An Evaluation of AM Data System Performance by Computer Simulation

By R. A. GIBBY

(Manuscript received February 9, 1960)

The mathematical relationships that describe an amplitude-modulated data system are developed in a form suitable for programming on a high-speed digital computer. These equations contain expressions that specify in general terms the transmission-frequency characteristics of a transmission medium. A data signal composed of a train of raised-cosine shaped pulses is generated in the stimulating process. The simultation provides a means for computing the resulting response of systems to pulse trains. The performance of a double-sideband AM data system is evaluated from measurements of the maximum vertical opening, or aperture, of the eye pattern formed by the received signal. This aperture is related to the system performance in terms of signal-to-noise ratio and error rate of the system. A verification of this technique is made by simultaing the conditions of an experimental laboratory data system on the computer and comparing computed and measured performance.

I. INTRODUCTION

The process of transmitting digital information and the problems relating to this type of communication have received much attention in recent years. Practical means of communication based on this type of information transmittal have been in use for a long time in the form of telegraph and teletype systems, but for both these systems the speed of operation is relatively slow, since the information is supplied either by a manual operator or by a mechanical device of some sort. Recently, however, many new sources of digital information have arisen for which the information rates may range from telegraph speed up to several orders of magnitude greater than that speed. This has created a need for new high-speed data transmission systems. To this need for higher speed has been added the requirement of greater accuracy — amounting to, in some cases, essentially error-free transmission. For a given bandwidth both high speeds and accurate transmission are often difficult to realize. For this reason, a great deal of effort has been spent in devising means for utilizing a channel as effectively as possible in order that both speed and accuracy may be maximized.

This effort may be divided into two areas. The first is concerned with what may be called the terminal problem, in which effort is directed toward developing data system terminal equipment that makes best use of the transmission channel. This has resulted in a number of competing schemes for data transmission based on different modulation methods, such as double and vestigial sideband amplitude modulation, frequency modulation and phase modulation. Also, a number of different detection methods are available for these various systems.

The second area of effort may be referred to as the transmission problem. Here, work is directed toward determining what influence the various transmission factors such as the transmission-frequency characteristics and noise have on the transmitted data signal, and what conditions should be maintained for a particular data system that is to operate at a given speed and accuracy.

In order to answer questions that arise in these two areas, it is necessary that a method for evaluating data system performance be devised. Good performance is associated with both high speed and accurate transmission, but quality of performance is not necessarily preserved by a direct interchange of these two quantities. It is of little value to send information at very high speeds with very low accuracy. The optimum performance is obtained, therefore, by transmitting data over a given channel at the highest rate possible, consistent with a given error criterion as dictated by the system requirements.

Because of the many factors that contribute to the over-all performance, the problem of evaluating performance is a difficult one. The most direct method and the one that has been most widely used is that of measuring performance on a data system under actual operating conditions. Some results from such experimental tests have been given[1] and made general, at least to some extent. The experimental results, however, are of necessity limited in their generality, and the task of obtaining results by this method is expensive and time-consuming. In addition, for design purposes it is desirable to know how a data system will perform before it is built as well as after.

Another approach to evaluating performance is by means of a theoretical analysis of the problems involved. A considerable amount of work has been done in this area.[2,3,4,5] While this work is of great value, the complexity of the problems make a general solution untractable, except under conditions of rather restricted and idealized assumptions. For conditions of practical significance, however, a system analysis using mathematical models is useful, provided the task of a numerical computation can be carried out.

This paper is the result of an effort to evaluate the performance of an AM data system by methods of numerical analysis and simulation of the system on a digital computer. In this work, the general method of analysis and simulation is outlined, a criterion for performance is defined and some results that have been obtained by the use of this method are given. While the problem is analyzed in rather general terms, the particular results presented apply to the specific problem of evaluating the degradation in performance due to delay distortion.

II. BASIC DATA SYSTEM

The basic data system considered here is shown in Fig. 1. The input function $f(t)$ is the electrical representation of the random sequence of binary information to be transmitted. This signal is made up of a succession of identical raised-cosine shaped pulses interspersed with gaps or spaces where no pulse occurs, forming a pattern of the type illustrated in Fig. 2. In this pattern the occurrence or absence of a pulse is governed by the occurrence of a one or a zero in the binary information. These

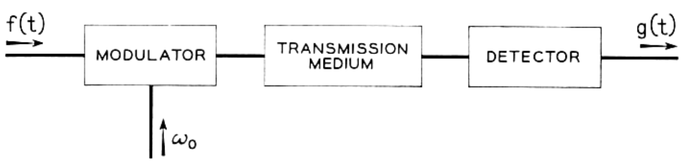


Fig. 1 – Data transmission system.

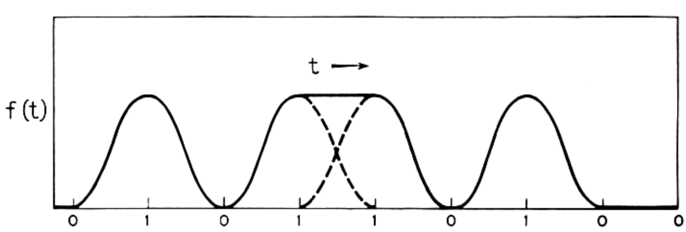


Fig. 2 – Binary information and equivalent pulse pattern.

binary values correspond in time to the points at which pulses may occur in the pattern, and they are separated by a uniform time interval called the *sampling interval*, which is the reciprocal of the bit speed of the system.

The basic pulse used to form the pattern is shown in Fig. 3. It is chosen to have a maximum width t_0, which is equal to two bit intervals at its base. Because of this fact, and because of the raised-cosine shape, pulses that are adjacent in the pattern of Fig. 2 add together over the interval that separates them to give a constant value equal to the maximum value of a single pulse. Also, adjacent gaps produce a constant value equal to zero. For the case of alternating values of one and zero, transitions are produced that are portions of a sinusoid.

For purposes of analysis it will be useful to consider a single pulse of the type used in forming the pulse pattern of Fig. 2. To do this a particular pattern is used, as shown in Fig. 3(a), in which $f(t)$ is formed by the binary sequence 010 followed by a sequence of zeros and then repeated periodically at intervals of T. This method of specifying the signal permits the analysis of the basic pulse to be made on a Fourier series basis that is conveniently adapted to numerical computation on the digital computer.

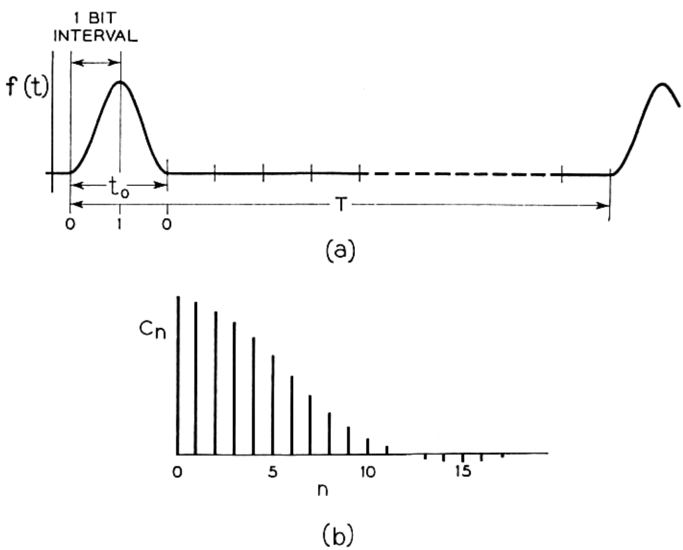


Fig. 3 – Basic pulse used to form pattern: (a) raised-cosine pulse; (b) associated spectrum.

The spectral density function resulting from a Fourier analysis of the periodic pulse pattern of Fig. 3(a), is shown in Fig. 3(b), and is expressed mathematically as

$$C_n = \frac{\sin \frac{n\pi t_0}{T}}{n\pi \left[1 - \left(\frac{nt_0}{T} \right)^2 \right]}. \tag{1}$$

We can imagine a perfect data system in which the signal $f(t)$ of Fig. 2 is transmitted without alteration. This will be considered a reference condition for measuring performance. Under this condition the binary information can be extracted from the received signal $g(t)$, which is identical to $f(t)$, by examining the pulse pattern at the sampling times to determine if the signal is greater than or less than some intermediate-value called the *slice level.* (In most cases this level is equal to one half the value of the maximum pulse height. Under certain conditions, however, it will be advantageous to use other values.)

For a system with ideal transmission-frequency characteristics, errors are produced if a sample is made during the time when a noise disturbance is large enough to cause the signal to make an erroneous excursion across the slice level. In order to meet system requirements the value of the signal-to-noise ratio must be such that the number of errors in a given time does not exceed a certain maximum. If the signal-to-noise ratio is less than this value the system is considered as failing. If the signal-to-noise ratio is such that the error criterion is just met, then a reduction in the noise (with the signal held constant) results in an operating margin against failure equal to the amount of the noise reduction. If the system were operating without noise, the amount of noise necessary to bring it to the failing condition would be called the *noise margin* for the system.

As the transmission-frequency conditions depart from ideal, the output $g(t)$ will no longer be identical to the input $f(t)$. This lack of fidelity is evidence of system degradation. It will be shown later how this degradation can be expressed quantitatively as a reduction in the noise margin for the system.

Transmission of the data signal over most facilities requires some form of modulation. It will be assumed here that the signal is amplitude-modulated by an ideal product modulator, and that an ideal envelope detector is used at the receiver. Also, the system will be considered, for the present, as being noise-free, and the impairment that the data signal suffers in transmission will be due entirely to the frequency characteristics of the transmission medium.

III. TRANSMISSION CHARACTERISTICS

The basic relationship that describes the transmission-frequency characteristics of a transmission system is given as

$$T(\omega) = A(\omega)e^{-j\psi(\omega)}, \tag{2}$$

in which $A(\omega)$ is the attenuation and $\psi(\omega)$ the phase characteristic. In the work to follow we will make use of the fact that the attenuation characteristic has even symmetry $A(\omega) = A(-\omega)$ and the phase characteristic has odd symmetry $\psi(\omega) = -\psi(-\omega)$ about the zero frequency axis.

To accommodate the use of a Fourier series development of the problem, the transmission-frequency characteristics can be expressed in terms of their values at discrete radian frequencies corresponding to the Fourier series components, rather than on a continuous basis. For this purpose, (2) becomes

$$T\left(\frac{n2\pi}{T}\right) = A\left(\frac{n2\pi}{T}\right)e^{-j\psi(n2\pi/T)}. \tag{3}$$

The phase curve of Fig. 4(a) has a shape typical of the characteristics of many transmission facilities and can be approximated by a sine-wave departure from linearity as given by

$$\psi(\omega) = a\omega - b\sin(\omega\tau + \theta), \tag{4}$$

where:

b is the maximum departure of the phase from linearity,

τ is the reciprocal of the period of the phase curve and

θ is the displacement of the phase curve from symmetry about the carrier.

The corresponding envelope delay is shown in Fig. 4(b) and is expressed analytically as the derivative of the phase:

$$\frac{d\psi}{d\omega} = a - b\tau\cos(\omega\tau + \theta). \tag{5}$$

The linear term in the phase expression results in only a constant delay, and the distortion in transmission is therefore produced entirely by the sine-wave departure from linearity. The three parameters b, τ and θ give a complete description of any sinusoidal delay or phase characteristic. A wide variety of such curves can be realized by letting these three parameters range over values of interest (it will be shown later that, with a sinusoidal representation of phase, a very close approximation to actual transmission characteristics can be achieved). With this repre-

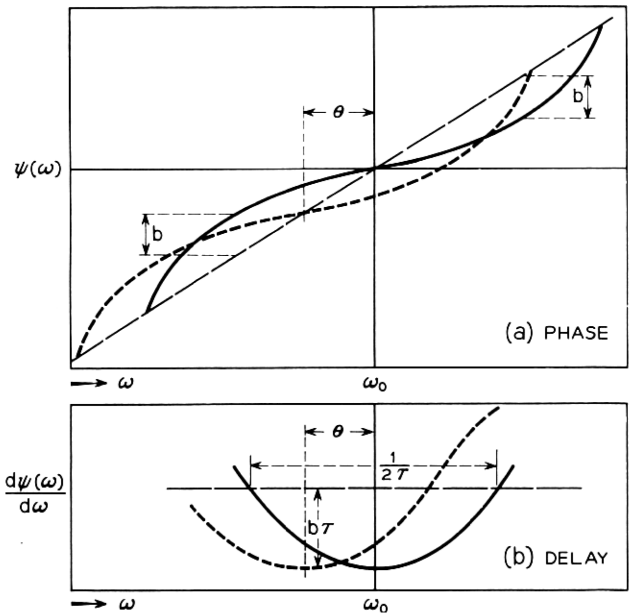


Fig. 4 – Sinusoidal characteristics: (a) phase; (b) delay.

sentation of the transmission characteristics, the problem of determining the performance of a data system in terms of the parameters b, τ and θ will now be considered.

IV. ANALYSIS OF AN AM DATA SYSTEM

In this section,* the analysis of an AM system is made in order to obtain the expression for the output function $g(t)$ in terms of the input function $f(t)$ and the parameters b, τ and θ. For a periodic input pulse as shown in Fig. 3(a) the output response will also be periodic. If the period T is chosen large enough to allow the transient response of the output to be restored essentially to zero there will be no intersymbol interference between these periodic pulses. We will then have a condition in which both $f(t)$ and $g(t)$ can be represented by a Fourier series.

The signal that results from an amplitude modulation of $f(t)$ by a cosine carrier is given by

$$h(t) = f(t) \cos \omega_0 t. \tag{6}$$

* This analysis is based on an unpublished memorandum of R. G. Segers. It includes some slight modification for use in the present study.

Expressing $f(t)$ by a complex Fourier series representation and using the complex form of the cosine function, we have

$$h(t) = \frac{e^{j\omega_0 t} + e^{-j\omega_0 t}}{2} \sum_{n=-\infty}^{\infty} C_n e^{jn(2\pi/T)t}, \tag{7}$$

where

$$C_n = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-jn(2\pi/T)t} \, dt. \tag{8}$$

Equation (7) may be expressed as

$$h(t) = \frac{1}{2} \sum_{n=-\infty}^{\infty} C_n e^{j[n(2\pi/T)+\omega_0]t} + \frac{1}{2} \sum_{n=-\infty}^{\infty} C_n e^{j[n(2\pi/T)-\omega_0]t}. \tag{9}$$

The modulated signal expressed by (9) is composed of two doubly infinite spectra, one distributed about the positive carrier of radian frequency ω_0, and the other about the negative carrier of radian frequency $-\omega_0$. It is apparent that this representation is redundant, and for purposes of efficient computation the number of terms can be reduced by a factor of two. This will be done in a way that preserves certain desired symmetries in the equations and allows the results to be interpreted meaningfully. First, however, it is necessary to determine the carrier response by introducing into (9) the expression for the transmission-frequency characteristics. The transmission-frequency characteristics given by (3) relate to baseband conditions; in order to properly relate this expression to carrier transmission it is necessary to translate these characteristics by an amount of $+\omega_0$ and $-\omega_0$. Making this transformation in (3) and substituting the results into (9) gives the following:

$$\begin{aligned} g(t) = {} & \frac{1}{2} \sum_{n=-\infty}^{\infty} C_n A\left[n\left(\frac{2\pi}{T}\right) + \omega_0\right] e^{j[n(2\pi/T)+\omega_0]t} e^{-j\psi[n(2\pi/T)+\omega_0]} \\ & + \frac{1}{2} \sum_{n=-\infty}^{\infty} C_n A\left[n\left(\frac{2\pi}{T}\right) - \omega_0\right] e^{j[n(2\pi/T)-\omega 0]} e^{-j\psi[n(2\pi/T)-\omega_0]}, \end{aligned} \tag{10}$$

which is the expression of the response of the carrier signal to the transmission system.

Use is now made of the fact that $f(t)$ is an even function. The quantity C_n, which can, therefore, be expressed by taking twice its value over one half of its range, is given as,

$$C_n = \frac{2}{T} \int_0^{T/2} f(t) \cos n\left(\frac{2\pi}{T}\right) t \, dt. \tag{11}$$

When this, and also the fact that the attenuation must have even symmetry and the phase must have odd symmetry, are taken into account, (10) can be rewritten for $n \gtreqless 0$, giving

$$r(t) = \sum_{n=0}^{\infty} C_n A\left[n\left(\frac{2\pi}{T}\right) + \omega_0\right] \cos\left\{\left[n\left(\frac{2\pi}{T}\right) + \omega_0\right] t - \psi\left[n\left(\frac{2\pi}{T}\right) + \omega_0\right]\right\} + \sum_{n=1}^{\infty} C_n A\left[n\left(\frac{2\pi}{T}\right) - \omega_0\right] \cos\left\{\left[n\left(\frac{2\pi}{T}\right) - \omega_0\right] t - \psi\left[n\left(\frac{2\pi}{T}\right) - \omega_0\right]\right\}. \tag{12}$$

Rather than computing with (12) as it stands, it will be more useful to extract the high-frequency carrier by the use of elementary trigonometric relationships, giving

$$\begin{aligned} r(t) = {} & \cos \omega_0 t \sum_{n=0}^{\infty} C_n A\left[n\left(\frac{2\pi}{T}\right) + \omega_0\right] \cdot \cos\left\{n\left(\frac{2\pi}{T}\right) t - \psi\left[n\left(\frac{2\pi}{T}\right) + \omega_0\right]\right\} \\ & - \sin \omega_0 t \sum_{n=0}^{\infty} C_n A\left[n\left(\frac{2\pi}{T}\right) + \omega_0\right] \cdot \sin\left\{n\left(\frac{2\pi}{T}\right) t - \psi\left[n\left(\frac{2\pi}{T}\right) + \omega_0\right]\right\} \\ & + \cos \omega_0 t \sum_{n=1}^{\infty} C_n A\left[n\left(\frac{2\pi}{T}\right) - \omega_0\right] \cdot \cos\left\{n\left(\frac{2\pi}{T}\right) t - \psi\left[n\left(\frac{2\pi}{T}\right) - \omega_0\right]\right\} \\ & + \sin \omega_0 t \sum_{n=1}^{\infty} C_n A\left[n\left(\frac{2\pi}{T}\right) - \omega_0\right] \cdot \sin\left\{n\left(\frac{2\pi}{T}\right) t - \psi\left[n\left(\frac{2\pi}{T}\right) - \omega_0\right]\right\}. \end{aligned} \tag{13}$$

Equation (13) is a general expression for the received carrier signal. An expression for the received response at the detector output, $g(t)$, can be obtained by considering only the envelope of this carrier. This allows $g(t)$ to be evaluated for any arbitrary attenuation or phase characteristic. In applying these results in the work to follow, the complete generality of (13) will not be preserved, since a special type of transmission characteristic is being considered. By sacrificing some of the generality, however, an advantage is gained in the facility with which the results can be logically organized and utilized.

In evaluating the way a data system's performance varies as the transmission-frequency characteristics are altered, the results will be more lucid and systematic if the phase and attenuation are considered separately. While these two qualities are not independent in physical transmission facilities, it is not entirely academic to treat them independently, since this approach is often used in equalizing facilities.

In this study the influence of sinusoidal phase variation as given by (4) will be considered in some detail, and, in order to isolate the effects of this influence, the attenuation will be considered equal to unity over the frequency band of interest. (A similar study could be made of the effects of attenuation by specifying linear phase characteristics and allowing the attenuation to vary systematically.) For computational purposes we will use an approximate expression for $r(t)$ by considering a finite number of frequency components. Applying these conditions to (13) and rewriting this equation for notational convenience, the carrier response may be expressed as

$$\begin{aligned} r(t) &= (M + N) \cos \omega_0 t + (R + S) \sin \omega_0 t \\ &= [(M + N)^2 + (R + S)^2]^{\frac{1}{2}} \cos [\omega_0 t + \varphi(t)], \end{aligned} \tag{14}$$

where

in-phase components:

$$\begin{aligned} M &= \sum_{n=0}^{N} C_n \cos \left[\frac{n2\pi}{T} t + b \sin \left(\frac{n2\pi}{T} \tau + \theta \right) \right], \\ N &= \sum_{n=1}^{N} C_n \cos \left[\frac{n2\pi}{T} t + b \sin \left(\frac{n2\pi}{T} \tau - \theta \right) \right]; \end{aligned}$$

quadrature components:

$$\begin{aligned} R &= - \sum_{n=0}^{N} C_n \sin \left[\frac{n2\pi}{T} t + b \sin \left(\frac{n2\pi}{T} \tau + \theta \right) \right], \\ S &= \sum_{n=1}^{N} C_n \sin \left[\frac{n2\pi}{T} t + b \sin \left(\frac{n2\pi}{T} \tau - \theta \right) \right] \end{aligned} \tag{15}$$

and

$$\varphi(t) = \tan^{-1} \left(\frac{R + S}{M + N} \right). \tag{16}$$

In (14), the expression for the carrier response, the envelope of the carrier signal that is the output of the envelope detector is given by

$$g(t) = \sqrt{(M + N)^2 + (R + S)^2}. \tag{17}$$

The desired expressions, which relate the output of the detector $g(t)$ to the input signal $f(t)$ (as manifest in C_n), and the three parameters b, τ and θ, which describe the phase characteristics of the transmission path, are given by (15) and (17). It is these equations that form the basis of the computational work on the digital computer. A careful examination of them will give not only a valuable insight into the influence of sinusoidal phase characteristics on AM transmission, but will also serve as a guide in computational work and help in interpreting results. It is evident from (14) that the terms M and N may be considered as modulating a cosine carrier, and are, therefore, referred to as the *in-phase components* of the received signal, while the terms R and S modulate a sine carrier and are designated as *quadrature components*. The expression for the envelope response, (17), indicates that these two quantities cannot be superimposed directly, but must be added at right angles. Also it can be seen from (15) that for $b = 0$ the quadrature components cancel and the in-phase components add together, and the received signal is an exact replica of the transmitted signal. For $b \neq 0$ but $\theta = 0$, the quadrature terms cancel and the in-phase terms add to give a distorted replica of the transmitted signal at the receiver. For the case $b \neq 0$ and $\theta \neq 0$, the received signal is a distorted replica of the transmitted signal that contains both in-phase and quadrature components.

It can be seen from (3) and Fig. 3 that a reversal of the sign in b amounts to inverting the phase curve with respect to the linear phase line. It has been shown[1] that such a reversal in the phase curve results in a reversal in time of the received signal. This important fact will be utilized later in interpreting the performance results. Another useful observation that can be made from examining (15) is that a reversal in the sign of θ in these equations leaves the received signal unaltered.

V. COMPUTED PULSE RESPONSE

Equations (1), (15) and (17) provide a means for computing the output of the envelope detector of an AM transmission system subject to a sinusoidal variation in the phase characteristics, when the periodically recurring raised cosine pulse of Fig. 3 is applied at the input. These equations were programmed for numerical evaluation on the IBM 704 computer. In evaluating these equations, the computations were carried out for a specific bit speed and specific values of b, τ and θ. The results from these computations, however, can be given on a normalized basis. To do this it is necessary to express the frequency scale of the phase or delay curves in terms of the bit speed rather than in cycles per second, and also to express the time scale of the response function and the delay curve in terms of the bit length rather than in seconds.

A number of normalized pulse responses are shown in Fig. 5, together with the normalized delay that produces the responses. The curves of Fig. 5(b) are for the case of symmetrical delay ($\theta = 0$) as shown in Fig. 5(a). For these cases, the quadrature component is zero, as would be expected from (15). The response curves in this figure are plots of the in-phase component of the received signal.

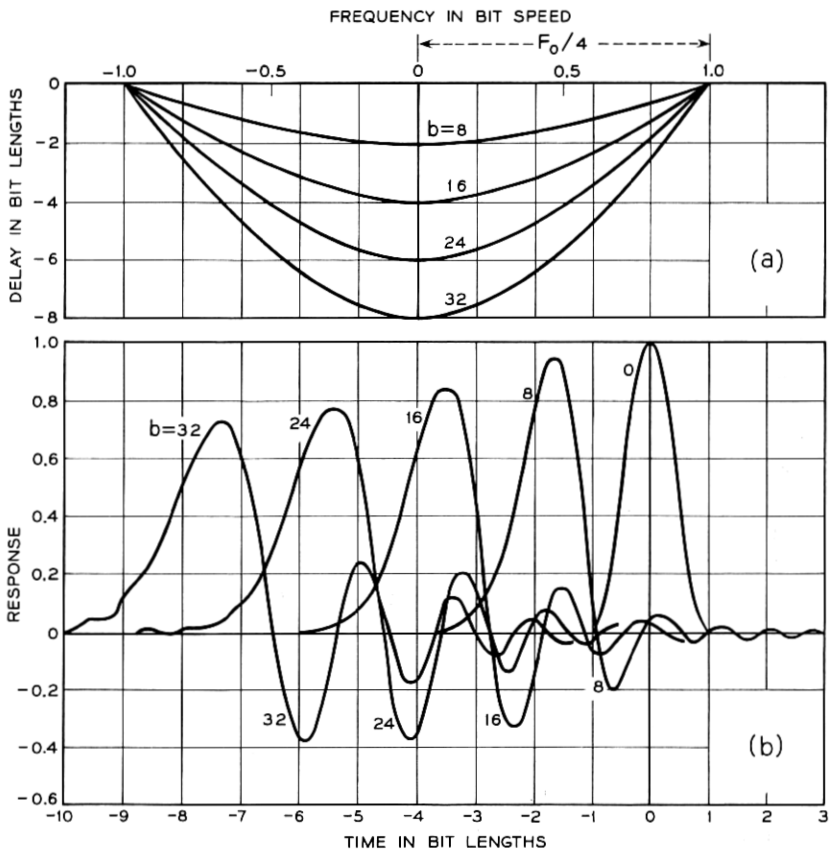


Fig. 5 – (a) Normalized delay; (b) corresponding pulse response.

To illustrate how these normalized results can be related to specific cases, consider an input pulse transmitted at a bit speed of 1200 bits per second (bit length = 833.3 microseconds). From Fig. 5(a) the period of the delay curve F_0 is equal to four times the bit speed or 4800 cycles, giving a value of $\tau = 1/F_0 = 208.33$ microseconds. The curve for $b = 8$ in Fig. 5(a) has a maximum delay excursion of $b\tau = 8 \times 208.3$ microseconds = 1666.3 microseconds = 2 bit lengths. The pulse response corresponding to this delay is shown in Fig. 5(b) for $b = 8$. These same results may also be applied to other bit speeds. For a bit speed of 2400 bits per second (bit length = 416.7 microseconds), $F_0 = 9600$, giving $\tau = 104.2$. Again the curve for $b = 8$ has a maximum delay of two bit lengths = $b\tau$ = 833.3 microseconds, and the same pulse response is produced here as for the low bit speed case.

In Fig. 6 pulse response curves are shown for conditions of asymmetry in the delay characteristics. This gives rise to the quadrature component in the received signal, and the resulting effect on the envelope response is shown. Here again the curves are normalized with respect to bit length. For the curve of Fig. 6(a), $\theta = \pm 45°$; for Fig. 6(b), $\theta = \pm 90°$. These curves illustrate the fact pointed out previously: that the response is independent of the sign of the angle θ.

These distorted pulse response curves indicate a degradation in transmission manifest in a general lowering of the pulse peak value and a spreading of the pulse in time. For the cases of larger distortion this results in a considerable amount of intersymbol interference between adjacent pulses in a data signal. This, however, is only a qualitative picture of the influence of delay distortion on system performance. In order to make this picture more precise, it is necessary to consider a signal consisting of a number of pulses making up a pattern such as that illustrated in Fig. 2. This extension to the basic operation of computing pulse response was achieved by combining the received response of a single pulse (before envelope detection) in accordance with the desired pattern. In doing this, the in-phase components from adjacent pulses were added together. Similarly, the quadrature components were added together, and the resulting envelope response was determined in accordance with

$$g(t) = \sqrt{[\Sigma(M + N)]^2 + [\Sigma(R + S)]^2}, \tag{18}$$

where M, N, R and S are defined by (15). By this method, the results previously discussed can now be extended so that the output response for pulse patterns can be obtained.

An example of pulse-pattern response showing the resultant in-phase

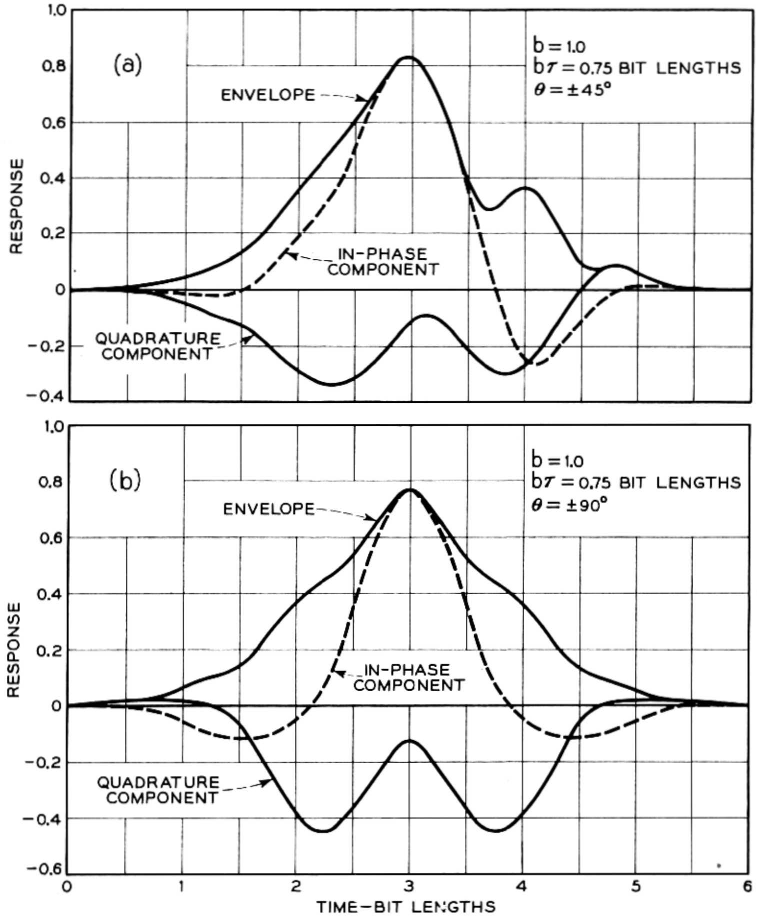


Fig. 6 – Pulse response curves.

and quadrature components and the envelope response is given in Fig. 7. In this figure a pattern of binary characters is shown together with the corresponding output response curve for conditions $b = 1$, $b\tau = 0.75$ bit length and $\theta = 90°$. In order to recover the binary information from the signal, the envelope response is examined at the sampling-time points to determine if the value of the curve at each of these points is less than or greater than the value of the slice level. The distance from the signal to the slice level for a given time point (as indicated by the arrows) is a measure of the noise margin for that point. The amount by which the noise margin is reduced by the transmission characteristic can be de-

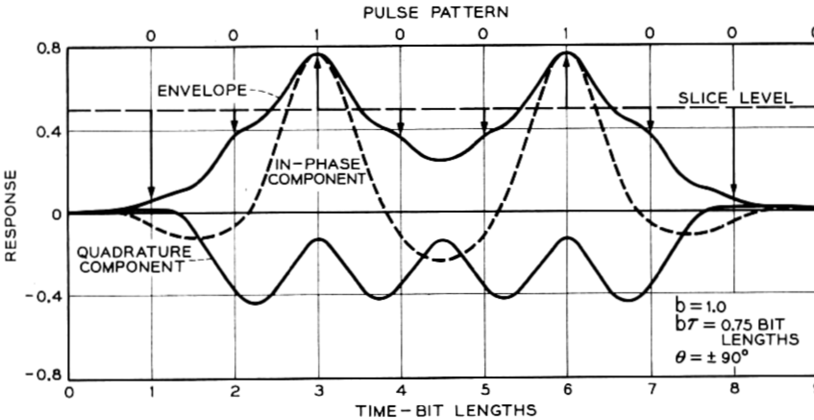


Fig. 7 – Basic method for evaluating performance.

termined by comparing these values with the corresponding values for an undistorted or reference output resulting from ideal transmission characteristics.

This procedure makes possible a quantitative evaluation of data system performance. It is evident, however, that for longer pulse patterns this procedure would soon become tedious and impractical if carried out in the manner shown. In order to alleviate this difficulty, the basic method for evaluating performance, as illustrated in Fig. 7, was included in the computer operation for any desired pulse pattern up to a certain maximum length. As the delay distortion increases, the range of intersymbol interference of a single pulse becomes larger; it was found necessary in this analysis to include cases for which the intersymbol interference extended over as many as nine bit lengths. In order to take this influence into account it was necessary to use a pattern that contained all possible combinations of binary characters that could occur in a nine-bit word. This results in a maximum pattern length of $2^9 = 512$ bits. A binary pattern of this type was used in the computer simulation work that was carried out.

VI. EYE PATTERN

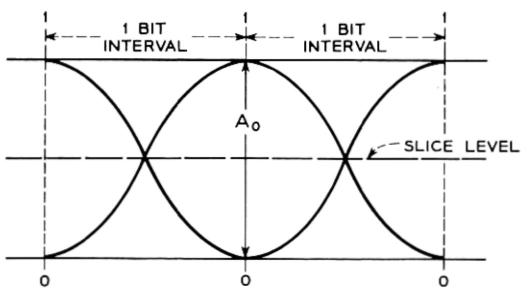
One way of portraying this procedure for evaluating data system performance for a long sequence of pulses is by means of the "eye" pattern. The manner in which this pattern is made can be illustrated by dividing the output response signal from the system into a number of segments each of duration equal to two bit lengths and plotting all these segments over the same two-bit time intervals. Such a plot for a reference or undistorted eye is shown in Fig. 8(a). Here the eye pattern is made up of traces that are straight lines at the top and bottom of the eye, corresponding to a sequence of ones and zeros, respectively, and traces that have sinusoidal shape in between, representing transitions from one to zero and zero to one. Since the pulses are undistorted, all traces of a particular type fall on top of one another, and the resulting eye pattern is made up of well-defined lines. In this case, the opening or aperture A, of the eye is equal to A_0, which is the maximum value of a single pulse. The quantity A is a measure of the noise margin and will be used as a criterion for data system performance.

For conditions of transmission distortion, the resulting eye pattern is illustrated by Fig. 8(b). Here the various individual traces have been distorted and no longer form well-defined lines of transition, but are spread out into bands. This results in a reduction of the aperture, which can be expressed numerically as an impairment in the system performance:

$$\text{transmission impairment} = 20 \log_{10} \frac{A_0}{A}. \tag{19}$$

There are two factors that determine the validity with which such an

(a)



(b)

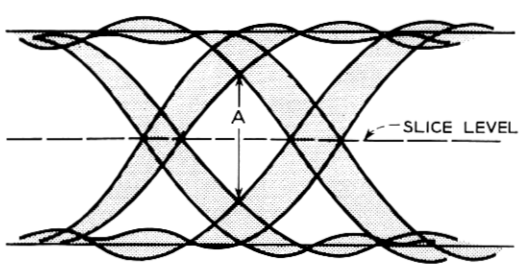


Fig. 8 – (a) Reference eye pattern; (b) received eye pattern with transmission distortion.

expression of performance describes an actual system. First, it must be acknowledged that the effective opening of the eye may change from one trace to another, and the quantity A as shown in Fig. 8(b) is a minimum bound of this opening. Since an opening of this magnitude will occur relatively infrequently this quantity is then a pessimistic estimate of the system performance; getting an exact measure would involve taking account of the statistical distribution of the quantity A. The probability of an error occurring would then be determined by convoluting this distribution with the distribution of noise. Such a refinement is rather complicated, and may not be necessary except in cases of very small aperture, since, for the larger values of A, the traces across the eye are rather densely packed.

The other fact to be considered is that in this simulation the degradation in performance due to timing recovery is not included. These results then specify the performance under conditions of ideal timing recovery, or for systems operating with ideal transmitted timing. Again, this is a good approximation to system operation with timing recovery, except for small aperture values, where timing jitter is unavoidable.

VII. EXPERIMENTAL VERIFICATION OF COMPUTER SIMULATION

It is of interest to determine how closely the performance results from this simulation procedure check with measured performance from an actual data system. The aperture value of an AM double-sideband data system operating in the voice band frequency range was measured at several bit speeds over a given transmission channel. The transmission characteristics of the channel were then applied to the computer simulation of the data system for the same bit speeds. The results from these two tests expressed in terms of aperture of the eye pattern are shown in Fig. 9. The fact that the computer simulation performance is slightly better over the range of bit speeds tested is to be expected, because of the assumption of idealization in the modulator and detector in making the simulation. These results tend to confirm the validity of the simulation technique as a means for evaluating data system performance.

VIII. RESULTS

A criterion for data system performance has been specified as A, the aperture of the eye pattern formed by the received signal of a data system. The results of computing performance based on the data system simulation are presented here as a portrayal of the manner in which A varies with a systematic variation of the parameters describing the

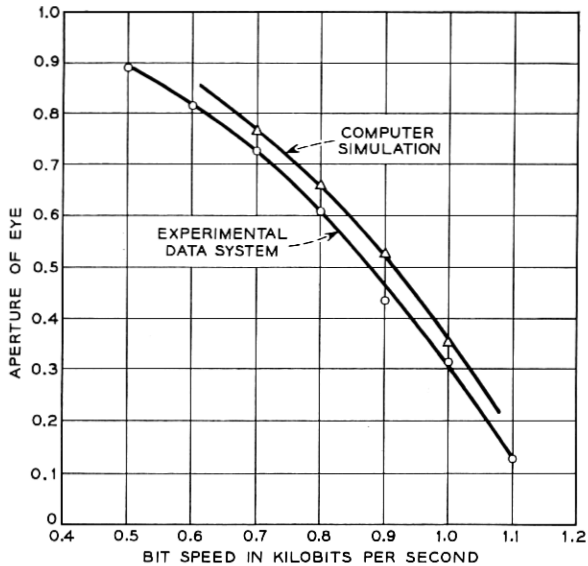


Fig. 9 – Comparison of computer simulation performance with experimental verification.

transmission characteristics of the system. To make these results more general, a normalized parameter δ, will be defined as $\delta = \tau \times$ bit speed. Results showing the manner in which the quantity A varies with b, δ and θ are given in Figs. 10 through 20, where b is expressed in radians and θ in degrees and δ is a quantity equal to the reciprocal of the period of the sinusoidal phase characteristics normalized with respect to bit speed.

In these figures, performance curves are plotted with the dependent variable b ranging over values sufficient to cause the aperture to vary from a reference value of unity to complete closure. The quantity δ varies from 0.2 bit length to 1.4 bit lengths. The degree of asymmetry is specified by the angle θ, which varies as a parameter from zero to 90 degrees. It is an apparent feature of each set of curves that the aperture value is reduced more rapidly, as b increases, for asymmetric phase conditions than for symmetrical conditions. This is an evidence of the influence of the quadrature component on data system performance. From these curves, it can also be seen that the relative influence of the quadrature component is much more pronounced for small values of δ than for larger values.

While the performance curves cover only one quadrant of the sinusoi-

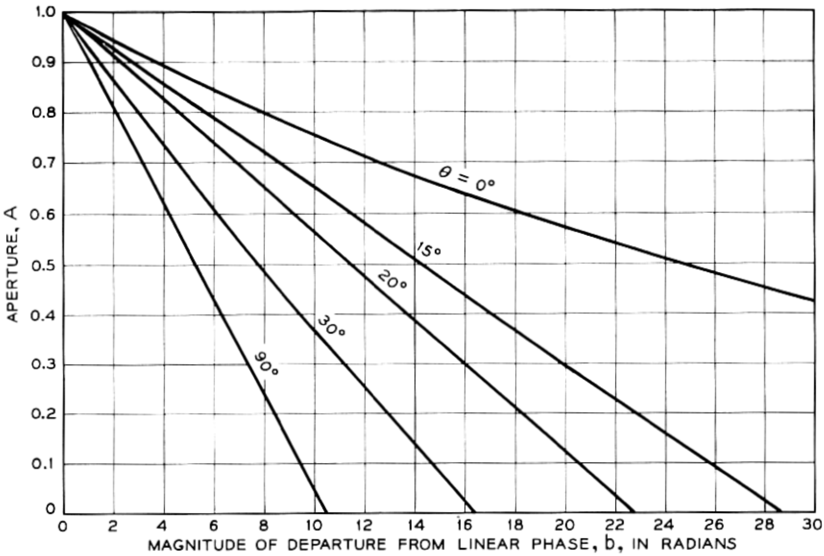


Fig. 10 – Plot of aperture vs. b for $\delta = 0.2 = \tau \times$ bit speed.

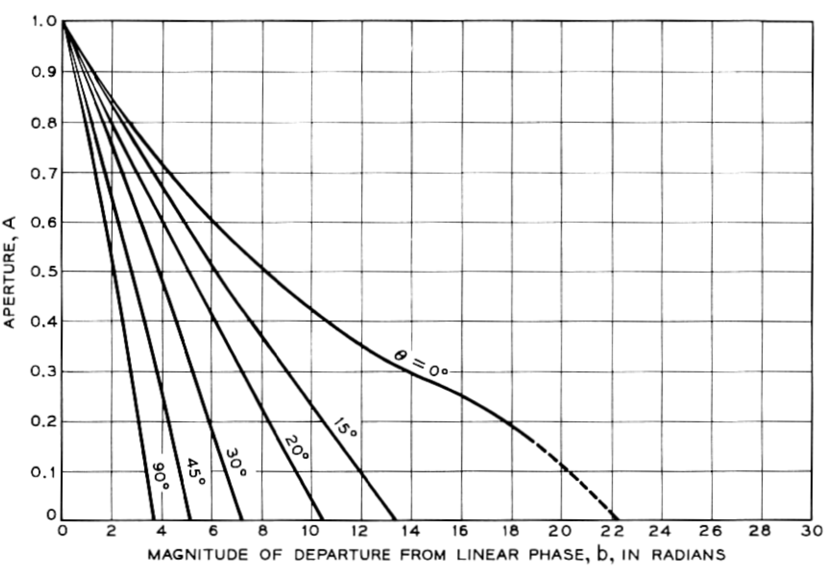


Fig. 11 – Plot of aperture vs. b for $\delta = 0.3 = \tau \times$ bit speed.

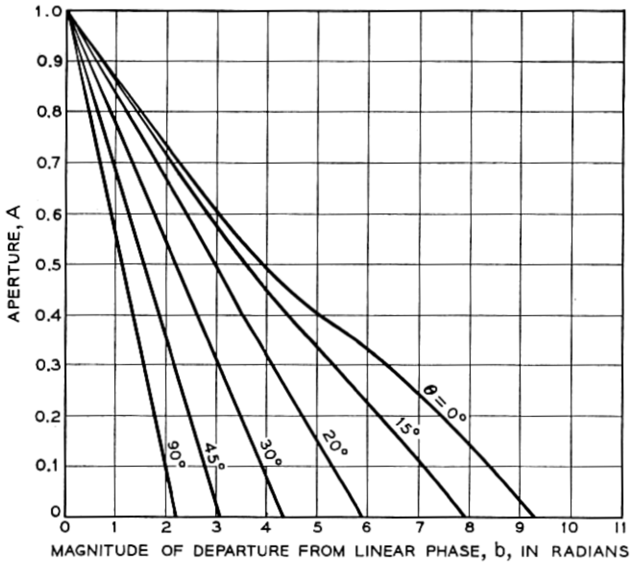


Fig. 12 – Plot of aperture vs. b for $\delta = 0.4 = \tau \times$ bit speed.

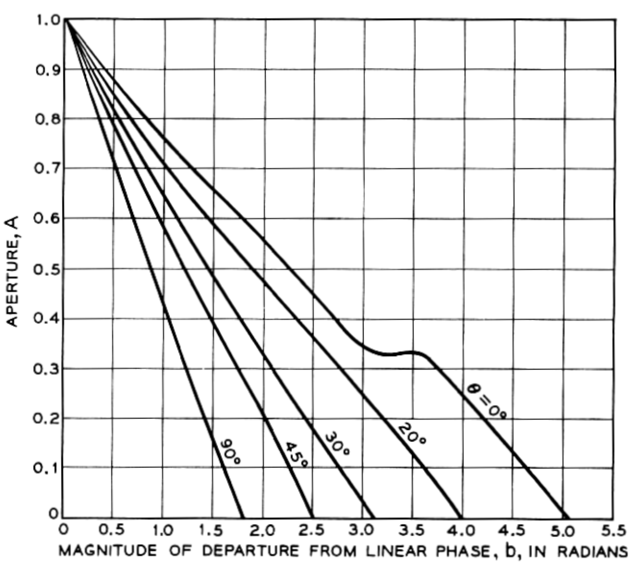


Fig. 13 – Plot of aperture vs. b for $\delta = 0.5 = \tau \times$ bit speed.

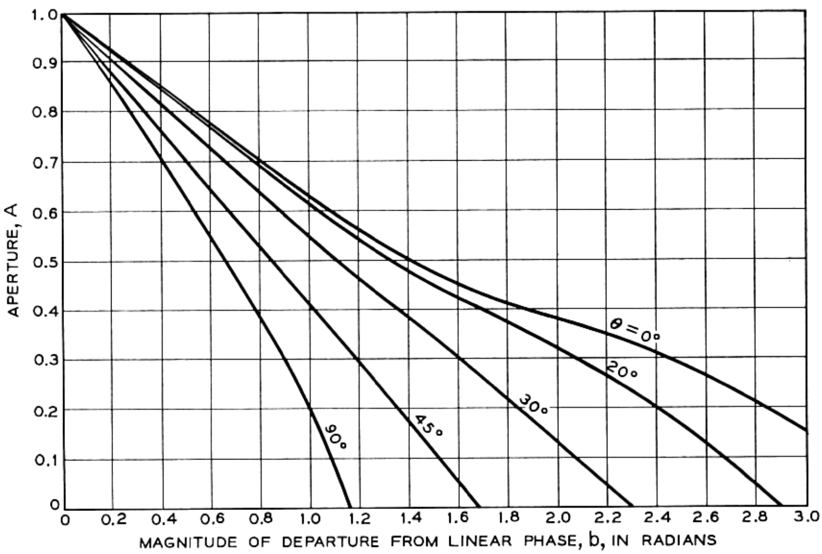


Fig. 14 – Plot of aperture vs. b for $\delta = 0.6 = \tau \times$ bit speed.

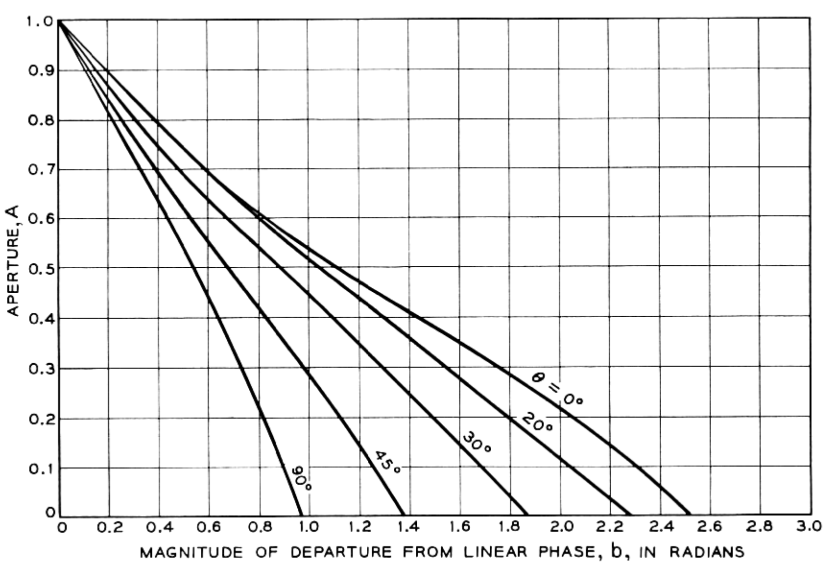


Fig. 15 – Plot of aperture vs. b for $\delta = 0.7 = \tau \times$ bit speed.

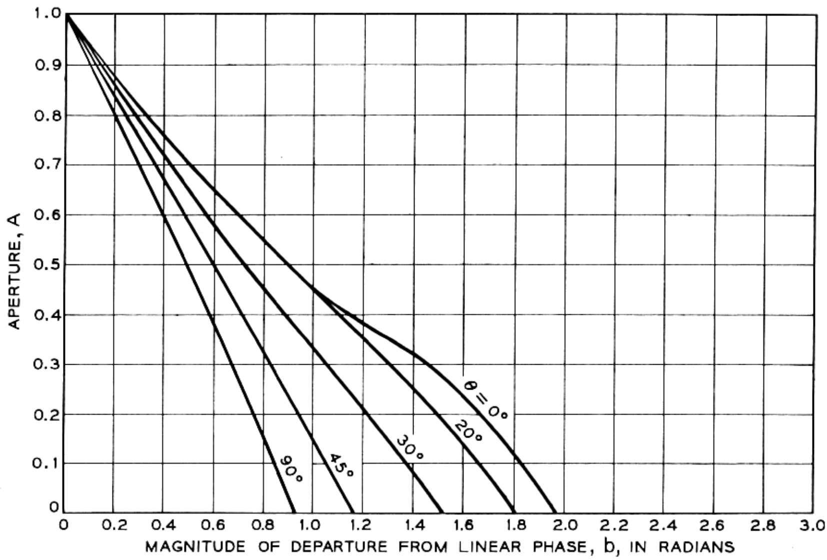


Fig. 16 – Plot of aperture vs. b for $\delta = 0.8 = \tau \times$ bit speed.

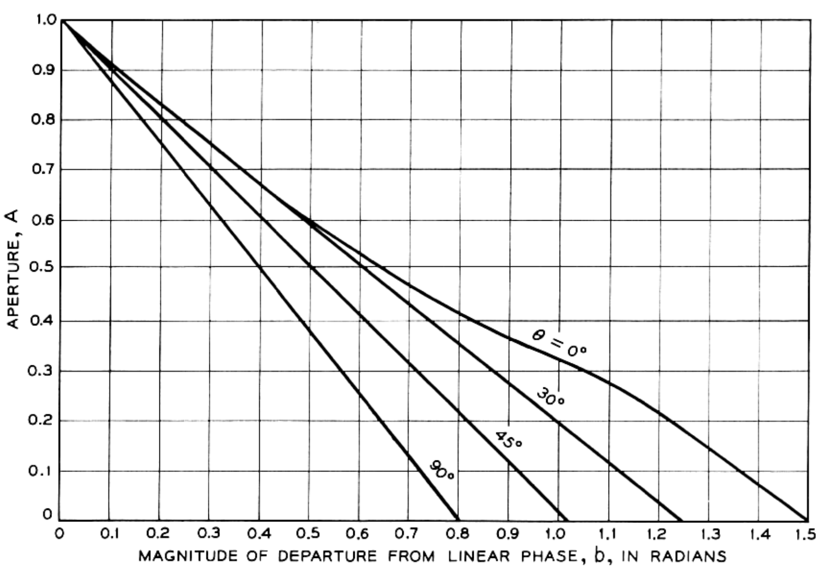


Fig. 17 – Plot of aperture vs. b for $\delta = 0.9 = \tau \times$ bit speed.

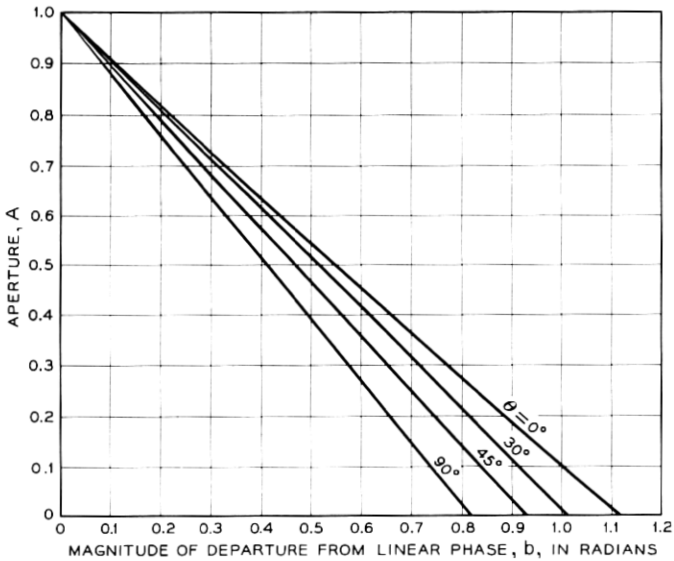


Fig. 18 – Plot of aperture vs. b for $\delta = 1.0 = \tau \times$ bit speed.

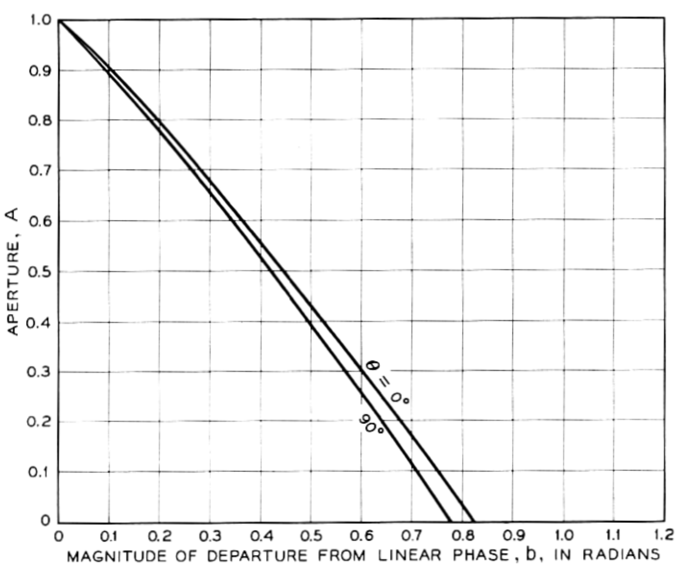


Fig. 19 – Plot of aperture vs. b for $\delta = 1.2 = \tau \times$ bit speed.

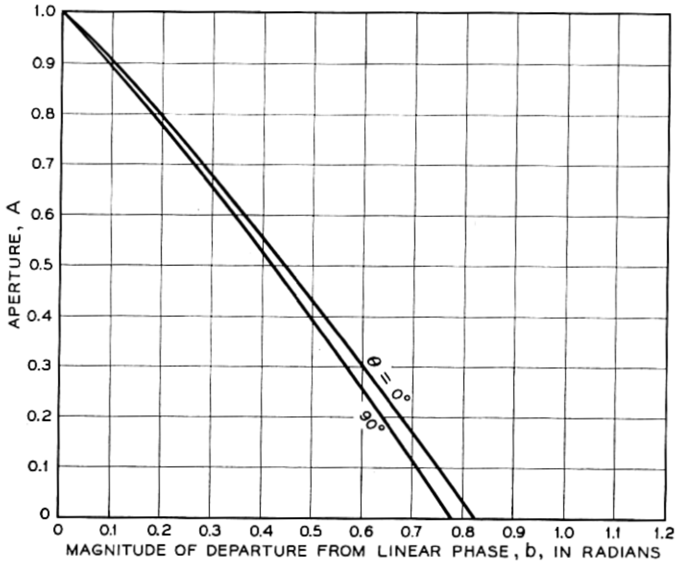


Fig. 20 – Plot of aperture vs. b for $\delta = 1.4 = \tau \times$ bit speed.

dal phase curve, the other three quadrants are duplicates of the curves shown. This can be seen from the facts previously stated, which indicated that a change in the sign of θ leaves the output response unchanged, and that a change in the sign of b reverses the time variable of the response and therefore reverses time in the eye pattern but does not change the aperture value.

The information given on the performance curves of Figs. 10 through 20 provides a means for evaluating the aperture of a data system for a wide range of sinusoidal phase characteristics, either by direct means or by interpolation. These results can be displayed in another manner, which gives a more concise picture of how the performance varies. To do this the performance information for a particular value of δ is plotted in polar form with b and θ as the radial and angular coordinates respectively. Such a plot is shown in Fig. 21 for $\delta = 0.75$, with contours of constant aperture shown as a parameter. The outer contour has a value of $A = 0$ and is the locus of points for which the aperture is completely closed. Any sinusoidal phase characteristic whose period corresponds to this value of δ can be located on the polar diagram of Fig. 21 as a point. If the magnitudes of b and θ are such that the point lies within the outer contour, the aperture value can be determined either directly or by interpolation. It is now possible to plot a path on this diagram showing how the value of A varies for changes in the values of b and θ.

This type of plot can be made for other values of δ. By combining such plots, the information given in Figs. 10 through 20 results in a very useful three-dimensional model with coordinates b, δ and θ, which depicts the performance of the data system. Such a model is shown diagrammatically in Fig. 22. In this three-dimensional space any sinusoidal phase (or delay) curve is represented as a point. A number of curves that illustrate the particular delay that occurs at various points are shown by Fig. 22(a), (b), (c) and (d). The surface of the model indicated by the dashed line in Fig. 22 represents the locus of points for which the aperture

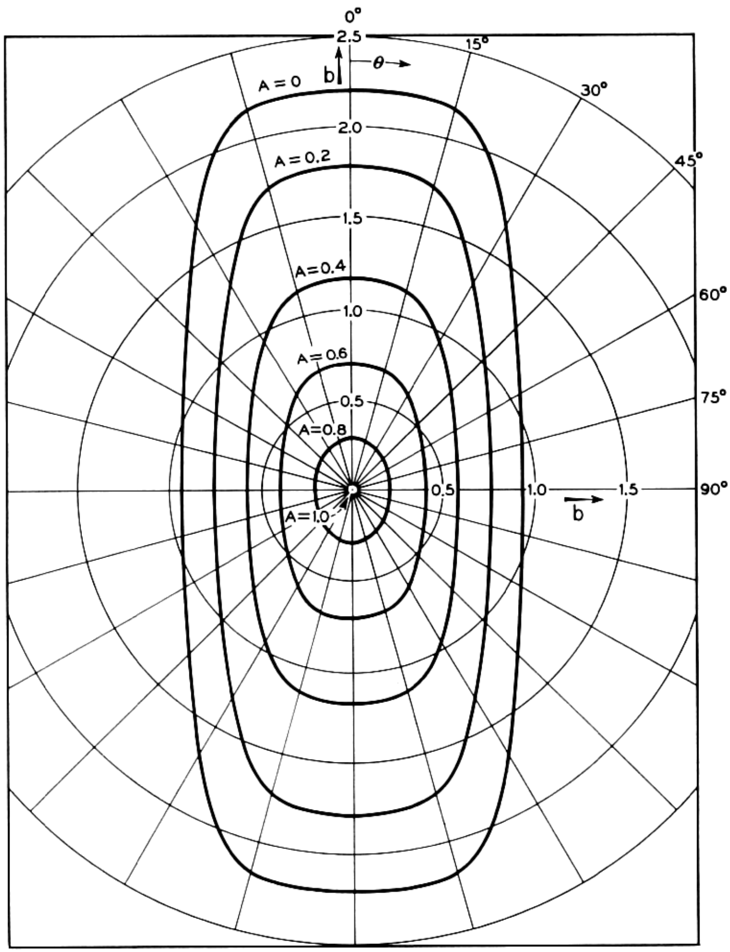


Fig. 21 – Polar diagram of b vs. θ for $\delta = 0.75$ with aperture as a parameter.

value is zero. For points inside this surface the aperture is greater than zero. Other surfaces within this model can be described by connecting all points having a common value of A. The points lying along the vertical axis represent conditions of zero distortion, or maximum aperture value A. A photograph of a performance model constructed from the numerical results obtained from the computer simulation is shown in Fig. 23.

IX. APPLICATION OF RESULTS TO ACTUAL TRANSMISSION FACILITIES

In the simulation of the data system that has been presented the phase characteristics have been assumed to be sinusoidal in shape, for reasons of mathematical convenience. Some rather general results have been given in terms of the parameters that describe the phase characteristic. In this section several specific transmission facilities will be considered in the light of this simulation technique.

The validity with which the sinusoidal phase (or cosine-shaped delay) represents actual transmission characteristics will be investigated, and computed performance resulting from data system operation over these facilities is given.

In Fig. 24, the delay characteristics of four message facilities are shown by the solid lines, together with a cosine approximation to these delay curves given by the dashed lines. It can be seen from Fig. 24 that the cosine-shaped delay curves give a very good approximation to the actual

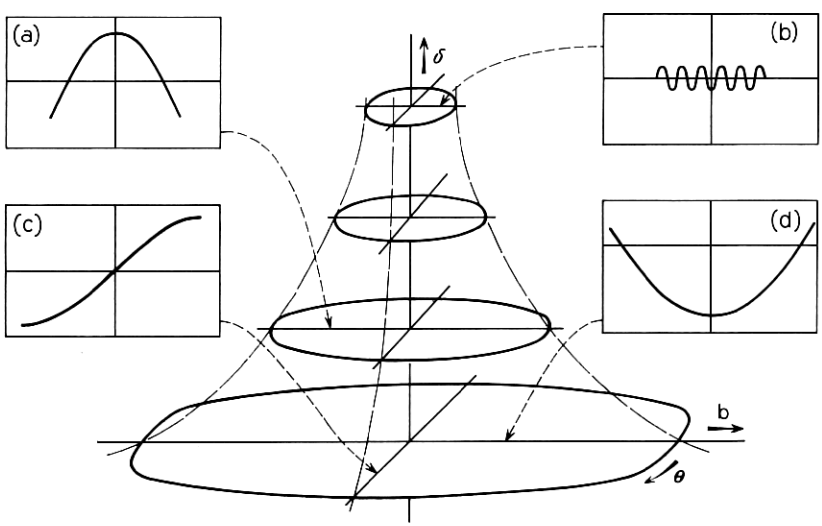


Fig. 22 – Three-dimensional model of data system performance.

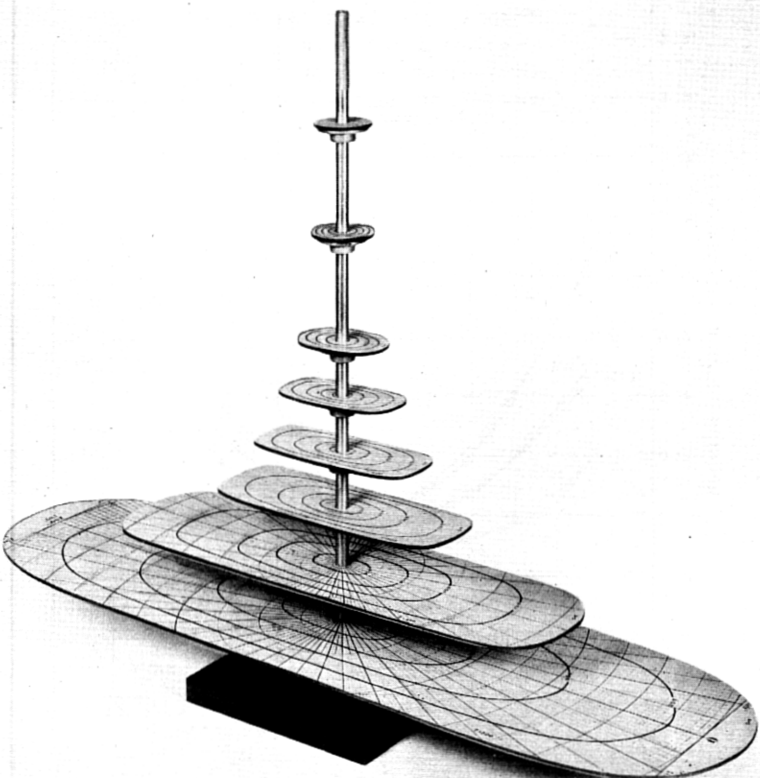


Fig. 23 – Performance model constructed from numerical results of computer simulation.

delay over the frequency range of 1000 to 2600 cycles per second, the maximum departure of the approximating curve from the actual facilities being less than 100 microseconds over this frequency range. The approximating curves can then be specified by the three parameters b, τ and θ (or when the bit speed of the data system is given by b, δ and θ). These values are given in Table I, for the cases of a data system operating with an 1800-cycle carrier, and at bit speeds of 1200 and 1500 bits per second. The resulting aperture values and transmission impairment figures produced are also given in Table I. Aperture values were determined by direct computer computation; these values can also be determined to a good approximation by interpolation from Figs. 10 and 11.

The method shown here for determining the impairment values in

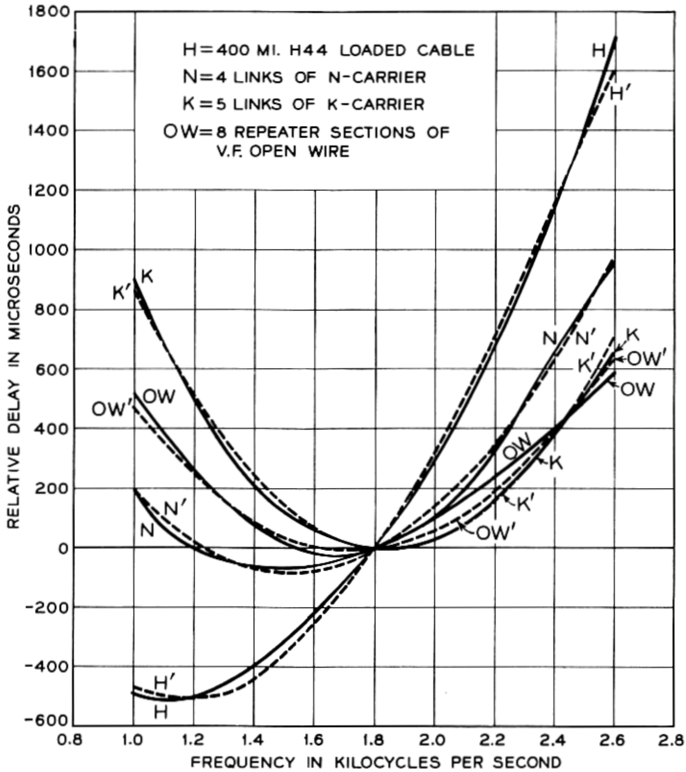


Fig. 24 – Delay characteristics of four message facilities.

Table I illustrates how data system performance can be established for actual transmission facilities from the results of the computer simulation. This procedure can also be used to determine how performance varies with some of the parameters that are important in the design and engineering of data systems, such as bit speed, location of carrier frequency and length and type of transmission facility. For example, as the bit speed is increased from 1200 to 1500 bits per second, the impairments of the facilities are increased by the amount shown. (The loaded cable changes from marginal operation to complete failure.)

The poor performance of the loaded cable facility is due to the fact that the location of the carrier frequency results in an asymmetrical disposition of the delay curve about the carrier, resulting in a considerable amount of quadrature component. If the location of the carrier were changed from 1800 to 1500 cycles per second, the angle θ would be

TABLE I

Facility	b, radians	τ, μ sec	θ, degrees	1200 bits per second			1500 bits per second		
				δ, bit lengths	A*	Transmission impairment, db	δ, bit lengths	A*	Transmission impairment, db
400-mile H44 loaded cable	9.26	194	43.8	0.233	0.092	20.7	0.292	—	—
4 links of N carrier	11.8	163	15.8	0.196	0.60	4.4	0.244	0.38	8.4
5 links of K carrier	9.89	190	−3.1	0.228	0.67	3.5	0.285	0.50	6.0
8-repeater sections of V.F. open wire	11.2	161	3.6	0.193	0.76	2.4	0.242	0.61	4.7

* These aperture values were determined by direct computation. They can also be obtained by interpolation from Fig. 10 and Fig. 11, with the exception of the loaded cable (which has an aperture value of either zero or very nearly zero).

changed from 43.8° to 22.8°. For a bit speed of 1200 bits per second, this gives an aperture value of 0.41 (by interpolation from Figs. 10 and 11) and a resulting transmission impairment of 7.8 db.

Performance for many other transmission conditions of interest can be determined with similar ease by this same procedure.

X. CONCLUSIONS

It is apparent that the data system simulation technique that has been developed and used here for evaluating performance is a powerful and valuable tool. Wherever possible, experimental and theoretical verifications of the results obtained from this simulation have been made. These tests indicate that this method is valid and accurate. The performance curves given have been used to illustrate the ease and effectiveness with which many of the complex problems relating to delay distortion can be answered in a rather general way.

Several uses can be made of this simulation technique. First, it provides a means of obtaining quantitative answers to a broad category of problems relating to data transmission, such as the influence on performance of the variation of bit speed, the location of carrier frequency and the length and type of transmission facility used. In addition, a performance model based on a sinusoidal phase characteristic and normalized with respect to bit speed has been produced for the data system under consideration. The surface of this model describes the boundary (in terms of variation of phase distortion) within which the system can operate. Similar models can be established for other types of data systems to provide a means for comparing the performance on a much broader basis than is presently possible.

The investigation presented here was directed to the problem of determining what influence phase (or delay) distortion has on data system performance. It is clear that the simulation method developed could be extended to include other problems of data transmission, such as a general study of the influence of the attenuation characteristics, the problem of recovering timing and the influence of noise on system transmission. Additional work in this general area is now being carried out.

XI. ACKNOWLEDGMENTS

The author wishes to acknowledge the assistance received in carrying out this work. The simulation program for this study was worked out in collaboration with Miss Geraldine Nelson, who carried out the many programming details. Much of the actual computational work on the computer and many other details involved in the preparation of this material were carried out by Miss Juanita Sacks.

REFERENCES

1. Fowler, A. D., and Gibby, R. A., Assessment of Effects of Delay Distortion in Data Systems, Comm. & Electronics, no. 40, 1959, p. 918.
2. Sunde, E. D., Theoretical Fundamentals of Pulse Transmission, B.S.T.J., **33,** 1954, p. 721; 987.
3. Sunde, E. D., Ideal Binary Pulse Transmission by AM and FM, B.S.T.J., **38,** 1959, p. 1357.
4. Oliver, B. M., Pierce, J. R. and Shannon, C. E., The Philosophy of PCM, Proc. I.R.E., **36,** 1948, p. 1324.
5. Mertz, P., Transmission Line Characteristics and Effects on Pulse Transmission, in Proc. Symp. on Information Networks, Polytechnic Inst. of Brooklyn, Brooklyn, N. Y., 1954.