THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 53, No. 3, March 1974
Printed in U.S.A.

Adaptive Coding for Coherent Detection of Digital Phase Modulation

By C. L. RUTHROFF and W. F. BODTMANN

(Manuscript received October 3, 1972)

Although coherent phase-shift keying (CPSK) is an efficient means of transmitting digital signals over carrier systems, it has not enjoyed widespread use at microwave and millimeter wavelengths because of the difficulty of recovering an accurate reference carrier for coherent detection.

In this paper, a system is described which requires only a narrow-band phase-locked-oscillator filter for reference carrier recovery. This is accomplished by block-coding and decoding the pulse sequence at the terminals; the recovery of a baseband timing wave is also facilitated by the coding process. It is also shown that: (i) for an arbitrary random input sequence, accurate carrier recovery cannot be achieved with just a narrow-band filter, (ii) for the system described, any input pulse sequence is acceptable, and (iii) there is a maximum error in the phase of the recovered reference carrier which can be controlled by choosing the number of pulses in the coding block and the bandwidth of the recovery filter.

I. INTRODUCTION

Coherent phase-shift keying (CPSK) is one of the most efficient means of modulation for the transmission of digital information over carrier systems. In particular, CPSK is at least as efficient as frequency-shift keying or differentially coherent phase-shift keying.[1–3] Equally important from the point of view of hardware realization, CPSK is suited to operation with amplifiers which operate most efficiently in a nonlinear regime; this class of amplifiers includes those using traveling-wave tubes, varactor up-converters, and tunnel or IMPATT diodes used as power amplifiers or as injection-locked oscillator amplifiers.

Traveling-wave-tube amplifiers have been proposed for use in satellite repeaters, and there is considerable current work directed toward the application of millimeter-wave integrated circuit injection-locked oscillator amplifiers in digital radio and waveguide transmission systems.[4–16] The CPSK method described here is suitable for those applications.

For the type of operation envisaged for these systems, the statistics of the digital sources are usually unknown. To achieve maximum operational flexibility, it was assumed at the outset that the system must operate with any input pulse sequence. With this arrangement, the statistics of the signal source need not be restricted.

In the system to be described, the recovery of the reference carrier phase—a major problem in CPSK transmission—is accomplished with the aid of a block-coding and decoding of the pulse sequence at the terminals. This coding allows recovery of the reference carrier with a narrow-band phase-locked-oscillator filter.

Another important problem in digital systems with unrestricted pulse sequences is the recovery of the timing wave for use in the regeneration process. The same coding process which affords reference carrier recovery for all sequences also assures timing wave recovery for all sequences.

In this paper, the block-coding, the reference carrier recovery, and the timing wave recovery are described for binary and multilevel CPSK systems.

II. COHERENT DIGITAL PHASE MODULATION

2.1 The baseband and modulated carrier signal formats

A diagram of the block-coded CPSK carrier system is shown in Fig. 1. Ignoring the block coder for the moment, the input to the radio system is a baseband sequence of discrete amplitudes as illustrated by a binary sequence of ones and zeros in Fig. 2a. The ones are coded as positive pulses and the zeros as negative pulses as shown in Fig. 2b; this sequence is the input to the phase modulator. Although raised-cosine pulses are used for illustration, other pulse shapes can also be used.

An m-level baseband sequence of raised-cosine pulses shown in Figure 2b is written

$$v(t) = V_o \sum_n a_n p(t - nT), \tag{1}$$

where T is the pulse interval, V_o the peak pulse amplitude, $a_n = \pm 1$, and

$$p(t) = \begin{cases} \dfrac{1}{2}\left(1 + \cos \dfrac{2\pi t}{T}\right), & |t| \leqq \dfrac{T}{2} \\[2ex] 0, & |t| > \dfrac{T}{2}. \end{cases}$$

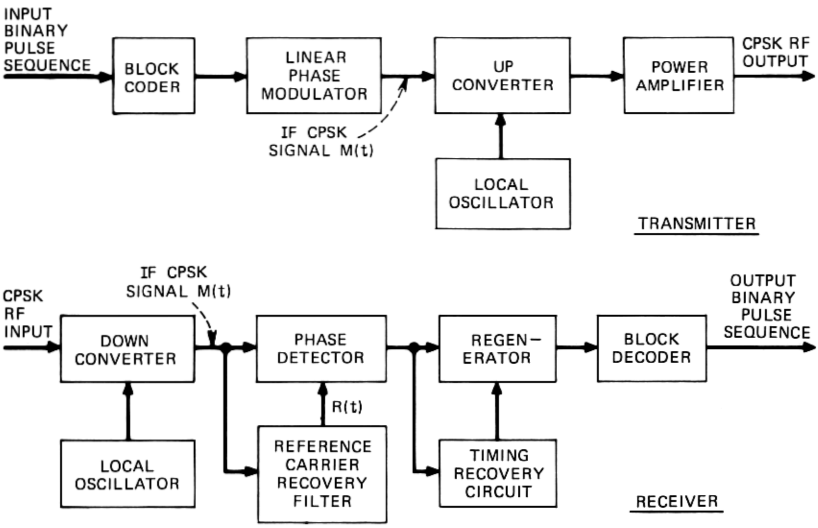

Fig. 1—Block-coded CPSK carrier terminals.

This baseband signal is used to phase modulate a sinusoidal carrier. The output of the phase modulator is

$$M(t) = A_c \cos \left[\omega_c t + \sum_n a_n p(t - nT)\right], \tag{2}$$

where $a_n = k\pi/m$, $k = \pm 1, \pm 3, \cdots, \pm(m - 1)$. The pulse sequence of Fig. 2b represents both the baseband signal voltage of (1) and the phase modulation in (2). The peak baseband voltage V_o produces a peak phase deviation of $\pi/2$ radians for the binary case illustrated.

A vector representation of the modulated signal is shown in Fig. 3a. The carrier amplitude is A_c and the unmodulated phase of the carrier

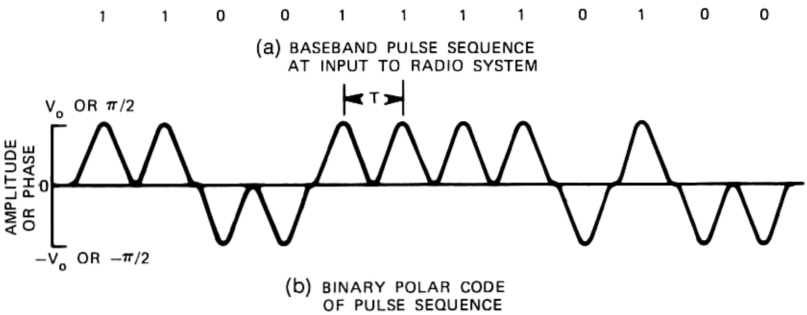

Fig. 2—Binary polar signal format.

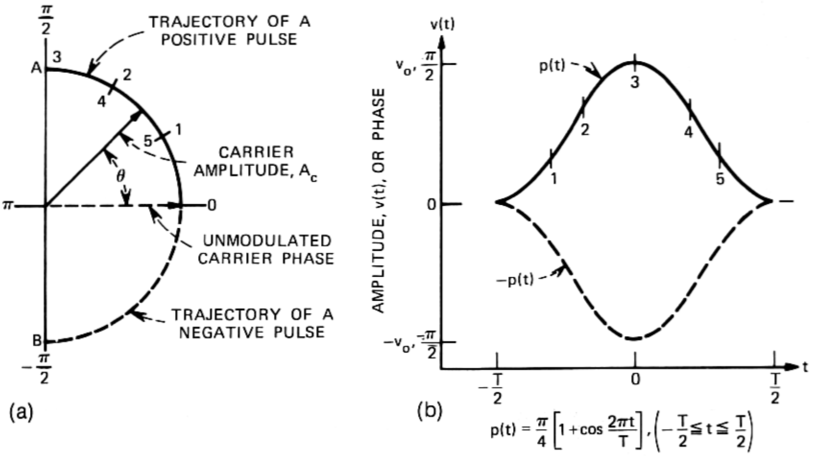

Fig. 3—Phase plane representation of a binary signal.

is zero. When pulses modulate the phase of the carrier the amplitude remains constant and the phase follows the modulating signal voltage. Trajectories for the positive and negative raised-cosine pulses of Fig. 3b are shown in Fig. 3a for the binary case.

It is worth noting that double-sideband suppressed-carrier modulators and switched delay-line modulators are sometimes regarded as phase modulators. A justification for this interpretation is that the pulses are sampled at the receiver only when the phase is at the peak value. However, they differ from phase modulators in the amount of amplitude modulation that is generated. The trajectory of the double-sideband suppressed-carrier modulation is the vertical axis in Fig. 3a between the points A and B; the trajectory of the switched delay-line modulation may be intermediate between the vertical trajectory and the circular trajectory of a phase modulator. Since future systems are expected to have power amplifiers operating in the region of saturation, the distortion caused by large variations in amplitude can be avoided by restricting consideration to phase modulators. Phase modulators suitable for this purpose are described elsewhere.[15]

2.2 A description of coherent phase detection

Let the input to the phase detector of the receiver in Fig. 1 be the phase-modulated signal $M(t)$. The phase detector requires a local reference signal with the proper phase. This reference signal is written

$$R(t) = -2A_R \sin\left[\omega_c t + \epsilon(t)\right], \tag{3}$$

where $\epsilon(t)$ is any error in the reference phase. The output of the phase detector is the low-frequency part, $V_R(t)$, of the product of the input signal and the reference signal.

$$V_R(t) = A_C A_R \sin \left[\sum_n a_n p(t - nT) - \epsilon(t)\right]. \tag{4}$$

If the phase error, $\epsilon(t)$, is zero, and if the output of the phase detector is sampled at times $t = nT$, the output will be $\pm A_C A_R$ accordingly as $a_n = \pm 1$ and the transmitted pulse sequence is recovered.

If the reference phase error is not zero, the signal output amplitude will be reduced by the factor $\cos \epsilon$. For example, if $\epsilon = \pi/4$, the baseband pulse amplitude will be reduced 3 dB. An important function of the system to be described is to recover the reference phase in such a manner that $\epsilon(t)$ is small.

III. REFERENCE CARRIER RECOVERY WITH A PHASE-LOCKED OSCILLATOR

The reference carrier recovery filter is assumed to be a phase-locked oscillator with a locking bandwidth much smaller than the bandwidth of the modulating pulse sequence. The analysis presented here applies to an injection-locked oscillator or a first-order phase-locked loop; the noiseless case will be considered.*

Let the input signal be

$$M(t) = A_c \cos \left[\omega_c t + \theta(t)\right]. \tag{5}$$

The differential equation describing the locking behavior of a negative resistance sine-wave oscillator has been derived in several forms.[17–19] With the present notation the equation is

$$\frac{d\epsilon(t)}{dt} = (\omega_o - \omega_c) - \Delta \sin \left[\epsilon(t) - \theta(t)\right], \tag{6}$$

where ω_o is the unlocked oscillator frequency, $|\omega_o - \omega_c| \ll \omega_c$, 2Δ is the locking bandwidth, and $\epsilon(t)$ is the reference phase error.

Since the oscillator is being used to recover the reference phase, the phase error, $\epsilon(t)$, should be as small as possible. For this reason it is necessary that the locking bandwidth of the locked oscillator be much smaller than the bandwidth of the signal, $\theta(t)$. This assumption, in its most useful form, means that $\Delta T \ll 1$. Following Adler, expression (6) is rearranged as follows:

$$\frac{d\epsilon(t)}{dt} = \Delta K - \Delta \sin \left[\epsilon(t) - \theta(t)\right], \tag{7}$$

* Eisenberg has presented a related analysis which includes additive thermal noise.[16]

where $K = (\omega_o - \omega_c)/\Delta$. The term K represents any initial difference between the free-running frequency of the oscillator and the input frequency; in the region of interest $|K| < 1$.

We are interested in deriving an unambiguous reference carrier for multilevel digital modulation in which each pulse is time-limited to a single interval of duration T.

$$\theta(t) = \sum_n a_n p(t - nT). \tag{8}$$

During a single pulse the variation in $\theta(t)$ is much larger than the variation in $\epsilon(t)$ because $\Delta T \ll 1$. Therefore, the phase error at the end of the nth pulse can be found by integrating (7) over the nth pulse with $\epsilon(t)$ held constant at the value of the phase error at the beginning of the nth pulse. Writing $\epsilon_n = \epsilon(nT)$, we have, from (7) and (8),

$$\begin{aligned}\epsilon_{n+1} - \epsilon_n &= \Delta \int_{nT}^{(n+1)T} \{K - \sin[\epsilon_n - \theta(t)]\} dt \\ &= \Delta T \left[K - \frac{\sin \epsilon_n}{T} \int_{nT}^{(n+1)T} \cos \theta(t) dt \right. \\ &\qquad \left. + \frac{\cos \epsilon_n}{T} \int_{nT}^{(n+1)T} \sin \theta(t) dt \right]. \end{aligned} \tag{9}$$

Each pulse is nonzero in a single interval of duration T so we have

$$\int_{nT}^{(n+1)T} \cos \theta(t) dt = \int_{nT}^{(n+1)T} \cos a_n p(t - nT) dt = \int_0^T \cos |a_n| p(x) dx,$$

and

$$\begin{aligned}\int_{nT}^{(n+1)T} \sin \theta(t) dt &= \int_{nT}^{(n+1)T} \sin a_n p(t - nT) dt \\ &= \frac{a_n}{|a_n|} \int_0^T \sin |a_n| p(x) dx.\end{aligned}$$

Simplifying the notation, we write

$$b_n \equiv a_n/|a_n|, \qquad C_n \equiv \frac{1}{T} \int_0^T \cos |a_n| p(x) dx,$$

and

$$S_n \equiv \frac{1}{T} \int_0^T \sin |a_n| p(x) dx. \tag{10}$$

Expression (9) becomes

$$\epsilon_{n+1} - \epsilon_n = \Delta T[K - C_n \sin \epsilon_n + b_n S_n \cos \epsilon_n]. \tag{11}$$

As shown in (10), C_n and S_n are functions of the shape of the pulse. For the digital signal described by (8) the peak deviation is less than π radians for any number of levels and, for the class of pulse shapes of interest, S_n is positive. This is not true of C_n—it can be positive, negative, or zero. Eisenberg[16] has derived C and S for several pulse shapes of interest. The sign of C_n has an important effect upon the phase of the recovered reference carrier; in order that the recovered reference carrier have an unambiguous phase near zero degrees, it will be shown that a pulse shape must be used for which $C_n > 0$.

For the binary case, (11) can be written

$$\frac{\epsilon_{n+1} - \epsilon_n}{\Delta T} = K - \sqrt{C^2 + S^2} \sin \left(\epsilon_n - b_n \tan^{-1} \frac{S}{C} \right). \tag{12}$$

Let the probability that $b_n = +1$ be p and the probability that $b_n = -1$ be $(1 - p)$. The average phase, ϵ_o, will be such that the error due to p positive pulses is equal in amplitude and opposite in sign from the error due to $(1 - p)$ negative pulses. From (12) we get

$$p \left[K - \sqrt{C^2 + S^2} \sin \left(\epsilon_o - \tan^{-1} \frac{S}{C} \right) \right] = -(1 - p) \left[K - \sqrt{C^2 + S^2} \sin \left(\epsilon_o + \tan^{-1} \frac{S}{C} \right) \right],$$

and solving for the average phase, we get

$$\epsilon_o = \sin^{-1} \frac{K}{\sqrt{C^2 + (2p - 1)^2 S^2}} + \tan^{-1} (2p - 1) \frac{S}{C}. \tag{13}$$

Under the best circuit adjustment, $K = 0$ and the average phase error is given by the second term in (13). When C is positive, the average phase is in the first or fourth quadrant and when $p = \frac{1}{2}$ the average phase is zero. On the other hand, when C is negative, the average phase is in the second or third quadrant and when $p = \frac{1}{2}$ the average phase is π.

A switch of the reference carrier phase from near zero to near π can happen in a multilevel system. Consider a 4-level system with rectangular pulses and peak phase deviations of $\pm\pi/4$ and $\pm 3\pi/4$. Suppose that for a time the pulses alternate between $\pm\pi/4$. From (10), $C = \cos \pi/4 = 1/\sqrt{2} > 0$ and the average phase is zero. If the pulse sequence then changes to alternate between $\pm 3\pi/4$, $C = \cos 3\pi/4 = -1/\sqrt{2} < 0$ and the average phase becomes π. This very undesirable situation can be avoided by using pulse shapes for which $C > 0$. In the rest of this paper we assume that, in all cases, a pulse shape is chosen for which $C_n > 0$.

It is highly desirable that the average phase of the reference carrier be near zero. This means that in addition to requiring that $C_n > 0$, it is also necessary that $K \approx 0$ and $p \approx \frac{1}{2}$ as may be seen from (13). The parameter K can be kept near zero by setting the rest frequency of the phase-locked oscillator equal to the signal carrier frequency. The system is required to operate with any input sequence so it is unlikely that p will always be near one-half. The input sequence can be coded—by a block coder to be described in Section V—into a transmitted sequence with $p = \frac{1}{2}$, thus insuring $\tan^{-1}(2p - 1)S/C \approx 0$. Even with coding, however, the reference phase will fluctuate about zero and it is necesssary to insure that these fluctuations do not cause substantial degradation in performance relative to the performance which would be obtained with a perfect reference carrier. It has often been thought that the problem in the recovery of reference phase is that the occurrence of long sequences of identical pulses drives the recovered phase beyond reasonable limits and that if the sequences of pulses were sufficiently random this problem would go away. Random sequences are therefore of great interest. In the next section the variance of the reference phase error is derived and the results illustrated by an example.

IV. REFERENCE CARRIER PHASE ERROR FOR RANDOM SEQUENCES

The differential equation which describes the phase-locked oscillator is nonlinear and therefore difficult to solve. We begin by noting that (7) can be solved exactly on a pulse-by-pulse basis if the pulses are rectangular. For pulses with other shapes an equivalent rectangular pulse can be derived. Then, linearizing the equation, and recognizing that the phase error is approximately normally distributed, the variance can be estimated for the equivalent rectangular pulses. The binary case is considered.

Let the pulses be rectangular with peak deviation $\pm\theta_n$. Then, rearranging (7) and setting $\theta(t) = \theta_n$, we have

$$\int_{nT}^{(n+1)T} \frac{d\epsilon(t)}{-K + \sin[\epsilon(t) - \theta_n]} = \int_{nT}^{(n+1)T} -\Delta dt.$$

The solution to the integral on the left can be found in many tables of integrals. After some algebra the result can be written

$$\epsilon_{n+1} = \theta_n + 2\tan^{-1} \times \left[\frac{R \tan \frac{\epsilon_n - \theta_n}{2} + \left(K - \tan \frac{\epsilon_n - \theta_n}{2}\right) \tanh\left(\frac{\Delta T}{2} R\right)}{R + \left(1 - K \tan \frac{\epsilon_n - \theta_n}{2}\right) \tanh\left(\frac{\Delta T}{2} R\right)} \right], \tag{14}$$

where $\epsilon_n = \epsilon[(n+1)T]$ and $R = \sqrt{1-K^2}$. Equation (14) is exact for rectangular pulses and, if an input sequence of rectangular pulses is specified, the exact phase error can be computed. We will estimate the variance for a linearized version of (14). When $K = 0$, (14) can be written

$$\epsilon_{n+1} = \theta_n + 2\tan^{-1}\left(e^{-\Delta T}\tan\frac{\epsilon_n - \theta_n}{2}\right). \tag{15}$$

Approximating the tangent by its argument,

$$\epsilon_{n+1} \approx \theta_n(1 - e^{-\Delta T}) + \epsilon_n e^{-\Delta T}. \tag{16}$$

Applying (16) repeatedly we get

$$\epsilon_{n+k} \approx \epsilon_n e^{-k\Delta T} + (1 - e^{-\Delta T})\sum_{m=0}^{k-1}\theta_{n+m}e^{-(k-1-m)\Delta T}.$$

When k is large, the phase error is independent of n and becomes

$$\epsilon_k \approx (1 - e^{-\Delta T})\sum_{m=0}^{k-1}\theta_m e^{-m\Delta T}, \tag{17}$$

where the pulses have been rearranged to simplify the notation.

From (12) it may be seen that the error due to a shaped pulse when $\epsilon_n \approx 0$ is given by

$$\epsilon \approx \Delta T\sqrt{C^2 + S^2}\sin\left(\tan^{-1}\frac{S}{C}\right), \tag{18}$$

where a pulse of positive polarity is assumed. Comparison of (18) with the error in (17) due to the pulse θ_o suggests that the equivalent rectangular pulse is obtained by letting

$$\theta_m = b_m\sqrt{C^2 + S^2}\sin\left(\tan^{-1}\frac{S}{C}\right) \tag{19}$$

in (17) where $b_m = +1$ with probability p, $b_m = -1$ with probability $(1-p)$.

The variance of (17) can be found by straightforward means;[20] the mean, from (13), and the variance for shaped pulses are:

$$\begin{aligned}\mu &\approx \tan^{-1}(2p-1)\frac{S}{C}\\ \sigma^2 &\approx 2\left[\sqrt{C^2+S^2}\sin\left(\tan^{-1}\frac{S}{C}\right)\right]^2\Delta T p(1-p).\end{aligned} \tag{20}$$

The reference error is approximately normally distributed and the probability that the reference phase error will exceed a specified value

ϵ_s is[21]

$$P(|\epsilon| \geqq \epsilon_s) \approx Q\left(\frac{\epsilon_s - m}{\sigma}\right) + Q\left(\frac{\epsilon_s + m}{\sigma}\right). \tag{21}$$

These results will be illustrated by an example. Let the pulse rate be 100 megabits per second and the locking bandwidth 0.5 MHz. For raised-cosine pulses with a peak deviation $\pm\pi/2$,

$$\mu \approx \tan^{-1}(2p - 1)$$
$$\sigma^2 \approx 2\Delta T S^2 p(1 - p)$$

with $S = 0.6021947$. For $p = \frac{1}{2}$, $\mu = 0$ and $\sigma \approx 0.0213$ radian rms. The fraction of time that the phase error exceeds 0.1 radian is

$$P(|\epsilon| \geqq 0.1) \approx 2Q(4.697) = 2.75 \times 10^{-6}.$$

In some applications, changes in pulse pattern density may occur which are reflected in fluctuations in p. In this event the probability of a pulse being positive will not be constant at $p = \frac{1}{2}$ but will wander slowly about this value. Suppose, in the foregoing example, p increases by five percent from $p = \frac{1}{2}$ to $p = 0.525$. Then, $\mu = 0.05$ and

$$P(|\epsilon| \geqq 0.1) \approx Q(0.235) + Q(7.04) \approx 0.41.$$

Two important conclusions are illustrated by this example.

(*i*) For practical circuit parameters, the probability of exceeding reasonable phase errors is uncomfortably large even for well-behaved random input sequences.

(*ii*) Small variations in pulse pattern density can cause large increases in the probability of exceeding reasonable phase errors.

It should be understood that the above example is but one of many possible examples; however, the filter bandwidth assumed is a practical value for systems operating at millimeter wavelengths. It should also be noted that the pulse sequences assumed in the example are highly idealized and may or may not approximate sequences from real sources.

V. THE BINARY BLOCK CODER

The coder described in this section is a digital adaptation of a coder invented by F. K. Bowers.[22]

The operation of the block coder will be described with the aid of Fig. 4. The block counter is an up/down counter which counts each

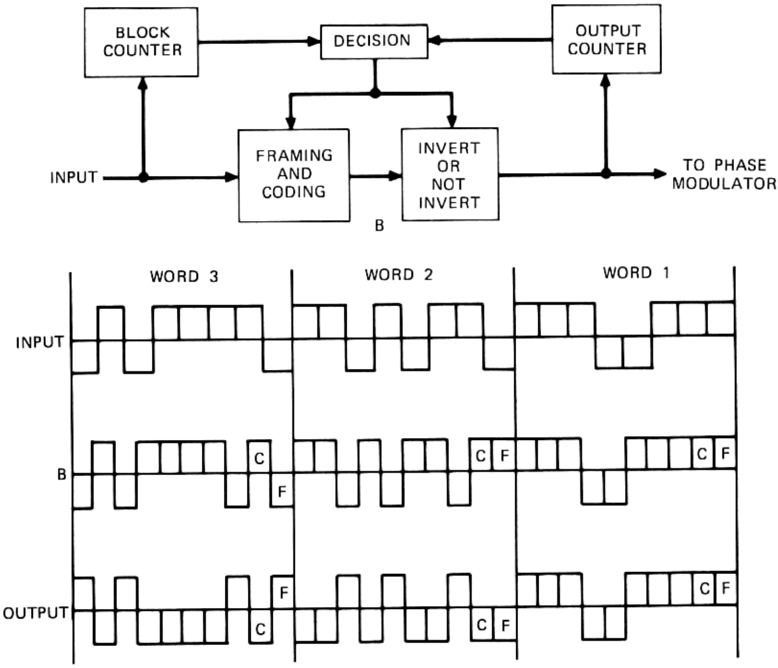

Fig. 4—Binary coder.

successive block of M pulses in the input sequence and indicates on its output terminal whether that block contains more positive than negative pulses. M is an even integer. The output counter is also an up/down counter which counts all pulses transmitted and indicates whether a surplus of positive or negative pulses has been transmitted since the start of transmission. The outputs of the two counters are used in the decision circuit to invert or not invert the block of M pulses just counted, the decision always being made to equalize the number of positive and negative pulses transmitted.

In addition to the framing pulses, a coding pulse is added to each block of M pulses and is used in the receiver to re-invert those blocks which were inverted at the transmitter. Figure 4 shows the operations of a block coder on a sequence of binary input pulses. In this configuration a framing pulse and a coding pulse have been added to each block of M input pulses. While a coding pulse is necessary for each block of M input pulses, fewer framing pulses can be used if desired.

The decoding process at the receiving terminal is illustrated in Fig. 5. The position of the coding pulses in the input sequence are known

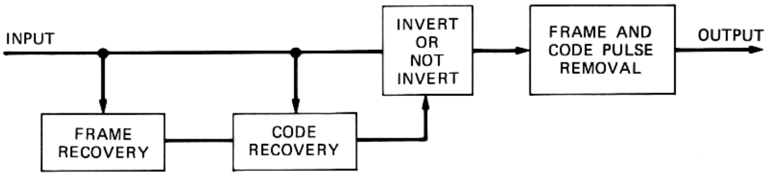

Fig. 5—Binary decoder.

relative to the position of the framing pulse. When framing is established, the coding pulses can be detected and the proper block inversions made so that the output sequence will be identical to the input sequence at the transmitting terminal.

A detailed analysis of the coding operation reveals the following results for M even.

(*i*) The output count, and hence the sum of the output sequence, cannot exceed $\pm(1 + 3M/2)$.

(*ii*) At the end of each frame of $M + 1$ output pulses the output count cannot exceed $\pm(M + 1)$; whatever the count at the end of a frame, the count at the end of the next frame will have moved in the direction of zero by a count of at least one.

(*iii*) The maximum number of pulses between a zero in the output counter and the next zero is $(M + 1)(M + 2)$.

(*iv*) The maximum number of identical pulses is $2 + 5M/2$ and the output count at the end of such a sequence is $\pm(1 + 3M/2)$.

In deriving these properties it is necessary to adopt a convention as to the output indicated by the output counter when the count is zero. If the count approached zero from the negative side it will indicate that a surplus of negative pulses has been sent and the converse is true if the zero count is approached from the positive side. Suppose the output counter indicates that a surplus of positive pulses has been transmitted. The coding pulse is counted as a positive pulse at the input counter making an odd number of pulses counted. At the end of the frame the input counter indicates that a surplus of positive or negative pulses is contained in the block. The block is inverted or not so that the output count goes toward zero. Since there is always a surplus of at least one pulse in each block, the output counter counts toward zero at least one count at the end of every frame; the output count can pass through zero in this process. Now suppose that the output count is zero and that this count was approached from the negative side. The output counter indicates that a surplus of negative pulses has been transmitted. If the next block has all positive pulses, the block will not be inverted and the output count will go to $(M + 1)$. This is the maximum count which can occur at the end of a frame since it has already been shown that at the end of the next frame the count must go toward zero by at least one. Property (*ii*) has therefore been demonstrated.

The example can be continued to demonstrate property (*i*). Let the count be $(M + 1)$ at the end of a frame. At the end of the next frame the count cannot exceed M so that the maximum number of positive pulses that can be added to the count during the frame is $M/2$. Thus, the maximum count is $M + 1 + M/2 = 1 + 3M/2$ and this is property (*i*).

The maximum number of pulses between zeros of the output count is found by achieving the maximum count of $(M + 1)$ in the first block and reducing the count by the minimum of one in successive blocks until zero is reached. There are just $(M + 2)$ blocks necessary to reach the next zero and $(M + 1)$ pulses per block so the maximum number of pulses between zeros is $(M + 1)(M + 2)$. This is property (*iii*). Finally property (*iv*) is achieved by letting the count at the end of a frame be $-(M + 1)$. The next frame has all pulses positive which brings the output counter to zero from the negative direction. The next $(1 + 3M/2)$ pulses can be positive bringing the total number of successive positive pulses to $(M + 1) + (1 + 3M/2) = 2 + 5M/2$ as stated.

When the coder is in operation the transmitted sequence contains equal numbers of positive and negative pulses. The resulting average phase is given by (13) with $p = \frac{1}{2}$.

$$\epsilon_o = \sin^{-1} \frac{K}{C}.$$

The fluctuations about ϵ_o can be determined from (17). The number of pulses between zeros is $(M + 1)(M + 2)$ and if ΔT is sufficiently small that $(M + 1)(M + 2)\Delta T \ll 1$, the exponential terms in (17) are approximately unity. Then, since the maximum sum of the output sequence is $1 + 3M/2$, an upper bound on the phase error results.

$$|\epsilon_{\max}| \leqq \sin^{-1} \frac{|K|}{C} + \Delta T \left(1 + \frac{3M}{2}\right) \theta_m, \tag{22}$$

where θ_m is the equivalent rectangular pulse as given in (19).

A graphic example of the effect of coding is given in Fig. 6. A sequence of 200 random pulses from the Rand table of random numbers

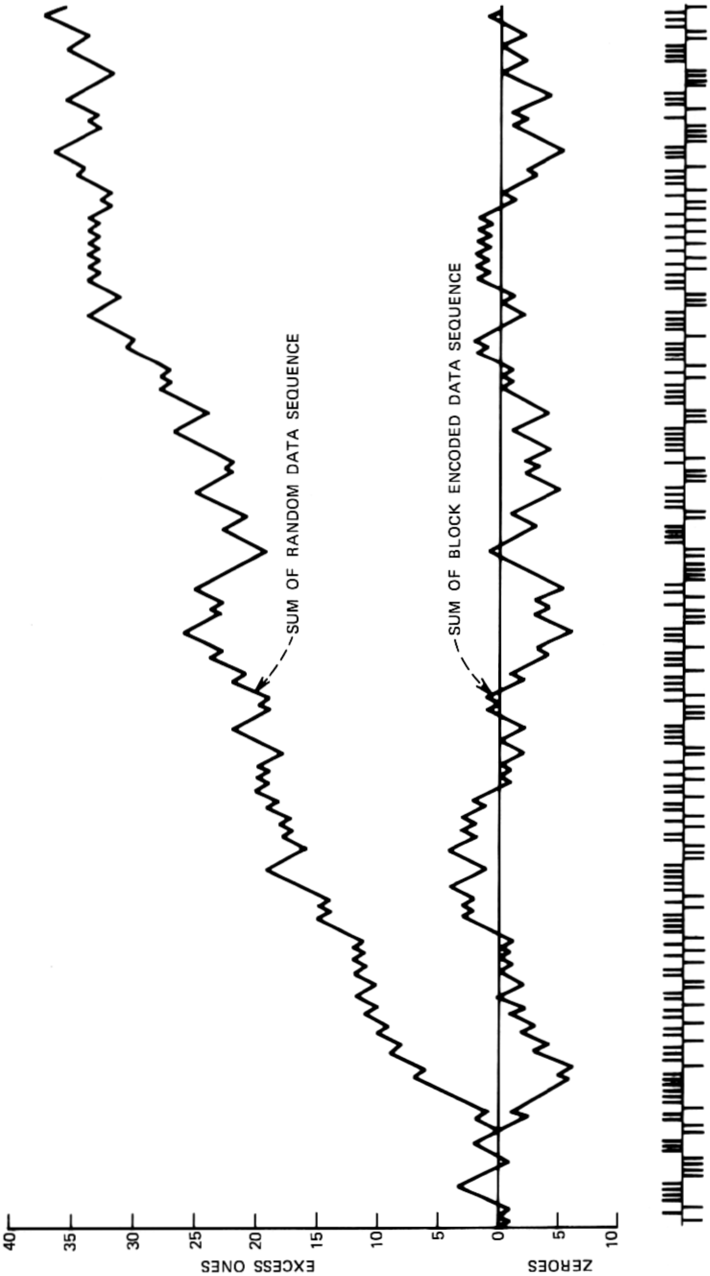

Fig. 6—The effect of coding a random sequence.

is shown at the bottom of the figure.[23] The sum of these digits, $\sum_n b_n$, is the upper plot. Note that although there are many transitions between positive and negative pulses the sum remains above zero most of the time. The slow drift of this sum illustrates the manner in which the phase error wanders.

The same input pulse sequence is shown after coding in a block coder with $M = 8$. The framing and coding pulses are not present. In the uncoded sequence the maximum error for $\Delta T = 0.01$ is 0.36 radian (20.6 degrees) for rectangular pulses with $\pi/2$ radian deviation, whereas the maximum phase error in the coded sequence is 0.06 radian (3.4 degrees).

The original sequence in Fig. 6 is not a rare case. As shown, $\sum_n b_n$ reaches 36 and the probability of this is

$$P_{200}(|\sum_n b_n| \geqq 36) \approx 2Q\left(\frac{36}{\sqrt{200}}\right)$$
$$\approx 2Q(2.54) = 0.011.$$

Thus, about one out of a hundred sequences of 200 pulses each has a sum at least as great as the one shown in Fig. 6.

The price paid for the recovery of the reference carrier with a small phase error is an increase in the transmission rate by the factor $(M + 1)/M$.

VI. TIMING WAVE RECOVERY

It has been shown that, by coding the transmitted pulse sequence, the reference carrier can be recovered accurately. As shown in Fig. 1, the reference carrier is used to drive the phase detector in which the baseband pulse sequence is recovered. In a self-timed system, such as the one depicted in Fig. 1, it is necessary to recover a timing wave for use in regenerating the pulse sequence. Block-coding also helps in this process.

Bennett has shown that a timing wave can be recovered by suitable nonlinear operations even if a spectral line at the timing frequency does not exist; the method requires a suitable number of transitions between signal polarities.[24] But the block coding discussed in Section V insures that the largest number of pulses between signal transitions is $2 + 5M/2$. Therefore, the block coding insures the recovery of both the reference carrier and the timing wave for any sequence of pulses whatever.

For the sequence of pulses shown in Fig. 6 it is instructive to note that, although the sum $\sum_n b_n$ fails to cross the axis for a string of 186 pulses, there are frequent transitions between signal states and for each transition the timing recovery filter will receive a timing pulse.[24] There are 98 transitions in all and the maximum number of identical pulses between transitions is six. This is typical of the behavior of random sequences and is the reason that the recovery of the reference phase is usually more difficult than the recovery of the timing wave.

VII. MULTILEVEL BLOCK-CODED CPSK

The binary coding scheme described in Section V can be extended to 4, 8, 16, and higher numbers of levels. In each case the coder operates to equalize the numbers of pulses with equal amplitudes and conjugate phases. For example, the 4-level coder illustrated in Fig. 7 equalizes the numbers of pulses with $\pi/4$ radian peak deviation and opposite signs, and equalizes the numbers of pulses with $3\pi/4$ radian peak deviation and opposite signs. The 4-level decoder is shown schematically in Fig. 8.

Because the multilevel coder equalizes the numbers of positive and negative pulses for each pair of levels the computation of bounds on the phase error reduces to the binary case. A bound can be computed for each pair of levels and the largest bound applies; this will usually be the bound computed for the pair of levels with the largest deviation.

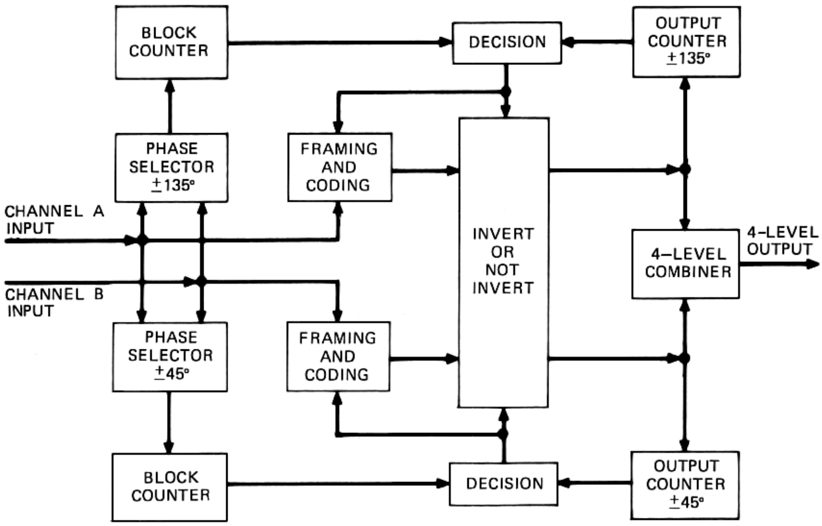

Fig. 7—Four-phase block encoder.

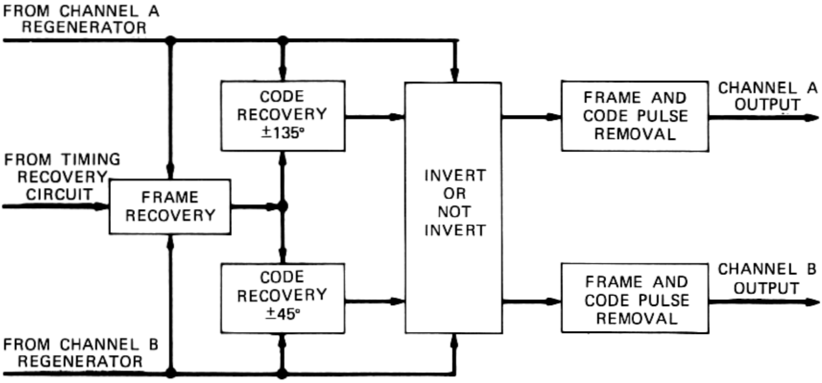

Fig. 8—Four-phase block decoder.

VIII. CONCLUSION

The CPSK system which results from block-coding the input digital sequence as described in this paper has the following properties:

(*i*) The system places no restrictions on the pulse sequence accepted from the source; any sequence whatever can be transmitted.
(*ii*) Recovery of the reference carrier at repeater points is accomplished with a narrow-band filter.
(*iii*) A timing wave can be recovered for any sequence of pulses.
(*iv*) Any pulse shaping required can be done at baseband.
(*v*) The phase-modulated carrier is suited to operation with nonlinear amplifiers; in some applications RF filters are not required to shape the spectrum.

The costs of providing these features are:

(*i*) A block coder must be supplied at the transmitting terminal and a decoder at the receiving terminal.
(*ii*) The transmission rate is increased by the factor $(M + 1)/M$ where M is the number of pulses in the coding block. In principle, M can be very large; in practice, it will be limited by the frequency stabilities of the RF oscillators used in the system.
(*iii*) The error rate is increased by the factor $2(1 - P_e)$ because an error in a coding pulse causes M errors in the signal sequence. This increase in error rate is of little practical importance.

IX. ACKNOWLEDGMENTS

Some of the ideas in this paper grew out of discussions with B. S. Glance; his papers and those of M. V. Schneider, W. W. Snell, and R. F. Trambarulo form the foundation for this work.

We are also indebted to L. J. Greenstein and V. K. Prabhu for valuable critical comments; V. K. Prabhu derived the increase in error rate.

REFERENCES

1. C. R. Cahn, "Performance of Digital Phase-Modulation Communication Systems," IRE Trans. Commun. Syst., *CS-7*, May 1959, pp. 3–6.
2. W. R. Bennett and J. R. Davey, *Data Transmission*, New York: McGraw-Hill, Inc., 1965, pp. 225–229.
3. V. K. Prabhu, "On the Performance of Digital Modulation Systems That Expand Bandwidth," B.S.T.J., *49*, No. 6 (July-August 1970), pp. 1033–1057.
4. M. V. Schneider, B. S. Glance, and W. F. Bodtmann, "Microwave and Millimeter Wave Hybrid Integrated Circuits for Radio Systems," B.S.T.J., *48*, No. 6 (July-August 1969), pp. 1703–1726.
5. B. S. Glance, "Power Spectra of Multilevel Digital Phase-Modulated Signals," B.S.T.J., *50*, No. 9 (November 1971), pp. 2857–2878.
6. B. S. Glance, "Microstrip Impatt Oscillator with High Locking Figure of Merit," Proc. IEEE, *57*, No. 11 (November 1969), pp. 2052–2053.
7. C. L. Ruthroff, "Injection-Locked-Oscillator FM Receiver Analysis," B.S.T.J., *47*, No. 8 (October 1968), pp. 1653–1661.
8. T. L. Osborne and C. H. Elmendorf, "Injection-Locked Avalanche Diode Oscillator FM Receiver," Proc. IEEE, *57*, No. 2 (February 1969), pp. 214–215.
9. B. S. Glance, "Digital Phase-Demodulator," B.S.T.J., *50*, No. 3 (March 1971), pp. 933–949.
10. M. V. Schneider, "Microstrip Lines for Microwave Integrated Circuits," B.S.T.J., *48*, No. 5 (May-June 1969), pp. 1421–1444.
11. M. V. Schneider, "Dielectric Loss in Integrated Microwave Circuits," B.S.T.J., *48*, No. 7 (September 1969), pp. 2325–2332.
12. M. V. Schneider, "A Scaled Hybrid Integrated Multiplier from 10 to 30 GHz," B.S.T.J., *50*, No. 6 (July-August 1971), pp. 1933–1942.
13. W. W. Snell, Jr., "Low-Loss Microstrip Filters Developed by Frequency Scaling," B.S.T.J., *50*, No. 6 (July-August 1971), pp. 1919–1931.
14. R. F. Trambarulo, "A 30-GHz Inverted-Microstrip Circulator," IEEE Trans. Microwave Theory and Techniques, *MTT-19*, No. 7 (July 1971), pp. 662–664.
15. C. L. Ruthroff and W. F. Bodtmann, "A Linear Phase Modulator for Large Baseband Bandwidths," B.S.T.J., *49*, No. 8 (October 1970), pp. 1893–1903.
16. M. Eisenberg, "Almost-Coherent Detection of Phase-Shift-Keyed Signals Using an Injection-Locked Oscillator," unpublished work.
17. R. Adler, "A Study of Locking Phenomena in Oscillators," Proc. IRE, *34*, No. 6 (June 1946), pp. 351–357.
18. R. V. Khoklov, "A Method of Analysis in the Theory of Sinusoidal Self-Oscillations," IRE Trans. Circuit Theory, *CT-7*, No. 4 (December 1960), pp. 398–413.
19. A. J. Viterbi, *Principles of Coherent Communication*, New York: McGraw-Hill, Inc., 1966, pp. 10–75.
20. A. Papoulis, "Narrow-Band Systems and Gaussianity," IEEE Trans. Inform. Theory, *IT-18*, No. 1 (January 1972), pp. 20–27.
21. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, N.B.S., Washington, D. C.: U.S. Government Printing Office, 1964, pp. 966–977.
22. F. K. Bowers, U.S. Patent No. 2,957,947, issued October 25, 1960.
23. The Rand Corporation, *A Million Random Digits With 100,000 Normal Deviates*, New York: The Free Press, 1966, pp. 66, lines 27 through 30.
24. W. R. Bennett, "Statistics of Regenerative Digital Transmission," B.S.T.J., *37*, No. 6 (November 1958), pp. 1501–1542.