H. Kobayashi

Coding Schemes for Reduction of Intersymbol Interference in Data Transmission Systems

Abstract: Various coding schemes and their effects on intersymbol interference in pulse amplitude modulation systems are discussed. First, the relation between imperfections in the baseband-equivalent channel and intersymbol interference is clarified and applied to explain the effect of correlative level coding and Gorog's frequency concept codes in reducing intersymbol interference. Another coding scheme is then introduced: construction of codes in the time domain with intersymbol interference directly in mind. A decimal code of length 4 and an alphanumeric code of length 6 are proposed as practical codes and their properties are discussed. Simulation results are presented to give quantitative comparison of these coding techniques. Curves of the vertical eye-opening vs transmission rate have been produced and we show that codes designed in the time domain achieve better performance than both the frequency concept codes and the conventional codes for a wide class of channel characteristics.

Introduction

Intersymbol interference is one of the most severe limitations encountered in high-speed data transmission systems.[1,2] This paper reports an investigation of the relation between channel imperfections and intersymbol interference in a baseband pulse amplitude modulation (PAM) system to determine the effects of various coding schemes on the reduction of intersymbol interference. When the input binary (or in general multilevel) sequences are completely random, i.e., when there is no restriction on the level of the successive digits, equalization of the channels is crucial and one cannot send sequences faster than the Nyquist rate. However, this is not the case when some restriction is imposed on the values that successive digits can take.

In this paper we assume that the channel is at baseband. In a real communication system this is seldom the case; some type of modulation is adopted to transport the signal energy to the frequency band best suited to the transmission medium and, at the receiver, the corresponding demodulation process translates the signal energy back to baseband. It is known that linear modulation techniques are well suited for high-speed data transmission because of their efficient use of available bandwidth, and the effect of linear modulation can be considered as a portion of the baseband channel.[1] The communication model we use is illustrated in Fig. 1; the message sequence to the signal generator is represented by a sequence of amplitude-

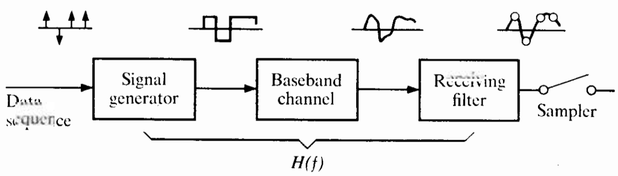


Figure 1 Baseband pulse amplitude modulation (PAM) system.

modulated impulses. This PAM system is characterized by a transfer function $H(f)$, which summarizes the overall frequency characteristics of the signal generator, the baseband channel and the receiving filter (including the equalizer).

Let $h(t)$ be the inverse Fourier transform of the system transfer function $H(f)$ and let T be the interval in seconds between successive symbols. The discrete impulse response function at the sampler output is

$$h^*(t) = \sum_{n=-\infty}^{\infty} h(nT)\,\delta(t - nT), \tag{1}$$

where an asterisk is used to denote a sampled function; the Fourier spectrum of Eq. (1) is

$$H^*(f) = T^{-1} \sum_{m=-\infty}^{\infty} H(f - mT^{-1}). \tag{2}$$

It is clear from Eq. (2) that $H^*(f)$ is periodic with period T^{-1}. Figure 2(a) is an example of $H(f)$ and the corre-

The author is located at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598.

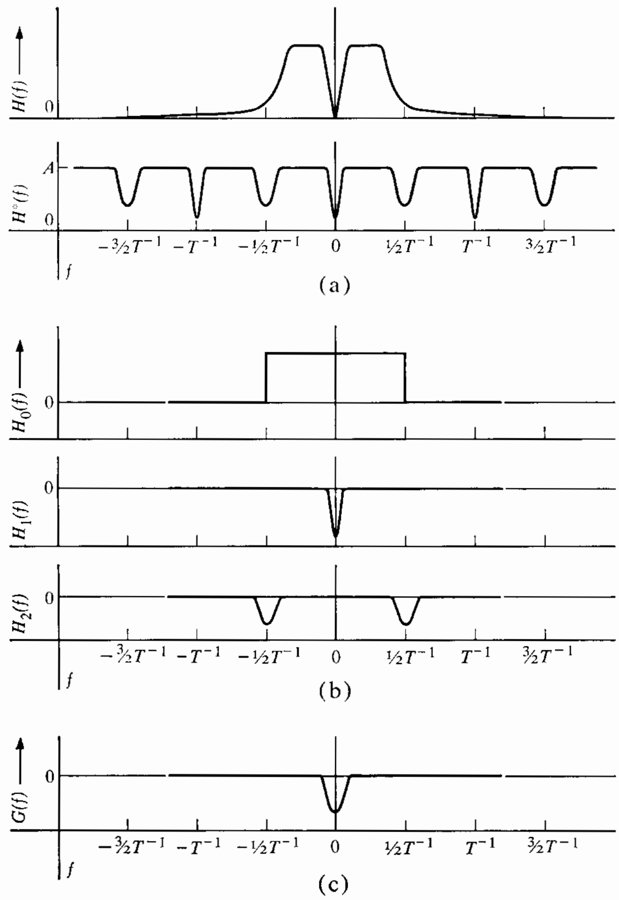


Figure 2 (a) The system transfer function $H(f)$ and the sampled function $H^*(f)$; (b) the equivalent components $H_0(f)$, $H_1(f)$ and $H_2(f)$ of $H(f)$; (c) the low-pass equivalent $G(f)$ of $H(f)$.

sponding $H^*(f)$. If $H^*(f)$ were to take on a constant value A for the entire frequency region $-\infty < f < \infty$, then

$$h^*(nT) = A\delta_{n,0} \tag{3}$$

and no intersymbol interference would occur.

In most practical cases heavy distortion exists at the high frequency end of the channel. Distortion also exists near the zero frequency end of the (equivalent baseband) channel when single sideband modulation is used. Even when vestigial sideband modulation is adopted, a pilot carrier tone is usually added in quadrature so that the carrier frequency and its phase can be recovered at the receiving end and, for satisfactory recovery, the data spectrum near the carrier frequency must be suppressed. These operations result in distortion at the zero frequency end of the corresponding baseband channel. It is convenient to introduce the following transfer functions, which are shown in Fig. 2(b):

$$H_0(f) = \begin{cases} AT, & |f| \le \frac{1}{2}T^{-1}; \\ 0, & |f| > \frac{1}{2}T^{-1}; \end{cases} \tag{4}$$

$$H_1(f) = \begin{cases} T[H^*(f) - A], & |f| \le \frac{1}{4}T^{-1}; \\ 0, & \text{elsewhere}; \end{cases} \tag{5}$$

and

$$H_2(f) = \begin{cases} \frac{1}{2}T[H^*(f) - A], & \frac{1}{4}T^{-1} \le |f| \le \frac{3}{4}T^{-1}; \\ 0, & \text{elsewhere}. \end{cases} \tag{6}$$

Here $H_0(f)$ is an ideal low-pass filter and $H_1(f)$ and $H_2(f)$ represent the channel imperfection at the lower and upper frequency ends, respectively. We note that

$$H^*(f) = T^{-1} \sum_{m=-\infty}^{\infty} [H_0(f - mT^{-1}) + H_1(f - mT^{-1}) + H_2(f - mT^{-1})]. \tag{7}$$

This equation implies that the given channel $H(f)$ is equivalent to the channel $H_0(f) + H_1(f) + H_2(f)$, as far as the sampled output is concerned.

The function $H_2(f)$ is representable in terms of its low-pass equivalent shown in Fig. 2(c):

$$H_2(f) = \tfrac{1}{2}G(f + \tfrac{1}{2}T^{-1}) + \tfrac{1}{2}G(f - \tfrac{1}{2}T^{-1}) \tag{8}$$

with

$$G(-f) = \tilde{G}(f), \tag{9}$$

where the tilde denotes the complex conjugate. One can infer from Eq. (8) that

$$h_2(t) = g(t)\cos(\pi t T^{-1}), \tag{10}$$

and therefore the sampled response is given by

$$h_2(nT) = (-1)^n g(nT). \tag{11}$$

An important practical situation is one in which the deviation from a perfectly equalized channel occurs in a narrow region at the high frequency end. In this case the function $g(t)$ is slowly varying compared with the sampling rate T^{-1} and therefore $h_2(nT)$ is an alternating sequence except for pairs of values adjoining the zeros of $g(t)$. Similarly, in the case in which the width of $H_1(f)$ is small compared with the value of $\frac{1}{2}T^{-1}$, $h_1(nT)$ changes sign infrequently.

Coding schemes for reduction of intersymbol interference

In this section the relation established in the previous section between channel imperfections and the channel response function is used to evaluate the effects of various coding schemes in reducing intersymbol interference. The schemes we discuss include the correlative level coding devised by Lender,[3] van Gerwen[4] and Kretzmer[5]; Gorog's frequency concept codes[6]; and time-domain-designed alphabets of the present author.[7]

• *Correlative level coding*

Let us assume that we are given a low-pass channel whose distortion exists at the high frequency end only. Then the sequence describing the intersymbol interference associated with this channel is $h_2(nT)$ of Eq. (11). Hence, the sequence

$$\tfrac{1}{2}\{h_2(nT) + h_2[(n-1)T]\}, \qquad (12)$$

which is the arithmetic mean of adjacent intersymbol interference, is a sequence with a small magnitude. This new sequence of reduced intersymbol interference can be obtained by use of correlative level coding[3] or the partial response channel.[5] Consider a linear filter with the impulse response function

$$\tfrac{1}{2}[\delta(t) + \delta(t-T)]. \qquad (13)$$

A binary impulse sequence of $+1$'s and -1's is transformed by this filter into a three-level sequence of $+1$'s, 0's and -1's and correlation is introduced among output digits. Let us denote the transformation, Eq. (13), of a two-level sequence into a three-level sequence [in general, an m-level sequence into a $(2m - 1)$-level sequence] by "correlative level coding, type I." To recover the original binary sequence from the received three-level signal without error propagation, a precoding operation is performed at the transmitter side on the input sequence.[3-5]

Two successive $+1$'s (or -1's) always have an even number of intervening zeros and successive $+1$ and -1 (or -1 and $+1$) always have an odd number of intervening zeros. Thus direct transition between top and bottom levels never occurs and, in fact, this is a well-known explanation[3] of why little intersymbol interference is observed with this type of correlative level coding. Another way of viewing the effect of correlative level coding on intersymbol interference is in the frequency domain. Since the transfer function of the correlative level coder (13) is $\cos(\pi fT)\exp(-j\pi fT)$, the frequency characteristic which contributes to intersymbol interference,

$$H_2(f)\cos(\pi fT)\exp(-j\pi fT), \qquad (14)$$

is much smaller in magnitude than the original $H_2(f)$.

Now let us assume that we are given a channel like that shown in Fig. 2(a), i.e., a channel containing imperfections at both ends of the given bandwidth. In this case we are faced with the two types of intersymbol interference $h_1(t)$ and $h_2(t)$ discussed in the introduction. A simple method that can reduce both types of interference simultaneously is the one which takes the difference between digits that are separated by one digit; that is, a linear filter with the response function

$$\tfrac{1}{2}[\delta(t) - \delta(t-2T)]. \qquad (15)$$

Binary sequences are again transformed into three-level sequences and we call this transformation "correlative level coding, type II." Reference 7 contains a more detailed analysis of intersymbol interference in the correlative level coding scheme. The disadvantage of this method comes from the increased number of signal levels, which results in a reduction of the signal-to-noise ratio in those applications in which the peak power or the average power of the transmitted signal is fixed in magnitude.

• *Codes designed in the frequency domain*

In a recent paper[6] Gorog proposed a group of codes called "frequency concept codes" that are designed to allow an efficient transmission of digital data. Let $a_1, a_2, \cdots, a_N$ be a sequence of length N. The frequency spectrum of this finite sequence is

$$S(f) = \sum_{n=1}^{N} a_n \exp[-j2\pi f(n-1)T]. \qquad (16)$$

Clearly $S(f)$ is a periodic function of frequency f with period T^{-1}.

If an input sequence is chosen so that it contains relatively little energy in the imperfect regions of the channel frequency characteristic, the amount of distortion to the sequence will not be great. This is the basis for the construction of the codes. Since the spectrum $S(f)$ is a continuous function, a sequence with the conditions

$$S(0) = 0 \quad \text{or} \quad \sum_{n=1}^{N} a_n = 0 \qquad (17)$$

and

$$S(\tfrac{1}{2}T^{-1}) = 0 \quad \text{or} \quad \sum_{n=1}^{N} (-1)^{n-1} a_n = 0 \qquad (18)$$

provides the desired property. By specifying the possible amplitude levels and the length N, we can determine a class of sequences that satisfy the conditions (17) and (18).

If the channel characteristic contains imperfections just in the upper frequency region, only the property (18) is required of the input sequence. For the case in which N is an odd number and $\{a_n\}$ is a binary sequence, the condition (18) should be modified to

$$\sum_{n=1}^{N} (-1)^{n-1} a_n = -1 \text{ (or } +1). \qquad (19)$$

Then a composite sequence of length $2N$, which consists of two such sequences of length N, satisfies the original condition (18). An interesting class of such sequences is Gorog's decimal code of length 5, which is designed to replace the conventional decimal codes. This class of codes is given in Table 1. Note that the binary representation of 1 and 0 is adopted in the table instead of $+1$ and -1.

Now consider the performance of these codes from the viewpoint of intersymbol interference. Let us assume that $H_2(f)$, the distortion characteristic in the high fre-

Table 1 Gorog's decimal code of length 5.

Decimal numeral	*Binary code*[a]
0	0 0 0 0 0
1	0 0 0 1 1
2	0 0 1 1 0
3	0 1 0 0 1
4	0 1 1 0 0
5	0 1 1 1 1
6	1 0 0 1 0
7	1 1 0 0 0
8	1 1 0 1 1
9	1 1 1 1 0

[a] The 1's and 0's correspond to the +1's and −1's, respectively, of the analysis.

Table 2 Gorog's alphanumeric code of length 8.

Alpha-numeric	*Binary code*[a]	*Alpha-numeric*	*Binary code*[a]
0	1 1 0 0 1 1 0 0	I	1 1 0 0 0 1 1 0
1	0 1 1 0 0 1 1 0	J	0 1 0 1 1 0 1 0
2	0 0 1 1 0 0 1 1	K	0 0 1 0 1 1 0 1
3	1 0 0 1 1 0 0 1	L	1 0 0 1 0 1 1 0
4	1 1 1 0 0 1 0 0	M	0 1 0 0 1 0 1 1
5	0 1 1 1 0 0 1 0	N	0 1 1 0 0 0 1 1
6	0 0 1 1 1 0 0 1	O	1 0 1 1 0 0 0 1
7	1 0 0 1 1 1 0 0	P	1 1 1 0 0 0 0 1
8	0 1 0 0 1 1 1 0	Q	1 1 0 0 0 0 1 1
9	0 0 1 0 0 1 1 1	R	1 1 0 1 1 0 0 0
A	1 0 0 1 0 0 1 1	S	0 1 1 0 1 1 0 0
B	1 0 1 0 0 1 0 1	T	0 0 1 1 0 1 1 0
C	1 1 0 1 0 0 1 0	U	1 0 0 0 0 1 1 1
D	1 1 0 0 1 0 0 1	V	0 0 0 0 1 1 1 1
E	0 0 0 1 1 0 1 1	W	0 0 0 1 1 1 1 0
F	0 1 1 0 1 0 0 1	X	0 0 1 1 1 1 0 0
G	1 0 1 1 0 1 0 0	Y	0 1 1 1 1 0 0 0
H	1 0 0 0 1 1 0 1	Z	1 1 1 1 0 0 0 0

[a] The 1's and 0's correspond to the +1's and −1's, respectively, of the analysis.

quency region, has a width that is small compared with the value of $(NT)^{-1}$, which is the frequency of occurrence of every codeword. Then $h_2(nT)$ is an alternating sequence whose magnitude changes slowly over the length of a codeword, and the intersymbol interference is effectively canceled because of condition (18).

If the channel characteristic contains imperfections at both the lower and upper frequency ends, sequences satisfying Eqs. (17) and (18), or equivalently

$$\sum_{n=1}^{N/2} a_{2n-1} = 0 \tag{20}$$

and

$$\sum_{n=1}^{N/2} a_{2n} = 0, \tag{21}$$

yield reduced intersymbol interference. For example, in the case of $N = 8$, there are 36 binary codewords that satisfy the two conditions. This code is given in Table 2 and has been suggested by Gorog[6] for use as an alphanumeric code.

On examining this class of codes, we note the following shortcomings: When $H_2(f)$ and $H_1(f)$ are not narrow compared with $(NT)^{-1}$, the reduction of intersymbol interference is not significant because the interference from neighboring digits dominates that from remote digits. Furthermore, the effect on intersymbol interference of the outer digits is not taken into account in codes that are selected using the conditions (20) and (21). These shortcomings are discussed further in subsequent sections.

- *Codes designed in the time domain*

In designing codes in the time domain, restrictions are placed in the following way on the sequence allowed for transmission. Consider a number of possible states. In each state only certain symbols (1 or 0 in the binary case) can be generated. When one of these symbols has been generated, the state changes to a new state depending on both the old state and the symbol generated. For example, consider the generation of binary data to be sent over a channel which has an imperfection at the high frequency end. Sequences containing 0 and 1 alternatively over many consecutive digits should be avoided because these sequences match the distortion sequence of Eq. (11) and yield heavy intersymbol interference at the channel output. The system which generates sequences under such constraints can be called the "discrete noiseless channel" according to Shannon's terminology[8] and the constraint can be indicated best in a linear graph as shown in Fig. 3, where the circles represent the possible states. The constraint in Fig. 3(a) requires that alternation of 1 and 0 does not occur over more than four consecutive digits. The maximum information rate generated is given by the channel capacity

$$C = \lim_{L\to\infty} L^{-1} \log_2 N(L) \quad \text{(in bits/sec)}, \tag{22}$$

where $N(L)$ is the number of allowed codewords of length L (sec). Using the computation formula given by Shannon,[8] we calculate the capacity of the sequence generator with the constraint of Fig. 3(a) as $C = 0.9469T^{-1}$ bits/sec, where T is the duration in seconds of one digit.

We are primarily interested in codes of relatively short length and if, for example, we limit ourselves to block codes of length 4, then the set of 10 codewords listed in Table 3 satisfies the constraint of Fig. 3(a). Since the information rate of this code is $R = \log_2 [(10/4)T^{-1}] = 0.8305T^{-1}$ bits/sec, this code achieves 87.6% of the maximum capacity of the sequence generator. An immediate application of this code is as a replacement for

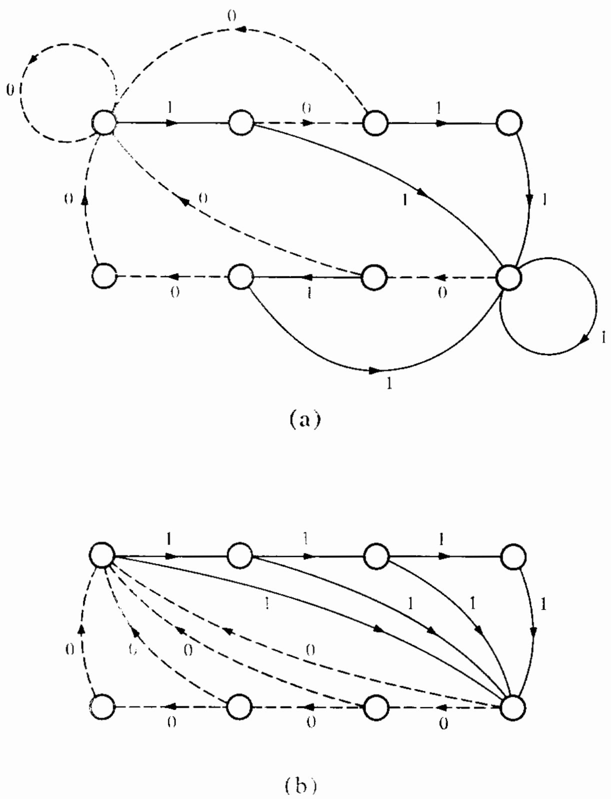


Figure 3 Graphic representation of constraints on binary sequences.

Table 3 The 5211-weighted decimal code of length 4.

Decimal numeral	*Binary code*[a]
0	0 0 0 0
1	0 0 0 1
2	0 0 1 1
3	0 1 1 0
4	0 1 1 1
5	1 0 0 0
6	1 0 0 1
7	1 1 0 0
8	1 1 1 0
9	1 1 1 1

[a] The 1's and 0's correspond to the $+1$'s and -1's, respectively, of the analysis.

Table 4 A selected alphanumeric code,[a] that is, 36 codewords of length 6 that satisfy the binary-sequence constraint shown in Fig. 3(b).

0 0 0 0 1 1	0 1 0 0 1 0	1 0 0 0 0 1	1 1 0 0 0 0
0 0 0 1 1 0	0 1 0 0 1 1	1 0 0 0 1 1	1 1 0 0 0 1
0 0 0 1 1 1	0 1 0 1 1 0	1 0 0 1 0 0	1 1 0 0 1 0
0 0 1 0 0 1	0 1 1 0 0 0	1 0 0 1 0 1	1 1 0 0 1 1
0 0 1 0 1 1	0 1 1 0 0 1	1 0 0 1 1 0	1 1 0 1 0 0
0 0 1 1 0 0	0 1 1 0 1 0	1 0 0 1 1 1	1 1 0 1 1 0
0 0 1 1 0 1	0 1 1 0 1 1	1 0 1 0 0 1	1 1 1 0 0 0
0 0 1 1 1 0	0 1 1 1 0 0	1 0 1 1 0 0	1 1 1 0 0 1
0 0 1 1 1 1	0 1 1 1 1 0	1 0 1 1 0 1	1 1 1 1 0 0

[a] The 1's and 0's correspond to the $+1$'s and -1's, respectively, of the analysis.

the natural binary-coded decimal (NBCD) code. The code digits in Table 3, incidentally, can be assigned weights of 5, 2, 1 and 1 from left to right, respectively. Comparison of the performance of this 5211-weighted decimal code with the performance of the NBCD code and Gorog's decimal code is made in the section on simulation results.

If the transfer function of the channel has imperfections in both the high and low frequency ends, sequences containing large runs of 1's or 0's and sequences containing 1 and 0 alternatively over many consecutive digits yield a large amount of intersymbol interference. However, the interference functions $h_1(t)$ and $h_2(t)$ cancel each other to some extent at the sampling points $t = (2n + 1)T$ while adding together at $t = 2nT$, where n is an integer; therefore, the most undesirable sequences are those that contain long runs of 1 (or 0) at every second digit. For example, the sequence $\cdots, 1, x_1, 1, x_2, 1, x_3, 1, x_4, 1, x_5, \cdots$ is undesirable regardless of the values of x_i. Codes that avoid this situation may be found as follows: Divide a sequence into two subsequences of odd-number digits and even-number digits, respectively; each subsequence is required to satisfy the constraint illustrated in Fig. 3(b), i.e., the run lengths of 1's or 0's in the subsequences may not be more than 4. The capacity of the sequence generator under this constraint is the same as that of Fig. 3(a), i.e., $C = 0.9469T^{-1}$ bits/sec.

If we restrict ourselves to a class of block codes of length 6, the constraint is equivalent to elimination of codewords such as $\{a_1, a_2, a_3, a_4, a_5, a_6\}$ in which $a_1 = a_3 = a_5$ or $a_2 = a_4 = a_6$. There are 36 codewords listed in Table 4. This class of codes of length 6 has the following properties: 1) The run lengths of 1's and 0's in any sequence of codewords are less than 8; 2) the alternating sequences of 1's and 0's in any sequence of the codewords have lengths less than 8; and 3) if $a_i = \pm 1$ in a codeword of length 6, where $+1$ and -1 correspond to 1 and 0 respectively in Table 4, then

$$\sum_{i=1}^{6} a_i = -2,\ 0 \quad \text{or} \quad +2 \quad \text{and}$$

$$\sum_{i=1}^{6} (-1)^i a_i = -2,\ 0 \quad \text{or} \quad +2.$$

Note that this property is similar to that of the frequency concept codes [see Eqs. (17) and (18)]. The information rate of this code is $R = \log_2 [(36/6)T^{-1}] = 0.8616T^{-1}$ bits/sec; its efficiency is 91.0%.

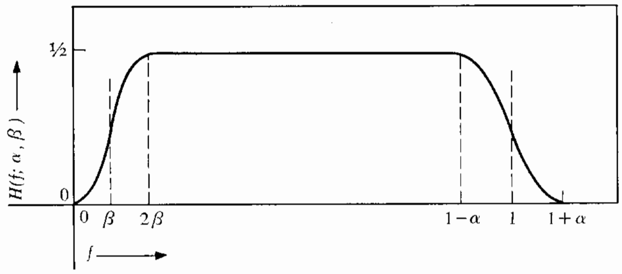


Figure 4 Frequency characteristic of the baseband channel.

Simulation results

In this section we present the results of simulating the performance of the coding schemes discussed in the previous section and make a quantitative comparison of these schemes from the viewpoint of the reduction in intersymbol interference. It is impractical to study all channels; the channels which we treat are assumed to be band-limited with sinusoidal roll-off (see Fig. 4); i.e.,

$$H(f;\alpha,\beta)=\begin{cases}\dfrac{1}{2}\left[1+\sin\dfrac{\pi(f-\beta)}{2\beta}\right], & 0<|f|<2\beta;\\ \frac{1}{2}, & 2\beta\le|f|<1-\alpha;\\ \dfrac{1}{2}\left[1-\sin\dfrac{\pi(f-1)}{2\alpha}\right], & 1-\alpha\le|f|<1+\alpha;\\ 0, & 1+\alpha\le|f|.\end{cases} \tag{23}$$

As can be seen from Eq. (23), we assume that the phase characteristic of the channel is completely linear. If a phase nonlinearity exists, $H_1(f)$ of Eq. (5) and $G(f)$ of Eq. (8) are not real functions and therefore $h_1(t)$ and $g(t)$ are not even functions; hence the intersymbol interference from the preceding digits and that from the following digits are not always equal. However, the general property still holds that $h_1(t)$ and $g(t)$ are slowly changing functions when $H_1(f)$ and $H_2(f)$ have narrow bandwidths and the results in this section are applicable to the general case. Thus, by carrying out extensive simulation experiments for the limited class of channels given by Eq. (23), our objective was achieved efficiently.

The impulse response function of the channel (23) is

$$h(t;\alpha,\beta)=\frac{\sin 2\pi t}{2\pi t}\,\frac{\cos 2\pi\alpha t}{1-(4\alpha t)^2}-\frac{\sin 2\pi\beta t}{2\pi t}\,\frac{\cos 2\pi\beta t}{1-(4\beta t)^2}. \tag{24}$$

If $\beta=0$, $H(f;\alpha,0)$ is the transfer function of a low-pass channel with sinusoidal roll-off and

$$h(\tfrac{1}{2}n;\alpha,0)=\delta_{n,0}, \tag{25}$$

where $n=0,\pm1,\cdots$ and $0\le\alpha\le1$. Therefore we can send an impulse sequence over the channel $H(f;\alpha,0)$ with no intersymbol interference at the rate $R=T^{-1}=2$ bits/sec. [In Fig. 4 the channel bandwidth is normalized to $(1+\alpha)$. If the channel characteristic is scaled by a factor of W on the frequency axis, i.e., if the channel bandwidth is $W(1+\alpha)$, the distortion-free transmission rate is $R=2W$ bits/sec.] However, when the transmission rate is not 2 bits/sec, intersymbol interference occurs. An appropriate quantitative measure of intersymbol interference is the eye opening in the usual eye pattern (Ref. 1, p. 61) and, in particular, if the timing of the sampling at the receiver is assumed to be perfect, the vertical eye opening represents the degradation caused by the intersymbol interference.

Let us consider a binary input sequence in a finite time interval $-T_{\max}\le t\le T_{\max}$, where the value of $T_{\max}$ is chosen so that the impulse response virtually dies out beyond $|t|=T_{\max}$. In the following calculation $T_{\max}=10$; it is sufficient to consider the finite-length input sequences $a_{-N}, a_{-N+1},\cdots,a_0,\cdots,a_{N-1},a_N$, where N is the integer such that $T_{\max}R\le N<T_{\max}R+1$, and R is the digit rate of the binary sequence. The input signal to the signal generator (see Fig. 1) is

$$x(t)=\sum_{n=-N}^{N}a_n\,\delta(t-nR^{-1}). \tag{26}$$

The signal received at the sampler in the absence of noise is

$$y(t)=\sum_{n=-N}^{N}a_nh(t-nR^{-1};\alpha,\beta). \tag{27}$$

Then for a binary input, the vertical eye opening at time $t=0$ is determined by the following two quantities:

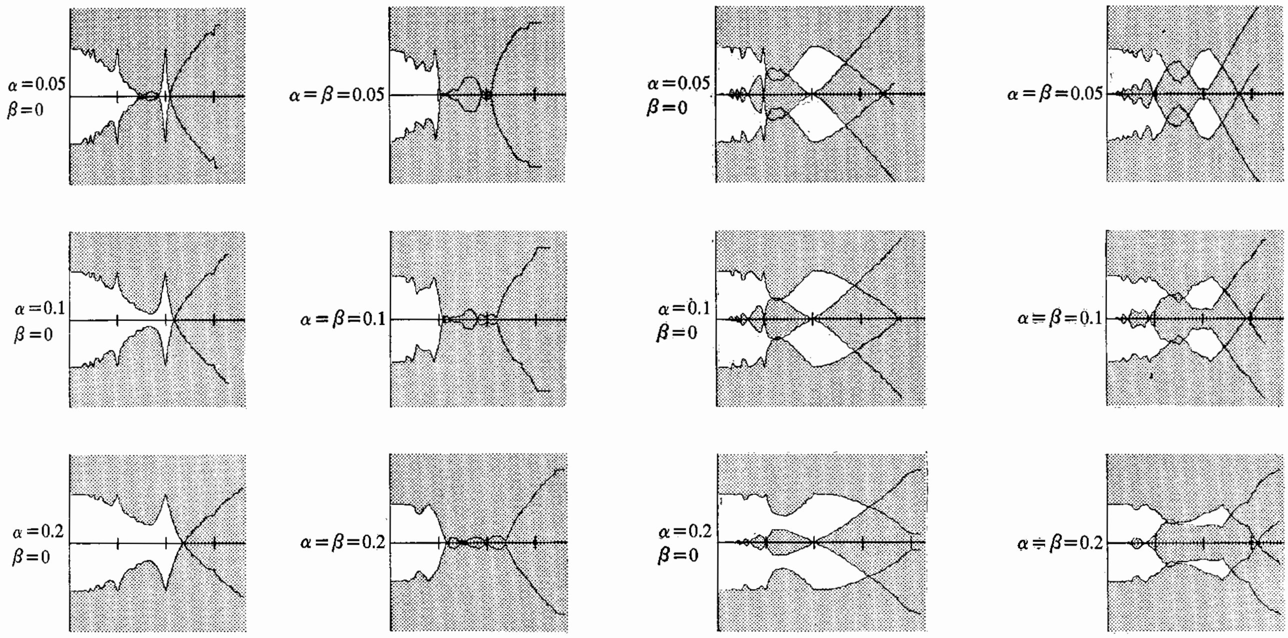
$$\begin{aligned}e^+(R)&=\min_{\{a_n;a_0=1\}}y(0)\\&=\min_{\{a_n;a_0=1\}}\sum_{n=-N}^{N}a_nh(-nR^{-1};\alpha,\beta)\\&=h(0;\alpha,\beta)-{\sum_{n=-N}^{N}}'|h(-nR^{-1};\alpha,\beta)|\end{aligned} \tag{28}$$

and

$$e^-(R)=\max_{\{a_n;a_0=-1\}}y(0)=-e^+(R), \tag{29}$$

where the prime on the summation sign indicates deletion of the term for $n=0$.

The curves in Fig. 5(a) indicate how the vertical eye opening varies as a function of the transmission rate R. The shaded areas indicate the regions in which $e^+(R)<e^-(R)$ (i.e., the eye is closed). Note that when $\beta=0$ the eye is completely open for rates $R=2/n$, $n=1,2,\cdots$.



(a) (b) (c)

Figure 5 Vertical eye opening as a function of transmission rate for binary input data: (a) without correlative level coding; (b) with correlative level coding, type I; (c) with correlative level coding, type II. The abscissa scales begin at 0 and the scale marks indicate $R = 1$, 2 and 3 bits/sec.

When $\beta \neq 0$, i.e., when the transmission channel excludes the zero frequency component, the eye is completely closed unless the transmission rate is very slow. It can be shown[7] that the transmission rate must be less than $(1 + \alpha)$ bits/sec for the eye to be open.

The curves of Figs. 5(b) and 5(c) correspond to the cases in which correlative level coding, type I [Eq. (13)] and type II [Eq. (15)] respectively, is used. There are two eyes because of the three-level signals. The vertical eye opening of the upper eye at the sampling instant is determined[7] by the quantities $e_1(R)$ and $e_2(R)$, and that of the lower eye by $e_3(R)$ and $e_4(R)$:

Type I coding

$$e_1(R) = h(0; \alpha, \beta) + h(R^{-1}; \alpha, \beta) - \sum_{n=1}^{\infty} |h(nR^{-1}; \alpha, \beta) + h[(n+1)R^{-1}; \alpha, \beta]|; \tag{30}$$

$$e_2(R) = \sum_{n=1}^{\infty} |h(nR^{-1}; \alpha, \beta) + h[(n+1)R^{-1}; \alpha, \beta]|; \tag{31}$$

$$e_3(R) = -e_2(R); \tag{32}$$

$$e_4(R) = -e_1(R); \tag{33}$$

Type II coding

$$e_1(R) = h(0; \alpha, \beta) - h(2R^{-1}; \alpha, \beta) - \sum_{n=1}^{\infty} |h(nR^{-1}; \alpha, \beta) - h[(n+2)R^{-1}; \alpha, \beta]|; \tag{34}$$

$$e_2(R) = \sum_{n=1}^{\infty} |h(nR^{-1}; \alpha, \beta) - h[(n+2)R^{-1}; \alpha, \beta]|; \tag{35}$$

$$e_3(R) = -e_2(R); \tag{36}$$

$$e_4(R) = -e_1(R). \tag{37}$$

When $e_1(R) < e_2(R)$ [which implies $e_3(R) < e_4(R)$], both the upper and the lower eye are closed; the shaded areas in Figs. 5(b) and 5(c) correspond to this situation. Comparing these curves with Fig. 5(a), we recognize a notable improvement in allowable transmission rate, particularly in the channels with $\beta \neq 0$, where the eye was originally closed for $R \approx 2$ and only allowed transmission at less than half the Nyquist rate. The eye opening is not sensitive to deviation of the transmission rate R from its optimum value; the reason for this is discussed later.

We now show how the binary codes in Tables 1 through 4 perform over the low-pass channel ($\beta = 0$). The vertical eye openings are illustrated in Fig. 6 as functions

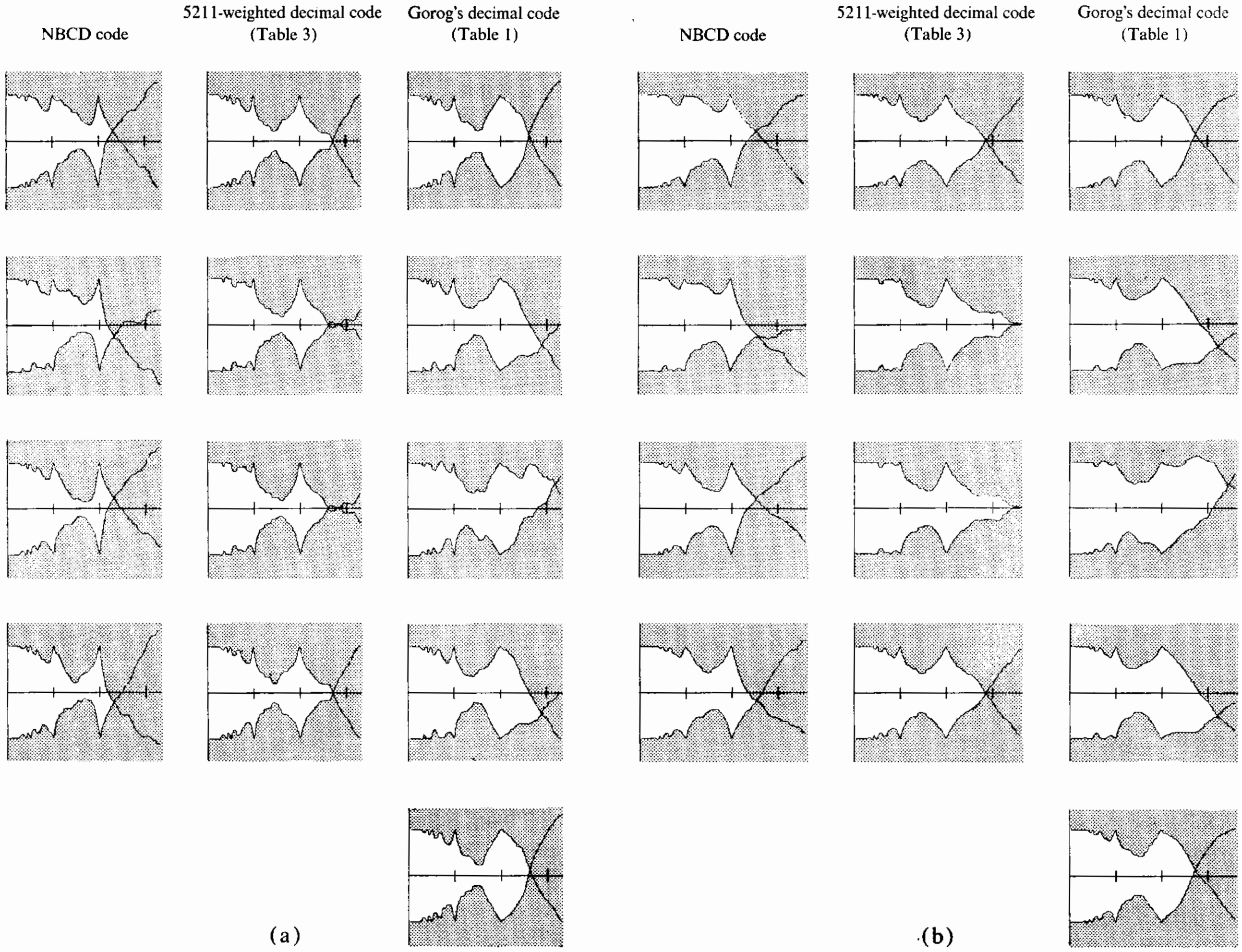


Figure 6 Vertical eye opening as a function of transmission rate over the low-pass channel ($\beta = 0$) for binary codes of lengths 4 and 5; (a) $\alpha = 0.1$; (b) $\alpha = 0.2$. The abscissa scales begin at 0 and the scale marks indicate $R = 1$, 2 and 3 digits/sec.

of the transmission rate R for the NBCD code, the 5211-weighted code (Table 3) and Gorog's decimal code (Table 1) as transmitted over channels with sinusoidal roll-off parameters $\alpha = 0.1$ [Fig. 6(a)] and $\alpha = 0.2$ [Fig. 6(b)]. Note that the interference pattern for a code of length N is represented by N different eye diagrams, since codeword synchronization is assumed.

The curves for the sharper cut-off ($\alpha = 0.1$) channel in Fig. 6 appear to indicate that the frequency concept code allows a faster transmission rate than the other codes. A careful observation, however, shows that this is not the case. Since we are comparing codes of different lengths, an appropriate measure is the ratio of the codeword transmission rate to the bandwidth, $J \equiv R_{\max}/[N(1 + \alpha)]$, where N is the length of the codeword and $R_{\max}$ is the maximum symbol rate at which the binary code can be transmitted while maintaining all N eye openings larger than a given amount. Figure 7 shows the relation between J and the parameter α with β fixed at 0. For the 5211-weighted decimal code and Gorog's decimal code, the four curves correspond to the cases in which the eye is closed, 25%, 50% and 100% open, respectively.

In Fig. 8 are shown the vertical eye openings for three alphanumeric codes: 1) a conventional binary decimal code of length 6 developed without taking intersymbol interference into account; 2) the selected code of length 6 (Table 4); and 3) Gorog's code of length 8 (Table 2). Comparison of these curves shows clearly that selection of an appropriate code can significantly reduce the intersymbol interference. Gorog's alphanumeric code performs better than the code of length 6 if the two are compared at the same symbol rate R; however, under the performance measure J, the selected alphanumeric code of length 6 is superior when the roll-off parameter of the channel is small. For channels with heavy distortion, e.g., $\alpha = 0.5$ and $\beta > 0.1$, an eye opening even 50% as large as that of the interference-free case cannot be obtained for the code of length 6. Similarly, a 25%-open eye is not possible[7] for the 6 digit code for channels with $\beta \geq 0.15$. Therefore Gorog's alphanumeric code is preferable to the others when the channel is heavily distorted at the low frequency end.

The vertical eye openings of the correlative level coding signal (Fig. 5) and of Gorog's decimal code (Fig. 6) are less sensitive to a change in the transmission rate around $R = 2$ than are the eye openings of the other codes. This fact can be explained as follows. The two functions $e_1(R)$ and $e_2(R)$ that determine the vertical eye opening of the correlative level coding signal are defined by Eqs. (30) and (31) for type I, and by Eqs. (34) and (35) for type II, respectively. Let us assume $\beta = 0$ and use the shorter notation $h(t)$ in place of $h(t; \alpha, 0)$. On taking the derivative of $h(nR^{-1})$ with respect to its argument and setting $R = 2$, we obtain

$$h'(\tfrac{1}{2}n) = \begin{cases} 0, & \text{for } n = 0; \\ \dfrac{(-1)^{n-1} g_n}{2\pi n}, & \text{otherwise}; \end{cases} \tag{38}$$

where

$$g_n \equiv \frac{\cos n\pi\alpha}{1 - (2n\alpha)^2}. \tag{39}$$

It follows that

$$\frac{d}{dR}\left[h\left(\frac{n}{R}\right)\right]\Bigg|_{R=2} = -\tfrac{1}{4}nh'(\tfrac{1}{2}n), \tag{40}$$

and also

$$\frac{d}{dR}\left[h\left(\frac{n}{R}\right) + h\left(\frac{n-1}{R}\right)\right]\Bigg|_{R=2} = \frac{(-1)^n}{8\pi}(g_n - g_{n+1}). \tag{41}$$

Since the sequence $\{g_n\}$ is slowly changing for small values of α, the difference $(g_n - g_{n+1})$ is quite small for all values of n. Similarly, the differences of g's appearing in

$$\frac{d}{dR}\left[h\left(\frac{n}{R}\right) - h\left(\frac{n+2}{R}\right)\right]\Bigg|_{R=2} = \frac{(-1)^n}{8\pi}(g_n - g_{n+2}) \tag{42}$$

have small values. Although the derivatives of $e_1(R)$ and $e_2(R)$ are, in general, discontinuous at $R = 2$, the left and right derivatives of each exist; these derivatives take on only small values, through their dependence on Eqs. (41) and (42).

The two functions $e^+(R)$ and $e^-(R)$, which determine the vertical eye opening of Gorog's decimal code, are defined by

$$e^+(R) = \min_{\{a_n;a_0=1\}} \sum_{n=-\infty}^{\infty} a_n h\left(-\frac{n}{R}\right) \tag{43}$$

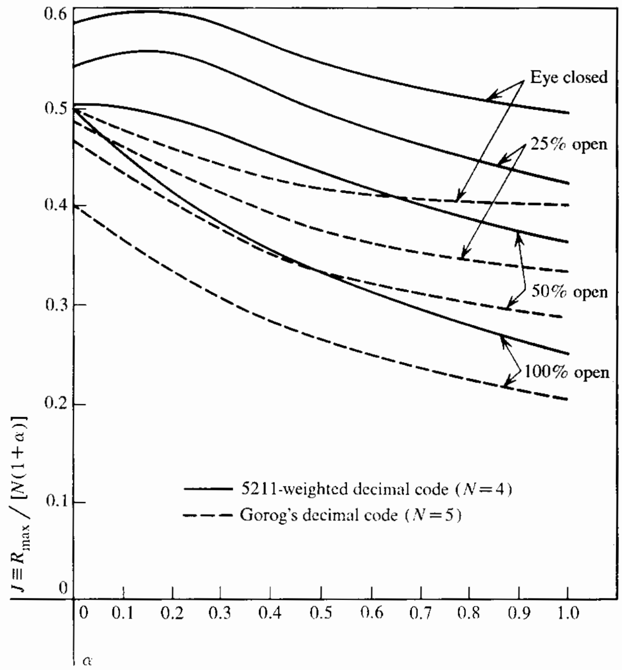


Figure 7 Codeword transmission-rate-to-bandwidth ratio as a function of the channel-shape parameter.

and

$$e^-(R) = \max_{\{a_n;a_0=-1\}} \sum_{n=-\infty}^{\infty} a_n h\left(-\frac{n}{R}\right). \tag{44}$$

[For computer simulation, the limits were given in Eqs. (28) and (29) as $\pm N$, where N is a large but finite integer.] The maximum and the minimum are to be taken under the constraint of the particular code. Thus

$$\frac{de^{\pm}(R)}{dR}\Bigg|_{R=2} \propto \frac{d}{dR}\left[\sum_{n=-\infty}^{\infty} a_n h\left(-\frac{n}{R}\right)\right]\Bigg|_{R=2} = \frac{1}{8\pi}\sum_{n=-\infty}^{\infty}{}' (-1)^n a_n g_n. \tag{45}$$

The expression (45) has a small value when the sequence $\{a_n\}$ satisfies Eq. (18), since the weighting factor g_n is changing slowly compared with the alternating sequence $\{(-1)^n\}$. Thus the left and right derivatives of $e^+(R)$ and $e^-(R)$ at $R = 2$ are small in magnitude, and the vertical eye opening changes little for transmission rates near $R = 2$.

Summary and remarks

The relation between imperfections of a data transmission channel and intersymbol interference of baseband PAM systems was clarified for situations in which the principal distortions (in phase as well as in amplitude) are near the lower and upper frequency ends of the channel. This result was used to explain the effect of correlative level coding in reducing intersymbol interference from a new

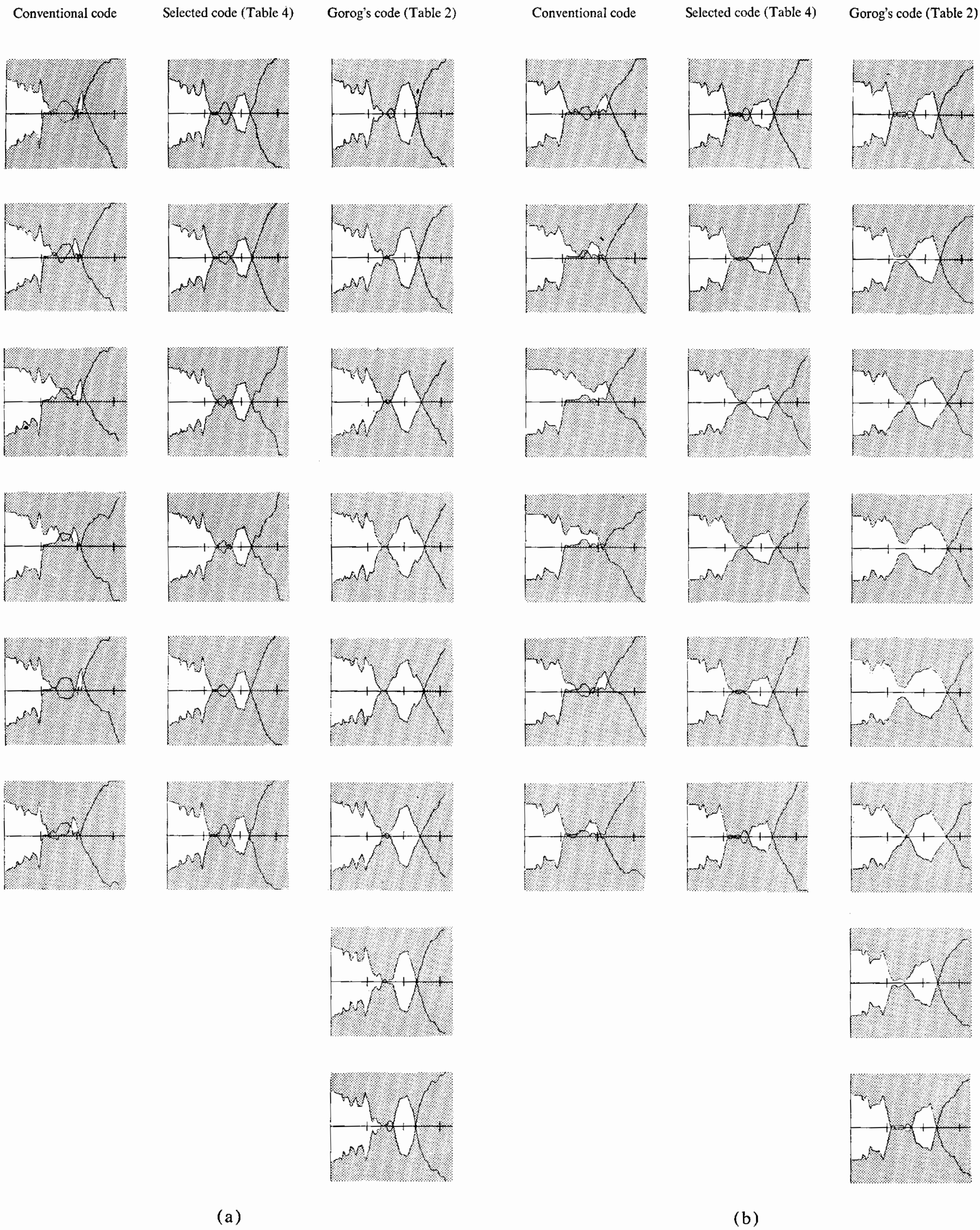


Figure 8 Vertical eye opening as a function of transmission rate for three alphanumeric codes of lengths 6 and 8: (a) $\alpha = 0.05$, $\beta = 0.05$; (b) $\alpha = 0.1$, $\beta = 0.1$. The abscissa scales begin at 0 and the scale marks indicate $R = 1$, 2 and 3 digits/sec.

point of view. Gorog's frequency concept codes were evaluated from a similar viewpoint and some shortcomings associated with the codes were pointed out. Another coding scheme for reduction of intersymbol interference was introduced: Codes that eliminate the worst intersymbol interference patterns were constructed under constraints in the time domain. Specifically, the decimal code of length 4 (5211-weighted code) and the alphanumeric code of length 6 were proposed and their properties discussed.

A quantitative comparison of these coding schemes was made by presenting the simulation results. The formulas for the vertical eye openings were obtained; the vertical eye openings were plotted vs the symbol transmission rate for various channel-shape parameters. It was shown that codes designed in the time domain achieve better performance than both frequency concept codes and conventional codes for a wide class of channel characteristics. The criterion adopted was the codeword transmission-rate-to-bandwidth ratio for a given eye opening.

We were concerned primarily with detection based on digits. However, the frequency concept codes proposed by Gorog have the additional property that the Hamming distance[9] is two and this property should be taken into account to make the best use of these codes. Reference 7 discusses the reception of the sequence using the matched filters for codewords and some simulation results are presented.

The design of codes with a time-domain constraint or with a frequency constraint can be extended to multilevel codes; Franaszek[10] recently developed an algorithm for constructing constrained codes of variable length, as well as of fixed length, and Tang[11] has discussed extensively the run-length-limited codes, which are also a class of codes with constraints.

Acknowledgments

The author is indebted to E. Gorog, whose work on the frequency concept codes stimulated this study. Particular thanks are due to P. D. Welch and C. D. Cullum for the helpful suggestions and criticism they provided. Thanks are also extended to the reviewer for some constructive comments on the previous version of this paper.

References

1. R. W. Lucky, J. Salz and E. J. Weldon, *Principles of Data Communication,* McGraw-Hill Book Co., Inc., New York 1968.
2. W. R. Bennet and J. R. Davey, *Data Transmission,* McGraw-Hill Book Co., Inc., New York 1965.
3. A. Lender, "Correlation Level Coding for Binary Data Transmission," *IEEE Spectrum* **3,** No. 2, 104 (1966).
4. P. J. van Gerwen, "On the Generation and Application of Pseudo-Ternary Codes in Pulse Transmission," *Phillips Res. Repts.* **20,** 469 (1965).
5. S. K. Becker, E. R. Kretzmer and S. H. Sheehan, "A New Signal Format for Efficient Data Transmission," *Bell System Tech. J.* **45,** 755 (1966).
6. E. Gorog, "Redundant Alphabets with Desirable Frequency Spectrum Properties," *IBM J. Res. Develop.* **12,** 234 (1968).
7. H. Kobayashi, "A Simulation Study of the Performance of Correlative Level Coding Schemes and Special Signaling Alphabets," *Research Report RC 2229,* IBM Thomas J. Watson Research Center, Yorktown Heights, New York 1968.
8. C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Tech. J.* **27,** 379 (1948).
9. W. W. Peterson, *Error Correcting Codes,* Massachusetts Institute of Technology Press, Cambridge, Mass. 1961.
10. P. A. Franaszek, "Determination of the Existence of Codes for Synchronous Transmission of Binary Data over Discrete Constrained Channels," *Research Report RC 2468,* IBM Thomas J. Watson Research Center, Yorktown Heights, New York 1968.
11. D. T. Tang, "Run-Length-Limited Codes for Synchronization and Compaction," *Research Report RC 1883,* IBM Thomas J. Watson Research Center, Yorktown Heights, New York 1967.

Received January 9, 1969; revision, August 15, 1969