Signal-to-Noise and Idle Channel Performance of Differential Pulse Code Modulation Systems – Particular Applications to Voice Signals

By R. A. McDONALD

(Manuscript received April 25, 1966)

Analysis leading to a figure of merit for differential pulse code modulation (DPCM) systems with linear feedback networks is presented. It is shown that the figure of merit can be optimized. Simple DPCM has a 6-dB advantage in signal/quantizing noise ratio over pulse code modulation (PCM) for speech. Optimization yields at most 4 dB more. Computer simulation of the system using actual speech samples leads to data supporting the figure of merit as a useful measure of performance for DPCM systems with four digits or more. The simulation also provides data on the error spectrum as a function of quantizer loading and on the probability density of the quantizer input as a function of loading. Performance of the optimum system as a function of increasing feedback network complexity is also shown.

Idle channel performance of a particular system is analyzed, indicating the presence of inband oscillations in many cases. The best quantizer bias from the point of view of idle channel performance is found. The level of idle channel noise in DPCM is shown to be approximately equivalent to that in PCM.

I. INTRODUCTION

Digital techniques for transmitting analog signals such as voice, television, or facsimile have been known for a long time, and technology has reached the point where some of these methods are commercially feasible. Since cost is a critical factor in determining applicability of these systems, there has been from the beginning an attempt to improve the efficiency of analog-to-digital conversion by reducing the bit rate required for a given accuracy of reproduction. One of the principal methods involves removing inherent signal redundancy through the use

of feedback around the quantizer, and has led to a wide variety of schemes which may all be classed as differential systems. The origins of differential pulse code modulation (DPCM) stem from patents by the N. V. Phillips Company in 1951[1] and by C. C. Cutler in 1952.[2] The ideas also appear in several papers of about that time.[3,4,5] Since that time, considerable research and development work has been reported, and one has only to look at our reference list, which is certainly not complete, to be convinced that the problems have been examined at great length.

The work to be reported here is the result of a fairly extensive investigation of the potential advantages and pitfalls of voice transmission by practical DPCM systems and by alternatives which are essentially variations on the basic theme of PCM or DPCM. The problems are handled analytically as far as is possible. But rather than dilute the result by using an over-simplified model for the input signal, a computer simulation is used to advantage in more than one place. Optimum as well as simple suboptimum systems are considered.

Some of the analysis reported here is applicable to systems other than ones for voice transmission, but the one application is considered throughout since it provided the motivation for the entire project. A similar project was carried out independently by J. B. O'Neal[6] of Bell Telephone Laboratories, but with special consideration given to television signals. The special considerations introduced by the speech signal include the need to investigate performance for a wide range of input signal levels and a need to investigate idle channel performance.

There is considerable overlap with the work of Nitadori.[7] The work in Sections III, IV, and V was influenced heavily by his original work, but is based on broader assumptions. The validity of our assumptions is checked by means of the computer simulation described in Section VI. This simulation may also be construed as a check on the assumptions used by Nitadori and others. Our optimum linear network is developed from a viewpoint different from that of Nitadori.

The analytical results are also essentially parallel to those of Oliver[8] although his work is not directly applicable to the differential systems investigated here.

II. SYSTEM DESCRIPTION

The pulse code modulation (PCM) system shown in Fig. 1 will serve as the basis of comparison for all the others. The input and output shown are sequences of samples, since all the systems under consideration will require sampling. In the PCM system, one high-speed quantizer and

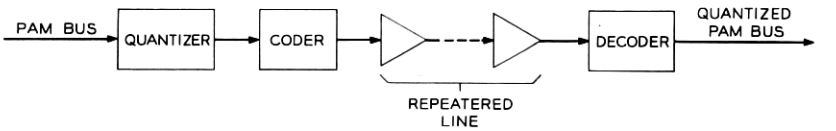

Fig. 1 — Basic PCM system.

coder can be shared among many channels by time division multiplexing the pulses representing the analog samples. There will be greater difficulty multiplexing the inputs to the differential systems, thereby introducing higher costs associated with the terminal portion of the system. It is this fact which controls the economics of the system. If, under an equiperformance criterion, the differential system requires fewer digits per unit time on the transmission portion of the system, but requires a more expensive terminal, there will be a net advantage whenever the repeatered line costs are a large enough portion of the total costs (i.e., long haul systems). It will be assumed throughout the paper that the controlling source of impairment is the quantization noise introduced by the quantizer with a finite number of steps of finite size. The overload noise is hence included here. The measure of performance will be the ratio of the mean squared signal to mean squared noise, or in the case of the idle channel, the mean squared noise alone.

The basic DPCM system which we shall consider is shown in Fig. 2. Without going into the details of operation of the system at this point, we note that the diagram actually represents a wide class of systems, different members of which are obtained with different prediction networks. In actual fact, we shall be restricted in our investigations to linear prediction networks, but this still leaves a rather broad class of systems.

The configuration of Fig. 2 bears a resemblance to several somewhat different systems described in the literature. We refer particularly to the work of Kimme,[9] Kimme and Kuo,[10] and of Spang and Schultheiss.[11]

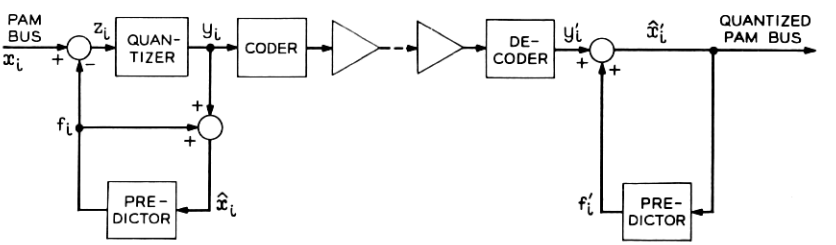

Fig. 2 — Basic DPCM system.

These systems involve quantization noise feedback rather than predictive feedback, and are thought of as shaping the spectrum of the noise rather than removing signal redundancy. It has been shown by Kimme[9] that there is an equivalence between a noise feedback system with predistortion and post-distortion filters and a DPCM system with predistortion and post-distortion filters. That is, given one configuration, there is a transformation which yields transfer functions for the blocks of the other configuration such that the performances are identical. However, we have found the predictive feedback point of view useful in its own right.

III. SIGNAL-TO-NOISE RATIO IMPROVEMENT

Notation needed for the algebraic analysis of DPCM appears in Fig. 2. A stochastic model is assumed for the speech samples, x_i, with a symmetrical zero-mean distribution not dependent on i. The primed quantities on the receiving end differ from the unprimed quantities only when the repeatered line introduces digital errors. For the most part, we shall ignore digital errors, and deal only with the unprimed quantities.

First, the quantizing error is defined as

$$e_i = z_i - y_i. \tag{1}$$

Note that in (1) and in the equations to follow, the index i, which denotes the time order of the samples, decreases to indicate samples further in the past. Unless otherwise stated, it is meant that the equations hold for all integers i.

The other fundamental relationships indicated by the block diagram are

$$z_i = x_i - f_i \tag{2}$$

$$\hat{x}_i = f_i + y_i \tag{3}$$

$$f_i = \sum_{j=1}^{\infty} h_j \hat{x}_{i-j}, \tag{4}$$

where the coefficients h_j are characteristics of the prediction filter. Note that in the last equation only the samples of the output of the assumed linear filter are indicated. This filter may have a continuous time response so long as the samples conform to (4). The absence of an h_0 term in (4) implies the presence of some delay around the loop.

Substitution of (2) and (3) into (1) gives

$$e_i = x_i - \hat{x}_i. \tag{5}$$

Comparison of (1) and (5) indicates a very important point about DPCM. The quantizing error samples, as defined in (1) are identical to the error samples for the overall system, in the absence of digital transmission errors. When quantizing is relatively fine, the successive quantizing error samples are statistically uncorrelated to a good approximation. Therefore, the signal reconstructed from the error samples has a power spectral density which is flat to a good approximation, as in PCM. According to (5), these same statements also hold for the overall system. There is no contradiction here, because e_i in $\hat{x}_i$ is not simply the response of a linear network to the error e_i in y_i. In fact, the feedback introduces error terms in z_i, and these combine with the quantizing error to produce the total error in $\hat{x}_i$. The flat spectrum does not hold for coarse quantization nor when the probability of overload is high. In these cases, neither the PCM nor the DPCM error spectrum would be flat, in general, nor would the two spectra be the same. The spectrum of error which results is discussed in detail later, and is determined by a computer simulation in Section VI.

In order to determine properties of the quantizing error, e_i, it is necessary to determine properties of the quantizer input, z_i. Substitution of (4) and (5) into (2) yield an equation for z_i

$$z_i = x_i - \sum_{j=1}^{\infty} h_j x_{i-j} + \sum_{j=1}^{\infty} h_j e_{i-j} . \tag{6}$$

It is obvious that even if the last term in (6) were neglected, the statistical properties of z_i, and in particular the probability density, depend on joint statistics of the input and past samples of the input. In the case of voice transmission, there exists empirical data on the probability density[12] and spectrum[13,14] of speech signals, but a good model for even the joint statistics of a pair of samples is not known to the author.

At this point, let us discuss the properties of e_i which it is desired to find. The spectral properties of e_i are already known, as mentioned earlier, provided relatively fine quantizing with low overload probability nolds. The probability density of e_i is not considered important, since there is no evidence to indicate a strong dependence of subjective quality on this property. But the most often needed property is the variance of e_i. A well-known[15,16] expression for the error variance of an L step quantizer is in terms of the probabilities of the various quantizer steps, p_{zj}, and the step sizes, Δ_j.

$$E\{e_i^2\} \mid_{\text{DPCM}} = \sum_{j=1}^{L} p_{zj} \frac{\Delta_j^2}{12} . \tag{7}$$

To emphasize the dependence of the step probabilities on the input variable statistics, we use the input variable as a subscript in addition to the step index. This expression is approximate since among other things overload is neglected, but it is most accurate for fine quantizing and low probability of overload. We are interested in comparing this with the quantizing noise in a PCM system with input x_i and step sizes Δ_j'. With the same type of notation the expression is

$$E\{e_i^2\} \mid_{\text{PCM}} = \sum_{j=1}^{L} p_{xj} \frac{\Delta_j'^2}{12}. \tag{8}$$

Although the ratio of the two quantities given by (7) and (8) is a complicated function of the probability densities of z and x, and also of the choice of step sizes, a rough understanding of what determines this ratio can be found in simpler terms. Suppose the step sizes Δ'_j are chosen to have a fixed ratio with the step sizes Δ_j ; that ratio being the same as the ratio of the rms values of the two inputs. Then, to the extent that the probabilities p_{xj} and p_{zj} are the same, the variances of the errors will be in the same ratio as the variances of x and z. The probabilities in question will be the same if the probability densities of the normalized variables $x/\sqrt{\overline{x^2}}$ and $z/\sqrt{\overline{z^2}}$ are the same.

Whereas an analytic expression for the probability density of z cannot be derived without a model for the joint statistics of x, empirical evidence will be given later to show a strong similarity between the probability density of normalized speech, and that for normalized z in one important case. It is hence natural to use as a figure of merit the ratio of $\overline{x^2}$ to $\overline{z^2}$, which we shall refer to as SNR IMPROVEMENT.

$$\text{SNR IMPROVEMENT} \equiv \frac{E\{x_i^2\}}{E\{z_i^2\}}. \tag{9}$$

We now return to (6). Under the assumption of vanishingly small statistical correlation among the error samples, and between the error and the input signals, the variance of z_i may be written

$$E\{z_i^2\} = E\left\{\left(x_i - \sum_{j=1}^{\infty} h_j x_{i-j}\right)^2\right\} + E\{e_i^2\} \sum_{j=1}^{\infty} h_j^2. \tag{10}$$

It may be noted that the last term in (10) becomes a negligible fraction of the total for high enough signal-to-noise ratios. We note also that the figure of merit depends on the ability to predict x_i with a linear sum of past samples. In fact, we can optimize the figure of merit by choosing the h_j to be the optimum linear prediction coefficients in the sense of minimum mean square error (see Papoulis, Ref. 17). On the other hand, it

should be noted that the optimum coefficients provide a best match to a particular set of signal properties. But speech signal statistics are not constant from speaker to speaker, nor even for one speaker. Therefore, it is best to investigate as well some suboptimal systems with parameters not dependent on signal properties. We also note in passing, that adaptively controlled prediction coefficients might provide an even better solution to the problem. We do not treat the adaptive case in this paper.

IV. SIMPLE, NONOPTIMAL, DPCM

Historically, most of the investigations of predictive feedback systems have not included general feedback networks. One of the most common systems has an integrator or accumulator in the feedback path. That is,

$$h_1 = 1$$
$$h_j = 0 \qquad j \neq 1. \tag{11}$$

It is easy to show that in this case,

$$f_i = \sum_{j=1}^{\infty} y_{i-j} \tag{12}$$

and

$$z_i = x_i - x_{i-1} + e_{i-1}\,. \tag{13}$$

Then, by (10),

$$E\{z_i^2\} = E\{x_i^2\}\,[2(1 - \rho_1)] + E\{e_{i-1}^2\} \tag{14}$$

where

$$\rho_1 = \frac{E\{x_i x_{i-1}\}}{E\{_i^2\}}\,.$$

Neglecting the last term in (14), the figure of merit becomes

$$\text{SNR IMPROVEMENT} \cong \frac{1}{2(1 - \rho_1)}\,. \tag{15}$$

Hence, the figure of merit is greater than unity whenever the normalized adjacent sample correlation of the input signal exceeds 0.5. This result is identical to results obtained by Oliver,[8] Nitadori,[7] and O'Neal,[6] although derived under different assumptions. Empirical work to demonstrate the validity of (15) will be shown in a later section.

This scheme has the advantage that there are no parameters dependent on signal statistics. On the other hand, performance does depend on

signal statistics. If ρ_1 drops below 0.5, the performance is actually worse than PCM. Another disadvantage to this system is that digital channel errors introduce a permanent change in the dc level of $\hat{x}_i'$. In fact, the dc level of $\hat{x}_i'$ will, in the presence of random channel errors, execute an unrestricted random walk until the output saturates. Further discussion of this problem will be found in the next section. In spite of these difficulties, this is the scheme most widely investigated in the literature. In fact, the computer simulation to be reported later will use this system.

V. OPTIMUM LINEAR FEEDBACK NETWORK

As was mentioned previously, the linear feedback coefficients h_j may be optimized in order to minimize the variance of z_i, thus maximizing the figure of merit given by (9). Note that our assumptions have been such as to eliminate the effect of the quantizer nonlinearity from the expressions, and that the solutions given here are optimum only for the cases where our assumptions hold. The problem is somewhat simplified by assuming that the sums in (10) terminate at $j = N$. Since the mutual information between samples usually becomes zero when the samples are remote from each other, the coefficients h_j will approach zero for large j. Therefore, the truncation at $j = N$ does not limit the applicability of the result in cases of interest. Differentiation of the right side of (10) with respect to the variables h_j, and setting the resulting expressions equal to zero gives the following set of linear algebraic equations.

$$\begin{aligned}
\rho_1 &= \left(1 + \frac{1}{K}\right)h_1 + h_2\rho_1 + h_3\rho_2 + \cdots + h_N\rho_{N-1} \\
\rho_2 &= h_1\rho_1 + \left(1 + \frac{1}{K}\right)h_2 + h_3\rho_1 + \cdots + h_N\rho_{N-2} \\
&\vdots \\
\rho_N &= h_1\rho_{N-1} + h_2\rho_{N-2} + h_3\rho_{N-3} + \cdots + \left(1 + \frac{1}{K}\right)h_N,
\end{aligned} \tag{16}$$

where

$$K = E\{x_i^2\}/E\{e_i^2\}$$

is the signal-to-noise ratio, assumed constant, and $\rho_j = E\{x_i x_{i-j}\}/E\{x_i^2\}$. With the exception of the coefficients on the principal diagonal, this is identical to the equations given by Papoulis[17] for determining the optimum linear prediction coefficients. By dividing each equation through by the coefficient on the principal diagonal, the equations are again nor-

malized, with new correlation coefficients. The problem is of course the classical Wiener-Kolmogorov prediction problem in discrete form.

For small N, the algebraic expressions for the solutions are easily obtained. For larger N, a computer solution is more suitable. These solutions may then be put back into the original expression for the error. It is easy to show[17] that the minimum variance of z_i may always be written in the form:

$$E\{z_i^2\}\,|_{\min} = E\{x_i^2\}\left[1 - \sum_{j=1}^{N} h_j\left(\rho_j \Big/ \left(1 + \frac{1}{K}\right)\right)\right]. \quad (17)$$

Hence, the optimum figure of merit becomes

$$\text{SNR IMPROVEMENT}\Big|_{\text{opt.}} = \frac{1}{1 - \sum_{j=1}^{N} h_j\left(\rho_j \Big/ \left(1 + \frac{1}{K}\right)\right)}. \quad (18)$$

It is clear that with fixed sampling rate the minimum error will either remain constant or be monotonically reduced for progressively larger N. That is, each sample further in the past can only add information on which to base a prediction. However, since speech is not perfectly predictable from past samples, it is to be expected that the minimum variance will approach a finite, nonzero, limit as N becomes large. A numerical example showing this relationship, with data from actual speech signals, is given later.

The stability of the closed loop which is present in the DPCM transmission terminal has not been studied in detail. However, the following reasoning clarifies the issue to some extent. Suppose the quantizer granularity is neglected, *i.e.*, assume $y_i \equiv z_i$. Then, using (2), (3), and (4), it is easy to show that

$$x_i \equiv \hat{x}_i$$

$$y_i = z_i = x_i - \sum_{j=1}^{\infty} h_j x_{i-j}.$$

Then the system is stable under a simple criterion requiring the sum of the magnitudes of the h_j's to be finite, and possibly under some weaker criteria. This simple case depends on a *precisely* unity gain amplifier in place of the quantizer. If the gain should remain linear but drift from unity, there exists a forward path from f_i to $\hat{x}_i$. If the gain of this path is sufficient, instability could result. If the granular characteristic of the quantizer is considered, it can be seen that oscillations are possible, in the manner of the bang-bang servo. Some of these cases are investigated

in Section VII. In the cases studied there, the oscillation is bounded, and this means stability in an operational sense.

Because the $N = 1$ case results in considerable simplification, it will be examined a little further. For this case,

$$h_1 = \rho_1 \Big/ \left(1 + \frac{1}{K}\right) \tag{19}$$

and the figure of merit becomes

$$\text{SNR IMPROVEMENT}\Big|_{N=1} = \frac{1}{1 - \left[\rho_1 \Big/ \left(1 + \frac{1}{K}\right)\right]^2}. \tag{20}$$

For large K this is approximately

$$\text{SNR IMPROVEMENT}\Big|_{N=1} \cong \frac{1}{1 - \rho_1^2}. \tag{21}$$

Comparison of (21) and (15) shows a slight advantage for the optimum case over the nonoptimum case, both using one past sample for prediction. Note also that the optimum case always holds an advantage over PCM whereas the nonoptimum case holds an advantage only when $\rho_1 > 0.5$. However, the optimum case requires a parameter adjusted to the assumed signal statistic ρ_1, whereas the nonoptimum case has no such parameters.

It should be noted that when $\rho_1 > 0$, the prediction network for the optimum case with $N = 1$ is merely an attenuation with delay. The overall response of the network from y_i to f_i is that of a "leaky" integrator with delay. This system is one that has been proposed as a means of reducing the problem created by digital channel errors. The effects of digital transmission errors decay exponentially. Hence, the system output does not execute an unrestricted random walk.

VI. COMPUTER SIMULATION — AN EXAMPLE

Because an adequate mathematical model for speech is not available, it was necessary to resort to simulation of the system on the computer, using as input digitized sampled speech, recorded on computer tape. The tape was kindly provided by J. F. Kaiser of Bell Telephone Laboratories. Measuring devices for probability density, variance, and autocorrelation were also simulated on the computer. Knowledge of the autocorrelation function alone is sufficient for evaluation of the figure of merit in each case. But statistics of the derived random variable z are

useful in checking the assumptions leading to our results. Autocorrelation of the quantizing error is also useful in checking our earlier assertions concerning error spectrum.

First, consider the properties of the signal contained on the input tape. The original digitization for computer purposes is linear quantizing with 11 digits. The quantizing error thereby incurred is ignored in further work because quantization introduced in the simulations is much more coarse than this. Table I(a) indicates the normalized autocorrelation of the samples. Note that sampling is at the rate of 9.6 kHz. The spectrum, constructed from this data using a hanning window,[18] is shown in Fig. 3(a) on linear coordinates.

Shown for comparison purposes in Fig. 3(b) is a spectrum constructed as follows. Speech spectra from Dunn and White[14] for men and women are averaged, and the sum multiplied by the attenuation characteristic of a local loop with a 500 telephone set.[19] Although the spectra in Figs. 3(a) and (b) are not precisely the same, the general characteristics are sufficiently close to ascertain the worth of the particular sample. It should be noted that the sample represents only 5 seconds of speech in real time, and cannot be expected to provide a representative average statistic with high precision. However, there is close enough agreement with published statistics to make our point.

Fig. 4 shows the probability density function of the speech samples normalized relative to their RMS value. This was drawn using data obtained by computer processing the input tape. The curve is shown only for positive values of the samples and actually represents an average of both halves of the data. Shown for purposes of comparison are the Laplacian distribution, often used as a speech model,[16] and the Gamma distribution proposed by Richards.[20] The presence of intersyllable and interword quiet time and the presence of low level unvoiced

TABLE I — AUTOCORRELATION COEFFICIENTS — ρ_N

N	(a) 9.6-kHz Samples	(b) 8.0-kHz Samples
0	1.0000	1.0000
1	0.9035	0.8644
2	0.6683	0.5570
3	0.3875	0.2274
4	0.1311	−0.0297
5	−0.0632	−0.1939
6	−0.1939	−0.2788
7	−0.2695	−0.3030
8	−0.3003	−0.2823
9	−0.2980	−0.2208
10	−0.2659	−0.1330

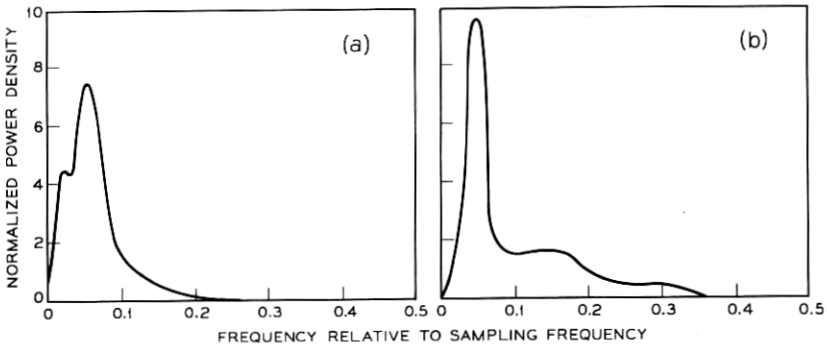

Fig. 3 — Speech spectra; (a) data used in simulation; (b) published data.

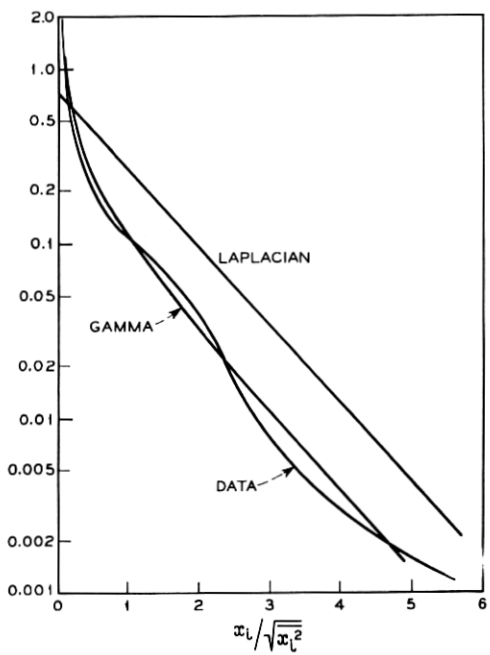

Fig. 4 — Normalized probability density of speech. Symmetrical average of "+" and "−" data.

consonant sounds account for the sharp spike at the origin. The general shape conforms reasonably well to data presented by Davenport.[12]

However, the main reason for using the speech samples is that the second-order statistics are supposed to be representative. It is necessary to take this on faith by extrapolating the favorable comparisons of the first-order statistics and the spectra.

The theory presented in Section IV predicts that the figure of merit for the simple, nonoptimal, DPCM is given by (15). Using ρ_1 from Table I(a), we get

$$10 \log_{10} (\text{SNR IMPROVEMENT}|_{9.6\ \text{kHz}}) = 7.14\ \text{dB}. \tag{22}$$

Under the assumptions developed in Section III, this is the amount by which the overall signal-to-noise ratio will be improved in comparison to PCM, with the same quantizer, comparably loaded.

In Fig. 5(a) are shown curves, determined by the simulation, of signal-to-noise ratio in dB versus input level for PCM and DPCM. The quantizers are 4 digit in each case, and have nonuniform steps conforming to the $\mu = 100$ logarithmic nonlinearity of Smith.[16] In both cases the 0 dB reference input level was determined by trial and error such that the total probability of the two largest quantizer output levels (plus and

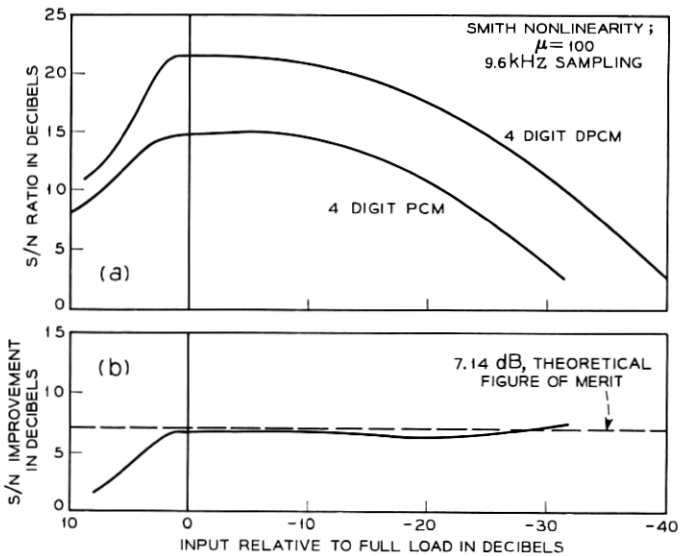

Fig. 5 — (a) Signal-to-noise ratio by simulation. (b) Actual improvement compared to theoretical.

minus) is 0.005. This arbitrary rule was suggested to me by Mr. J. F. Kaiser of Bell Laboratories. The noise includes overload noise, and hence does not conform to the approximate equation (7). Fig. 5(b) shows the difference in dB between the two curves of Fig. 5(a) and also a straight line representing the 7.14 dB difference predicted by the theory. The results indicate that whereas the theory does not predict the exact improvement it does so within about a dB over an input range of about 30 dB. Some improvement in the prediction made by the theory would be expected with quantizers with larger numbers of steps, since the quantizing error term in (14) would then be smaller. This would reduce the discrepancy created by dropping that term in arriving at the figure of merit in (15). But the prediction will never be good in the overload region, because under that condition there is a high probability of large errors, regardless of quantizer step sizes, and the error samples become correlated with each other and with the signal. In addition, some of the discrepancy must be due to the fact that the normalized probability density of z is not the same as that of x. That assumption was made in determining that the figure of merit represented the improvement in signal-to-noise ratio. The assumption that the normalized probability density of z has a particular shape will be poorest under lightly loaded conditions, because the quantizing error becomes a large fraction of the signal z. Even when quantizing error is a negligible component of z, the result depends on the second-order statistics of the input variable x, and these statistics have been guessed at but are not known. The simulation was also carried out using another companding characteristic, with similar results, but those results are not shown here.

Observe in Figs. 6(a), (b), (c), and (d) the normalized probability density of z under various conditions of loading. These curves were obtained from the computer simulation. Note that under progressively lighter loads the probability density changes to bimodal. This can be explained in terms of the oscillation present in DPCM systems under light load when the quantizer is of the midriser type. (See the next section.) The fact that the shape is well maintained in the overload region has not been explained on intuitive or theoretical grounds, but note that in spite of the shape assumption being met, the variances of x and z do not maintain the ratio predicted in (15) because of the error term in z.

Further evidence related to our assumptions is shown in Figs. 7(a), (b), (c), and (d). There the normalized quantizing error power spectral densities for progressively decreased load on the quantizer are shown. In the very lightly loaded case, approximating the idle channel, the oscillation at the half sampling frequency is clearly shown. This was

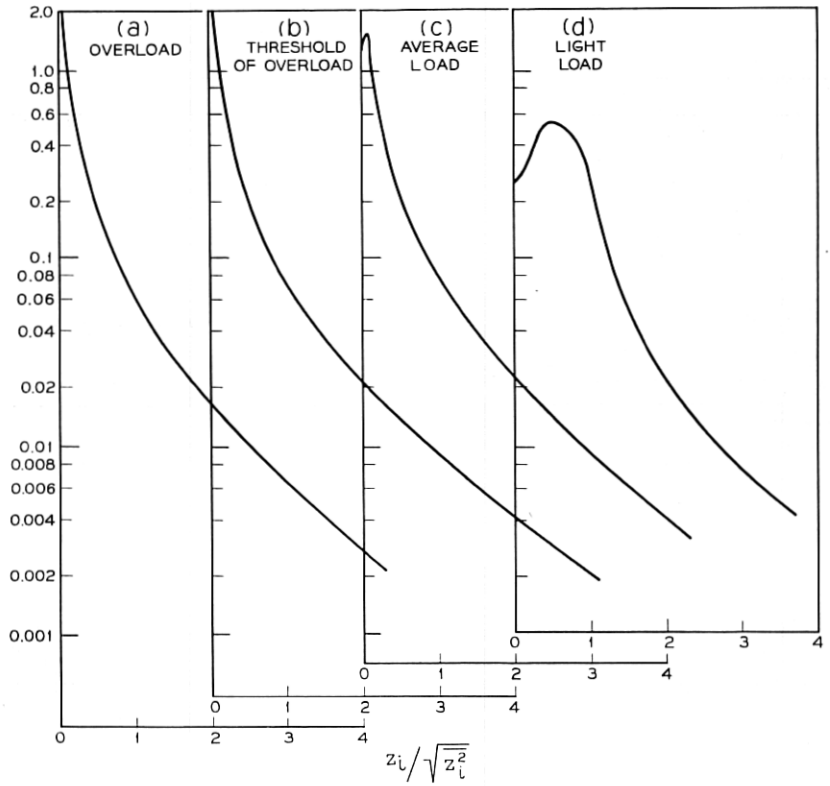

Fig. 6 — Normalized probability densities of z. Symmetrical average of "+" and "—" data; (a) overload; (b) threshold of overload; (c) average load; (d) light load.

mentioned earlier, and is described in full in the next section. The error spectrum remains approximately flat under changing load until the overload phenomena begin. In overload a sharp concentration of energy at low frequencies occurs. No analytical explanation of this has been developed. However, this is not due to a signal correlated component of error because that component was removed computationally in arriving at the spectra shown. It may be due to a statistical dependence more complex than linear correlation, however. The fine structure present on the curves should be ignored, since it is due to the truncation in time of the autocorrelation data used.

Finally, we note one additional point. The simulation was done with a sampling rate of 9.6 kHz. The sampling rate in Bell System voice-fre-

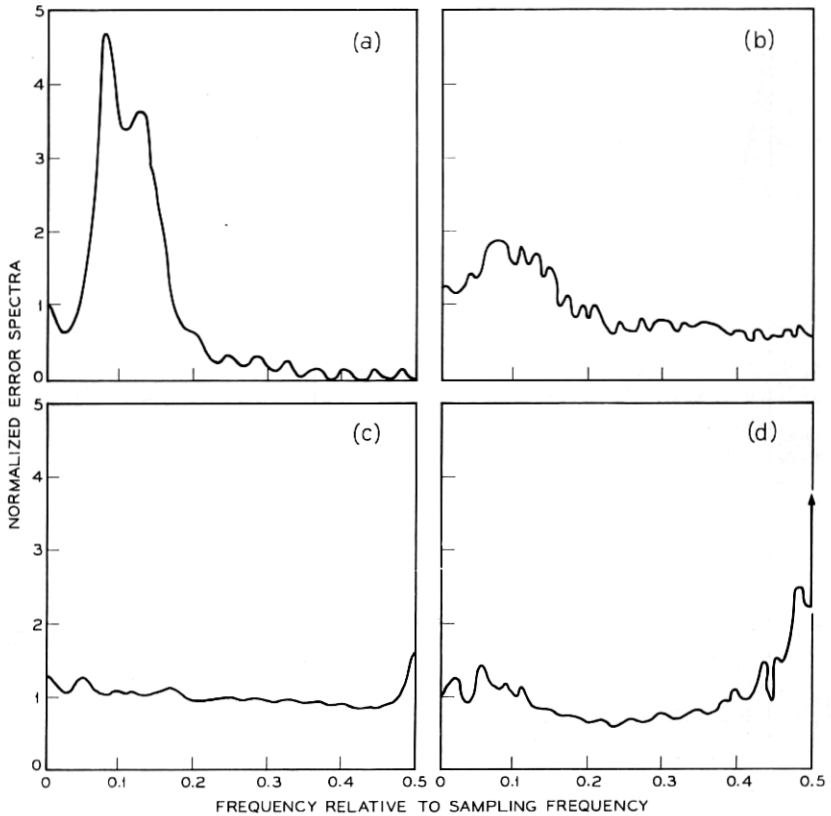

Fig. 7 — Normalized error spectra versus frequency relative to sampling frequency; (a) overload; (b) threshold of overload; (c) average load; (d) light load.

quency PCM equipment, such as the T1 Carrier System, is approximately 8 kHz.[19,21] If the value of ρ_1 for 8-kHz sampling can be determined, a prediction of the advantage of DPCM systems with this sampling rate can be determined. If the autocorrelation function of speech is reconstructed by means of the cardinal series, the value at 0.125 μsec is determined as

$$\rho_1 \mid_{8\text{ kHz}} = 0.8644.$$

All the correlation coefficients determined in this way appear in Table I(b). The figure of merit for nonoptimal DPCM is then

$$10 \log_{10} (\text{SNR IMPROVEMENT} \mid_{8\text{ kHz}}) = 5.7 \text{ dB}. \qquad (23)$$

This means that for the 8-kHz sampling rate, approximately one digit per sample can be saved by using DPCM.

Not considered in the simulations described here are the variations in speech statistics which are known to occur. The usual way of handling the volume variation is to use companding to shape the curve shown in Fig. 5(a). However, the characteristic number ρ_1 will also undoubtedly vary among talkers, perhaps correlated in some way with volume. No statistics on this are known to the author.

With the correlation coefficients given in Table I(b), it is possible to compute the optimum linear feedback coefficients, from (16). Since slightly greater generality is obtained, this will be done for $K \to \infty$. With a signal-to-noise ratio of 30 dB or more, one would expect practically negligible difference. Of great interest is the figure of merit calculated by (18) as a function of N. Fig. 8 shows this relationship. Note that for $N = 1$, the figure of merit is practically 6 dB, just over that attained by the nonoptimal case. For large N the improvement levels off at just over 10 dB, less than 2 digits better than PCM. Better than 9 of the 10 dB are available using only $N = 2$. No simulations have been run for this type of system; hence no check has been made of the assumptions as in the nonoptimal case.

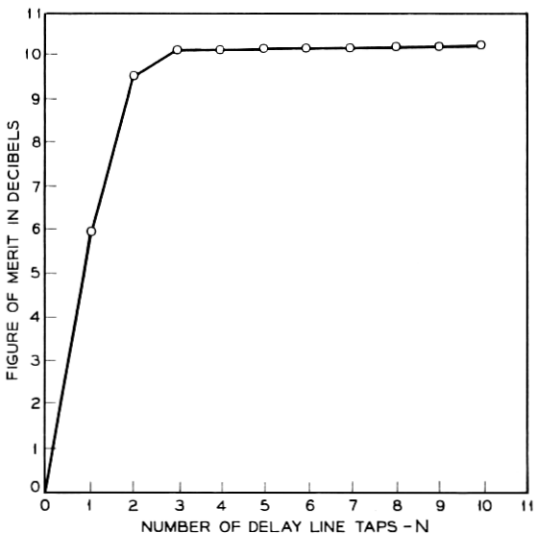

Fig. 8 — Figure of merit vs predictor complexity for 8-kHz sampling.

VII. IDLE CHANNEL PERFORMANCE

Of importance in the design of PCM systems is the so called idle channel performance. Shennum and Gray[22] calculated the output noise of a PCM system with low level thermal noise input, as a function of step size relative to rms noise, and as a function of the bias of the quantizer thresholds nearest the origin. This noise can be larger or smaller than the input thermal noise causing it. A phenomenon with similar causes but different effects occurs in DPCM systems, and these effects are analyzed here.

In this analysis we restrict ourselves to the nonoptimal DPCM described earlier, and represented, for purposes of analysis, in Fig. 9. The Gaussian independent thermal noise samples x_i are the input, and the samples f_i are the output. It is convenient, for analytical purposes, to consider uniform quantizer steps with unit step width and height. It is, of course, common for speech quantizers to be nonuniform, but the steps are generally almost uniform near the origin, the region with which we are concerned. The values computed here for idle channel noise should be compared with those for PCM when the step sizes are the same. Under this condition the systems are approximately equivalent with respect to quantizing noise performance.

The static characteristic of the uniform quantizer we have assumed is shown in Fig. 10, indicating the decision levels b_j, the representation levels a_j, and the biases A and B. The quantizer is assumed to have an infinite number of levels. The biases A and B can represent either drift in the quantizer or a deliberate design, or both. In the overall system, the decoder at the system output, may exhibit a third bias different from A, but that problem is separate from the ones being considered here. There is another factor, similar to these biases, which must be considered; it is the initial condition on the accumulator output at the outset of the idle period. With the exception of the *dc* level created in f_i, and the algebraic sign, the initial condition is identical in its effect to the bias B. Hence, it is no restriction to arbitrarily choose B, if we wish to ignore the *dc*

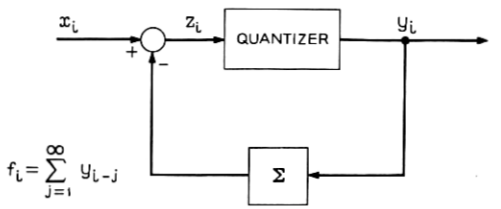

Fig. 9 — Simplified DPCM model.

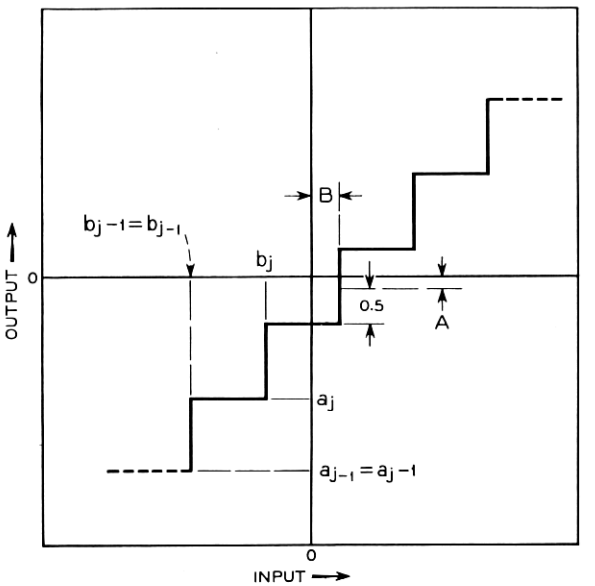

Fig. 10 — Quantizer definitions.

level, as long as the full range of initial conditions is considered. For convenience, choose $A = B$. Then the only two remaining independent parameters are the bias A and the initial value, I.

Under the above set of assumptions we note that the following relations hold.

$$a_j = \frac{b_{j-1} + b_j}{2} \tag{24}$$

$$b_j = a_j + \tfrac{1}{2}; \qquad b_{j-1} = a_j - \tfrac{1}{2} \tag{25}$$

$$y_i = a_j \quad \text{if} \quad a_j - \tfrac{1}{2} \leqq z_i < a_j + \tfrac{1}{2}. \tag{26}$$

For further convenience, let us assume a particular set of a_j's for reference. Let

$$a_j' = j + \tfrac{1}{2} \qquad j = \cdots -2, -1, 0, 1, \cdots. \tag{27}$$

Then

$$a_j = a_j' + A. \tag{28}$$

It is sufficient to study cases covering the range $-\frac{1}{2} \leqq A \leqq \frac{1}{2}$, and in

fact symmetry of the system precludes the necessity of studying half this range.

Of interest is the conditional one step transition probability of the output value f_i. We note this with two subscripts, the second denoting the initial value, the first the subsequent value. The superscript indicates the number of time periods that elapse in the transition.

$$p_{I+a_j,I}^{(1)} = \text{prob}\{a_j - \tfrac{1}{2} \leqq x_i - I < a_j + \tfrac{1}{2}\}. \tag{29}$$

In terms of our previous notations, this may also be written

$$p_{I+A+\frac{1}{2}+j,I}^{(1)} = \text{prob}\{j \leqq x_i - I - A < j + 1\}. \tag{30}$$

Equation (30) is the fundamental equation describing the generation of the samples f_i. However, it still does not give the mean squared value of f_i nor any other statistical characteristics in which we are interested. We may note that the sequence of samples f_i is first order Markov, but this doesn't bring us closer to a solution for our problem.

Let us note that following the initial value I, in one step, there is an enumerable set of possible values of f_i, of the form $\{I + a_j\}$. (Only a small finite subset of this enumerable set have significant probability.) Following each of these possibilities in one more step is another enumerable set of possible outputs; but in general the set of possible outputs is different following each member of the set $\{I + a_j\}$. This makes the state diagram for the sequence of values of f_i extremely complicated in general. It will be necessary to resort to computer simulation to discover what happens in these cases. But under some special assumptions it will be possible to carry the analysis further, and we take those cases first.

Let us first consider the special case, $A = -\frac{1}{2}$, called the midtread case. Then by (28), $a_j = j$, where j is any integer. Under these conditions, the set of possible outputs following I in one step is $\{I + j\}$, including I itself. Following any member of this set in one more step is any member of the *same* set of possible outputs, since the sum of integers is an integer. See Fig. 11 (*a*) for a flow graph to further clarify this case. Only three output states are shown although an enumerable number is required in general. Any member of the output set may be identified by the integer appearing in the expression for that member. The conditional one step transition probability from any member of the set to any other is written

$$p_{jk}^{(1)} = \text{prob}\{j - k - \tfrac{1}{2} \leqq x_i - I - k < j - k + \tfrac{1}{2}\}. \tag{31}$$

Equation (31) denotes a matrix of values of the conditional transition probabilities, which will be called the one step transition matrix, $P^{(1)}$.

$$P^{(1)} = \begin{bmatrix} \vdots & \vdots & \vdots & \\ \cdots\ p_{-1-1} & p_{-10} & p_{-11} & \cdots \\ \cdots\ p_{0-1} & p_{00} & p_{01} & \cdots \\ \cdots\ p_{1-1} & p_{10} & p_{11} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}. \tag{32}$$

In general, $P^{(1)}$ has enumerable dimensionality, but the probabilities become negligible in cases of interest for large enough magnitude of the indices, hence truncation of the range of the indices creates negligible error.

Similarly, let us consider the special case, $A = 0$, called the midriser case. Here, $a_j = j + \frac{1}{2}$. The first set of possible outputs following I is $\{I + j + \frac{1}{2}\}$, a set which does not include I. The next set of possible outputs, following any member of the above set is $\{I + k\}$ where k is any integer. The third set of possible outputs, following any member of $\{I + k\}$ is the same as the previous set $\{I + j + \frac{1}{2}\}$. Hence, there are two subsets of the total output set, and the output alternates between these subsets. Observe that the flowgraph for this case, Fig. 11 (*b*), has arrows only between the two subsets, none within the subsets. (For convenience the number of members of each subset has been truncated at three.) We may solve the problem of indexing the members of these sets by defining the index in the form

$$l = 2(j + \tfrac{1}{2}) = 2j + 1$$

for the first subset, and

$$l = 2k$$

for the second subset. Now with this notation we may write the one step conditional transition probabilities as before.

$$p_{jk}^{(1)} = \text{prob}\left\{\frac{j-k-1}{2} \leqq x_i - I - \frac{k}{2} < \frac{j-k+1}{2}\right\}, \qquad j - k \text{ odd}$$
$$= 0, \qquad j - k \text{ even}. \tag{33}$$

Because of the alternation between members of two sets, about twice as many possible outputs have to be used in this case as in the first case to make the truncation error negligible.

Cases with larger numbers of output subsets may be found under the

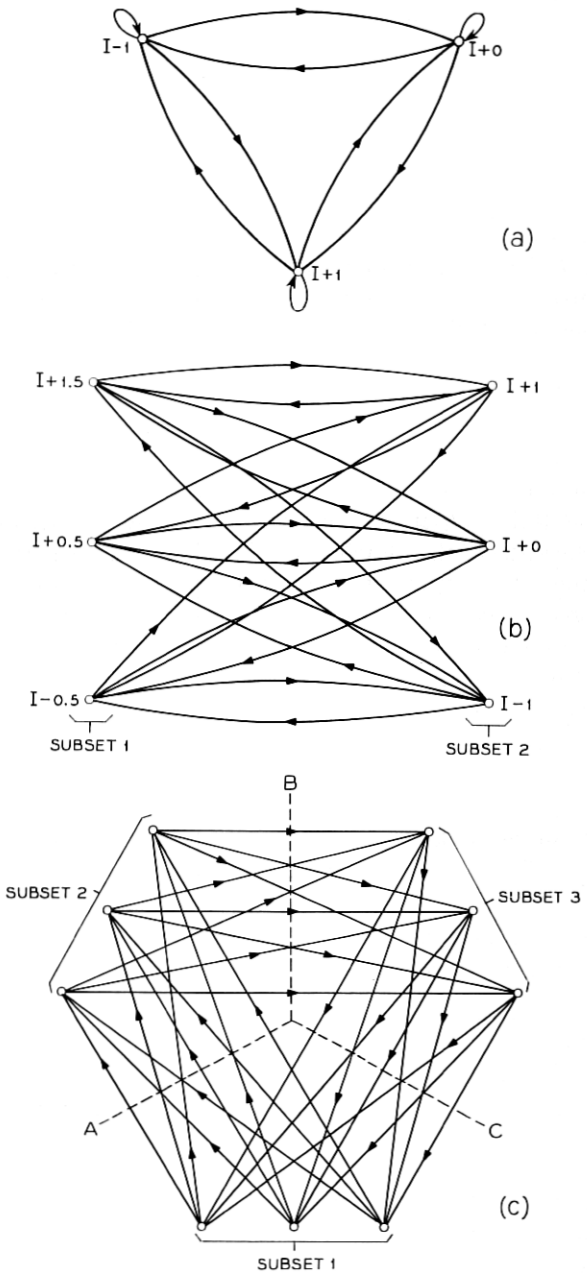

Fig. 11 — (a) Midtread case; $A = -\frac{1}{2}$, $M = 1$. (b) Midriser case; $A = 0$, $M = 2$. (c) $A = -\frac{1}{6}$, $M = 3$.

following conditions. Suppose A is any rational number of the form $\pm m/n$ where $| m/n | \leqq \frac{1}{2}$. Then the number of output subsets is M if M is the smallest integer satisfying

$$M\left(\frac{1}{2} \pm \frac{m}{n}\right) = \text{integer.}$$

The cases for $M = 3, 4, 5$, etc., could be worked out algebraically as we have for $M = 1, 2$, but those cases become unwieldy by the method being used. For clarity, the case $A = -\frac{1}{6}$ $(M = 3)$ is shown in Fig. 11(c), truncated at 3 members of each of the three subsets.

Let us return to the midtread and midriser cases. It is a simple matter to write the conditional probability equations governing the output sequence, and then to arrive at expressions for the output power and autocorrelation function. We write

$$P^{(v)} = P^{(1)} P^{(v-1)}. \tag{34}$$

Equation (34) is a form of the Smoluchowski equation. By iteration, one can easily solve for $P^{(v)}$.

$$P^{(v)} = [P^{(1)}]^v. \tag{35}$$

Let the *a priori* probability of the ith member of the output set be designated C_i, and let the square matrix C be formed with major diagonal elements C_i and other elements zero. The probabilities C_i may be determined from $P^{(1)}$ by simultaneous solution of the following equations:

$$\begin{aligned} C_i &= \sum_{\text{all } k} C_k p_{ik}^{(1)} \qquad \text{for all } i \\ 1 &= \sum_{\text{all } k} C_k . \end{aligned} \tag{36}$$

Equations (36) are an overdetermined system, but in general there is a unique solution, obtainable by not using one of the first group of equations. Now the joint probability of the output at a given step and the output v steps later may be written as the matrix

$$Q^{(v)} = P^{(v)} C \tag{37}$$

and the autocorrelation as a function of v may be written as the quadratic form

$$R(v) = f Q^{(v)} f', \tag{38}$$

where f is a row vector for which the elements are the members of the set of output values, indexed as indicated when developing these sets.

f' is the corresponding column vector. For purposes of defining $R(0)$, $P^{(0)}$ is defined as the unit matrix. All the terms on the right side of (38) may be evaluated by means of the computer, once the transition probability matrix is evaluated. The transition probabilities given by (31) or (33) may be evaluated in terms of the probability function assumed for the samples x_i. It is here that the ratio of variance of x_i to the step size (unity) enters as a parameter. Since the x_i are assumed to be samples of thermal noise, the probability density is assumed Gaussian with zero mean. The most important computed result is $R(0)$, the mean squared output noise relative to the squared step size. The dependence of $R(0)$ on the parameters A, I, and $\overline{x_i^2}$ is shown in Fig. 12. The values shown in Fig. 12 vary between the same maximum and minimum values calculated for PCM by Shennum and Gray.[22] Note that the noise exhibits much more fluctuation as a function of I in the midtread case ($A = -\frac{1}{2}$), although its minimum value is smaller. However, the initial condition I is a parameter which cannot be controlled in the design; hence, midriser

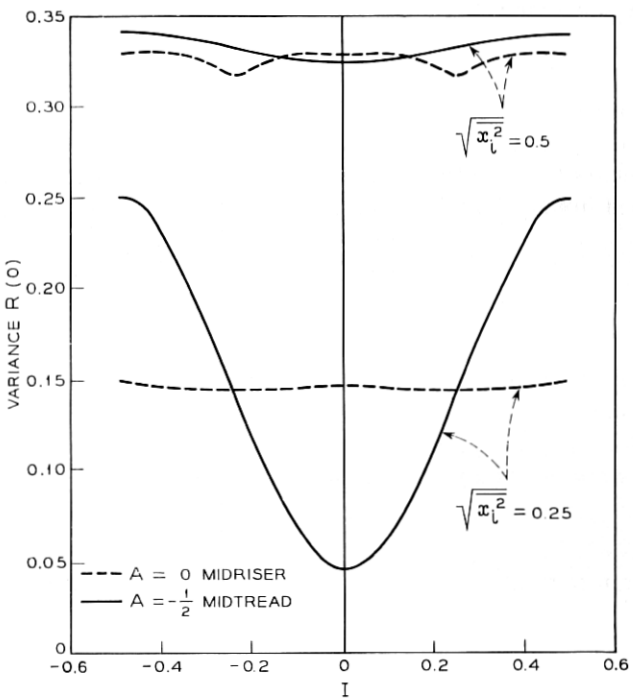

Fig. 12 — Idle channel noise in DPCM system.

operation is a considerably better choice. The other values of $R(v)$ contribute little by way of additional information except to show periodicities and dc values. In the cases studied to this point, these effects were small. Of these two, the midriser case is the only one exhibiting a periodicity. The amplitude of the periodicity depends on the initial value I. The frequency is the half sampling frequency and is easily filtered out in practical systems. This is the same periodicity which showed up in the simulation described in Section VI, and which is indicated in the noise spectrum shown in Fig. 7 (a). Aside from the small dc and periodic components, the output noise samples were uncorrelated.

Other cases, that is those with other values of A, were solved by simulation rather than algebraically. This is because the possible output values become more closely spaced and much higher dimensionality of the matrices is required for accuracy. The simulation is actually quite simple, since all the blocks shown in Fig. 9 are already shown as mathematical operations. The input samples are taken from a so-called Gaussian random number generator program, which produces essentially uncorrelated samples, and having a distribution which is quite accurately Gaussian but which truncates at six times the rms value. The output noise level showed no marked difference from the two cases already presented. Hence, the detailed results will not be shown except for one very interesting case. As indicated earlier, there are particular values of A for which the possible outputs are divided into 3, 4, 5, $\cdots$ etc., subsets. In these cases, the output sequences through these subsets in a definite order, although the particular member is chosen randomly. Clearly there is in general more than one rational fraction A for a given period M, and the pattern followed by the sequence of output subsets may be different for different rational fractions that go with a given value of M. As a consequence, there is a periodicity at $\frac{1}{3}$, $\frac{1}{4}$, $\frac{1}{5}$, $\cdots$ etc., of the sampling frequency respectively. This periodicity falls in the band below the half sampling frequency, and cannot be filtered out as in the midriser case. The sample sequences of the output are not periodic in general because of the randomness of the choice of a particular member of each subset. However, if in each subset there is one highly probable member and others of very small probability, the sample sequences are almost periodic. In the limit of zero input noise, the output sequence is a periodic function. In all cases, the mean and variance are periodic functions of time. As an example, a flowgraph for the case $M = 3$ is shown in Fig. 11(c). At cut A, all the arrows are from subset 1 to subset 2. No other arrows leave subset 1. Similarly, cut B intersects all the arrows from

subset 2 to subset 3, and cut C intersects all those from subset 3 to subset 1.

For the members of a given output subset, *a priori* probabilities may be computed by the methods indicated in (36), and these probabilities can be used to compute the mean and variance for the places in the sequence where that subset applies. The mean and variance for a case where the periodicity is $\frac{1}{20}th$ of the sampling frequency is shown in Fig. 13. In the case of the variance, both the theoretical curve computed from the probabilities and the curve obtained from the simulation are presented. Good agreement is shown.

If the value of A is irrational, the sequence of output subsets does not close on itself as in the cases indicated here. Hence, the period is not an integral multiple of the sampling interval. However, a similar phenomenon takes place with respect to periodicity of the mean and variance of a continuous signal reconstructed from the output sample sequence.

With the simulation method it is possible to study cases with nonuniform quantizing as well as the uniform cases studied to this point. One case with a small amount of nonlinearity was tried but no significantly different phenomena appeared in the results.

It is clear that the model used here for idle channel noise is not ade-

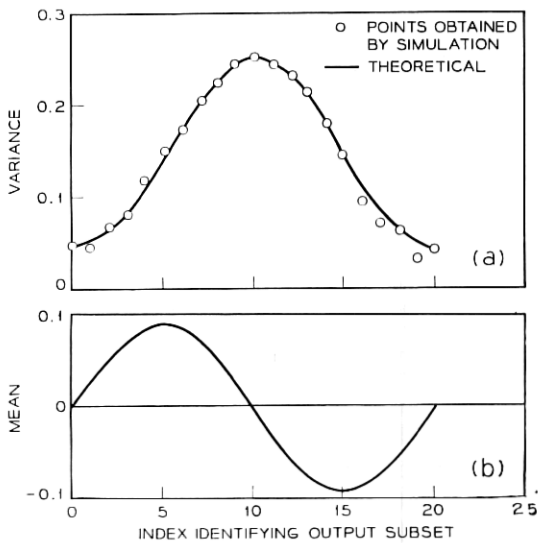

Fig. 13 — Mean and variance of idle channel noise $A = -0.45$; $\sqrt{\overline{x_i^2}} = 0.25$.

quate to describe what goes on in other systems such as DPCM with optimal feedback.

By way of summary, let us note that all evidence presented here points to the midriser case ($A = 0$) as the best one from the system design viewpoint. This is because of the inband periodic component present in the output of all of the other cases except the midtread case. The annoyance created by these inband periodicities will depend on their amplitude and frequency. This could be further studied by the simulation methods, and evaluated with subjective tests. But the midriser design should prove satisfactory, and additional investigation has not been undertaken. The midtread case is the only one showing no periodicity, but it is also the only one showing a high degree of dependence of the output variance on the initial condition I. Since the value of I cannot be controlled by the designer, the midtread case is also unsatisfactory. We note that it is only the quantizer output bias which must be controlled closely in the practical system, since the input bias change is equivalent to a change of initial condition, and we have found no great dependence on this parameter except in the midtread case.

It may be that the midriser design would no longer appear best when one considers crosstalk into an idle channel with a shared coder. Since this has not been investigated, no conclusions are presented on this point.

VIII. PRE-EMPHASIS

Equation (13) shows that in simple DPCM, it is the difference between adjacent input samples which forms the principal component of quantizer input. This leads one to the idea that a similar performance advantage is to be gained by using a pre-emphasis network with PCM. The network should approximate a differentiator. The principal qualitative difference in performance of this system is that an integrator is needed at the output to restore the original signal. This destroys the independence of the error samples and creates a subjective change in the output noise. It is not known whether frequency weighting of the noise will adequately account for the subjective changes. This problem was examined briefly, but is not reported in detail here. The pre-emphasis filter can be optimized, subject to a frequency weighted error criterion. It was found that, using this objective performance measure, an advantage nearly the same as that of simple DPCM can be attained.

Pre-emphasis (and de-emphasis) can also be used with DPCM systems. However, the differential aspect of the system makes use of most of the

advantage to be gained, leaving little additional for the filtering. In other words, the effects are not disjoint.

IX. SUMMARY

The results presented have fallen into approximately three categories. First, an analysis of signal to quantizing noise ratio has been presented, indicating the advantages to be gained by the use of various forms of DPCM, including simple DPCM and optimal DPCM with varying amounts of memory. The analytical results are discussed in the light of results obtained by other authors and the assumptions used. Second, a computer simulation was used to check the assumptions implicit in the present work and that of others. The probability density of the quantizer input and the quantizing error spectrum were studied by the simulation technique. The computer was also used to evaluate the performance to be expected when DPCM is used for speech transmission. It is shown that approximately 6 dB or one digit per sample advantage over PCM is attained by the simplest DPCM system. With optimal linear prediction, 10 dB or less than two digits per sample advantage over PCM is attained. Finally, the performance of the DPCM idle channel is investigated. It is shown that periodicities in the output of the idle channel sometimes are present. Amplitude and frequency depend on the bias of the quantizer output. It is pointed out that the most satisfactory design is the so-called midriser case, where the periodicity is at the half sampling frequency and can be filtered out in a practical system. The idle channel noise of DPCM varies in a different way from that of PCM. The level of the idle channel noise is approximately the same in both PCM and DPCM, when the quantizing noise performance is the same.

X. ACKNOWLEDGMENT

The author wishes to express his appreciation to the many individuals who were of assistance to him in this work. Among these were J. F. Kaiser who provided the computer tape and some programming advice, Miss E. G. Cheatham who helped with the programming, and M. R. Aaron who provided insight and later helped edit the manuscript.

REFERENCES

1. N. V. Phillips, Gloeilampenfabrieken of Holland, French Patent No. 987, 238. Applied for May 23, 1949; Issued August 10, 1951.
2. Cutler, C. C., Differential Quantization of Communication Signals, U. S. Patent No. 2,605,361, July 29, 1952.
3. Schouten, J. F., De Jager, F., and Greefkes, J. A., Delta Modulation, A New

Modulation System for Telecommunication, Phillips Technical Review, March, 1952.
4. Van De Weg, H., Quantizing Noise of a Single Integration Delta Modulation System with an N-digit Code, Phillips Research Reports, *8,* 1953, pp. 367–385.
5. Zetterburg, L., A Comparison Between Delta and Pulse Code Modulation, Ericsson Technics, *2,* No. 1, 1955, pp. 95–154.
6. O'Neal, J. B., Predictive Quantizing Systems (Differential Pulse Code Modulation) for Transmission of Television Signals, B.S.T.J., *45,* May–June, 1966, pp. 689–721.
7. Nitadori, K., Statistical Analysis of Δ-PCM, Electronics and Communications in Japan, *48,* No. 2, February, 1965.
8. Oliver, B. M., Efficient Coding, B.S.T.J., *31,* July, 1952.
9. Kimme, E. G., Methods of Optimal System Design for a PCM Video System Employing Quantization Noise Feedback, Unpublished work.
10. Kimme, E. G. and Kuo, F. F., Synthesis of Optimal Filters for a Feedback Quantization System, IEEE Trans. on C. T., *CT-10,* No. 3, September, 1963, pp. 405–413.
11. Spang, H. A. and Schultheiss, P. M., Reduction of Quantizing Noise by the Use of Feedback, IRE Trans. on Commun. Syst., December, 1962, p. 373.
12. Davenport, W. B., An Experimental Study of Speech-Wave Probability Distributions, J. Acoust. Soc. Amer. *24,* No. 4, July, 1952, pp. 390–399.
13. Purton, R. E., A Survey of Telephone Speech Signal Statistics and Their Significance in the Choice of PCM Companding Law, IEE Paper No. 3773E, January, 1962.
14. Dunn, H. K. and White, S. D., Statistical Measurements on Conversational Speech, J. Acoust. Soc. Amer., *11,* No. 1, January, 1940.
15. Panter, P. F. and Dite, W., Quantization Distortion in Pulse-Count Modulation with Nonuniform Spacing of Levels, Proc. IRE, *39,* January, 1951, pp. 44–48.
16. Smith, B., Instantaneous Companding of Quantized Signals, B.S.T.J., *27,* July, 1948, pp. 446–472.
17. Papoulis, A., *Probability, Random Variables, and Stochastic Processes,* McGraw-Hill Book Company, Inc., 1965, see Chapter 11, p. 385 ff.
18. Blackman, R. B. and Tukey, J. W., *The Measurement of Power Spectra,* Dover Publications, Inc., New York, 1958.
19. Members of Technical Staff of Bell Telephone Laboratories, *Transmission Systems for Communication,* Bell Telephone Laboratories, 1964.
20. Richards, D. L., Statistical Properties of Speech Signals, Proc. IEE, *111,* No. 5, May, 1964.
21. Fultz, K. E. and Pennick, D. B., The T1 Carrier System, B.S.T.J., *44,* September, 1965, p. 1405.
22. Shennum, R. H. and Gray, J. R., Performance Limitations of a Practical PCM Terminal, B.S.T.J., *41,* January, 1962, pp. 143–171.
23. Brainerd, R. C., Subjective Measurement on PCM Noise Feedback Coder for Video Transmission, Unpublished Memorandum.
24. Graham, R. E., Predictive Quantizing of Continuous Sources, Unpublished Memorandum.
25. McDonald, H. S. and Jackson, L. B., A Study of Single Frame Encoding of Pictures by Differential PCM Incorporating a Lossy Integrator, Unpublished Memorandum.
26. Miller, R. L., The Possible Use of Log Differential PCM for Speech Transmission in Unicom, Paper delivered at Globecom Conference in Chicago, 1961.
27. Miller, R. L. and Mounts, F. W., A Comparison of Differential PCM with Normal PCM for Speech Transmission, Unpublished Memorandum.