**}).

Thus, from these two examples, we might conclude that A^{**} is the best overall approximation, while $A_{E,1}$ has slightly better tail behavior.

6.3 Case: D/D₂/1 (D/G/1)

We briefly consider here the $D/G/1$ case and a simple numerical example which illustrates some interesting properties of our approximations.

For the $D/G/1$ case, we have

$$K(u) = \begin{cases} B(u + d); & u \geq -d \\ 0; & u < -d' \end{cases}$$

Table IV—P(Delay ≤ x) for case M/E₂/1

x	True	$A_{E,0}$	$A_{E,1}$	$A_{0,1}$	A^*	A^{**}
0	0.9000	0.9014	0.8760	0.9000	0.9014	0.9000
0.5	0.9220	0.9323	0.9148	0.9257	0.9268	0.9261
1.0	0.9414	0.9534	0.9414	0.9448	0.9460	0.9454
5.0	0.9962	0.9977	0.9971	0.9950	0.9950	0.9951
10.0	0.9999	0.9999	0.9999	0.9997	0.9997	0.9998
$\bar{w}$	0.1667	0.1314	0.1654	0.1680	0.1654	0.1654

Note: $\rho = 0.1$, True $W(x) = 1 - 0.1667e^{-0.7500x} + 0.06667e^{-1.200x}$.

Table V—P(Delay ≤ x) for case M/E₂/1

x	True	$A_{E,0}$	$A_{E,1}$	$A_{0,1}$	A^*	A^{**}
0	0.5000	0.5119	0.4666	0.5000	0.5119	0.5000
0.5	0.5644	0.5922	0.5547	0.5742	0.5859	0.5776
1.0	0.6272	0.6593	0.6277	0.6374	0.6488	0.6431
5.0	0.9084	0.9192	0.9117	0.8998	0.9058	0.9073
10.0	0.9848	0.9866	0.9854	0.9799	0.9818	0.9828
$\bar{w}$	1.500	1.357	1.483	1.556	1.483	1.483

Note: $\rho = 0.5$, True $W(x) = 1 - 0.5532e^{-0.3596x} + 0.05317e^{-1.390x}$.

Table VI—P(Delay ≤ x) for case M/E₂/1

x	True	$A_{E,0}$	$A_{E,1}$	$A_{0,1}$	A^*	A^{**}
0	0.1000	0.1057	0.0919	0.1000	0.1057	0.1000
0.5	0.1244	0.1354	0.1220	0.1278	0.1349	0.1296
1.0	0.1509	0.1641	0.1511	0.1548	0.1632	0.1582
5.0	0.3497	0.3617	0.3518	0.3426	0.3584	0.3557
10.0	0.5359	0.5446	0.5374	0.5198	0.5397	0.5388
$\bar{w}$	13.50	13.26	13.46	14.33	13.46	13.46

Note: $\rho = 0.9$, True $W(x) = 1 - 0.911e^{-0.06745x} + 0.01110e^{-1.483x}$.

where d is the constant interarrival time. Thus,

$$K(0) = B(d)$$

$$\hat{K}_-(a) = \int_{-\infty}^0 e^{ay} dK(y) = 1 - \int_0^\infty e^{ay} dK(y) = 1 - \hat{K}_+(a)$$

$$\overline{u_+} = \int_0^\infty x dK(x) = \hat{K}'_+(0) \, ,$$

where a satisfies

$$\int_{-\infty}^\infty e^{ay} dK(y) = \hat{K}_-(0) + \hat{K}_0(0) = 1.$$

This class of systems $(D/G/1)$ is particularly important in the analysis of a certain class of schedules for computer systems with real time applications. For example, see Refs. 1 and 2 where the approximations given here are applied to the study of these schedules. The service time distribution for these systems are generally discrete in nature. Here we look at a simple special case with

$$A(t) = U(t - d)$$

$$B(\tau) = p_1 U(\tau - s_1) + p_2 U(\tau - s_2),$$

where $U(x)$ is the unit step function, $p_1 + p_2 = 1$ and $s_1 < d < s_2$. From these, we readily obtain

$$K(0) = p_1$$

$$\hat{K}_-(a) = p_1 e^{a(s_1 - d)}$$

$$\hat{K}_+(a) = p_2 e^{a(s_2 - d)}$$

$$\overline{u_+} = \hat{K}'_+(0) = p_2(s_2 - d) \, ,$$

and a_E satisfies

$$\hat{K}_-(a) + \hat{K}_+(a) = \bar{e}^{ad}(p_1 e^{as_1} + p_2 e^{as_2}) = 1 \, .$$

Thus, (7) and (9) readily yield

$$P_D = C_0(a_E) = \frac{1 - p_1}{1 - p_1 e^{a_E(s_1 - d)}} \tag{63}$$

$$\bar{w}_{E,1} = \frac{C_1(a_E)}{a_E} = \frac{p_2(s_2 - d)}{p_1 - p_1 e^{a_E(s_1 - d)}}. \tag{64}$$

For simplicity, we consider the special case $s_1 = 0$, $s_2 = 2d$. In this case, the exact solution satisfies the difference equation

$$P_i = p_1 P_{i+1} + p_2 P_{i-1}, \tag{65}$$

where $P_i = P$(work in the system = id at an arbitrary arrival epoch), i.e., $P_i = P$(arrival delayed id). Taking $d = 1$, the solution to (65) is readily found to be

$$P_i = P(\text{Delay} = i) = (1 - r)r^i; \qquad r = \frac{p_2}{p_1}. \tag{66}$$

Hence,

$$P_D = r = \frac{p_2}{p_1} \tag{67}$$

$$\bar{w} = \frac{r}{1 - r} = \frac{p_2}{p_1 - p_2} \tag{68}$$

and $a_E = \ln(r^{-1})$.

Now (63) and (64) yield

$$C_0(a_E) = \frac{1 - p_1}{1 - p_2} = \frac{p_2}{p_1} \tag{69}$$

$$\bar{w}_{E,1} = \frac{C_1(a_E)}{a_E} = \frac{p_2}{p_1 - p_2}, \tag{70}$$

i.e., $A_{E,0}$ gives the *exact* value for P_D and $A_{E,1}$ gives the *exact* value $\bar{w}$. In particular, this implies that the approximation $A_{E,0}$ is *exact* at grid points—$W_{A_{E,0}}(x)$ is exact for $x = i$, an integer.

Table VII shows the resulting delay distributions for several of our approximations. We see here that A^* seems to be the best overall approximation. Note that $A_{0,1}$ (and hence A^{**}) result in somewhat poorer approximations for this case. This may indicate that A^* is potentially more robust.

As another comparison for the mean delays predicted by our approximations, we have included on the table (denoted by $\bar{w}[K - L]$) the results of using an approximation for the mean delay given by

Table VII—P(Delay $\leq i$) for case $D/D_2/1$

i	True	$A_{E,0}$	$A_{E,1}$	A^*	$A_{0,1}$
0	0.1818	0.1818	0.0970	0.1818	0.1282
1	0.3306	0.3306	0.2612	0.3178	0.2331
2	0.4523	0.4523	0.3955	0.4312	0.3254
3	0.5519	0.5519	0.5054	0.5258	0.4066
4	0.6334	0.6334	0.5953	0.6046	0.4780
5	0.7000	0.7000	0.6689	0.6704	0.5408
6	0.7546	0.7546	0.7291	0.7252	0.5960
7	0.7992	0.7992	0.7784	0.7709	0.6446
8	0.8357	0.8357	0.8187	0.8089	0.6874
9	0.8655	0.8655	0.8516	0.8407	0.7250
10	0.8900	0.8900	0.8786	0.8672	0.7581
$\bar{w}$	4.50	4.08	4.50	4.50	6.80

Note: $\rho = 0.9$, True $P(\text{Delay} = i) = \left[1 - \left(\frac{9}{11}\right)\right]\left(\frac{9}{11}\right)^i$, $\bar{w}(K - L) = 6.43$.

Kramer and Largenbach-Belz.[11] This approximation is essentially a heuristic extension of an approximation originally obtained by Heyman[12] via diffusion techniques. The rather simple form of the approximation $\bar{w}(K - L)$ given below makes its use very appealing.

$$\bar{w}(K - L) = \frac{\bar{\tau}\rho e^{-\gamma}(C_\tau^2 + C_t^2)}{2(1 - \rho)},$$

where

$$\gamma = \begin{cases} \dfrac{2(1 - \rho)(1 - C_t^2)^2}{3\rho(C_\tau^2 + C_t^2)} & C_t^2 \leq 1 \\[2ex] \dfrac{(1 - \rho)(C_t^2 - 1)}{4C_\tau^2 + C_t^2} & C_t^2 \geq 1 \end{cases}$$

and C_τ, C_t are, respectively, the coefficients of variation of the service time and the interarrival time. (Note that $\bar{w}(K - L)$ is exact for an $M/G/1$ system.)

6.4 Case: $H_2/E_3/1$

As a last numerical example, we consider an $H_2/E_3/1$ system. Specifically,

$$\hat{A}(s) = \frac{p_1\lambda_1}{\lambda_1 + s} + \frac{p_2\lambda_2}{\lambda_2 + s}$$

$$\hat{B}(s) = \frac{1}{(1 + s)^3}.$$

Again, the exact waiting time is relatively easy to find via standard techniques (see Ref. 10), which yield

$$W(x) = 1 - C_1 e^{-a_1 x} - C_2 e^{-a_2 x} - \bar{C}_2 e^{-\bar{a}_2 x},$$

where the a_i are the roots of the equation

$$\hat{K}(s) = \hat{A}(-s)\hat{B}(s) = 1,$$

$\bar{a}$ denotes the complex conjugate of a, and the C_i's are the corresponding residues which are readily obtainable (see Ref. 10).

For the approximations of the preceding sections, we find that

$$\bar{u} = 3 - \frac{p_1}{\lambda_1} - \frac{p_2}{\lambda_2}$$

$$\sigma_u^2 = 3 + 2\left(\frac{p_1}{\lambda_1^2} + \frac{p_2}{\lambda_2^2}\right) - \left(\frac{p_1}{\lambda_1} + \frac{p_2}{\lambda_2}\right)^2$$

$$K(0) = p_1 q_1 + p_2 q_2$$

$$\hat{K}_-(a) = \frac{p_1 q_1 \lambda_1}{\lambda_1 + a_1} + \frac{p_2 q_2 \lambda_2}{\lambda_2 + a_1},$$

Table VIII—$P(\text{Delay} \leq x)$ for case $H_2/E_3/1$

x	True	A^*	A^{**}	$A_{0,1}$	$A'_{0,1}$
0	0.7834	0.7897	0.7835	0.7835	0.7835
0.5	0.8121	0.8265	0.8225	0.8212	0.8080
1	0.8406	0.8569	0.8544	0.8523	0.8377
5	0.9731	0.9694	0.9702	0.9680	0.9740
10	0.9977	0.9955	0.9959	0.9953	0.9978
20	0.9999844	0.9999054	0.9999223	0.9998966	0.9999849
$\bar{w}$	0.5536	0.5458	0.5458	0.5662	0.5523

Note: $\rho = 0.2$, True $P(\text{Delay} \leq x) = 1.0 - 0.3260e^{-0.4974x} + e^{-1.215x}[0.10944 \cos(0.3129x)$
$- 0.08673 \sin(0.3129x)]$, $W_{A'_{0,1}}(x) = 1 - 0.3161e^{-0.4974x} + 0.0995e^{-1.215x}$, $\bar{w}(K - L) = 0.5476$.

Table IX—$P(\text{Delay} \leq x)$ for case $H_2/E_3/1$

x	True	A^*	A^{**}	$A_{0,1}$	$A'_{0,1}$
0	0.5728	0.5867	0.5730	0.5730	0.5727
0.5	0.6145	0.6408	0.6305	0.6280	0.6104
1	0.6578	0.6877	0.6803	0.6752	0.6550
5	0.8997	0.8982	0.8996	0.8913	0.9008
10	0.9806	0.9749	0.9764	0.9723	0.9808
20	0.9993	0.9985	0.9987	0.9982	0.9993
40	0.9999990	0.99999997	0.9999998	0.9969756	0.9999990
$\bar{w}$	1.497	1.475	1.475	1.561	1.494

Note: $\rho = 0.4$, True $P(\text{Delay} \leq x) = 1.0 - 0.5176e^{-0.3285x} + e^{-1.265x}[0.09038 \cos(0.3793x)$
$- 0.06644 \sin(0.3793x)]$, $W_{A'_{0,1}}(x) = 1 - 0.5124e^{-0.3285x} + 0.0851e^{-1.265x}$, $\bar{w}(K - L) = 1.488$.

Table X—$P(\text{Delay} \leq x)$ for case $H_2/E_3/1$

x	True	A^*	A^{**}	$A_{0,1}$	$A'_{0,1}$
0	0.1785	0.1906	0.1791	0.1791	0.1785
0.5	0.2041	0.2253	0.2122	0.2122	0.2024
1.0	0.2325	0.2586	0.2489	0.2440	0.2314
5.0	0.4609	0.4779	0.4738	0.4562	0.4616
10.0	0.6588	0.6632	0.6627	0.6398	0.6592
50.0	0.9912	0.9899	0.9904	0.9866	0.9912
100	0.9999093	0.9998741	0.9998874	0.9997827	0.9999094
$\bar{w}$	9.285	9.230	9.230	9.967	9.928

Note: $\rho = 0.8$, True $P(\text{Delay} \leq x) = 1.0 - 0.8517e^{-0.0915x} + e^{-1.321x}[0.03026 \cos(0.4548x)$
$- 0.02125 \sin(0.4548x)]$, $W_{A'_{0,1}}(x)$;q $1 - 0.8517^{-0.0915x} + 0.0302e^{-1.321x}$, $\bar{w}(K - L) = 9.266$.

where

$$q_1 = \hat{B}(\lambda_i) = \frac{1}{(\lambda_i + 1)^3}.$$

Tables VIII, IX, and X compare our various approximations for this case with $p_1 = 0.25$, $\lambda_2 = 2\lambda_1$ and $\rho = 0.2$, 0.4, and 0.8.[†] (The approximation denoted by $A'_{0,1}$ will be discussed shortly.) Here we see again that A^* and A^{**} yield quite good approximations over a wide range.

Again, for comparison, we have included the approximation $\bar{w}(K -$

[†] The roots needed for the exact solution (and the approximations) were obtained by using a program developed by A. E. Eckberg.

L). Unlike the $D/D_2/1$ case, we see that $\bar{w}(K - L)$ provides a good approximation to the mean delay for these cases.

VII. EXTENSIONS

As we have seen, even with the assumption of only a single exponential for the form of $W(x)$, we are led to a wide variety of approximations by using the Lindley integral equation (1) to determine the two unknown coefficients. We consider here some extensions for the case where we wish to choose a more complicated form for $W(x)$.

7.1 Method of moments for hyperexponentials

In many cases, $W(x)$ admits the expansion

$$W(x) = 1 - \sum_{i=1}^{\infty} c_i e^{-a_i x}, \tag{71}$$

where the coefficients C_i, a_i are not necessarily real. In such cases, an approximation of the form

$$W_A(x) = 1 - \sum_{i=1}^{m} C_i e^{-a_i x} \tag{72}$$

seems most appropriate. In general, of course, we do not know the form of $W(x)$. However, if the equation

$$\hat{K}(-s) = \hat{A}(s)\hat{B}(-s) = 1 \tag{73}$$

can be solved more than one nonzero root, then it seems reasonable to attempt to incorporate additional roots in an approximation of the form (72). One can discretize (1) to obtain a set of (implicit) equations for the needed C_i; however, we show here that the method of moments introduced in Section III can readily be extended to obtain a set of linear equations for these quantities. For this purpose, it is convenient to consider the equation for $W^c(x) = 1 - W(x)$ corresponding to (1), i.e.,

$$L[W^c](x) = W^c(x) - 1 + K(x) - \int_{-\infty}^{x} W^c(x - y)dK(y) = 0. \tag{74}$$

We denote the Lth moment of (74) by μ_L:

$$\mu_L = \int_0^{\infty} x^L dL[W^c](x). \tag{75}$$

Using (74) in (75) we obtain

$$\mu_0 = \int_0^{\infty} dW^c(x) + \int_0^{\infty} dK(x) + \int_{-\infty}^{0} W^c(-y)dK(y)$$

$$\mu_L = \int_0^\infty x^L dW^c(x) + \int_0^\infty x^L dK(x)$$

$$+ L \int_0^\infty x^{L-1} \int_{-\infty}^x W^c(x-y) dK(y) dx, \qquad L > 0. \qquad (76)$$

With the notation

$$\overline{w_L^c} = \int_0^\infty x^L dW^c(x), \qquad \overline{u_L} = \int_0^\infty x^L dK(x)$$

(76) can be written

$$\mu_0 = \overline{w_0^c} + \overline{u_0} + \int_{-\infty}^0 W^c(-y) dK(y)$$

$$\mu_L = \overline{w_L^c} + \overline{u_L}$$

$$+ L \int_0^\infty x^{L-1} \int_{-\infty}^x W^c(x-y) dK(y) dx, \qquad L > 0. \qquad (77)$$

[Note $\bar{u}_0 = 1 - K(0)$ and $\bar{u}_1 = \bar{u}_+$ of the preceding sections.]
   Now using the form

$$W_A^c(x) = \sum_i C_i \bar{e}^{a_i x} \qquad (78)$$

$\mu_0 = 0$ leads to

$$\sum_i C_i(K_i - 1) = -\overline{u_0}, \qquad (79)$$

where

$$K_i = \int_{-\infty}^0 e^{a_i y} dK(y).$$

After some algebra, $\mu_L = 0$ for $L > 0$ yields

$$\sum_i C_i \left(\frac{L! K_i}{a_i^L} + \frac{L \bar{u}_{L-1}}{a_i} + \sum_{k=1}^{L-1} \frac{L(L-1) \cdots (L-k)}{a_i^{k+1}} \bar{u}_{L-1-k} - \frac{L!}{a_i^L} \right)$$

$$= -\bar{u}_L. \qquad (80)$$

The empty sum is taken to be zero.
   Thus, given a set of roots a_i, $i = 1, \cdots, m$ of (73) eqs. (79) and (80)
allow us to readily compute the desired coefficient C_i. (Note that this
method is identical to the common techniques of "method of moments"
frequently used for other classes of integral equations [e.g., see Ref.
13).] We refer to the resulting approximation as

$$A_{0,1,\cdots,m-1}^E.$$

Table XI—Dominant residue for case $H_2/E_1/1$

ρ	True	$C_{E,0}$	$C_{E,1}$	$C_{0,1}$	$C_1(A'_{0,1})$
0.2	0.3260	0.2103	0.2714	0.2164	0.3161
0.4	0.5176	0.4133	0.4841	0.4270	0.5124
0.8	0.8517	0.8094	0.8444	0.8209	0.8517

7.2 Numerical example ($H_2/E_3/1$)

Note that if we apply the results of (71) to the $H_2/E_3/1$ example of Section VI and include all three roots, a_1, a_2, a_3, then the resulting delay distribution will be exact [that is, if all roots are given, (80) is equivalent to the standard methods of determining the appropriate residues]. To see how we can use more structure to improve our approximations, we will use the true dominant root a_1 as one of our roots, but take $a'_2 = \text{Re}(a_2)$ as another root, the resulting approximation is shown on Tables VIII, IX, and X, where it is denoted by $A'_{0,1}$. We see that the inclusion of this additional term results in an approximation that more closely captures the structure of the true delay distribution. This is perhaps best illustrated by Table XI. Here we have given the values for the dominant residue, i.e., the coefficient of the exponential with the dominant exponent, for the various approximations. We see that of the three single exponential approximations $A_{E,1}$ has a C value nearest the true value. This shows why $A_{E,1}$ tends to have better tail behavior than the other single exponential approximation. However, we see that $A'_{0,1}$ resulted in an excellent approximation to the true dominant residue, even though we did not use another exact root. Hence, the excellent agreement on Tables VIII to X for the extreme tails of the distribution. This behavior can be very important in studying computer systems with dedicated real-time applications where often criteria are specified in the 10^{-5} probability range, i.e., the probability of delay greater than T shall be less than 10^{-5}.

VIII. CONCLUSIONS

The basic idea we have exploited is to choose a functional form for an approximation $W_A(x)$ to a true delay distribution, say, $W(x)$, and use the well-known Lindley integral equation to find the undetermined coefficients. For the case where $W_A(x)$ is exponential—$W_A(x) = 1 - Ce^{-ax}$—we have used this technique to develop several approximations, some of which make use of the explicit structure of the relevant service and interarrival time distributions, while others require only moment information.

Although not always the best choice, A^* seems to provide the robustness that one would require of a good approximation. The resulting approximation for the mean delay is excellent and the result-

ing probability of delay quite good, although A^{**} provides a better value for the latter. For predicting tails of distributions, we see that $A_{E,1}$ is the best of the simple exponential approximations. We have also seen that increasing the complexity of the forms of the waiting time distribution assumed in the approximation, e.g., using more than one exponential, can result in extremely accurate predictions of the tails of the delay distribution.

IX. ACKNOWLEDGMENTS

APPENDIX
Summary of Notations and Formulas

$$\tilde{t} = \text{interarrival time}$$
$$A(t) = \text{interarrival time distribution}$$
$$\text{L.S.T.} = \text{Laplace-Stieltes Transform}$$
$$\hat{A}(s) = \text{L.S.T. of } A(t)$$
$$E = \text{Expected value}$$
$$\bar{t} = E(\tilde{t}) = 1/\lambda$$
$$\tilde{\tau} = \text{service time}$$
$$B(\tau) = \text{service time distribution}$$
$$\hat{B}(s) = \text{L.S.T. of } B(\tau)$$
$$\bar{\tau} = E(\tilde{\tau}) = 1/\alpha$$
$$\rho = \lambda/\alpha$$
$$\tilde{u} = \tilde{\tau} - \tilde{t}$$
$$K(u) = \text{distribution function of } \tilde{u}$$
$$\hat{K}(s) = \text{L.S.T. of } K(u) = \hat{A}(-s)\hat{B}(s)$$
$$\bar{u} = E(\tilde{u}) = 1/\alpha - 1/\lambda$$
$$\sigma_u^2 = \text{variance of } u$$

$$\hat{K}_-(s) = \int_{-\infty}^{0} e^{su} dK(u)$$

$$\hat{K}_+(s) = \int_{0}^{\infty} e^{su} dK(u)$$

$$\hat{K}(s) = \hat{K}_-(-s) + \hat{K}_+(-s)$$

$$a_E = \text{positive real root of } \int_{-\infty}^{\infty} e^{au} dK(u) = 1$$

$$a_i = \text{ordered roots of characteristic equation,}$$
$$\hat{K}(-a) = A(s)B(-s) = 1(a_1 = a_E)$$

$$\bar{u}_+ = \int_0^\infty x dK(x) = \hat{K}'_+(0)$$

$$K(0) = \hat{K}_-(0)$$
$$a_H = -2\bar{u}/\sigma_u^2$$
$$p = \hat{A}(\alpha)$$
$$q = \hat{B}(\lambda)$$
$$W(x) = \text{waiting time distribution}$$
$$P_D = \text{true probability delay greater than zero}$$
$$\bar{u} = \text{true mean delay}$$
$$W^C(x) = 1 - W(x) = \text{complimentary waiting time distribution}$$
$$L[W^C](x) = \text{Lindley integral equation (complimentary)}$$

$$L[W^C](x) = W^C(x) - 1 + K(x) - \int_{-\infty}^x W^C(x-y)dK(y)$$

$$\mu_L = \int_0^\infty x^L dL[W^C](x)$$

$$K_i = \hat{K}_-(a_i)$$

Approximate waiting time distribution (single exponential):

$$W_A(x) = 1 - Ce^{-ax}$$

$$C = \text{approximate probability of delay}$$

$$\bar{w}_A = C/a = \text{approximate mean delay}.$$

Approximate $A_{E,0}$:

$$a = a_E$$

$$C = C_0(a_E) = \frac{1 - K(0)}{1 - \hat{K}_-(a_E)} = \frac{1 - \hat{K}_-(0)}{1 - \hat{K}_-(a_E)}.$$

Approximation $A_{E,1}$:

$$a = a_E$$

$$\bar{w}_{E,1} = \frac{C_1(a_E)}{a_E} = \frac{\bar{u}_+}{K(0) - \hat{K}_-(a_E)} = \frac{\hat{K}'_+(0)}{\hat{K}_-(0) - \hat{K}_-(a_E)}.$$

Approximation $A_{0,1}$:

$$a = a_{0,1} \text{ solution of}$$

$$a = \frac{[(K(0) - \hat{K}_-(a)][1 - K(0)]}{\bar{u}_+[1 - \hat{K}_-(a)]} = \frac{[\hat{K}_-(0) - \hat{K}_-(a)][1 - \hat{K}_-(0)]}{\hat{K}'_+(0)[1 - \hat{K}_-(a)]}$$

$$C = C_{0,1} = \frac{1 - K(0)}{1 - \hat{K}_-(a_{0,1})} = \frac{1 - \hat{K}_-(0)}{1 - \hat{K}_-(a_{0,1})}.$$

Approximation A^*:

$$C = C^* = C_0(a_E) \text{ (from } A_{E,0})$$

$$\bar{w}^* = \frac{C^*}{a^*} = \frac{C_1(a_E)}{a_E} \text{ (from } A_{E,1})$$

$$\text{determines } a^* = \frac{C_0(a_E)}{C_1(a_E)} a_E.$$

Approximation A^{**}:

$$C = C^{**} = C_{0,1} \text{ (from } A_{0,1})$$

$$\bar{w}^{**} = \frac{C^{**}}{a^{**}} = \frac{C_1(a_E)}{a_E} \text{ (from } A_{E,1})$$

$$\text{determines } a^{**} = \frac{C_{0,1}}{C_1(a_E)} a_E.$$

Approximation $A_{H,0}$:

$$a = a_H = \frac{-2\bar{u}}{\sigma_u^2}$$

$$C = C_{H,0} = C_0(a_H) = \frac{1 - K(0)}{1 - \hat{K}_-(a_H)} = \frac{1 - \hat{K}_-(0)}{1 - \hat{K}_-(a_H)}.$$

Approximation $A_{H,1}$:

$$a = a_H$$

$$\bar{w}_{H,1} = \frac{C_{H,1}}{a_H} = \frac{C_1(a_H)}{a_H} = \frac{\bar{u}_+}{K(0) - \hat{K}_-(a_H)} = \frac{K'_+(0)}{\hat{K}_-(0) - \hat{K}_-(a_H)}.$$

Approximation A_{HT}:

$$a = a_H$$

$$\bar{w}_{HT} = \frac{C_{HT}}{a_H} = \frac{1}{a_H}.$$

Approximation $A_{G,0}$:

$$a = a_H$$

$$C_{G,0} = \frac{1 - \text{Erf}\left(\dfrac{-\bar{u}}{\sqrt{2}\sigma_u}\right)}{1 + \text{Erf}\left(\dfrac{-\bar{u}}{\sqrt{2}\sigma_u}\right)}.$$

Approximation $A_{G,1}$:

$$a = a_H$$

$$\bar{w}_{G,1} = \frac{C_{G,1}}{a_H} = \frac{\dfrac{\sigma_u}{\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{\bar{u}}{\sigma_u}\right)^2} + \dfrac{\bar{u}}{2}\left[1 - \operatorname{Erf}\left(\dfrac{-\bar{u}}{\sqrt{2}\sigma_u}\right)\right]}{\operatorname{Erf}\left(\dfrac{-\bar{u}}{\sqrt{2}\sigma_u}\right)}.$$

Approximation A_{LT}:

$$C_{LT} = 1 - K(0) = 1 - \hat{K}_-(0)$$

$$\bar{w}_{LT} = \frac{C_{LT}}{a_{LT}} = \frac{\bar{u}_+}{K(0)} = \frac{\hat{K}'_+(0)}{\hat{K}_-(0)}.$$

Approximate waiting time distribution (multiple exponentials):

$$W_A(x) = 1 - \sum_{i=1}^{m} C_i e^{-a_i' x}.$$

Approximation $A^E_{0,1,\ldots,m}$:

$$a_i \text{ roots of } \hat{K}(-a) = 1$$

$$a_i' = a_i, \qquad i = 1, \cdots, m$$

$$C_i \text{ determined from}$$

$$\sum_{i=1}^{m} C_i \left(\frac{L! K_i}{a_i^L} + \frac{L \bar{u}_{L-1}}{a_i} + \sum_{k=1}^{L-1} \frac{L(L-1)\cdots(L-k)}{a_i^{k+1}}\, \bar{u}_{L-1-k} - \frac{L!}{a_i^L} \right) = -\bar{u}_L,$$

where the empty sum is taken to be zero.
Approximation $A'_{0,1}$:

$$a_i \text{ roots of } \hat{K}(-s) = \hat{A}(s)\hat{B}(-s) = 1$$

$$a_1' = a_1 = a_E$$

$$a_2' = \operatorname{Re}(a_2)$$

$$C_i \text{ determined from}$$

$$(1 - K_1)C_1 + (1 - K_2)C_2 = \overline{u}_0 = 1 - K(0)$$

$$\frac{[K(0) - K_1]}{a_1'} C_1 + \frac{[K(0) - K_2]}{a_2'} C_2 = \bar{u}_1 = \bar{u}_+.$$

REFERENCES

1. A. A. Fredericks, "Analysis of a Class of Schedule with Real Time Applications," *Performance of Computer Systems*, M. Arrato, A. Butrimenko, and E. Gelinbe, eds., Amsterdam, New York, and Oxford: North-Holland Publishing Company, pp. 201–16.
2. A. A. Fredericks, "Analysis and Design of Processor Schedules for Real Time

Applications," *Proceedings of the Conference on Applied Probability and Computer Science, The Interface,* R. Disney and T. J. Ott, eds., Boston: Birkhauser, March, 1982.

3. D. V. Lindley, "The Theory of Queues with a Single Server," Proc. Cambridge Philosophical Soc., *48* (1952), pp. 277–89.

4. J. F. C. Kingman, "The Heavy Traffic Approximation in the Theory of Queues," *Congestion Theory,* W. L. Smith, and W. E. Wilkinson, eds., Chapel Hill: Univ. N. Carolina Press, 1964, pp. 137–69.

5. W. Feller, *An Introduction to the Theory of Probability and its Applications,* New York: John Wiley, 1966, 1971.

6. J. F. C. Kingman, "Inequalities in the Theory of Queues," J. Royal Statist. Soc. *B32* (1970), pp 102–10.

7. H. J. Rossberg, and C. Siegel, "The Importance of Kingman's Integral-Inequalities ...," Math. Operationsforsh. Statist. *5* (1974), pp. 687–99.

8. S. M. Ross, "Bonds on the Delay Distribution in $GI/G/1$ Queues," J. Appl. Prod. *11* (1974), pp. 417–21.

9. R. Bergmann, and D. Stoyan, "Exponential Bounds for the Waiting-Time Distribution Function in $GI/G/1$," J. Appl. Prob., *13* (1976), pp. 411–7.

10. D. Gross and C. M. Harris, *Fundamentals of Queueing Theory,* New York: John Wiley, 1974.

11. W. Kramer and M. Lagenbach-Belz, "Approximate Formulae for the Delay in the Queueing System $GI/G/1$," Congressbook, 8th Int. Teletraffic Conf., Melbourne, 1976, pp. 235.1–235.8.

12. D. P. Heyman, "A Diffusion Approximation for the $GI/G/1$ Queues in Heavy Traffic," B.S.T.J., *54* (1975), pp. 1637–46.

13. A. Yu. Luchka, *The Method of Averaging Functional Corrections, Theory and Applications,* New York: Academic Press, 1965.

A General Class of Zero- or Minimum-Delay Fractional Rate Change Circuits

By S. V. AHAMED

Rate changing occurs whenever sequences of data undergo transformations in rate without undergoing a change in the order of sequence. When the ratio of transformation is not an integer, fractional rate changes are necessary. These are generally, a prerequisite for the time-compression multiplexing mode of data transmission. Zero or minimal delay is a desirable characteristic, for example, in reducing the annoyance from the far-end echo whenever voice is encoded and transmitted. Conventional fractional rate changing entails an inherent delay in the rate change circuits. Segmenting shift registers reduces the delay of the last bit without completely eliminating it, unless the shift-register length is reduced to one bit. In this paper, a method of partitioning the shift registers by logarithmic counts is developed to reduce the complexity of the gating and the counting circuits. Zero last-bit delays are attainable in all cases where the rate increase is greater than two or, conversely, the rate reduction is less than half. For the remaining cases, the compromise between circuit complexity and the last-bit delay is outlined.

I. INTRODUCTION

Under the time-compression mode of data transmission, the round trip delay time is critically important in controlling the echo from the far end, as well as the direct transmission delay. The excess delay of the last bit, in changing the rate from the primary (or terminal) rate to the secondary [or time-compression multiplexing (TCM)] rate and vice versa, which can be a significant fraction of the overall delay, in any particular block is reduced to zero by the circuits presented in this paper. In general, it is shown that the last-bit delay can be reduced to zero whenever the required rate change is more than two, while increasing the number of shift registers (SRs) and the gating functions

logarithmically. Further, the gating signals are shown to be derived as combinations of logarithmic counts of the primary (terminal) clock or those of the secondary clock. The general principle of fractional rate changing is extended to fractional rate collating circuits. The compromise between zero last-bit delay and the complexity of the circuit for fractional rate changing between half and one is delineated.

In 1971, a general class of rate change circuits was presented[1] for use in the magnetic domain technology. The constraint on the design of such circuits was that all the individual bits of information (domains) be propagated by one period (the physical spacing between one pattern to the next in domain technology) in one clock cycle of the rotating magnetic field. To conform to this design requirement in domain technology, the number of patterns in different paths of the rate change circuit were arranged to follow a geometric progression. The topological arrangement of SRs and gates, if operated in conventional semiconductor technology, will perform satisfactorily. However, it imposes two unnecessary restrictions: (i) this version in semiconductor technology ignores the capability of SRs to shift-in (SI) at one rate and shift-out (SO) at another rate, and (ii) fractional rate change is a two-step procedure, thus, demanding a large number of SR locations. The essential feature retained in these circuits is that the delay between the reception of the last bit of block of data and its transmission would be zero. Zero last-bit delay fractional rate change is not possible with a conventional arrangement* of SRs, unless the number of independent registers is increased to the number of bits in the block, thus, demanding extremely complex arrangements of gating and shifting functions.

In this paper, we present circuits that retain the zero last-bit delay characteristics and the simplicity of the gating function *without* linearly increasing the number of SRs, or the complexity of gating and shifting functions. Even though most of the emphasis is placed on fractional rate changing, integral rate changing is equally easily accomplished by the circuit arrangements presented here.

II. ZERO-DELAY FRACTIONAL RATE INCREASING

For the cases presented in this section, let the proportional rate increase be greater than two. Cases where the rate change is between one and two are discussed in Section IV.

2.1 Fractional rate between two and three

Consider a block 42 bits long where the rate change ratio is 3:7. Forty-two, being a multiple of 3 and 7, generates a situation where one

---

* Conventional arrangement consists of having one or more SRs in which data are sequentially shifted in at the incoming rate and shifted out at the output rate.

block of data 42 bits long experiences the same set of gating functions* that was performed for the last 42-bit data block or the next 42-bit data block. Next, consider an arrangement of 7 SRs shown in Fig. 1. Data are received uniformly one bit ever t seconds. Data are generated in a burst of 42 bits. Burst repeats every $42t$ seconds and lasts for $18t$ seconds. Effective rate change is 3:7 during the burst. Here, the SRs follow a sequence 2^0, 2^1, 2^2, 2^3, 2^4, $(42 - 2^5)$. The signals C and C' are clock signals at t and t', where t denotes the primary (terminal) clock and t' denotes the secondary clock durations in seconds. The signals a_6, a_5, —, a_1, a_0 are generated every $42t$ seconds and last for $(42 - 2^5)$, 2^4, —, —, 2^0, 1 t seconds, etc. Similarly, b_6 through b_0 are generated by identical circuits operated at the secondary clock t' but delayed for the $24t$ of the $42t$ seconds cycle time. The operation of the SRs is summarized in Table I.

It can be seen that all SRs have a positive duration between the finish of the SI and start of SO (see column 6, Table I). As proved in Appendix A, the kth SR experiences a delay of:

$$d_k = 2^{k-1}t - 2^k t' \text{ seconds.} \tag{1}$$

If t is defined as $>2t'$, d_k is always positive. In the example presented here, the delays for SRs 1 to 5 can be equated to d_k in (1) by substituting $t' = \frac{3}{7}t$ both in (1) and in the sixth column of Table I. The delay of the 42nd bit is zero since the ending instants of received and the transmitted bits coincide at the end of the block.

The four distinct characteristics of the circuit configuration presented in this section can, thus, be summarized as follows:

(i) The SI-SO duration is always positive for all SRs.

(ii) The gating signals a_6 through a_0 and b_6 through b_0 are generated simply by identical binary (in cases where the fractional rate change ratio $r > 2$ and < 3) counters each being driven by a primary (terminal) clock and by a secondary clock, respectively.

(iii) The delay in the circuit as measured by the difference between the end of arrival of the last data bit and the end of transmission of that same data bit at the higher clock is zero.†

(iv) The number of SR locations required is always the bit length of the block reduced by one.

2.2 Fractional rate increasing between three and four

In Section 2.1, the number of locations in the SR progresses as 2^0, 2^1, $\cdots$, 2^i, $(N - 2^{i+1})$, depending on the size of the block (N).

---

* Systems considerations usually require such a repetitive set of functions for continuous operations.

† It is possible to advance the operation by an extra $(t - t')$ second by advancing t' by this period and loading the last bit into an SR. But, we have ignored this situation as it needs only a trivial modification of the circuits shown here.
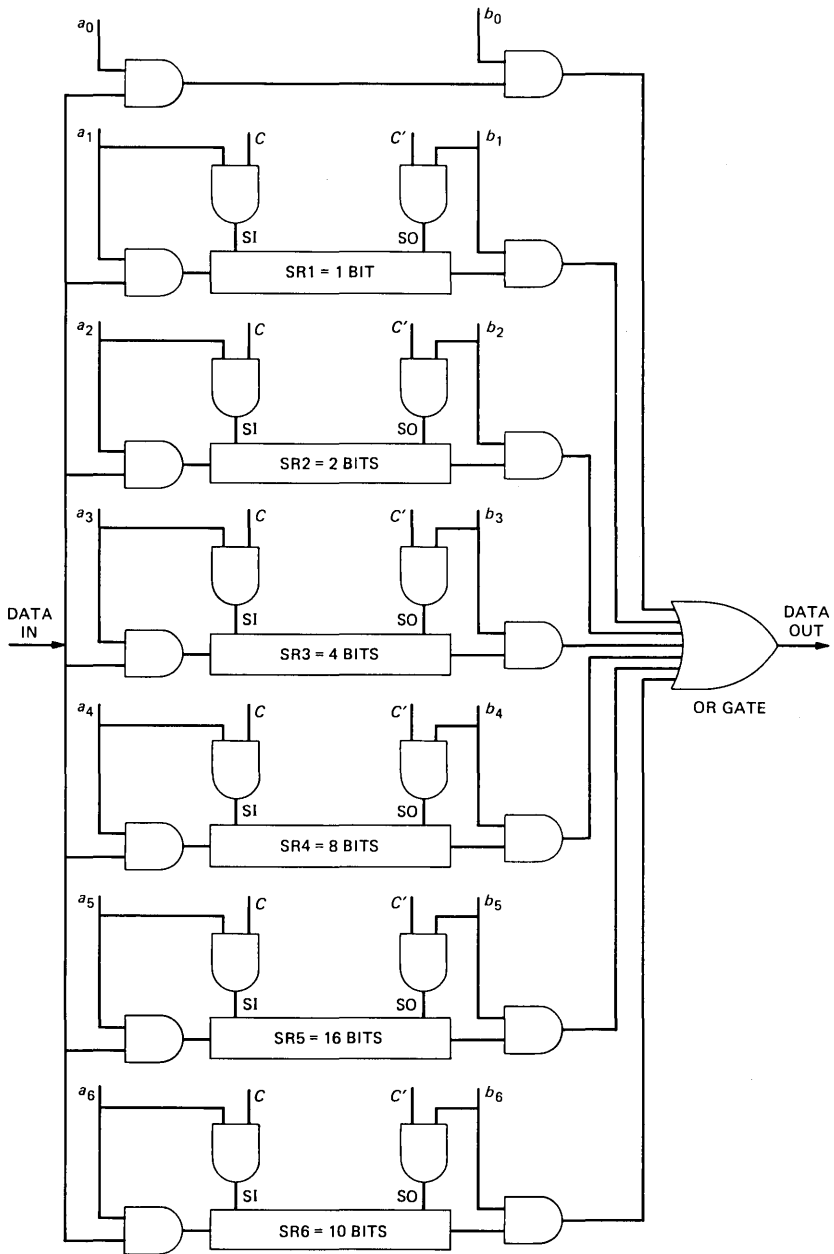
Fig. 1—Shift register arrangement for a 42-bit data block.

Table I—Details of 42-bit block—3:7 rate increasing circuit

SR	Start of SI (a)	Finish of SI (b)	Start of SO (c)	End of SO (d)	SI-SO Interval (c − b)*	On Time SI	On Time SO
SR6	0	10t	$\Delta = 24t$	$\Delta + 10t'$	14t	10t	10t'
SR5	10t	26t	$\Delta + 10t'$	$\Delta + 26t'$	$10t' - 2t$	16t	16t'
SR4	26t	34t	$\Delta + 26t'$	$\Delta + 34t'$	$26t' - 10t$	8t	8t'
SR3	34t	38t	$\Delta + 34t'$	$\Delta + 38t'$	$34t' - 14t$	4t	4t'
SR2	38t	40t	$\Delta + 38t'$	$\Delta + 40t'$	$38t' - 16t$	2t	2t'
SR1	40t	41t	$\Delta + 40t'$	$\Delta + 41t'$	$40t' - 17t$	t	t'
Signal	Start	Finish	Signal	Start	Finish	Finish b_0 −Finish a_0	
a_0	41t	42t	b_0	$\Delta + 41t'$	$\Delta + 42t'$	Zero	

* This interval can be always verified as being positive at $t = 7/3 \times t'$.
Notes:
1. Shift-in takes place during one clock cycle at t.
2. Shift-out takes place during one clock cycle at t'. For maintaining zero delay, it is assumed that the end of the 42nd clock cycle at t coincides with the end of the 98th cycle at t'. This condition ascertains that a_0 and b_0 end simultaneously at the end of $42t$ or $98t'$.

When the fractional rate change ratio r is between 3 and 4, all the essential characteristics of this class of rate change circuit may be maintained by changing the numbered locations in the SRs to follow the progression

$$3^0, 3^1, 3^2, \cdots, 3^i, \left(N - 1 - \sum_{i=0}^{i} 3^i \right).$$

Consider a rate change ratio of 7:23 with a 161-bit block. The SR progression becomes a 1-, 3-, 9-, 27-, 81-, 39-bit position constituting 160 locations. The six SR circuit is shown in Fig. 2, and the shifting times are tabulated in Table II.

The generalization of this arrangement for an N-bit block with any fractional rate change r is presented in Appendix B.

III. FRACTIONAL RATE DECREASING CIRCUITS

When the decrease in rate equals or becomes less than 0.5, we have a situation where the circuit arrangement presented in Section II can be used to decrease the rate. Two changes however, are necessary: (i) the direct transmission of data between a_0 and b_0 (Figs. 1 or 2 have to be replaced by a flip-flop to load and hold the first data bit and, (ii) the signals $a_0, \cdots, a_n$ and $b_0, \cdots, b_n$ must start at the beginning of the block instead of ending at the end of the block.

Consider a data block 680 bits long where the desired rate change ratio is 0.425. The minimum size of the data block is 680, being the least common multiple of 17 and 40, even though equally satisfactory arrangements can be conceived for all blocks that are multiples of 680 bits long. Once again the SR sizes (Fig. 3) can be written down as 1, 2^0,
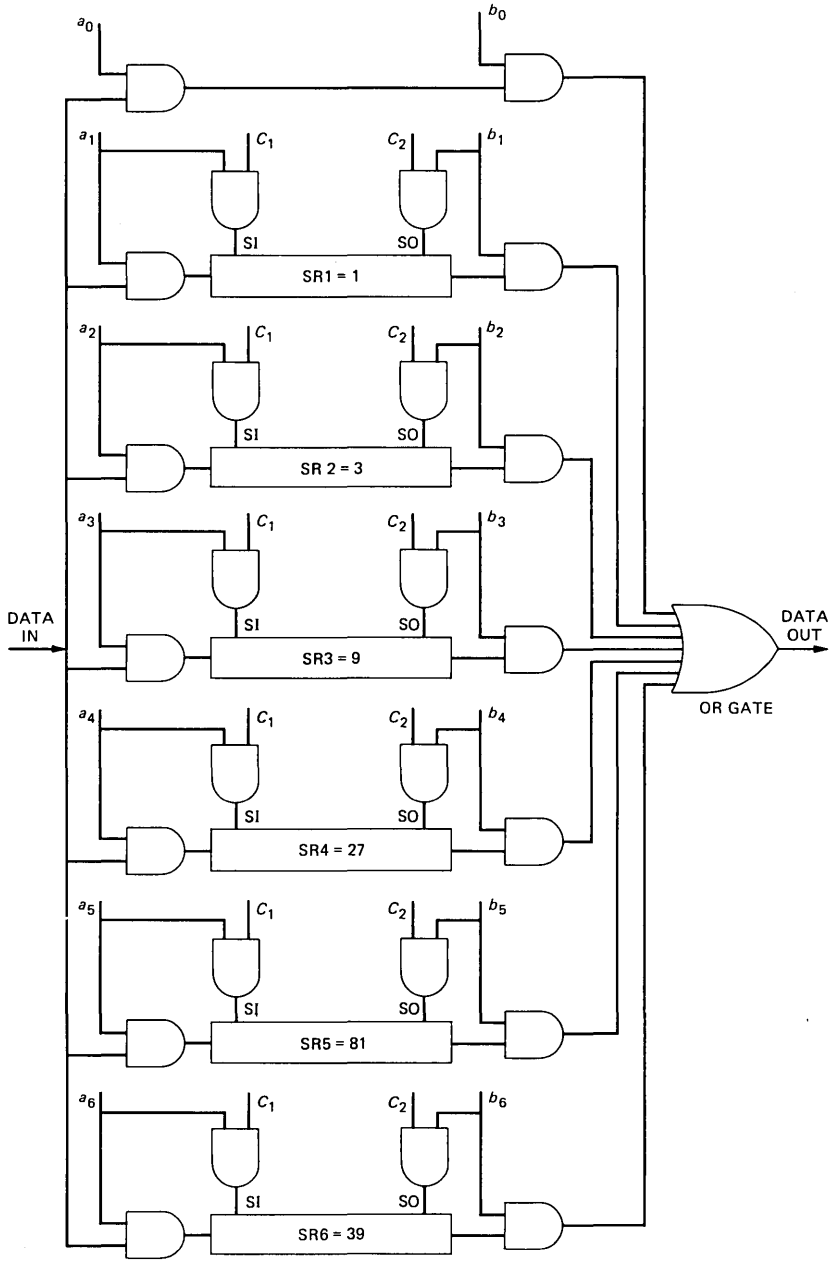
Fig. 2—Shift register arrangement for a 161-bit block. Rate change ratio is 7:23.

Table II—Details of 161-bit block—7:23 rate increasing circuit

SR	Start of SI (a)	Finish of SI (b)	Start of SO (c)	Finish of SO (d)	SI-SO Interval $(c - b)^*$	On Time	
						SI	SO
SR6	0	39t	$\Delta = 112t$	$\Delta + 39t'$	73t	39t	39t'
SR5	39t	120t	$\Delta + 39t'$	$\Delta + 120t'$	$39t' - 8t$	81t	81t'
SR4	120t	147t	$\Delta + 120t'$	$\Delta + 147t'$	$120t' - 35t$	27t	27t'
SR3	147t	156t	$\Delta + 147t'$	$\Delta + 156t'$	$147t' - 44t$	9t	9t'
SR2	156t	159t	$\Delta + 156t'$	$\Delta + 159t'$	$156t' - 47t$	3t	3t'
SR1	159t	160t	$\Delta + 159'$	$\Delta + 160t'$	$159t' - 48t$	t	t'
Signal	Start	Finish	Signal	Start	Finish	Finish b_0 Finish a_0	
a_0	160t	161t	b_0	$\Delta \pm 160t$	$\Delta + 161t'$	Zero	

* This number can be verified as always being positive when $t = 23/7 \times t'$.

2^2, 2^3, 2^4, 2^5, 2^6, 2^7, 2^8, $(680 - 2^9)$. The clocks C and C' are in the proportion of 17:40. The signals a_0, a_1 $\cdots$ a_8, a_9 and b_0, b_1 $\cdots$ b_8, b_9 are generated for $(1, 2^0, 2^1, \cdots, 2^8, 168)t$ and $(1, 2^0, 2^1, \cdots, 2^8, 168)t'$ seconds,* starting at the beginning of the block. The starting instants of these signals are $(0, 1, 2^0, 2^1, \cdots, 2^8)t$ seconds and $(0, 1, 2^0, 2^1, \cdots, 2^8)t'$ seconds from the starting position of the block for a_0, a_1, $\cdots$, a_9 and b_0, b_1, $\cdots$, b_9, respectively. The circuit arrangement is presented in Fig. 3, and the gating times are tabulated in Table III. The principle may be extended to an N-bit block whose effective rate reduction is r, and a circuit configuration similar to that shown in Fig. 6 may be derived.

IV. FRACTIONAL RATE CHANGES BETWEEN ONE AND TWO

A zero-delay circuit configuration for these rate changes may also be derived by extending the general principles of Sections II and III. However, the number of SRs increases $(1^i = 1)$ to the number of bits in the block, and the circuit configuration becomes trivial. On the other hand, it is possible to compromise the zero delay requirement slightly and obtain the advantages of reducing the number of SRs and simplifying the gating functions. These considerations are discussed next.

4.1 Rate increasing ratios between one and two

Consider a data block 84 bits long. If the desired rate increase is 4:7, one can arrange a circuit shown in Fig. 4a that is consistent with the arrangement shown in Fig. 1. However, if SR7 is emptied after a delay of $36t^\dagger$ seconds, then SR6 would not have completely shifted in before it becomes necessary to shift it out, and it is here that the compromise becomes necessary. If the shifting out sequence is now delayed by a

---

* Durations t and t' are cycle times at faster (TCM) and slower (terminal) clocks, respectively.

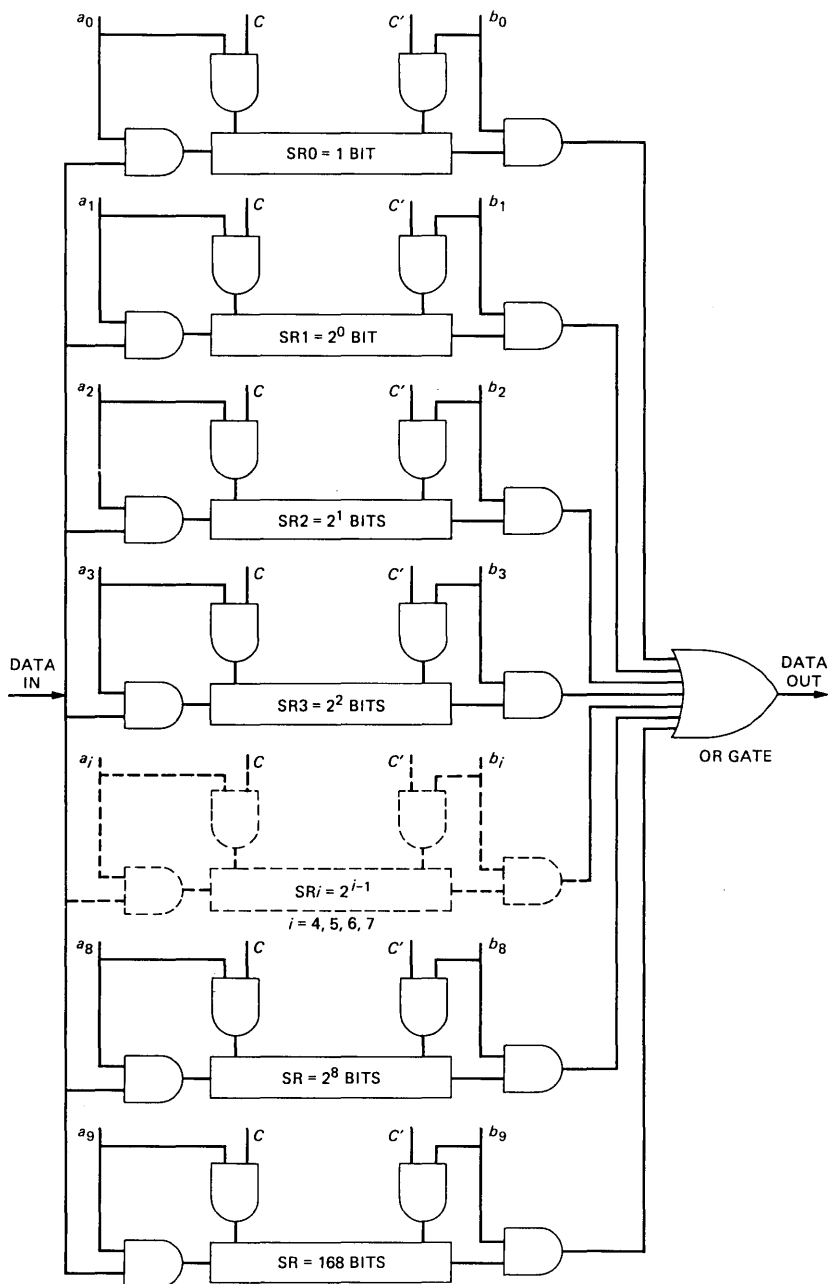† The duration t is a clock cycle time at the terminal (primary) rate and $36 = 84(1-\frac{4}{7})$.

Fig. 3—Rate-reducing circuit for a 680-bit data block. Rate change ratio is 17:40.

Table III—Details of 680-bit block—40:17 rate decreasing circuit

SR	Start of SI (a)	Finish of SI (b)	Start of SO (c)	Finish of SO (d)	SI-SO Interval	On Time	
						SI	SO
SR_0	0	0*	0*	t'	0	0*	t'
SR_1	t	$2t$	t'	$2t'$	$t'-t$	t	t'
SR_2	$2 \times t$	$4 \times t$	$2t'$	$4t'$	$2(t'-t)$	$2t$	$2t'$
SR_i	$2^{i-1} \times t$	$2^i \times t$	$2^{i-1}t'$	$2^i t'$	$2^{i-1}(t'-t)$	$2^{i-1}t$	$2^{i-1}t'$
SR_8	$2^7 \times t$	$2^8 \times t$	$2^7 \times t'$	$2^8 t'$	$2^7(t'-t)$	$2^7 t$	$2^7 t'$
SR_9	$2^8 \times t$	680	$2^8 t'$	$680t'$	$2^8(t'-t)$	$168t$	$168t'$

* Indicates shift-in time of SR_0. This is $\ll t$ or t' and can be used effectively to mark the beginning and end of a 680-bit data block.

minimal amount to just accommodate the binding requirement on this SR, then it can be seen that the delay should be $4.571t$* seconds. Further, this minimal delay depends on the length of the block, the fractional rate change ratio r, and the SR position in the hierarchy of registers. The generalized calculations for the delays are presented in Appendix C.

For the particular case presented in Fig. 4 for the 84-bit block, it is now possible to compute the gating times and the delays for the overall functioning of the circuit. These results are tabulated in Table IV, and Fig. 4b depicts the timings. The minimum delay of $4.571t$ seconds plays a central role in the operation of the circuit. This delay being essential for the functioning of SR6 appropriately, gets fragmented into $2.286t$, $1.143t$, $0.571t$, $0.286t$, $0.143t$, $0.071t$ and, finally, $0.071t$ for SR6 through SR0 to add up to the net delay of $4.571t$ seconds for the circuit. Once this delay is generated by an additional circuitry, the functioning of the gates at the higher clock rate can be normal binary counters, delayed by this appropriate amount. This feature eliminates the need for complicated gating circuits, even when the fractional rate increasing is less than two.

4.2 Compromise between circuit complexity and minimum delay

It can be seen that the entire delay of $4.571t$ seconds can be cut into half by partitioning SR6 into two sixteen-bit registers. In this case, SR7 is delayed by $2.286t$ and the SI finish coincides with the SO start for SR5, and the delay reappears as the summation $1.143t$, $0.571t$, $0.286t$, $0.143t$, $0.071t$, and $0.071t$ for the remaining SRs at the end of the block. If one extends this reasoning indefinitely, then all the SRs become partitioned to single-bit registers for zero delay. It is at this stage that the compromise between SR complexity and minimum delay becomes obvious. All the same, these circuits outperform conventional circuit

---

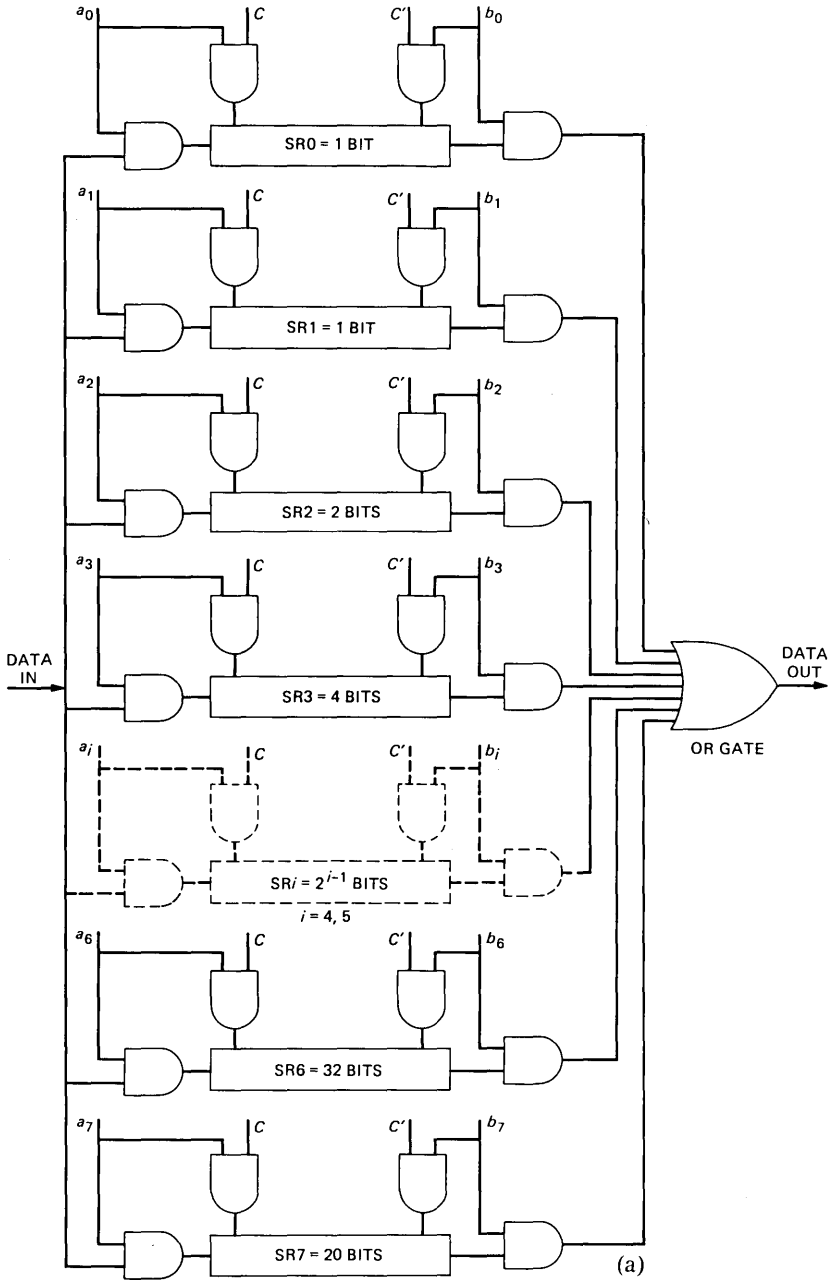* $4.571 = (20 + 32) - (36 + 20 \times \frac{4}{7})$.

Fig. 4a—Rate-increasing circuit for an 84-bit block. Rate change ratio is 4:7.

Table IV—Details of 84-bit block—4:7 rate increasing circuit

SR	Start of SI (a)	Finish of SI (b)	Start of SO (c)	Finish of SO (d)	SI-SO Interval e = (c − b)	On Time SI	On Time SO
SR7	0	20t	Δ = 40.571t	Δ + 20t'	20.571t	20t	20t'
SR6	20t	52t	Δ + 20t'	Δ + 52t'	0	32t	32t'
SR5	52t	68t	Δ + 52t'	Δ + 68t'	2.285t	16t	16t'
SR4	68t	76t	Δ + 68t'	Δ + 76t'	3.429t	8t	8t'
SR3	76t	80t	Δ + 76t'	Δ + 80t'	4.000t	4t	4t'
SR2	80t	82t	Δ + 80t'	Δ + 82t'	4.286t	2t	2t'
SR1	82t	83t	Δ + 82t'	Δ + 83t'	4.429t	t	t'
SR0	83t	84t	Δ + 83t'	Δ + 84t'	4.571t*	t	t'

* Minimum delay = $4.571t$ and $\Delta = (36 + 4.571)t$. In the case of SR0, $(c = d - b)$ to prove that the end of the block at the faster rate ends as late as the minimum delay imposed upon shifting the first data bit out of SR7. (See Fig. 4a.)
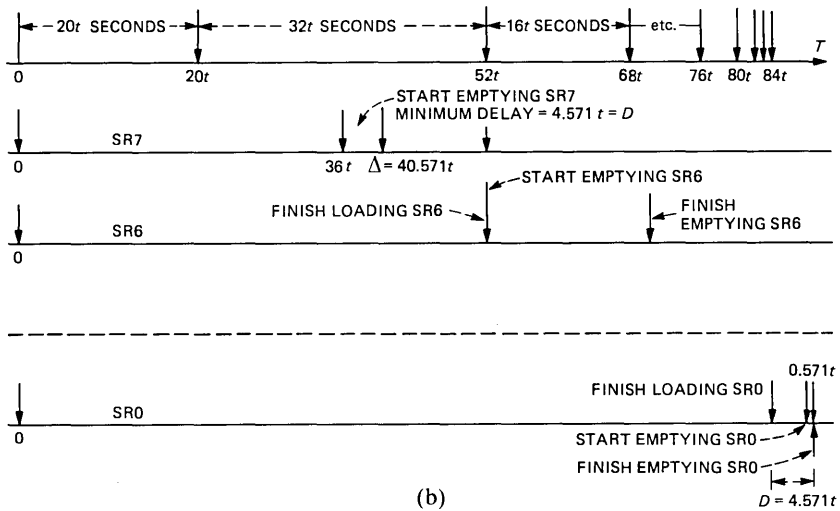

(b)

Fig. 4b—Timing diagram for an 84-bit data block in a 4:7 rate-increasing circuit configuration.

arrangements because the delay time of 7 SRs, each 12 bits long, would be a $12t$ second compared to $4.571t$ for the arrangement suggested here with 8 SRs, or of $2.286t$ with 9 SRs. The extent of gain depends on the length of the block and the rate change ratio. The circuit complexity has to be chosen by a designer depending on the particular requirements.

V. FRACTIONAL RATE COLLATING CIRCUITS

To illustrate the example of fractional rate collating, consider 17 independent data lines each carrying independent data uniformly at a primary rate of one bit every t seconds. The 18th line also carries

binary data but a tertiary rate of one bit every t_1 seconds. Let the block time be T seconds, corresponding to 7 bits at the primary rate t, and 4 bits at the tertiary rate t_1 (i.e., $7t = 4t_1$). The collating circuit will be able to present the first bit of the first channel followed by the first bit of the second channel etc., until the 17th channel, then the second bit of the first channel, followed by the second bit of the second channel, etc., and so on, until the 7th bit of the 17th channel. The final section of the output data block would be the four bits of the 18th channel. The output data rate will be at a uniform secondary bit rate of t' (i.e., $T/123$).

A collating circuit to perform this function is presented in Fig. 5a. Data arrive uniformly every t seconds on 17 channels, $C_1, \cdots, C_{17}$, and t_1 seconds on channel C_{18} with $7t = 4t_1$ data are transformed to 123-bit blocks. The first bit of C_1 is collated with first bit of C_2, etc., until C_{17} the second bit of C_1 is collated with the second bit of $C_2, \cdots$, etc., and the last four bits of the block are bits received on C_{18}. The output rate is t' with $123t' = 7t = 4t_1 = T$. The first bits of all the 17 channels are collated into SR1, the second bits of all the 17 channels are collated into SR2, etc. The 17 blocks of SI signals $m_1, \cdots, m_{17}; m_{18}, \cdots, m_{34}, \cdots$, etc., $\cdots, m_{119}$ are generated by ANDing signals $t, 2t, 3t, \cdots, 7t$, with t'—the signal t being long enough to accommodate 17 full cycles of t'. The SO signals $n_1, \cdots, n_{17}; n_{18}, \cdots, n_{34}; \cdots$, etc., $\cdots, n_{123}$, are generated by delaying the secondary block t' by 21 (i.e., $17 + 4$) full cycles of the secondary clock t'. The SI signals are in two classes: $m_1, \cdots, m_{119}$ are generated as indicated earlier, and $m_{120}, \cdots, m_{123}$ are generated at the tertiary rate of one signal every t_1 seconds. The signal timings are shown in Fig. 5a. Shift-in signals $m_1, \cdots, m_{119}$ are generated by ANDing t with 17 of t'. Signals $m_{120}, \cdots, m_{123}$ are generated by ANDing t with one of t'. Shift-out signals $n_1, \cdots, n_{123}$ are generated at t' after delaying clock by $21t'$ from the start of the block. It is interesting to note that SR1 has a delay time of $4t'$ seconds between SI and SO. However, this time gradually diminishes by $\frac{4}{7}t'$ as one progresses through SR2 to SR7, with SR7 itself shifting out as soon as it is shifted in. (See Fig. 5b.) It is this limit that specifies the minimum delay of the collating circuits.

Other collating circuits can be generated by the principles indicated, and the generalization of such circuits is also possible but omitted here.

VI. DISCUSSION OF RATE CHANGING AND COLLATING CIRCUITS

Conventional SR arrangements lead to a delay in the transmission of data, unless the size of the register diminishes to one. For circuits described here, there is a geometric progression in the sizes of the registers and in the logarithmic number of registers. The base of the
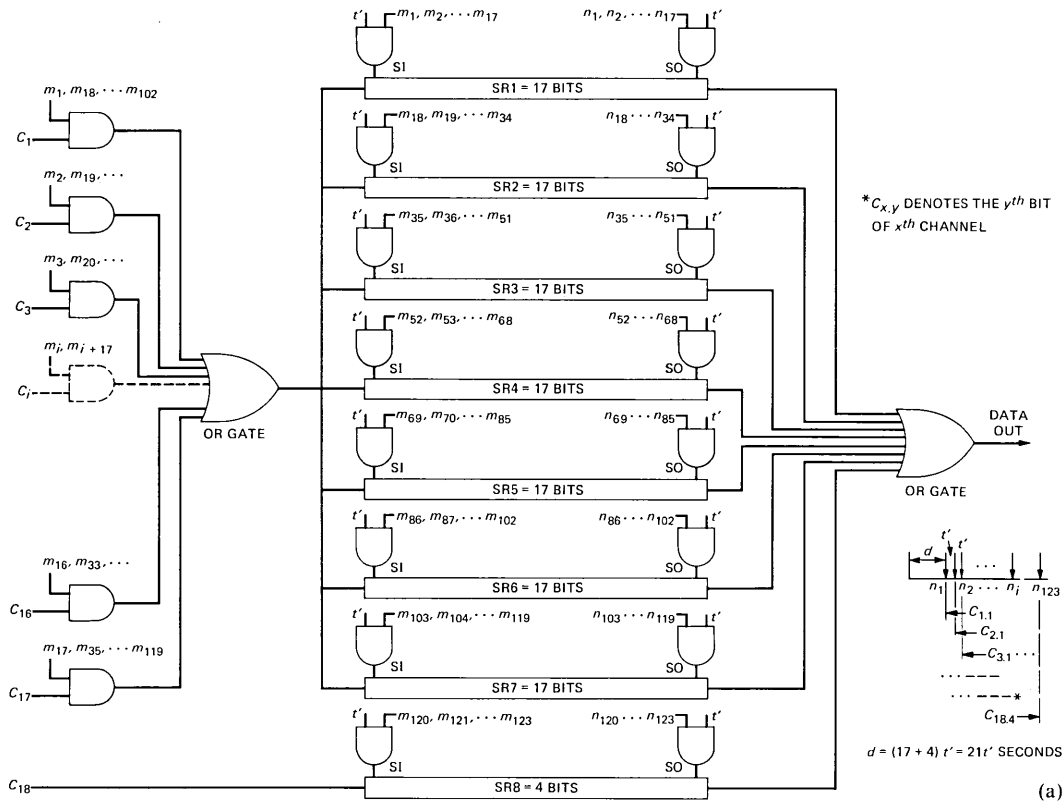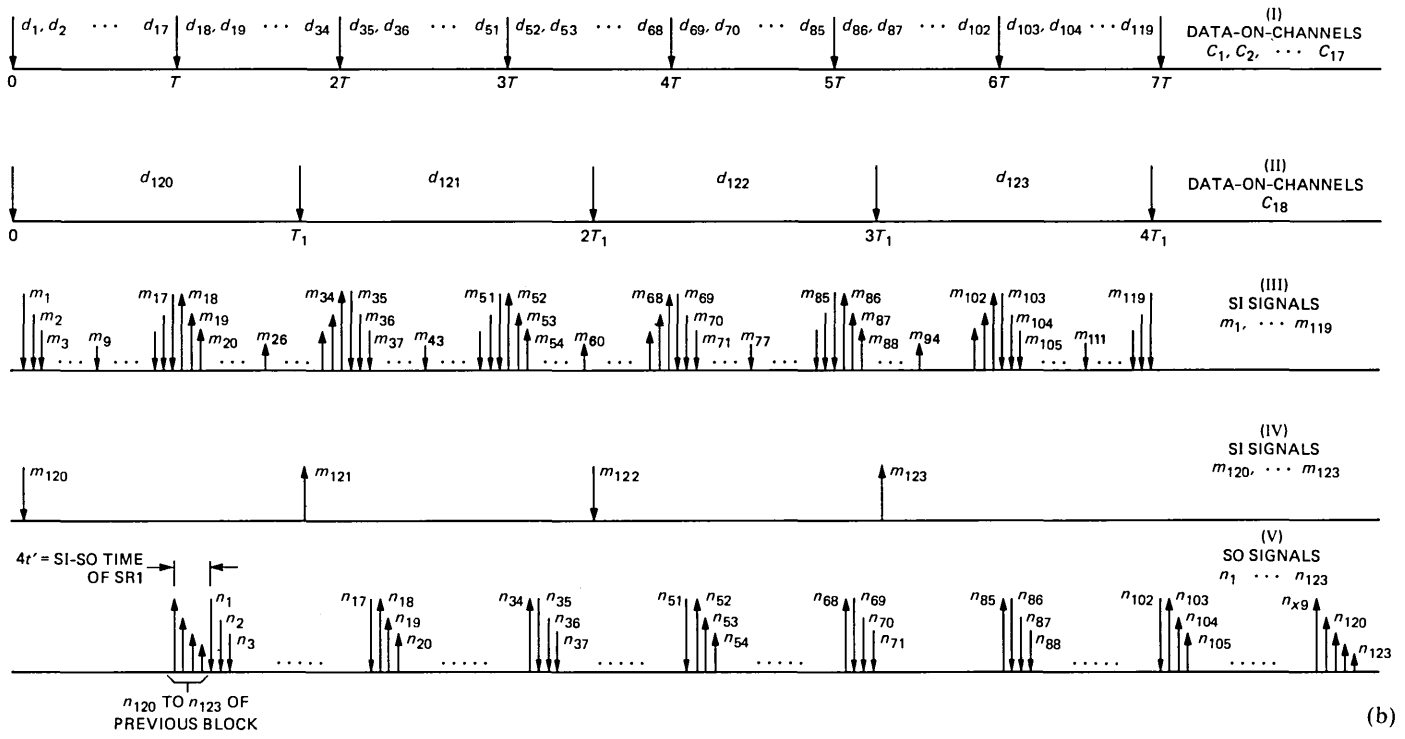
Fig. 5a—Fractional rate collating circuit for a 123-bit data block.

$d_1, d_2 \quad \cdots \quad d_{17}$ | $d_{18}, d_{19} \quad \cdots \quad d_{34}$ | $d_{35}, d_{36} \quad \cdots \quad d_{51}$ | $d_{52}, d_{53} \quad \cdots \quad d_{68}$ | $d_{69}, d_{70} \quad \cdots \quad d_{85}$ | $d_{86}, d_{87} \quad \cdots \quad d_{102}$ | $d_{103}, d_{104} \cdots d_{119}$

(I)
DATA-ON-CHANNELS
$C_1, C_2, \quad \cdots \quad C_{17}$

0   T $2T$ $3T$ $4T$ $5T$ $6T$ $7T$

d_{120} d_{121} d_{122} d_{123}

(II)
DATA-ON-CHANNELS
C_{18}

0   T_1 $2T_1$ $3T_1$ $4T_1$

(III)
SI SIGNALS
$m_1, \quad \cdots \quad m_{119}$

(IV)
SI SIGNALS
$m_{120}, \quad \cdots \quad m_{123}$

m_{120} m_{121} m_{122} m_{123}

(V)
SO SIGNALS
$n_1 \quad \cdots \quad n_{123}$

$4t' = \text{SI-SO TIME}$
OF SR1

n_{120} TO n_{123} OF
PREVIOUS BLOCK

(b)

Fig. 5b—Timing diagram for circuit shown in Fig. 5a.

geometric series is chosen to be an integer just under the fractional rate change ratio. The time for shifting out any particular register in the series is made to overlap with the time for shifting in of the next lower-sized register of the series. Two definite advantages arise by this arrangement of SRs: (*i*) the number of registers can be reduced exponentially (see Appendices A through C) for a given delay, and (*ii*) the gating can be performed very simply by counters that block out exponentially varying sizes of time at the primary or secondary clocks. These features are absent in the conventional arrangement of registers.

When the fractional rate change falls between one and two, the arrangement suggested still outperforms conventional arrangement as far as delay is concerned for a given number of SRs. In a limit, the two systems converge to a minimum delay with one-bit registers. The compromise between circuit complexity and minimum delay is offered to the designer.

Collating circuits also retain some features of minimizing the delay. Consider the circuit presented in Section V. Here, conventional arrangements would have two extremes of implementation. First, consider 17-seven and 1-four bit SRs, all shifting in and then shifting out. The delay is obviously $123t'$ since all the registers have to be shifted in before shifting out. Second, consider 123 one-bit registers. With an immense amount of complexity in gating and clocking, one-bit delay may be achieved. The first arrangement does not compare favorably, from the delay consideration, with the configuration of the 18 registers—the delay being $123t'$ versus $21t'$. The second arrangement offers an advantage in delay time (t' versus $21t'$) but at an enormous complexity (123 timing circuits and SRs versus 18 such circuits). As indicated in Section 4.2, a compromise between circuit complexity and minimum delay also exists for collating circuits.

VII. CONCLUSIONS

Fractional rate changing may be achieved with zero last-bit delay *without* increasing the number of SRs to the number in the bits in the block. In the circuit arrangement presented, the number of SRs increases logarithmically. The base for the logarithmic increase is the highest integer just below the rate change ratio. When the rate change ratio is between one and two, the simplicity of the SR arrangement and gating may be retained by accepting a slight last-bit delay in the circuit. Such an arrangement, even though not reducing the last-bit delay to zero, still outperforms a conventional arrangement by a large margin. The exact compromise between complexity and delay is left to the circuit designer.

Fractional rate collating of data blocks have some of the advantages so far as the simplicity of SR arrangement and gating is concerned.

VIII. ACKNOWLEDGMENT

APPENDIX A

Shift Register Delays

Consider an N ($>2^{j-1}$ and $<2^j$) bit block in which a rate change of r (≥ 2 and <3) is being accomplished. The arrangement of a circuit with j srs consists of 2^0, 2^1, $\cdots$, 2^{j-2}, $(N - 2^{j-1})$ bits each. The size of the kth ($k < j - 1$) sr being 2^{k-1} finishes at an instant of:

$$d_i = \left(N - 1 - \sum_{i=0}^{i=k-2} 2^i \right) t \text{ seconds.}^*$$ (2)

Similarly, the start of the so of the kth sr does not take place until an instant

$$d_o = N(t - t') + \left(N - 1 - \sum_{i=0}^{i=k-1} 2^i \right) t' \text{ seconds.}$$ (3)

The duration between the instant at which so starts and the instant at which si ends can be obtained by subtracting (2) from (3) leading to

$$d_k = 2^{k-1}t - 2^k t' \text{ seconds,}$$ (4)

since

$$2^k = 1 + \sum_{i=0}^{i=k-1} 2^i.$$ (5)

For this Appendix, by definition[†] we have $t \geq 2t'$ and d_k always remains positive.

APPENDIX B

Generalized Fractional Rate Increasing Circuit

Let the rate of an N-bit block be increased by a fractional rate of r. The general configuration of the circuit is shown in Fig. 6. The number j of srs is chosen to satisfy the condition that:

$$\sum_0^{j-1} r_1^i < N - 1 \leq \sum_0^j r_1^i,$$ (6)

where r_1 is an integer just less than r. The first

$$N - 1 - \sum r_1^{j+1}$$

---

* In eq. (2), t corresponds to the duration of a clock cycle at the terminal rate and t' corresponds to the duration at the burst rate. The origin of time is chosen to coincide with the start of the block.
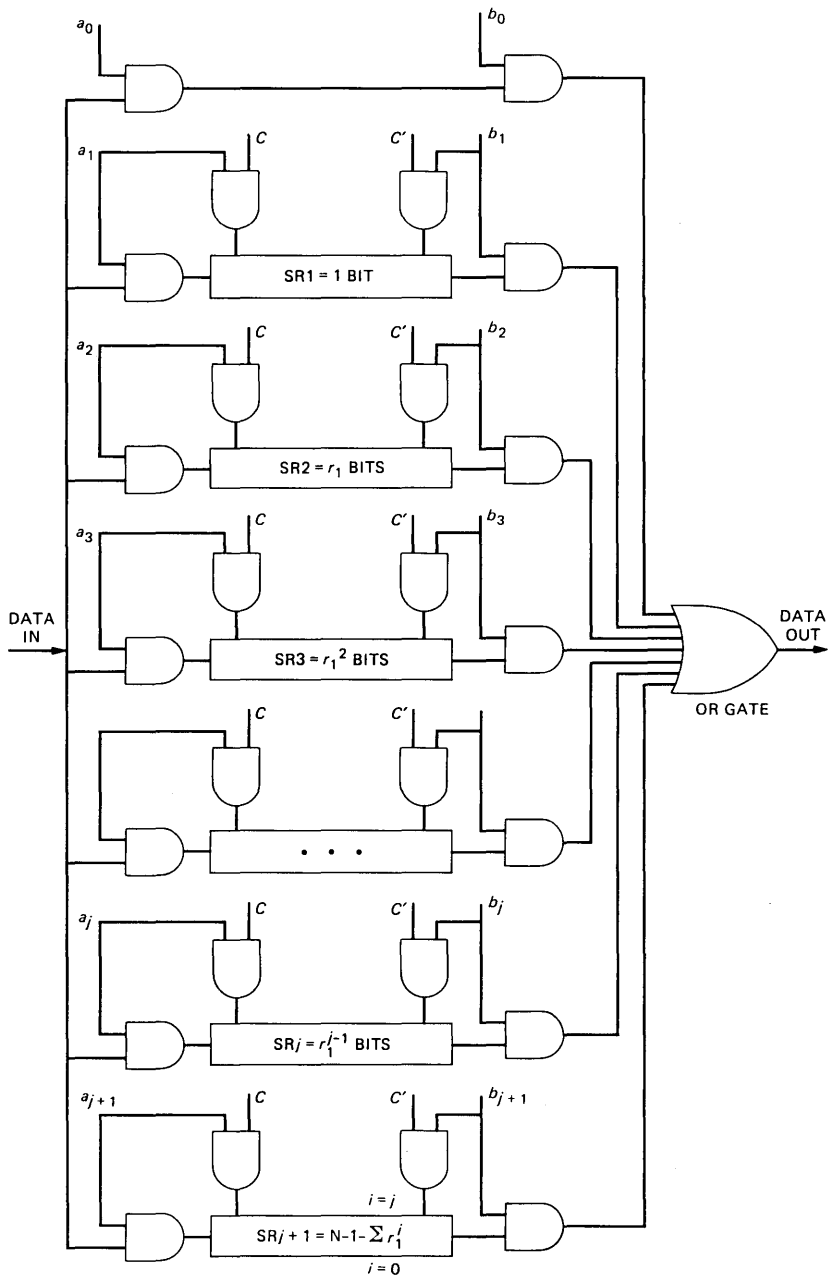
† The fractional change in rate is ≥ 2.

Fig. 6—Generalized fractional rate change circuit.

bits of the block are channeled into SR_{j+1} and the rest r_1^j bits into SR_j, etc. Now consider the kth SR, r_1^{k-1} bits long. The ending instant of SI for this register takes place at:

$$d_i = \left(N - 1 - \sum_{i=0}^{i=k-2} r_1^i \right) t \text{ seconds,} \tag{7}$$

where t denotes the duration at the primary clock and the start of the SO time takes place at an instant:

$$d_o = N(t - t') + \left(N - 1 \sum_{i=0}^{i=k-1} r_1^i \right) t' \text{ seconds.} \tag{8}$$

The duration between the instant at which the SO starts and the instant at which the SI ends is given by

$$d_{ol} = t - t' + \sum_{i=0}^{i=k-2} r_1^i t - \sum_{i=0}^{i=k-1} r_1^i t'$$

$$= \left(1 + \sum_{i=0}^{k-2} r_1^i \right) t - \left(1 + \sum_{i=0}^{i=k-1} r_1^i \right) t'. \tag{9}$$

In the limiting case where $t = r_1 t'$, we have:

$$d_{ol} = r_1 t' - 2t'. \tag{10}$$

The value of d_{ol} becomes zero when $r_1 = 2$ and eq. (10) becomes the same as eq. (4) in its limiting case. However, d_{ol} always remains equal to or greater than zero, thus, validating the basic law of all SRs—that they may not start to SO, unless they are completely shifted in.

APPENDIX C
Delay Requirement on Fractional Rate Increasing Between One and Two

Consider an N-bit block with a rate change ratio of r (<2 and >1). Let r' be the reciprocal of r. Then it can be seen that an arrangement of SRs shown in Fig. 7, and zero delay leads to a situation where the jth register would not have been shifted in before it is ready to be shifted out. To avoid this and to assure that SO takes place immediately after shifting in, the delay necessary can be calculated as

$$d_{in} = [(N - 2^{k+1}) + 2^k - N(1 - r') - (N - 2^{k+1})r']t \text{ seconds}$$

$$= 2^k(2r' - 1)t \text{ seconds.} \tag{11}$$

The first two terms in (11) denotes the time required to fill the last two SRs with $(N - 2^{k+1})$ and 2^k register positions at the primary rate of t seconds per clock cycle. The third term denotes the length of the block at a difference of clock rates between the primary clock rate of t second and a secondary clock of t' ($= r't$) seconds. The fourth term
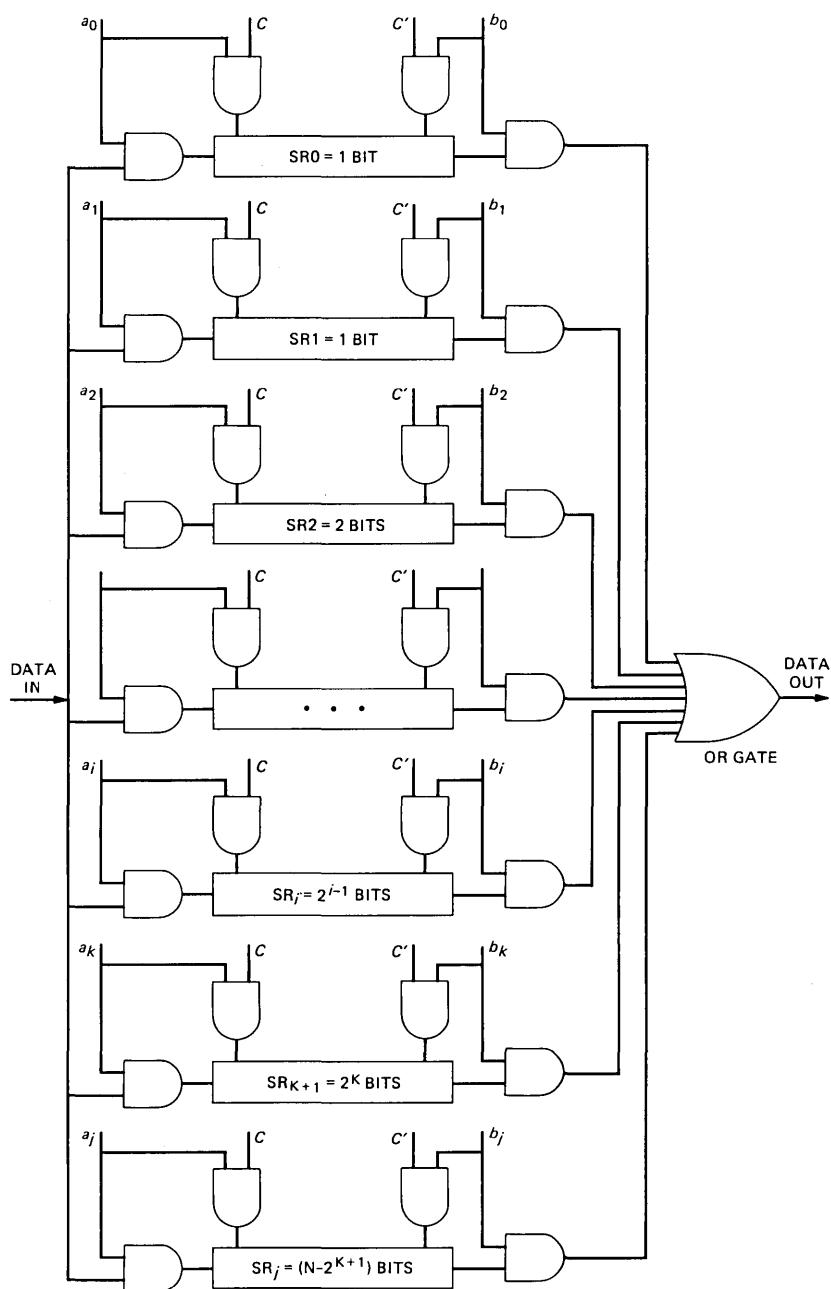
Fig. 7—Generalized fractional rate change circuit for ratios between 1 and 2.

represents the time to empty the last register at the secondary rate. It can be seen that in the limiting case when r' tends to 0.5 the delay necessarily tends to become zero.

The physical significance of this delay is to adjust the coincidence of the finish of the SI instant with the start of the SO of the last but one SR, thus, finishing the binary sequences of register locations. It is also interesting to note that the time delayed in emptying the first SR reappears at the end of the block as a summation of delays in the various other registers before the last register. For instance, the delay between shifting in and shifting out of the $(i + 1)$st; $(0 < i < k)$ SR can be written as:

$$d_{(i)} = \left[r' \sum_{j=i}^{j=k} 2^j - \sum_{j=i-1}^{j=k-1} 2^j \right] t \text{ seconds.} \tag{12}$$

The delay for SR1 becomes

$$d_1 = [r' \times 2(2^k - 1) - (2^k - 1)]t \text{ seconds.} \tag{13}$$

The delay for SR0 causes an incremental increase in delay of $(2r' - 1)t$ seconds. Hence, the total delay of the last bit coming out of SR0 is:

$$d_o = (r' \times 2^{k+1} - 2^k)t \text{ seconds,} \tag{14}$$

which is consistent with eq. (11).

REFERENCE

1. S. V. Ahamed, "A General Class of Rate Change Circuits," B.S.T.J., *50*, No. 10 (December 1971), pp. 2177–93.

Peak Signal-to-Noise Ratio Formulas for Multistage Delta Modulation With RC-Shaped Gaussian Input Signals

By R. STEELE

(Manuscript received July 28, 1981)

A multistage delta modulation (MSDM) encoder contains a number of delta modulation (DM) stages, where each delta modulator encodes the band-limited error of the previous delta modulator. The DM binary outputs are then multiplexed for transmission. By this technique, substantial gains in s/n compared to a single-stage DM can be achieved at high transmitted bit rate to message bandwidth ratios (f_p/f_c). For Gaussian input signals having band-limited resistance-capacitance (RC) spectra, the peak s/n performance of MSDM as a function of (f_p/f_c) and the number of DM stages is presented. It is shown that like PCM, MSDM exchanges s/n with (f_p/f_c) on an exponential basis.

I. INTRODUCTION

Delta modulation (DM) has been extensively studied,[1-8] and following its integration onto a chip,[9] it is being increasingly used in industrial applications. The salient advantages of DM are robustness to transmission errors; tolerance to clock jitter; simple filtering requirements; suitability for encryption; and low complexity resulting in inexpensive implementation. In typical applications, the ratio f_p/f_c is <10, where f_p is the transmitted bit rate and f_c is the bandwidth of the message signal. However, DM does not efficiently improve its s/n with increasing f_p/f_c, particularly when compared to pulse code modulation (PCM). As a consequence, DM is rarely used to encode high-quality audio signals because of the excessive f_p/f_c ratios required.

In DM, the quantization noise is dependent on the error $e(t)$ between the input signal $x(t)$ and a locally reconstructed version $y(t)$ (formed by locally decoding the transmitted bit stream). The $y(t)$ signal essentially tracks $x(t)$, and the polarity of the transmitted bit is identical to

the polarity of the tracking error $e(t)$ at any sampling instant. To increase s/n by increasing f_p, requires $e(t)$ to decrease. It might be supposed that the reduction in $e(t)$ would be enhanced if adaptive delta modulation[1] (ADM) is used rather than linear DM.[1] In ADM, the changes in $y(t)$ per clock period, i.e., the step sizes, are not constant as in linear DM. Various step-size algorithms have been used in ADM, but all occasionally produce inappropriate step-sizes resulting in "over-shoot-noise."[10] In fact, ADM is generally not used to increase peak s/n but, rather, to greatly extend the dynamic range of nonadaptive DM. However, with care the peak s/n can be enhanced at high clock rates, but not by significant amounts.[11]

The basic problem with any form of DM is that the encoder generates information which is only dependent on the polarity of the error. No description of the magnitude of the error is available at the receiver. Das and Chatterjee[12] made a proposal to overcome this defect by conceiving an encoder composed of many DM stages, each encoding the band-limited error signal of the preceding stage. In this way, a more accurate description of the tracking error is available at the receiver, and the exchange of s/n with transmitted bit-rate is greatly enhanced. This method of modulation is multistage delta modulation (MSDM).

Initially it was claimed[12] that MSDM had a better coding efficiency than conventional PCM for the same information rate, but subsequent work[13] using computer simulation up to 3-DM stages showed that although MSDM is better than DM, it does not perform as well as PCM operating with "4σ loading." A theory of MSDM was presented by Franks, Schachter, and Shilling,[14] together with computer simulation of a two-stage MSDM. Chakravarthy and Faruqui[15] constructed a two-stage MSDM using adaptive DM, and refined the expressions for s/n previously propounded.[14]

In spite of these endeavors, there appeared to be a need for more precise formulations of the peak s/n of MSDM. These expressions were found to involve the summation of the peak s/n of numerous DM stages; therefore, it became necessary to develop accurate, yet simple, expressions of s/n for a DM encoder. This was done, and the findings published separately[8] as part of the theory on linear DM. We now use these results for their intended purpose, namely, to provide simple equations for the peak s/n of MSDM when encoding Gaussian input signals with band-limited RC spectra. In the pursuit of this goal, we hope to provide new insight into the behavior of MSDM. Later in Section IV, we compare the s/n performance of MSDM with DM, PCM, and differential pulse code modulation (DPCM). It should be noted that we do not play advocate for MSDM, but merely endeavor to place its s/n performance in perspective. No attempt is made to judge its

relative complexity. We commence with a description of the principle of MSDM.

II. PRINCIPLE OF MSDM

The simplest form of MSDM uses two delta modulators in the encoder as shown in Fig. 1. The analog input signal is band-limited by filter F_1 to give the MSDM input signal $x(t)$. The first delta modulator DM_1 encodes $x(t)$ producing binary information representing the polarity of the tracking error $e_1(t)$. In delta modulation, as distinct from MSDM, the error in the recovered signal is the filtered $e_1(t)$ signal, namely $\tilde{e}_1(t)$. Two-stage MSDM reduces this error by encoding $\tilde{e}_1(t)$ by a second delta modulator DM_2 and, thereby, making it available at the receiver. By this means, the overall error is reduced to the in-band error $\tilde{e}_2(t)$ of DM_2. Observe that in our nomenclature a tilde ($\tilde{\ }$) above a symbol means that it has been low-pass filtered by a filter having a linear phase-frequency characteristic.

Thus, in the two-stage MSDM, the signals $x(t)$ and $\tilde{e}_1(t)$ are encoded by DM_1 and DM_2 to yield binary signals $L_1(t)$ and $L_2(t)$. Typically, both delta modulators will be clocked at the same rate f_s, resulting in a transmitted bit-rate of $f_p = 2f_s$. At the receiver, the $L_1(t)$ and $L_2(t)$ signals are demultiplexed and decoded to give $y_1(t)$ and $y_2(t)$, respectively. A delay D is introduced in the first channel to compensate for the delay resulting from the band-limiting of $e_1(t)$ by filter F_I at the input to DM_2. The delayed decoded signal $z_1(t)$ is added to $y_2(t)$, and the noise residing outside the highest frequency f_c in $x(t)$ is removed by the final filter F_o to yield a recovered signal $m(t)$ that is a close approximation to $x(t)$. We assume that the binary signals $L_1(t)$ and $L_2(t)$ are generated without error.

The scheme is extendable to N delta modulators, each encoding the band-limited error signal of the preceding modulator and operating at a clock rate f_{si}, $i = 1, 2, \cdots, N$. Figure 2 shows an N-stage MSDM system. Signals $x(t)$ and $\tilde{e}_k(t)$, $k = 1, 2, \cdots, N - 1$, are encoded by delta modulators DM_k, $k = 1, 2, \cdots, N$, into binary signals $L_k(t)$, $k = 1, 2, \cdots, N$. The binary signals are multiplexed, and transmitted at a bit-rate f_p whose value is the sum of the DM sampling frequencies. After demultiplexing, each of the $L_k(t)$ signals are decoded into $y_k(t)$, $k = 1, 2, \cdots, N$ and delayed by the networks designated D_k, $k = 1, 2, \cdots, N - 1$, in Fig. 2, with the exception of $y_N(t)$. We will assume that filters F_1, F_I, and F_o are identical linear filters that impose a signal delay of t_o seconds, whereas the delays associated with the networks D_k are integer multiples of t_o, being $(N - 1)t_o$ for the first channel ($k = 1$), and reducing by t_o for subsequent channels. The locally decoded signals in each channel, suitably compensated by the
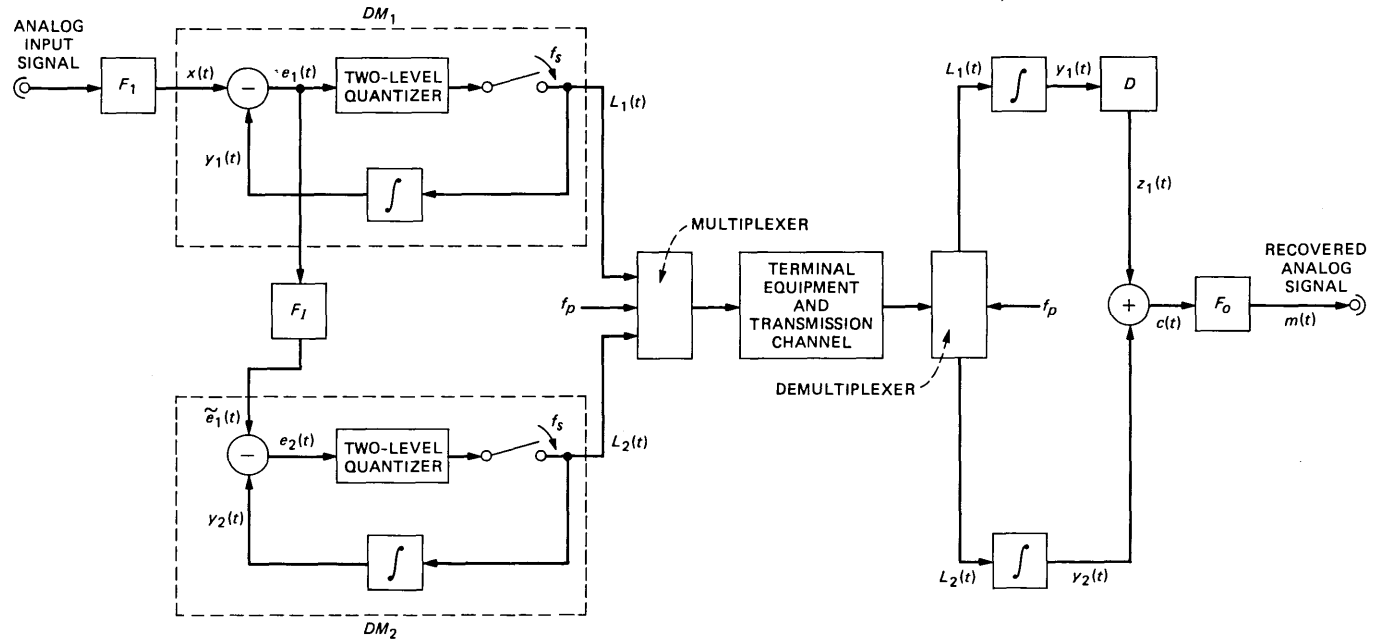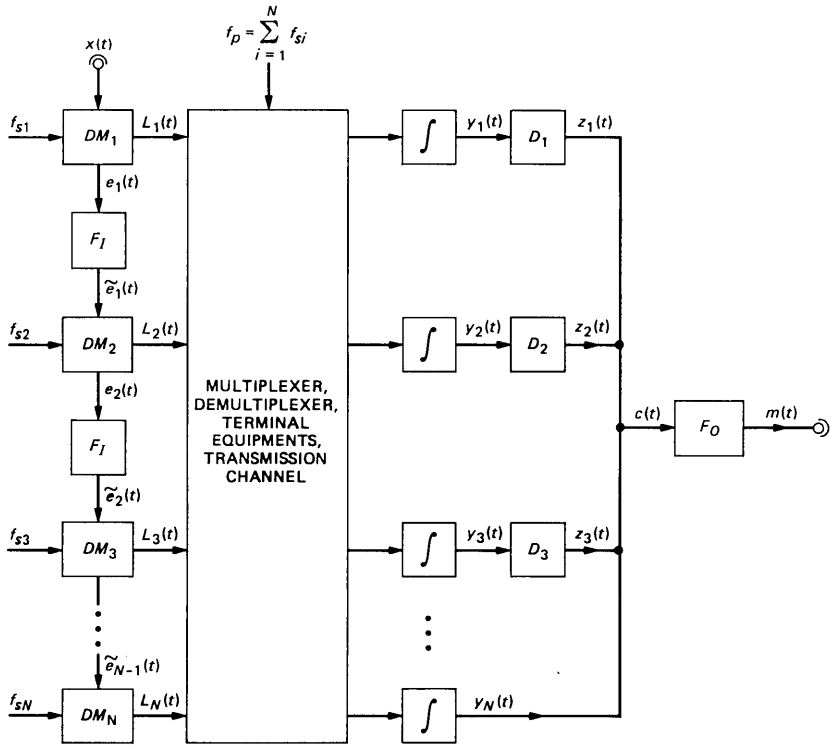
Fig. 1—Two-stage MSDM system.

Fig. 2—N-stage MSDM system.

delays D_k for the delays caused by filters F_I in the encoder, are combined to give

$$c(t) = y_N(t) + \sum_{k=1}^{N-1} z_k(t),$$ (1)

where $z_k(t)$ is the signal at the output of D_k. Now the locally decoded signals at the outputs of the integrators are

$$y_1(t) = x(t) - e_1(t)$$

and

$$y_k(t) = \tilde{e}_{k-1}(t) - e_k(t), \qquad k = 2, 3, \cdots, N - 1$$

and after delay compensation they become

$$z_1(t) = x(t - (N - 1)t_o) - e_1(t - (N - 1)t_o)$$ (2)

and

$$z_k(t) = \tilde{e}_{k-1}(t - (N - k)t_o) - e_k(t - (N - k)t_o),$$

$$k = 2, 3, \cdots, N - 1, \quad (3)$$

respectively. Substituting $z_1(t)$ and $z_k(t)$ from eqs. (2) and (3) into eq. (1) yields

$$c(t) = x[t - (N-1)t_o] - e_N(t)$$
$$- \sum_{k=1}^{N-1} \{e_k[t - (N-k)t_o] - \tilde{e}_k[t - (N-k-1]t_o)\}, \quad (4)$$

and upon filtering $c(t)$ by filter F_o to remove out-of-band noise, the recovered signal is obtained,

$$m(t) = x(t - Nt_o) - \tilde{e}_N(t). \quad (5)$$

This signal $m(t)$ is composed of the input signal $x(t)$ delayed by Nt_o, and a noise component that is the filtered error signal of the Nth stage DM. Observe that the error signals in each DM stage, with the exception of the last stage, cancel out because of the choice of delays D_k in the MSDM decoder channels.

Therefore, the s/n of MSDM is

$$s/n = \frac{\langle x^2(t) \rangle}{\langle \tilde{e}_N^2(t) \rangle}, \quad (6)$$

where $\langle (\cdot) \rangle$ means time averaging of $(\cdot)$. Alternatively, eq. (6) can be expressed as

$$s/n = \frac{\langle x^2(t) \rangle}{\langle \tilde{e}_1(t) \rangle} \frac{\langle \tilde{e}_1(t) \rangle}{\langle \tilde{e}_2(t) \rangle} \cdots \frac{\langle \tilde{e}_{N-1}(t) \rangle}{\langle \tilde{e}_N(t) \rangle}$$
$$= \prod_{i=1}^{N} s/n_i, \quad (7)$$

the product of the s/n's of each DM stage. The s/n in dBs is

$$s/N = \sum_{i=1}^{N} s/N_i, \quad (8)$$

where

$$s/N_i = 10 \log_{10} s/n_i, \quad (9)$$

The upper case s/N in eq. (9) is in dBs, and the lower case s/n is a ratio. Thus, the s/N of an MSDM system is the sum of the s/N's of each DM stage. The next problem is to determine these s/N's in terms of DM parameters.

III. PEAK S/N OF MSDM

We will assume that each DM stage has a step-size which produces peak s/n for that stage. For a Gaussian input signal $x(t)$ band-limited to frequency f_c, the peak s/n for the first DM stage in the MSDM encoder may be expressed as[8]

$$s\hat{/}n_{\text{RC}} = \frac{0.059}{C_{\text{opt,RC}}^2} \left(\frac{\mu}{\beta}\right)\left(\frac{f_{si}}{f_c}\right)^3, \qquad 0.01 \leq \beta \leq 0.5 \tag{10}$$

and

$$s\hat{/}n_F = \frac{0.177}{C_{\text{opt},F}^2}\left(\frac{f_{si}}{f_c}\right)^3, \tag{11}$$

where the subscripts RC and F refer to RC and flat Gaussian input signals, respectively. The flat Gaussian signal is band-limited white noise occupying the frequency band $\pm f_c$, and by RC filtering this signal, the RC Gaussian signal is obtained. The sampling frequency in eqs. (10) and (11) is f_{si}; the break frequency of the RC Gaussian input signal is f_1; and β and μ are given by

$$\beta = \frac{f_1}{f_c} \tag{12}$$

and

$$\mu = \frac{\tan^{-1}(1/\beta)}{1 - \beta\tan^{-1}(1/\beta)}. \tag{13}$$

The optimum slope loading factors are

$$C_{\text{opt,RC}} = 1.3\left[\log_e\left(\frac{f_{si}}{f_c}\right)\right]^{0.72} \tag{14}$$

and

$$C_{\text{opt},F} = 0.5 + 0.722\log_e\left(\frac{f_{si}}{f_c}\right). \tag{15}$$

Consequently, if the first delta modulator stage DM_1 is encoding an RC Gaussian input signal the peak s/n can be expressed from eq. (10) as

$$s\hat{/}n_{\text{RC}} = \theta_{\text{RC}}f_{s1}^3, \tag{16}$$

where θ_{RC} is a DM parameter whose value is evident from eq. (10), and f_{s1} is the clock frequency for DM_1. Observe that θ_{RC} has a relatively weak dependence on f_{s1} [see eq. (14)] and this dependency will be considered later in the calculation of the peak s/n of MSDM.

Irrespective of the spectrum of $x_1(t)$, the signals applied to subsequent DM stages will be assumed to be flat Gaussian signals. This is a reasonable assumption as each DM, excluding the first, encodes the band-limited error signal from the preceding stage. To substantiate this assumption, we make the following points. Granular noise is known[8] to dominate slope overload noise when a delta modulator operates at its maximum s/n. Computer simulation results for granular noise when the s/n is close to its peak value have shown[7] that the

noise spectrum is flat over the message bandwidth. Models have also been proposed[16] for DM that have flat noise spectra when the encoder is optimally loaded. Let us now consider the probability density function (PDF) of the filtered error signal $\tilde{e}_1(t)$. Although the unfiltered DM error signal $e(t)$ is uniformly distributed over the range of twice the DM step-size,[8] the act of filtering $e(t)$ by a filter F_I of bandwidth f_c is to produce a signal $\tilde{e}(t)$ whose PDF becomes increasingly Gaussian as the ratio of f_{si}/f_c is increased. For the values of f_{si}/f_c considered here, the Gaussian PDF assumption of the $\tilde{e}(t)$ signal is reasonable, and this, coupled with its flat spectral properties, supports our assertion that $\tilde{e}(t)$ is a flat Gaussian signal. Finally, we note that even if the PDF of $\tilde{e}(t)$ were not Gaussian, we would still be justified in treating $\tilde{e}(t)$ as a flat Gaussian signal. This is because we only use the power properties of the signal in our calculations. The PDF of $\tilde{e}(t)$ need not be considered because of the DM noise being predominately granular.

From eq. (11), we will express the peak s/n of a delta modulator encoding a flat Gaussian input signal as

$$\widehat{\text{s/n}}_{F,j} = \theta_F f_{sj}^3, \qquad j = 2, 3, \cdots, N, \tag{17}$$

where θ_F is a DM parameter having a weak dependence on f_{sj} [see eq. (15)]. The value of j in eq. (17) denotes the DM stage number and f_{sj} is the sampling frequency for the jth stage.

The s/n in dB will be written in upper case letters. Thus, from eq. (8) and the preceding discussion, we have for the N stage MSDM,

$$\text{S/N} = \text{S/N}_{\text{RC}} + \sum_{j=2}^{N} \text{S/N}_{F,j} \tag{18}$$

and from eqs. (16), (17), and (18)

$$\text{S/N} = 10 \log_{10} \theta_{\text{RC}} + (N - 1)10 \log_{10} \theta_F + \sum_{i=1}^{N} 30 \log_{10} f_{si}$$

$$= 10 \log_{10} \theta_{\text{RC}} + (N - 1)10 \log_{10} \theta_F + 30 \log_{10} \lambda, \tag{19}$$

where

$$\lambda = \prod_{i=1}^{N} f_{si}. \tag{20}$$

The transmitted bit rate is

$$f_p = \sum_{i=1}^{N} f_{si}. \tag{21}$$

If θ_{RC} and θ_F were independent of the DM sampling rates f_{si}, $i = 1, 2, \cdots, N$, then the S/N of the MSDM would be maximized by maximizing λ, subject to the constraint that f_p in eq. (21) is constant. This would occur when

$$f_{s1} = f_{s2} = \cdots = f_{sN}. \tag{22}$$

However, θ_{RC} and θ_F depend on $C_{opt,RC}$ and $C_{opt,F}$, and although these slope loading factors are functions of the sampling frequency f_{si}, the variation of θ_{RC} and θ_F with f_{si} has only a minor effect on S/N. Thus, with each DM operating at the same sampling frequency, the S/N for the MSDM is close to its peak value. To each DM stage we will, therefore, assign the clock frequency

$$f_s = \frac{f_p}{N}, \tag{23}$$

enabling the S/N of eq. (19) to have a peak value of

$$\widehat{s/N} = \widehat{s/N}_{RC} + (N-1)\widehat{s/N}_F. \tag{24}$$

The $\widehat{s/N}_{RC}$ and $\widehat{s/N}_F$ terms are the values of $\widehat{s/n}_{RC}$ and $\widehat{s/n}_F$ in dBs, where f_{si} in eqs. (10) and (11) is replaced by f_p/N. Observe that $\widehat{s/N}_F$ is independent of the RC Gaussian input signal $x(t)$ applied to the MSDM encoder, provided that DM_1 tracks it to produce a flat error signal spectrum. Thus, the $\widehat{s/N}$ of MSDM is calculable from eqs. (10) and (11).

Substituting the result of eq. (23) into eq. (19) enables the $\widehat{s/N}$ to be expressed in terms of parameters θ_{RC} and θ_F, the transmitted bit rate f_p, and the number of stages N, namely,

$$\begin{aligned}
\widehat{s/N} &= 10 \log_{10}\theta_{RC} + (N-1)10 \log_{10}\theta_F \\
&\quad + 30 \log_{10}\left(\frac{f_p}{N}\right)^N \\
&= 10 \log_{10}\theta_{RC} + (N-1)10 \log_{10}\theta_F \\
&\quad + 30N \log_{10} f_p - 30N \log_{10} N.
\end{aligned} \tag{25}$$

Chakravarthy and Faruqui[15] assumed the $\widehat{s/N}$ to be N times the $\widehat{s/N}$ for each DM stage. In their presentation, θ was not explicitly derived; instead it was a system parameter. From their measurements, they concluded that the assumed $\widehat{s/N}$ was too high and, accordingly, it was reduced by a factor $H(N-1)$, where H is an empirical constant, namely,

$$\widehat{s/N} = 10N \log_{10}\left\{ \theta_{RC}\left(\frac{f_p}{N}\right)^3 \right\} - H(N-1). \tag{26}$$

The exponent 3 is a system parameter α in their formula as they considered delta modulators having local decoders composed of either single or double integration stages. From eqs. (25) and (26) we can specify their value of H for the MSDM using linear DM stages as

$$H = 10 \log_{10}\left(\frac{\theta_{RC}}{\theta_F}\right) = 10 \log_{10}\left(\frac{\widehat{s/N}_{RC}}{\widehat{s/N}_F}\right), \tag{27}$$

or from eqs. (10) and (11),

$$H = 10 \log_{10}\left[\frac{1}{3}\left(\frac{\mu}{\beta}\right)\left(\frac{C_{\mathrm{opt},F}}{C_{\mathrm{opt},\mathrm{RC}}}\right)^2 \right].\qquad(28)$$

As μ/β is independent of f_{si}/f_c, and $C_{\mathrm{opt},F}/C_{\mathrm{opt},\mathrm{RC}}$ is nearly independent[8] of f_{si}/f_c, we can appreciate why H was introduced as a constant.

Returning to eq. (25) and replacing θ_{RC} and θ_F by the DM parameters, the peak S/N of MSDM can be expressed in terms of the normalized transmitted bit rate f_p/f_c, namely,

$$\widehat{\mathrm{S/N}} = -4.77 - 7.52N - 20\log_{10}C_{\mathrm{opt},\mathrm{RC}} - 20(N-1)\log_{10}C_{\mathrm{opt},F}$$

$$+ 10\log_{10}\left(\frac{\mu}{\beta}\right) + 30N\log_{10}\left(\frac{f_p}{f_c}\right) - 30N\log_{10}N,\qquad(29)$$

where $C_{\mathrm{opt},\mathrm{RC}}$ and $C_{\mathrm{opt},F}$, given by eqs. (14) and (15) have sampling frequencies f_{si} equal to f_p/N. The variation of the peak S/N of MSDM as expressed by eq. (29), is presented in Fig. 3 as a function of f_p/f_c for different values of N. The value of β used is 0.235, corresponding to frequency parameters f_1 and f_c of 800 and 3,400 Hz, respectively. This value of β is often used[3,8] when RC Gaussian signals are employed as
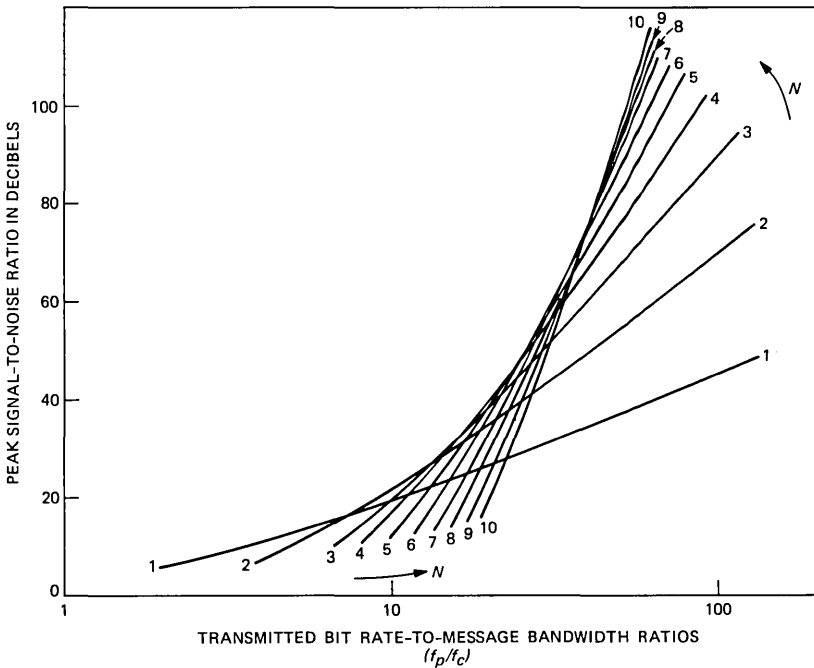

Fig. 3—The S/N of MSDM as a function of f_p/f_c for N having values 1 to 10, and $\beta = 0.235$.

an approximation to band-limited speech signals in telephony. For monochrome luminance signals, $\beta = 0.01$ may be used,[3,8] and although this smaller value of β increases the S/N of DM_1, the shape of the curves in Fig. 3 are essentially the same. Thus, it is f_p/f_c and $\hat{N}$ that are the cardinal parameters governing the S/N of MSDM.

3.1 Selecting N for maximum S/N

From Fig. 3, we observe that for any f_p/f_c there is a particular number of DM stages to maximize the S/N. When $f_p/f_c < 7.5$, only one stage should be used (linear DM), and as f_p/f_c is increased the number of stages N should be increased appropriately. The maximum value of the normalized frequency associated with a given N for the maximum s/̂N is, for $\beta = 0.235$,

$$\left(\frac{f_p}{f_c}\right)_{max} = \frac{N + 0.4}{0.187}; \qquad 1 \le N \le 4. \tag{30}$$

For example, $N = 3$ gives the maximum s/̂N over the frequency range $12.8 < f_p/f_c < 18.2$. The range of f_p/f_c associated with a value of N to give maximum s/̂N becomes progressively smaller as N is increased. Practical MSDM systems are unlikely to be produced with $N > 4$. Thus, given that N can be varied to maximize the s/̂N for any f_p/f_c, we have from the curves of Fig. 3,

$$s/̂N_{max} = 5 + 1.7\left(\frac{f_p}{f_c}\right). \tag{31}$$

IV. COMPARING S/N OF MSDM WITH DM, PCM, AND DPCM

4.1 Multistage delta modulation (MSDM)

The s/̂N of MSDM, given by Eq. (29), increases at a rate of approximately $7N$ dB/octave increase in f_p/f_c, for $32 < f_p/f_c < 128$. At lower values of f_p/f_c, the variation of s/̂N with f_p/f_c is approximately $5.3N$ dB/octave. Selecting N to peak the s/̂N for any f_p/f_c, yields the s/̂N$_{max}$ of eq. (31).

4.2 Delta modulation (DM)

The s/̂N of linear DM is found by putting $N = 1$ in eq. (29),

$$s/̂N_{DM} = -12.3 + 10 \log_{10}\left(\frac{\mu}{\beta}\right) - 20 \log_{10} C_{RC} + 30 \log_{10}\left(\frac{f_p}{f_c}\right), \tag{32}$$

i.e., we may view DM as a special case of MSDM. In DM, the rate of improvement of s/̂N with f_p/f_c varies from 5 dB/octave for $2 < f_p/f_c < 5$, to 9 dB/octave for $32 < f_p/f_c < 128$. This rate of improvement is significantly less than in MSDM [see eq. (31)].

4.3 Pulse code modulation (PCM)

The input signal is sampled at the Nyquist rate of $2f_c$, and the transmitted bit rate f_p is $2f_c n$, where n is the number of bits in the code words. For "4σ loading," the $\hat{s/N}$ in dB of linear PCM is[2]

$$\hat{s/N}_{PCM} = -7.3 + 6n = -7.3 + 3\left(\frac{f_p}{f_c}\right). \tag{33}$$

We observe that PCM is more efficient at exchanging $\hat{s/N}$ with f_p/f_c, or n, compared to both DM and MSDM. However, at low values of f_p/f_c, PCM has a lower $\hat{s/N}$ than DM.

4.4 Differential pulse code modulation (DPCM)

The $\hat{s/N}$ in dB of linear DPCM is

$$\hat{s/N}_{DPCM} = -A + 6n + G_p = -A + 3\left(\frac{f_p}{f_c}\right) + G_p, \tag{34}$$

where A is a constant, and G_p is the prediction gain factor. The constant A depends on the probability density function of the error signal and can often be taken as $\simeq 7$ dB, i.e.,[17]

$$\hat{s/N}_{DPCM} = \hat{s/N}_{PCM} + G_p. \tag{35}$$

The prediction gain factor is

$$G_p = 10 \log_{10}(\text{SFM}), \tag{36}$$

and SFM is the ratio of the arithmetic mean to the geometric mean of the discrete spectrum of the input signal.[17] For a first-order optimum predictor

$$G_p = -10 \log_{10}(1 - \rho^2), \tag{37}$$

where ρ is the correlation between adjacent samples of the input signal. The autocovariance function of an RC Gaussian input signal band-limited to f_c has been found by O'Neal,[2] enabling us to formulate an approximate expression for ρ as

$$\rho = \left[\exp(-2\pi f_1/f_s) \right.$$

$$\left. - 4\frac{f_1}{f_s}\left\{\frac{\cos(2\pi f_c/f_s)}{(2\pi f_c/f_s)} + \text{Si}(2\pi f_c/f_s) - \frac{\pi}{2}\right\}\right] \Big/ 1 - (2/\pi)\beta, \tag{38}$$

where β is defined by eq. (12), f_s is the sampling rate, and Si is the sine integral function. When the input signal is sampled at the Nyquist rate, i.e., $f_s = 2f_c$, eq. (38) reduces to

$$\rho = \frac{e^{-\pi\beta} + 0.0751\beta}{1 - (2/\pi)\beta}. \tag{39}$$

For a value of β of 0.235, $\rho = 0.54$, resulting in $G_p = 1.12$ dB.

Thus, linear DPCM has an advantage G_p over linear PCM, but it has the same type of dependence on f_p/f_c. Therefore, we conclude that MSDM, where the number of stages are selected according to eq. (30), has an $\widehat{\text{S/N}}$ of the same form as PCM and DPCM, namely,

$$\widehat{\text{S/N}} = P_1 + P_2\left(\frac{f_p}{f_c}\right), \tag{40}$$

where P_1 and P_2 are parameters that depend on the type of modulation, and for MSDM, DM, and DPCM, on the input signal. The MSDM operating with the optimum number of stages to peak the S/N is, therefore, more efficient at exchanging $\widehat{\text{S/N}}$ with f_p/f_c than DM [see eq. (32)].

The variation of peak $\widehat{\text{S/N}}$ as a function of f_p/f_c for an RC Gaussian input signal having $\beta = 0.235$ is shown in Fig. 4 for DM; MSDM with $N = 4$; MSDM employing the optimum number of stages; and PCM with "4σ loading" and sampling at the Nyquist rate. The curve of DPCM for $\beta = 0.235$ and Nyquist sampling is not displayed as it has the same shape as that of PCM, but with $\widehat{\text{S/N}}$ increased by G_p.
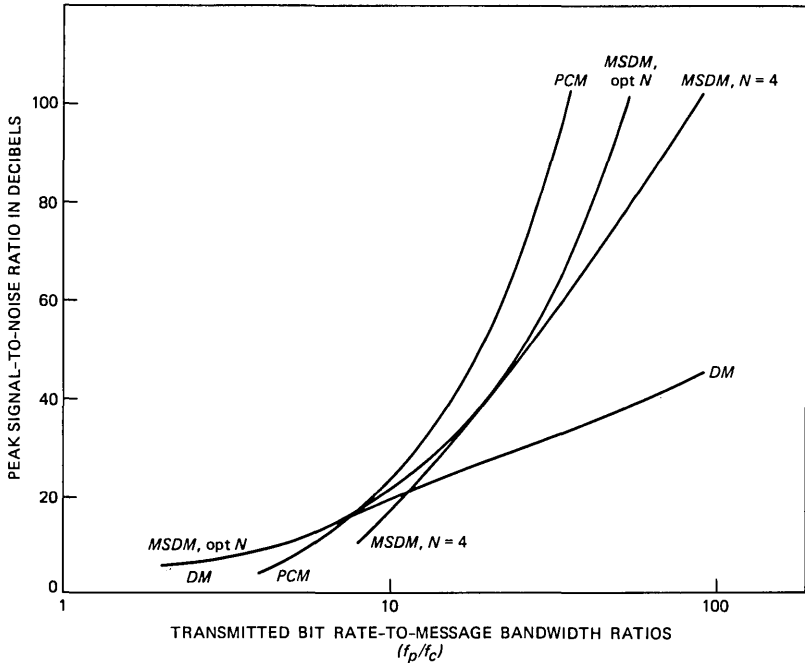

Fig. 4—The $\widehat{\text{S/N}}$ as a function of f_p/f_c for linear DM, $\beta = 0.235$; MSDM, $N = 4$, $\beta = 0.235$; MSDM with optimum N, $\beta = 0.235$; and PCM, $f_p = 2f_c$, "4σ loading."

V. THE MSDM USING ADM STAGES

An MSDM codec is required to operate over a wide range of input levels. In Section III, we considered the delta modulators to be linear with their step sizes adjusted to maximize their S/N values. When the input power of such an MSDM encoder is reduced causing DM_1 to operate in the granular region with a slope loading factor $C_{RC} < 8$, the noise generated in DM_1 is substantially independent of the input signal power.[1] As the filtered error signal of DM_1 is the input to DM_2, and the power of this signal is approximately unaltered by the reduction of the input power to DM_1, the S/N of DM_2 remains the same. Because the power level to DM_2 is unchanged, the S/N's of subsequent DM stages are therefore similar, particularly for $N \leq 4$. As the input power to DM_1 is further reduced, the noise generated in DM_1 will eventually increase,[3] causing some overloading in DM_2 which, in turn, will reduce the S/N's in subsequent DM stages. Thus, if excessive granular noise is generated in DM_1, slope-overloading in the remaining delta modulators ensues. However, this slope-overloading of the second and subsequent DM stages for a reduction in input level to DM_1 operating in the granular region, is far less severe than when the input power is increased, causing DM_1 to become slope-overloaded. When DM_1 is slope-overloaded, all the stages in the MSDM experience slope-overload, and the S/N of the MSDM rapidly deteriorates with increasing input power.

The range of input power over which the MSDM can operate, while providing an acceptable S/N, i.e., the dynamic range, can be greatly enhanced by using adaptive delta modulators (ADMs) instead of linear delta modulators. The diversity of ADMs is considerable,[1] but all have the property of extending the dynamic range, while retaining an approximately constant S/N. The quantization noise power in ADM is, therefore, proportional to the signal power. As the input signal power varies, the filtered error signal applied to DM_2 varies, and if this encoder is not to be overloaded, it must also be adaptive. The same argument applies to subsequent stages, and hence the complete MSDM codec is composed of ADM stages.

If the N ADM encoders in the MSDM codec are syllabically companded delta modulators[1] [currently available on a chip[9] in the form of continuously variable slope delta (CVSD) codecs], and they use single integrators in the local decoding process, then the peak S/N of each ADM stage is a close approximation to that given by eqs. (10) or (11). Further, the dynamic range of each codec where the S/N is maintained near its peak S/N is wide,[15] typically 40 dB. Thus, the maximum S/N of MSDM given by eq. (29) is a good approximation for MSDM having ADM stages over a wide range of input power.

VI. DISCUSSION

Easily evaluated equations of peak S/N for MSDM have been derived in terms of normalized bit rate, number of DM stages, and the shape of the spectrum of the RC Gaussian input signal. A suitable choice for the sampling rate for each delta modulator was found to be f_p/N. For a given input signal bandwidth f_c, and transmission bit rate f_p, there is an optimum number of stages which maximizes S/N. The variation of this $\widehat{S/N}$ with f_p/f_c is given by eq. (29), and we showed in Section IV that this has the same form as for PCM and DPCM. Thus, MSDM is more efficient than DM at exchanging $\widehat{S/N}$ for f_p/f_c, but the $\widehat{S/N}$ of MSDM is generally lower than that of PCM and DPCM. At very low values of f_p/f_c, DM performs better than the other modulation methods considered (see Fig. 4). In Section V, we discussed MSDM with CVSD stages, concluding that the $\widehat{S/N}$ of eq. (29) remains valid, provided that single-stage integrators are used in the CVSD codecs. Further gains in $\widehat{S/N}$ are attainable if double stage integrators[1,15] replace the single-stage integrators. The CVSD codecs enable the MSDM to have a wide dynamic range and, therefore, we envisage MSDM codecs to be constructed[15] with adaptive, rather than linear, delta modulators.

Although this work has been concerned with the derivation of $\widehat{S/N}$ of MSDM, we conclude with the observation that MSDM having four CVSD codecs might have a role to play in a variable bit-rate transmission system. For example, each CVSD stage could operate at 16 kb/s giving a transmitted bit rate of 64 kb/s when the four stages are in use. In a time division multiplex (TDM) system, the MSDM codecs would attempt to operate at 64 kb/s, but as traffic increased they could discard the higher order CVSD stages, decreasing their bit rates from 64 to 48 to 32 kb/s, until when the system is at maximum capacity, each MSDM would behave as a 16-kb/s CVSD codec. By using MSDM instead of a single-stage DM operating at the same bit rate, we are able to enhance the quality of the recovered signal as the bit rate increases from 16 kb/s.

VII. ACKNOWLEDGMENTS

REFERENCES

1. R. Steele, "Delta Modulation Systems," London: Pentech Press, New York: Halstead Press, 1975.
2. J. B. O'Neal, "Delta Modulation Quantizing Noise Analytical and Computer Simulation Results for Gaussian and Television Signals," B.S.T.J., *45*, No. 1 (January 1966), pp. 117–41.

3. J. E. Abate, "Linear and Adaptive Delta Modulation," Proc. IEEE, *55*, No. 3 (March 1967), pp. 298–308.
4. D. J. Goodman, "Delta Modulation Granular Noise," B.S.T.J., *48*, No. 5 (May–June 1969), pp. 1197–218.
5. D. Slepian, "On Delta Modulation," B.S.T.J., *51*, No. 10 (December 1972), pp. 2101–37.
6. L. J. Greenstein, "Slope Overhead Noise in Linear Delta Modulators with Gaussian Inputs," B.S.T.J., *52*, No. 3 (March 1973), pp. 387–421.
7. M. Passot and R. Steele, "Application of the Normal Spectrum Technique to the Calculation of Distortion Noise in Delta Modulators," The Radio and Elect. Eng., *44* (October 1974), pp. 545–52.
8. R. Steele, "SNR Formula for Linear Delta Modulation with Bandlimited Flat and RC Shaped Gaussian Signals," IEEE Trans. Commun., *COM-28*, No. 12 (December 1980), pp. 1977–84.
9. R. Steele, "Chip Delta Modulators Revive Designers' Interests," Elect. (October 13, 1977), pp. 86–93.
10. F. T. Sakane and R. Steele, "Estimation of Signal-to-Noise Ratio in High Information and Constant Factor Delta Modulation Systems," IEEE Trans. Comms., *COM-25*, No. 12 (December 1977), pp. 1441–8.
11. R. Steele and S. D. Cridge, "Upper Bound of Signal-to-Noise Ratio for Instantaneously Companded Delta Modulation Systems," Elect. Lett., *11*, No. 7 (April 3, 1975), pp. 145–7.
12. J. Das and P. K. Chatterjee, "Optimized Δ—Δ Modulation System," Elect. Lett., *3*, No. 6 (June 1967), pp. 286–7.
13. P. K. Chatterjee and V. Rama Rao, "Digital Computer Simulation Results of Multistage Delta Modulation Systems," Proc. IEE, *120*, No. 11 (November 1973), pp. 1379–82.
14. J. Frank, H. Schachter, and D. Shilling, "*N*-th Order Delta Modulation, an Improved Differential Coding Technique," IEEE Symp. Digest, 1970 Canadian Symp. Commun., Montreal, November 1970.
15. C. V. Chakravarthy and M. N. Faruqui, "A Multidigit Adaptive Delta Modulation (ADM) System," IEEE Trans. Comms. (August 1976), pp. 931–5.
16. N. S. Jayant, "A First-Order Markov Model for Understanding Delta Modulation Noise Spectra," IEEE Trans. on Commun., *COM-26*, No. 8 (August 1978), pp. 1316–18.
17. J. L. Flanagan et al., "Speech Coding," IEEE Trans. Commun., *COM-27*, No. 4 (April 1979), pp. 710–37.

Note on Some Factors Affecting Performance of Dynamic Time Warping Algorithms for Isolated Word Recognition

By L. R. RABINER

To implement a dynamic time warping (DTW) algorithm for isolated word recognition, a number of factors must be specified. These factors include endpoint constraints, local path constraints, global path constraints, axis orientation, and local distance measure. Although a number of studies have been made to decide how best to choose these factors, several unresolved issues remain. The problems with choosing these factors become more complicated if the word reference patterns are not created from whole words, but instead are built up from subword units, e.g., syllables or demisyllables. In this paper, we consider an isolated word recognizer in which the reference patterns are obtained from a set of demisyllables as specified in a user-supplied lexicon. Different pronunciations of a word are represented by multiple entries in the lexicon. Because of the inherent boundaries in the reference patterns (at each demisyllable junction), it was felt that local constraints could influence performance more than for the standard whole-word case. Further, it was felt that some flexibility in the endpoint path constraint at the end of the word would be helpful, in general, for isolated word recognition caused by high variability at the ending of words with stops. A DTW algorithm was programmed and tested on an 1109-word vocabulary. Results showed small accuracy improvements when the path endpoint was allowed to vary across four frames (of either test or reference pattern). Loosening the local path constraints, however, had a significant degrading effect on the performance.

I. INTRODUCTION

The technique of using dynamic time warping (DTW) to align (in time) a test and reference pattern for isolated word recognition has

been shown to be effective in a wide variety of recognition systems.[1-6] Although a great deal of investigation has been made into "optimal" DTW algorithms, there still remains uncertainty as to how best to specify the factors of the DTW implementation to achieve the highest recognition accuracy.

In this paper, we consider two of these factors, namely, the endpoint constraint at the end of the warp and the local path constraints. Previous investigations were made by Rabiner et al.[4] and Myers et al.[6] on the effects on word recognition accuracy of both loosening endpoint constraints and using different local path constraints. However, the work[4] on DTW algorithms with relaxed endpoint constraints considered only two specific variations, namely, the unconstrained endpoint case (the UE2-1 algorithm), and the local minimum case (the UELM algorithm). Neither of these algorithms considered relaxing just the endpoint constraint at the end of the word. This constraint is very important since replications of isolated words generally have the most variability at the end of the word. Similarly, although various local path constraints were considered in Ref. 6, the use of subword units in the reference patterns raises again the question as to whether increased flexibility in the choice of warping path leads to improvements in recognition scores. We show that by loosening the endpoint constraint at the end of the utterance, a small but consistent increase in recognition accuracy is obtained. It is also shown that the Itakura local path constraints yield higher recognition accuracies than generalized Itakura local constraints (Type III, Ref. 6). This result corroborates previous findings[4,6] which show that, in general, opening up of the region for DTW matching of words generally leads to degraded performance, since the improved score for the correct matches is offset by the improved scores for incorrect matches.

The outline of this paper is as follows. In Section II, we briefly review the DTW implementation, and describe the factors which were studied. In Section III, we present and discuss results of a recognition test using a 1109-word vocabulary, and in Section IV, we summarize the findings.

II. THE DTW IMPLEMENTATION

Assume that we are given a test pattern **T**, consisting of a sequence of N vectors, i.e.,

$$\mathbf{T} = \{T(1), T(2), \cdots, T(N)\}, \tag{1}$$

where the vector $T(i)$ is a spectral representation of the ith frame of the test word. The vectors, $T(i)$, are a set of 9 autocorrelations (from which an 8th order linear predictive coding model is derived). The

duration of the test word is N frames, where each frame represents 45 ms of speech, and adjacent frames are spaced 15 ms apart.

For a given vocabulary of V words, we denote the reference pattern for the vth word as $\mathbf{R}_v$, and we represent each reference pattern as a sequence of M_v vectors, i.e.,

$$\mathbf{R}_v = \{R_v(1), R_v(2), \cdots, R_v(M_v)\}, \qquad (2)$$

where each vector is again a spectral representation of the corresponding frame within the word.

To optimally align the time scales of the test (the n index) and reference (the m index) patterns via a DTW algorithm, we must solve for a warping, or path alignment function of the form

$$m = w(n) \qquad (3)$$

and thereby seek to minimize the average distance

$$D = \frac{1}{N} \sum_{n=1}^{N} \tilde{d}\{T(n), R[w(n)]\} \qquad (4)$$

over all possible $w(n)$, where $\tilde{d}$ is the local distance between test frame n and reference frame $m = w(n)$. To solve the DTW problem requires specification of the following:[6]

    (*i*) endpoint constraints
    (*ii*) local path constraints
    (*iii*) global path constraints
    (*iv*) axis orientation
    (*v*) local distance measure.

In this paper, we have considered the effects, on recognition accuracy, of two of the DTW factors—the local endpoint constraints, and the local path constraints. In particular, we have considered the following DTW specifications:

Endpoint Constraints—We assume the word endpoints of the test and reference patterns satisfy the path constraints

$$w(1) = 1 \qquad (5a)$$

$$M - \delta_R \leq w(n) \leq M, \qquad N - \delta_T \leq n \leq N, \qquad (5b)$$

where δ_R is a range of frames, at the end of the reference pattern, and δ_T is a range for frames, at the end of the test pattern, where the warp path can terminate. Thus, the warp path begins at the point (1, 1) and ends in a region of δ_R frames from the end of the reference, and δ_T frames from the end of the test.

Local Path Constraints—We assume that the local path satisfies one of the two sets of constraints shown in Fig. 1, namely, the Itakura[2]

path constraints (Fig. 1a) and the Myers et al.[6] Type III path constraints (Fig. 1b). Both of these local path constraints satisfy the continuity equations

$$0 \le w(n) - w(n-1) \le 2 \tag{6a}$$

$$w(n) - w(n-1) = 0 \Rightarrow w(n-1) - w(n-2) > 0. \tag{6b}$$

The difference between the Itakura and the Type III path constraints is a subtle, but an important one, and it is related to the look-ahead (or look-back) capability of the local path constraints in finding the best path to a given grid point. This difference is illustrated in Fig. 2 which shows three possible local paths to grid point (n, m). Path 1 comes initially from grid point $(n-2, m-1)$ and goes through grid point $(n-1, m)$ before terminating at grid point (n, m). Path 2 goes from grid point $(n-1, m-1)$ to grid point (n, m). Path 3 goes from grid point $(n-1, m-2)$ to grid point (n, m). For the Itakura constraints (with 1 level of look-ahead logic), if the best path to the intermediate grid point $(n-1, m)$ came from grid point $(n-2, m)$— that is, it came across—then the possibility of path 1 being the best path to grid point (n, m) is not considered, whereas for Type III constraints (with two levels of look-ahead logic) path 1 is considered along with paths 2 and 3. At the end of a path (and often within the path), these differences can be significant.

Global Path Constraints—The endpoint and local path equations constrain the range of frames of the reference pattern (the m-axis) for which the basic DTW recursion is done to:

$$M_L(n) \le m \le M_H(n), \tag{7}$$

where

$$M_H(n) = \min\{2(n-1) + 1, M - (N - \delta_T - n)/2, M\} \tag{8a}$$

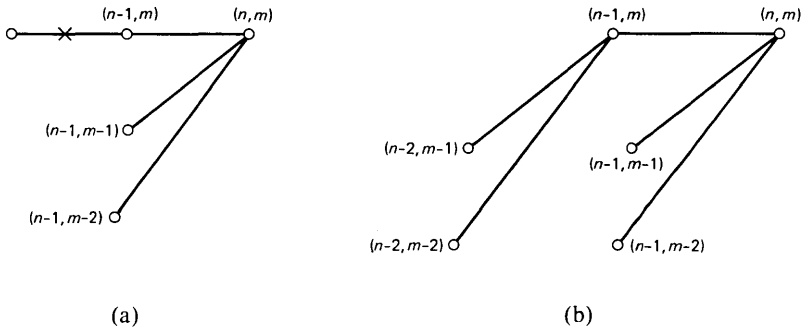(a)   (b)

Fig. 1—Two sets of DTW local path constraints. (a) Itakura constraints. (b) Type III constraints.
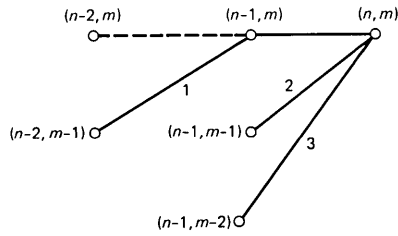
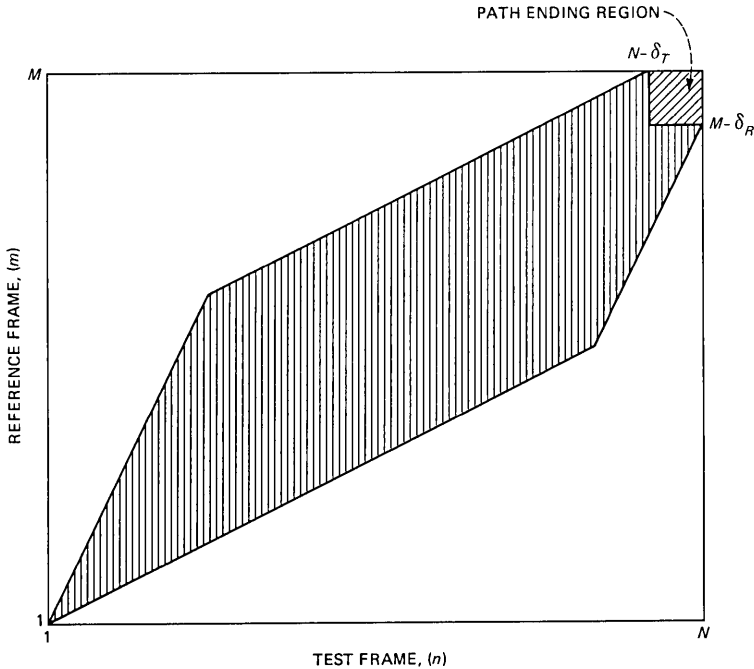Fig. 2—Possible local paths for Itakura and for Type III local constraints.

Fig. 3—The region in the warping plane in which the dynamic time warping path can lie with ending uncertainty of δ_T frames along the test and δ_R frames along the reference.

$$M_L(n) = \max\{(n-1)/2 + 1, M - 2(N - n) - \delta_R, 1\}. \qquad (8b)$$

These global path constraints essentially define a parallelogram (as shown in Fig. 3) with lines of slope 2 and ½ emanating from the points $m = 1, n = 1$, and from $m = M - \delta_R, n = N - \delta_T$.

Axis Orientation—We assume that the test sequence index n is always mapped to the abscissa and the reference sequence index m is always mapped to the ordinate. Experience has shown that this orientation leads to the best performance in isolated word recognition systems.[4,6]

Local Distance Measure—The local distance measure used in this study is the Itakura log likelihood ratio[2] which is implemented in the form

$$\tilde{d}[T(n), R(m)] = \log[T(n) \cdot R(m)], \tag{9}$$

i.e., a log of the dot product of the two vectors $T(n)$ and $R(m)$.

2.1 Word recognition algorithm

Word recognition is achieved by computing the optimal warping path and distance for each word in the vocabulary, giving

$$D_v = \min_{w(n)} \min_{N-\delta_T \leq \hat{N} \leq N} \left[\frac{1}{\hat{N}} \sum_{n=1}^{\hat{N}} \tilde{d}\{T(n), R_v[w(n)]\} \right] \tag{10}$$

and using the nearest neighbor rule to choose v^* as the best candidate where

$$v^* = \operatorname*{argmin}_{v}[D_v]. \tag{11}$$

An ordered set of word distances is also maintained for statistical analysis purposes.

2.2 Creation of reference patterns

The reference patterns, R_v, for the vth vocabulary word was created from a lexicon which contained one or more demisyllable-based specifications for each word.[7] The entries in the lexicon were obtained by converting a standard dictionary word pronunciation to a set of demisyllable tokens which could be concatenated to create the word. Variant word pronunciations were represented by multiple lexical entries.

A set of 955-demisyllable tokens were used as the set of basis units from which each word was created. An 1109-word vocabulary (Basic English[8]) was used as the recognition vocabulary and a total of 1773 lexical entries were used for the vocabulary.

In creating the individual reference patterns from the demisyllable prototypes, a distinct boundary is created between each pair of demisyllables. To minimize boundary effects, a nonlinear smoother was used to interpolate (in a minimum mean-squared error sense) the vocal tract log area ratios (i.e., the ratio between the areas of adjacent sections of a nine-section vocal tract model for each frame of speech) in a four-frame vicinity of the boundary. Because of the presence of one or more boundaries within each reference pattern, it was felt that the local path constraints could potentially have more effect on the recognition results than for the case of reference patterns created from isolated words.

2.3 Summary of DTW factors that were studied

As discussed in this section, there are two DTW factors that were studied in the context of isolated word recognition from reference patterns created out of a corpus of demisyllables. These factors were as follows:

(i) *The range at the end of the test pattern,* δ_T—Two values of δ_T were chosen, namely, $\delta_T = 0$ (the standard case) and $\delta_T = 4$ frames (60 ms), corresponding to the duration of a release of a burst, etc.

(ii) *The range at the end of the reference pattern,* δ_R—For reasons similar to those in the choice of values for δ_T, values of 0 and 4 were studied for δ_R.

(iii) *The local path constraint*—Both the Itakura and the Type III constraints were studied in conjunction with variations of δ_R and δ_T. Only one simple experimental evaluation was made, and the results are given in Section III.

III. EXPERIMENTAL EVALUATION

To study the effects of the two factors discussed at the end of Section III on the recognition accuracy, an experiment was performed in the following way. The system was run as a speaker-trained, isolated word recognizer with a testing vocabulary of 1109 words (i.e., one replication of each word of the vocabulary) and a reference set of 1773 patterns obtained from the lexical description of the 1109 word-vocabulary. The DTW algorithm (with the flexibility of varying the DTW factors) was programmed in FORTRAN on a Data General Eclipse S230 computer. (Normally, a CSP MAP 200 array processor is used to process the data at high speed; however, the flexible version of the DTW algorithm was not available on the MAP.) Each DTW alignment took on the order of 1 second on the Eclipse; hence, a complete test on a *single* set of DTW factors would have taken on the order of 5000 hours of computing. Clearly, this amount of computation is prohibitive. Thus, an alternative testing procedure was derived in which a "standard" DTW was run on the CSP MAP (where $\delta_R = \delta_T = 0$ and the Itakura local constraints were applied). The complete matrix of distances (1109 words × 1773 references) was scanned and for each test word, the most highly confused alternative word was found (i.e., the word reference different from the test token with the smallest distance). A second "word recognition" test was now performed on the Eclipse in which, for each test word, only two DTW distances (with the variable factors included) were measured. The DTW distances were, thus, obtained for a reference corresponding to the spoken word, and for a reference corresponding to the vocabulary word most similar to the spoken word. In this manner, the eight combinations of δ_R, δ_T, and

local constraints were varied and their effects on word recognition accuracy were measured.

Before proceeding on to the results, some justification of this non-standard testing methodology must be given. It should be clear that for test words which were correctly recognized, and which had one or fewer close alternative words, such a procedure is acceptable for studying the effects of DTW parameters. It can also be seen that for cases in which the DTW factors degrade performance using the two candidate algorithms, performance with all candidates will be degraded even further. Hence, the only cases in which we cannot place total reliance on the results are those words with several close candidates for recognition. Hopefully, the number of such cases is small and will not greatly affect the results presented.

3.1 Experimental results

For each set of DTW factors (i.e., specification of δ_R, δ_T, and LC—local constraints) a set of three quantities were measured, namely:

(i) *Overall recognition accuracy*—This quantity assumes that use of the two references best matching the test word was adequate when the DTW factors were varied. To the extent that this is the case, this measurement is the performance index of most concern.

(ii) *Average distance separation*—This quantity measures the average difference of the distances from the test to the reference pattern of the spoken word, and to the reference pattern of the closest competitor. It should be noted that the average distance separation was computed over *all* vocabulary words (including errors); hence, for errors the distance separation is negative and for correct recognitions it is positive. The average distance separation is one measure of separation of the distances for the correct words, and distances for nearest competitors.

(iii) *Number of occurrences of maximum distance separation over the eight factors*—This quantity counted the number of times the particular set of DTW factors provided the best distance separation (over the eight sets of factors) for a given word, if it was recognized correctly. Hence, this measurement is a marginal count (conditioned on correct recognition) of the superiority (in terms of distance separation) of one set of factors over the seven alternative sets.

The results of the word recognition experiment, in terms of the above three measurements, are given in Table I. The code $LC = 0$ is used to denote the Itakura local constraints; similarly, $LC = 1$ is used to denote the Type III local constraints. The following results can be seen from Table I:

(i) The overall recognition accuracy using $LC = 0$ is from 2 to 3 percent higher than when using $LC = 1$ for all choices of δ_R and δ_T.

Table I—Performance results for DTW factors

DTW Factors			Percent of Overall Recognition Accuracy	Average Distance Separation	Number of Occurrences of Maximum Distance Separation
LC	δ_R	δ_T			
0	0	0	76.4	0.044	356
0	4	0	76.4	0.043	355
0	0	0	76.9	0.045	483
0	4	4	77.2	0.044	445
1	0	0	74.4	0.039	291
1	4	0	73.7	0.039	294
1	0	4	74.8	0.040	377
1	4	4	74.2	0.039	343

(ii) The variation in recognition accuracy for a fixed value of LC (and varying δ_R and δ_T) is small.

(iii) The "best" recognition accuracy scores are for $LC = 0$, $\delta_R = 0$ or 4, and $\delta_T = 4$.

(iv) The average distance separations for $LC = 0$ are about 10 percent higher than for $LC = 1$.

(v) The best performances in terms of occurrences of maximum distance separation are for $LC = 0$, $\delta_T = 4$, and $\delta_R = 4$ and 0. Similarly for $LC = 1$, the cases $\delta_T = 4$ and $\delta_R = 0$ and 4 give the best performance in terms of this measurement.

3.2 Discussion

The results presented in Section 3.1 lead to the following two conclusions:

(i) The Itakura local constraints ($LC = 0$) yield better performance than the more general Type III local constraints for *all* choices of δ_R and δ_T, and for all three performance measurements.

(ii) The selection of $\delta_T = 4$ (with $\delta_R = 0$ or 4) is marginally better, in terms of recognition accuracy and average distance separation, than the selection of $\delta_T = 0$. However, in terms of maximum distance separation, the selection of $\delta_T = 4$ is significantly better than the selection of $\delta_T = 0$.

The first conclusion is expected in the sense that a great deal of earlier research has shown that any broadening of path regions invariably degrades performance for DTW algorithms,[4,6] since the improvement in score for the correct reference is generally more than offset by the improvement in score for the nearest incorrect reference. Although it was anticipated that the special construction of the reference patterns (i.e., from concatenated demisyllable units) might mitigate this general result, the data of Table I shows that this is not the case. It is especially notable that even for the case of two carefully chosen reference patterns in the recognizer, this general conclusion is still

valid. This result tends to lend some credence to the somewhat unusual testing methodology.

The second conclusion, although anticipated, is somewhat more difficult to explain. One expectation was that opening up the ending region (via nonzero values of δ_R and for δ_T) would lead to the same effect noted in conclusion one, namely, degraded performance. This was definitely not the case here. A second expectation was that if opening up the ending region was a good thing to do, *both* δ_R and δ_T would have to be nonzero. Although nonzero values of both δ_R and δ_T led to the best performance, it is seen that $\delta_R = 0$ was an almost equally good choice. The reason for this is the assymetry in the DTW implementation in which every single test frame is used in the warping path, but any given reference can be omitted entirely (since the path *can* jump by 2 frames). As such, at the end of the path, 2 of the 4 ending reference frames could be skipped with $\delta_R = 0$. Thus, the importance of a nonzero value of δ_R is greatly lessened, whereas a nonzero value of δ_T clearly leads to improved performance.

Additional analyses of the recognition results were performed to see if any simple correlations existed between cases in which improved recognitions occurred and specific linguistic events of the ends of the words, e.g., stops, nasals, fricatives, etc. No consistent and meaningful correlations were found. Hence, we conclude that the results of Table I are probably independent of the vocabulary items and the fact that the vocabulary was created from demisyllable tokens.

A final comment concerns the recognition accuracy achieved for the 1109-word vocabulary. Using whole word speaker-trained templates, Rabiner et al.[9] achieved a recognition accuracy of 79.2 percent across 6 talkers for the same 1109-word vocabulary. Thus, the best accuracy achieved using demisyllable-based templates (77.2 percent) compares favorably with the accuracy from whole word templates.

IV. SUMMARY

We have examined briefly the effects on isolated word recognition of three factors in the DTW alignment procedure. It was found that the Itakura local path constraints provided the best performance and that relaxing the ending constraint on the dynamic path also provided small, but consistent, improvements in performance.

REFERENCES

1. H. Sakoe and S. Chiba, "Dynamic Programming Algorithm Optimization for Spoken Word Recognition," IEEE Trans. Acoustics, Speech, and Signal Processing, *ASSP-26*, No. 6 (December 1978), pp. 575–82.
2. F. Itakura, "Minimum Prediction Residual Principle Applied to Speech Recognition," IEEE Trans. Acoustics, Speech, and Signal Processing, *ASSP-23*, No. 1 (February 1975), pp. 67–72.

3. G. M. White and R. B. Neely, "Speech Recognition Experiments with Linear Prediction, Bandpass Filtering, and Dynamic Programming," IEEE Trans. Acoustics, Speech, and Signal Processing, *ASSP-24,* No. 2 (April 1976), pp. 183–8.
4. L. R. Rabiner, A. E. Rosenberg, and S. E. Levinson. "Considerations in Dynamic Time Warping for Discrete Word Recognition," IEEE Trans. Acoustics, Speech, and Signal Processing, *ASSP-26,* No. 6 (December 1978), pp. 575–82.
5. H. F. Silverman and N. R. Dixon, "State Constrained Dynamic Programming (SCDP) for Discrete Utterance Recognition," Proc. Int. Conf. ASSP (April 1980), pp. 169–72.
6. C. S. Myers, L. R. Rabiner, and A. E. Rosenberg, "Performance Tradeoffs in Dynamic Time Warping Algorithms for Isolated Word Recognition," IEEE Trans. Acoustics, Speech, and Signal Processing, *ASSP-28,* No. 6 (December 1980), pp. 622–33.
7. A. E. Rosenberg, et al., "A Preliminary Study on the Use of Demisyllables in Automatic Speech Recognition," Proc. 1981 Int. Conf. Acoustics, Speech, and Signal Processing (March 1981), pp. 967–70.
8. C. K. Ogden, *Basic English: International Second Language*, New York: Harcourt, Brace, and World Inc., 1968.
9. L. R. Rabiner et al., unpublished work.

LETTER TO THE EDITOR

Comments on "Fractionally Spaced Equalization: An Improved Digital Transversal Equalizer," by R. D. Gitlin and S. R. Weinstein*

(Communication received November 1981)

In the Gitlin and Weinstein paper, Fig. 3 and Formula (15) give the following result: the equalized signal, called $\tilde{u}(t)$, is analytic, i.e., its real and imaginary parts are a Hilbert transform pair. This result is contradictory with the fact that in QAM, $\text{Re}\{\tilde{u}(t)\}$ contains in-phase data a_n, and $\text{Im}\{\tilde{u}(t)\}$ contains independent quadrature data b_n. It is worthwhile to make the three following points:

(i) $\tilde{q}(t)$ in Formula (13), is not analytic since ω_c is within the band of $s(t)$. Therefore, $\tilde{q}(t)$ cannot be written as $q(t) + j\check{q}(t)$ as it was done on page 281. Consequently, $\tilde{u}(t)$ is complex but not analytic and its real and imaginary parts are not Hilbert transform. $\tilde{q}(t)$ and $\tilde{u}(t)$ are analytic if ω_c is on the edges of the band of $s(t)$ or outside this band as is the case in single sideband systems.

(ii) In Formula (9), $\tilde{s}(t)$ is analytic if and only if $P(\omega) = 0$; $|\omega| \geq \omega_c$.

(iii) On page 280, it should be noted that $\tilde{f}_B(t)$ should be given by the convolution of $\frac{1}{2}\tilde{x}_B(t)e^{-j\theta}$ with $p(t)$, where $\theta = \angle X(\omega_c)$. The factor $\frac{1}{2}$ was indeed mentioned in the Appendix, Formula (62).

G. Kawas-Kaleh
Ecole Nationale Supérieure de Télécommunication
Department Systemes et Communications
46 Rue Barrault, 75634 Paris, CEDEX 13

Author Response

Mr. Kawas-Kaleh is correct in pointing out that $\tilde{q}(t)$ is not an analytic signal. Since $\tilde{r}(t) = r(t) + j\check{r}(t)$ is an analytic signal centered about w_c, the demodulated signal $\tilde{q}(t)$ will be a complex signal having frequency components centered about the origin, and hence cannot be an analytical signal.

The condition that $P(w) = 0$ when $|w| \geq w_c$, which is implied though not explicitly stated in the paper, is standard practice.

It should be noted that Mr. Kawas-Kaleh's comments concerning analytic signals in no way affect the performance attributes of the fractionally spaced equalizer reported in the paper.

R. D. Gitlin

CONTRIBUTORS TO THIS ISSUE

Syed V. Ahamed, B.E., 1957, University of Mysore; M.E., 1958, Indian Institute of Science; Ph.D., 1962, University of Manchester, U.K.; Post-Doctoral Research Fellow, 1963, University of Delaware; Assistant Professor, 1964, University of Colorado; M.B.A. (Economics), 1978, New York University; Bell Laboratories, 1966–1981. At Bell Laboratories, Mr. Ahamed has worked in computer-aided engineering analysis and design of electromagnetic components, designed and implemented minicomputer software and hardware interfacing, applied algebraic analysis to the design of domain circuits, and investigated computer aids to the design of bubble circuits. He has investigated new varacter designs for microwave power in the C-band and developed hardware and software interfacing for audio frequency codecs. Beginning in 1975, he optimized codec designs, encoding techniques, and speech-encoded data storage and manipulation by minicomputers. Since 1977, he has been simulating the overall performance of the loop plant, while carrying bidirectional digital data in the range of 56 to 324 kbaud. He is currently a visiting professor in the Electrical Engineering Department at the University of Hawaii at Monoa, Hawaii and is on a leave of absence from Bell Laboratories.

Ruth Chang, B.A. (Physics/Astronomy), 1981, Wellesley College; Pennsylvania Power and Light Company, 1981—. Miss Chang was a participant in the Bell Laboratories Summer 1980 Research Program for Minorities and Women. Under this program, she assisted with the video colorization diagnostics in optical telecommunications project. Currently, she is a Mathematician at Pennsylvania Power and Light Company.

Albert A. Fredericks, B.S. (Mathematics), 1962, Fairleigh Dickinson University; M.S., 1965, Ph.D. (Mathematics) 1970, Courant Institute, New York University; Bell Laboratories, 1961—. Mr. Fredericks is supervisor of the Performance Analysis Methods group. His responsibilities include the development of methods for analyzing the performance of processor-based systems, including auxiliary microprocessors, mini-computer-based operations systems, data networks and stored program control switching system. Member IEEE, MAA, SIAM.

Dan L. Philen, B.S. (Chemistry), 1968, Auburn University; Ph.D. (Physical Chemistry), 1975, Texas A&M University; Georgia Institute of Technology, 1976–1979; Bell Laboratories, 1979—. Since joining Bell Laboratories, Mr. Philen has been engaged in exploratory measure-

ments on optical fiber properties. Member, American Chemical Society, Optical Society of America, Sigma Xi, Sigma Pi Sigma, Phi Lambda Upsilon.

Herman M. Presby, B.A., 1962, and Ph.D., 1966, Yeshiva University; Research Scientist, Columbia University, 1966–1968; Assistant Professor Physics, Belfer Graduate School of Science, Yeshiva University, 1968–1972; Bell Laboratories, 1972—. Mr. Presby is engaged in studies on the properties of optical fiber waveguides.

Lawrence R. Rabiner, S.B. and S.M., 1964, Ph.D. (Electrical Engineering), 1967, Massachusetts Institute of Technology; Bell Laboratories, 1962—. From 1962 through 1964, Mr. Rabiner participated in the cooperative plan in electrical engineering at Bell Laboratories. He worked on digital circuitry, military communications problems, and problems in binaural hearing. Presently, he is engaged in research on speech communications and digital signal processing techniques. He is coauthor of *Theory and Application of Digital Signal Processing* (Prentice-Hall, 1975) and *Digital Processing of Speech Signals* (Prentice-Hall, 1978). Member, IEEE ASSP Technical Committee on Digital Signal Processing, IEEE ASSP Proceedings Editorial Board, Acoustical Society Executive Council, Eta Kappa Nu, Sigma Xi, Tau Beta Pi. Fellow, Acoustical Society of America, IEEE.

Raymond Steele, B.S., 1959, Ph.D. (Electrical Engineering), 1975, Durham University, Durham, England; E. K. Cole, Ltd., 1959–1961; Cossor Radar, Electronics, Ltd., 1961–1962; The Marconi Company, 1962–1965; Royal Naval College, 1965–1968; Loughborough University of Technology, 1968–1979; Bell Laboratories, 1979—. At E. K. Cole, Ltd., Cossor Radar, Electronics, Ltd., and The Marconi Company, all located in Essex, England, Mr. Steele was engaged in research and development. As a member of the Lecturing Staff at the Royal Naval College in London, he lectured on telecommunications for the NATO and external London University degree courses. At the Loughborough University of Technology in Loughborough, England, he directed a research group in digital encoding of speech and television signals, in addition to serving as Senior Lecturer. Before joining Bell Laboratories on a full-time basis, Mr. Steele served as a part-time consultant to the Laboratories' Acoustics Research Department. Presently, he is a member of the Communications Methods Research Department. He is the author of *Delta Modulation Systems* published in 1975.

PAPERS BY BELL LABORATORIES AUTHORS

COMPUTING/MATHEMATICS

Representations of Multivariate Normal Distributions with Special Correlation Structures. F. B. Six, Commun in Stat, *A10,* No. 13 (August 1981), pp 1285–96.

Summation of Second-Order Recurrence Terms and Their Squares. D. L. Russell, The Fibonacci Quarterly, *19,* No. 4 (October 1981), pp 336–40.

On the VAX/VMX Time-Critical Process Scheduling. Y. T. Wang, Performance Evaluation Review, *10,* No. 3 (Fall 1981), pp 11–18.

ENGINEERING

CCIS Network Performance and New Capabilities. R. F. Frerking, N. A. Marlow, M. A. McGrew, and J. M. Sebeson, Proc Int Switching Symp, *2,* (September 21–25, 1981), No. 24B2.

Color Center Laser Measurements of O-H Diffusion In Optical Fiber Cores. D. L. Philen, Tech Digest of Third Int Conf Integrated Optics and Optical Fiber Commun, Paper TUG1 (April 27–29, 1981), pp 62–3.

Earthquakes and Telecommunications Lifelines. J. W. Foss, *Earthquakes and Earthquake Engineering—Eastern United States,* Vols 1 and 2, edited by J. E. Beavers, Ann Arbor, Michigan: Ann Arbor Science Publ Inc, *1,* (1981), pp 475–94.

The Effects of Backdriving Digital Integrated Circuits During In-Circuit Testing. L. J. Sobotka, Proc 1981 Bell Syst Conf on Elec Testing (September 22–24, 1981), pp 114–34.

Evolution of Fault Recovery Techniques. N. O. Whittington, Proc Nat Elec Conf, *35* (October 25, 1981), pp 48–9.

An Improved Single Mode Fiber Design Exhibiting Low Loss, High Bandwidth, and Tight Mode Confinement, Simultaneously. P. D. Lazay, A. D. Pearson, P. J. Lemaire, and W. A. Reeds, Proc Conf Lasers and Electro-optics—Postdeadline Papers, (June 10, 1981), pp WG6-1-3.

Large-Area Back-Illuminated InGaAs/InP Photodiodes for Use at 1 to 1.6 μm Wavelength. C. A. Burrus, A. G. Dentai, and T. P. Lee, Optics Commun, *38,* No. 2 (July 15, 1981), pp 124–6.

Modulated Barrier Photodiode: A New Majority-Carrier Photodetector. C. Y. Chen, A. Y. Cho, P. A. Garbinski, C. G. Bethea, and B. F. Levine, Appl Phys Lett, *39,* No. 4 (August 15, 1981), pp 340–2.

Phasing Multitone Signals to Minimize Peak Factors. L. J. Greenstein and P. J. Fitzgerald, IEEE Trans Commun, *COM-29,* No. 7 (July 1981), pp 1072–4.

Required Transmit Filter Bandwidths in Digital Radio Systems. L. J. Greenstein and D. Vitello, IEEE Trans Commun, *COM-29,* No. 9 (September 1981), pp 1405–8.

A Single Chip CMOS PCM Codec With Filters. D. G. Marsh, B. K. Ahuja, T. Misawa, M. R. Dwarakanath, P. E. Fleischer, and V. R. Saari, IEEE J Solid State Circuits, *SC-16,* No. 4 (August 1981), pp 308–15.

The Structure of Quantization Noise from Sigman-Delta Modulation. J. C. Candy and O. J. Benjamine, IEEE Trans Commun, *COM 29,* No. 9 (September 1981), pp 1316–23.

Technology Transfer: A Case History Construction of a Multijoule Laser at Bell Labs Shows How Technology Developed for Government Programs can be Transferred to Other Organizations. R. H. Storz, Laser Focus, *17,* No. 9 (September 1981), pp 65–6.

A Voiceband Codec With Digital Filtering. J. C. Candy, B. A. Wooley, and O. J. Benjamine, IEEE Trans Commun, *COM-29,* No. 6 (June 1981), pp 815–30.

PHYSICAL SCIENCES

The Absence of Photochromic Switching in Semiconducting Films of Thioindigo. S. C. Dahlberg and C. B. Reinganum, J Chem Phys, 75 (September 1981), pp 2429-31.

Active Mode Locking of Double Heterostructure Lasers in an External Cavity. J. P. van der Ziel, J Appl Phys, 52, No. 7 (July 1981), pp 4435-46.

An AES Study of the Oxidation of a 97Pb-3Sn Alloy. R. P. Frankenthal and D. J. Siconolfi, Corrosion Sci, 21, No. 7 (July 1981), pp 479-86.

Alloyed Tin-Gold Ohmic Contacts on n-type InP. P. A. Barnes and R. S. Williams, Solid State Elec, 24, No. 10 (1981), pp 907-13.

CO Emission from Seyfert Galaxies. J. H. Beiging, L. Blitz, C. J. Lada, and A. A. Stark, Astrophys J, 247, No. 2 (July 15, 1981), pp 443-8.

CO in M31. A. A. Stark, R. A. Linke, and M. A. Frerking, Bull Astron Soc, 13, No. 2 (1981), pp 535.

Compensation in Ge-doped p-Type Ga$_{1-x}$Al$_x$As Grown by Liquid-Phase Epitaxy. V. Swaminathan, P. J. Anthony, J. L. Zilko, M. D. Sturge, and N. E. Schumaker, J Appl Phys, 52, No. 9 (1981), pp 5603-7.

Covercoat Retardation of Permeation Through Sheet Molding Compound. T. E. Graedel and J. P. Franey, J Appl Polymer Sci, 26, No. 11 (1981), pp 3933-8.

Direct Measurement of Scattering Losses in Single-Mode and Multimode Optical Fibers. D. L. Philen and F. T. Stone, *Phys Fiber Opt, Vol. 2: Advances in Ceramics,* edited by B. Bendow and S. Mitra, Columbus, Ohio: Amer Ceramic Soc (July 1981), pp 237-46.

Electron Impact Ionization Cross Section of Metastable N$_2$(A$^3\sum_u^t$). P. B. Armentrout, S. M. Tarr, A. Dori, and R. S. Freund, J Chem Phys, 75, No. 6 (September 15, 1981), pp 2786-94.

Extensional Flow of Linear and Star-Branched Hydrogenated Polybutadiene with Narrow Molecular Weight Distribution. V. R. Raju, V. S. Au-Yeung, and C. W. Macosko, J Rheology, 25, No. 4 (August 1981), pp 445-52.

Ferroelectric Behavior and Phase Transition at 715K in SrAlF$_5$. S. C. Abrahams, J. Ravez, A. Simon, and J. P. Chaminade, J Appl Phys, 52, No. 7 (July 1981), pp 4740-3.

Free-to-Bound Transitions in S-doped Epitaxial Ga$_{1-x}$Al$_x$As. V. Swaminathan, M. D. Sturge, and J. L. Silko, J Appl Phys, 52, No. 10 (October 1981), pp 6306-11.

Ga$_a$O$_3$ – G$_e$O$_2$ – As$_2$O$_5$ Glasses. W. H. Grodkiewicz, H. M. O'Bryan, L. Pressman, S. Singh, L. G. VanWitert, and G. Zydzik, J Non-Crystalline Solids, 44, No. 2-3 (June 1981), pp 405-8.

Galvanic Corrosion of Metals Coupled to Carbon-Black Filled Polyethylene. G. Schick, Proc Int Symp Honoring Prof. H. H. Uhlig on his 75th Birthday—Corrosion and Corrosion Protection, Corrosion Div Elec Soc, 81-8 (August 1981), pp 259-65.

Internal Structure of Phase-Slip Centers in Superconducting Microstrips. W. J. Skocpol and L. D. Jackel, Physica B, 108B-C, No. 2 (August-September 1981), pp 1021-2.

On the Oxidation of Tin-Lead Alloys. R. P. Frankenthal and D. J. Siconolfi, *Corrosion and Corrosion Protection,* edited by R. P. Frankenthal and F. Mansfeld, Pennington, N. J.: The Electrochemical Soc Inc, pp 276-84.

Pulsating Output of Separate Confinement Buried Optical Guide Lasers Due to the Deliverate Introduction of Saturable Loss. J. P. van der Ziel, W. T. Tsang, R. A. Logan, and W. M. Augustyniak, Appl Phys Lett, 39, No. 5 (September 1981), pp 376-8.

Radiative Recombination in Crystalline As$_2$S$_3$. K. Murayama and M. A. Bosch, Phys Rev B, 23, No. 12 (June 15, 1981), pp 21-3.

Raman Study of Evaporated and Sputtered Ge$_x$Se$_{1-x}$ Glass Films. J. E. Griffiths and W. R. Sinclair, Appl Phys Lett, 39, No. 7 (October 1, 1981), pp 551-3.

The Replication, Generation, and Reduction of Dislocations During Iso-Epitaxy on InP Substrates. S. Mahajan, V. G. Keramidas, A. K. Chin, S. N. G. Chu, W. A. Bonner, and D. D. Manchon, Jr., Inst Phys Conf, Ser No. 59, Chapter 8 (September 8-12, 1980), pp 413-8.

Specific Heat of the Concentrated Spin Glasses Mn_3Ge_2 and MnGe.   J. J. Hauser and F. S. L. Hsu, Phys Rev, *24,* No. 3 (August 1, 1981), pp 1550–1.

Velocity Field Characteristics of Minority Carriers (Electrons) in p-$In_{0.53}Ga_{0.47}As$.   J. Degani, R. F. Leheny, R. E. Nahory, and J. P. Heritage, Appl Phys Lett, *39,* No. 7 (October 1, 1981), pp 569–72.

SOCIAL AND LIFE SCIENCES

What Does Ghiselin Mean by "Individual"?   J. B. Kruskal, The Behavioral and Brain Sciences, *4,* No. 2 (June 1981), pp 294–5.

CONTENTS, APRIL 1982

No. 10A REMOTE SWITCHING SYSTEM

B.S.T.J. BRIEF

Experimental Verification of Ultra-Wide Bandwidth Spectra in Double-Clad Single-Mode Fiber

By S. J. JANG,* L. G. COHEN, W. L. MAMMEL, and M. A. SAIFI*

The low-loss and low-dispersion properties of single-mode fibers make them obvious choices for wide bandwidth system applications with very long repeater spans. This brief describes the fabrication procedure and transmission properties of a double-clad single-mode fiber which is capable of wide bandwidth (greater than 100 GHz-km for laser sources with 4-nm emission bandwidths) transmission over the widest wavelength range (1.45 μm to 1.73 μm) thus far reported in the literature. This range completely covers the lowest-loss wavelength window for fused-silica optical fibers. Double-clad lightguides[1-4] were formed by using an inner cladding to form an index well between the core and a pure silica outer cladding. A computer-aided analytical procedure was used to choose the proper fiber diameter so that waveguide dispersion effects could be used to cancel material dispersion at predetermined wavelengths.[5]

The modified-chemical-vapor-deposition (MCVD) technique was used for preform fabrication. Sixty outer cladding layers, a composition of GeO_2-P_2O_5-SiO_2 and fluorine, having the same refractive-index as pure silica were deposited by the MCVD method inside a 16- by 20-mm fused silica tube. The composition of six inner cladding layers is fluorine-doped silica in order to maintain a 0.35 percent negative index difference relative to the outer cladding. Then six core layers of GeO_2-SiO_2

---

* Western Electric, Engineering Research Center, Princeton, N.J.

composition having a 0.55 percent positive index difference relative to the outer cladding were deposited. The refractive-index profile and diameter of the preform were measured by the laser beam refraction technique.[6] Figure 1 shows the refractive-index profile of a double-clad lightguide preform. The usual dopant burn-off from the center of the core during the high temperature collapsing step caused profile deviations from an intended step-index shape. The fiber was drawn from the preform at 2100°C using an RF-induction zirconia furnace and coated by a 50-μm-thick UV-curable acrylate resin (EA-II) immediately after exiting from the furnace.

The preform index profile data in Fig. 1 were used in a computer analysis program[5] to predict dispersion and bandwidth spectra for single-mode fibers. The optimal fiber diameter, 2a = 11 μm, was calculated to achieve the highest possible transmission bandwidths over the widest possible band of wavelengths within the lowest loss window for fused silica optical fibers. Calculated results in Fig. 2 show how dispersion effects caused by the structure of the double-clad waveguide (curve . . .) can be apportioned to cancel dispersion effects caused by germania and fluorine-doped fused silica materials (curve - - - -). Results obtained by summing points on those two curves predict the total chromatic dispersion (curve ——) if the preform profile in Fig. 1 is drawn into a fiber with a diameter of 2a = 11 μm.

Figure 3 plots measurements of group delay versus wavelength obtained from 1-km-long fibers that were drawn from the preform characterized in Fig. 1. Fiber No. 1 had a diameter of 2a = 11 μm and Fiber No. 2 had a diameter of 2a = 13.2 μm. The curves were fitted to the data using a least-mean-square-fit method.[7] Figure 4 shows chromatic dispersion spectra obtained by differentiating the curve shown in Fig. 3. The predicted (curve . . . .) and measured (curve ——) values of chromatic dispersion for the optimal fiber core diameter (2a = 11 μm) agree closely. Notice that the fiber dispersion is less than 0.665 ps/ km × nm between the two spectral zero crossings at λ = 1.495 μm and 1.666 μm. This low-dispersion spectral range, for Fiber No. 1, is nearly 2.5 times wider than for the curve, labeled Miya et al.,[4] corresponding to the best previously published result. Figure 5 shows transmission bandwidth spectra transformed[7] from chromatic dispersion spectra in Fig. 4. Results are applicable if Fibers No. 1 and No. 2 are excited by a laser source with a 4-nm linewidth and can be rescaled for different linewidths. Notice that bandwidth values for Fiber No. 1 are larger than 100 GHz-km for all wavelengths spanning the region from 1.45 μm to 1.73 μm. Low loss values were maintained simultaneously with low dispersion. Specific loss values were 0.8 dB/km at λ = 1.3 μm, 0.4 dB/km at λ = 1.6 μm, and 4 dB/km near the center of the water absorption peak at λ = 1.39 μm.

The importance of choosing optimal lightguide parameters can be appreciated by making comparisons between the dispersion and bandwidth spectra of Fiber No. 1 (solid line) and Fiber No. 2 (dashed line). Fiber No. 2, which has a cladding diameter (2a) about 2.2 μm (or 20 percent) larger than the optimal value, shows only one zero dispersion wavelength. As a result, 100 GHz-km bandwidth values can be maintained only within a 0.026-μm wavelength range that is an order of magnitude narrower than the corresponding 0.28-μm wavelength range for the optimal Fiber No. 1.

In conclusion, this brief reports the fabrication procedure and transmission properties of a double-clad fiber which has potential for wavelength-division-multiplexing applications within wide wavelength ranges. A computer analysis program was used to determine the optimal lightguide diameter which demonstrated high bandwidths over the widest range of wavelengths (1.45 μm to 1.73 μm) published to date. Further efforts are concentrating on fabricating fibers whose high bandwidth spectra cover the band from 1.3 μm to 1.55 μm.

ACKNOWLEDGMENTS

REFERENCES

1. S. Kawakami and S. Nishida, "Characteristics of a Doubly Clad Optical Fiber with a Low-Index Inner Cladding," IEEE J. Quan. Elect., *QE-10,* (1974), 879–87.
2. K. Okamoto et al., "Dispersion Minimization in Single-Mode Fibres Over a Wide Spectral Range," Electron. Letters, *15,* (1979), 729–31.
3. L. G. Cohen, W. L. Mammel, and S. Lumish, "Tailoring the Shapes of Dispersion Spectra to Control Bandwidths in Single-Mode Fibers," Paper 3.3, Tech. Dig., 7th Eur. Conf. on Opt. Comm., Copenhagen, Denmark, September 8–11, 1981.
4. T. Miya et al., "Fabrication of Low Dispersion Single-Mode Fibers Over a Wide Spectral Range," IEEE J. of Quan. Elect., *QE-17,* (1981), 858–61.
5. W. L. Mammel and L. G. Cohen, "Prediction of Fiber Transmission Characteristics from Arbitrary Refractive-Index Profiles," Appl. Optics (February 1982).
6. L. S. Watkins, "Laser Beam Refraction Traversely Through a Graded-Index Preform to Determine Refractive Index Ratio and Gradient Profile," Appl. Optics, *18,* (1979), 2214–22.
7. L. G. Cohen, W. L. Mammel, and S. Lumish, "Dispersion and Bandwidth Spectra in Single-Mode Fibers," IEEE J. Quan. Elect. (January 1982).
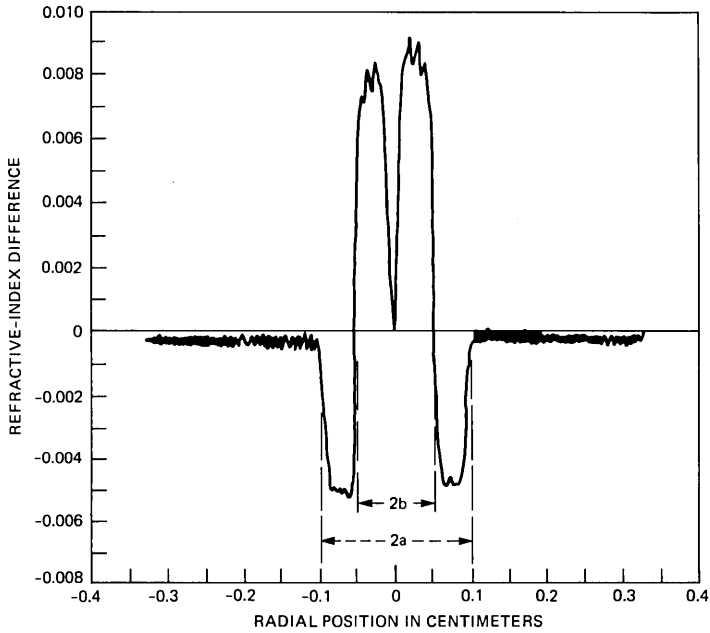
Fig. 1—Refractive index profile of a double-clad single-mode fiber preform measured by the laser beam refraction technique.
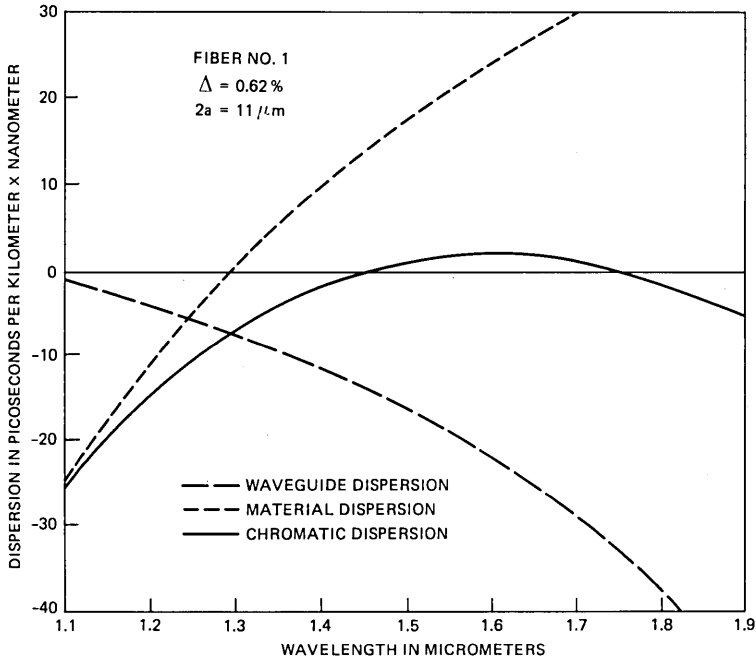

Fig. 2—Dispersion spectral components predicted if the preform profile in Fig. 1 is drawn into a fiber with 2a = 11 μm diameter.
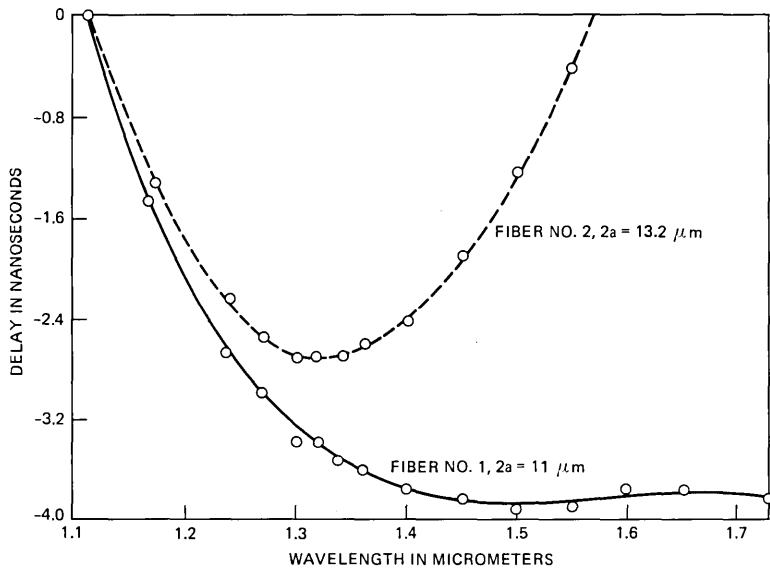
Fig. 3—Group delay spectra from 1-km length of Fibers No. 1 (2a = 11 μm) and No. 2 (2a = 13.2 μm).
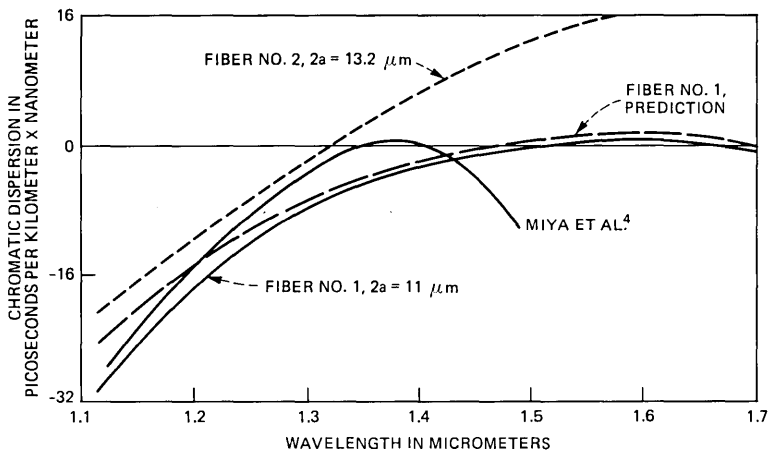
Fig. 4—Chromatic dispersion spectra calculated from Fig. 3 and predicted chromatic dispersion spectrum for diameter 2a = 5.5 μm (dotted line).
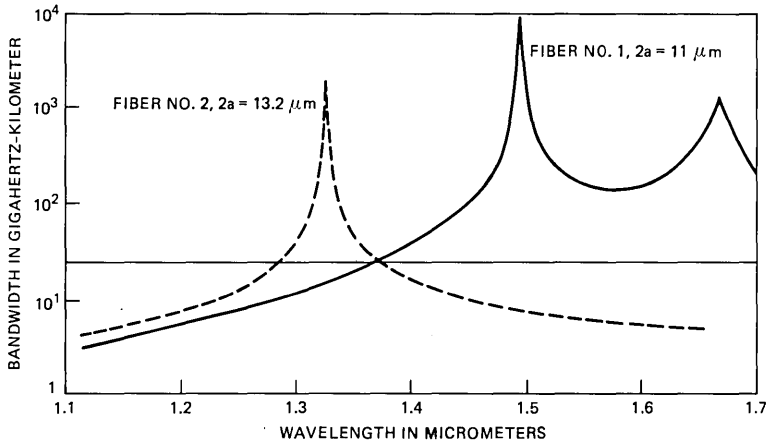
Fig. 5—Bandwidth spectra calculated from Fig. 4.