CONTINUOUS-PHASE MODULATION

Continuous-Phase Modulation

Continuous-phase modulation (*CPM*) is a general form of digital phase modulation. The earliest digital phase modulation is the phase-shift keying (*PSK*) and was established by the mid-1960s. The modulation method is both linear and memoryless. A *PSK* signal is generated by shifting the phase of a carrier signal at signaling interval boundaries between a predetermined set of phase values to represent the digital information that is being transmitted. The information-carrying phase is constant in each interval except for discontinuous jumps at the beginning of each signaling interval. However, such abrupt switching from one phase to another results in relatively large spectral side lobes, and consequently, requires a large frequency band to transmit the signal. But, in many practical applications, the available radio spectrum is limited. This fact naturally imposes the development of spectrum efficient digital signaling techniques. Various methods were proposed in 1970s in the interest of reducing the bandwidth, constituting a general class of digital phase-modulation techniques (1,2). In this class, the carrier phase is constrained to change in a variety of continuous fashion from symbol to symbol unlike that of *PSK* signals, hence the name continuous phase modulation. The continuous nature of their phase provides spectral efficiency compared to the *PSK* modulated signals, with a varying degree of spectral roll-off factors. *CPM* signals are named with respect to the shape of their continuous phase transitions. Examples include minimum-shift keying (*MSK*), tamed frequency modulation (*TFM*), and the well-known Gaussian minimum-shift keying (*GMSK*), which is adopted by global system for mobile communications (*GSM*).

The continuous-phase nature of *CPM* signals also introduces memory into the transmitted signal and offers power gain over *PSK* signals. *CPM* achieves all these (spectral and power performance) yet with constant envelope, which is suitable for nonlinear amplification and is an important attribute in satellite, terrestrial digital radio, and mobile radio communications. This is another advantage of *CPM* that made it a popular transmission technique in the last two decades. Consequently, intrinsic to *CPM* is its continuous phase and the resulting memory, providing spectral efficiency and power gain over their *PSK* counterparts, respectively.

Signal and System Descriptions

It is essential for a proper understanding of *CPM* to resort to its mathematical description, and it is necessary to introduce some terms to establish a systematic classification. To begin with, let us describe the transmitted constant envelope *CPM* signal at the nth signaling interval after Aulin and Sundberg (3) as

$$s(t, \alpha_n) = \sqrt{\frac{2E}{T}} \cos[2\pi f_0 t + \phi(t, \alpha_n) + \varphi_0] \qquad (1)$$

where $\phi(t, \alpha_n)$ is the information carrying phase and given by

$$\phi(t, \alpha_n) = 2\pi h \int_{-\infty}^{t} \sum_{i=-\infty}^{n} \alpha_i \, g(\tau - iT) \, d\tau,$$
$$nT \le t \le (n+1)T \qquad (2)$$

in which α_n is the M-ary modulation symbol corresponding to the nth input x_n, drawn from the set { $\pm 1, \pm 3,$ $\ldots, \pm (M-1)$} with the relation $\alpha_n = 2x_n - M + 1$. For example, the modulation symbols for a binary ($M = 2$) CPM are ± 1 while taking values $\pm 1, \pm 3$ for a quarternary system ($M = 4$). That is, for a sequence of quarternary input data $x = \{0, 1, 2, 3\}$, the modulation symbols are assigned as $\alpha = \{-3, -1, 1, 3\}$. E denotes the symbol energy, T the symbol interval, f_0 the carrier frequency, and π_0 is an arbitrary constant phase shift, which without loss of generality can be set to zero in the case of coherent transmission. The vaable h is referred to as modulation index, which is related to the size of phase variation $\phi(t, \alpha_n)$, and $g(t)$ is the instantaneous frequency pulse function. Defining the waveform $q(t)$ as baseband phase pulse

$$q(t) = \int_{-\infty}^{t} g(\tau) d\tau \qquad (3)$$

the information-carrying phase of the *CPM* signal $\phi(t, \alpha_n)$ in Eq. (2) can be written as

$$\phi(t, \alpha_n) = 2\pi h \sum_{i=n-L+1}^{n} \alpha_i q(t - iT) + \pi h \sum_{i=-\infty}^{n-L} \alpha_i,$$
$$nT \le t \le (n+1)T \qquad (4)$$

where it is assumed that the pulse $g(t)$ is of finite duration and has length L, hence occupying the interval $(0, LT)$. The first term on the right-hand side of Eq. (4) governs the change of the phase in response to the modulation symbol α_n. The second term is the accumulated phase prior to the input α_n, which is constant over the symbol interval. As seen from Eq. (4), the information-carrying phase $\phi(t, \alpha_n)$ at the nth symbol interval depends not only on the present modulation symbol α_n but also all the transmitted symbols prior to α_n. More precisely, the initial value of the phase depends on all the previously transmitted symbols, while the shape of its transition at the nth interval is influenced by α_n as well as the $L - 1$ symbols prior to α_n. It is clear from Eqs. (3) and (4) that the nature of the frequency or phase pulse defines a particular *CPM* scheme. It is therefore apparent that an infinite variety of *CPM* signals can be generated by choosing different pulse shapes $g(t)$ and modulation index h and the alphabet size M. For example, continuous-phase frequency shift keying (*CPFSK*) is a special case of *CPM*, which uses the rectangular pulse $g(t)$. Another example is a special case of *CPFSK* with modulation index $h = 0.5$, which is the scheme known as *MSK*. However, it is worth nothing here that $g(t)$ and $q(t)$ are closely related as seen in Eq. (3), and can be used interchangeably to define a *CPM* scheme. In order for the modulation to be causal, the phase pulse is chosen such that $q(t) = 0$ for $t \le 0$. The frequency pulse is also normalized so as to integrate to $\frac{1}{2}$, that is, $q(t) = \frac{1}{2}$ for $t \ge LT$. From Eq. (4), the maximum phase change in a symbol interval would then be ($M - 1)h\pi$ radians.

CPM schemes are classified into two categories depending on the length of their frequency pulse $g(t)$. If $g(t)$ extends over only one symbol interval, that is, $g(t) = 0, t > T$, the *CPM* signal is called full-response and corresponds to the case $L = 1$. It is called partial-response *CPM* when $g(t)$ spills over more than one interval, which corresponds to the case $L > 1$. To illustrate the structure of *CPM* and its performance, mainly the rectangular (*REC*), the raised-cosine (*RC*), and the half-cycle sinusoid (*HCS*) frequency pulseunctions are used in the subsequent sections. These pulse names are commonly used as *LREC*, *LRC*, and *LHCS*, with

$$
\text{LREC} \quad g(t) = \begin{cases} \dfrac{1}{2LT} & 0 \le t \le LT \\ 0 & \text{otherwise} \end{cases}
$$

$$
q(t) = \begin{cases} 0 & t < 0 \\ \dfrac{t}{2LT} & 0 \le t \le LT \\ 1/2 & t \ge LT \end{cases}
$$

$$
\text{LRC} \quad g(t) = \begin{cases} \dfrac{1}{2LT}\left[1 - \cos\left(\dfrac{2\pi t}{LT}\right)\right] & 0 \le t \le LT \\ 0 & \text{otherwise} \end{cases}
$$

$$
q(t) = \begin{cases} 0 & t < 0 \\ \dfrac{1}{4\pi}\left[\dfrac{t}{2LT} - \sin\left(\dfrac{2\pi t}{LT}\right)\right] & 0 \le t \le LT \\ 1/2 & t \ge LT \end{cases}
$$

$$
\text{LHCS} \quad g(t) = \begin{cases} \dfrac{\pi}{4LT}\left[\sin\left(\dfrac{\pi t}{LT}\right)\right] & 0 \le t \le LT \\ 0 & \text{otherwise} \end{cases}
$$

$$
q(t) = \begin{cases} 0 & t < 0 \\ \dfrac{1}{4}\left[1 - \cos\left(\dfrac{\pi t}{LT}\right)\right] & 0 \le t \le LT \\ 1/2 & t \ge LT \end{cases}
$$

Fig. 1. Various frequency and phase pulse functions. Phase pulse or its derivative frequency pulse characterize *CPM* schemes with their error and spectral performance.
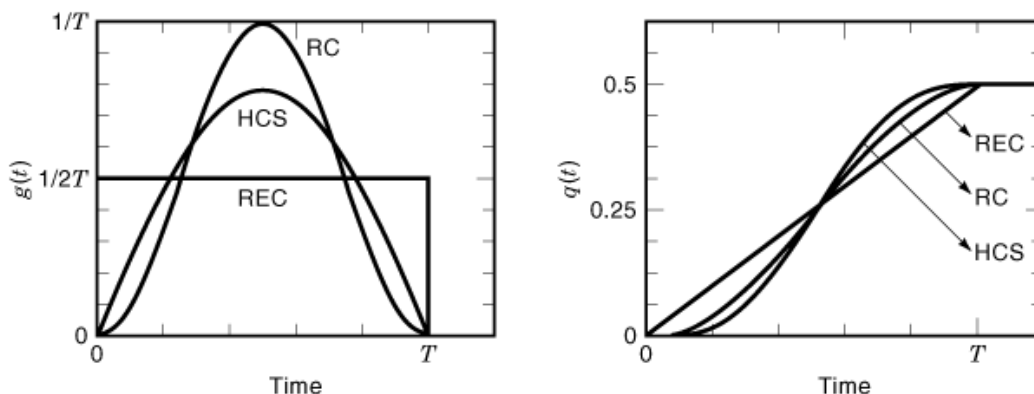
Fig. 2. Frequency and the corresponding phase pulse shapes given in Fig. 1 with $L = 1$, corresponding to the full-response case.

L representing the length of the pulses in terms of symbol interval T. For instance, an *RC* pulse extending over three symbol intervals is represented as $3RC$. Figure 1 shows the equations for different frequency pulse functions $g(t)$ with their corresponding phase functions $q(t)$. To visualize these pulses, Fig. 2 shows their corresponding shapes for the 43ull-response ($L = 1$) case.

    Figure 3 illustrates both the full- and the partial-response *RC* frequency and phase pulse shapes for comparison. Note that $g(t)$ frequency pulses are normalized such that in both cases they integrate to 0.5. This means that for a binary scheme ($M = 2$), the maximum phase change in a symbol interval will be only πh radians.

The Structure of CPM Signals

Thus far, we have given a general description of *CPM*. Now, let us gain better insight into the *CPM* signal and to the spectral performance it offers over *PSK* modulation. For this purpose, we have chosen a binary *PSK* (*BPSK*)
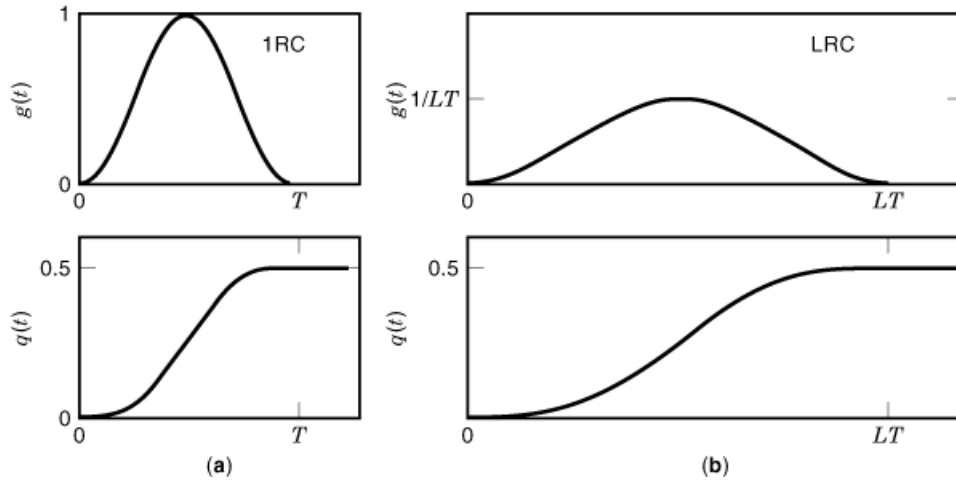
Fig. 3. (a) The full-response *RC* frequency pulse and its corresponding phase pulse; (b) the partial-response *RC* frequency pulse with its corresponding phase pulse. Note that in both cases, the phase accumulates to 0.5 at $t = LT$ defining the maximum change of phase associated with the input.
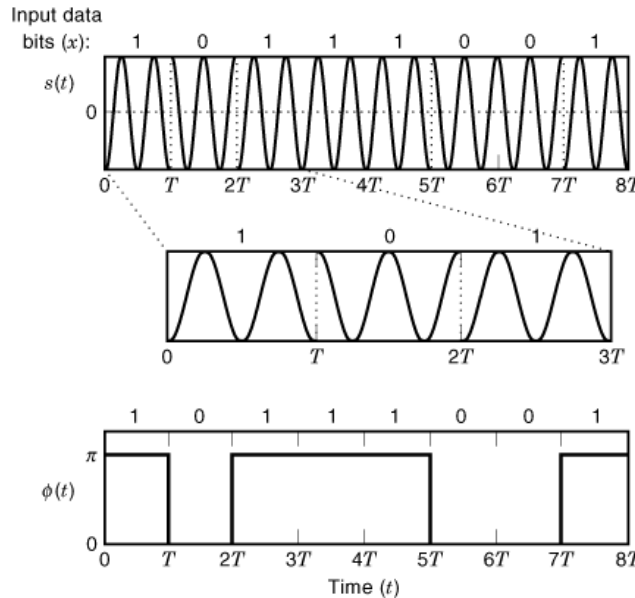
Fig. 4. The time waveform and the phase $\phi(t)$ of a pure *BPSK* signal. Abrupt changes are observed in the modulated signal phase, in which the information is carried.

signaling and a binary *1REC CPM* with modulation index $h = 0.75$ scheme to illustrate the main point behind this improvement. Figure 4 shows the transmitted *BPSK* waveform $s(t)$ and its information-carrying phase $\phi(t)$ for 8-bit intervals. The transmitted data bits x are $x = \{10111001\}$. As seen from the figure, the phase of the signal abruptly switches between 0 and π at data bit interval boundaries. The discontinuous jumps in the signal phase is the characteristic of the modulation. Now, to highlight the difference of *CPM*, let us examine
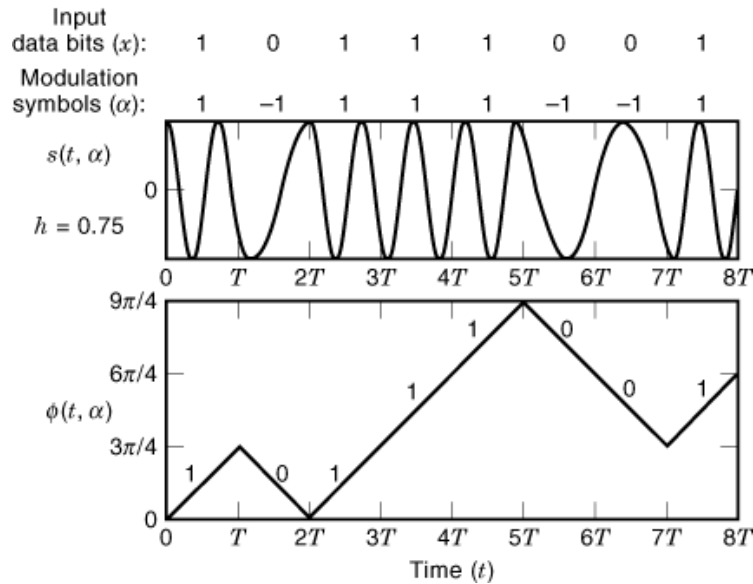
Fig. 5. The time waveform and the phase $\phi(t, \alpha)$ of a *CPM* signal. The scheme is 1*REC* and the information bearing phase changes in a continuous manner from interval to interval. This may result in a better error and spectral performance compared to the *PSK* case.

the waveforms when the same data bits are transmitted by the chosen *CPM* scheme. Figure 5 illustrates the time waveform of the transmitted binary *CPM* signal, and the information carrying phase $\phi(t, \alpha)$ corresponding to the data bits x for 8-bit intvals. The transmitted modulation symbols α corresponding to the input data bits x are $\alpha = \{1 - 1111 - 1 - 11\}$ as shown in the figure. The envelope of the transmitted signal is constant as in the *BPSK* case, but the transmitted signal phase no longer exhibits abrupt changes. As seen from the figure, the signal phase associated with the input data changes continuously, and hence yields the underlying spectral performance. We refer readers to have a close inspection of Fig. 5 to have a good understanding of Eqs. 1 2 3 4.

Let us give another example to illustrate why of the spectral performance offered by *CPM*, in which the 2*RC* ($L = 2$) frequency pulse is used to transmit the data. As is clear from the pulse shape, the *CPM* signal to transmit the data is of a partial-response type in this case. Figure 6 shows the individual smoothing phase pulses for the transmission of the same input data used in Fig. 5, and the resulting transmitted phase $\phi(t, \alpha)$ in Eq. (4) for 2*RC* scheme with $h = 0.75$ in (a) and (b), respectively. In producing $\phi(t, \alpha)$ in the figure, the input bit prior to $t = 0$ is assumed to be 1. For the same data being transmitted, Fig. 6 also illustrates the 1*REC* scheme, with dotted lines for comparison. One can see that the phase transitions in 2*RC* case is much smoother compared to that in 1*REC* case. Now, let us have a look at the power spectral density (*PSD*) of the illustrated *BPSK*, 1*REC*, and 2*RC CPM* schemes to see their individual spectral performance. Figure 7 shows the *PSD* for the three schemes. It is clear from the figure that the 2*RC* scheme has better spectral side lobes. Abrupt changes observed in the phase of the *BPSK* signal imposed larger spectral occupancy. But the continuous nature of the *CPM* schemes improves the spectral side lobes, and the smoother the phase trajectory the better the spectral performance, as observed. It is therefore clear that as much as the continuity, also the smoothness of the transmitted phase pulses plays an important role on the spectral performance, making their choice to be of crucial importance. Comparing Figs. 4, 5, 6, 7 provides good insight and helps to better understand the nature of *CPM* and the spectral improvement it offers over *PSK* modulation.

(a)

**(b)**

2RC pulse

1 REC pulse

Time (t)

$h = 3/4 = 0.75$
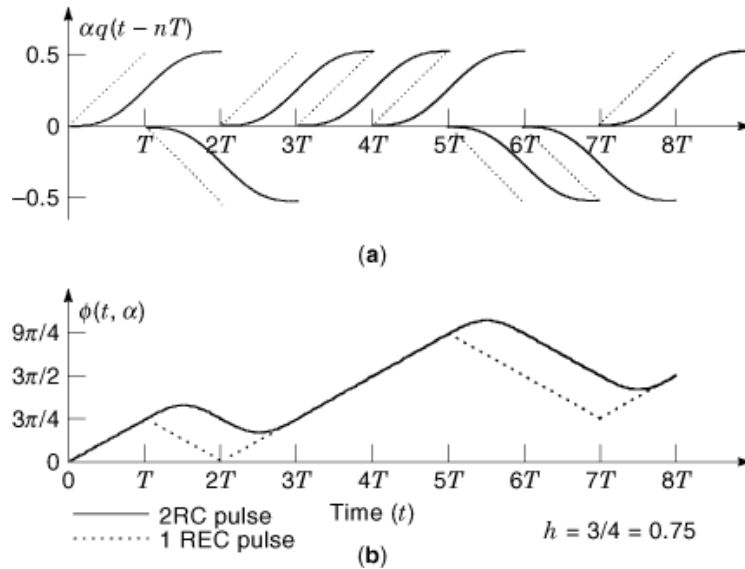
Fig. 6. (a) Illustration of the individual partial-response ($L = 2$) phase pulses corresponding to each input bit, (b) the transmitted resultant phase. Each input bit defines its own contribution to the total transmitted signal phase.
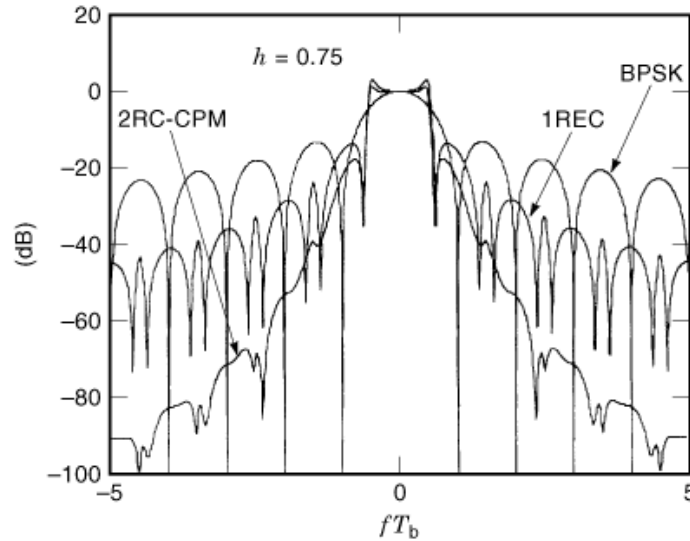

Fig. 7. Power spectral density of different schemes. Spectral improvement is clear with partial-response signaling. The 2*RC* scheme is superior to both 1*REC CPM* and the *BPSK*.

Phase Tree.   The information-carrying phase $\phi(t, \alpha)$ illustrated before, such as in Fig. 5, corresponds to a unique data stream over $8T$ intervals. It is instructive to consider all possible values of the modulation symbols α over an interval, and sketch the set of phase trajectories corresponding to these symbols commencing with a common initial phase at time $t = 0$. This is a useful way of visualizing *CPM*, which results in a so-called phase
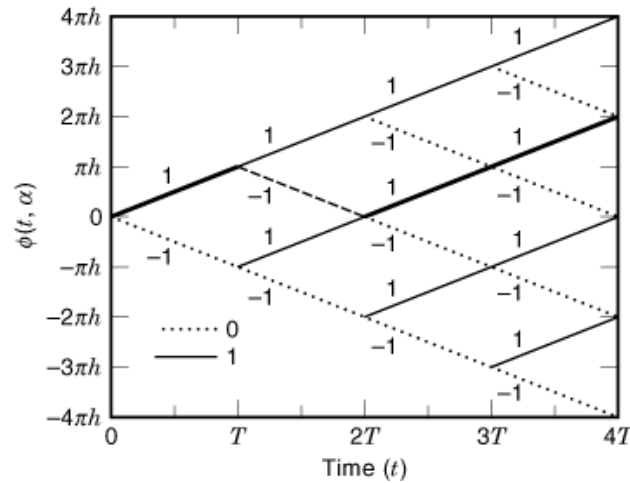
Fig. 8. Phase tree for binary ($M = 2$) 1*REC CPM* scheme. The darker lines correspond to the phase transitions corresponding to the input sequence 1 0 1 1. Signal phase is assumed to be 0 at $t = 0$. Pture shows all the possible phase transitions over *4T* signal intervals. First merger event happens at $t = 2T$.
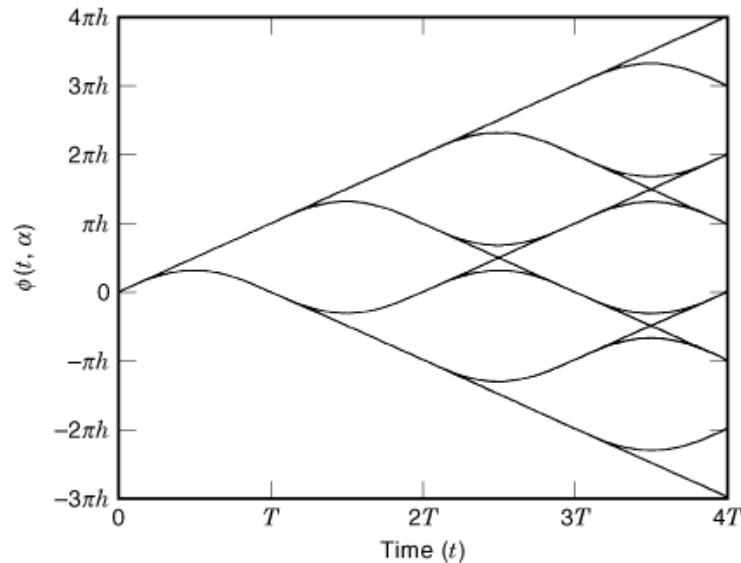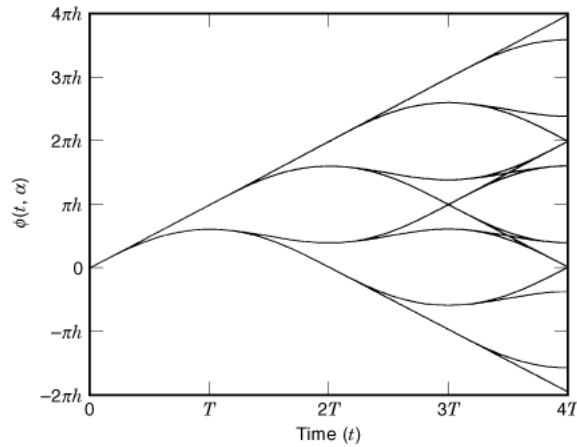

Fig. 9. Phase tree for binary ($M = 2$) 2*RC-CPM* scheme. The phase transitions are smoother compared to the 1*REC* scheme in Fig. 8, therefore offering better spectral performance. The first merger event happens at $t = 3T$. For the phase tree, the data prior to $t = 0$ is assumed to be all +1.

tree. It is an important tool in analyzing the distance properties of *CPM* signals. For example, Fig. 8 shows the set of phase trajectories for the case of binary CPFSK scheme over the interval [0, $4T$] beginning at time $t = 0$. In the figure, the dotted lines represent the phase for the input bit 0, and the solid ones show the phase for the input 1. For instance, the input data suence {1 0 1 1} traces the phase trajectory shown in the figure by darker lines. Smoother phase trees can be obtained by using pulses such as *RC* or *HCS*. For comparison, Figs. 9, 10, 11, 12 illustrate smoother phase trees, which are for 2*RC* scheme with $M = 2$, 2*RC* scheme with
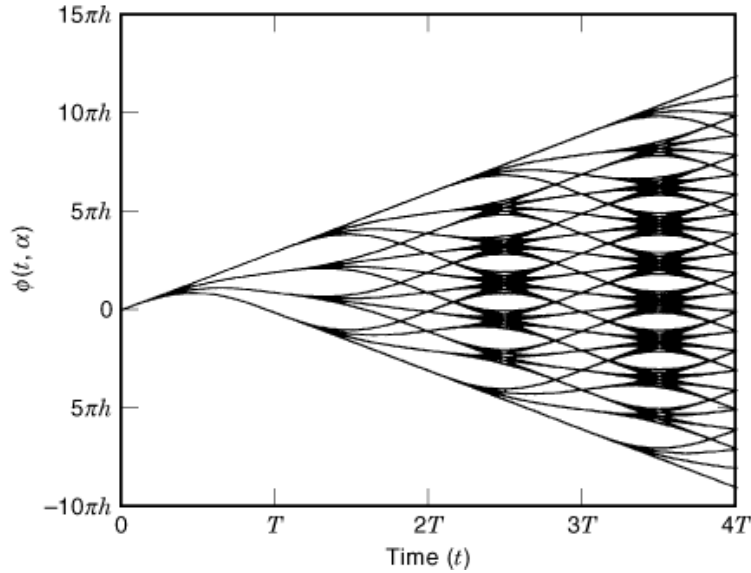
Fig. 10. Phase tree for quarternary ($M = 4$) 2*RC-CPM* scheme. The phase transitions are smoother compared to the 1*REC* scheme. The maximum change of phase in an interval increased compared to the schemes with $M = 2$ in Figs. 8 and 9. The first merger event happens at $t = 3T$ as in the case of $M = 2, L = 2$ in Fig. 9. Note that the first merger depends on only the parameter L. For the phase tree, the data prior to $t = 0$ is assumed to be all $+3$.


Fig. 11. Phase tree for binary ($M = 2$) 3*RC-CPM* scheme. The phase transitions are smoother compared to the schemes with $L = 1$ and $L = 2$ in Figs. 8 to 10, and the first merger is at $t = 4T$. For the phase tree,he data prior to $t = 0$ is assumed to be all $+1$.

$M = 4$, 3*RC* scheme with $M = 2$, and 3*RC* scheme with $M = 4$. Note that the phase axis is scaled in terms of $2\pi h$ by treating the modulation index h as a real number, so that the trees shown can be employed for any modulation index.

Note from the figures that the phase tree exhibits M-ary branching from each phase trajectory every T seconds, and therefore there are M^N distinct phase paths at the end of N symbol intervals. It is clear that
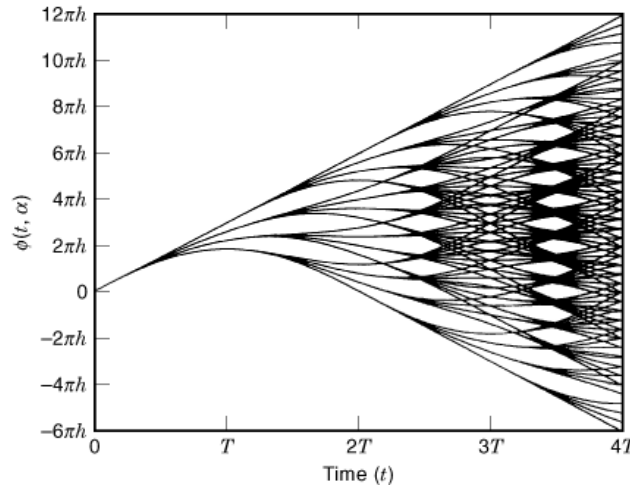
Fig. 12. Phase tree for quarternary ($M = 4$) $3RC$-CPM scheme. Phase transitions are smooth. First merger happens at $t = 4T$ as in Fig. 11, but the number of possible phase trajectories increased with time due to $M = 4$. The data prior to $t = 0$ is assumed to be all $+3$.
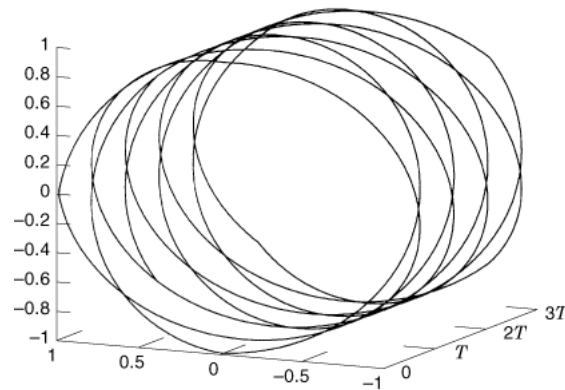

Fig. 13. Phase cylinder for $1REC$-CPM with $M = 2$. All phase paths in Fig. 8 is shown on the surface of a cylinder with $h = \frac{3}{4}$ over three symbol intervals. As the transitions that look apart in the phase tree may actually coincide, the phase cylinder shows the actual phase trajectories better.

the number of phase trajectories grows exponentially with time. In fact, the phase of the carrier is only in the range $[\,0, 2\pi\,]$ or, equivalently, in $[-\pi, \pi\,]$. This implies that the phases in the tree should be considered modulo 2π. When they are viewed modulo 2π, we notice that two phase trajectories appearing far from each other in the tree may in fact be close or even coincide and thus produce identical signals. This can be visualized more clearly by wrapping a phase tree onto a cylinder. But, to do this, the modulation index must be specified, and can be viewed as a stretching factor in wrapping onto a cylinder. Figures 13, 14, 15 show such cylinders, onto which the phase tree for the schemes $1REC$, $2RC$, and $2RC$ are wrapped with modulation indices $h = \frac{3}{4}$, $h = \frac{3}{4}$, and $h = \frac{2}{3}$, respectively.

Phase Trellis.   Since the number of branches in the phase tree corresponding to each data symbol grows exponentially, a receiver operating on the phase tree will require infinite memory. Since this is impractical,
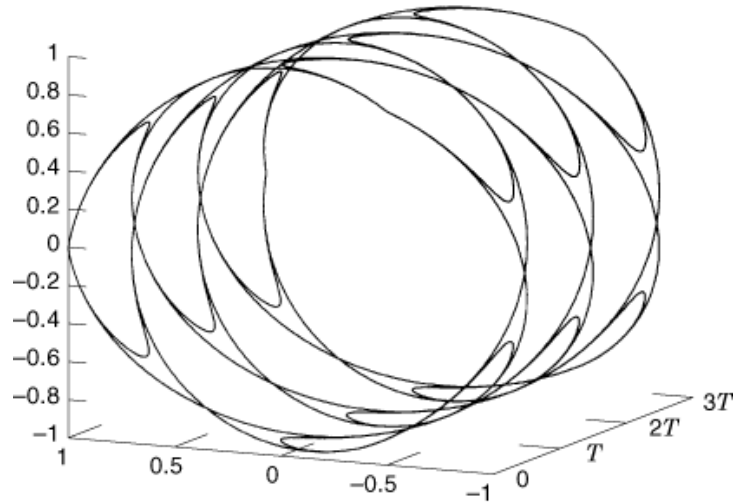
Fig. 14. Phase cylinder for 2RC-CPM scheme in Fig. 9 with $h = \frac{3}{4}$ over three symbol intervals. Compare the phase cylinder with that in Fig. 13 to see the effect of partial response on the shape of actual phase trajectories.
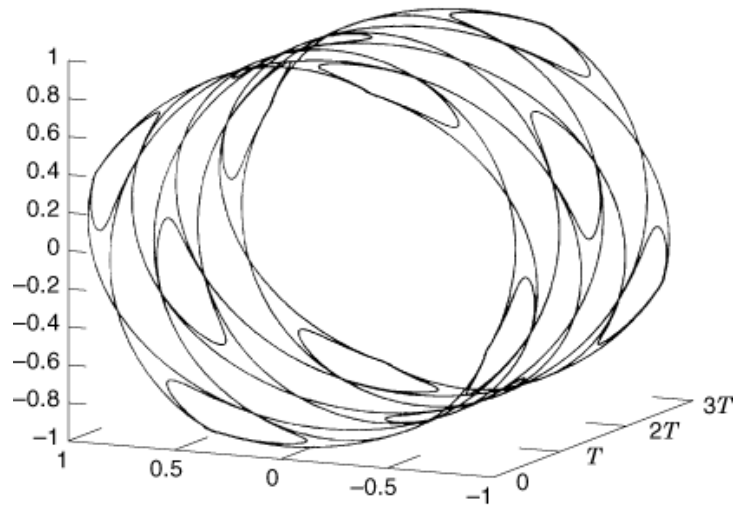
Fig. 15. Phase cylinder for 2RC-CPM with $M = 2$ and $h = \frac{2}{3}$. Compare the phase cylinder with that of Fig. 14 to see the effect of modulation index h on actual phase trajectories.

the modulation index is restricted to be a rational number as $h = p/q$ where p and q are integers that have no common factors. This constraint ensures that phase trees collapse into a finite state phase trellis, which can then be explored by the Viterbi algorithm in the form of a trellis decoder Aulin (4). To explore the state description for *CPM*, let us rewrite Eq. (4) as

$$\phi(t, \alpha) = \theta(t, \alpha_n) + \theta_n, \qquad nT \leq t \leq (n+1)T \qquad (5)$$

where

$$\theta(t, \alpha_n) = 2\pi h \sum_{i=n-L+1}^{n} \alpha_i q(t - iT) \qquad (6)$$

and

$$\theta_n = \pi h \sum_{i=-\infty}^{n-L} \alpha_i \qquad (7)$$

which is the time-varying phase at $nT \le t \le (n+1)T$ corresponding to the input symbol α_n and the accumulated phase up to time $t = nT$, respectively. It is clear that $\phi(t, \alpha_n)$ depends on the L most recent symbols, one of which is the current symbol and the remaining are the $L - 1$ previous symbols called correlative states. θ_n is called phase state, which is constant over the symbol interval. One can see that for any symbol interval, say nth interval, the phase $\phi(t, \alpha)$ is completely defined by the present data symbol α_n, the correlative state vector ($\alpha_{n-1}, \alpha_{n-2}, \ldots, \alpha_{n-L+1}$) as the $L - 1$ most recent symbols prior to the present symbol α_n, and the phase state θ_n. Thus the total state of the transmitted phase is defined as the L-tuple (5)

$$\sigma_n = (\theta_n, \alpha_{n-1}, \alpha_{n-2}, \ldots, \alpha_{n-L+1}) \qquad (8)$$

Therefore, from Eq. (8), the states of the phase trellis for full-response signals are defined in terms of only the cumulative phase states θ_n, while defined in terms of θ_n and correlative states α_n for partial-response signals. Note that for an M-ary scheme, the number of correlative states is equal to M^{L-1}. For a rational h, a full-response CPM signal at time $t = nT$ will have the phase states

$$\theta_n = \left\{ 0, \frac{\pi p}{q}, \frac{2\pi p}{q}, \ldots, \frac{(q-1)\pi p}{q} \right\} \qquad (9)$$

when p is even, and

$$\theta_n = \left\{ 0, \frac{\pi p}{q}, \frac{2\pi p}{q}, \ldots, \frac{(2q-1)\pi p}{q} \right\} \qquad (10)$$

when p is odd. It is clear from Eqs. (9) and (10) that the phase trellis have q states when p is even and $2q$ states when p is odd. For instance, if $h = \frac{1}{2}$, a full-response CPM signal will have four states as $\theta = \{0, \pi/2, \pi, 3\pi/2\}$ and for $h = \frac{2}{3}$ it will have three states as $\theta = \{0, 2\pi/3, 4\pi/3\}$. But, for a partial-response CPM, the number of distinct states increases to qM^{L-1} when p is even and $2qM^{L-1}$ when p is odd. The transmitted present modulation symbol α_n determines the transition from the state σ_n to the next state σ_{n+1}. Notice that the state complexity of a given CPM scheme depends only on h, M, and L, and not on the particular shape of the pulse shape. Figure 16 shows the state-transition diagram for a binary full-response CPM with $h = \frac{2}{3}$, in which there are 3 states (since the numerator of the modulation index is even) labeled with cumulative modulator phase states and the transitions are labelled with the present data that causes it. A finite-state trellis description is to depict all possible state transitions of a CPM modulator. Figure 17 shows the states with all the possible transitions and the phase trellis for a binary full-response scheme with $h = \frac{2}{3}$. As seen in the figure,

$$\sigma_n = \theta_n \qquad M = 2 \qquad L = 1 \qquad h = 2/3$$
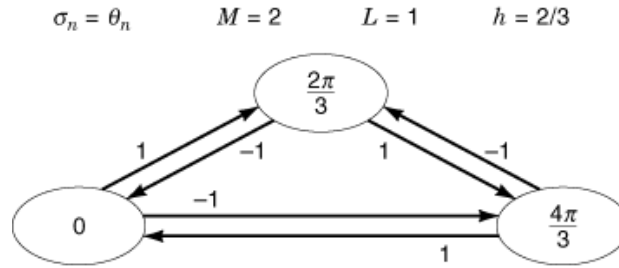
Fig. 16. State-transition diagram for a binary full-response *CPM* with $h = \frac{2}{3}$. There are only three phase states since the numerator of the modulation index is even. When the signal phase at the beginning of an interval is 0, at the end of the signal interval becomes $4\pi/3$ if the input bit is 0 as illustrated. Recall that the modulation symbol $\alpha = -1$ corresponds to the data input 0.

$$\sigma_n = (\theta_n) \qquad \sigma_{n+1} \qquad nT \qquad (n+1)T \qquad (n+2)T \qquad (n+3)T \qquad (n+4)T \qquad (n+5)T$$
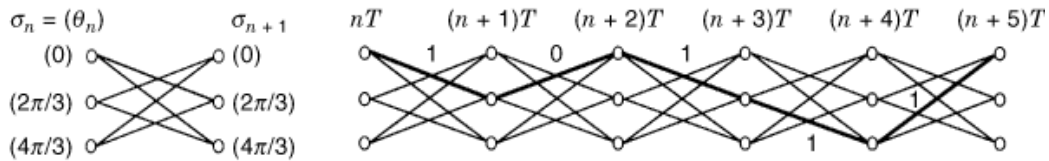
Fig. 17. The state transitions and the phase trellis for binary *1REC* scheme with $h = \frac{2}{3}$ in Fig. 16. Phase trellis is shown for the scheme, on which the influence of the input 1 0 1 1 1 on the change of phase state is traced.

$$\sigma_n(\theta_n, \alpha_{n-1}) \qquad M = 2 \qquad L = 2 \qquad h = 2/3$$
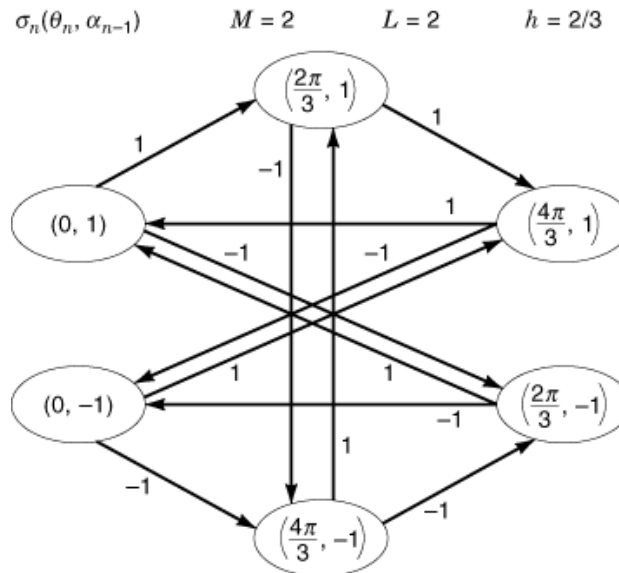
Fig. 18. State-transition diagram for the binary scheme in Fig. 17 with a change of only $L = 2$, making the scheme partial-response. Note that the number of phase states are the same, but the total number of states due to the correlative states are now 6. Transition from one state to another depends on the present phase state and the transmitted symbol in the previous interval.
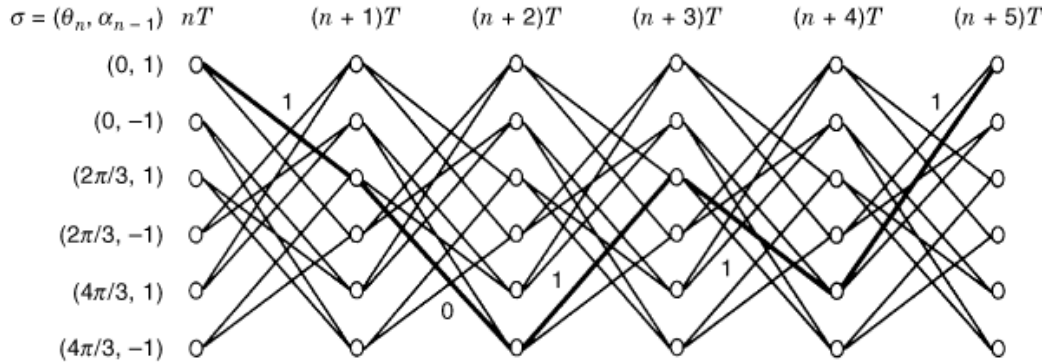
Fig. 19. The phase trellis for the binary partial-response *CPM* with $L = 2$ and $h = \frac{2}{3}$ in Fig. 18. The phase trellis has six states whereas there were only three states in the full-response case in Fig. 17. The trace shows the phase transition corresponding to the input 1 0 1 1 1.
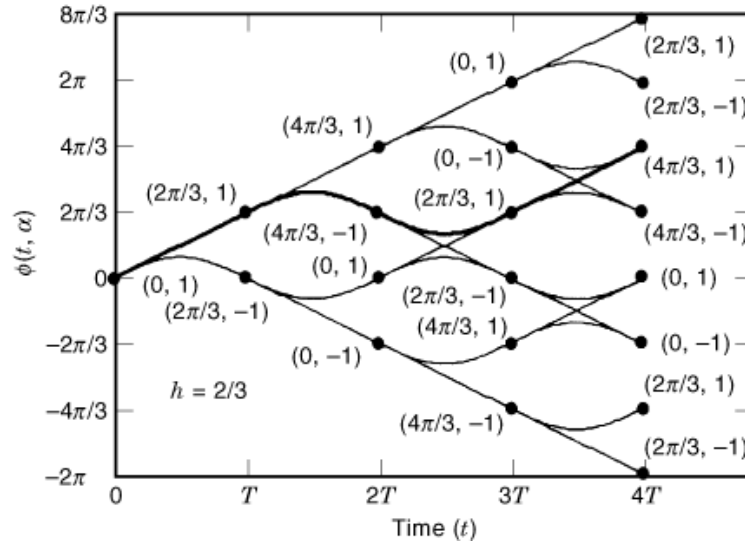

Fig. 20. Phase tree for binary 2RC scheme with $h = \frac{2}{3}$. This is for the scheme in Fig. 18 with *RC* pulse shape. The phase and the correlative states are marked on the tree. The traced trajectory corresponds to the input data 1 0 1 1.

a trellis structure called phase trellis is obtained by appending this state trellis side by side, each of which represents a symbol interval. In the figure, a trellis path is also traced that represents the transition of the transmitted phase corresponding to the input data (1 0 1 1 1). Let us consider another example for which the parameters are $L = 2$, $M = 2$, and $h = \frac{2}{3}$, and illustrate its state-transition diagram. Figure 18 shows the state-transition diagram for the considered scheme, in which the number of states is $qM^{L-1} = 3 \times 2^{2-1} = 6$. In this case, note that the states are labeled with a cumulative state and a data component (the data prior to the present one). For instance, if the modulator state is $\sigma_n = (2\pi / 3, -1)$ and the current input is a symbol $+1$, then the next cumulative state becomes 0 radians, and the next data state (correlative state vector) is $+1$. Namely, the next state vector is $\sigma_n = (0, 1)$. Figure 19 illustrates the phase trellis for a binary partial-response scheme with $L = 2$ and $h = \frac{2}{3}$. The same input data example is emphasised also in this figure, assuming the

$h = 1/4$ $h = 2/3$
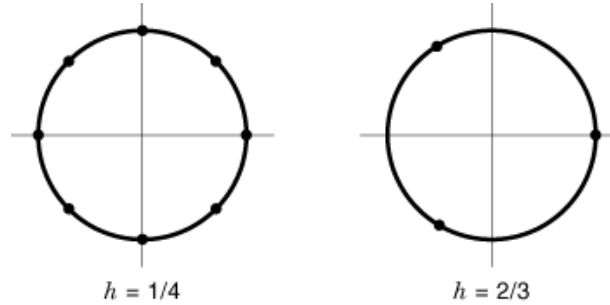
Fig. 21. Signal-space diagram for a full-response *CPM* signal with $h = \frac{1}{4}$ and $h = \frac{2}{3}$. There are only eight phase states for $h = \frac{1}{4}$ and three phase states for $h = \frac{2}{3}$. Since the scheme is of constant envelope, start and the end points of the possible phases reside on a circle.

accumulated phase $\theta_n = 0$ and $\alpha_{n-1} = 1$. Finally, to clarify the connection between phase tree, phase trellis, and the impact of modulo 2π and rational modulation index, let us redraw the phase tree for the binary 2RC example with $h = \frac{2}{3}$. Figure 20 shows the tree labeled with the states assigned according to Eq. (8). Note that in the figure, the phase trajectory corresponding to the input data in Fig. 19 is also traced. If the figure is closely inspected, one can notice that there is repetition among the states, and there are only six of those that are distinct. By writing down all the states that can be reached from each state, one ends up with the phase trellis shown in Fig. 19. In this way, the phase tree is converted into a phase trellis showing the phase transitions from one state to another associated with the input data.

Geometric Representation of CPM.   Unlike the *PSK* signals, *CPM* signals cannot be represented as discrete points in signal space, because of their continuous and time-varying phase. However, for constant-envelope *CPM* signals, only the possible starting and ending points of the phase trajectories can be marked on a circle. The number of marks will represent the number of cumulative phase states. Figure 21 shows a sample signal-space diagram for a full-response *CPM* signal with $h = \frac{1}{4}$ and $h = \frac{2}{3}$.
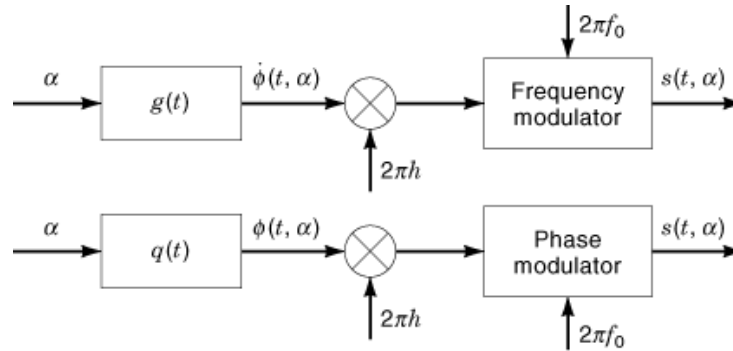
Basic CPM Transmitter

Considering Eqs. (2) and (3), it is seen that the transmitted waveform in Eq. (1) can be regarded as either a frequency-modulated signal whose frequency is shaped by $g(t)$ or a phase-modulated signal whose phase is smoothed by $q(t)$. Figure 22 illustrates this by a conceptual block diagram. For instance, this can be most easily illustrated by using a *CPM* signal of a 1REC pulse with binary symbols by writing Eq. (1) as

$$s(t) = \sqrt{\frac{2E}{T}} \cos\left(2\pi f_0 t \pm \frac{t}{2T}\right) \qquad (11)$$

with a constant carrier frequency f_0 and a time-varying phase, or as

$$s(t) = \sqrt{\frac{2E}{T}} \cos\left[\left(2\pi f_0 \pm \frac{1}{2T}\right)t\right] \qquad (12)$$

with two different frequencies deviating from f_0 and a constant phase. However, in the sequel, we consider only the phase-modulation approach. It is of interest to translate this generic transmitter structure into a practical

Fig. 22. Schematic *CPM* transmitter. *CPM* signal can be generated by using either a phase- or a frequency-modulation approach as illustrated.

solution. There are various methods to do this with their own limitations, one of which is the read-only memory (ROM) structure as the most straightforward form of implementation. The following introduces a general implementation of such a structure with some practical aspects. For this, the transmitted RF (radio-frequency) band-pass *CPM* signal in Eq. (1) can also be written by dividing it into its quadrature components as

$$s(t, \alpha) = \sqrt{\frac{2E}{T}} [I(t) \cos(2\pi f_0 t) - Q(t) \sin(2\pi f_0 t)] \qquad (13)$$

where the quadrature components are

$$I(t) = \cos[\phi(t, \alpha)]$$
$$Q(t) = \sin[\phi(t, \alpha)] \qquad (14)$$

All the possible values of these quadrature $I(t)$ and $Q(t)$ pulse shapes for the scheme over one symbol interval are then stored in I-channel and Q-channel ROMs, to be modulated by quadrature carriers of frequency f_0 as in Eq. (13) after being converted to analog. Figure 23 illustrates this general way of implementing a *CPM* transmitter. In the figure, the present data bit x_n with the latest $L - 1$ bits and the phase state address the I-channel and the Q-channel ROMs to fetch the corresponding $I(t)$ and $Q(t)$. A phase-state ROM provides the address for the phase state at the beginning of the signaling interval with a delay of one symbol interval (T), since the phase state at the beginning of the signaling interval is the phase state at the end of the previous interval. One should note here that the I-channel and the Q-channel ROMs must be clocked with an m multiple of the symbol rate (m/T), where m is the number of samples per symbol interval. In general, each signal in the ROM is identified by an address of M-ary data symbols of length L and a phase state number. For example, consider a *CPM* signal with parameters $M = 4, L = 2, h = 3$ using 8 samples/symbol representing each sample by 8 bits. Since there are $2qM^L$ signal alternatives in each signal interval, for which $h = p/q$ where p is odd as in the example, 6×4^2 signal alternatives exist. As 8 samples/symbol are used and each sample is represented by 8 bits, the total size of each of the I-channel ROM and the Q-channel ROM is then $6 \times 4^2 \times 8 \times 8 = 6144$ bits. However, to reduce the size needed for the ROMs, by using Eqs. (6) and (7), $I(t)$ and $Q(t)$ in Eq. (14) can be
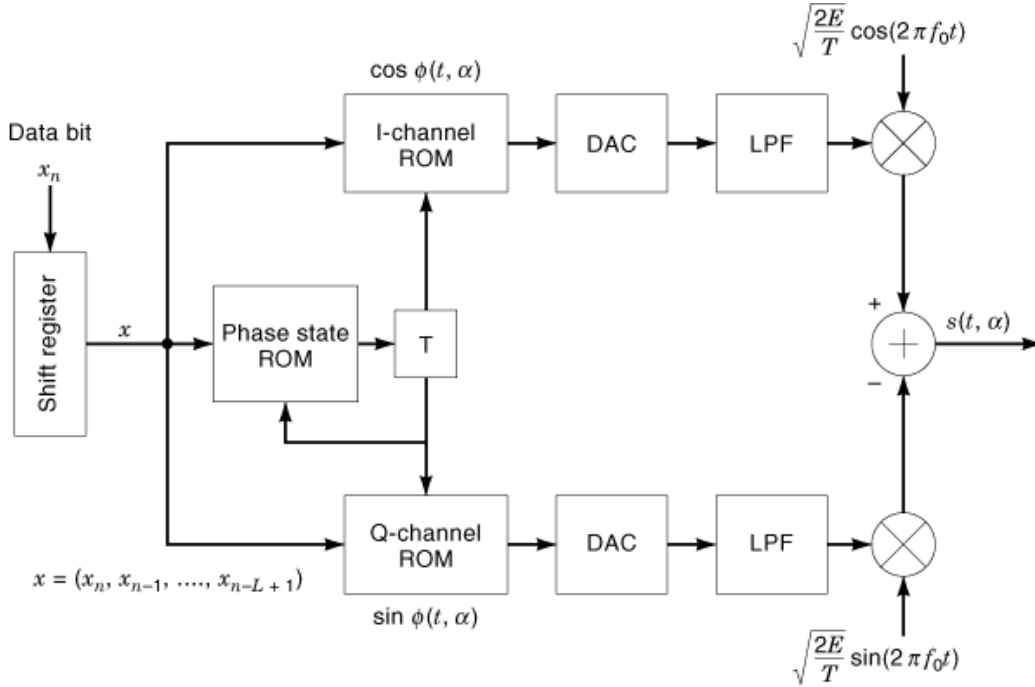
Fig. 23. Generic *CPM* ROM transmitter structure. Input bits and the phase state provided by the phase state ROM select the signal to be transmitted from the ROMs. I-channel and Q-channel signals are then swept out at a rate of m/T, where m is the number of samples/symbol, converted to analog, and anti-aliasing filtered by a low-pass filter (LPF) to modulate the quadrature carriers.

written as

$$\begin{bmatrix} I(t) \\ Q(t) \end{bmatrix} = \begin{bmatrix} \cos\theta(t, \alpha_n) & -\sin\theta(t, \alpha_n) \\ \sin\theta(t, \alpha_n) & \cos\theta(t, \alpha_n) \end{bmatrix} \begin{bmatrix} \cos\theta_n \\ \sin\theta_n \end{bmatrix} \qquad (15)$$

This indicates that the phase transitions starting from only the zero phase state $\theta_n = 0$ can be stored. Other pulse shapes starting from different phase states can be computed by a simple transformation as given in Eq. (15). This would consequently reduce the size of each quadrature channel ROM by the number of phase states. For the example above, a size of 1024 bits will be enough for each ROM in this case. Note that this reduction in size may be quite significant depending on the value of h, especially for h values with large denominators. But the penalty for this reduction is four multipliers and 2 small ROMs of $6 \times 8 = 48$ bits for $\cos(\theta_n)$ and $\sin(\theta_n)$ as in Eq. (15). From a practical point of view, such small ROMs are not available. But, as far as the very large scale integrated circuit (*VLSI*) implementation of the modulator is concerned, the required number and size of ROMs can be implemented on-chip. However, Anderson et al. (5) give a variety of methods for implementing *CPM* transmitters.

Performance of CPM

The continuous trajectory of the phase of the transmitted *CPM* signal, shown in Fig. 5, builds memory into the signal by introducing a certain dependency between transmitted phases, and thus improves the decoding performance and side lobes. A power gain is obtained over pure *PSK* signals with a varying degree of spectral improvement depending on the *CPM* signal parameters. The power gain is defined as $10 \log_{10}(d^2{}_{min}/d^2{}_{ref})$. Here, $d^2{}_{min}$ is the minimum Euclidean distance of the modulation scheme concerned and $d^2{}_{ref}$ is the minimum Euclidean distance for the reference pure (unfiltered) *PSK* modulation operating with the same energy per bit. For instance, the normalized $d^2{}_{min}$ for the pure *BPSK* is 2 and is 2.43 for the binary 1*REC* scheme with $h = 0.715$. Using this definition, the 1*REC* scheme obtains a 0.85 dB power gain over *BPSK*. This means that the 1*REC* scheme needs 0.85 dB less energy per bit to obtain the same error performance compared to that for the *BPSK* modulation for the same noise power. In the following, we discuss the error and the spectral performance of *CPM* signals.

     Error Performance of CPM. For a digital communication system, the Euclidean distance (*ED*) is a useful parameter that gives a good measure of system performance (symbol or bit error probability). A commonly used criterion for the system performance is the probability of symbol error, which is dominated by the quantity corresponding to the minimum Euclidean distance. Assuming a Gaussian channel characteristics with a large signal-to-noise ratio (*SNR*) E_b/N_0, Anderson et al. (5) gives the error rate performance as

$$P_e \sim Q \left(\sqrt{d^2_{min} \cdot \frac{E_b}{N_0}} \right) \qquad (16)$$

where

$$Q(u) = \frac{1}{\sqrt{2\pi}} \int_u^\infty e^{-x^2/2} \, dx \qquad (17)$$

In Eq. (16), $d^2{}_{min}$ is referred to as normalized squared minimum *ED* in the signal space, E_b denotes the bit energy, and N_0 represents the single-sided *PSD* of white Gaussian noise. To develop an *ED* expression for a *CPM* signal, let us suppose two signals $s(t, \alpha)$ and $s(t, \beta)$ with their corresponding phase trajectories $\phi(t, \alpha)$ and $\phi(t, \beta)$ of length N symbol intervals. The squared *ED* between these two signals, d^2, is defined as

$$\begin{aligned}
d^2 &= \int_0^{NT} [s(t, \alpha) - s(t, \beta)]^2 \, dt \\
&= \int_0^{NT} s^2(t, \alpha) \, dt \\
&\quad + \int_0^{NT} s^2(t, \beta) \, dt - 2 \int_0^{NT} s(t, \alpha)s(t, \beta) \, dt \\
&= 2NE - 2 \left(\frac{2E}{T} \right) \int_0^{NT} \cos[2\pi f_0 t \\
&\quad + \phi(t, \alpha)] \cos[2\pi f_0 t + \phi(t, \beta)] \, dt
\end{aligned} \qquad (18)$$

assuming $2\pi f_0\, T \gg 1$

$$d^2 = \frac{2E}{T} \int_0^{NT} \{1 - \cos[\phi(t, \alpha) - \phi(t, \beta)]\}\, dt \qquad (19)$$

It is seen from Eq. (19) that the *ED* is related to the difference between the phase trajectories and the symbol energy E . However, it is convenient and desirable to express the distance in terms of the bit energy E_b and to normalize the *ED* with respect to twice the bit energy as $d^2/2E_b$, so that different schemes with different alphabet size can be compared properly. So, we need to obtain the normalized squared *ED*, D^2, from Eq. (19). Noting that $E = E_b \log_2 (M)$ and writing the phase difference after Aulin and Sundberg (6) as $\phi(t, \gamma) = \phi(t, \alpha) - \phi(t, \beta)$ in which $\gamma = \alpha - \beta$ and $\gamma_i = \alpha_i - \beta_i$, D^2 is obtained from Eq. (19) as

$$D^2 = \log_2(M)\left(\frac{1}{T} \int_0^{NT} \{1 - \cos[\phi(t, \gamma)]\}\, dt \right) \qquad (20)$$

Since the minimum of the normalized squared Euclidean distances, d^2_{min}, is the dominant parameter on error performance, one needs to find the minimum normalized squared Euclidean distance among all the possible combinations of symbol sequences α and β of length N (for $N \to \infty$). Technically speaking,

$$d^2_{min} = \min_{\alpha, \beta} \{D^2\} \qquad (21)$$

in which α and β differs in the first symbol producing a nonzero difference symbol. It would be advantageous to know the ultimately achievable d^2_{min} for a *CPM* scheme before a system is actually designed, therefore saving computation power and time. In the following, for the error performance of *CPM* $d^2_{min} (h)$ and its upper bounds are considered for full-response and partial-response cases, respectively. We will plot the upper bounds on the distances that each system can achieve, so that they can be compared to assess the error performance of both schemes.

Full-Response Minimum Euclidean Distance Bounds.   A helpful tool to compute d^2_{min} is the phase tree. The phase tree for a binary 1*REC* scheme over the interval (0, 4*T*) is shown in Fig. 24, where the modulation index is general. In the figure, the dotted lines corresponds to the phase trajectory corresponding to input symbol 0 with $\alpha = -1$, and the solids to input symbol 1 with $\alpha = 1$. For the calculation of d^2_{min} for a length of N symbol intervals by using Eqs. (20) and (21), all possible difference sequences corresponding to the phase trajectory pairs must be considered, and the pairs of phase trajectories leaving the root node at $t = 0$ must not coincide, namely $\gamma_0 \neq 0$. The minimum of the all quantities obtained by using Eq. (21) is the desired result for an observation interval of N symbols.

Good candidates for a bound are infinitely long pairs of signals that merge as soon as possible. For the binary 1*REC* scheme, a pair of a sequences are traced in Fig. 24 such that the two phase trajectories coincide for all $t \geq 2T$ contributing zero quantity to the distance. This is called the first merger event, and determines the upper bound on d^2_{min} for full-response schemes. Later mergers may also be considered in computing the upper bound, but they will all yield larger *ED* since the *ED* is a nondecreasing function of observation intervals. For a general M-level signals, Aulin and Sundberg (3) choose a difference symbol sequence as

$$\gamma = \gamma_0, -\gamma_0, 0, 0, \ldots; \quad \gamma_0 = 2, 4, 6, \ldots, 2(M - 1) \qquad (22)$$
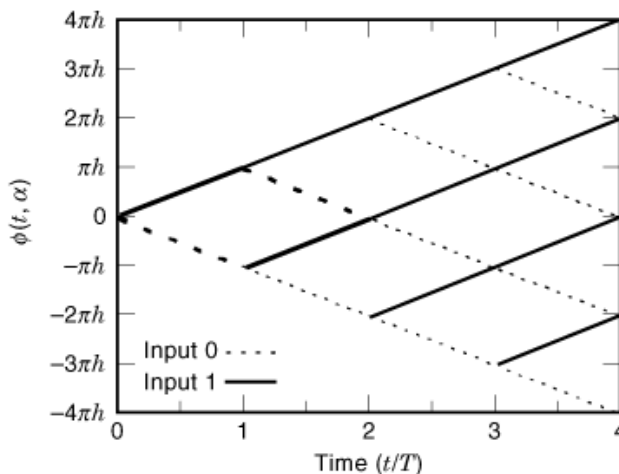
Fig. 24.   The phase tree for binary *1REC CPM* with general h. The first inevitable merger is denoted by darker lines. This merger determines the ultimately achievable minimum Euclidean distance for the scheme.

and obtain the upper bound on the normalized squared *ED*, $d^2_{UB}(h)$, for pulses having the symmetry property $g(t) = g(T - t)$, $0 \leq t \leq T$ (such as *REC*, *RC*, and *HCS* pulses), by using Eqs. (20) to (22) as

$$d^2_{UB}(h) = \log_2(M) \min_{1 \leq k \leq M-1} \left(2 - \frac{1}{T} \int_0^T \cos[4\pi hkq(t)]\, dt \right)$$

$$(23)$$

which for the *REC* pulse reduces to a simple analytic expression as

$$d^2_{UB}(h) = 2 \log_2(M) \min_{1 \leq k \leq M-1} \left(1 - \frac{\sin 2k\pi h}{2k\pi h} \right) \qquad (24)$$

As seen also from Fig. 24, first mergers happen at $t = 2T = (L + 1)T$ for all full-response schemes for any value of modulation index. Occurring independent of h, such mergers are called *inevitable mergers*. However, there are also mergers prior to the first inevitable mergers, depending on the value of some specific modulation indices. Such mergers occur if there is a phase difference that equals to a nonzero integer multiple of 2π, and the specific h value causing this merger is called *weak* modulation indices. Such modulation indices always exist. For instance, $h = 0.5$ is a weak modulation index for the four-level *1REC* scheme, producing a merger at $t = T$ prior to the first h-independent inevitable merger at $t = 2T$. At only such modulation indices, the $d^2_{min}(h)$ is always below $d^2_{UB}(h)$ (5). At a weak modulation index, if $d^2_{min}(h)$ is particularly far below the upper bound, the modulation index are called catastrophic.

One can compute the upper bounds $d^2_{UB}(h)$ for different pulse shapes by using the Eq. (23) with numerical integration. Figures 25, 26, 27 illustrate the $d^2_{UB}(h)$ for the three pulse shapes with M as parameter. As seen from the figures, no pulse shape always yields the better distance for all modulation indices. A scheme that is good for small h values ($h < 0.3$ in the figures) yields poorer for large h values. The performance obviously depends on the frequency pulse shape $g(t)$, the modulation index h, and the modulation level M. From the figures, an important result is observed, which is that the distance increases with increasing M.
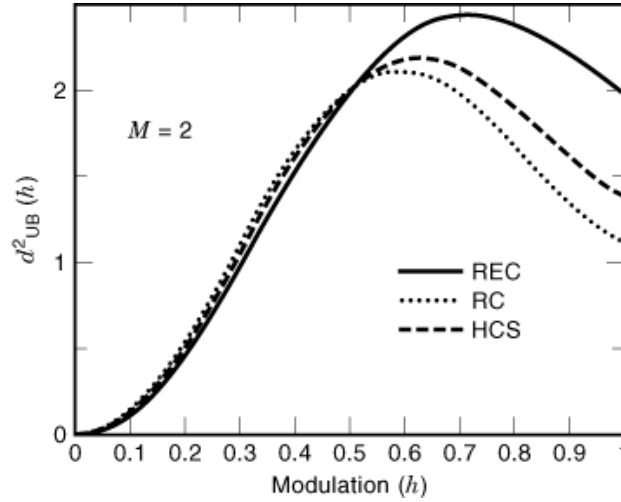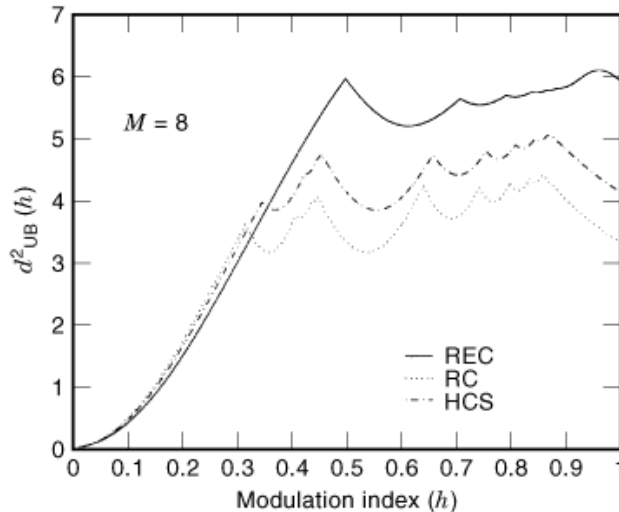
Fig. 25.   Upper bounds on $d^2_{min}(h)$ for binary 1REC, 1RC, and 1HCS systems with $M = 2$. The RC pulse performs better for $h < 0.5$, but the REC pulse improves for $h > 0.5$. The 1REC with h $= 0.5$ is known as MSK, and the d^2_{min} for MSK is 2 as seen.
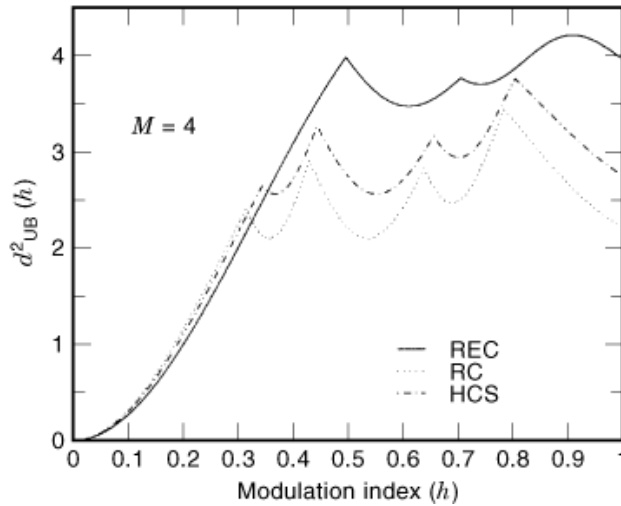
Fig. 26.   Upper bounds on $d^2_{min}(h)$ for binary 1REC, 1RC, and 1HCS systems with $M = 4$. With a slight difference, distance for the REC pulse is inferior to the others. However, for the h values greater than approximately 0.3, the REC scheme performs better. Note that the distance bounds are greater for all h values compared to that for $M = 2$ case in Fig. 25.

It is interesting to know the value of N for which the $d^2_{min}(h)$ is reached, in other words, the smallest N at which $d^2_{min,N}(h) = d^2_{min}(h)$. The growth of the $d^2_{min,N}(h)$ with N is also an important parameter for system design. A fast-growing $d^2_{min}(h)$ with N is preferred since the receiver complexity and the time to demodulate the signal increase with N. To assess the distance properties of a scheme, it would be interesting to visualize the behavior of its $d^2_{min,N}(h)$. From Eq. (20), $d^2_{min}(h)$ for a given observation interval length of N can be calculated

Fig. 27. Upper bounds on $d^2{}_{\min}$ (h) for binary $1REC$, $1RC$, and $1HCS$ systems with $M = 8$. Note that the distance bounds are greater for all h values compared to that for $M = 2$ and $M = 4$ case in Figs. 25 and 26.

by using the difference sequences γ_N through

$$d^2_{\min,N}(h) = \log_2(M) \min_{\gamma_N} \left(N - \frac{1}{T} \int_0^{NT} \cos[\phi(t, \gamma_N)] dt \right) \quad (25)$$
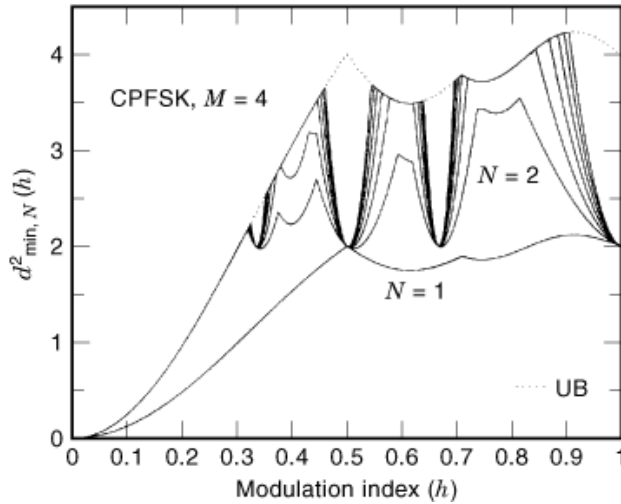
with

$$\gamma_i = \begin{cases} 0, & i < 0 \\ 2, 4, 6, \ldots, 2(M-1), & i = 0 \\ 0, \pm 2, \pm 4, \ldots, \pm 2(M-1), & i = 1, 2, \ldots, N-1 \\ 0, & i > N \end{cases} \quad (26)$$

Aulin et al. (7) gives an iterative tree algorithm to compute Eq. (25) with Eq. (26). Using the algorithm, we produce numerical results on $d^2_{\min,N}(h)$. Figures 28 to 1 illustrate these results for $1REC$ scheme as a function of N and h with parameter M, in which the maximum value of N is 7. Note from the figures that the weak modulation indices becomes clear, and the value of the distance stays the same for increasing N. Inspection of the figures also reveal that the number of weak modulation indices (at which the distance dramatically decrease) increase with increasing modulation level M.

Mulligan and Wilson (8) provide an improved algorithm to evaluate the $d^2_{\min,N}(h)$ by exploiting the trellis structure with rational h values using the well-known Viterbi algorithm. For a better and much meaningful presentation, using the algorithm, we illustrate some results on the distance profile $d^2_{\min,N}(h)$ (the growth of the minimum ED with observation interval N) for 8-level $1REC$ scheme over some h values in Table 1. As seen from the table, before a decision is made, one needs to observe the received signal for some length of symbol intervals to make sure that the minimum ED is reached for the scheme. For example, since the $d^2{}_{\min}$ $(0.35) = 3.792$ for 8-level $1REC$ CPM, one needs to observe the received signal for $N = 8$ symbol intervals to ensure the error performance of the scheme is dominated by this distance. Otherwise, if a decision is made at every fourth
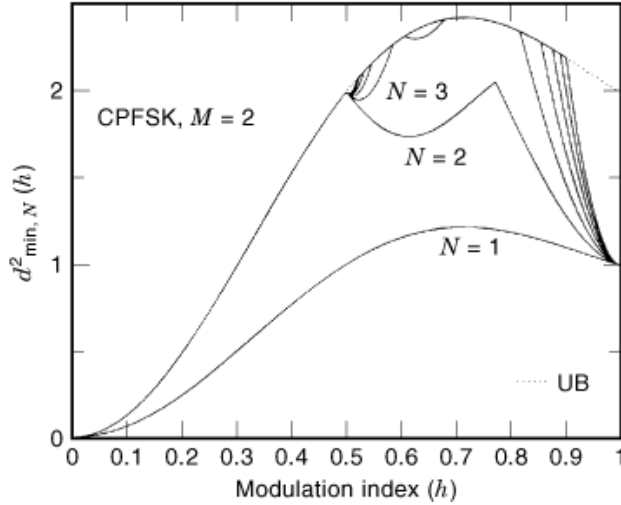
Fig. 28.   Euclidean distance profile $d^2_{\min,N}(h)$ for CPFSK signal with $M = 2$. Minimum Euclidean diance increases with the observation interval N. For $N = 1$, the distance is well below the upper bound. However, for $N \geq 2$, the distance for the scheme mostly achieves the upper bound. $h = 1$ is a week modulation index for the scheme, as the distance suddenly drops down to 1. This is because an early merger happens at $t = T$. Note the steep decline on distance near the week modulation index $h = 1$.

Fig. 29.   Euclidean distance profile $d^2_{\min,N}(h)$ for CPFSK signal with $M = 4$. For the observation interval $N \geq 2$, upper bound is achieved for most values of h except for and near weak modulation indexes. Note that the number of weak modulation index has increased compared to the scheme with the $M = 2$ case in Fig. 28.

symbol interval, the performance of the scheme will be determined by the distance 3.3. The length over which the signal must be observed to ensure the minimum *ED* for the system is called decision depth. As seen from the table that the decision depth for the signal depends on the modulation index.

Table 1. $d_{min,N}^2(h)$ Distance Profile for Eight-Level CPFSK[a]

N	h = 0.15	h = 0.35	h = 0.65	h = 0.80	h = 0.95
1	0.4248	1.8964	2.6507	2.8581	3.0583
2	0.8496	3.0063	3.2225	3.000	3.3021
3	0.8596	3.1531	3.3693	——	3.4490
4	——	3.3000	3.5162		3.5958
5		3.4468	3.6630		3.7426
6		3.5936	3.8098		3.7426
7		3.7405	3.9566		4.0363
8		3.7927	4.1035		4.1831
9		——	4.2503		4.3300
10			4.3971		4.4768
11			4.5440		4.6236
12			4.6908		4.7704
13			4.8376		4.9173
14			4.9845		5.0641
15			5.1313		5.0641
16			5.2781		5.3578
17			5.3014		5.5046
18			——		5.6514
19					5.7983
20					5.9451
21					6.0919
22					6.1165

[a] The numbers in each column represent the growth of the minimum Euclidean distance with the observation intervals. The last number in each column is the distance this scheme can achieve. For $h = 0.65$, the minimum Euclidean distance this scheme can achieve is 5.3014, which is obtained at $N = 17th$ observation interval. Increasing the observation intervals beyond $N = 17$ will not change the distance. That is, at the receiver site, we need to observe the signal at least 17 interval before making any decisions on the received data to achieve the expected performance determined by the Euclidean distance 5.3014. Observing the received signal less than 17 intervals will decrease the error performance of the scheme, depending on the length of the observation interval.

Partial-Response Minimum Euclidean Distance Bounds. The procedure for constructing an upper bound on the minimum *ED* for partial-response schemes is similar to that of full-response schemes. The phase tree is again used to identify sequences that yield early mergers at a specific time and coinciding ever after. Since the two paths coincide after the merge event, note that the contribution of such paths to the *ED* after the merger will be zero. Recall from the full-response case that the time instant after which a phase difference of two trajectories diverging at $t = 0$ can be made zero ever after is in general $t = (L + 1)T$. This time instant of first inevitable merger for the full-response case ($L = 1$) is $2T$, while it is T for a partial response scheme with $L = 3$. However, weak modulation exist as in the case of full-response schemes. One noticeable difference of partial-response schemes in constructing the upper bounds on *ED* is that the upper bound is determined by the first merger event in full-response case, while also later mergers contribute to the upper bound in partial-response case. The upper bound is found by taking the minimum of Euclidean distances associated with many mergers. Aulin and Sundberg (7) state that no mergers are needed to be considered later than the *L*th.
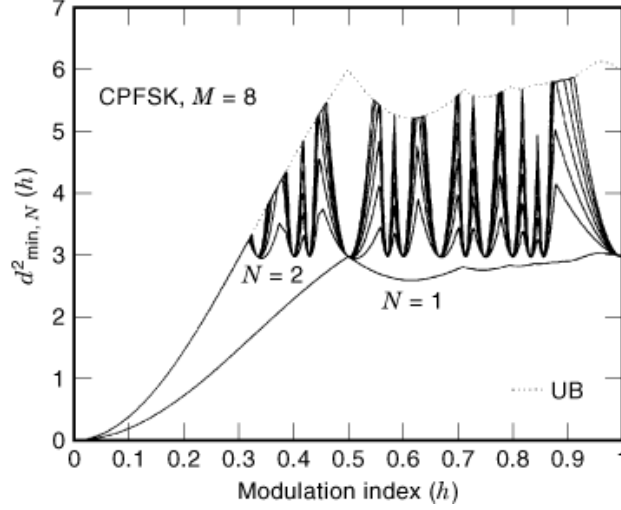
Fig. 30. Euclidean distance profile $d^2_{\min,N}(h)$ for CPFSK signal with $M = 8$. For the h values greater than about 0.3, upper bound is mostly not achieved as the number of weak modulation indexes has increased compared to that for the $M = 2$ and $M = 4$ cases in Figs. 28 to 29.

Aulin and Sundberg (7) gives the difference sequence for the mth merger as

$$\gamma_i = \begin{cases} 0, & i < 0 \\ 2, 4, 6, \ldots, 2(M-1), & i = 0 \\ 0, \pm 2, \pm 4, \ldots, \pm 2(M-1), & 0 < i < m+1 \\ 0, & i > m+1 \end{cases} \quad (27)$$

satisfying $\Sigma_{i=0}^{m} \gamma_i = 0$. To obtain the upper bound for a partial response, Eqs. (20) and (27) is used for $m = 1, \ldots, L$ with $[\, 0, (m+1)T\,]$ integration intervals. Taking the minimum of all *ED*s yields the result. Figures 31, 32, 33 illustrate the upper bounds for *LREC*, *LRC*, and *LHCS* partial-response schemes with $M = 4$ and L as parameters. One can notice from the figures that the upper bounds on the *ED* increase with increasing level M as in the case of full-response systems. From the figures, another point to be made here is that the upper bounds becomes smaller for small h values with increasing L, but get larger for larger modulation indices. However, partial-response schemes achieves generally greater *ED* compared to the full-response schemes.

Multi-h CPM. To explain the use of multi-h *CPM*, let us confine our attention to the full-response case only. Remember that the upper bounds on the *ED* for full-response schemes is determined by the first inevitable merger, which occurs at $t = 2T$ except at weak modulation index values. Since the *ED* is a nondecreasing function of observation interval N, one way of increasing the *ED* is to delay the first merger. This generally leads to an increased distance. This can be done by the use of properly chosen more than one modulation index, by cyclically changing them in successive signaling intervals, hence the name multi-h. This idea was first introduced by Miyakawa et al. (9) and generalized by Anderson and Taylor (10).
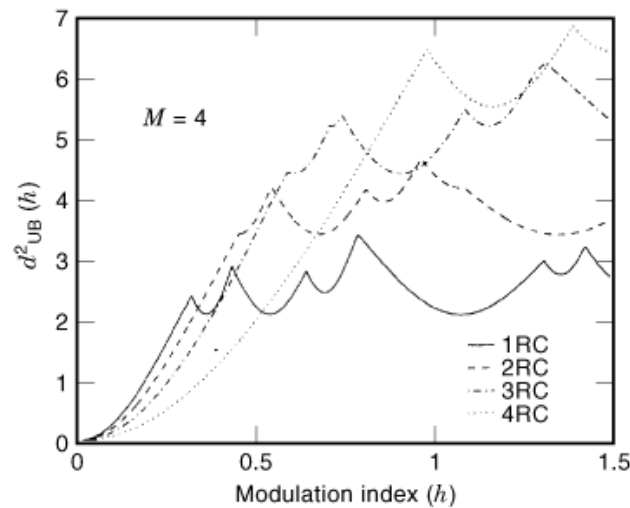
Modulation indices in multi-h schemes are chosen from a set of rational numbers with a common denominator q as

$$H_K = \{h_1, h_2, \ldots, h_K\} = \left\{ \frac{p_1}{q}, \frac{p_2}{q}, \ldots, \frac{p_K}{q} \right\} \quad (28)$$
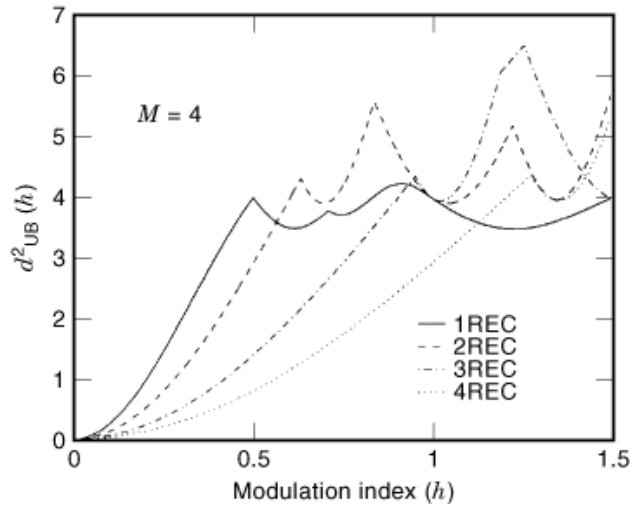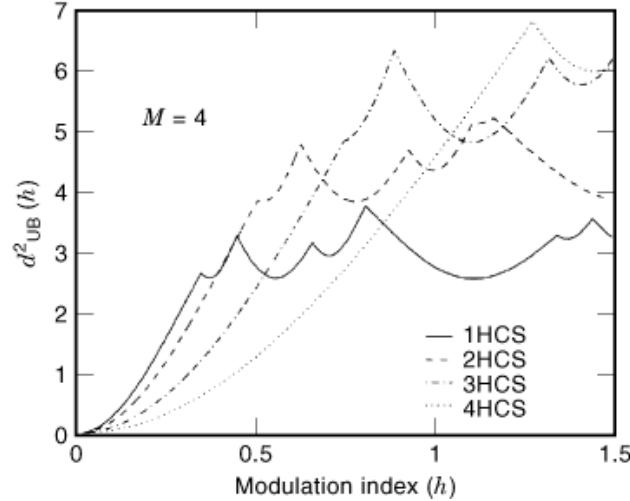
Fig. 31.   Upper bounds on $d^2_{\min}(h)$ for partial-response 2REC–4REC systems with $M = 4$. For $L = 2$, the bound is mostly above that for the full-response case. The bounds become smaller for $L > 2$ and $h < 1$.

Fig. 32.   Upper bounds on $d^2_{\min}(h)$ for partial-response 2RC–4RC systems with $M = 4$. For $L \geq 2$, the bound is above that for the full-response case. The bounds are smaller for $L > 1$ and h smaller than about 0.5. Compare the case with that in Fig. 31.

where q and p_i $(i = 1, \ldots, K)$ are small integers and h_i values are restricted to be multiples of $1/q$ so that the multi-h system collapses into a trellis structure, which can then be explored by the Viterbi algorithm. As Anderson and Taylor (10) states, multi-h schemes can also be viewed as a coding operation because an extra memory is introduced into the signal by the use of multiple modulation indices, in addition to the memory due to the continuous nature of the phase. We call K in Eq. (28), the number of modulation indices, as the constraint length of the multi-h code.

Fig. 33. Upper bounds on $d^2_{min}(h)$ for partial-response 2HCS–4HCS systems with $M = 4$. Bound for $L \geq 2$ are mostly above that for the full-response case, as similar to that in Fig. 32. Compare the case with that in Fig. 32.

To illustrate the idea of using multiple h, Fig. 34 shows the phase tree on the left for a binary 1RC scheme with $h = 2/11$, and a binary multi-h scheme for 1RC with $h = (2/11, 3/11)$ on the right. It is seen from the figure that the first inevitable merger happens at $t = 2T$ for the single-h system, but in the multi-h case on the right the first inevitable merger is delayed to $t = 3T$, introducing a possible improvement on distance. A crucially important point here in the multi-h case is that the *ED* depends on which node is taken as a root node as different from the full-response single-h systems. K different node must be considered, where K is the number of modulation indices. The minimum of the Euclidean distances obtained in each case is the *ED* for the scheme. Aulin and Sundberg (6) constructs upper bounds on the *ED* for also multi-h CPFSK signals.

The algorithm by Mulligan and Wilson (8) to find the minimum *ED* for a given scheme can also be used for multi-h schemes. By using the algorithm, Fonseka and Mao (11) give some numerical multi-h results on the minimum Euclidean distance for different pulse shapes and lengths. One can notice that the cyclic use of modulation indices in successive symbol intervals leads to a time-varying trellis, imposing an additional multi-h synchronization besides the symbol and carrier synchronization. Fonseka (12) introduces a different use of multi-h, in which the multi-h synchronization overhead is not needed. We refer readers to an intensive survey of multi-h systems by Sasase and Mori (13).

Spectral Performance of CPM. As a natural consequence of the fact that the width of the available radio frequency-spectrum is a limited source and its utilization is organized by international law, the spectral width of a signal is an important parameter in communications system designs, as important as its error performance. It is therefore of paramount importance that a designed signal occupies as narrow portion of the electromagnetic spectrum as possible. In the following, we will illustrate the spectral performance offered by *CPM* signals due to their continuous phase, and discuss the effect of signal parameters on their spectra.

Although there is no single definition of bandwidth, commonly used definition is the spectral width in which some specified portion of the total power is confined. This is called in-band power and may be specified as the 90%, 99%, or 99.9% of the total transmitted power. Wilson (14) uses the shorthand notation to represent the in-band RF power normalized with the data rate in bits/s as B_{90}, B_{99}, and $B_{99.9}$, respectively. For instance, $B_{99} = 1.2$ for *MSK* signal while it is 8.2 for OQPSK modulation (15). The *MSK* signal is generally used as a reference signal with its distance and spectral properties. Proakis (15) gives some simple expressions that are
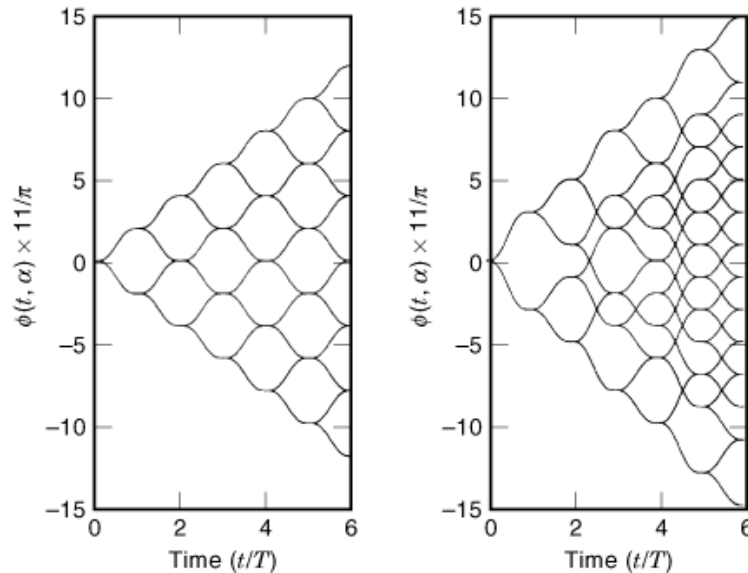
Fig. 34. Phase tree for full-response single-h 1RC scheme with $h = 2/11$ on the left, and the phase tree for full-response multi-h 1*RC* scheme with $h = (2/11, 3/11$ on the right. The modulation index for the single-h scheme is the same in all signal intervals; hence a regular phase trajectory structure as seen. However, in the multi-h scheme, different modulation index is used at consecutive intervals in a cyclic manner. Above, the modulation index pattern starts at the first interval with $h = 3/11$, and periodically changes between the two in consecutive intervals. Note that the first merger at $t = 2T$ on the left is delayed to $t = 3T$ in the multi-h case.

available for commonly used simple modulations. But, most *CPM* signals do not generally have an analytical closed-form representation, although some complicated formulas exist. However, Aulin and Sundberg (16) give an excellent method to calculate the power spectral density (*PSD*) for *CPM* signals, which is based on the autocorrelation function and the Fourier transform.

Anderson et al. (5) shows that an increase in h leads to an expansion of bandwidth for a fixed pulse shape and modulation level. At the same time, when h and pulse shape are kept the same, increasing M also results in increase of bandwidth. But, for a given modulation index and level, different pulse shapes give different spectral performance. Figure 35 shows the *PSD* of *BPSK* and *MSK* modulations, both of which have $d^2_{min} = 2$. This means that the two schemes have the same error performance. But, from the figure, the advantage of *MSK* over *BPSK* modulation is apparent with its narrower main lobe and faster decaying side lobes due to the continuous nature of its phase. Now, let us see what happens if the pulse shape is changed from rectangular to *RC* shape by imposing the same h and M . Figure 36 illustrates the *PSD* for the new pulse shape and for the *MSK* signal for comparison. As clear from the figure, the *RC* pulse shape exhibits some broadening in the main lobe, but trades this for significantly reduced side lobes. As seen, each pulse shape has its own main and side lobe spectral performance. Now, let us see the effect of increasing the pulse length L on spectral performance by holding the frequency pulse $g(t)$, modulation index h, and level M the same Figure 37 illustrates the impact of increasing L by using binary 1*RC*, 2*RC*, and 3*RC* pulses with $h = 0.5$. The improvement in the main lobe and as well as in the side lobes are clearly seen from the figure. It is therefore concluded that increasing L results in a faster roll-off factor on side lobes. Examples on selecting smoothing pulses can be found in Aulin et al. (17).

One aspect of a communication system concerns the spillage of far side of the signal spectrum into a neighbor channel, which is known as the adjacent channel interference. It is therefore of also interest to know
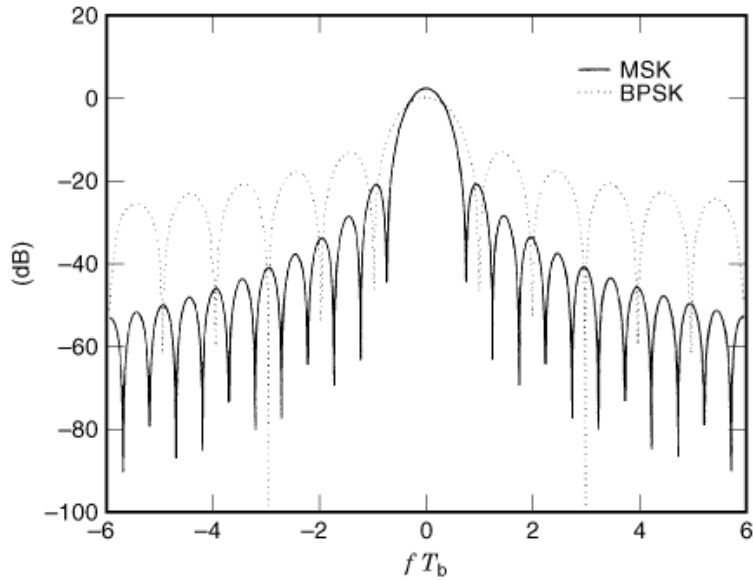
Fig. 35. Power spectral density for *MSK* and *BPSK* systems. Due to the continuity of its phase, *MSK* has lower spectral sidelobes and it drops faster than that of *BPSK*.
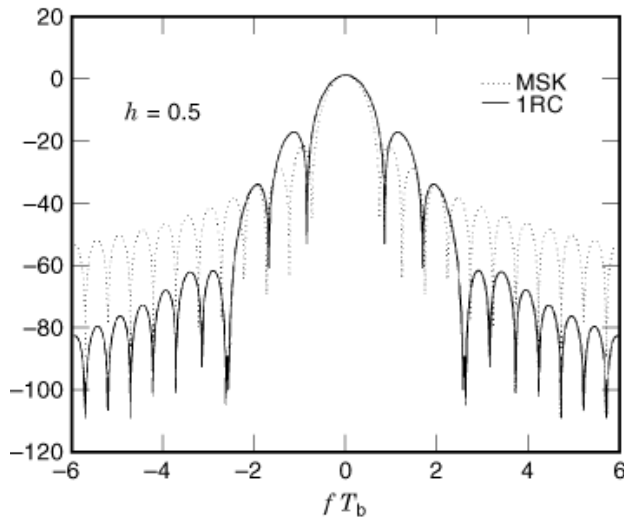
Fig. 36. Power spectral density for *MSK* and 1*RC* scheme with $h = 0.5$. Both systems have the same modulation index. The only difference is the pulse shape. *RC* scheme has slightly larger first sidelobe, but exhibits a faster roll off in the later sidelobes.

the behavior of the spectrum in the side lobes for a complete description of the signal spectrum. Aulin et al. (3) define the fractional out of band power as a useful plot to show the behavior of the energy density outside the band concerned. The out of band power $P_{\text{OB}}(B)$ is defined as

$$P_{\text{OB}}(B) = 1 - \int_{-B}^{B} S(f)df \qquad (29)$$
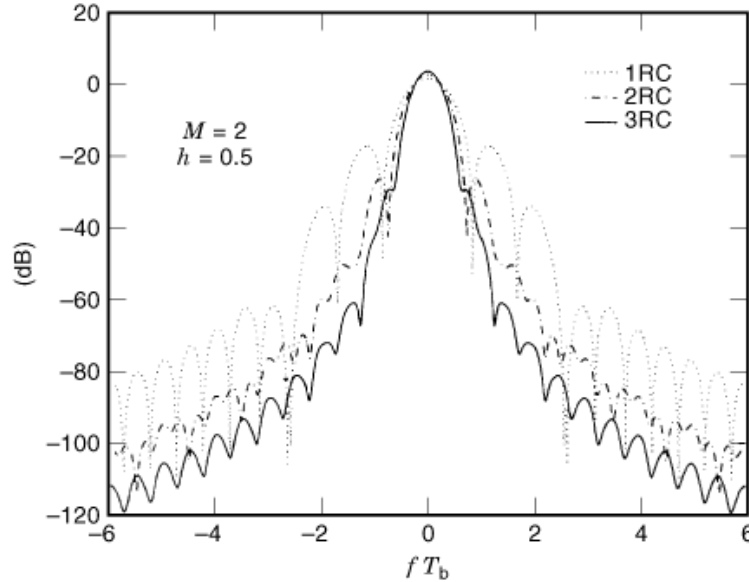
Fig. 37.   Power spectral density for LRC scheme with $h = 0.5$ and $L = 1, 2$, and 3. Note that the spectral sidelobes improve when the pulse length increase. See the superiority of the $3RC$ scheme to the $1RC$, in terms of spectral sidelobes.

in which B is the band of interest, $S(f)$ is the *PSD*, and the total power is normalized to unity. Figure 38 shows an example of out-of-band power for *MSK*, 1*REC*, and 2*REC* pulses with $M = 2$ and $h = 0.5$. In the figure, the horizontal axis shows the RF bandwidth normalized with the bit period as BT_b . Notice that BT_b = 0 corresponds to the total transmitted power, which is unity. The vertical axis shows the amount of signal power out side the $2BT_b$ normalized band relative to the total power. Note the superior performance of 3*RC* pulse over that of *MSK*, and realize (recalling Fig. 36) why the curves for *MSK* and 1*RC* schemes cross over. It is observed that partial-response signals ($L > 1$) exhibit better spectral properties. There are two special cases of partial-response signaling with compact spectrum, one of which is the well known Gaussian minimum shift keying (*GMSK*) and the other is the so called tamed frequency modulation (*TFM*). These two types of modulation have a frequency pulse of infinite duration, which are symmetrically truncated to some length and are named according to the length of their pulses. For instance, *GMSK*4 or TFM4 means that signals have frequency pulses truncated to a length of $4T$ (four-bit intervals). In the following, we complete the spectral considerations of *CPM* signals with *GMSK* and *TFM* as special types of *CPM*.

   Gaussian Minimum Shift Keying.   Murota and Hirade (18) describes that *GMSK* is a special type of partial-response *CPM* which uses a premodulation Gaussian low-pass filter on *MSK* signal with the impulse response $g(t)$ as

$$g(t) = \frac{1}{2T}\left[Q\left(2\pi B\frac{t - T/2}{\sqrt{\ln(2)}}\right) - Q\left(2\pi B\frac{t + T/2}{\sqrt{\ln(2)}}\right)\right] \quad (30)$$

where

$$Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-\tau^2/2}\, d\tau \quad (31)$$
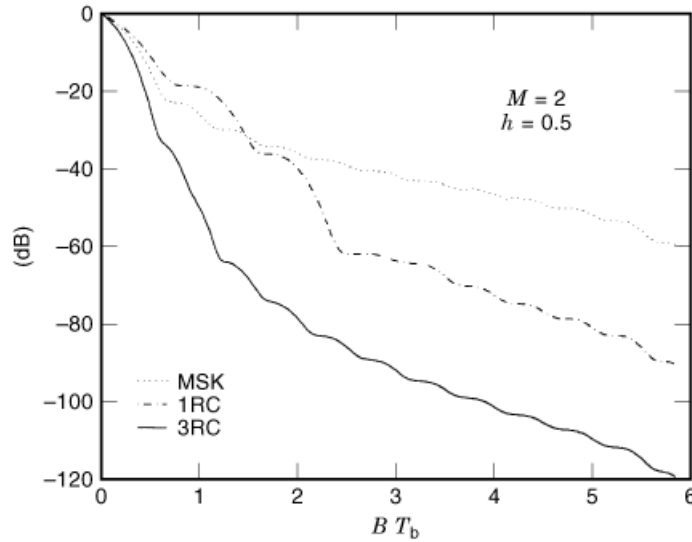
Fig. 38.   The out of powers for binary systems: *MSK*, 1*RC*, and 3*RC* with $h = 0.5$. The signal power contained outside a band of interest is a useful measure of the spectral performance of signals. A faster decaying of P_{OB} (B) indicates a faster decaying sidelobes, and hence a less adjacent channel interference. See the superiority of 3*RC* scheme in terms of the fractional out of band power.
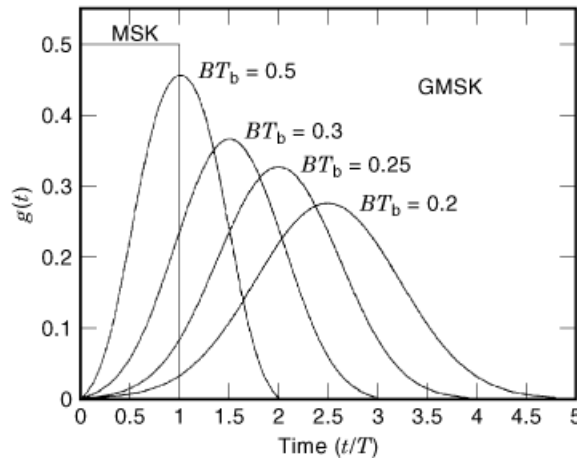
Fig. 39.   Frequency pulse shapes for *GMSK* modulation for various BT_{b} products. *MSK* has a pulse length T and corresponds to $BT_{\mathrm{b}} =$. When the BT_{b} product decreases, the length of the frequency pulse increases. For instance, the pulse with $BT_{\mathrm{b}} = 0.2$ has a length of $5T$, whereas the pulse with $BT_{\mathrm{b}} = 0.5$ has length $2T$.

The parameter B in Eq. (30) is the bandwidth of the Gaussian low-pass filter, and T is the bit period T_{b} . Rectangular frequency pulses of *MSK* signal is thus smoothed to a Gaussian shape. As seen, the impulse response $g(t)$ has infinite duration. In practical implementations, its length is truncated symmetrically L symbol intervals. The modulation is always pronounced with the BT_{b} product. A small BT_{b} results in longer impulse response and a larger one yields shorter response, therefore producing different spectral performance. The Gaussian premodulation filter prevents the instantaneous changes of frequency inherent in the *MSK* signal,
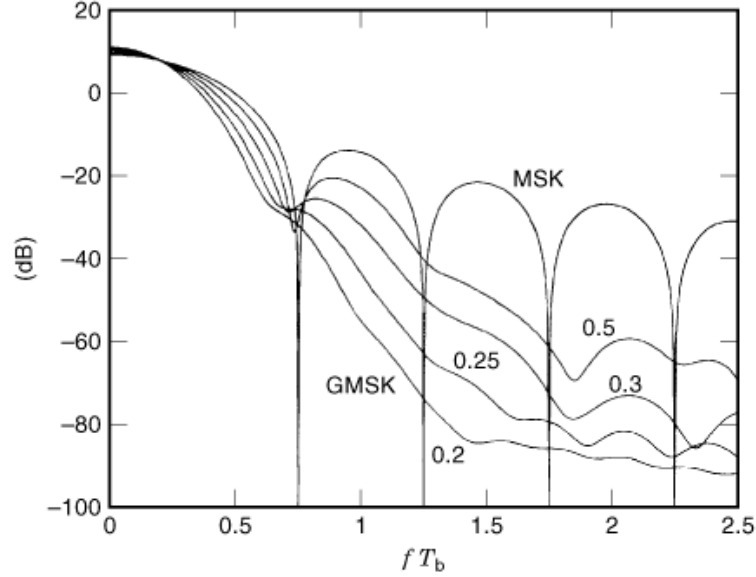
Fig. 40. Power spectral density for *MSK* and *GMSK* modulation with various BT_b products. A decrease in the BT_b product, which corresponds to a longer pulse length, improves the spectral performance.
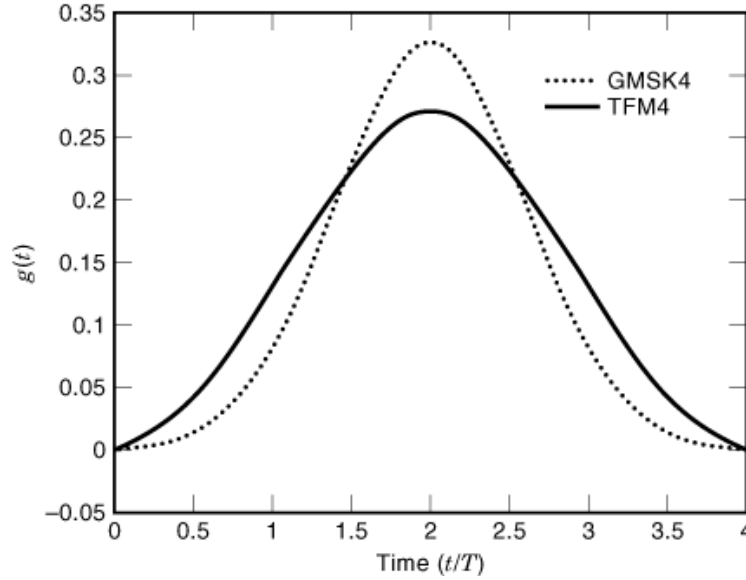
producing a power spectrum with smaller side lobes than that of *MSK*. Figure 39 illustrates the frequency pulse of *GMSK* signals in Eq. (30) for different values of BT_b products, and their spectral performances are illustrated in Fig. 40 in comparison to that of *MSK* signal. As seen from the figure, while a smaller value of the BT_b product results in a more compact *PSD*, the introduced intersymbol interference degrades the error performance. However, this leads to a tradeoff in the choice of BT_b. For instance, the global system for mobile communications (*GSM*) system uses *GMSK* modulation with the parameter $BT_b = 0.3$. As seen from Fig. 39, a *GMSK* signal with $BT_b = 0.3$ product has the pulse length 3, and the signal is also referred to as *GMSK3*.

Tamed Frequency Modulation. Jager and Dekker (19) introduces the tamed frequency modulation (*TFM*) as a special case of partial-response binary *CPM*. The frequency shaping function $g(t)$ in this case is

$$g(t) = \frac{1}{8}\left[g_0(t-T) + 2g_0(t) + g_0(t+T)\right] \qquad (32)$$

where

$$g_0(t) = \frac{1}{T}$$
$$\left(\frac{\sin\left(\frac{\pi t}{T}\right)}{\frac{\pi t}{T}}\right.$$
$$\left.-\frac{\pi^2}{24}\frac{2\sin\left(\frac{\pi t}{T}\right) - \frac{2\pi t}{T}\cos\left(\frac{\pi t}{T}\right) - \left(\frac{\pi t}{T}\right)^2\sin\left(\frac{\pi t}{T}\right)}{\left(\frac{\pi t}{T}\right)^3}\right)$$
$$(33)$$

Fig. 41.   Frequency pulse shape for *TFM* with length 4 and for *GMSK* with the same length.

Figure 41 shows the frequency pulse shape for *TFM* of length 4 (TFM4), with that of *GMSK* having the same pulse length 4 (*GMSK*4) or, equivalently, the parameter $BT_b = 0.25$ for comparison. Figure 42 illustrates the *PSD* for TFM3 and *GMSK*3, in which the *TFM* signal has narrower main but a larger side lobes. However, Chung (20) gives an extension of *TFM* called generalized tamed frequency modulation (*GTFM*), which provides flexibility in selecting frequency pulse $g(t)$.

General Remarks on CPM.

- $d^2_{min}(h)$ increases with increasing M, and so does the signal bandwidth.
- $d^2_{min}(h)$ generally increases with increasing h, and so does the signal bandwidth.
- Number of weak modulation indices increases with M, and most of them are catastrophic, and larger values of N are required to reach $d^2_{UB}(h)$.
- The first merger determines the upper bound on $d^2_{min}(h)$ for full-response systems with single-h except at weak modulation indices, but also later mergers must be taken into account for multi-h and partial-response systems.
- $d^2_{min}(h)$ is a function of the pulse shape $g(t)$, and the *REC* pulse generally gives greater distance results.
- No pulse shape performs better for the whole range of modulation indices.
- *CPM* signals possess finite state with rational modulation indices, $h = p/q$.
- K different case has to be considered to calculate the d^2_{min} for multi-h schemes having K different modulation indices.
- Better d^2_{min} is generally obtained with close h values, which requires larger q values and therefore leads to an increase in system complexity.
- Larger values of N are required to reach $d^2_{UB}(h)$ for the modulation indices close to weak ones, which leads to an increase in system complexity.
- Increasing L results in improved power spectrum for a given h, M, and $g(t)$.
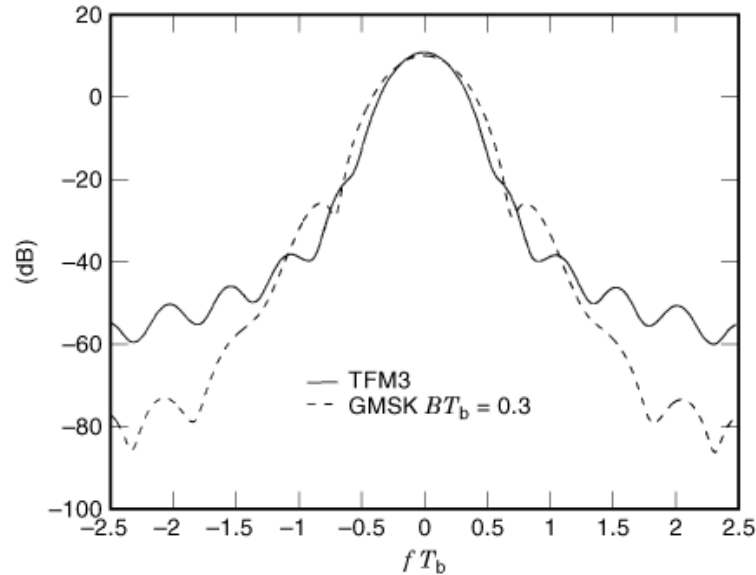
Fig. 42. Power spectral density for *TFM* and *GMSK* modulation both with $L = 3$. With this parameter, *TFM* has a slightly narrower main lobe whereas the *GMSK* has smaller sidelobes.

Topics in CPM

Rimoldi (21) introduced that any *CPM* system can be decomposed into a continuous-phase encoder as a linear time-invariant sequential circuit and a memoryless time-invariant modulator. Campanella et al. (22) presents an analytical procedure to minimize the effective bandwidth of a binary full-response *CPM* signal with respect to the shape of the frequency pulse for a given Euclidean distance. Campanella et al. (23) extends the work to the partial-response case. Asano et al. (24) calculate the optimal signal shapes for full- and partial-response signals for various receiver observation intervals, using the performance measures of effective bandwidth and minimum distance. It has been also of interest to combine *CPM* signals with an external encoder to increase the power–bandwidth performance of *CPM* schemes. In this way, extra memory has been introduced into *CPM* signals in addition to their own memory due to the continuity of their phase, improving power as well as spectral performance. Anderson and Sundberg (25) give an intensive survey of the advances in coded *CPM* modulation. Bhargava et al. (26) presents the calculation of orthogonal basis functions for a vector representation of *CPM* signals in the signal space. Ertas and Poon (27) reports multi-h codes for rate $\frac{2}{3}$ trellis codes combined with CPFSK signals for a better power–bandwidth performance. Ertas and Ali (28) construct upper bounds on the free distance for Ungerboeck-type trellis codes combined with CPFSK signals.

The detection of *CPM* signals is a complex task, and received a great deal of attention in the literature. As the detection process is outside the scope of this article, we refer readers to Sasae and Mori (13) and the references therein Anderson and Sundberg (25) and the references therein, and the Refs. 29 to 36 for the synchronization, reduced complexity, coherent and noncoherent detection of multi-h, full-response and partial-response *CPM* signals.

BIBLIOGRAPHY

1. F. Amoroso Pulse and spectrum manipulation in the minimum (frequency) shift keying (*MSK*) format, *IEEE Trans. Commun.*, **COM-24**: 381–384, 1976.
2. M. Rabzel S. Pasupathy Spectral shaping in minimum shift keying (*MSK*) type signals, *IEEE Trans. Commun.*, **COM-26**: 189–195, 1978.
3. T. Aulin C.-E. W. Sundberg Continuous phase modulation. Part I: Full response signaling, *IEEE Trans. Commun.*, **COM-29**: 196–209, 1981.
4. T. Aulin Viterbi detection of continuous phase modulated signals, *Natl. Telecommun. Conf.*, Houston, TX, pp. 14.2.1–14.2.7, 1980.
5. J. B. Anderson T. Aulin C.-E. Sundberg *Digital Phase Modulation*, 1st ed., New York: Plenum, 1986.
6. T. Aulin C.-E. Sundberg On the minimum Euclidean distance for a class of signal space codes, *IEEE Trans. Inf. Theory*, **IT-28**: 43–55, 1982.
7. T. Aulin N. Rydbeck C.-E. W. Sundberg Continuous phase modulation. Part II: Partial response signaling, *IEEE Trans. Commun.*, **COM-29**: 210–225, 1981.
8. M. G. Mulligan S. G. Wilson An improved algorithm for evaluating trellis phase codes, *IEEE Trans. Inf. Theory*, **IT-30**: 846–851, 1984.
9. H. Miyakawa H. Harashima Y. Tanaka A new digital modulation scheme: Multi-mode binary CPFSK, *3rd Int. Conf. Digital Satellite Commun.*, Kyoto, pp. 105–112, 1975.
10. J. B. Anderson D. P. Taylor A bandwidth-efficient class of signal-space codes, *IEEE Trans. Inf. Theory*, **IT-24**: 703–712, 1978.
11. J. P. Fonseka R. Mao Multi-h phase codes for continuous phase modulation, *Electron. Lett.*, **28**: 1495–1497, 1992.
12. J. P. Fonseka Nonlinear continuous phase frequency shift keying, *IEEE Trans. Commun.*, **COM-39**: 1473–1481, 1991.
13. I. Sasase S. Mori Multi-h phase-coded modulation, *IEEE Commun. Mag.*, **29** (12): 46–56, 1991.
14. S. G. Wilson R. C. Gaus Power spectra of multi-h phase codes, *IEEE Trans. Commun.*, **COM-29**: 250–256, 1981.
15. J. G. Proakis *Digital Communications*, 2nd ed., New York: McGraw-Hill, 1989.
16. T. Aulin C.-E. Sundberg An easy way to calculate power spectra for digital FM, *IEE Proc., Part F*, **130**: 519–526, 1983.
17. T. Aulin G. Lindell C.-E. Sundberg Selecting smoothing pulses for partial-response digital FM, *Proc. Inst. Electr. Eng., Part F*, **128**: 237–244, 1981.
18. K. Murota K. Hirade *GMSK* modulation for digital mobile radio telephony, *IEEE Trans. Commun.*, **COM-29**: 1044–1050, 1981.
19. F. Jager C. B. Dekker Tamed frequency modulation—A novel method to achieve spectrum economy in digital transmission, *IEEE Trans. Commun.*, **COM-26**: 534–542, 1978.
20. K.-S. Chung Generalized tamed frequency modulation and its application for mobile radio communications, *IEEE J. Select. Areas Commun.*, **JSAC-2**: 487–497, 1984.
21. B. E. Rimoldi A decomposition approach to *CPM*, *IEEE Trans. Inf. Theory*, **IT-34**: 260–270, 1988.
22. M. Campanella U. L. Faso G. Mamola Optimum bandwidth-distance performance in full response *CPM* systems, *IEEE Trans. Commun.*, **COM-36**: 1110–1118, 1988.
23. M. Campanella, *et al.* Optimum bandwidth-distance performance in partial response *CPM* systems, *IEEE Trans. Commun.*, **COM-44**: 148–151, 1996.
24. D. K. Asano H. Leib S. Pasupathy Phase smoothing functions for continuous phase modulation, *IEEE Trans. Commun.*, **COM-42**: 1040–1049, 1994.
25. J. B. Anderson C.-E. W. Sundberg Advances in constant envelope coded modulation, *IEEE Commun. Mag.*, **29** (12): 36–45, 1991.
26. V. K. Bhargava *et al. Digital Communications by Satellite: Modulation, Multiple Access, and Coding*, New York: Wiley, 1981.
27. T. Ertas F. S. F. Poon Trellis coded multi-h *CPM* for power and bandwidth efficiency, *Electron. Lett.*, **29**: 229–230, 1993.
28. T. Ertas F. H. Ali Upper bounds on the free distance for trellis coding combined with CPFSK signals, *Electron. Lett.*, **33**: 1438–1440, 1997.
29. W. P. Osborne M. B. Luntz Coherent and non-coherent detection of CPFSK, *IEEE Trans. Commun.*, **COM-22**: 1023–1036, 1974.

30. T. Aulin C. -E. Sundberg Partially coherent detection of digital full response continuous phase modulated signals, *IEEE Trans. Commun.*, **COM-30**: 1096–1117, 1982.
31. A. Stevenson T. Aulin C.-E. Sundberg A class of reduced complexity viterbi decoders for partial response continuous phase modulation, *IEEE Trans. Commun.*, **COM-32**: 1079–1087, 1984.
32. J. Huber W. L. Liu An alternative approach to reduced complexity *CPM* receivers, *IEEE J. Select. Areas Commun.*, **JSAC-7**: 1437–1449, 1989.
33. A. Stevenson Reduced state sequence detection of partial response continuous phase modulation, *Proc. Inst. Electr. Eng., Part-I*, **138**: 256–268, 1991.
34. S. Bellini G. Tartara Efficient discriminator detection of partial response continuous phase modulation, *IEEE Trans. Commun.*, **COM-33**: 883–886, 1985.
35. M. K. Simon D. Divsalar Maximum likelihood block detection of noncoherent continuous phase modulation schemes using decision feedback, *IEEE Trans. Commun.*, **COM-41**: 90–98, 1993.
36. J. Huber W. L. Liu Data aided synchronization of coherent *CPM* receivers, *IEEE Trans. Commun.*, **COM-40**: 178–189, 1992.

TUNCAY ERTAŞ
Uludag University