Abstract—The paper treats linear and nonlinear methods that refine the delta modulation process. These methods improve the quality of the decoded signal at a given transmission rate. Two linear processes are analyzed. The first one matches the source signal to the coder for minimum weighted noise in the decoded signal. The second one influences the signal which is fed to the comparator; it shapes the spectrum of the quantization noise desirable. Two nonlinear methods used in the delta modulation process are discussed. The first achieves a best possible prediction of the future signal from past binary decisions and thus a reduction of the error signal. The methods of the second type make the coder adaptive to variations in the power level of the source signal. This results in a large dynamic range of the coding system. A comparison of the various methods is made and a coding system for speech signals, based on a suitable combination of these methods, is described.

I. INTRODUCTION

DELTA MODULATION is a method for the encoding of analog source signals; it is characterized by the facts that only one binary symbol is produced in each sampling interval and that it can be implemented with simple hardware [1], [2].

In spite of the simplicity of delta-coding equipment, delta modulation, in its early days, did not find widespread application because its dynamic range was too limited. This lack in dynamic range was due to the fixed step sizes used in the delta-modulation process. In the past years many suggestions have been made for a refinement of the simple delta-modulation process in order to achieve an acceptable dynamic range [3]-[4]. They are all based on matching a variable step size to the power level of the source signal. Two basic approaches exist for the control of the step size: in the first one a control signal is derived from the analog signal and transmitted separately to the receiver (forward control) [5]; in the second one, a control signal is derived, in some way, from the digital signal produced by the coder (feedback control). With delta modulation exhibiting adequate dynamic range, it has been applied in many public and military systems [15]-[18].

Compared to pulse-code modulation (PCM), the mathematical description of the delta-modulation coding process is more difficult. In recent years many contributions have been made to the analysis of granular- and slope-overload noise [19]-[31]. Most theoretical investigations have treated the case of single integration, fixed step-size delta modulation.

In this paper three aspects of delta modulation are treated. Assuming stationary signals, it will first be shown how linear processes can be introduced in delta modulation and how they are chosen for optimum coding performance. Two linear processes are considered. A first one, the prefilter, treats the source signal before it is fed to the coder. It will be optimized for a maximum ratio of signal-to-weighted-noise power of the decoded signal. The second linear process acts on the signal fed to the comparator; it shapes the spectrum of the quantization noise suitably.

Secondly, the case of nonlinear delta modulation is investigated; from the analysis based on prediction theory, a nonlinear prediction algorithm is derived. It will be demonstrated that a simplification of the algorithm causes little loss in coding quality. The nonlinear algorithm, also optimized for stationary signals, can be described in form of a state diagram. Finally, assuming nonstationary source signals, adaptive algorithms, superimposed on the algorithms described, will be discussed and their stability analyzed. These adaptive algorithms adjust the step size (or a range of step sizes) to the power level of the signal and thus enhance the dynamic range of the coding system appreciably. The step-size control mechanism will be adapted to the dynamic behavior of speech.

The ratio of signal power to weighted-noise power will be used as an optimality criterion throughout this paper; this criterion is suitable for a mathematical analysis and it describes rather well the subjectively judged quality of a coding system.

II. LINEAR DELTA MODULATION

A. The Concept of Delta Modulation

A basic delta-modulation system operating with fixed step sizes is shown in Fig. 1. It consists of a coder, a digital transmission channel and a decoder. An analog signal having a spectrum $S_{an}$ is transformed by the coder into a sequence of binary symbols. This is achieved by a comparator continuously comparing the analog source signal to a reconstructed signal, and by a sampler periodically sampling the comparator decisions. The binary symbols, in the form of positive and negative impulses, are fed to a linear network $H(\omega)$; it reconstructs a waveform that follows closely the analog source signal.

The most simple network $H(\omega)$ consists of an ideal
integrator (single integration delta modulation). Fig. 2 shows analog and reconstructed waveforms for this case.

With a more general network \( H(\omega) \) operation of the coder and the reconstructed waveforms becomes more complex. In order to make the coding process suitable for analytical treatment, the block diagram of the delta-modulation system is reconstructed according to Fig. 3.

The reconfiguration is based on dividing the network \( H(\omega) \) into two sections, i.e., an ideal integrator and a network \( G(\omega) \). The network \( G(\omega) \) is moved, within the feedback loop, to the position in front of the comparator. Thus, it will influence the comparator decisions and therefore it is called decision filter. If the spectrum of the source is shaped by a filter \( G'(\omega) \), the two configurations of Fig. 3 perform identically. For the analysis, a general filter \( F'(\omega) \) (prefilter), matching the source spectrum to the coder, will be assumed.

The decoder integrates the received sequence of binary symbols. A postfilter and an output-filter follow the integrator. The postfilter \( F'^{-1}(\omega) \) compensates for the response of the prefilter while the postfilter \( A(\omega) \) characterizes the response of a transducer which makes the information visible or audible, assuming picture or speech inputs to the encoder, respectively. A noise-weighting filter characterizes the subjective weighting of coding noise.

In the coding process two types of quantization errors are generated. These are granular noise and slope-overload distortion (Fig. 2). The two types of noise are not equally noticeable to the human ear. Slope overload [32], being correlated to the input signal, is less disturbing than granular noise at equal noise power levels [33].

### B. The Choice of the Prefilter for a Maximum Ratio of Signal Power to Weighted-Noise Power of the Decoded Signal

The prefilter that will lead to a maximum ratio of signal power to weighted-noise power of the decoded signal will be determined under the following assumptions.

**Assumption 1:** The source is Gaussian with spectrum \( S_{XX}(\omega) \), bandlimited to a frequency \( W \).

**Assumption 2:** The sampling rate of the coder \( f_s \) far exceeds the effective bandwidth \( f_e \) of the source signal (\( f_s > 64 \cdot f_e \)).

**Assumption 3:** A slope loading not exceeding \( 4\sigma \) is assumed; this means that the maximum slope that the coder is capable of reconstructing is equal to (or larger than) four times the rms value of slopes of the signal fed to the coder. This loading leads to a probability of slope overload of less than \( 4 \cdot 10^{-5} \). Thus we consider an optimization of the granular portion of the total coding noise.

As derived in the Appendix, these assumptions ensure that the quantization noise is not dependent on the spectrum of the signal fed to the coder. This statement is true for a large class of signal spectra.

As shown in Fig. 4, the coding system can be replaced by a slope limiting device and a noise source with spectrum \( S_{xx}(\omega) \). The noise source represents the granular quantization noise generated in the coding process. We assume that the output filter \( A(\omega) \) and the noise-weighting filter \( D(\omega) \) are given. Since the postfilter \( F'^{-1}(\omega) \) compensates the response of the prefilter \( F(\omega) \), the signal power at the input of the noise-weighting filter is constant, i.e., it does not depend on the choice of the prefilter. Therefore we shall search for the prefilter that minimizes the weighted-noise power under the condition of fixed slope loading (see Fig. 4):

\[
f_s = \left\{ \frac{\int_{-\infty}^{\infty} p S_{xx}(f) df}{\int_{-\infty}^{\infty} S_{xx}(f) df} \right\}^{1/2}
\]
Fig. 4. Equivalent circuit for the determination of the prefilter.

\[ \int_0^\infty S_{nn}(f) \cdot |F(f)|^2 \cdot |A(f)|^2 \cdot |D(f)|^2 \, df \rightarrow \text{minimum}. \]

The condition of fixed slope loading can be written

\[ 4\pi^2 \cdot \left[ \int_0^W S_{xx}(f) \cdot |F(f)|^2 \cdot \left[ f_0^2 + f^2 \right] \, df \right] / Q(|F|[2], f) = \text{constant}. \]

For practical reasons a leaky integrator was assumed, since constant slope loading with a perfect integrator would allow low-frequency components of arbitrary amplitude. \( f_0 \) is the corner frequency of the leaky integrator.

The solution to this problem can be found by making use of variational calculus [34]. The key equation in this calculus is the Euler differential equation; it can be applied to find the function \( F(f) \) that minimizes the weighted-noise power. The equation states (with the derivative of the function to be determined not appearing in our case)

\[ -S_{nn}(f) \cdot |A(f)|^2 \cdot |D(f)|^2 \cdot |F(f)|^4 \]
\[ + \lambda \cdot S_{xx} \cdot \left[ f_0^2 + f^2 \right] = 0. \]

Thus

\[ -S_{nn}(f) \cdot |A(f)|^2 \cdot |D(f)|^2 \cdot |F(f)|^4 \]
\[ + \lambda \cdot S_{xx} \cdot \left[ f_0^2 + f^2 \right] = 0. \]

This leads to

\[ |F(f)|^2 = |A(f)|^2 \cdot |D(f)| \cdot \left[ \frac{1}{\lambda S_{xx}(f) \cdot \left[ f_0^2 + f^2 \right]} \right]^{1/2}. \]

Inserting \( |F(f)|^2 \) into the condition for constant slope loading yields

\[ 4\pi^2 \cdot \left[ \int_0^W S_{xx}(f) \cdot |A(f)|^2 \cdot |D(f)| \cdot \left[ f_0^2 + f^2 \right] \right] \cdot \left[ \frac{1}{\lambda S_{xx}(f) \cdot \left[ f_0^2 + f^2 \right]} \right]^{1/2} \]
\[ \cdot \left[ \frac{1}{\lambda S_{xx}(f) \cdot \left[ f_0^2 + f^2 \right]} \right]^{1/2} \, df = \text{constant}. \]

Resolving for \( \lambda \) and inserting \( \lambda \) into the equation describing \( |F(f)|^2 \) leads to

\[ |F(f)|^2 = \text{const.} \cdot \left[ \int_0^W S_{xx}(f) \cdot S_{nn}(f) \cdot |A(f)|^2 \cdot |D(f)|^2 \cdot \left[ f_0^2 + f^2 \right] \right]^{1/2} \cdot \left( \frac{S_{xx}(f) \cdot |A(f)|^2 \cdot |D(f)|^2}{S_{xx}(f) \cdot \left[ f_0^2 + f^2 \right]} \right)^{1/2}. \]

Thus

\[ |F(f)|^2 \sim \left( \frac{1}{S_{xx}(f) \cdot \left[ f_0^2 + f^2 \right]} \right)^{1/2}. \]

The response of the prefilter leading to minimum weighted noise in the decoded signal has been found. This filter depends mainly on the source spectrum, the noise spectrum and on output and weighting filters. For the case of white noise (single integration delta modulation) and flat output and weighting filters, \( |F(f)|^2 \) becomes

\[ |F(f)|^2 \sim \left( \frac{1}{S_{xx}(f) \cdot \left[ f_0^2 + f^2 \right]} \right)^{1/2}. \]

It is interesting to compare this result to the response of a prefilter that maximizes the entropy of the decoded signal (derived in [31]). In that case

\[ |F(f)|^2 \sim \left( \frac{1}{S_{xx}(f) \cdot \left[ f_0^2 + f^2 \right]} \right). \]

This second (entropy-maximizing) prefilter matches the source spectrum to the overload characteristic of the coder, whereas the first prefilter (maximizing weighted

\[ ^\dagger \text{The overload characteristic is defined as the maximum sine-wave amplitude that does not cause slope overload, as a function of sine-wave frequency.} \]
signal-to-noise (S/N) ratio) performs only a partial matching. Furthermore, the noise spectrum as well as the output and weighting filter influence only the prefilter that minimizes the weighted noise.

The ratios of signal power to weighted-noise power (SNR) can be determined for the two cases in a straightforward manner. The results are

\[ \text{SNR}_N = \frac{1}{4\pi^2} \text{slope loading}^2 \left[ \int_0^W S_{xx}(f) |A(f)|^2 \, df \right] \sqrt{\text{SNR}} \]

and

\[ \text{SNR}_E = \frac{1}{4\pi^2 \cdot W} \text{slope loading}^2 \left[ \int_0^W S_{xx}(f) |A(f)|^2 \, df \right] \sqrt{\text{SNR}} \]

with

\[ \text{SNR}_N \quad \text{ratio of signal power to weighted-noise power, prefilter optimized for minimum weighted noise.} \]
\[ \text{SNR}_E \quad \text{ratio of signal power to weighted-noise power, prefilter optimized for maximum entropy of decoded signal.} \]

The two formulae are rather similar. They become identical if the expressions under the integrals in the denominators are independent of frequency. In that case the entropy-maximizing prefilter also maximizes the ratio of signal to weighted-noise power. The formulae furthermore show that a suitable shaping of the noise spectrum \( S_{xx}(f) \) might increase the SNR. It will be shown in the following section how this spectrum can be suitably shaped by means of a decision filter.

C. The Noise Spectrum and Choice of the Decision Filter

Linear filtering of the error signal fed to the comparator, i.e., the introduction of a decision filter, influences the behavior of the coder in several ways. We shall see that, for a slope loading not exceeding 4\( \pi \), an increased ratio of signal to weighted noise can be achieved. However, the recovery from slope overload may become longer and in some cases undesirable idling patterns are generated by the coder. The spectrum of the quantization noise can be shaped by the decision filter for minimum weighted-noise power. This is achieved at the expense of increased noise power outside the range of interest. The out-band noise however can be eliminated by a filter following the decoder.

In contrast to related studies on delta modulation [30], we consider here the case of signal reconstruction with fixed step sizes but error-dependent comparator decisions.

In the case of coding without the decision filter, the comparator makes decisions based on the instantaneous differences between the input signal and the reconstructed signal, as shown in Fig. 5,

\[ b_n = \text{sgn} (y_n - r_{n-1}) = \text{sgn} (v_n) \]

with

\[ b_n \quad \text{binary decision at time } nT, \text{ i.e., } \pm 1 \]
\[ \text{sgn} (x) \quad \text{sign function } = x / |x| \]

The comparator always decides such that the smaller of the two possible error values is chosen. The error \( e_n \) is defined

\[ e_n = y_n - r_n. \]

In the presence of a decision filter, the comparator makes a binary decision \( b_n^* \) as follows:

\[ b_n^* = \text{sgn} [y_n + g_0 e_n + g_1 e_{n-1} + g_2 e_{n-2} + \ldots ]; \]

and now also the past error samples influence the comparator decision. The values \( g_0, g_1, g_2, \ldots \) represent the sampled impulse response of the decision filter of time 0, T, 2T, 3T, \ldots. We shall assume \( g_0 = 1 \). If this impulse response is properly chosen, the comparator will make binary decisions such that the contributions to the spectrally weighted-noise power are minimal. In order to simplify the analysis, we assume that the sampling rate is much higher than the Nyquist rate. As shown in the Appendix the error signal does not contain component that are correlated to the source signal, if

\[ f_s > 42fw. \]

For this range of sampling frequencies, the total error signal power appears as noise to a human observer. The symbols are

\[ f_s \quad \text{sampling rate} \]
\[ f_w \quad \text{effective bandwidth of source signal.} \]

The weighted-noise power \( \eta_\theta \) is determined as follow (for \( f_s > 42fw \)):

\[ \eta_\theta = 2T \int_0^W S_{ss^*}(f) |N(f)|^2 \, df, \]

with

\[ W \quad \text{bandwidth of source signal} \]
\[ T \quad \text{sampling interval} \]
\[ S_{ss^*} \quad \text{spectrum of the error samples} \]
\[ N(f) \quad \text{frequency response of postfilter, output filter and noise-weighting filter combined.} \]

The spectrum of the error samples \( S_{ss^*}(f) \) can be describe using the autocovariance coefficients
Fig. 5. Coding process and error samples under the influence of the decision filter.

$$S_{e*}(f) = \frac{1}{T} \sum_{-\infty}^{\infty} R_{ee}(\mu T) \delta(t - \mu T) \exp(-j\omega t) \, dt$$

$$= \sum_{\mu} R_{ee}(\mu T) \exp(-j\omega \mu T),$$

with

$$R_{ee}(\mu T) = E\{e_n \cdot e_{n+\mu}\}.$$ 

The symbol $E$ stands for the expected value of the product $\{e_n \cdot e_{n+\mu}\}$. Since the autocovariance coefficients are symmetrical,

$$R_{ee}(\mu T) = R_{ee}(-\mu T),$$

we obtain

$$S_{e*}(f) = R_{ee}(0) + 2 \sum_{\mu=1}^{\infty} R_{ee}(\mu T) \cos(\omega \mu T).$$

With these relations the weighted-noise power can be determined:

$$\eta_0 = 2T \int_0^W R_{ee}(0) + 2 \sum_{\mu=1}^{\infty} R_{ee}(\mu T) \cos(\omega \mu T) \, |N(f)|^2 \, df$$

$$= 2T \left[ \int_0^W |N(f)|^2 \, df \right] R_{ee}(0)$$

$$+ 4 \cdot T \sum_{\mu=1}^{\infty} \left[ \int_0^W \cos(2\pi \mu T) \, |N(f)|^2 \, df \right] R_{ee}(\mu T).$$

The expression for the weighted-noise power is of the form

$$\eta_0 = K \cdot \sum_{j=0}^{\infty} a_j R_{ee}(jT) + \sum_{j=0}^{\infty} a_j R_{ee}(jT).$$

with

$$K = 2T \int_0^W |N(f)|^2 \, df.$$ 

The coefficients $a_i, a_2, a_3, \ldots$ are given by

$$a_j = 2 \left\{ \int_0^W \cos(2\pi j \cdot f \cdot T) \, |N(f)|^2 \, df \right\} \int_0^W |N(f)|^2 \, df$$

$$j = 1, 2, \ldots$$

with

$$a_0 = 1.$$

The expression for the weighted-noise power will lead to the unit impulse response of the decision filter, assuming negligible probability of slope overload. In the derivation it will be assumed that the decision filter always causes two consecutive binary decisions to deviate from the decisions taken without the filter (Case a in Fig. 5), and that events like Case b in Fig. 5 occur with negligible probability. This last assumption allows an analytical treatment of the problem; it can be verified by computer simulation once the decision filter is found.

The contribution to the weighted-noise power at time $nT$ can be derived from (6):

$$\Delta \eta_0 = K\{e_n \cdot e_n + a_0 e_n \cdot e_{n-1} + a_2 e_n \cdot e_{n-2} + \cdots\}.$$ 

If a pair of consecutive comparator decisions deviates from that taken without decision filter, a single error sample changes as follows (Fig. 5):

$$e_n^* = e_n \pm 2 \cdot d.$$ 

The symbols are:

- $e_n^*$: error sample, coding process with decision filter
- $d$: step size of delta coder.

The new error samples $e_n^*$ (coding process with decision filter) are composed of the original error samples $e_n$ and sporadic impulses $z_i$ of size $+2d$ or $-2d$ (Fig. 5):

$$e_n^* = e_n \pm 2 \cdot d.$$ 

The autocovariance coefficients of the new sequence of error samples can now be written

$$R_{ee*}(jT) = E\{e_n \cdot e_n + a_0 e_n \cdot e_{n-1} + a_2 e_n \cdot e_{n-2} + \cdots\}.$$ 

The weighted-noise power now becomes

$$\eta_0^* = K\left[ \sum_{j=0}^{\infty} a_j R_{ee}(jT) + \sum_{j=0}^{\infty} a_j R_{ee}(jT) \right] + \sum_{j=0}^{\infty} a_j R_{ee}(jT).$$

At time $nT$ the comparator should take a decision that leads to the smallest possible contribution to the weighted-noise power. With
binary decision at time \( nT \), coding process without
decision filter

**\( b_n^* \)** binary decision at time \( nT \), coding process with
decision filter

the two possible contributions can be determined.

If

\[
b_n = b_n^*
\]

then

\[
z(nT) = 0
\]

and the contribution to the weighted-noise power \( \Delta \eta \) becomes

\[
\Delta \eta(z_n = 0) = \epsilon_n [a_0 \epsilon_n + a_1 (\epsilon_{n-1} + z_{n-1})
+ a_2 (\epsilon_{n-2} + z_{n-2}) + \cdots].
\]

On the other hand, if

\[
b_n \neq b_n^*,
\]

then

\[
z(n) \neq 0,
\]

and the contribution \( \Delta \eta^* \) is

\[
\Delta \eta^*(z_n \neq 0) = \epsilon_n [a_0 (\epsilon_n + z_n)
+ a_1 (\epsilon_{n-1} + z_{n-1}) + \cdots].
\]

The difference between the two contributions to the weighted-noise power is

\[
\Delta \Delta \eta = \Delta \eta^* - \Delta \eta = z_n [a_0 (2\epsilon_n + z_n)
+ a_1 (\epsilon_{n-1} + z_{n-1}) + \cdots].
\]

According to Fig. 5 the following relation is true:

\[
e_n^* = \epsilon_n + z_n = v_n + \frac{z_n}{2}
\]

therefore

\[
2\epsilon_n + z_n = 2v_n.
\]

Thus

\[
\Delta \Delta \eta = 2z_n \left[ a_0 v_n + a_2 \frac{(\epsilon_{n-1} + z_{n-1}) + \cdots}.
\]

In order to achieve minimum weighted-noise power the comparator must make decisions according to the following rule:

\[
b_n^* = -b_n \quad \text{if} \quad \Delta \Delta \eta < 0
\]

and

\[
b_n^* = b_n \quad \text{if} \quad \Delta \Delta \eta > 0.
\]

This is equivalent to

\[
b_n^* = b_n \cdot \text{sgn} [\Delta \Delta \eta].
\]

By making use of the relationship

\[
z_n = |z_n| \text{sgn}(v_n) = 2db_n
\]

one obtains

\[
b_n^* = \text{sgn} \left\{ 2 \cdot K \cdot |z_n| \cdot b_n \left( v_n + \frac{a_1}{2} \epsilon_{n-1}^* + \frac{a_2}{2} \epsilon_{n-2}^* + \cdots \right) \right\} \cdot b_n
\]

or

\[
b_n^* = \text{sgn} \left\{ v_n + \frac{a_1}{2} \epsilon_{n-1}^* + \frac{a_2}{2} \epsilon_{n-2}^* + \cdots \right\}
\]

since \( K > 0 \). (8)

By comparing the above equation with (3), the sampled unit impulse response of the decision filter is easily obtained:

\[
g_0 = 1
\]

\[
g_j = \int_0^w \cos (2\pi f j T) N(f) |^2 df \int_0^w N(f) |^2 df
\]

for \( j = 1, 2, \cdots \)

\[
g_j = 0
\]

for \( j < 0 \). (9)

This decision filter optimizes the performance of the coder for a given noise-weighting criterion [which is contained in \( N(f) \)]. The optimization applies only to a coding process with suitable slope loading.

However, if the probability of slope overload is not negligible or if the coder is idling certain optimal decision filters can cause problems. In the first case, recovery from overload may become excessively long and in the second case the coder may generate undesirable idling patterns. Both conditions may be unacceptable. Thus, the recovery from overload as well as the idling behavior must be examined while optimizing the coder performance.

**D. The Idling and Overload Behavior of the System**

In the absence of an input signal the coder is in the idling state. It will generate repetitive idling patterns (i.e., repetitive binary sequences) which depend on the decision filter and on an initial charge condition in the integrating capacitor. In view of adaptive algorithms that will be discussed in Section IV and that increase the system's dynamic range, certain idling patterns may not be desirable. Therefore, a criterion for the acceptability of the idling behavior must first be defined.

For the definition the concept of the variation of the digital sum (DSV) will be used.

The variation of the digital sum is the difference between the largest and the smallest value of the running sum \( S_k \) of consecutive binary symbols

\[
S_k = \sum_{j=1}^k b(jT); \quad b(jT) = \pm 1,
\]

and

\[
\text{DSV} = S_{\max} - S_{\min}.
\]
It will become clear in Section IV, that “good” adaptive algorithms depend on the nonexistence of idling patterns containing many equal consecutive binary symbols; that means that the maximum variation of the digital sum within an idling pattern should be limited. On the other hand, it can be shown that the decision filter can only be effective (with respect to noise spectrum shaping), if the DSV is larger than 1. Therefore, we shall define the acceptability criterion as follows.

The system is acceptable if the variation of the digital sum of all possible idling patterns is less than or equal to 2.

With this criterion the acceptability of the system’s idling behavior can be tested. According to Fig. 6, a coder can generate a repetitive, binary sequence of length XL, i.e., a sequence \( \ldots b_n b_{n+1}, \ldots \) if for all sampling intervals \( nT \) the following condition is true:

\[
b_{n+i}' = b_{n+i}
\]

with \( b_{n+i}' \) being the response of the open-loop delta coding system to the sequence \( \ldots b_{n-1}, b_n \).

The above condition can be expressed in the form of a set of inequalities. With \( h_j \) being the unit impulse response of integrator and decision filter combined combined \( h_i = \sum_{i=0}^{j} g_i \),

We can write

\[-\text{sgn} \{ b_n h_0 + b_{n-1} h_2 + b_{n-2} h_3 + \ldots \} = b_{n+1}.\]

Through mathematical manipulation one obtains

\[-\text{sgn} \{ \sum_{j=0}^{\infty} b_{n-j} - b_n \otimes k \} = b_{n+1},\]

with

\[ k_j = \sum_{i=j}^{\infty} g_i \]

and \( b_n \otimes k \): convolution of the vectors \( b \) and \( k \), with

\[ k = k_1, k_2, k_3, \ldots \]

and

\[ b = b_n, b_{n-1}, b_{n-2}, \ldots \]

The term

\[ h_n \sum_{j=0}^{\infty} b_{n-j} = U_n \]

represents an initial condition of the system at time \( t = nT \). The coder can only generate a sequence \( \ldots b_n b_{n+1}, \ldots \), if the following conditions hold:

\[-\text{sgn} \{ U_n + \sum_{i=1}^{j} b_{n+i} - b_{n+j} \otimes k \} = b_{n+j+1}\]

for all \( j \geq 0 \).

This is equivalent to the following condition:

\[ b_{n+j+1} \{ U_n + \sum_{i=1}^{j} b_{n+i} - b_{n+j} \otimes k \} < 0 \]

for all \( j \geq 0 \). (10)

Idling sequences can exist only if there is a positive range of initial conditions \( U_n \) for which all the above inequalities (with \( j = 0, 1, 2, \ldots \) etc.) hold. On the other hand, if for a certain repetitive pattern that range is found to be negative, nonexistence of this pattern is guaranteed.

For various decision filters this range has been determined for repetitive binary patterns up to a length of 20 using a computer program. Two examples will be discussed in more detail; starting with a noise-weighting criterion the decision filter will be determined and the idling behavior of the coder examined.

In a first example, the combined postfilter, output-filter and noise-weighting filter \( N(f) \) is of the form (see Fig. 9)

\[ N_{\gamma}(f) = \frac{1}{1 + (f/f_0)^\gamma} \]

By applying (9) and an integral transformation \([35]\), the unit-impulse response of the decision filter is found:

\[ g_1(nT) = \exp (-nT/\tau) \]

with

\[ \tau = \frac{1}{2\pi f_0}. \]

The unit-impulse response \( q_1(nT) \) of a network combining the decision filter with the integrator is

\[ q_1(nT) = \sum_{j=0}^{n} g(jT) = d[1 - \exp (-T/\tau) \cdot \exp (-nT/\tau)] \]

with

\[ q_1(\infty) = d. \]

The corresponding network is identical to the network described in the literature as “double integration” \([2]\). Its form together with its frequency- and unit-impulse
response are shown in Fig. 7. For this example the acceptability of the system's idling behavior can be derived analytically [31]. It leads to the following condition:

\[ f_0 > 0.029 \cdot f_s \]

with

- \( f_s \) sampling frequency

and

- \( f_0 \) corner frequency between 6 dB/oct and 12 dB/oct slopes of frequency response (Fig. 7).

The diagram, demonstrating the existence of idling patterns, is shown in Fig. 8(a). For a noise-weighting criterion characterized by \( f_0 = 2 \) kHz, the idling behavior is acceptable at a sampling rate of 64 kHz.

In a second example the combined output- and noise-weighting filter is chosen to be that of the C-message weighting filter [36]. Only the section above 1000 Hz is used, however, since the encoding of low frequencies does not pose a problem in delta modulation. The source spectrum is chosen to be the long time spectrum of speech [37]. For the determination of the decision filter, the postfilter and thus the prefilter must be known. However, the prefilter that minimizes the weighted noise depends on the spectrum of the noise which again depends on the decision filter used. In order to avoid this complication, the entropy-maximizing prefilter \( F_E(f) \) will be used instead; it is only dependent on the source spectrum and on the leaky integrator, see (1b):

\[ |F_E(f)|^2 = S_{x*}(f^2 + f_0^2) \]

Thus the filter combining the post-, output-, and noise-weighting filter function \( N_z(f) \) has the response (Fig. 9)

\[ N_z^2(f) = C^2(f) \cdot S_s[f_n^2 + f^2] \]

with

- \( C(f) \) C-message criterion for \( f \geq 1000 \) Hz; C-message criterion at 1000 Hz for \( f < 1000 \) Hz.

From the filter function \( N_z^2(f) \), the unit-impulse response of the decision filter is derived which is used to test the existence of idling patterns. The result [Fig. 8(b)] shows unacceptable idling of the coder. Repetitive patterns exhibiting a variation of their digital sum exceeding 2 can exist up to a length of 18 sampling intervals.

The shaping of the noise spectrum by the two decision filters is shown in Fig. 10. The spectra were determined at 4x loading using a Gaussian source signal. The coding process without a decision filter generates a spectrum that is approximately white. The decision filter type [derived from the first noise-weighting criterion \( N_1(f) \)] shows a large reduction of the noise power below 6000 Hz. The decision filter type 2 [derived from \( N_2(f) \)] achieves an additional noise reduction of approximately 2 dB.

The above result applies to 4x-loading causing negligible slope overload. However, if the coder is slope overloaded, recovery from overload takes longer if an effective decision filter is used. The recovery is best analyzed by testing the unit-step response of the system. The result is seen in Fig. 11. In the absence of a decision filter, immediate recovery from overload takes place. The decision filter type 1 causes a recovery time of eight sampling intervals whereas with the decision filter type 2 the recovery tim
SCHINDLER: DELTA MODULATION

Fig. 10. Noise spectra at 4σ slope loading.
Curve A: No decision filter.
Curve B: Decision filter based on \( N_1(f) \) ("double integration").
Curve C: Decision filter based on \( N_2(f) \).

is extended to about 22 intervals. The poor recovery from slope overload does not justify the small noise advantage of the second decision filter. Finally, Fig. 12 shows the ratio of signal and weighted quantization noise (including slope-overload noise) as a function of slope loading (at \( f_s = 60 \text{ kHz} \)). The source spectrum had a 6 dB/oct roll-off above 300 Hz and was band-limited to 4 kHz. At 4σ loading and below, a gain in signal- to weighted-noise ratio is achieved by the decision filters. As expected, this gain is lost at slope-loading factors that cause nonnegligible slope overload.

III. NONLINEAR DELTA-MODULATION ALGORITHMS

We have seen how linear processes in delta modulation can improve the coding performance. Now the refinement of the coding process by nonlinear means will be investigated.

For the analysis we assume a stationary, Gaussian source, with a spectrum that rolls off at 6 dB/oct above 300 Hz and is band-limited to a frequency \( W \). The nonlinear algorithm to be found will be based on the concept of prediction: the history of the digits produced by the coder is used for the prediction of the future signal [38], [39]. As a consequence of oversampling in the delta-modulation process such a prediction is expected to be very effective. Fig. 13 shows the prediction mechanism. The binary coder decisions are known up to time \( nT \) and the signal level at time \((n + 1)T \) is predicted. Initially, the past five binary symbols \( b(n) \cdots b(n - 4) \) will be used for predic-
tion. Prediction will be with minimum rms error if the expected value of the signal at time \((n + 1)T\) is predicted [38]. Under very general assumptions the binary digit produced at time \((n + 1)T\) will then carry a maximum of information, it will be a \(-1\) or a \(+1\) with equal probability and thus be uncorrelated to the past binary symbols [31]. Fig. 13 illustrates the above. For given statistical properties of the source, each binary combination (SEQ) of the past and the present five binary comparator decisions will lead to an expected value of the source signal \(\tilde{q}\) at time \((n + 1)T\). Of interest is, of course, the step \(\tilde{q}(n)\) that makes the reconstructed signal at time \((n + 1)\) equal to this expected signal value.

For the prediction of \(\tilde{q}(n)\), the following algorithm (with yet unknown coefficients) is chosen:

\[
\tilde{q}(n) = k_0 \cdot b(n) \cdot q(n - 1) + k_1 \cdot q(n - 1) + k_2 \cdot q(n - 2)
\]

\[+ k_3 \cdot q(n - 3) + k_4 \cdot q(n - 4)
\]

\[+ k_{5\times5} \cdot b(n) \cdot q(n - 1) \cdot b(n - 1) \cdot b(n - 2)
\]

\[+ k_{6\times3} \cdot b(n) \cdot q(n - 1) \cdot b(n - 1) \cdot b(n - 3)
\]

\[+ \ldots
\]

\[+ k_{8\times2} \cdot q(n - 2) \cdot b(n - 3) \cdot b(n - 4)
\]

\[+ k_{9\times2} \cdot b(n) \cdot q(n - 1) \cdot b(n - 1) \cdot b(n - 2) \cdot b(n - 3) \cdot b(n - 4)
\]

with

\[q(j) = \text{step at time } jT.
\]

The prediction algorithm contains 16 terms. The first five terms represent the linear portion of the prediction process. This portion is basically the sum of the weighted past steps chosen by the coder for signal reconstruction. As can be seen, the first term deviates in its form from the other terms; this is because only the binary comparator decision \(b(n)\) is known at time \(nT\) and not yet the step size, which is to be determined. The unknown step size \(|q(n)|\) at time \(nT\) is replaced by a first estimate which is the step size used one sampling interval earlier, i.e., \(|q(n - 1)|\).

The nonlinear terms are products of a step size and past binary coder decisions. They provide higher order prediction. Let us assume that the signal statistics do not change if the polarity of the signal is inverted (symmetrical signal statistics). This means that pairs of waveforms which have equal magnitude but opposite polarity will be encoded by binary sequences of opposite polarity (see Fig. 13). Also the predicted step size \(\tilde{q}(n)\) will change polarity. For this reason the algorithm (11) cannot contain terms that are the product of an even number of the binary decision values \(b(n)\) through \(b(n - 4)\).

The change in step size can be represented as follows:

\[
\frac{|\tilde{q}(n)|}{|q(n - 1)|} = k_0 + k_1 b(n) \cdot b(n - 1)
\]

\[+ k_2 b(n) \cdot b(n - 2) \cdot \frac{|q(n - 2)|}{|q(n - 1)|}
\]

\[+ k_3 b(n) \cdot b(n - 3) \cdot \frac{|q(n - 3)|}{|q(n - 1)|}
\]

\[+ k_4 b(n) \cdot b(n - 4) \cdot \frac{|q(n - 4)|}{|q(n - 1)|}
\]

\[+ k_{5\times5} b(n - 1) \cdot b(n - 2) + \text{etc.} \quad (12)
\]

The prediction algorithm contains 16 unknown coefficients. They can be determined as follows; using compute simulation, the relative step sizes \(|\tilde{q}(n)|/|q(n - 1)|\) and the rms deviation \(\sigma_{\tilde{q} - q}\) around the predicted signal value can be determined for all combinations SEQ and for various slope-loading factors (SLF's), assuming constant step sizes up to time \(nT\). If we assume symmetrical signal statistics, only 16 such relative step sizes exist; five binary symbols are considered; this is due to the fact that symmetrical binary sequences (obtained through symbol inversion) yield identical step sizes. With the sixteenth values for the step sizes, the coefficients can be determined by solving the following set of equation (written in matrix representation):

\[
|B| \cdot |K| = |Q|
\]

with \(B\) being the matrix consisting of the following products:

\[
b(n) \cdot b(n), b(n) \cdot b(n - 1), b(n) \cdot b(n - 2), \ldots
\]

\[
b(n) \cdot b(n - 1) \cdot b(n - 2) \cdot b(n - 3), \ldots
\]

where each line in the matrix corresponds to one of 16 combinations of binary symbols. \(K\) is a column vector consisting of the coefficients

\[k_0, k_1, k_2, \ldots, k_{5\times5}, k_{9\times2} \ldots
\]

and \(Q\) is a column vector consisting of the values

\[|q(n)|
\]

...
two binary symbols as compared to prediction based on an infinite number of binary symbols. Accepting such a small loss, the prediction algorithm can be simplified to the point where only coefficients $k_0$ and $k_1$ remain. They are shown (with dotted lines) as a function of the SLF in Fig. 14. The simplified prediction algorithm is now of the form

$$\frac{|q(n)|}{|q(n-1)|} = k_0 + k_1 \cdot b(n) \cdot b(n-1) \quad (14)$$

or

$$\frac{|\tilde{q}(n)|}{|q(n-1)|} = k_0 + k_1 \quad \text{if } b(n) = b(n-1)$$

$$k_0 - k_1 \quad \text{if } b(n) \neq b(n-1).$$

This result looks similar to the algorithm of the adaptive delta modulation with a one-bit memory described in [8], but in our case the coefficients $k_0$ and $k_1$ are still dependent on the SLF, which is defined as

$$\text{SLF} = \frac{|q(n-1)|}{\sigma_F}$$

with $\sigma_F$ = rms value of signal slopes. These coefficients vary in general from sampling interval to sampling interval. By applying these rules and assuming a step-size quantization of 1 dB a state diagram for the coding process emerges (Fig. 16).

The state diagram shows that equal consecutive binary coder decisions cause a very rapid increase in step size; on the other hand, the decrease in step size due to unequal consecutive binary decisions is very gradual. Assuming a coarser quantization (e.g., 2 dB), a state diagram of simpler form results. Finally, the block diagram of such a delta-coding system, based on nonlinear prediction, is shown in Fig. 17. In addition to the linear delta coder, the system contains a discrete multiplier and a step-size logic that is the implementation of the algorithm derived above. The coding process based on the state-diagram algorithm was simulated on a computer. In the decoder an interpolator was introduced (see Fig. 17). In contrast to the algorithm of the coder, the interpolator in the decoder can make use of past and future binary symbols. The interpolation algorithm was derived by applying methods similar to those developed for the determination of the coder algorithm.

The computer simulation yielded several results. A first one showed that for a prediction based on five binary symbols, the binary signal produced by the coder was almost completely decorrelated; this proves the effectiveness of the nonlinear algorithm. A second result, the ratio of signal-to-weighted-noise power as a function of SLF is shown in Fig. 18. The SLF $\sigma_F/\sigma_{AL0}$ is normalized to 0 dB for those signal slope values that were used in the derivation of the nonlinear prediction algorithm ($\sigma_{AL0}$). Curve A shows the S/N ratio produced by the coder-decoder with interpolator and output filter removed. As expected, the S/N ratio reaches its peak where $\sigma_F$ is
Fig. 16. State diagram of nonlinear delta coder. Gaussian source, spectrally matched to coder but band-limited to 4 kHz. Sampling rate $f_s = 60$ kHz. Sequence length $ML = 2$. Quantization $= 1$ dB. The decibel scale is with respect to the rms slope values.

Fig. 17. Discrete, nonlinear delta modulation: block diagram.

equal to $\sigma_{d,B}$. With interpolator and output filter (4 kHz low-pass) inserted (curve C), the peak is reached at a lower slope-loading level; in this case the step sizes are generally too large, causing anticorrelation of consecutive binary coder decisions. This results in a shaping of the noise spectrum, with high-frequency noise being emphasized. Since this portion of the noise is eliminated by the output filter, an overall increased S/N ratio is observed.

A comparison of these results with those obtained by a linear coder (based on the decision filter) shows that a linear delta coder achieves a higher ratio of signal-to-weighted noise. The spectrum-shaping capability of the decision filter is more effective than the nonlinear prediction algorithm. However, in the case of nonlinear delta modulation, the weighted SNR degrades more slowly, if the signal power is reduced. This statement applies to delta-coding systems operating at a high ratio of sampling rate to signal bandwidth (e.g., $f_s = 64$ kb/s, $W = 4$ kHz).

Fig. 18. S/N ratio of nonlinear delta coder as a function of slope loading. Gaussian source, spectrally matched to coder, band-limited to 4 kHz. Sequence length $ML = 2$. Sampling rate $= 60$ kHz. Curve A: No interpolation. Curve B: With interpolation. Curve C: With interpolation and low-pass filtering (4 kHz).

IV. ADAPTIVE ALGORITHMS FOR THE ENCODING OF NONSTATIONARY SIGNALS

The delta-coding methods analyzed so far were based on the assumption of a stationary signal source. However, many signals, like speech, are not stationary. Speech levels vary during a conversation and also depend on the speaker. Level variations often exceed 20 dB. Linear and nonlinear delta coders (Figs. 12 and 18) maintain an optimum S/N ratio only over a small range of input levels. Signals whose levels lie outside this range are encoded either with excessive granular or slope-overload noise. That means that these systems lack dynamic range.

The design of delta-coding systems having a wide dynamic range is possible by dynamically adapting the step size to the rms slope value of the source signal. This adaption process can either be forward controlled, with a control signal transmitted separately to the receiver, or it can be feedback controlled. In the latter case the step-size control is derived from the binary output of the coder and no separate control information has to be transmitted. For this reason, only the feedback approach will be investigated further.

The question then arises: what are the properties of the binary sequence produced by the coder that best indicate proper, excessive or inadequate slope loading? A first property to be considered is the probability of occurrence of two equal consecutive bits, or related to this, the first autocovariance coefficient of the binary signal. A look at the idling behavior of the linear and nonlinear coder however, shows that this property is not suitable. In both coders, idling patterns can exist that contain two equal consecutive bits [see Fig. 8(a) for the linear coder; see Fig. 16 for the nonlinear coder with the following idling pattern: $+1+1-1-1+1+1\ldots$ and the associated step sizes $+3,+6,+3,+6,+3\ldots$ dB]. The selection of such patterns depends on an initial-charge condition in the integrating capacitor, which is not well defined in a practical coding system. With the probability of occurrence of two equal consecutive bits being finite
for input signals of zero amplitude, the considered property is ambiguous and not very useful for step-size control.

Considering three consecutive binary symbols, a more suitable criterion can be found. Fig. 19 shows the probability of occurrence of three equal consecutive bits $p_{\{LLL/000\}}$ as a function of the SLF for the linear and the nonlinear delta coder. In the same diagram the S/N ratio is shown. In the range of optimum S/N, the probability $p_{\{LLL/000\}}$ varies very rapidly with slope loading.

On the other hand, three equal consecutive bits never occur in idling patterns if the coder is designed properly (linear coder with DSV $\leq 2$). For these reasons, this criterion is used for dynamic control of the step size.

A. Adaptive Algorithms for Speech Signals

Speech belongs to the class of nonstationary signal sources. The spectral properties as well as the power level vary with time. While encoding speech waveforms, a delta coder with dynamic step-size control continuously adjusts its step size to the derivative of the source signal and therefore follows variations in the short-term spectrum and the power level.

It is known that the power level of speech can rise very rapidly after pauses in a conversation; the decay of speech levels is, in general, slower due to the mechanism of speech generation. In a delta coder designed for the encoding of speech, the adaption algorithm should be matched to these dynamic properties of speech. In order to determine such an algorithm, the logarithm of the absolute value of the differences between consecutive speech samples was analyzed as a function of time. This logarithmic value is a measure for the step size in decibels that is required by the coder in order to follow a speech waveform. Fig. 20 shows a typical response. It is taken from the first three words of the sentence “cats and dogs each hate the other,” spoken by a female speaker. This sentence was chosen for its very rapid transition from small to large slope values and its phonetical balance. Fig. 20(a) shows the envelope of all the logarithmic values taken at time intervals of $16.7 \mu s$ which corresponds to a sampling rate of 60 kHz; on the other hand, Fig. 20(b) shows details of the first, very rapid transition (“c” of the word “cat”).

From Fig. 20 the following can be concluded for a coder operating at 60 kb/s.

1) In order to follow rapid transitions in speech levels (e.g., explosives), the coder must be capable of increasing its step size by at least 2 dB per sampling interval.

2) In smooth sections of speech, a smaller increase in step size (like 1 dB), might be desirable, since it leads to reduced granular noise.

3) The decrease in step size may be more gradual and should be in the order of 0.1 dB per sampling interval.

These conclusions in conjunction with the criterion for the control of proper slope loading lead to the adaption algorithm shown in Table I.

The symbol X indicates that any binary symbol may take its place. This algorithm is based on four consecutive binary symbols and thus permits the distinction between smooth and rapid increase in step size. In the implementation of an adaptive delta-modulation system, the small changes in step size of $-0.125$ dB would be accumulated in a counter until they reach an integer decibel value. Only then would the step size actually be reduced. This adaption process is applicable to both linear and nonlinear delta coders. In the latter case, the whole range of step sizes of the state diagram is shifted under control of the adaption algorithm.

Using a computer, the linear adaptive and the nonlinear adaptive coding process were simulated and the
S/N ratio (rms speech/rms noise) was determined for real speech signals. Both coders’ adaption processes were limited to a ratio of largest-to-smallest step size (compression ratio) of 30 dB. The sentence used in the simulation was: “the bill was paid every third week.” It was low-pass filtered to 4 kHz. The sampling rates were 60 and 30 kHz.

The results can be seen in Fig. 21. For comparison purposes the corresponding curves for 7- and 8-bit PCM (A-law) are shown. At equal bit rates the performance of the linear adaptive delta coder (S/N ratio, dynamic range) is comparable to that of compacted PCM; this applies to the range of bit rates that is used in today’s toll quality telephony. The nonlinear adaptive delta coder achieves a S/N ratio that lies approximately 3 dB below that of the linear adaptive system.

On the other hand, the performance of delta modulation at 30 kHz shows that delta modulation degrades very slowly with a reduction in the bit rate. This is in contrast to PCM where the degradation in S/N ratio is much faster. At 30 kHz, the linear adaptive and the nonlinear adaptive delta coder perform about equally well at medium speech levels.

At low speech levels, nonlinear adaptive delta modulation shows for both bit rates an SNR advantage over linear adaptive delta modulation. This is due to the large range of step sizes available in the state diagram of nonlinear delta modulation. This range together with the range of the adaption process result in the extended dynamic range of nonlinear delta modulation (Fig. 21, curves B and D).

V. CONCLUSIONS

Starting with single-integration, fixed step-size delta modulation we have shown how delta coding can be refined by the introduction of linear and nonlinear processes. A special form of a nonlinear process is the adaptive step-size algorithm which provides a large dynamic range to the system.

In comparing the different refinements, one sees that the (linear) decision filter is the most effective method for obtaining a best possible ratio of signal-to-weighted-noise power. This applies to delta-coding systems operating at bit rates exceeding 30 kb/s and source signals (speech or Gaussian) with speech-like spectra. The nonlinear delta-modulation system does not perform as well in this range of bit rates. It shows a lower peak in signal-to-weighted-noise ratio; however this ratio degrades more slowly if the signal power is reduced.

The adaption process that is superimposed on the other refinements is most effective for providing a large dynamic range to the delta-coding systems. As was shown, the adaption process can be matched to the dynamic behavior of speech. A constant ratio of signal-to-weighted-noise power is obtained for a large range of signal levels. This is important if a coding system is designed for use in telephony.

Thus, in the range of bit rates exceeding 30 kb/s, a delta-modulation system based on a prefilter (together with a postfilter), a decision filter and an adaptive algorithm gives the best overall performance.

A comparison with compacted PCM (A-law) shows that linear adaptive delta modulation performs better at bit rates below about 60 kb/s; above 60 kb/s PCM achieves a higher S/N ratio. In subjective tests linear adaptive delta modulation is in general preferred over PCM, if both systems operate at equal bit rates in the range of 20 to 60 kb/s. This is due to several facts. First, the frequency response of a delta-modulation system may have a very gentle roll-off since the noise is reduced very little by sharp filtering. This provides a great naturalness
to delta-encoded speech. In contrast, PCM requires sharp filtering at half the sampling rate with most of the fricatives being removed. Secondly, adaptive delta modulation generates noise that increases with the slope value of the signal. Since the fricatives contain very high-slope values, the largest amount of noise is produced when these sounds are present. Noise in fricatives, however, is least disturbing since the character of fricatives is somewhat similar to that of noise.

At low bit rates (15 to 40 kHz), delta modulation proves to be especially useful. The quality of the decoded signal degrades much slower than that of PCM if the bit rate is reduced. The mathematical treatment of the coding process (except that of nonadaptive single integration delta modulation) however becomes extremely complex. Furthermore, the important criterion at these low bit rates is intelligibility and not necessarily S/N ratio or entropy. With heuristic optimization methods good results were achieved at bit rates as low as 15 kHz [40].

Today's applications of delta modulation lie in areas where low bit rate encoding of speech, a high tolerance to errors in the digital transmission path or inexpensive coder hardware (single-channel encoding) is required.

APPENDIX

D. J. Goodman describes in his paper "Delta Modulation Granular Quantizing Noise" [21] a statistical analysis of a single integration delta-modulation system in which slope-overload effects are negligible. He considers input signals that are stationary Gaussian. For the autocovariance terms of the error signal he finds [21, eq. (25)]

\[ Q_0 \sigma^2 = \frac{64\pi^2}{3F^2} \]

\[ Q_s \sigma^2 = \frac{256}{F^2} \sum_{k=1}^{\infty} \frac{(-1)^{sk}}{k^2} \exp \left( -\frac{F^2k^2}{64} \right) \sinh \left( \frac{F^2k^2\rho_s}{64} \right) \]

with

\[ F = f_s/f_e \]

\[ f_s = \left( \int_{-\infty}^{\infty} S_{xx}(f) \, df \right) \left( \int_{-\infty}^{\infty} S_{xx}(f) \, df \right)^{1/2} \]

The symbols are

- \( Q_0, Q_s \) autocovariance coefficients of the error signal \( e_i \) (see Fig. 4; decision filter replaced by a straight connection)
- \( \rho_s \) normalized autocovariance coefficients of input signal \( y_i \)
- \( \sigma_s^2 \) power of input signal \( y_i \) (Fig. 4)
- \( d \) step size of delta coder
- \( f_s \) sampling frequency
- \( f_e \) effective bandwidth of input signal

We shall show that, for certain assumptions concerning the signal source and for large values of \( F \) (\( F > 64 \)), the error samples are essentially uncorrelated and independent of the spectrum of the source signal. (\( F = 64 \) results from an effective signal bandwidth of 800 Hz, which is approximately equal to that of speech, and a sampling rate of 51 kHz.) We can write

\[ \frac{Q_s}{Q_0} \approx \frac{12}{\pi^2} \sum_{k=1}^{\infty} \frac{(-1)^{sk}}{k^2} \exp \left( -\frac{F^2k^2}{64} \right) \sinh \left( \frac{F^2k^2\rho_s}{64} \right) \]

\[ = \frac{6}{\pi^2} \sum_{k=1}^{\infty} \frac{(-1)^{sk}}{k^2} \exp \left[ -64k^2(1 - \rho_s) \right] \]

\[ - \exp \left[ -64k^2(1 + \rho_s) \right] \]

\[ \frac{Q_s}{Q_0} \leq 0.61 \sum_{k=1}^{\infty} \frac{1}{k^2} \exp \left[ -64k^2(1 - \rho_s) \right] \]

\[ + \exp \left[ -64k^2(1 + \rho_s) \right]. \]

Let us assume that the autocovariance coefficients of the source signal are bounded as follows:

\[ -0.95 \leq \rho_s \leq 0.95 \quad \text{for} \quad |\mu| = 1 \]

\[ -0.9 \leq \rho_s \leq 0.9 \quad \text{for} \quad |\mu| \geq 2. \]

Then

\[ 0.05 \leq (1 - \rho_s) \leq 1.95 \quad \text{for} \quad |\mu| = 1 \]

\[ 0.05 \leq (1 + \rho_s) \leq 1.95 \quad \text{for} \quad |\mu| = 1 \]

and

\[ 0.1 \leq (1 - \rho_s) \leq 1.9 \quad \text{for} \quad |\mu| \geq 2. \]

\[ 0.1 \leq (1 + \rho_s) \leq 1.9 \quad \text{for} \quad |\mu| \geq 2. \]

This leads to the following approximation:

\[ \frac{Q_s}{Q_0} \approx 0.61 \frac{1}{k^2} \exp \left[ -64k^2(1 - |\rho_s|) \right] \]

for \( \mu \neq 0 \)

and thus

\[ \left| \frac{Q_1}{Q_0} \right| \approx 0.024 \]

and

\[ \left| \frac{Q_s}{Q_0} \right| \approx 0.4 \cdot 10^{-4} \quad \text{for} \quad |\mu| \geq 2. \]

With the assumptions made on the signal source (they are not very restrictive) and on \( F \) we see that in single integration delta modulation the error samples are very weakly correlated. The power of the error samples is approximately that of a random variable, distributed uniformly over an interval of length \( 2d \). The spectrum of
the error samples is approximately flat (see Fig. 10, curve A) and highly independent of the source spectrum.

As described in Section II-C, the decision filter changes some comparator decision and therefore also some of the error samples. Since the complete decision process can be described on the basis of the error samples alone, the new spectrum of the error samples (obtained with decision filter) is again essentially independent of the spectrum of the signal source.

We shall now demonstrate that the error signal does not contain components that are correlated to the input signal if $F > 42$. We can write

$$Q_n = E[x_n y_{n+p}] = E[(y_n - r_n) (y_{n+p} - r_{n+p})]$$

$$= E[x_n y_{n+p}] - E[y_{n+p} r_n] + E[r_n r_{n+p}]$$

with

$y_n$ samples of the input signal at time $nT$.

$r_n$ samples of the reconstructed signal at time $nT$ (see Fig. 5). Note that $r_n$ corresponds to $x_n$ in Goodman’s paper.

$E$ expectation operator.

Resolving for the cross correlation between error signal and input signal

$$E[x_n y_{n+p}] = Q_n + E[y_{n+p} r_n] - E[r_n r_{n+p}].$$

From [21, eqs. (23)-(25)] we obtain for $F > 42$

$$E[y_{n+p} r_n] \approx \sigma^2 \rho_n$$

and

$$E[r_n r_{n+p}] \approx \sigma^2 \rho_n + Q_n.$$  

This leads to

$$E[x_n y_{n+p}] \approx Q_n + \sigma^2 \rho_n - \sigma^2 \rho_n - Q_n = 0.$$  

Thus it can be concluded that, for the assumption made, the total power of the error signal is audible as noise.

ACKNOWLEDGMENT

The author is deeply indebted to Prof. Dr. E. Baumann and Prof. H. Weber from the Swiss Federal Institute of Technology, Zurich, Switzerland, for their interest and their valuable suggestions. He wishes to thank K. E. Drangeid, Director of the IBM Zurich Research Laboratory, Rüschlikon, Switzerland, for his steady support. Also, Dr. E. Hopner’s continuous encouragement is gratefully acknowledged. Many colleagues at IBM in particular Dr. A. H. Frei and P. Vettiger, have contributed through stimulating discussions and numerous suggestions; a special thanks goes to them.

REFERENCES


Baseband Pulse Transmission in a Chain of Analog Repeaters

YOSHITAKA TAKASAKI, MEMBER, IEEE, AND NARIMICHI MAEDA, MEMBER, IEEE

I. INTRODUCTION

Pulse transmission in a chain of analog repeaters placed between digital repeaters has been studied and is called hybrid digital transmission [1], [2]. Two kinds of systems have been studied in this field; a passband hybrid digital transmission and a baseband one. The former utilizes existing analog transmission systems and provides a digital network compatible to analog systems [3]. It lays stress on compatibility and is not necessarily the best solution from the standpoint of economy. On the other hand, the latter is expected to provide an economical digital transmission line [4] by using analog repeaters specially designed for the system. In this system, however, accumulation of distortions in the chain of analog repeaters is so large that it has been difficult to realize economical analog repeaters that counteract these distortions.

The main causes of distortions introduced in analog repeaters are: 1) low-frequency cutoffs, 2) high-frequency cutoffs, 3) nonlinearities in amplifiers, and 4) equalization errors. In previous papers [5], [6] simple methods that compensate for causes 1) and 2) are reported. Cause 3) is studied in [7]. It seems, however, more investigation will be required in order to realize an economical baseband hybrid digital transmission system.