A Decomposition Approach to CPM

BIXIO E. RIMOLDI

Abstract—It is shown that any continuous phase modulation (CPM) system can be decomposed into a continuous-phase encoder and a memoryless modulator in such a way that the former is a linear (modulo some integer P) time-invariant sequential circuit and the latter is also time invariant. This decomposition is exploited to obtain alternative realizations of the continuous-phase encoder (and hence of continuous phase modulation) and also to obtain alternative forms of the optimum decoding algorithm. When P is a prime p so that the encoder is linear over the finite field GF(p), it is shown that cascading it with an outside convolutional encoder is equivalent to a single convolutional encoder. It is pointed out that the cascade of the modulator, the waveform channel which we assume is characterized by additive white Gaussian noise, and the demodulator that operates over one symbol interval yield a discrete memoryless channel which can be studied without the distractions introduced by continuous-phase encoding.

I. INTRODUCTION

For most applications requiring transmission of digital data over nonlinear and/or fading channels, constant-envelope digital modulation is desirable. Because of their simplicity, constant-envelope modulation schemes such as phase-shift keying (PSK) and frequency-shift keying (FSK) have often been used. The pressing need for better spectral utilization has led to the search for more bandwidth-efficient modulation schemes. It has long been known that the bandwidth of constant-envelope digital modulation schemes could be reduced by smoothing the variations of the information-carrying phase. This can be done by shaping the phase using an analog filter. The model generally used for this class of modulation schemes is shown in Fig. 1; such a modulation scheme is called continuous phase modulation (CPM) [1], [2].

Besides providing spectral economy, CPM systems exhibit a "coding gain" when compared to PSK modulation. This "coding gain" is due to the memory that is introduced by the phase-shaping filter and that can be exploited by the decoder. The description of CPM modulation by the processing shown in Fig. 1 is not entirely satisfactory. First of all, it is not well fitted to describing the key aspect of a digital modulator, i.e., the memory that it introduces, by means of an analog device, i.e., the phase-shaping filter. This is particularly true since the manner in which CPM modulation exhibits memory resembles in many ways the way in which a convolutionally encoded data sequence exhibits memory—in both cases a "trellis" can be used to display the possible output signals. As pointed out by Massey [7], "it is no accident that the optimum 'demodulators' for these (CPM) systems have turned out to incorporate a 'Viterbi decoder!'" Massey further suggested that it would be desirable to decompose a CPM modulation scheme as shown in Fig. 2, namely, as a continuous-phase encoder (CPE) that is a possibly nonlinear and possibly time-varying finite-state machine, followed by a (possibly time-varying) memoryless modulator (MM).

Such a decomposition would have two obvious advantages. First, it would permit the "encoding" operation to be studied independently of the modulation. This might suggest alternative realizations of the encoder (and hence alternative implementations of CPM) and might suggest alternative forms of the optimum decoding algorithms. Moreover, if the CPE should turn out to be time invariant and linear (in some appropriate finite algebraic structure), then the CPE could be studied by the same techniques that have been developed for convolutional encoders [5], [6] which are also time-invariant systems linear over a finite field. Moreover, if the CPE were linear over a finite field GF(p), then the CPE itself would be a convolutional encoder, and the cascade of an outside convolutional encoder with the CPE would reduce to an equivalent single convolutional encoder. The outside convolutional encoder could be designed to maximize the Euclidean distance (ED) between different output signals, whereas the CPE would serve "to shape the modulated signal's spectrum" [7]. Such outside convolutional encoders are currently being considered for use with CPM systems [8]–[11].

The second obvious advantage of such a decomposition of CPM is that the isolation of the MM would allow one to...
model the cascade of the MM, the waveform channel (that we assume is characterized by additive white Gaussian noise (AWGN)), and the demodulator that operates over one symbol interval, as a discrete memoryless channel. If this MM would turn out to be time invariant (in the sense that the set of possible phase trajectories in one symbol period is the set of time translates of the phase trajectories in any other symbol period), then this memoryless channel would be the usual discrete memoryless channel (DMC), the extensive theory for which [12] could then be applied to the study of CPM.

We show not only that such a decomposition of CPM systems is possible, but also that it can always be done in such a way that the CPE is linear modulo some integer P and time invariant, and such that the MM is also time invariant. In Section II we introduce the modified information-carrying phase for CPM and the tilted-phase trellis that is time invariant (in the sense defined earlier). The modified information-carrying phase is then used to carry out the decomposition of the CPM-system. Section III is devoted to simplifications that derive from the time invariance of the tilted-phase trellis. In Section IV examples are given to show how the encoder of conventional CPM can be modified and/or combined with usual convolutional encoders. The modification of the minimum shift keying (MSK) encoder leads naturally to differential MSK whose encoder and decoder are shown to be simpler than that for MSK and whose probability of error for practical signal-to-noise ratios (S/N) is shown to be reduced by a factor of 2.

II. DECOMPOSITION OF CPM

The transmitted signal for CPM systems can be described [1], [2] by

\[ s(t, \alpha) = \sqrt{2E/T} \cos(2\pi f_0 t + \varphi(t, \alpha) + \varphi_0), \quad t \geq 0 \]  

where the information-carrying phase is

\[ \varphi(t, \alpha) = 2\pi h \sum_{i=0}^{\infty} \alpha_i f(t - iT), \quad t \geq 0 \]

and where we assume without loss of essential generality that \( \varphi(t, \alpha) = 0 \) at time \( t = 0 \). The parameter \( h \) is referred to as the modulation index. For practical purposes, only rational modulation indexes of the form

\[ h = \frac{K}{P} \]

where \( K \) and \( P \) are relatively prime positive integers are of interest. Otherwise, the CPM system has an infinite number of states, and the optimum receiver has infinite complexity. The information sequence

\[ \alpha = (\alpha_0, \alpha_1, \cdots) \]

where

\[ \alpha_i \in \{ \pm 1, \pm 3, \cdots, \pm M - 1 \}, \quad M \text{ even}, i \geq 0 \]

\[ \alpha_i \in \{ 0, \pm 2, \pm 4, \cdots, \pm M - 1 \}, \quad M \text{ odd}, i \geq 0 \]

is an \( M \)-ary sequence. CPM is often formulated only for \( M \) even [1], [2], but Jackson [3] and Omura and Jackson [4] have considered the case where \( M \) is odd. The phase response \( f(t) \) satisfies

\[ f(t) = \begin{cases} 0, & t \leq 0 \\ 1/2, & t > LT \end{cases} \]

where \( L \) is a positive integer that we shall call the memory of the CPM scheme. The symbol energy is \( E \), the data rate is one symbol every \( T \) seconds, and the carrier frequency is \( f_0 \).

The ensemble of all phase trajectories \( \varphi(t, \alpha) \) is the phase tree which completely describes the CPM signal, except for the carrier frequency \( f_0 \) and the phase offset \( \varphi_0 \) (parameters that do not affect performance) and except for the symbol energy \( E \) (a design choice). Fig. 3(a) shows the phase tree for minimum shift keying [13], i.e., the CPM scheme in which

\[ h = 1/2 \]

\[ f(t) = \begin{cases} 0, & t \leq 0 \\ t/2T, & 0 < t \leq T \\ 1/2, & t > T \end{cases} \]

\[ M = 2. \]

The semi-open interval \([nT, nT + T)\) will be referred to as the \( n \)th symbol interval. Because we are interested in \( s(t, \alpha) \) only for \( t \geq 0 \), we shall hereafter assume always that \( n \geq 0 \). Phases that differ by an integer multiple of \( 2\pi \) are, of course, physically indistinguishable. We shall call the phase \( \theta \), when taken modulo \( 2\pi \) to remove this ambigu-
ity, the physical phase, and denote it by $\bar{\theta}$. Formally, we have

$$\bar{\theta} = R_{2\pi}[\theta]$$  \hspace{1cm} (6)$$

where $R_{2\pi}[\cdot]$ is the “modulo $2\pi$ operator” defined by

$$R_{2\pi}[\theta] = \theta - \frac{1}{2\pi} \theta \cdot 2\pi$$

and where $\lfloor \cdot \rfloor$ denotes the largest integer not exceeding the enclosed number. We shall later use the fact that $R_{2\pi} + R_{2\pi} = R_{2\pi}$.

The “phase trellis” of the physical phase $ij(t,a)$ for MSK is shown in Fig. 3(b). This trellis is seen to be time varying in the sense that the physical phase trajectories in the even-numbered symbol intervals are not time translates of those in the odd-numbered symbol intervals. However, if one measures phase relative to the lowest phase trajectory in Fig. 3(a) (the highest phase trajectory could also have been chosen), then this new phase $\psi(t,a)$, which is defined by $\psi(t,a) = \varphi(t,a) + \pi(1/2)t/T$, has the phase tree and physical phase trellis shown in Fig. 4(a) and (b), respectively. The trellis of Fig. 4(b) is now seen to be time invariant, i.e. (after the initial “transient”), the physical phase trajectories in any two symbol intervals are time translates of one another. This time-invariant choice of phase for MSK was introduced implicitly by Amoroso and Kivett [16] and quite explicitly by Morales-Moreno and Pasupathy [18].

We now show that for general CPM such a time-invariant phase trellis can be obtained by defining

$$\psi(t,a) = \varphi(t,a) + \pi h(M-1)t/T.$$  \hspace{1cm} (8)$$

We shall call $\varphi(t,a)$ and $\psi(t,a)$ the traditional phase and the tilted-phase, respectively, for CPM. To show that the physical tilted-phase $\psi(t,a)$ always has a time-invariant trellis, we first use (2), (5), and (8) to obtain

$$\psi(t,a) = \pi h \sum_{i=0}^{n-L} a_i + 2\pi h \sum_{i=-n-L+1}^{n} a_i f(t-2iT) + \pi h(M-1)t/T,$$

where, by way of convention, $a_{-n} = a_{-n+1} = \cdots = a_{-1} = 0$. It is now convenient to introduce the modified data sequence $U = (U_{-L}, U_{-L+1}, \cdots)$ defined by

$$U_i = (a_i + (M-1)/2).$$  \hspace{1cm} (10)$$

Note that, regardless of whether $M$ is even or odd, $U_i \in \{0,1,\cdots, M-1\}$ so that the modified data digits are just $M$-ary digits. Using (10) in (9) and with a slight abuse of notation (namely, writing $\psi(t,U)$ in place of $\psi(t,a)$), and letting $t = r + nT$, we obtain

$$\psi(r + nT, U) = 2\pi h \sum_{i=0}^{n-L} U_i + 4\pi h \sum_{i=0}^{L-1} U_{-i-1} f(t+iT) + \pi h(M-1)t/T$$

$$- 2\pi h(M-1) \sum_{i=0}^{L-1} f(t+iT) + (L-1)(M-1)\pi h, \hspace{0.5cm} 0 \leq r < T.$$  \hspace{1cm} (11)$$

We see that all time-dependent terms on the right side of (11) depend only on the translated time variable $r - t - nT$. Hence the possible physical phases $\psi(r + nT, U)$ in any two symbol intervals will differ only by time translations after an initial transient that allows the time-independent data-dependent term on the right side of (11) to assume all of its possible values modulo $2\pi$ provided that these possible values are the same in all subsequent symbol intervals. However, the time-independent data-dependent term on the right side of (11), taken modulo $2\pi$, gives

$$R_{2\pi}[2\pi h \sum_{i=0}^{n-L} U_i] = R_{2\pi}[2\pi(K/P) \sum_{i=0}^{n-L} U_i]$$

$$= R_{2\pi}[2\pi(K/P) R_{P} \sum_{i=0}^{n-L} U_i]$$  \hspace{1cm} (12)$$

where $R_{P}[\cdot]$ is the modulo $P$ operator for integers. That the sum of data digits in (12) can validly be reduced modulo $P$ follows from the fact that adding $P$ to this sum changes the argument of $R_{2\pi}[\cdot]$ by $2\pi K$. From (12) we see that there are only $P$ possible values of the time-independent data-dependent term from (11) reduced modulo $2\pi$ and, moreover, that all $P$ values are possible when $n$
satisfies
\[(n - L + 1)(M - 1) \geq P - 1 \quad (13)\]
since this ensures that the sum of \(n - L + 1\) \(M\)-ary digits can reach the largest required value \(P - 1\).

We have thus shown that the physical tilted-phase \(\psi(t, U)\) for CPM always has a time-invariant trellis. To complete the decomposition of a CPM modulator into a continuous-phase encoder and memoryless modulator, we need only to determine the \(MM\) input, i.e., to find the information needed to specify which of the physical phase trajectories should be emitted by the \(MM\) in the current symbol interval and then to find the recursion by which the CPE can update this \(MM\) input with the next data digit to produce the next \(MM\) input.

A. The Memoryless Modulator

Using (8) in (1), and replacing the tilted phase \(\psi(t, U)\) by the physical tilted phase \(\tilde{\psi}(t, U)\), we obtain
\[s(t, U) = \sqrt{2E/T} \cos(2\pi f_s t + \tilde{\psi}(t, U) + \psi_0) \quad (14)\]
where the new frequency variable \(f_s = f_0 - h(M - 1)/2T\) has been introduced to compensate for the offset between \(\psi(t, U)\) and \(\psi(t, U)\). From (6), (7), (11), and (12) the physical tilted phase is given by
\[\psi(t, U) = R_2\left[\psi(t + nT, U)\right]\]
\[= R_2\left[2\pi h \sum_{i=0}^{n-L} U_i + 4\pi h \sum_{i=0}^{L-1} U_{-i} f_i(t + iT) + W(\tau)\right] \]
\[= R_2\left[2\pi hR_p \sum_{i=0}^{n-L} U_i \right.\]
\[+ 4\pi h \sum_{i=0}^{L-1} U_{-i} f_i(t + iT) + W(\tau)\right],\]
\[0 \leq \tau < T \quad (15)\]
where
\[W(\tau) = \pi h(M - 1)\tau/T - 2\pi h(M - 1) \sum_{i=0}^{L-1} f_i(t + iT) + (L - 1)(M - 1)\pi h, \quad 0 \leq \tau < T \quad (16)\]
represents the data-independent terms. From (15) we see that the \(MM\) input, which completely specifies the physical phase and, therefore, completely specifies the output signal, can be defined as follows:
\[X_n = [U_{n-L+1}, \ldots, U_{n-1}, V_n] \quad (17)\]
where
\[V_n = R_p \sum_{i=0}^{n-L} U_i \quad (18)\]
accounts for the accumulated phase at time \(n\) due to the data input symbols from time 0 to \(n - L\), inclusive. Again slightly abusing the notation by writing
\[\tilde{\psi}(\tau, X_n)\] instead of \(\psi(\tau + nT, U)\), \(0 \leq \tau < T\) and
\[s(\tau, X_n)\] instead of \(s(\tau + nT, U)\), \(0 \leq \tau < T\), we obtain
\[s(\tau, X_n) = \sqrt{2E/T} \cos(2\pi f_s (\tau + nT) + \tilde{\psi}(\tau, X_n) + \psi_0), \quad 0 \leq \tau < T \quad (19)\]
For realization purposes, it is convenient to decompose (19) into in-phase and quadrature components, as proposed in [2]:
\[s(\tau, X_n) = I(\tau, X_n) \Phi_I(\tau) + Q(\tau, X_n) \Phi_Q(\tau) \quad (20)\]
where
\[I(\tau, X_n) = \sqrt{E/T} \cos(\tilde{\psi}(\tau, X_n)) \quad (21a)\]
\[Q(\tau, X_n) = \sqrt{E/T} \sin(\tilde{\psi}(\tau, X_n)) \quad (21b)\]
and
\[\Phi_I(\tau) = \sqrt{1/2 \cos[2\pi f_s (\tau + nT) + \psi_0]} \quad (22a)\]
\[\Phi_Q(\tau) = -\sqrt{1/2 \sin[2\pi f_s (\tau + nT) + \psi_0]} \quad (22b)\]
The conceptual diagram in Fig. 5 shows the MM defined by (15) and (20)–(22). Note from (17) and (18) that there are \(PM^L\) possible values for the input symbol \(X_n\). Note also from (15) that (after a transient whose length is specified by (13)) there can be up to \(PM^{L-1}\) possible values of the physical phase at the beginning of each symbol interval (there are exactly \(PM^{L-1}\) possible values if and only if the second term in (15), evaluated for \(\tau = 0\) and taken modulo \(2\pi\), gives different values for all of the \(M^{L-1}\) choices for \(U_{n-L+1}, \ldots, U_{n-1}\)).

B. The Continuous-Phase Encoder

As pointed out previously, the task of the CPE is to update the \(MM\) input \(X_n\) using the next data digit \(U_{n+1}\) to produce the next \(MM\) input \(X_{n+1}\). Replacing \(n\) by \(n + 1\) in (18) and making use of (7), we obtain
\[V_{n+1} = R_p \left[\sum_{i=0}^{n-L+1} U_i\right] = R_p \left[\sum_{i=0}^{n-L} U_i + U_{n-L+1}\right] \]
\[= R_p \left[R_p \left[\sum_{i=0}^{n-L} U_i + U_{n-L+1}\right]\right] \]
\[= R_p \left[V_n + U_{n-L+1}\right]. \quad (23)\]
Obviously, the updating of the first $L$ components of $X_n$ (see (17)) can be accomplished by shifting the last $L$ data digits which are stored in a shift register. A possible realization of the CPE is shown in Fig. 6, where the sum is taken modulo $P$. This configuration holds for any CPM described by (1)-(5). Note that the CPE performs a linear operation in the ring of integers modulo $P$. We define the state of the CPE to be the $L$-tuple

$$\sigma_n = \{U_{n-1}, \ldots, U_{n-L+1}, V_n\},$$  

which can take on $M^{L-1}P$ different values. The CPE and the MM for MSK are shown in Fig. 7; this CPE is precisely that given by Forney [15, p. 270], and this MM is equivalent to Forney's.

We are now able to decompose any CPM system into a continuous-phase encoder and a memoryless modulator. The encoder is a time-invariant linear sequential circuit (LSC) containing a modulo $P$ adder, where $P$ is the denominator of the modulation index $h$ when reduced to a ratio of relatively prime integers, and $L$ delays, where $L$ is the number of intervals under the varying part of the phase response $f(t)$.

III. CONSEQUENCES OF THE TILTED-PHASE TRELLIS

Because of their equivalence, everything that can be done with the tilted-phase trellis can also be done with the traditional phase trellis. Nevertheless, when one tries to get insight into a given process for purposes of analysis, simplification of related algorithms, realization, etc., one finds it easier to consider the time-invariant process described by the tilted-phase trellis. In the following, we consider some consequences of the time-invariant tilted-phase trellis on the MM and on the optimum receiver.

The MM of a particular CPM scheme must be able to generate the set of all different interval physical phase trajectories of the corresponding phase trellis. In a particular symbol interval, the number, say $V$, of different interval phase trajectories is the same in both the traditional and tilted-phase trellises. Because of time invariance, the set of $V$ interval phase trajectories in any interval of the tilted-phase trellis is the same, but this is not true for the traditional phase trellis. In the MSK case, there are only four instead of eight different interval phase transitions for the tilted phase as opposed to the traditional phase trellis, as can be seen from Figs. 4(b) and 3(b), respectively.

Now we give an example of a simplification that appears to be more easily seen from the time-invariant trellis than from the time-variant one. The simplification applies to the optimum coherent MSK decoder. Many solutions have been proposed to coherently decode MSK. In chronological order they are De Buda's in-phase and quadrature receiver [14] (also called the parallel technique), the Viterbi algorithm (VA) [15], the serial technique of Amoroso and Kivett [16], and Massey's receiver [17] which was proposed for differential MSK. All four techniques are optimal. However while the decoders in [17], [14], and [16] make a decision observing two symbol intervals of the received signal only, the VA decoder of [15] seems to need to
observe the complete sequence before taking an optimal decision, as is usual for the VA. Now we show that the VA, when used to decode MSK, can optimally decode with one unit delay, exactly as the other receivers. Furthermore, we show that a maximum likelihood (ML) state sequence receiver will result. From the state sequence, the encoder input sequence can then be easily found. Familiarity with the VA [15] is assumed.

Consider the received signal

\[ r(t) = s(t, a) + n(t) \]  

(24)

where \( s(t, a) \) is the output of an MSK modulation system and \( n(t) \) is AWGN. The metric to be minimized is the Euclidean distance (ED) between the received signal \( r(t) \) and the estimated signal sequence. Equivalently, because the energy per symbol interval of the transmitted signal \( s(t, a) \) is a constant, one can maximize the correlation between the received signal and the estimated signal [2].

Consider the trellis of Fig. 8, which is the state trellis of the MSK encoder (Fig. 7) with transitions labeled to represent specific modulator output signals. Note that for MSK,

\[ S_n(t) = -S_t(\tau) \quad S_1(t) = -S_0(\tau) \]  

(25)

The branch metric \( \lambda_n(S_i) \), \( i = 1, \ldots, 3 \), is the correlation between the received signal \( r(t) \) in the \( n \)th symbol interval and the modulator output signal corresponding to the considered branch, i.e.,

\[ \lambda_n(S_i) = \int_{nT}^{nT+T} r(t)S_i(t-nT) \, dt, \quad i = 0, \ldots, 3. \]  

(26)

Because of (25), we have

\[ \lambda_n(S_0) = -\lambda_n(S_3) \quad \lambda_n(S_1) = -\lambda_n(S_2). \]  

(27)

Our claim is the following: if the survivors at depth \( n+1 \) of the trellis go through the same state \( a_n \) at depth \( n \), then the survivors at depth \( n+2 \) go through the same state at depth \( n+1 \). The proof is broken down in two cases.

**Case 1:** Suppose that the survivors at depth \( n+1 \) in Fig. 9 go through the same state \( a_n = 0 \) at depth \( n \). The claim is proved for this case if and only if at depth \( n+2 \) the VA prunes both lower or both upper (dashed) branches reaching the states \( a_{n+1} \). The upper branch reaching \( a_{n+2} = 0 \) is pruned if and only if

\[ \lambda_n(S_0) + \lambda_{n+1}(S_0) > \lambda_n(S_1) + \lambda_{n+1}(S_1). \]  

(28)

The upper branch reaching \( a_{n+2} = 1 \) is pruned if and only if

\[ \lambda_n(S_0) + \lambda_{n+1}(S_0) > \lambda_n(S_1) + \lambda_{n+1}(S_1). \]  

(29)

However, (27) implies that (28) and (29) are the same condition, which proves the claim for case 1.

![Fig. 8. State trellis of MSK with output signals.](image)

**Case 2:** Suppose now that the survivors at depth \( n+1 \) go through the same state \( a_n = 1 \) at depth \( n \). The upper branches reaching \( a_{n+2} = 0 \) and \( a_{n+2} = 1 \) are pruned if and only if

\[ \lambda_n(S_3) + \lambda_{n+1}(S_3) > \lambda_n(S_2) + \lambda_{n+1}(S_2) \]  

(30)

and

\[ \lambda_n(S_3) + \lambda_{n+1}(S_3) > \lambda_n(S_2) + \lambda_{n+1}(S_2). \]  

(31)

respectively. Again (27) implies that (30) and (31) are the same condition so that the survivors at depth \( n+2 \) must go through the same state at depth \( n+1 \). This completes the proof.

Because at depth 1 all survivors go trivially through the same previous state \( a_0 = 0 \), we can conclude that survivors at any depth \( n+1 \) go through the same state at depth \( n \). The VA can, therefore, decode the state sequence until the \( n \)th node (or, equivalently, the input sequence until the \( (n-1) \)th digit) after observation of \( r(t) \) only up to the \( n \)th symbol interval, i.e., with a one symbol delay. Note that we have proved the following ML decoding rule for the state sequence of MSK:

\[ \delta_n = n \quad \text{if (28) is true, then } \delta_{n+1} = 0, \text{ else } \delta_{n+1} = 1 \]  

(32)

\[ \delta_n = 1 \quad \text{if (30) is true, then } \delta_{n+1} = 0, \text{ else } \delta_{n+1} = 1 \]  

(33)

where \( \delta_n \) denotes the estimate of \( a_n \). However, (32) and (33) are also the same! This also follows from property (27). Thus the decoding rule simplifies to

\[ \lambda_n(S_0) + \lambda_{n+1}(S_0) > \lambda_n(S_1) + \lambda_{n+1}(S_1), \]  

then \( \delta_{n+1} = 0, \) else \( \delta_{n+1} = 1. \)  

(34)

leading to the optimum state sequence receiver of Fig. 10. Given the state sequence, one can determine the symbol sequence by the inverse of the operation that generates the state \( a_{n+1} \). From Fig. 7 this reverse operation is seen to be

\[ U_n = a_{n+1} - a_n \]  

(35)

and thus

\[ \hat{U}_n = \delta_{n+1} - \delta_n \]  

(36)
where \( \hat{U}_n \) denotes the estimate of \( U_n \). The ML receiver for MSK is, therefore, the cascade of the circuits shown in Fig. 10 and in Fig. 11. Note that the ML state sequence decoder of Fig. 10 is the time invariant (and fully equivalent) version of Massey’s time-varying receiver \([17]\) proposed to optimally decode differential MSK.

![Fig. 10. Optimum state sequence receiver for MSK.](image)

![Fig. 11. Inverse of MSK state encoder.](image)

We conclude this section on the tilted-phase trellis by noting an interesting fact: the tilted-phase is a linear function of the input sequence \( U_n \) whose alphabet contains zero. Therefore, when the input sequence is the zero sequence, all data-dependent terms will vanish. There remain only the data-independent terms that we have explicitly separated in (15) and called \( W(\tau) \). If the derivative of \( W(\tau) \) depends on \( \tau \), the transmitted signal frequency (i.e., the derivative of the information carrying phase) will have a periodic (with period \( T \)) data-independent variation. However, such a variation increases the bandwidth without containing information. One might expect that good CPM schemes have a \( \hat{W}(\tau) \) with a constant derivative. Indeed, this is true for the most popular CPM classes, like those considered in \([9]\), where the derivative of \( f(t) \) has the form: rectangular over \( L \) symbols (LREC), or raised cosine over \( L \) symbols (LRC), or triangular over \( L \) symbols (LTRI), etc.

### IV. USEFULNESS OF THE CPE/MM SEPARATION

The most direct consequence (although not the most important one) of the separation just carried out is that it offers a new possibility to realize a CPM modulation system as shown in Figs. 5 and 6. A similar method has been proposed by Aulin et al. \([2]\). However, the complexity of both the CPE and MM, as given here, is reduced compared to the corresponding devices of \([2]\). Aulin et al. \([2]\) use the traditional phase, and thus their MM must either store more possible phase trajectories than does the MM for the tilted phase or it must equivalently compute additional phase trajectories if the reduced set of phase trajectories is stored. Aulin et al. \([2]\) use, in place of the \( P \)-ary state component \( V_n \) used here, a state component that (because of the time variation of the traditional phase trellis) can take on \( 2P \) different values.

Now we give some examples of the new capabilities offered by the CPE. To have a CPE with a well-defined algebraic structure, we restrict our attention hereafter to \( M \)-ary CPM’s having \( h = K/P \), where \( K \) and \( P \) are relatively prime positive integers, and \( M \) is such that

\[
M = P^k \tag{37}
\]

for some integer \( k \). This limitation is not very restrictive. Indeed, for almost all the coded systems considered in \([9]\) and \([11]\), the value of \( h \) maximizing the minimum Euclidean distance is such that (37) can be satisfied (two exceptions exist where, however, the minimum Euclidean distance is almost maximum if \( h \) is approximated to satisfy (37)). When (37) is satisfied, the encoder can always be defined over the ring of integers modulo \( P \). Indeed, the \( P^k \)-ary input variable \( U_n \) can be represented in radix \( P \) form:

\[
U_n = U_n^1P^{k-1} + U_n^2P^{k-2} + \cdots + U_n^k, \quad U_j^k \in \{0,1,\cdots,P-1\}, \quad j = 1,\cdots,k,
\]

and every \( U_j^k \) is independently encoded by the CPE as in Fig. 12. Note that, because of the modulo \( P \) operation, the radix \( P \) form of the state variable \( V_n \) differs from zero only in the least significant digit (LSD), i.e., \( V_n \) is represented by a single \( P \)-ary digit with the value

\[
V_n = R_P \left\lfloor \frac{a - b}{k} \right\rfloor
\]

**Example 1:** Fig. 12 shows the CPE for a conventional CPM scheme satisfying (37). Every input line is independently processed by what we shall call a subencoder. The input/output relation of the \( j \)th subencoder (starting from the upper subencoder) can be concisely expressed by

\[
X'(D) = U'(D)G'(D), \quad j = 1,\cdots,k \tag{38}
\]

where \( U'(D) = U_j^1 + U_j^2D + U_j^3D^2 + \cdots \) is the \( D \) transform of the input sequence to the \( j \)th subencoder, \( X'(D) \) is the \( D \) transform of the output \( X_j^1 \) from the \( j \)th subencoder, and \( G'(D) \) is the transfer function matrix of the \( j \)th subencoder. The transfer function matrices of the subencoders are

\[
G'(D) = \begin{cases} [1,D,\cdots,D^{k-1},0], & 1 \leq j < k \\ [1,D,\cdots,D^{k-1},D^{k-1}(1-D)], & j = k. \end{cases} \tag{39}
\]

Because the \( D \) transform \((1-D)\) always has an inverse over the ring of integers modulo \( P \), i.e., \(1/(1-D) = 1 + D + D^2 + \cdots\), we can define an equivalent \( k \)th subencoder as

\[
G_k'(D) = [1-D,D(1-D),\cdots,D^{L-1}(1-D),D^L] \tag{40}
\]

where every component has been multiplied by \((1-D)\). The subencoder (40) is equivalent to the subencoder (39) because the input \( U_k'(D) \) to the first generates the same output as the input \( U_k'(D)/(1-D) \) to the second. Such a
definition of equivalence is motivated by the fact that all input sequences are assumed equiprobable [5]. The subencoder with transfer function (40) has no feedback but requires \( L \) adders modulo \( P \). However, one sees that

\[
G_k^L(D) = G_k^L(D)A
\]

where

\[
G_k^L(D) = [1, D^k, \cdots, D^L-1, D^L]
\]

and where \( A \) is an invertible \((L + 1) \times (L + 1)\) matrix with components \( a_{ij} \) defined as

\[
a_{ij} = \begin{cases} 
1, & \text{if } i = j \\
-1, & \text{if } i = j + 1 \\
0, & \text{otherwise.}
\end{cases}
\]

The matrix \( A \) performs an invertible memoryless mapping (it contains no delay operator \( D \)) on the output of \( G_k^L(D) \) that can either be explicitly performed as shown by the dotted circuit on the right side of Fig. 13 or that can be incorporated into the mapping performed by the MM. In this case, the complexity of the MM is unchanged, but the encoder is simplified. The resulting modulation is a modified continuous phase modulation (MCPM) scheme with the same set of output signals, same spectrum, and same probability of Viterbi decoding error events [15] but with a different mapping from information digits to output signals and thus possibly a different bit error probability compared to conventional CPM. More insight into such differences will be gained from the next example.

Example 2: The CPE for MSK and the CPE for modified MSK obtained according to the transformation of Example 1 are

\[
G(D) = \left[1, D/(1 - D)\right]
\]

and

\[
G_m(D) = \left[1 + D, D\right] \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} = [1, D],
\]

respectively. Recall that if \( U_m(D) \) is the input to the modified MSK encoder (45), then the input to the MSK encoder (44) producing the same modulated signal will be \( U(D) = (1 - D)U_m(D) \). This shows that modified MSK is just differential MSK (DMSK), but this reduction to differential modulation does not hold for MCPM when \( k > 1 \). Consequently, the time-invariant decoder designed for MSK can be used to decode DMSK, provided that the MSK decoder output \( \hat{U}(D) \) is divided by \( 1 - D \), i.e.,

\[
\hat{U}_m(D) = \hat{U}(D)/(1 - D)
\]

where \( \hat{U}_m(D) \) is the optimum estimate of the information sequence belonging to the DMSK signal. However, rewriting (36) using the \( D \)-transform notation, i.e.,

\[
\hat{U}(D) = \hat{\delta}(D)(1 - D)
\]

where \( \hat{\delta}(D) \) is the \( D \) transform of the state sequence estimate \( \hat{\delta}_n \) of MSK, and inserting (47) into (46), we obtain

\[
\hat{U}_m(D) = \hat{\delta}(D).
\]

This is the output of the optimum state sequence receiver of MSK shown in Fig. 10. Thus with DMSK, not only is the encoder simpler, but it is also the optimum decoder.

What about the bit error probability \( P_b \)? It is well-known that, for practical signal-to-noise ratios, a typical error event occurs when the decoder makes a mistake by choosing the detour with the smallest Euclidean distance to the correct path [15]. However, with MSK and DMSK, the detour with the smallest ED is the shortest detour. Indeed, every split and every merge increases the ED by the same
Fig. 13. CPE for conventional $M = P^k$-ary CPM modulation. Note 1: deletion of this box causes CPM to be MCPM. Note 2: Memoryless function of this box can be incorporated into memoryless modulator (MM).

Fig. 14. Encoded (with usual convolutional encoder) CPM system. (a) Usual description. (b) New description. (c) Equivalent encoded CPM system.
amount. Therefore, the detour which splits in one interval and merges in the next interval must give the smallest ED. In terms of the state sequence estimate (the output in Fig. 10), a single error will therefore typically appear between a sequence of correct state estimates. To decode MSK signals, however, the circuit in Fig. 11, which follows the state estimator, converts a single error into a run of two errors. Equivalently, one can say that the MSK decoder will make bit errors in adjacent pairs where the DMSK decoder would make a single bit error. We conclude that for practical S/N, DMSK has only half the bit error probability of ordinary MSK.

Example 3: In recent years, to reduce error probability, CPM has been studied together with convolutional encoders [8]–[11]. In [9] and [11], among others, the system in Fig. 14(a) was considered. Let the modulation index \( h \) be 1/2 (among the \( h \) values considered in [9] and [11], \( h = 1/2 \) maximizes the minimum ED between distinct transmitted signals for this coded modulation system). From our result in Section II, we obtain the system shown in Fig. 14(b), where any addition is taken modulo 2. Hereafter, a convolutional encoder will be called a channel encoder (CE) to distinguish it from the CPE. Note that in our system no conceptual need exists for the natural mapper (denoted \( Q_1 \) in [9], [11]) because the CPE accepts the input in radix two form. The resulting single encoder in Fig. 14(b) has the transfer function

\[
G_r(D) = \left[ D^2, 1 + D + D^2, (1 + D + D^2)D/(1 - D) \right].
\]

As in Example 1, we can multiply (49) by 1 - \( D \) to obtain the equivalent feedback-free encoder

\[
G_e(D) = \left[ D^2 - D^3, 1 - D^3, D + D^2 + D^3 \right],
\]

which again can be decomposed into

\[
G_M(D) = [1, D, D^2, D^3],
\]

and a memoryless mapping matrix

\[
B = \begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
1 & 0 & 1 \\
1 & 1 & 1
\end{bmatrix}.
\]

The mapping \( B \) can be integrated into the lookup table of the MM. The resulting equivalent encoded CPM system is shown in Fig. 14(c).

V. DISCUSSION AND CONCLUSION

We started in Section II by observing that, in contrast to the trellises generally given in the literature, the phase trellis of CPM can be made time invariant. To show this, we introduced the modified information carrying phase and the tilted-phase trellis. After a few manipulations on the tilted phase, it became obvious how to decompose the CPM system into a continuous-phase encoder and memoryless modulator. The CPE is a sequential circuit linear (in the ring of integers modulo \( P \)) with input alphabet \( \{0, 1, \cdots, M - 1\} \), where \( M \) is the size of the CPM input alphabet, containing \( L \) delays and a modulo \( P \) adder, where \( P \) is the denominator of the modulation index \( h = K/P \). \( K \) and \( P \) are relatively prime positive integers, and \( L \) is the number of intervals under the varying part of the phase response \( f(t) \).

Section III was devoted to the physical tilted-phase trellis which is time invariant. It was shown that a MM that is based on the tilted-phase trellis has lower complexity compared with an equivalent solution based on the traditional phase trellis. To show the kind of advantages one can obtain by considering the decoder based on the tilted-phase trellis, we derived the optimum ML decoder for MSK. We showed that the Viterbi algorithm can make an optimum decoding decision with a one-unit delay. The resulting decision strategy led to an optimum receiver which was composed of an optimum state sequence receiver followed by the inverse of the MSK state encoder. The optimum state sequence decoder is the time-invariant version of Massey's time-varying receiver [17] designed to decode differential MSK.

We began Section IV by pointing out that the CPE is also simpler than a previous nonlinear encoder structure that can be recognized in [2]. Three examples followed to show how the CPE can be transformed and/or combined with usual convolutional encoders. To do so, a subclass of CPM systems was considered, leading to CPE's over the ring of integers modulo \( P \). In Example 1, we discussed a transformation leading to a modified CPM system (MCPM). Consequences of the MCPM system were analyzed in Example 2 for the MSK case. It was shown that both the encoder and the decoder are simpler than for conventional MSK, and that the bit error probability for practical S/N is reduced by a factor of two. The last example showed how to combine CPE with an outside convolutional encoder. The resulting single encoder was simplified by the same technique used in Example 1, and a simple structure resulted.

ACKNOWLEDGMENT

The author is much indebted to Prof. Dr. J. L. Massey for having encouraged this work and for many helpful discussions. He also wishes to acknowledge Prof. Dr. P. Leuthold for his support and Mr. H. Ochsner for linguistic assistance.

REFERENCES


