The Effects of Digital Errors on PCM Transmission of Compandored Speech

By I. DOSTIS

(Manuscript received August 10, 1965)

The recent interest in the use of pulse code modulation (PCM) for the transmission of speech has made it desirable to determine the effects of digital line errors on certain codes. Improved performance for low-level talkers has been generally obtained by the use of instantaneous compressors and expandors (compandors) in order to avoid using a large number of digits in the PCM coder-decoder (codec). A comparison of the effects of digital errors for a logarithmically compandored system transmitting speech is made on the basis of mean square distortion power and the distribution of error magnitude. Results are obtained for the binary, Gray (reflected binary), and folded binary codes. The results indicate that the binary and Gray codes have a "click" as well as a "noise" component of distortion while the folded binary code produces only noise at low-talker levels. "Clicks" have been defined as errors which have amplitudes greater than half the full range amplitude; the remaining errors are considered "noise". For the folded binary code, the most significant digit gives polarity information; the remaining digits represent the signal magnitude in binary code.

I. INTRODUCTION

A comparison of the effects of digital errors for a logarithmically compandored[1,2] system transmitting speech by PCM is made on the basis of mean square error power and the distribution of the error magnitude. A block diagram of the system considered in the analysis is shown in Fig. 1. The calculations are based on the assumption of at most one error per 8-digit word. The results obtained for binary and Gray codes indicate that there is a "click" as well as "noise" component of distortion at low talker levels. "Clicks" have been defined as errors which have amplitudes greater than half the full range, the remaining error

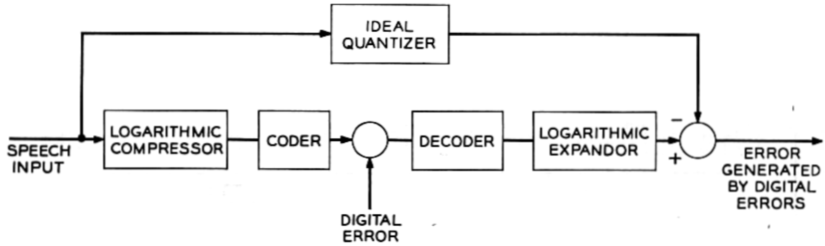

Fig. 1 — A block diagram of the system under consideration.

amplitudes are considered "noise". A folded binary code (see Fig. 2) essentially produces only the noise component at low talker levels. Low talker levels were considered to be 50 db below full load sine wave (50 db BFLSW) for all the cases calculated.

The results obtained for the folded binary code indicate that this code does not have a click problem at low talker levels when compared to the binary and the Gray codes. A question is raised, however, concerning the subjective effect at low levels of the noise components of distortion. If the noise is subjectively equivalent to quantizing error, an error rate of 10^{-6} will cause a 3-db impairment in the system performance. This occurs for speaker levels between 30 and 50 db BFLSW in a system using a logarithmic compression factor[1,2] (μ) of 100 and the folded binary code.

The conclusions obtained from the computed results for logarithmic companding are:

(*i*) The folded binary code produces fewer clicks than either the Gray or binary code at low speaker levels.

(*ii*) The Gray, binary, and folded binary click and noise components are comparable at higher speaker levels.

The combined subjective effect of noise and clicks occurring simultaneously has not been evaluated.

II. RESULTS

The results obtained for the probability of a given code word occurring are plotted in Fig. 3 for logarithmic compression factors[1,2] of 50, 100, and 200. The input signal was assumed to be Laplace distributed.[3,4,5] The figure indicates that compression has a major effect on the probability of a given code word occurring. The tendency of the compressor is to assure a more uniform distribution of the signal across the lower code levels for low and average level talkers. One notes that for a μ of

FOLDED BINARY CODE	BINARY CODE	GRAY CODE
11111	11111	10000
11110	11110	10001
11101	11101	10011
11100	11100	10010
11011	11011	10110
11010	11010	10111
11001	11001	10101
11000	11000	10100
10111	10111	11100
10110	10110	11101
10101	10101	11111
10100	10100	11110
10011	10011	11010
10010	10010	11011
10001	10001	11001
10000	10000	11000
00000	01111	01000
00001	01110	01001
00010	01101	01011
00011	01100	01010
00100	01011	01110
00101	01010	01111
00110	01001	01101
00111	01000	01100
01000	00111	00100
01001	00110	00101
01010	00101	00111
01011	00100	00110
01100	00011	00010
01101	00010	00011
01110	00001	00001
01111	00000	00000

Fig. 2 — 5-digit codes.

100 the average talker (30 db BFLSW) has a nearly uniform distribution over the range $-\frac{1}{4}$ to $+\frac{1}{4}$ full range.

The mean square value of the distortion caused by digital line errors is compared to 8-digit quantizing error power[1,2] for a 10^{-6} error rate, a 2-volt peak-to-peak input amplitude and compression factors (μ) of 50, 100, and 200 in Figs. 4 and 5. Altering the error rate simply results in a scale change for the line error distortion power. The results plotted in Fig. 4 include all error amplitudes whereas errors greater than half the full range amplitude (clicks) have been removed for the $\mu = 100$ case in Fig. 5. Preliminary results of subjective tests indicate that clicks are less objectionable than quantizing noise. The figures indicate the folded binary code yields superior click performance at low talker levels compared to the Gray or binary codes. The noise or non-click component for the folded binary is greater than that produced by the Gray code at low levels. For an error rate of 10^{-6} and 8-digit quantizing, the line error distortion power and quantizing error power are equal for the folded binary code in the 30 db to 50 db BFLSW range. If the line error distortion is subjectively equivalent to quantizing error, the system per-

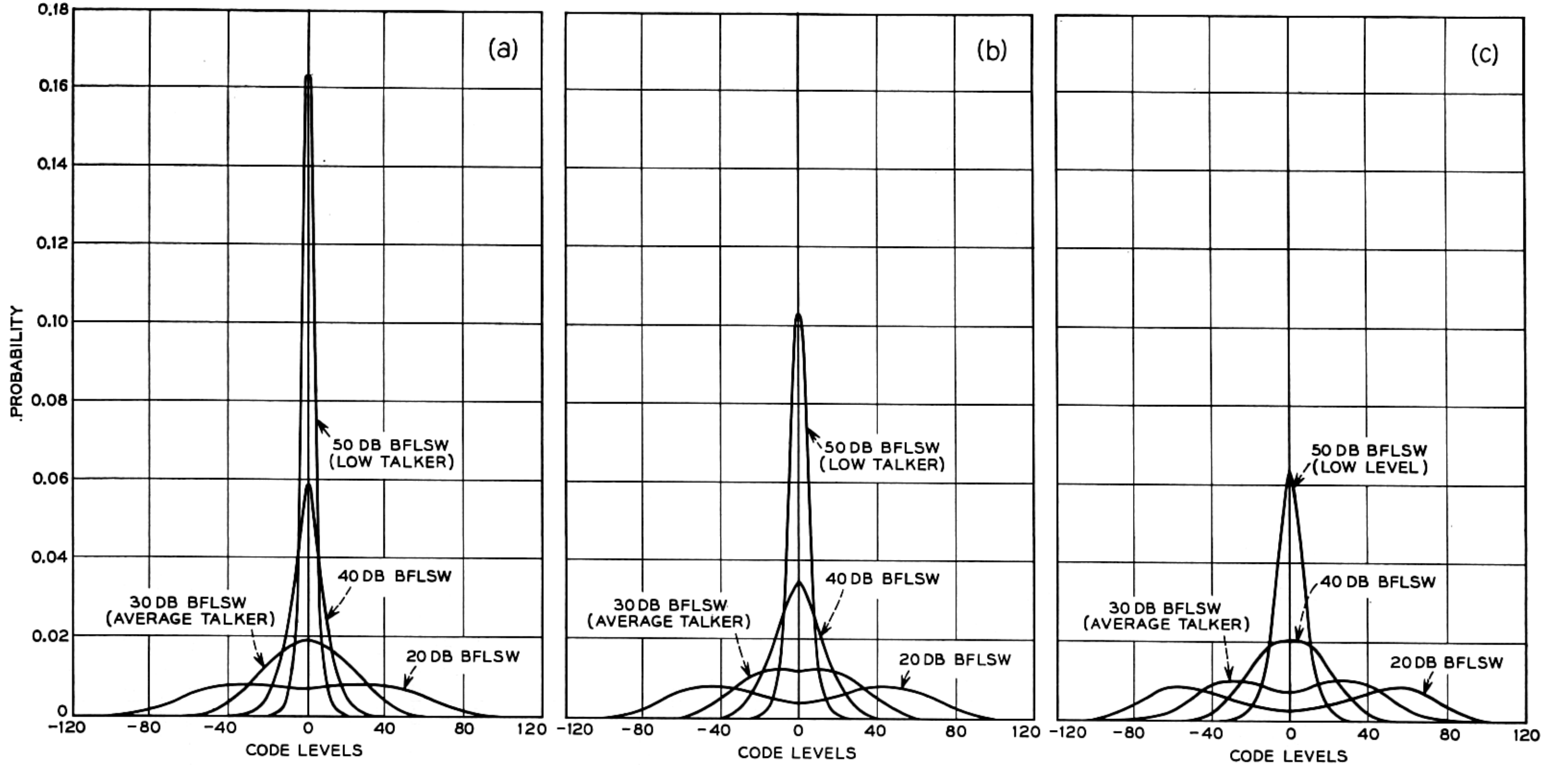

Fig. 3 — Word probability—8-digit coder Laplace distribution. (a) μ = 50 logarithmic compression. (b) μ = 100 logarithmic compression. (c) = 200 logarithmic compression.

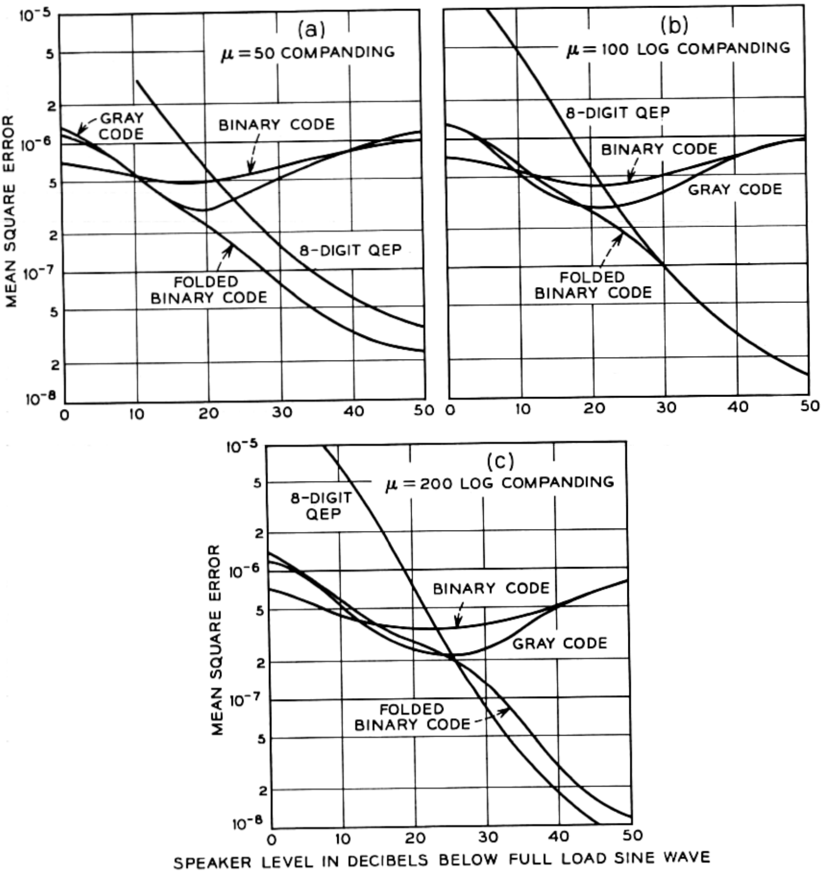

Fig. 4 — Mean square error vs speaker level-8 digit coder — 2-volt peak-to-peak full load amplitude. (a) $\mu = 50$ log companding. (b) $\mu = 100$ log companding. (c) $\mu = 200$ log companding.

formance will be degraded by 3 db. This result applies for a logarithmic compression of 100.

We conclude that the essential low-level problem is noise for the folded binary and clicks for the Gray and binary codes.

The probability of errors of a specified magnitude occurring are plotted in Figs. 6 through 9 for the binary, folded binary, and Gray codes. A clear demarcation between errors greater than and less than half the full amplitude is indicated in Figs. 6, 7, and 8 for low levels. These results led to the definition that errors greater than half the full load would

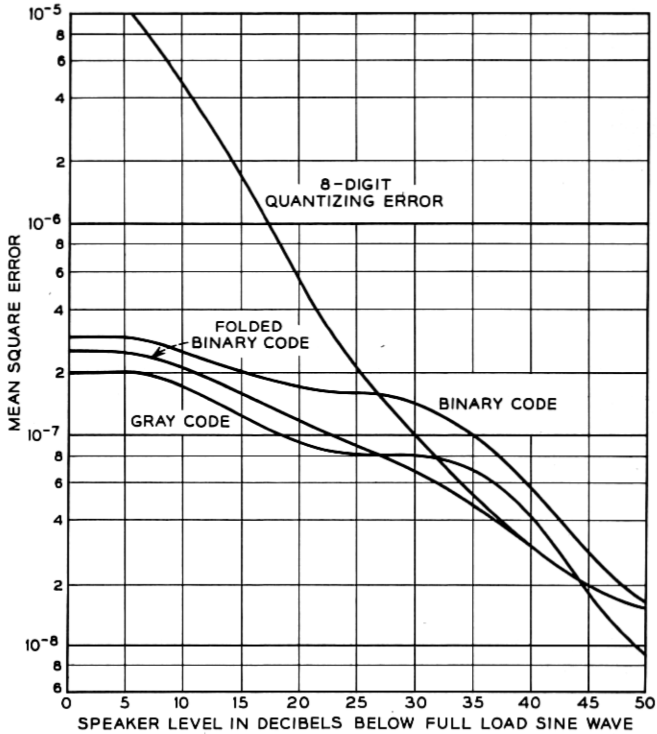

Fig. 5 — 8-digit codes (compression factor, $\mu = 100$ error rate 10^{-6}, clicks removed — 2-volt peak-to-peak full load amplitude).

be called clicks. The dividing line between clicks and noise is not clear for average and high-level speakers.

It is apparent that the folded binary code produces virtually no clicks at low levels. This leads to the conclusion that noise rather than clicks will be the dominant low-level problem in transmission of folded binary, whereas clicks are the problem for Gray and binary.

The mean square error caused by specified digit errors for various speaker levels are plotted in Fig. 10(a), (b) and (c). The results indicate that the largest mean square error for the folded binary code is caused by digit 2 errors. The largest mean square error for Gray and straight binary depends to some extent on the speaker level. The largest error for low and average levels occurs in the 1st digit for binary and the 2nd digit for Gray.

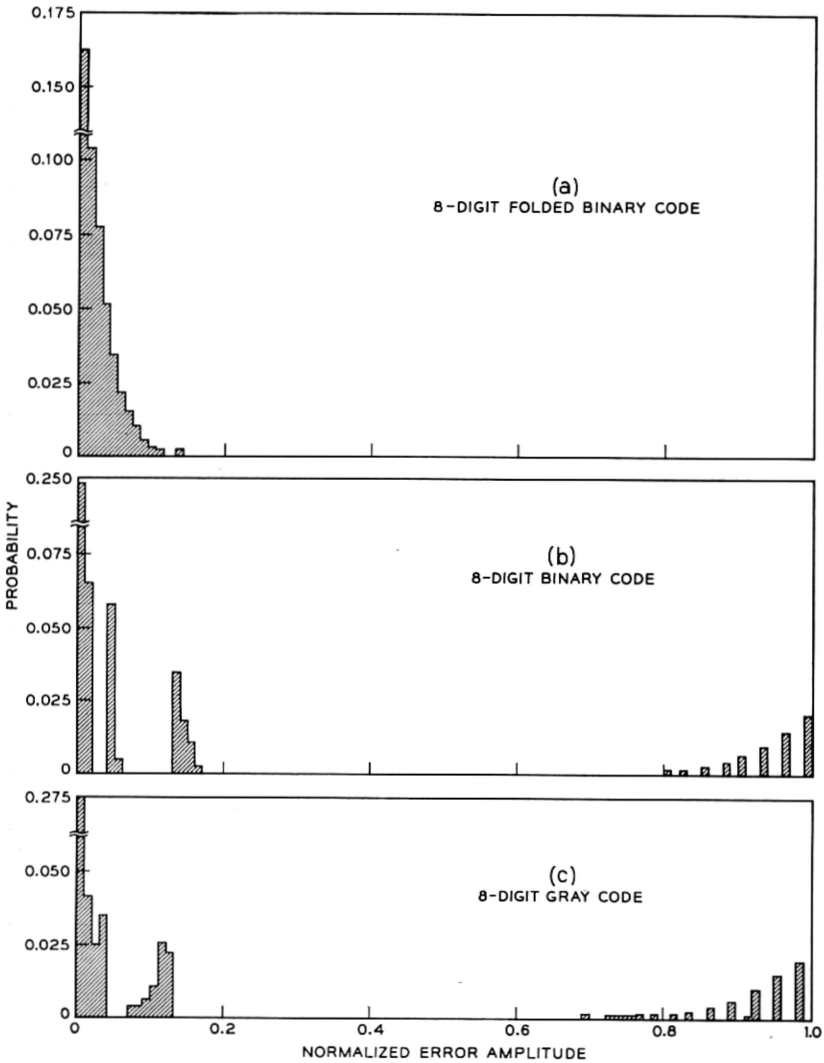

Fig. 6 — Probability of error (speaker level 50 db BFLSW, $\mu = 50$ log companding, symmetric for negative amplitudes, low-level talker). (a) 8-digit folded binary code. (b) 8-digit binary code. (c) 8-digit Gray code.

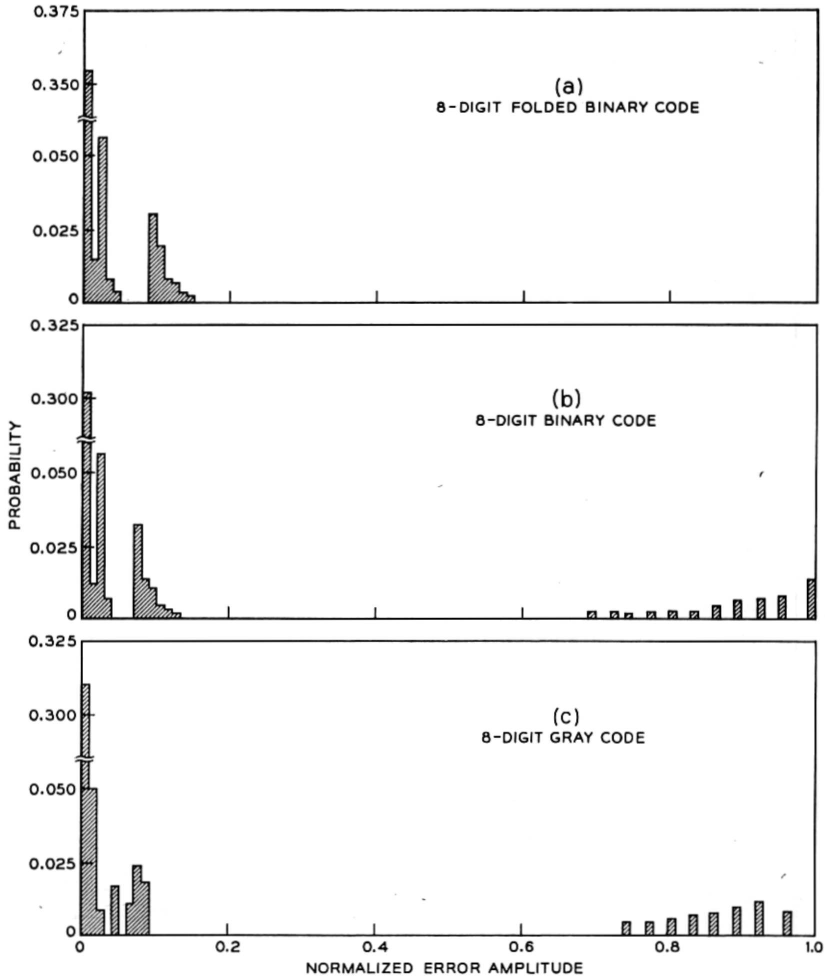

Fig. 7 — Probability of error (speaker level 50 db BFLSW, $\mu = 100$ log compandor, symmetric for negative amplitudes, low-level talker). (a) 8-digit folded binary code. (b) 8-digit binary code. (c) 8-digit Gray code.

III. ASSUMPTIONS AND PROCEDURES

3.1 *Assumptions*

The results obtained were calculated using the following assumptions:

(*i*) The terminals of the transmission system including the coder and decoder are ideal.

(*ii*) All digital errors are introduced in the transmission medium.

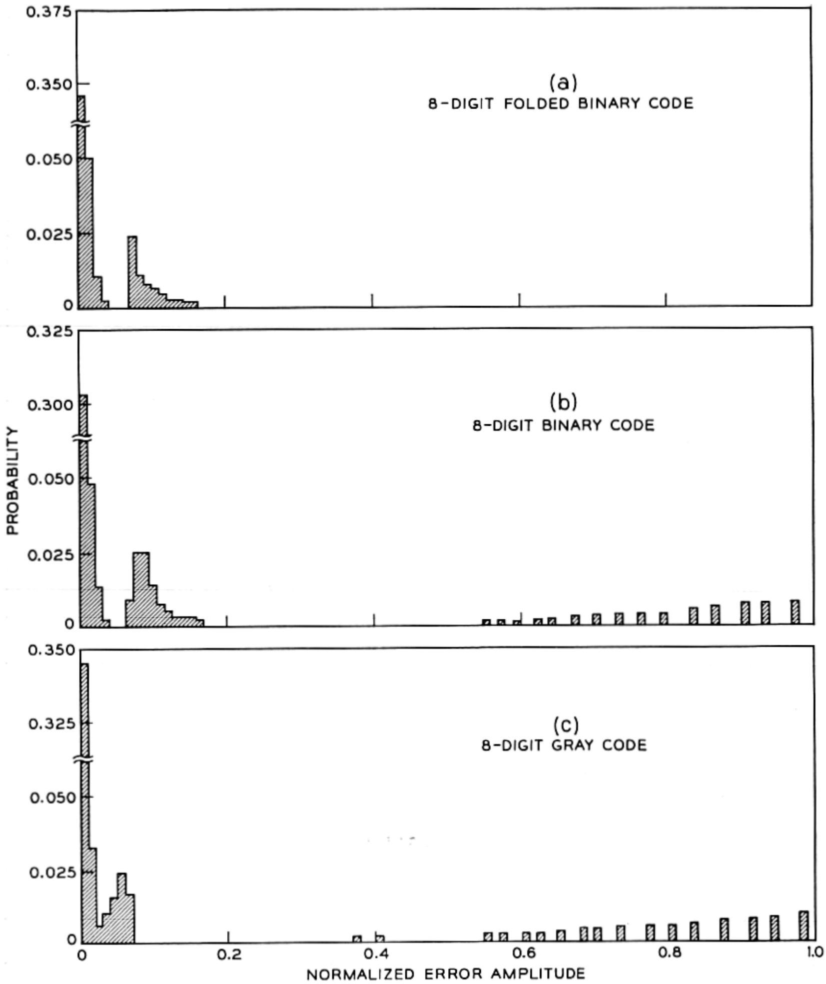

Fig. 8 — Probability of error (speaker level 50 db BFLSW, $\mu = 200$ log companding, symmetric for negative amplitudes, low-level talker). (a) 8-digit folded binary code. (b) 8-digit binary code. (c) 8-digit Gray code.

(*iii*) No more than one digital error occurs in any given code word.

(*iv*) The input signal is speech and is assumed to be Laplace distributed.[3,4,5]

(*v*) The mean value of the input signal corresponds to the mid-range of the coder-decoder.

(*vi*) Logarithmic companding[1,2] with $\mu = 50$, 100, and 200.

(*vii*) The effects of overload have been ignored.

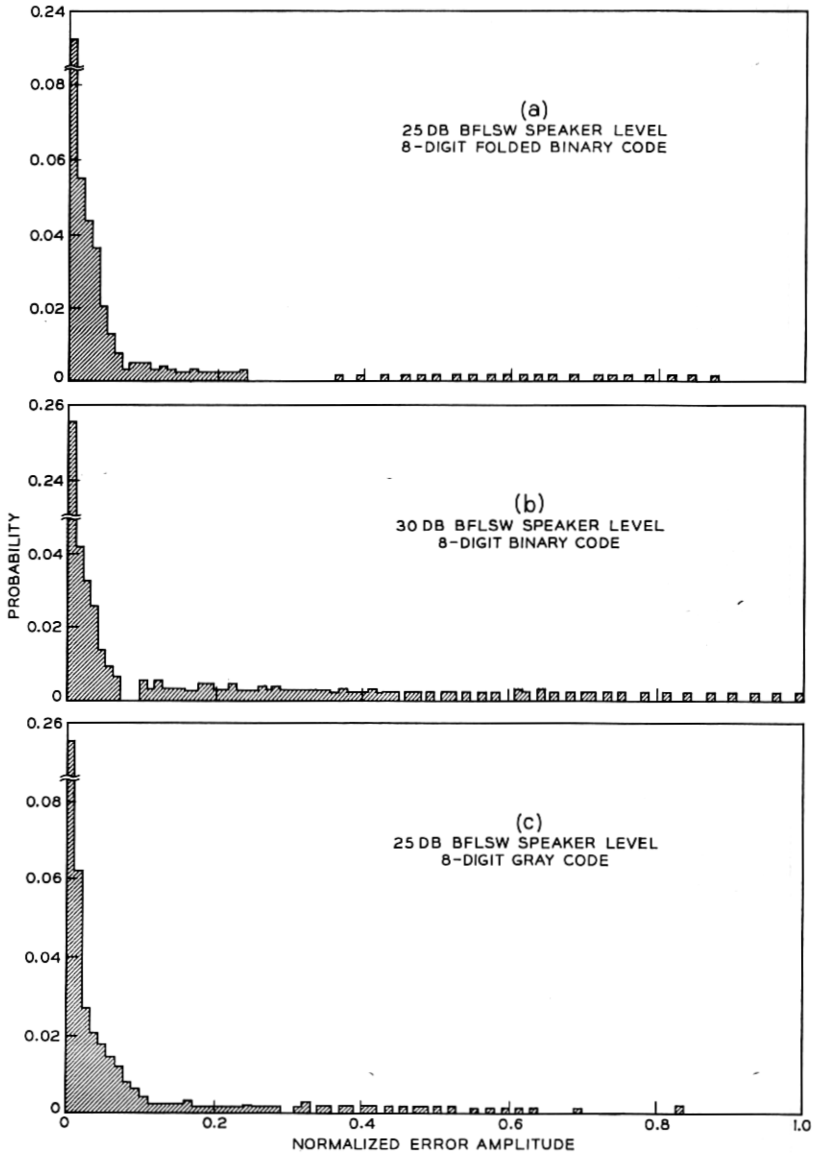

Fig. 9 — Probability of error — $\mu = 100$ log compandor symmetric for negative amplitudes average talker. (a) 8-digit folded binary code (speaker level 25 db BFLSW). (b) 8-digit binary code (speaker level 30 db BFLSW). (c) 8-digit Gray code (speaker level 25 db BFLSW).

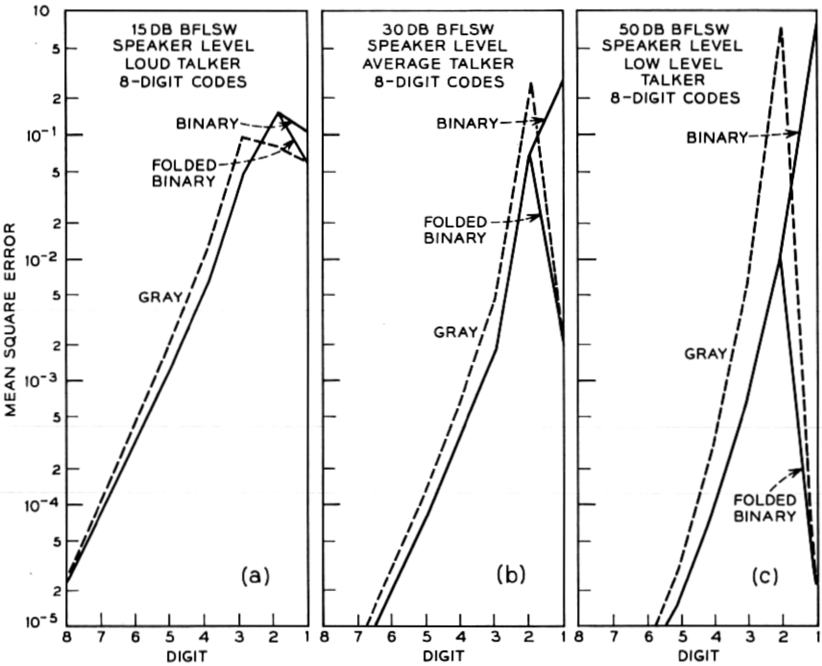

Fig. 10 — Mean square error vs digit — 2-volt peak-to-peak signal — (a) 8-digit codes (speaker level 15 db BFLSW — louder talker). (b) 8-digit codes (speaker level 30 db BFLSW — average talker). (c) 8-digit codes (speaker level 50 db BFLSW — low-level talker).

3.2 *Procedure*

3.2.1 *Mean Square Error*

A block diagram of the system assumed for the calculations performed is given in Fig. 1. The results were calculated using a probability of error of 10^{-6}. This error rate is sufficiently low to justify assumption (*iii*) above. The probability of a given word occurring on the line depends on the compression factor of the coder and the input statistics. The mean square error caused by digital line errors for the assumptions made is*

$$\overline{\eta^2} = \sum_{J=1}^{L} \sum_{\substack{K=-M \\ K\neq 0}}^{M} \epsilon_{KJ}^{\;2} P_K P_J P_\varepsilon \tag{1}$$

* Formulations have been presented in the literature, different from the above, which may be more suitable for determining an optimal nonredundant code.[6, 7]

where

ϵ_{KJ} is the amplitude of the error for an error in digit J of word K

P_K is the probability of a particular code word occurring

P_J is the probability of a particular digit of a given word being chosen*

P_ε is the probability of a digital error (error rate)

L is the number of digits in the coder

$M = 2^{(L-1)}$

Using the assumptions listed above, we reduce (1) to

$$\overline{\eta^2} = P_\varepsilon \sum_{J=1}^{L} \eta_J^2 P_J \tag{2}$$

where η_J^2 is the mean square error for a given digit and is defined in (3) as

$$\eta_J^2 \triangleq 2 \sum_{K=1}^{M} \epsilon_{KJ}^2 P_K . \tag{3}$$

The calculation of the values given in (2) and (3) are straightforward but quite tedious in the case of a compandored system with varying speech levels. The problem was programmed for the IBM 7094 computer and the results of the computation for (2) and (3) presented in Figs. 4, 5, and 10(a), (b), and (c).

3.2.2 *Word Probability*†

The probability of a given word K occurring for a given speaker level was calculated directly from (4)[2,3,4,5]

$$P_K = C \int_{B_K}^{B_{K+1}} \exp\left(-2C \mid x \mid\right) dx \tag{4}$$

where

$$C = \frac{RMS \text{ VALUE FULL LOAD SINE WAVE}}{RMS \text{ VALUE SIGNAL}} = \frac{E}{\sqrt{2\sigma^2}}$$

and the B_K are normalized coder levels for a specified compression characteristic. Normalization is with respect to full load.

The value of P_K given above is for a zero mean input signal. When the input signal and the coder do not have a zero mean, a shift can be performed easily. Using the assumptions given, the following result for P_K

* $P_J = 1/L$ for the case considered.

† An analytic approach is presented in Appendix A.

was obtained,

$$P_K = \tfrac{1}{2}\{\exp(-2CB_K) - \exp(-2CB_{K+1})\}. \quad (5)$$

The word probability for speaker levels of 20 db, 30 db, 40 db, and 50 db BFLSW have been plotted in Fig. 3 for μ = 50, 100, and 200.

3.2.3 *Line Errors*

The introduction of line errors was accomplished by complementing a given digit in a specific code word and determining the new word obtained. The results obtained obviously depend on the code being considered. A transition matrix was produced for each particular code considered, i.e., Gray, binary, and folded binary.

The rows of the transition matrix represent the originally transmitted code word and the columns the digit in error from most to least significant. The entry in column J and row K is the new word R produced by an error in digit J of word K. The transition matrix for a 3-digit binary code is given in Table I. This matrix is then used to determine the error caused by a specified digital error. The operation required to determine ϵ_{KJ} is decoding. If the output voltages corresponding to the code words K and R are denoted by D_K and D_R respectively, we obtain the following expression for ϵ_{KJ} ,

$$\epsilon_{KJ} = D_R - D_K. \quad (6)$$

One should note that the digit in error was specified as digit J in the above example.

TABLE I — TRANSITION MATRIX FOR 3-DIGIT BINARY CODE

Binary Word Representation	Word-in-Error	Digit-in-Error 1	Digit-in-Error 2	Digit-in-Error 3
000	0	4	2	1
001	1	5	3	0
010	2	6	0	3
011	3	7	1	2
100	4	0	6	5
101	5	1	7	4
110	6	2	4	7
111	7	3	5	6

The error matrix having elements ϵ_{KJ} was then used to calculate the mean square error for a given digit, the mean square error for a given error rate with and without the "click" component and the distribution of error amplitudes for each code considered. Criteria other than mean square can be easily adapted to the computation.

IV. REMARKS AND CONCLUSIONS

The results obtained indicate the folded binary code yields better click performance for a given probability of digital error when compared to Gray and binary transmission. The effects of clicks should be virtually nonexistent at low talker levels in the folded binary code when compared to binary and Gray code. The question arises, however, concerning the subjective effect of the line error distortion when compared to quantizing error since the clicks at low levels have been suppressed. A 3-db reduction in S/N will occur for an 8-digit, $\mu = 100$, 10^{-6} error rate system transmitting folded binary for speaker levels between 30 db and 50 db BFLSW. The system performance will also be degraded by 3 db if line error distortion is subjectively equivalent to quantizing error.

Further work in evaluating the subjective effect of clicks and noise occurring simultaneously is required.

V. ACKNOWLEDGMENT

The author wishes to acknowledge the invaluable advice and aid on programming given by Miss E. G. Cheathem.

APPENDIX

Calculation of Density Function and Word Probability After Compression

A.1 *Word Probability*

An analytic expression for the probability density function at the output of a logarithmic compressor is derived. The density function is derived for a speech input which can be approximated by a Laplacian distribution.[2,3,4,5] The density function used is given in (7) as

$$p(e_i) = C \exp(-2C \mid e_i \mid), \tag{7}$$

where, e_i is the input signal and C is defined by

$$C = \frac{RMS \text{ VALUE OF FULL-LOAD SINE WAVE}}{RMS \text{ VALUE OF SIGNAL}}.$$

Normalizing the result to full-load sine wave of unit amplitude results in a value $1/\sqrt{2\sigma_i^2}$ for C.

The characteristic of a logarithmic compressor normalized to a full-load unit sine wave is given by[2]

$$v_0 = \frac{\log(1+\mu e_i)}{\log(1+\mu)} \quad \text{for} \quad 0 \leqq e_i \leqq 1, \tag{8}$$

and

$$v_0 = -\frac{\log(1-\mu e_i)}{\log(1+\mu)} \quad \text{for} \quad -1 \leqq e_i \leqq 0. \tag{8b}$$

The output density function can be determined directly by transforming the input, i.e.,

$$p(v_0) = p(e_i)\left|\frac{de_i}{dv_0}\right|. \tag{9}$$

Performing the operation defined in (9) yields,

$$p(v_0) = \frac{C\log(1+\mu)}{\mu}(1+\mu)^{v_0}\exp\left[-\frac{2C}{\mu}\{(1+\mu)^{v_0}-1\}\right] \tag{10}$$
$$\text{for} \quad \mu > 0 \quad \text{and} \quad v_0 \geqq 0$$

and $p(-v_0) = p(v_0)$. The expression given in (10) is the probability density function of the compressor output.

Determining the word probability for a given code word K involves integrating the derived density between the lower and upper transition values for which this code word is emitted. For the case of a large number of digits (fine quantizing), an approximate expression can be derived. The step size for an n digit quantizer having a peak-to-peak amplitude of 2 is $1/2^{(n-1)}$. The value of the word probability $P(K)$ for word K is given by;

$$P(K) = \frac{C\log(1+\mu)}{\mu}\int_{(K-1)/2^{n-1}}^{K/2^{n-1}}(1+\mu)^{v_0} \cdot\exp\left[-\frac{2C}{\mu}\{(1+\mu)^{v_0}-1\}\right]dv_0 \tag{11}$$
$$\text{for} \quad K \geqq 1 \quad \text{and} \quad v_0 \geqq 0.$$

We can approximate this result for the case of fine step sizes by

$$P(K) \cong \frac{C \log (1 + \mu)}{2^{n-1}\mu} (1 + \mu) e^{(K-\frac{1}{2})/2^{n-1}} \cdot \exp\left[-\frac{2C}{\mu}\left\{(1 + \mu) e^{(K-\frac{1}{2})/2^{n-1}} - 1\right\}\right] \tag{12}$$

$$\text{for} \quad K \geqq 1 \quad \text{and} \quad v_0 \geqq 0,$$

or

$$P(K) \cong \frac{C \log (1 + \mu)}{2^{n-1}\mu} (1 + \mu) e^{(2K-1)/2^n} \cdot \exp\left[-\frac{2C}{\mu}\left\{(1 + \mu) e^{(2K-1)/2^n} - 1\right\}\right]. \tag{13}$$

For the case of $n = 8$, $\mu = 100$, $K = 1$, $C = 100$ (40 db BFLSW)

$$P(1) \cong \frac{4.61}{128} (1.002)e^{-2(0.002)} \cong \frac{4.61}{128} (1.002) \cong 0.036. \tag{14}$$

This agrees favorably with the value computed from the exact equation for this case of 0.0351.

A.2 *Determination of the Most Probable Code*

The most probable code level will correspond approximately to the peak(s) of the density function when the quantizing steps are small. For this condition we obtain,

$$\frac{dp(v_0)}{dv_0} = 0 = \frac{C \ln^2 (1 + \mu)}{\mu} (1 + \mu)^{v_{0m}} \exp\left[-\frac{2C}{\mu} \cdot [(1 + \mu)^{v_{0m}} - 1]\right] \times \left\{1 - \frac{2C}{\mu} (1 + \mu)^{v_{0m}}\right\} \tag{15}$$

$$\text{for} \quad v_{0m} \geqq 0$$

where v_{0m} is the peak value of v_0. Hence we have,

$$\frac{\mu}{2C} = (1 + \mu)^{v_{0m}}$$

or

$$v_{0m} = \frac{\log (\mu/2C)}{\log (1 + \mu)} \quad \text{for} \quad v_{0m} \geqq 0. \tag{16}$$

The results obtained using (16) compare favorably with those deter-

TABLE II — COMPARISON OF COMPUTED AND ANALYTICALLY OBTAINED VALUES

C (db BFLSW)		0	5.0	10.	15.	20.	25.	30.	35.	40.
v_{0m}	Theoretical	0.85	0.72	0.599	0.474	0.349	0.224	0.099	—	—
	Computer	0.85	0.72	0.60	0.47	0.35	0.22	0.10	0	0

mined on the computer as indicated in Table II. The value given for v_{0m} is the decimal fraction of full load code levels.

REFERENCES

1. Members of the Technical Staff of Bell Telephone Laboratories, Transmission Systems for Communications, Bell Telephone Laboratories, 1964.
2. Smith, B., Instantaneous Companding of Quantizing Signals, B.S.T.J., *36,* May, 1957, p. 653.
3. Davenport, W. B., An Experimental Study of Speech Wave Probability Distributions, J. Acoust. Soc. Amer., *24,* 1952, p. 390.
4. Purton, R. F., A Survey of Telephone Speech Signal Statistics and Their Significance in the Choice of a PCM Companding Law, IEE, January, 1952, p. 60.
5. Cramer, H., *Mathematical Methods of Statistics,* Princeton University Press, 1946.
6. Bado, J., Coding for Least RMS Error in Binary PCM Channels, Wescon, TR62-3, May, 1962.
7. Velichkin, A. I. and Grushko, I. I., Optimum Nonredundant Codes, Bull. Acad. Sci. USSR, Tech. Sci. — Energetics and Automation, *6,* 1962.