Exact and Approximate Construction of Digital Phase Modulations by Superposition of Amplitude Modulated Pulses (AMP)

PIERRE A. LAURENT

Abstract—**Minimum shift keying and offset QPSK are two well-known modulations which can be interpreted as a set of time/phase-shifted AM pulses.**

We show in this paper that any constant amplitude binary phase modulation can also be expressed as a sum of a finite number of time limited amplitude modulated pulses (AMP decomposition).

New methods for computing autocorrelation and power frequency spectrum are derived, which give very simple results for half-integer index modulations.

We also show that the signal can be built with good accuracy using only one optimized pulse ("main pulse"). This synthesis is particularly satisfactory for modulations that have good spectral characteristics and/or low index.

I. Introduction

It is a well-known fact that offset quadrature PSK (OQPSK) and minimum shift keying (MSK) can both be interpreted in terms of a set of time/phase-shifted AM pulses [1]–[3]. The same type of interpretation is also valid for modified versions of these modulations [4]–[6]. This type of decomposition has several advantages from a theoretical point of view (computation of autocorrelation and frequency spectrum), and from a practical one (design of modulators and demodulators).

However, no comparable interpretation is available for more complex modulations, like TFM [7], GTFM [8], or other optimized waveforms [9]–[13]. The analysis is quite different in the general case of continuous phase modulations [14], [15], and, for example, the power frequency spectrum is derived from the elementary phase function(s) only [16]–[18].

We show in this paper that any constant amplitude binary phase modulation can be expressed as a sum of a finite number of time-limited amplitude modulated pulses (AMP decomposition): this is a generalization of the interpretation of MSK. It is valid for any noninteger value of the modulation index (i.e., not restricted to quadrature modulations) and any (finite) length of the basic frequency pulse.

The components of the AMP decomposition are derived from the elementary phase variation and shifted in time and phase according to the transmitted binary information.

This interpretation has a fundamental advantage: the signal can be approximated with a fairly good degree of accuracy using only one optimized amplitude modulated waveform (named "main pulse"), exactly like OQPSK or MSK. For half-integer values of the modulation index, the main pulse is the first component of the decomposition; for other values, the main pulse is the sum of the first component and a weighted combination of the other ones.

The AMP interpretation leads to a new formulation of the power frequency spectrum and the autocorrelation of the signal. For half-integer values of the modulation index, computing these functions is particularly simple. In the general case, they can be approximated using the main pulse only, like the signal itself.

These properties reduce significantly the inherent complexity of interpretation of digital phase modulations.

The theoretical results presented here have already been used for the conception of modulators and demodulators, which thereby can be very similar to those implemented for MSK.

Paper approved by the Editor for Data Communications and Modulation of the IEEE Communications Society. Manuscript received January 7, 1985; revised June 5, 1985. This paper was presented in part at the GRETSI Conference, Nice, France, May 1983.

The author is with the Division Télécommunications, Thomson-CSF, 92231 Gennevilliers Cedex, France.

IEEE Log Number 8406879.

II. Exact AMP Representation of the Signal

A. Preliminary Definitions

Any binary controlled phase modulation can be defined by the phase shift function $\varphi(t)$ and the modulation index h. The phase shift function is assumed to be zero for negative values of time, and to have constant value for t greater than L times the bit duration T:

$$\varphi(t)=0 \qquad t\leqslant 0$$
$$\varphi(t)=h\pi \qquad t\geqslant LT. \tag{1}$$

The modulating bit stream is represented by the set $\{a_i\}$ of the signs of the phase variations associated with each of the bits.

The nth bit, a_n, appears at time $t = NT$, and causes a phase variation of $+\varphi(t)$ or $-\varphi(t)$ which adds to the phase variations associated to the previous bits. The CPM signal therefore has the following form, where θ_0 is a constant phase which will be ignored hereafter:

$$S(t)=\exp j\left[\theta_0+\sum_{n=-\infty}^{\infty} a_n\cdot\varphi(t-nT)\right]$$
$$\text{with } |a_n|=1. \tag{2}$$

This well-known representation of the signal is the starting point of all subsequent computations.

For the sake of simplicity, we will use shortened notations all through this paper. They are the following:

$$t=NT+\tau \qquad 0\leqslant\tau<T$$
$$\Phi=h\pi$$
$$C=\cos\Phi \quad S=\sin\Phi \quad J=\exp j\Phi$$
$$M=2^{L-1}. \tag{3}$$

B. Construction of the AMP Representation

Considering (1) and (2), and taking into account the fact that at time $t = NT + \tau$ the phase variation caused by the bits $N + 1, N + 2, \cdots$ is still 0, we can rewrite the signal as the following product of complex exponentials:

$$S(t) = J^{\sum_{n=-\infty}^{N-L} a_n} \times \prod_{i=0}^{L-1} \exp j[a_{N-i} \cdot \varphi(t-(N-i)T)]. \quad (4)$$

The next (and most important) step consists in replacing the complex exponential associated to the nth bit by an equivalent sum of two terms, the second one only depending upon the value of the said bit:

$$\exp j[a_n \cdot \varphi(t-nT)] = \frac{\sin [\Phi - \varphi(t-nT)]}{S} + J^{a_n} \frac{\sin [\varphi(t-nT)]}{S}. \quad (5)$$

This formulation is irrelevant for integer values of the modulation index, since, in these cases, $S = 0$. This is of no practical importance, since for "narrow-band" modulations h is generally smaller than 1.

Let us now introduce the generalized phase pulse function $\Psi(t)$, derived from the phase shift function $\varphi(t)$. It has nonzero values only for $0 \leqslant t \leqslant 2LT$; it is still more important than $\varphi(t)$ itself, and it is defined by the following relations:

$$\Psi(t) = \varphi(t) \qquad t < LT$$

$$\Psi(t) = \Phi - \varphi(t-LT) \qquad LT \leqslant t. \quad (6)$$

Its importance comes from the fact that it allows definition of the following functions, $S_n(t)$, which are the basis of the interpretation:

$$S_n(t) = \frac{\sin [\Psi(t+nT)]}{S} = S_0(t+nT). \quad (7)$$

In fact, the complex exponential associated to the nth bit, as expressed by (5), can be rewritten as follows:

$$\exp j[a_n \cdot \varphi(t-nT)] = S_{L-n}(t) + J^{a_n} S_{-n}(t). \quad (8)$$

Using this formulation in (4) leads to a new expression of the CPM signal:

$$S(t) = J^{\sum_{n=-\infty}^{N-L} a_n} \times \prod_{i=0}^{L-1} [S_{i+L-N}(t) + J^{a_{N-i}} S_{i-N}(t)]. \quad (9)$$

The right side of this equation is a sum of 2^L different terms. A detailed analysis of this sum shows that it actually consists of only $2^{(L-1)}$ distinct functions of time, named signal components. Each of them is a product of $S_0(t)$ and $L - 1$ different functions $S_i(t)$.

For example, with $N = 0$ and $L = 4$, $S(t)$ is a sum of 16 different products. By writing S_n instead of $S_n(t)$, one can find among these products $S_0 \cdot S_5 \cdot S_2 \cdot S_3$, $S_4 \cdot S_1 \cdot S_6 \cdot S_3$, and $S_4 \cdot S_5 \cdot S_2 \cdot S_7$; if these terms are rewritten $S_0 \cdot S_5 \cdot S_2 \cdot S_3$, $S_1 \cdot S_6 \cdot S_3 \cdot S_4$, and $S_2 \cdot S_7 \cdot S_4 \cdot S_5$, respectively, they appear to be the three nonzero consecutive time-shifted versions of $F(t) = S_0 \cdot S_5 \cdot S_2 \cdot S_3$, i.e., $F(t)$, $F(t + T)$, and $F(t + 2T)$.

It is then convenient to use for the Kth function of time [component $C_K(t)$] the binary representation of its index K:

$$K = \sum_{i=1}^{L-1} 2^{i-1} \cdot \alpha_{K,i} \qquad (0 \leqslant K \leqslant M-1)$$

$$\alpha_{K,i} = 0 \text{ or } 1. \quad (10)$$

This representation of K allows great simplifications in the expression of the Kth component of the signal which is nonzero only for a finite duration. This is a direct consequence of the definition of $\Psi(t)$ and $S_i(t)$ [see (6) and (7)]:

$$C_K(t) = S_0(t) \times \prod_{i=1}^{L-1} S_{i+L \cdot \alpha_{K,i}}(t) \qquad (0 \leqslant K \leqslant M-1)$$

$$0 \leqslant t \leqslant T \times \min_{i=1,2,\cdots,L-1} [L(2-\alpha_{K,i}) - i] = T \cdot L_K$$

$$C(t) = 0 \qquad \text{otherwise}. \quad (11)$$

It can then easily be shown that the M components have the following durations:

$$C_0(t) \text{-----------} (L+1)T$$

$$C_1(t) \text{-----------} (L-1)T$$

$$C_2(t),\ C_3(t) \text{-----------} (L-2)T$$

$$C_4(t),\ C_5(t),\ C_6(t),\ C_7(t) \text{-----------} (L-3)T$$

$$\cdots$$

$$C_{M/2}(t),\ \cdots\cdots,\ C_{M-1}(t) \text{-----------} T. \quad (12)$$

In addition, the above binary representation of K is used to define the value of the complex phase coefficient associated to the Kth component beginning at time $N \cdot T$:

$$\exp j\varphi_{K,N} = (\exp j\Phi)^{A_{K,N}} = J^{A_{K,N}}$$

$$A_{K,N} = \sum_{n=-\infty}^{N} a_n - \sum_{i=1}^{L-1} a_{N-i} \cdot \alpha_{K,i}$$

$$A_{K,N} = A_{0,N} - \sum_{i=1}^{L-1} a_{N-i} \cdot \alpha_{K,i}. \quad (13)$$

Finally, the CPM signal can be expressed as a sum of amplitude modulated pulses in the following general way, which is the canonical amplitude modulated pulse (AMP) formulation:

$$S(t) = \sum_{N=-\infty}^{\infty} \sum_{K=0}^{M-1} J^{A_{K,N}} C_K(t-nT). \quad (14)$$

This fundamental result shows that the interpretation of MSK can be generalized, whatever the phase pulse duration L or the modulation index h may be ($L = 1$ and $h = 0.5$ for MSK). This is very important, since any binary CPM can be considered as a superposition of rather simple AM waveforms: the "classical" representations of CPM, which are rather complex, are replaced by a simpler one, dealing only with amplitude modulated pulses, which are easy to handle.

C. Examples

1) General Case for $L = 1$: For $L = 1$, there exists only one component, $C_0(t)$. The signal is given by the following equation:

$$L = 1 \qquad (M = 1)$$

$$S(t) = \sum_{N=-\infty}^{\infty} J^{A_{0,N}} \times C_0(t-NT) \quad (15)$$

with

$$C_0(t) = S_0(t) = \frac{\sin [\Psi(t)]}{S} \qquad (0 \leqslant t \leqslant 2T).$$

For example, with MSK ($L = 1$), $C_0(t)$ is the well-known

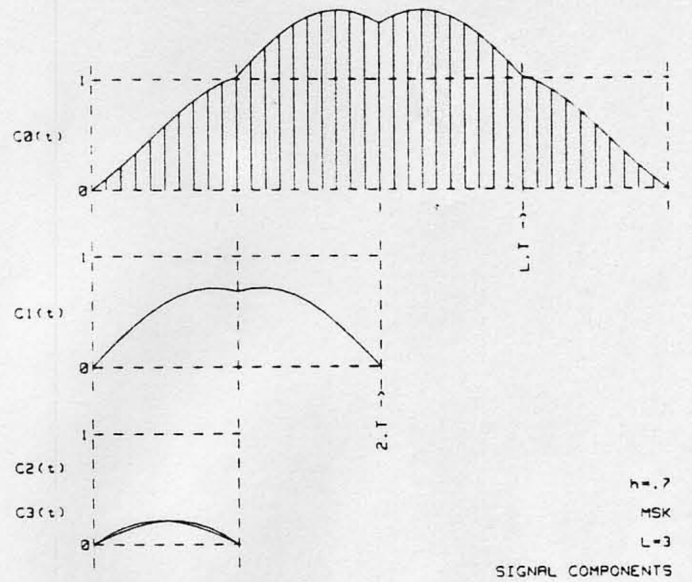

Fig. 1. Signal components of a linear FM with $L = 3$.

half-sine pulse, which can easily be shown to be equal to $S_0(t)$ over the whole duration ($2T$) of this pulse:

$$C_0(t)_{\mathrm{MSK}} = \sin\left[\frac{\pi t}{T}\right] \qquad 0 \leqslant t \leqslant 2T \ (L-1). \tag{16}$$

For offset QPSK, it can be shown that $L = 1$, $h = 0.5$, and $\varphi(t) = \pi/4$ for $0 \leqslant t \leqslant T$, so that the only component is given by

$$C_0(t)_{\mathrm{OQPSK}} = \frac{1}{\sqrt{2}} \qquad 0 \leqslant t \leqslant 2 \cdot T \ (L=1). \tag{17}$$

2) General Case for $L=3$: The number M of components of the signal, according to the AMP representation, increases exponentially with the length L of the phase variation. We take here the case where $L = 3$, for computations are still feasible by hand and the results clearly exhibit the basic properties of the AMP interpretation.

The $M = 4$ different component functions in the AMP representation are given by

$$L = 3 \quad (M=4)$$

$$K=0: \quad \alpha_{0,1}=0, \ \alpha_{0,2}=0;$$
$$C_0(t) = S_0(t) \cdot S_1(t) \cdot S_2(t) \qquad 0 \leqslant t < 4T$$
$$K=1: \quad \alpha_{1,1}=1, \ \alpha_{1,2}=0;$$
$$C_1(t) = S_0(t) \cdot S_4(t) \cdot S_2(t) \qquad 0 \leqslant t < 2T$$
$$K=2: \quad \alpha_{2,1}=0, \ \alpha_{2,2}=1;$$
$$C_2(t) = S_0(t) \cdot S_1(t) \cdot S_5(t) \qquad 0 \leqslant t < T$$
$$K=3: \quad \alpha_{3,1}=1, \ \alpha_{3,2}=1;$$
$$C_3(t) = S_0(t) \cdot S_4(t) \cdot S_5(t) \qquad 0 \leqslant t < T. \tag{18}$$

These components for a linear FM (lengthened MSK) where $\varphi(t) = \Phi t/LT$ and $h = 0.7$ are plotted in Fig. 1.

The exponents of J, which give the four phase coefficients, are the following:

$$A_{0,N} = \sum_{i=-\infty}^{N} a_i$$
$$A_{1,N} = A_{0,N} - a_{N-1} \quad A_{2,N} = A_{0,N} - a_{N-2}$$
$$A_{3,N} = A_{0,N} - a_{N-1} - a_{N-2}. \tag{19}$$

Finally, the CPM signal itself can be written:

$$S(t) = \sum_{N=-\infty}^{\infty} J^{A_{0,N}} \times [C_0(t-NT) + J^{-a_{N-1}} C_1(t-NT) + J^{-a_{N-2}} C_2(t-NT) + J^{-a_{N-1}-a_{N-2}} C_3(t-NT)]. \tag{20}$$

It is important to notice that this set of equations is still valid if the actual value of L is less than 3. For example, if $L = 2$ instead of 3, it can easily be shown that $C_2(t)$ and $C_3(t)$ have a length of 0 (because $S_5(t) = 0$), and therefore disappear. In this case, the numerical values of $C_0(t)$ and $C_1(t)$ are identical to those computed with $L = 2$, i.e., $S_0(t) \cdot S_1(t)$ and $S_0(t) \cdot S_3(t)$, respectively. Moreover, if L is not an integer number, the AMP decomposition of the signal can be obtained by replacing L with the immediately higher integer number.

This can be generalized: the AMP representation of CPM for a given value of L is also valid (although unnecessarily complicated) for lower values.

3) Generalized Tamed Frequency Modulation (GTFM): This modulation is derived from the TFM [7] and is described in detail in [8]. It has been selected here because of its apparent complexity.

The phase function $\varphi(t)$ is defined by its first derivative $G(t)$, which is the convolution product of an elementary function $H(t)$ and the impulse response of a three-tap transversal filter:

$$\frac{d\varphi(t)}{dt} = G(t) = H(t) * \left(\frac{1-B}{2} \cdot \partial(t) + B \cdot \partial(t-T) + \frac{1-B}{2} \cdot \partial(t-2T)\right). \tag{21}$$

$H(t)$ is the impulse response of a low-pass filter satisfying the third Nyquist criterion. Its frequency response can be written as

$$H(f) = \frac{\pi f T}{\sin(\pi f T)} \cdot W(f)$$

$$W(f) = \begin{cases} 1 & 0 \leqslant |f| \leqslant \dfrac{1-r}{2T} \\ \dfrac{1}{2}\left(1 - \sin\left(\dfrac{\pi}{r}\left(fT - \dfrac{1}{2}\right)\right)\right) & \dfrac{1-r}{2T} \leqslant |f| \leqslant \dfrac{1+r}{2T} \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

r is the "rolloff factor," in the range from 0 to 1.

In practice, $G(t)$, which is of infinite duration, is symmetrically truncated without noticeable consequence to the modulation performances. A value of 5 is chosen for L hereafter.

Fig. 2 represents the frequency pulse and the phase variation, with B and r set to 0.62 and 0.36, respectively.

The $M = 16$ components of the signal are shown in Fig. 3 for a modulation index $h = 0.7$, and in Fig. 4 for $h = 0.5$.

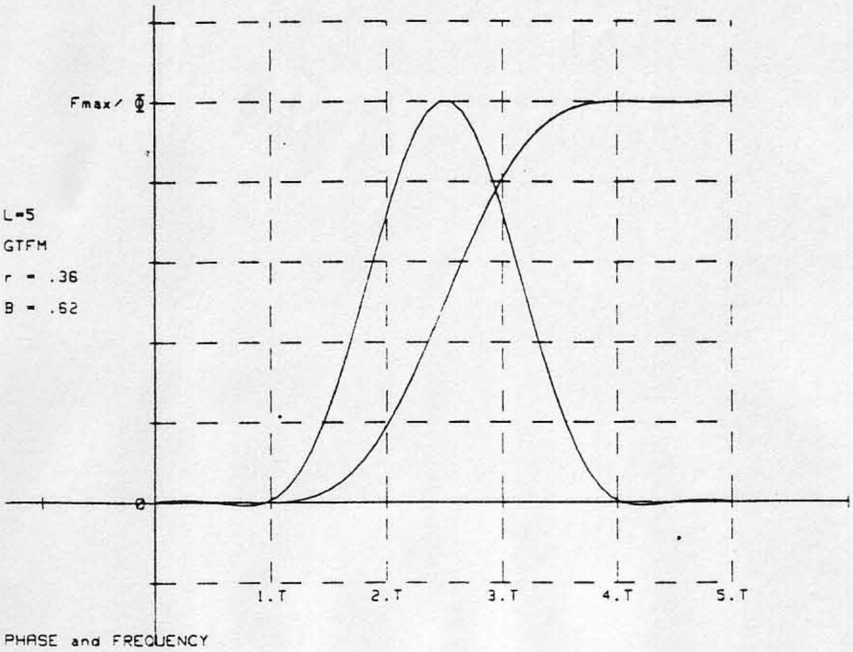

Fig. 2. Frequency pulse and phase variation for GTFM. $L = 5$, $B = 0.62$, $r = 0.36$.

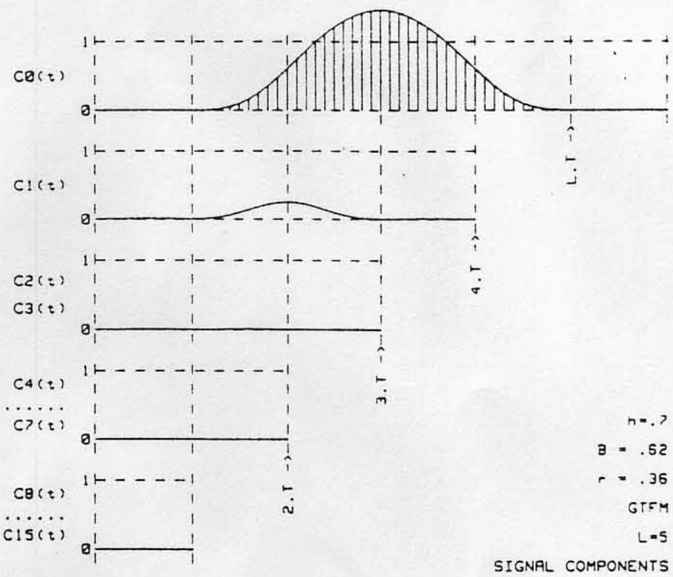

Fig. 3. Signal components of GTFM. $L = 5$, $B = 0.62$, $r = 0.36$ ($h = 0.7$).

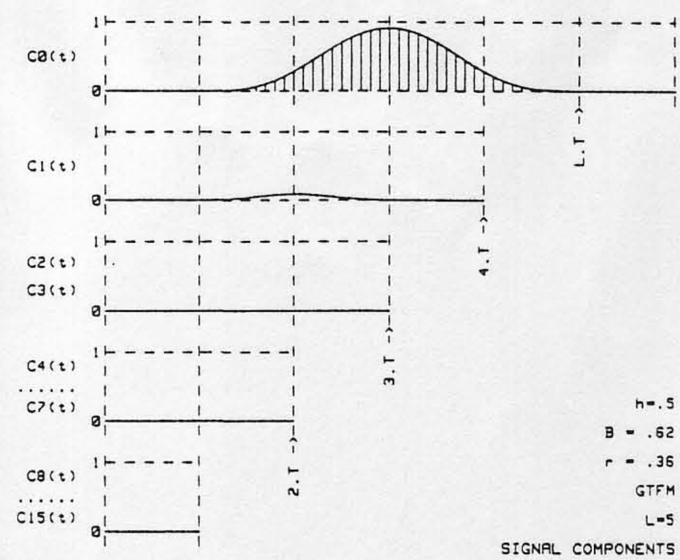

Fig. 4. Signal components of GTFM. $L = 5$, $B = 0.62$, $r = 0.36$ ($h = 0.5$).

The synthesis of GTFM is shown in Fig. 5: the upper half of the figure represents the components of the signal, with their actual phases and positions, in an in-phase/in-quadrature/time perspective; the lower half represents the final result.

A very important fact can be observed here (and probably with any other binary CPM): $C_0(t)$, which represents the pulse of longest duration (i.e., $(L + 1)T$), also happens to have the highest energy and is the most important component of the signal. This can be confirmed by comparison between a partial synthesis of the signal using only $C_0(t)$ (Fig. 6) or $C_0(t)$ and $C_1(t)$ (Fig. 7) for a modulation index of 0.5.

On the lower part of the figures, the dashed curve represents the approximate signal, and the solid line its exact value. With $C_0(t)$ alone, the synthesized signal is very close to the actual one; adding $C_1(t)$ yields a result which can hardly be distinguished from it. This seems to be a general property of CPM, and suggests a possibility of approximating AMP representation of the signal using only one pulse of length $(L + 1)T$. This approach is developed in Section IV.

III. Derivation of Autocorrelation and Spectrum of the Signal

The AMP decomposition of CPM signals into a sum of real functions of time $C_K(t)$ leads to a new formulation of the spectrum and the autocorrelation of the signal. This formulation is particularly simple for half-integer values of the

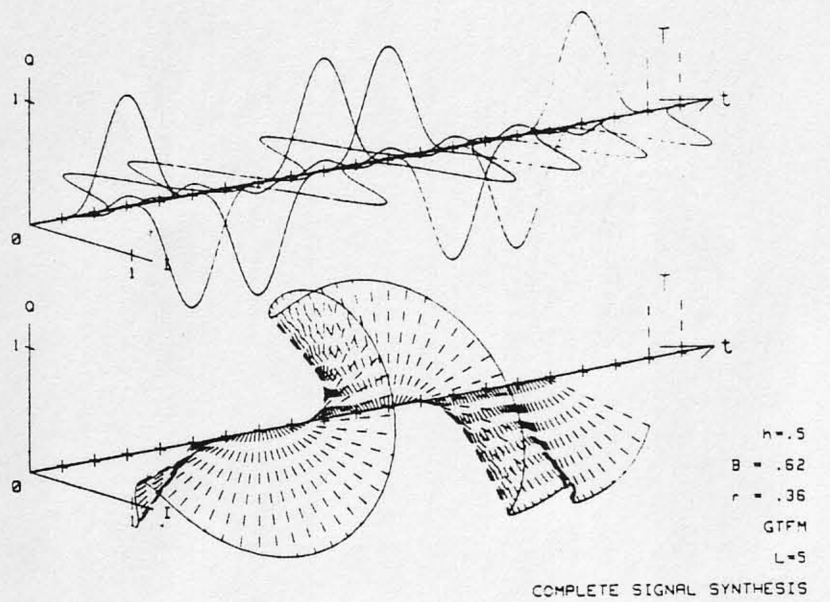

Fig. 5. Synthesis of GTFM (h = 0.5). Top: individual components. Bottom: sum of the components.

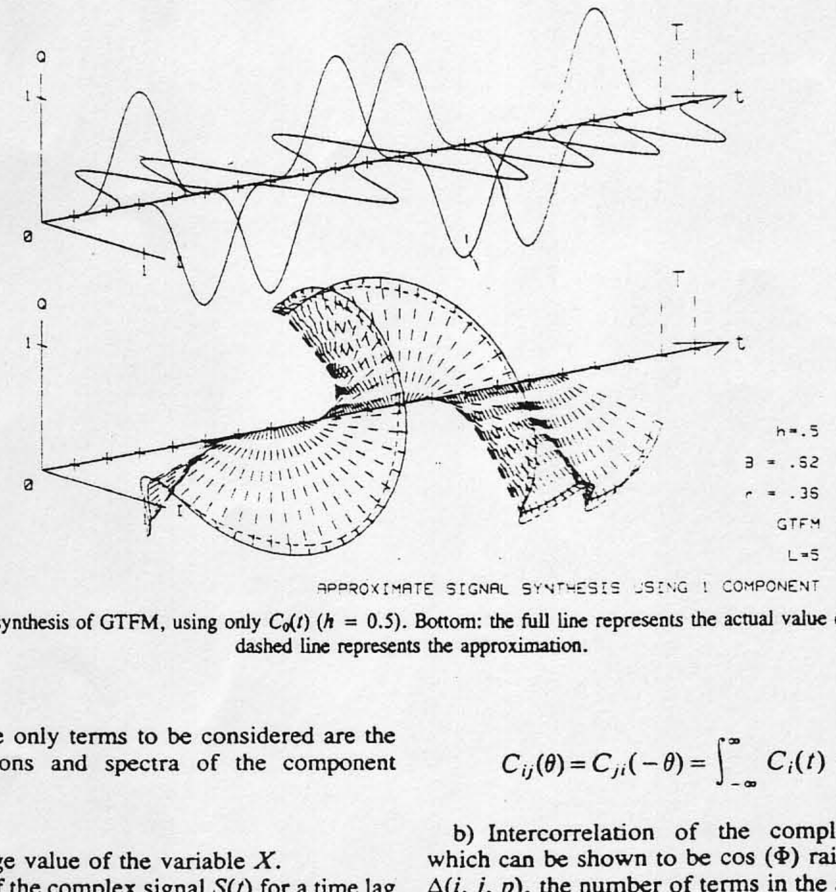

Fig. 6. Partial synthesis of GTFM, using only $C_0(t)$ (h = 0.5). Bottom: the full line represents the actual value of the signal; the dashed line represents the approximation.

modulation index h: the only terms to be considered are the individual autocorrelations and spectra of the component functions.

A. Autocorrelation

Let $\langle X \rangle$ be the average value of the variable X.

The autocorrelation of the complex signal $S(t)$ for a time lag θ is defined as

$$R(\theta)=\langle S(t) \cdot S^*(t+\theta)\rangle. \tag{23}$$

Given the AMP representation of the signal, the only quantities to be computed are the following.

a) Intercorrelation of the component functions of the signal:

$$C_{ij}(\theta)=C_{ji}(-\theta)=\int_{-\infty}^{\infty} C_i(t) \cdot C_j(t+\theta) \cdot dt. \tag{24}$$

b) Intercorrelation of the complex phase coefficients, which can be shown to be cos (Φ) raised to a power equal to $\Delta(i, j, p)$, the number of terms in the algebraic expressions of the difference $A_{i,n} - A_{j,n+p}$:

$$E_{ij}^{p}=E_{ij}^{*p}=E_{ji}^{-p}=E_{ji}^{*-p}=\langle J^{A_{i,n}} \cdot J^{-A_{j,n+p}}\rangle$$

$$E_{ij}^{p}=C^{\Delta(i,j,p)}$$

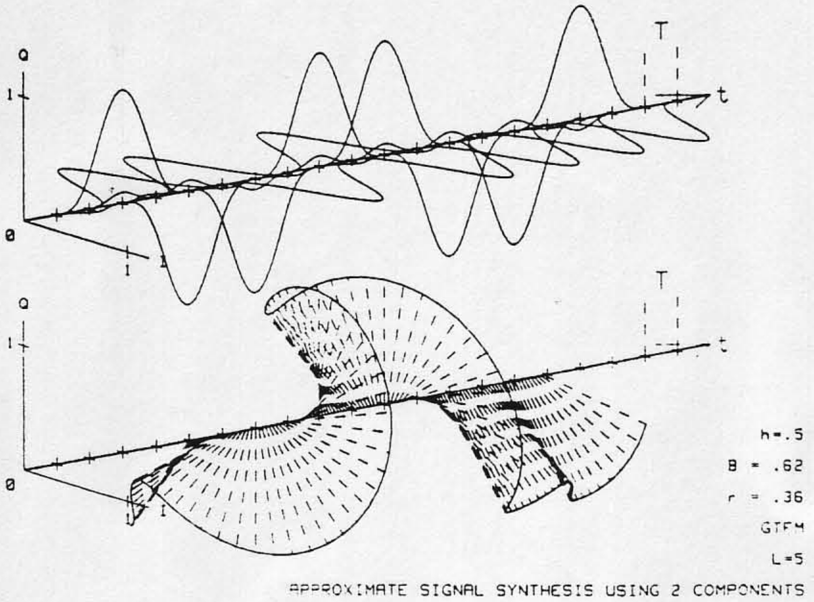

Fig. 7. Partial synthesis of GTFM, using $C_0(t)$ and $C_1(t)$ ($h = 0.5$). Bottom: same as Fig. 6.

with $\Delta(i, j, p)$ = number of terms in the algebraic expression of $A_{i,n} - A_{j,n+p}$. (25)

The computation of $\Delta(i, j, p)$ is straightforward, but somewhat tedious. We will only give here the final result, valid for any value of p:

$$\Delta(i, j, p) = |p| + \sum_{k=1}^{L-1} (\alpha_{i,k} + \alpha_{j,k})$$
$$-2\left[\sum_{\substack{k\geq 1}}^{\substack{k\leq -p-1\\ k\leq L-1}} \alpha_{i,k} + \sum_{\substack{k\geq 1}}^{\substack{k\leq p-1\\ k\leq L-1}} \alpha_{j,k} + \sum_{\substack{k\geq 1\\ k\geq 1-p}}^{\substack{k\leq L-1-p\\ k\leq L-1}} \alpha_{i,k}\cdot\alpha_{j,k+p}\right]. \quad (26)$$

If $|p|$ is greater than $L - 1$, this expression becomes

$$\Delta(i, j, p) = |p + \Delta(i, j, \infty)| \qquad (|p| \geq L) \quad (27)$$

with

$$\Delta(i, j, \infty) = \sum_{k=1}^{L-1} (\alpha_{i,k} - \alpha_{j,k}).$$

The final expression of autocorrelation, derived from the AMP representation (14) of the signal, is the following:

$$R(\theta) = R(-\theta) = \sum_{i=0}^{M-1}\sum_{j=0}^{M-1}\sum_{p=-\infty}^{\infty} E_{ij}^{p}\cdot C_{ij}(\theta - pT). \quad (28)$$

Taking into account the maximum length of the component functions as well as the simple expression of $\Delta(i, j, p)$ for $|p| \geq L$, simplifications arise for great values of $|\theta|$:

$$R(\theta) = C\cdot R(|\theta| - T) \qquad (|\theta| \geq (L+1)T). \quad (29)$$

If the modulation index h is an half-integer number, $C = \cos(\Phi) = 0$ and these equations can be simplified to

$$(h = H + 1/2,\ H: \text{integer})$$

$$R(\theta) = \sum_{i=0}^{M-1} C_{ii}(|\theta|) \qquad 0 \leq |\theta| \leq (L+1)T$$
$$R(\theta) = 0 \qquad |\theta| \geq (L+1)T. \quad (30)$$

B. Power Frequency Spectrum

The power frequency spectrum of the signal is derived from the autocorrelation. If $X(\omega)$ denotes the Fourier transform of some function of time $X(t)$, we use the well-known relation

$$|S(\omega)|^2 = R(\omega). \quad (31)$$

Given the AMP representation of the signal, the only quantities to be computed are the following.

a) The spectra of the intercorrelations of the component functions of the signal:

$$C_{ij}(\omega) = C_i^*(\omega)\cdot C_j(\omega) = C_i(-\omega)\cdot C_j(\omega). \quad (32)$$

b) The spectra of the intercorrelations of the complex phase coefficients:

$$E_{ij}(\omega) = E_{ji}^*(\omega) = E_{ji}(-\omega) = \sum_{p=-\infty}^{\infty} E_{ij}^{p}\cdot \exp j(-p\omega T). \quad (33)$$

The above spectra can be computed from (25)–(27).

Taking into account the simple expression of $\Delta(i, j, p)$ for $|p| \geq L$ allows great simplifications, so that only the terms for $|p| < L$ have to be explicitly computed. This fact is exploited by writing

$$E_{ij}(\omega) = \sum_{p=-\infty}^{\infty} \exp j(-p\omega T)\cdot C^{|p+\Delta(i,j,\infty)|}$$
$$+ \sum_{p=-L+1}^{L-1} \exp j(-p\omega T)\times[C^{\Delta(i,j,p)} - C^{|p+\Delta(i,j,\infty)|}]. \quad (34)$$

For $i = j = 0$, this expression is rather simple, since $\Delta(i, j, p) = |p|$. Otherwise, it can be shown after some computation

TABLE I
VALUES OF $W_{ij}(\omega)$ FOR $L \leqslant 3$ ($C = \cos(\Phi)$; $Z = \exp j(\omega T)$)

I	J	$W_{IJ}(\omega) = W^*_{JI}(\omega) = E_{IJ}(\omega) / E_{00}(\omega)$
0	0	1
1	0	CZ^2-C^2Z+C
2	0	$C^2Z^3-C^3Z^2+C$
3	0	$CZ^3-C^3Z+C^2$
1	1	$C^2Z^2-(C+C^3)Z+1+2C^2-(C+C^3)Z^{-1}+C^2Z^{-2}$
2	1	$C^3Z^3-(C^2+C^4)Z^2+(C+C^3)Z-C^3Z^{-1}+C^2Z^{-2}$
3	1	$C^2Z^3-C^3Z^2+C+C^3-(C^2+C^4)Z^{-1}+C^3Z^{-2}$
2	2	$C^3Z^3-C^4Z^2-CZ+1+2C^2-CZ^{-1}-C^4Z^{-2}+C^3Z^{-3}$
3	2	$C^2Z^3-(C^2+C^4)Z+C+C^3+C^4Z^{-1}-(C^3+C^5)Z^{-2}+C^4Z^{-3}$
3	3	$C^3Z^3-C^4Z^2-CZ+1+2C^2-CZ^{-1}-C^4Z^{-2}+C^3Z^{-3}$

that the result is an exact multiple of this value, $E_{00}(\omega)$. The result is as follows:

$$E_{ij}(\omega) = E_{00}(\omega) \times W_{ij}(\omega)$$

$$E_{00}(\omega) = \frac{S^2}{1+C^2-2C\cos(\omega T)}$$

$$W_{ij}(\omega) = \frac{1}{S^2} \times \Bigg(\sum_{p=-L}^{L} \exp j(-p\omega T)$$

$$\times [(1+C^2) \cdot C^{\Delta(i,j,p)} - C^{1+\Delta(i,j,p-1)} - C^{1+\Delta(i,j,p+1)}] \Bigg). \tag{35}$$

Table I gives the values of the 16 different polynomials $W_{ij}(\omega)$ for $L \leqslant 3$.

Finally, the spectrum of the signal is a direct consequence of the expression (28) of autocorrelation:

$$|S(\omega)|^2 = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} C_{ij}(\omega) E_{ij}(\omega)$$

$$|S(\omega)|^2 = E_{00}(\omega) \times \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} W_{ij}(\omega) C_i(-\omega) C_j(\omega). \tag{36}$$

Taking into account the symmetry properties of the different Fourier transforms allows simplifications of this expression: at most M spectra $C_i(\omega)$ have to be evaluated for a specific modulation; the $M(M+1)/2$ quantities $W_{ij}(\omega)$ with $j \geqslant i$ depend only upon the modulation index, and can be computed once for all.

For half-integer values of the modulation index, the general expression (36) is only the sum of the spectra of the M components of the signal, i.e.,

$$h = H + 1/2, \qquad H \text{ integer}$$

$$|S(\omega)|^2 = \sum_{i=0}^{M-1} C_{ii}(\omega) = \sum_{i=0}^{M-1} |C_i(\omega)|^2. \tag{37}$$

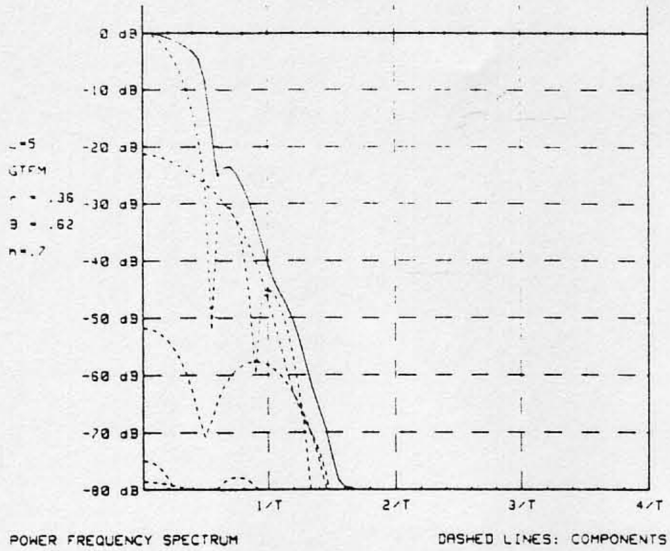

Fig. 8. Power frequency spectrum of GTFM ($h = 0.7$).

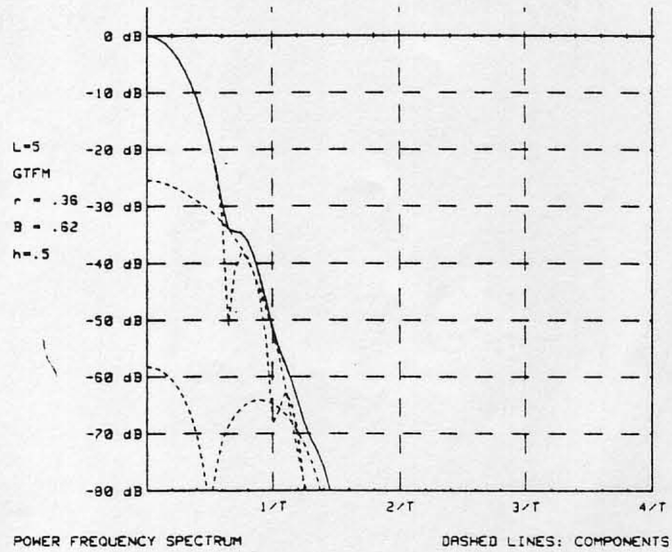

Fig. 9. Power frequency spectrum of GTFM ($h = 0.5$).

Finally, with $L = 1$, and for any value of h, the power frequency spectrum is given by

$$L = 1$$

$$|S(\omega)|^2 = E_{00}(\omega) \cdot |C_0(\omega)|^2$$

$$|S(\omega)|^2 = \frac{S^2}{1+C^2-2C\cos(\omega T)} \times |S_0(\omega)|^2. \tag{38}$$

C. Examples and Remarks

Figs. 8 and 9 represent the power frequency spectra of the component functions $C_K(t)$ and of the CPM signal $S(t)$, for GTFM as described in Section II-C-3; the values of h are 0.7 and 0.5, respectively.

A quite natural explanation of the rather chaotic appearance of the signal spectrum can be found in the fact that this spectrum is the result of the combination of M different spectra of unequal length pulses.

As could be expected, for $h = 0.5$, this spectrum is practically the sum of the spectra of the first two components,

TABLE II
VALUES OF $W_{0K}(Z)$ FOR L $\leqslant$ 4 ($C = \cos(\Phi)$; $Z = \exp j(\omega T)$)

K	$\alpha_{K,3}$	$\alpha_{K,2}$	$\alpha_{K,1}$	$W_{0K}(\omega) = E_{0K}(\omega) / E_{00}(\omega)$
0	0	0	0	1
1	0	0	1	$C \times [1 - CZ^{-1} + Z^{-2}]$
2	0	1	0	$C \times [1 - C^2Z^{-2} + CZ^{-3}]$
3	0	1	1	$C^2 \times [1 - CZ^{-1} + C^{-1}Z^{-3}]$
4	1	0	0	$C \times [1 - C^3Z^{-3} + C^2Z^{-4}]$
5	1	0	1	$C^2 \times [1 - CZ^{-1} + Z^{-2} - CZ^{-3} + Z^{-4}]$
6	1	1	0	$C^2 \times [1 - C^2Z^{-2} + Z^{-4}]$
7	1	1	1	$C^3 \times [1 - CZ^{-1} + C^{-2}Z^{-4}]$

since they provide a very close approximation of the signal (see Section II-C-3).

The general expressions of autocorrelation and spectrum of the signal are not very simple for any modulation index. In fact, they are almost as complex as in the classical formulations. However, they are very easy to compute for half-integer values of the modulation index [see (30) and (40)].

In practice, it is rarely necessary to compute all the terms of the equations: if the energy of a component is well below the required accuracy, it can be discarded without incidence on the final result. As a rule of thumb, one can consider that the energies of the components decrease rapidly with their rank, so that only the four spectra $C_K(\omega)$ ($0 \leqslant K \leqslant 3$) are to be computed in most cases.

IV. Approximate AMP Representation of the Signal

A. General Formulation

In the AMP formulation of the signal, the pulse described by the component function $C_0(t)$ is the most important among all other components $C_K(t)$ (and the only one, for $L = 1$): its duration is the longest one ($2 \cdot T$ more than any other component), and it conveys the most significant part of the energy of the signal. (We did not attempt to prove this fact mathematically, but it could be observed in every single case.) Trying to represent CPM using only one component, as in the case of MSK, is therefore quite a reasonable attempt.

Thus, the problem to solve is to find an amplitude modulated pulse $P(t)$, called "main pulse" in this paper, which is supposed to have the same phase as $C_0(t)$ and must provide by itself the best possible approximation of the signal.

We therefore suppose that the approximate signal can be represented by the following equation:

$$\hat{S}(t) = \sum_{n=-\infty}^{\infty} J^{A_{0,n}} P(t - nT). \tag{39}$$

An *a priori* good optimization criterion consists in minimizing the average energy of the difference between the complete signal and its approximation. The signal being statistically equivalent to itself at any bit position, this energy has the following value at time τ after the beginning of any bit:

$$\sigma^2(\tau) = \langle |\hat{S}(\tau) - S(\tau)|^2 \rangle \qquad (0 \leqslant \tau < T). \tag{40}$$

Without anticipating the properties of $P(t)$, especially its duration, it is convenient to use the following notation, for a given value of τ:

$$P_n(\tau) = P(\tau + nT) \qquad (0 \leqslant \tau < T). \tag{41}$$

If all the derivatives of the error energy with respect to these above quantities are zero, $P(t)$ is the "best" expression of the main pulse. So, $P(t)$ is defined by the (infinite) set of equations:

$$\frac{\partial \sigma^2(\tau)}{\partial P_m(\tau)} = 0 \qquad -\infty < m < \infty. \tag{42}$$

Using the definition (39) of the approximate signal, these equations become

$$\sum_{n=-\infty}^{\infty} P_n(\tau) \cdot C^{|m-n|} = \mathrm{Re}\ \langle J^{A_{0,-m}} \cdot S^*(t) \rangle. \tag{43}$$

Let $Rm(\tau)$ be the value of the right side of this expression. Multiplying each member of (43) by $\exp j(-m\omega T)$ and adding up for all the values of m gives

$$\sum_{m=-\infty}^{\infty} \exp j(-m\omega T) \cdot \left(\sum_{n=-\infty}^{\infty} P_n(\tau) C^{|m-n|} \right)$$

$$= \left(\sum_{k=-\infty}^{\infty} C^{|k|} \exp j(-k\omega T) \right)$$

$$\times \left(\sum_{n=-\infty}^{\infty} P_n(\tau) \exp j(-n\omega T) \right)$$

$$= \sum_{m=-\infty}^{\infty} R_m(\tau) \exp j(-m\omega T) \tag{44}$$

which can also be written [see (35)]

$$\sum_{n=-\infty}^{\infty} P_n(\tau) \exp j(-n\omega T))$$

$$= \frac{1}{E_{00}(\omega)} \sum_{m=-\infty}^{\infty} R_m(\tau) \exp j(-m\omega T). \tag{45}$$

Two methods are proposed hereafter to find the final expression of $P(t)$. The first one is rather straightforward, but cannot easily illustrate the important properties of the main pulse. The second one is based on the AMP representation of the signal and gives this information quite explicitly.

B. "Classical" Computation of the Main Pulse

Given the classical representation of the signal, the value of $Rm(\tau)$ is the following:

$$R_m(\tau) = \mathrm{Re} \left\langle \prod_{i=-\infty}^{-m} \exp j(a_i \Phi) \right.$$

$$\left. \times \prod_{i=-\infty}^{\infty} \exp j[-a_i \varphi(\tau - iT)] \right\rangle. \tag{46}$$

The bits a_i being $+1$ or -1, we can write

$$R_m(\tau) = \prod_{i=m}^{\infty} \cos\,(\Phi - \varphi(\tau + iT)) \times \prod_{i=-\infty}^{m-1} \cos\,(\varphi(\tau + iT)). \tag{47}$$

The value of $P_m(\tau)$ ($P(t)$ for $t = \tau + mT$) is the coefficient of $\exp j(-m\omega T)$ in the left side of (45), i.e.,

$$P_m(\tau) = \frac{1}{S^2}\,[(1 + C^2) \cdot R_m(\tau) - C \cdot (R_{m+1}(\tau) + R_{m-1}(\tau))]. \tag{48}$$

The final result, expressed as a function of $\varphi(t)$, is as follows:

$$\begin{aligned} P_m(\tau) = \frac{1}{S} \times & \left[\prod_{i=-\infty}^{m-2} \cos\,(\varphi(\tau + iT))\right] \\ & \times \left[\prod_{i=m+1}^{\infty} \cos\,(\Phi - \varphi(\tau + iT))\right] \\ & \times [\sin\,(\varphi(\tau + mT)) \cdot \cos\,(\varphi(\tau + (m-1)T)) \\ & - C \cdot \sin\,(\varphi(\tau + (m-1)T)) \cdot \cos\,(\Phi - \varphi(\tau + mT))]. \end{aligned} \tag{49}$$

As can be seen, computing $P(t)$ is comparatively easy: most of the terms equal 1, and only a few sines and cosines have to be evaluated. Analyzing this expression in detail shows that $P(t)$ is zero for $t < 0$ (or $m < 0$), and $t > (L + 1) \cdot T$ (or $m > L$), as expected, since $P(t)$ was supposed to replace $C_0(t)$, which has the same length.

It can be observed that the "configuration" of this expression changes from one bit to the following: this fact suggests that $P(t)$ may actually be a combination of elementary pulses appearing at each new bit. The second method of evaluation of $P(t)$ consists in building $P(t)$ using such pulses, namely the component functions $C_K(t)$.

C. AMP-Based Evaluation of the Main Pulse

The computation steps are the same as in the previous method. The difference lies in the representation of $S(t)$, which is now given by its AMP equation (14).

Thus, (47) is replaced by

$$R_m(\tau) = \mathrm{Re}\left(\sum_{n=-\infty}^{\infty}\sum_{K=0}^{M-1} J^{A_{0,-m} - A_{K,n}} C_K(\tau - nT)\right). \tag{50}$$

Taking into account the definition (25) of the coefficients used for computing autocorrelation leads to the expression

$$R_m(\tau) = \sum_{n=-\infty}^{\infty}\sum_{K=0}^{M-1} E_{0,K}^{m-n}\, C_K(\tau + nT). \tag{51}$$

Replacing $R_m(\tau)$ with this value in (45) and taking into account (33) and (35) results in

$$\sum_{n=-\infty}^{\infty} P_n(\tau)\, \exp j(-n\omega T) = \sum_{K=0}^{M-1} W_{0K}(\omega) \cdot \left(\sum_{n=-\infty}^{\infty} C_K(\tau + nT)\, \exp j(-n\omega T)\right). \tag{52}$$

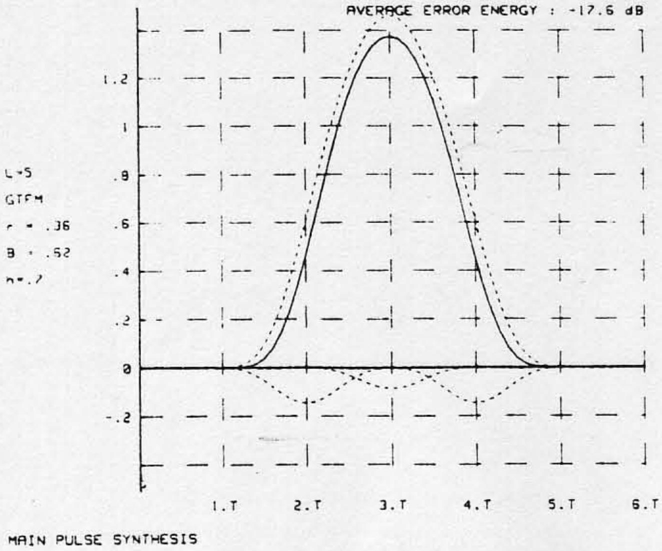

Fig. 10. Main pulse construction for GTFM ($h = 0.7$). Solid line: main pulse.

According to the definition of $W_{ij}(\omega)$, it can be shown that $W_{0K}(\omega)$ is given by

$$W_{0K}(\omega) = C^{\sum_{i=1}^{L-1} \alpha_{K,i}} \times \left[1 + (\exp j(-\omega T) - C) \times \sum_{p=1}^{L-1} \alpha_{K,p}\, \exp j(-p\omega T)\, C^{p+1-2\sum_{j=1}^{p} \alpha_{K,j}}\right]$$

$$W_{0K}(\omega) = \sum_{p=0}^{L} W_{0K}^{p}\, \exp j(-p\omega T). \tag{53}$$

An exact expression of the polynomials $W_{0K}(\omega)$ is given in Table II for $L \leqslant 4$.

Thus, the value of $P_m(\tau)$ ($P(t)$ for $t = \tau + mT$) is

$$P_m(\tau) = P(\tau + mT) = \sum_{K=0}^{M-1}\sum_{p=0}^{L} W_{0K}^{p} C_K(\tau + mT - pT). \tag{54}$$

Coming back to the complete main pulse instead of its different samples, we obtain the final expression of $P(t)$:

$$P(t) = \sum_{K=0}^{M-1}\sum_{p=0}^{L} W_{0K}^{p} C_K(t - pT). \tag{55}$$

This fundamental equation shows that $P(t)$ is actually a weighted superposition of time-shifted versions of the finite duration component functions $C_K(t)$. Moreover, considering the duration of each of the components together with the degree of the associated polynomial $W_{K0}(\omega)$, it can be shown that $P(t)$ is exactly of the same length as $C_0(t)$, i.e., $(L + 1)T$: this confirms the adequation of the approximate representation of the signal by the corresponding main pulse, as expressed in (39).

For $L = 1$ or for half-integer values of h, the main pulse is identical to the first component of the AMP representation of the signal, which is a proof of its usefulness:

$$L = 1 \text{ or } h = H + 1/2, \qquad H \text{ integer}$$

$$P(t) = C_0(t). \tag{56}$$

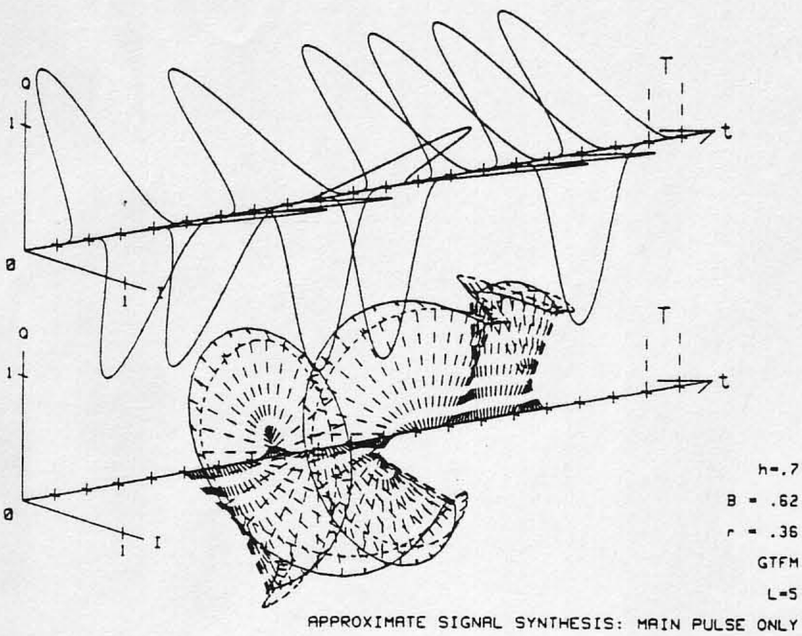

Fig. 11. Synthesis of GTFM using main pulse (h = 0.7). Bottom: same as Fig. 6.

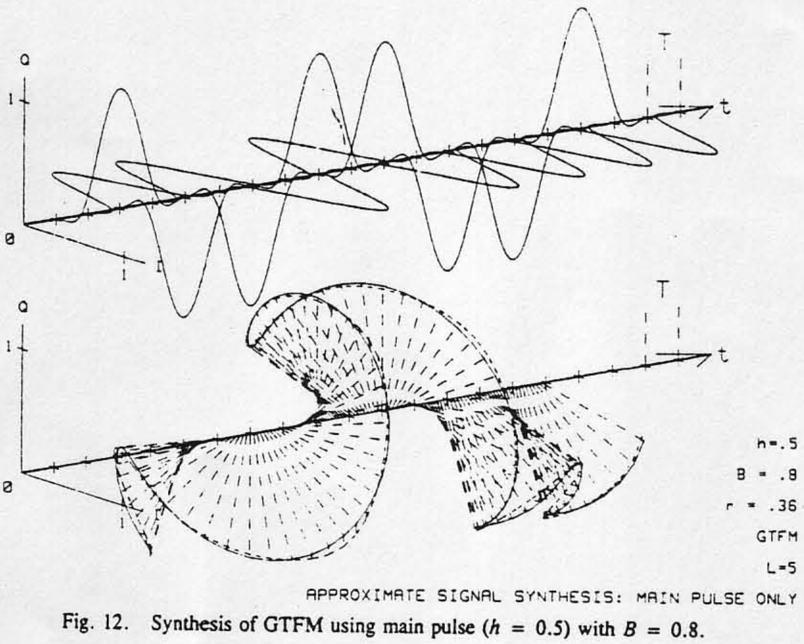

Fig. 12. Synthesis of GTFM using main pulse (h = 0.5) with B = 0.8.

In these cases, using only the first AMP component of the signal is the best—and simplest—possible approximation: for L = 1 and h = 0.5, the signal can be perfectly synthesized exactly like MSK or OQPSK, i.e., by amplitude modulating two quadrature CW's at half the bit rate.

D. Approximation Error

Considering (40)–(43), it can easily be shown that the average value of the error energy over a bit period, which expresses the quality of the approximation, is given by

$$\sigma^2 = \frac{1}{T}\int_0^T \sigma^2(\tau)\, d\tau$$

$$\sigma^2 = 1 - \frac{1}{T} \times \sum_{i=0}^{L}\sum_{j=0}^{L} C^{|i-j|} \int_0^T P(\tau+iT)\cdot P(\tau+jT)\cdot d\tau. \tag{57}$$

E. Examples and Remarks

Fig. 10 shows the synthesis of the main pulse for GTFM with h = 0.7. The solid line is the main pulse itself, which can be compared to the first component, $C_0(t)$ (closest dotted line).

Fig. 11 represents the approximate synthesis of the signal using only this pulse. (The case h = 0.5 was treated in Section II-C-3 (Fig. 6) since $P(t) = C_0(t)$.)

In the general case, one can observe that the quality of the approximate representation by main pulse increases with the degree of continuity of $\varphi(t)$. It is also better for low values of the modulation index (see Fig. 6) and/or when the effective phase change duration is small: in fact, for GTFM with $B = 0.62$, this duration is about $3 \cdot T$, which is rather long. The approximation is much better with $h = 0.5$ (low modulation index) and $B = 0.8$ (phase change duration approximately divided by 2); the synthesis is shown in Fig. 12.

V. Concluding Remarks

The AMP decomposition of MSK-like modulations ($h = 0.5$, any L) was already achieved a few years ago. Its form was slightly different from the one presented here, because it implicitly took into account the properties of $J = \exp j(\pi/2)$. Our generalization to any noninteger value of the modulation index was achieved shortly after, with the help of a desktop computer.

AMP interpretation of binary CPM is now in current use in our laboratories, owing to the important simplifications it leads to: spectrum computation, modulation analysis considering only the significant components, and conception of modulators and demodulators.

We generalized it recently to more complex modulations: in this improved version, it can be used for the study of the signal distortions in nonlinear transmitters and receivers. This generalization will be published in a subsequent paper.

Finally, extension to M-ary CPM (instead of binary CPM only) is comparatively easy. However, it is not so useful as in the binary case, because it is rather complicated and gives rise to many components of significant energy.

Acknowledgment

The author is grateful to the reviewers for their useful comments and suggestions.

References

[1] S. A. Gronemeyer and A. L. McBride, "MSK and offset QPSK modulation," *IEEE Trans. Commun.*, vol. COM-24, pp. 809–820, Aug. 1976.

[2] F. Amoroso and J. A. Kivett, "Simplified MSK signaling technique," *IEEE Trans. Commun.*, vol. COM-25, pp. 433–441, Apr. 1977.

[3] S. Pasupathy, "Minimum shift keying: A spectrally efficient modulation," *IEEE Commun. Mag.*, vol. 17, pp. 14–22, July 1979.

[4] K. Murota, K. Kinoshita, and K. Hirade, "GMSK modulation for digital mobile telephony," *IEEE Trans. Commun.*, vol. COM-29, pp. 1044–1050, July 1981.

[5] S. A. Rhodes, "FSOQ, a new modulation method that yields a constant envelope," in *Nat. Telecommun. Conf. Rec.*, Houston, TX, 1980, pp. 51.1.1–51.1.7.

[6] A. R. Hambley and O. Tanaka, "Generalized serial MSK modulation," *IEEE Trans. Commun.*, vol. COM-32, pp. 305–308, Mar. 1984.

[7] F. de Jager and C. B. Dekker, "Tamed frequency modulation—A novel method to achieve spectrum economy in transmission," *IEEE Trans. Commun.*, vol. COM-26, pp. 535–542, May 1978.

[8] K.-S. Chung, "Generalized tamed frequency modulation and its applications to mobile radio communications," *IEEE Trans. Veh. Technol.*, vol. VT-33, pp. 103–113, Aug. 1984.

[9] M. K. Simon, "A generalization of minimum-shift keying (MSK)-type signaling based upon input data symbol pulse shaping," *IEEE Trans. Commun.*, vol. COM-24, pp. 845–856, Aug. 1976.

[10] T. Le-Ngoc, K. Feher, and H. Pham Van, "New modulation techniques for low-cost power and bandwidth efficient satellite earth stations," *IEEE Trans. Commun.*, vol. COM-30, pp. 275–283, Jan. 1982.

[11] M. C. Austin and M. U. Chang, "Quadrature overlapped raised-cosine modulation," *IEEE Trans. Commun.*, vol. COM-29, pp. 237–249, Mar. 1981.

[12] S. Kato and K. Feher, "XPSK, a new cross-correlated phase-shift keying modulation technique," *IEEE Trans. Commun.*, vol. COM-31, pp. 701–707, May 1983.

[13] D. Muilwijk, "Correlative phase shift keying—A class of constant envelope modulation techniques," *IEEE Trans. Commun.*, vol. COM-29, pp. 226–236, Mar. 1981.

[14] T. Aulin and C.-E. Sundberg, "Continuous phase modulation (CPM)—Part I: Full response signaling," *IEEE Trans. Commun.*, vol. COM-29, pp. 196–209, Mar. 1981.

[15] T. Aulin, N. Rydbeck, and C.-E. Sundberg, "Continuous phase modulation—Part II: Partial response signaling," *IEEE Trans. Commun.*, vol. COM-29, pp. 210–225, Mar. 1981.

[16] T. Aulin and C.-E. Sundberg, "Numerical calculation of spectra for digital FM signals," in *Nat. Telecommun. Conf. Rec.*, New Orleans, LA, 1981, pp. D8.3.1–D8.3.6.

[17] G. J. Garrison, "A power spectral analysis for digital FM," *IEEE Trans. Commun.*, vol. COM-23, pp. 1228–1243, Nov. 1975.

[18] L. J. Greenstein, "Spectra of PSK signals with overlapping baseband pulses," *IEEE Trans. Commun.*, vol. COM-25, pp. 523–530, May 1977.

[19] H. Ganem and R. Vallet, "Démodulation des signaux numériques à enveloppe constante," in *GRETSI'83 Conf. Rec.*, Nice, France, May 1983, pp. 413–418.

[20] P. A. Laurent, "Interprétation des modulations d'indice demi-entier—Extension à des indices voisins et applications," in *GRETSI'83 Conf. Rec.*, Nice, France, May 1983, pp. 503–509.

★

Pierre A. Laurent was born near Bordeaux, France, in 1947. He received the diploma from the Ecole Centrale des Arts et Manufactures de Paris, Paris, France, in 1969.

Currently, he is working in the signal processing research team of Thomson-CSF, Division Télécommunications (DTC), Gennevilliers, France. His professional interests include modulations, error-correcting codes, vocoders, and everything relating to advanced telecommunication systems.