original

Reprinted from The Deep Space Network Progress Report 42-49
November and December 1978
Jet Propulsion Laboratory
Pasadena, California
February 15, 1979

Comparison of Phase Modulation Systems

J. L. Massey
Consultant to JPL

Comparison of the energy-to-noise ratio necessary to obtain a given modulation bit error probability has indicated that three-phase modulation is about 0.75 dB superior to four-phase. It is shown that this apparent superiority results entirely from the greater bandwidth required by the three-phase system for the same bit transmission rate. Two further comparison criteria are proposed, which are based on the cut-off rate R_0 of the discrete channel created by the modulation system. For the criterion which constrains both the bandwidth and transmitted power, it is shown that four-phase modulation is always superior to three-phase modulation. The conclusion is that three-phase modulation offers no practical advantage over four-phase modulation.

I. Introduction

In a recent memorandum (Ref. 1), Pierce has shown a substantial superiority for three-phase modulation, compared to two-phase and four-phase, for signaling in the presence of additive white Gaussian noise (AWGN). In this memorandum, we will show that Pierce's criterion of goodness is only one among three criteria that can reasonably be chosen, and is in fact the least "fair" of the three in that it both gives a larger bandwidth to three-phase modulation and also presumes that coding will not be used. We will show that if the criterion is changed to permit coding, then the superiority of three-phase modulation is negligible in spite of its requirement for greater bandwidth. We will show further that if coding is permitted and bandwidth is specified, then four-phase modulation is always superior to three-phase.

We shall in fact consider the general case of M-ary phase modulation in AWGN where a sinusoid with phase $2\pi\, m/M$ radians is transmitted in each signaling interval, or *baud*, where $m \in \{0, 1, 2, \ldots, M-1\}$. The following notation will be used:

E = transmitted energy per baud

E_0 = transmitted energy per modulation bit

P_0 = bit error probability at output of a "hard decision" demodulator

P_e = baud error probability for a "hard decision" demodulator

N_0 = one-sided noise power spectral density

R_0 = (also denoted by R_{comp}) cut-off rate per baud of the discrete memoryless channel created by the modulator, waveform channel, and an "infinitely-soft decision" demodulator

S = transmitted power

B = transmitter bandwidth measured in bauds per second

There are three useful relations among the above quantities, namely

$$E_0 = \frac{E}{\log_2 M} \tag{1}$$

$$S = EB \tag{2}$$

which are always true, and

$$P_0 \simeq \frac{P_e}{\log_2 M} \tag{3}$$

which is a good approximation when P_e is small. Equation (3) presumes, when $M = 2^i$, that the i modulation bits in each baud are gray-coded into the M phases so that baud errors by the demodulator, which with high probability are errors between adjacent phases, result in single bit errors with high probability. When $M \neq 2^i$, we presume that the modulation bits are appropriately mapped over several bauds so that Eq. (3) is still approximately valid when P_e is small (or, equivalently, when E_0/N_0 is large).

We shall consider only $M \geqslant 3$ for the following reason. Four-phase modulation can be thought of as two independent two-phase modulations in each baud. Thus, it is superfluous to consider two-phase modulation. Moreover, we thus avoid the question of whether B is the appropriate measure of bandwidth for two-phase modulation. We note, however, that B is certainly the appropriate measure of bandwidth when $M \geqslant 3$, since then both the "sine" and "cosine" quadrature components must be available in each baud.

II. The Comparison Criteria

Criterion 1: Fix P_0, the demodulator bit error probability, at some small value and compare the necessary E_0/N_0, the modulation bit energy-to-noise ratio.

Criterion 1 makes sense only if coding is prohibited so that P_0 is the error probability in the transmitted information bits, which is the only "error probability" of real interest to the user of the communication system. Note, however, that Eqs. (1) and (2) imply

$$B = \frac{S}{E_0 \log_2 M} \tag{4}$$

so that, for a specified S and E_0, the bandwidth decreases as M increases. In particular

$$\frac{(B)_3}{(B)_4} = \frac{\log_2 4}{\log_2 3} = 1.262 \tag{5}$$

(where the subscripts on the parentheses here and hereafter denote the value of M) which shows that three-phase modulation will use 1.01 dB more bandwidth than four-phase modulation when the transmitted power, S, and modulation bit energy, E_0, are fixed. Criterion 1 is indifferent to this increased bandwidth requirement for three-phase modulation, so it is perhaps not surprising that three-phase modulation will emerge as optimum for this criterion.

We remark that Criterion 1 is equivalent to that used by Pierce (Ref. 1), who argued that $P_e/\log_2 M$ should be fixed (at some small value) so that all systems would deliver the same "symbol error probability" for "symbols" consisting of a fixed number of binary digits. The equivalence to our Criterion 1 follows from Eq. (3).

Criterion 2: Fix E_0/N_0, the modulation bit energy-to-noise ratio, and compare the resulting cut-off rate per modulation bit, $R_0/\log_2 M$.

There are strong arguments supporting the claim that R_0 is the best measure of quality for a modulation system that may be used together with coding (Refs. 2, 3). In a very real sense, R_0 is the "practical upper limit" of information bits per baud for reliable transmission, just as channel capacity is the theoretical upper limit. Criterion 2 thus shows how many reliable information bits per modulation bit one is buying per unit of modulation bit energy.

Because the modulation bit energy rather than the baud energy is fixed in Criterion 1, it follows that the bandwidth of the different M-ary phase modulation systems for the same transmitted power will again be inversely proportional to $\log_2 M$. Thus, because the three-phase system is given more bandwidth, it is again not surprising that it appears superior to four-phase modulation under Criterion 2.

Criterion 3: Fix E/N_0, the baud energy-to-noise ratio, and compare the resulting cut-off rate per baud, R_0.

In our opinion, Criterion 3 is the most appropriate one for comparing M-ary phase modulation systems with different values of M. By fixing the energy per baud, one assures that the systems compared will have the same bandwidth when they use the same transmitted power. Moreover, the use of R_0, rather than P_e, in the criterion has the other advantages discussed above under Criterion 2. Criterion 3 thus directly measures the quality of M-ary modulation systems when both the transmitted power and bandwidth are specified – and this seems to be the fairest comparison possible.

The fact that three-phase modulation is always inferior to four-phase modulation under Criterion 4 strongly suggests that three-phase modulation is devoid of practical interest.

III. Analysis

A. Criterion 1

We use the signal space approach given in Ref. 4. The error probability p for the maximum-likelihood decision between two signal vectors at euclidean distance d from one another in the presence of AWGN with variance $N_0/2$ in each dimension is very closely approximated (and always over-bounded) by

$$P \approx \frac{1}{\sqrt{\frac{\pi d^2}{N_0}}} e^{-d^2/(4N_0)} \tag{6}$$

when $p \ll 1$ (Ref. 4). In M-ary phase modulation, the distance between signal vectors for adjacent phases is just

$$d = 2\sqrt{E}\sin\left(\frac{\pi}{M}\right) \tag{7}$$

Moreover, when p is small, almost all errors occur between the transmitted phase and its two adjacent phases so that $2p$ is a very close approximation to the baud error probability for $M \geqslant 3$. Thus, using Eqs. (1), (6), and (7), we obtain, for $M \geqslant 3$,

$$P_e \approx \frac{1}{\sqrt{\pi \log_2 (M) \sin^2\left(\frac{\pi}{M}\right)\frac{E_0}{N_0}}} e^{-\log_2 (M) \sin^2 (\pi/M) E_0/N_0} \tag{8}$$

where the approximation is extremely tight when P_e is small.

Let

$$\gamma_1 = \frac{(E_0)_4}{(E_0)_M} \tag{9}$$

be the *efficiency factor under Criterion 1* for M-ary phase modulation relative to four-phase modulation when $(P_0)_M = (P_0)_4$ or, equivalently by Eq. (3), when

$$\frac{(P_e)_M}{\log_2 M} = \frac{(P_e)_4}{2} \tag{10}$$

In the high signal-to-noise-ratio (SNR) limit where E_0/N_0 is extremely large, we see that Eq. (10) will hold only if the corresponding two exponents in Eq. (8) are (approximately) equal. Thus, in the high SNR limit, the efficient factor in Eq. (9) becomes simply

$$\gamma_1 = \log_2 (M) \sin^2\left(\frac{\pi}{M}\right) \tag{11}$$

which is tabulated in Table 1. Three-phase modulation, in the high SNR limit, is thus seen to be 0.75 dB better than four-phase modulation, *according to Criterion 1.* (Pierce's memo in Ref. 1 states the advantage as 0.5 dB, but 0.75 is, in fact, consistent with the curves given in Ref. 1.)

To find γ_1 elsewhere than for the high SNR limit, we can proceed as follows: first, choose a value γ_1 smaller than the high SNR limit, then substitute Eq. (8) into Eq. (10) and solve for $(E_0)_4/N_0$; finally, compute $P_0 = (P_e)_4/2$ from Eq. (8). Defining

$$A = \sqrt{\log_2 \frac{(M)}{\gamma_1}} \sin\left(\frac{\pi}{M}\right) \tag{12}$$

we can write the solution for $(E_0)_4/N_0$ as

$$\frac{(E_0)_4}{N_0} = \frac{\log_e\left(\frac{A \log_2 M}{2}\right)}{1 - A^2} \tag{13}$$

In Table 2, we show the bit energy-to-noise ratio $(E_0)_4/N_0$ and corresponding bit error rate P_0 at which three-phase modulation achieves the indicated efficiency factor. (The numbers in Table 2 agree with the curves in Pierce's memo, Ref. 1.)

B. Criterion 2

For any M-ary signal set (such as the M-ary phase modulation signal set) for which the uniform probability distribution maximizes the defining expression for R_0, R_0 for the AWGN case is given (Ref. 3) by

$$R_0 = \log_2 M - \log_2 \left[1 + \frac{1}{M}\sum_{i=1}^{M}\sum_{\substack{j=1\\ j\neq i}}^{M} e^{-|s_i - s_j|^2/(4N_0)}\right] \tag{14}$$

where $s_1, s_2, \cdots, s_M$ are the signal vectors. Using the symmetry of the M-ary phase modulation signal set, we can rewrite Eq. (14) as

$$R_0 = \log_2 M - \log_2 \left[1 + \sum_{j=1}^{M-1} e^{-E \sin^2 (j\pi/M)/N_0} \right] \quad (15)$$

where we have also used the fact that

$$|s_i - s_j|^2 = 4E \sin^2 \left[(j - i) \frac{\pi}{M} \right] \quad (16)$$

We define the efficiency factor under Criterion 2 for M-ary phase modulation relative to four-phase modulation to be

$$\gamma_2 = \frac{\left(\frac{R_0}{\log M} \right)_M}{\left(\frac{R_0}{2} \right)_4} \quad (17)$$

when the modulation bit energy-to-noise ratio is fixed, i.e., when $(E_0)_M/N_0 = (E_0)_4/N_0$.

In Table 3, we show this efficiency factor for three-phase modulation relative to four-phase modulation, over an interesting range of modulation bit energy-to-noise ratios, as found from Eqs. (15) and (17). From Table 3, we see that three-phase modulation offers only a slight advantage over four-phase modulation – the maximum advantage being only 0.16 dB – even though, under Criterion 2, three-phase modulation uses 1.26 times as much bandwidth compared to four-phase modulation. We see that this additional bandwidth is not being exploited very efficiently by three-phase modulation.

C. Criterion 3

We define the efficiency factor under Criterion 3 for M-ary phase modulation relative to four-phase modulation to be

$$\gamma_3 = \frac{(R_0)_M}{(R_0)_4} \quad (18)$$

when the baud energy-to-noise ratio is fixed, i.e., when $(E)_M/N_0 = (E)_4/N_0$. In Table 4, we show this efficiency factor, for three-phase modulation relative to four-phase modulation, over an interesting range of baud energy-to-noise ratios. We see that three-phase modulation, under Criterion 3 which constrains both the bandwidth and transmitted power of the compared systems to be the same, is *always inferior* to four-phase modulation, although by a small amount. We have also included the cut-off rate for eight-phase modulation in Table 4 to illustrate the fact that eight-phase modulation is always superior to four-phase. Does this suggest that one should always choose eight-phase modulation over four-phase? The answer to this question is an emphatic no! For baud energy-to-noise ratios, E/N_0, of 3.0 (5 dB) or less, the loss in R_0 suffered by four-phase modulation relative to eight-phase is negligible; thus, the greater simplicity of four-phase modulation makes it the practical choice. For large E/N_0, however, one should seriously ask whether the increased R_0 offered by eight-phase (or even higher order phase modulation) is sufficient to justify choosing the more complex modulation rather than four-phase modulation. However, we do argue that the Criterion 3 comparison does dictate that *four-phase modulation should always be chosen over three-phase modulation.* Our reasoning is that three-phase modulation is actually more difficult to implement than four-phase so that the former offers no compensating advantage for its loss in cut-off rate, R_0.

IV. Reconciling Criteria 1 and 2

It may appear puzzling that Criteria 1 and 2 lead to such different measures of the improvement that three-phase modulation offers over four-phase modulation, since both criteria give three-phase the same bandwidth advantage. One might conclude that this discrepancy is due to the use of P_0 (the demodulation bit error probability) in Criterion 1 versus the use of $R_0/\log_2 M$ (the cut-off rate per modulation bit) in Criterion 2. But we shall now show that this difference is not the cause of the discrepancy, after which we shall indicate the true cause.

Consider the following criterion:

Criterion 2′: Fix $R_0/\log_2 M$, the cut-off rate per modulation bit, and compare the necessary E_0/N_0, the modulation bit energy-to-noise ratio.

Notice that Criterion 2′ differs from Criterion 2 only as to which of the parameters, $R_0/\log_2 M$ and E_0/N_0, is held fixed. We find the *efficiency factor under Criterion 2′* for M-ary phase modulation relative to four-phase modulation to be

$$\sigma_{2'} = \frac{(E_0)_4}{(E_0)_M} \quad (19)$$

when $(R_0)_M/\log_2 M = (R_0)_4/2$.

Using Eq. (1) in Eq. (15) and solving for E_0/N_0, we find

$$\frac{(E_0)_4}{N_0} = -\log_e \left[2^{1-(R_0)_4/2} - 1 \right] \quad (20)$$

and

$$\frac{(E_0)_3}{N_0} = -\frac{4}{3\log_2 3}\log_e\left\{\frac{1}{2}\left[3^{1-(R_0)_3/\log_2 3} - 1\right]\right\} \quad (21)$$

In Table 5, we show the efficiency factor σ_2' for three-phase modulation relative to four-phase modulation, over an interesting range of modulation bit energy-to-noise ratios. In particular, the values of $(E_0)_4/N_0$ in the upper half of Table 5 are the same as those in Table 3 – we see from these that three-phase modulation has a much greater advantage over four-phase when measured under Criterion 2′ rather than Criterion 2. In fact, *the advantage measured under Criterion 2′ coincides almost exactly with that measured under the error probability Criterion 1* – as can be seen from the lower half of Table 5 where the values of $(E_0)_4/N_0$ are the same as those in Table 2.

We are now in position to see why Criteria 1 and 2′ give, in our opinion, a misleadingly large advantage to three-phase modulation over four-phase. As can be seen from Table 5, the apparent advantage (under Criterion 2′) of three-phase increases markedly as $R_0/\log_2 M$ approaches its asymptotic value of 1 for large energy-to-noise ratios. Criterion 2′ thus shows three-phase to be most favorable when modulation energy is being "wasted" to drive R_0 extremely close to 1. In our view, the purpose of the modulation system is to maximize the number of information bits that can be transmitted reliably per unit of modulation energy, i.e., to maximize $R_0/\log_2 M$ per unit of E_0. This dictates operation at a fairly low E_0/N_0 so that operation is in the range where R_0 is linear with E_0. The burden of actually achieving the desired reliability should be placed on the coding system, not the modulation system! This is the reason we argue that, even if three-phase is given a bandwidth advantage over four-phase as with Criteria 1, 2, and 2′, one should fix E_0/N_0 and compare $R_0/\log_2 M$ instead of vice versa. That is, for the same modulation energy one should see how good a channel is created for coding, rather than starting from some arbitrary channel and measuring how much energy is needed to create it. We note that the former kind of comparison is impossible with a modulation bit error probability criterion – there is no natural way to compare different error probabilities in "dBs" – and this dooms, in our opinion, most modulation comparisons based on error probability to be misleading.

We have argued that the apparent 0.75 dB superiority of three-phase modulation over four-phase at high energy-to-noise ratios is an advantage that results only when the modulation system is misused to deliver reliable bits to the user, rather than to pave the way for the coding system to do this. In the next and final section, we give an example of how a very simple coding system can make four-phase modulation appear superior to uncoded three-phase modulation when both systems use the same bandwidth. This example illustrates very clearly, we believe, why Criteria 1, 2, and 2′ are all misleading, although Criterion 2 is the least so.

V. Remarks and Conclusions

We have argued that Criterion 1 is an inappropriate criterion for comparing phase-modulation systems since it ignores bandwidth requirements and prohibits the use of coding. We now give an example to buttress our case against Criterion 1.

Suppose that four-phase modulation is encoded by adding one parity bit to each four information bits to be transmitted, the parity bit being the modulo-two sum of the 4 information bits. This expands the bandwidth by a factor of 5/4 = 1.25 so that, for the same energy per information bit, the coded four-phase modulation system has the same bandwidth as uncoded three-phase modulation (which requires 1.26 times the bandwidth of uncoded four-phase modulation). For the same information bit error probability, a calculation similar to that in Section III A above shows that the coded four-phase system is 1.29 dB superior to the uncoded three-phase system with the same bandwidth, for large values of the modulation bit energy-to-noise ratio. Moreover, the necessary encoder and decoder for the four-phase system are probably no more complex than the additional equipment needed to implement three-phase rather than four-phase modulation. Thus, when attention is given to bandwidth, the alleged superiority of three-phase modulation vanishes entirely.

We argue that, because it is the only one of the three criteria above which constrains both bandwidth and power, Criterion 3 is the most useful for comparing phase modulation systems. We have shown that Criterion 3 indicates that four-phase modulation should always be preferred over three-phase.

For some channels, notably the deep-space channel, bandwidth is of secondary importance so that there is a strong case for comparing modulation systems under a power constraint only. We argue that, in this instance, the appropriate choice is Criterion 2, rather than Criterion 1 (bit error probability) or Criterion 2′, for the following reason. Criterion 2 compares the quality of two modulation systems when both have the *same* transmitter power; whereas Criteria 1 and 2′ fix the quality and *then* compare the necessary power – this leads to misleading comparisons in the "inefficient" (or "saturated") modula-

tion region where small improvements in quality require large increases of power. Thus, the latter two criteria actually constrain channel quality, *not* transmitted power. We have seen that, under Criterion 2, three-phase modulation is superior to four-phase, but the maximum advantage (about 0.16 dB) is so slight that it generally would not be worth the extra complexity required to implement three-phase rather than four-phase modulation.

References

1. Pierce, J. R., "Three-Phase vs Two-Phase and Four-Phase Modulation, California Institute of Technology Internal Memorandum, Electrical Engineering Department, October 10, 1977.
2. Wozencraft, J. M. and Kennedy, R. S., "Modulation and Demodulation for Probabilistic Coding," *IEEE Trans. Info. Th.*, Vol. IT-12, pp. 291–297, July, 1966.
3. Massey, J. L., "Coding and Modulation in Digital Communications," in *Proc. 1974 Int. Zurich Seminar on Digital Comm.*, pp. E2(1)–E2(4), March 12–15, 1974.
4. Wozencraft, J. M. and Jacobs, I. M., *Principles of Communication Engineering*, New York, Wiley, 1967.

Table 1. Efficiency factor γ_1 by Criterion 1 for M-ary phase modulation in the high SNR limit

M	γ_1	γ_1 in dB
3	1.189	+0.752
4	1	0
5	0.802	−0.957
6	0.6462	−1.896
7	0.5285	−2.770
8	0.4393	−3.572
16	0.1522	−8.175

Table 2. Demodulator bit error probability P_0 at which the efficiency factor γ_1 is obtained for three-phase modulation (relative to four-phase modulation)

γ_1	γ_1 in dB	$(E_0)_4/N_0$	$(E_0)_4/N_0$ in dB	P_0
1.10	0.414	2.403	0.381	1.65×10^{-2}
1.13	0.531	3.988	6.01	2.62×10^{-3}
1.15	0.607	6.416	8.07	1.82×10^{-4}
1.16	0.645	8.900	9.49	1.29×10^{-5}
1.17	0.682	14.04	11.47	6.01×10^{-8}
1.18	0.719	30.97	14.91	1.80×10^{-15}
1.189	0.752	995.0	30.0	$<10^{-99}$

Table 3. Efficiency factor γ_2 for three-phase modulation (compared to four-phase modulation under Criterion 2) vs the modulation bit energy-to-noise ratio

E_0/N_0	$R_0/\log_2 M$		γ_2	γ_2 in dB
	$M = 4$	$M = 3$		
0.01	0.0072	0.0072	1.000	0
0.10	0.0703	0.0707	1.001	0.02
0.30	0.2002	0.2031	1.014	0.06
1.00	0.5481	0.5670	1.035	0.147
1.30	0.6523	0.6767	1.037	0.159
1.50	0.7094	0.7362	1.038	0.161
1.70	0.7579	0.7860	1.037	0.159
3.00	0.9299	0.9499	1.022	0.093
10.0	0.9999	1.0000	1.000	0

Table 4. Efficiency factor γ_3 for three-phase modulation (compared to four-phase modulation under Criterion 3) vs the band energy-to-noise ratio

E/N_0	R_0		γ_3	γ_3 in dB	R_0
	$M = 4$	$M = 3$			$M = 8$
0.01	0.0072	0.0072	1.000	0	0.0072
0.10	0.0712	0.0712	1.000	0	0.0712
0.30	0.2083	0.2081	0.999	0	0.2083
1.0	0.6321	0.6254	0.989	−0.05	0.6326
3.0	1.4188	1.3090	0.923	−0.35	1.444
10.0	1.9806	1.5834	0.799	−0.97	2.4380
30.0	2.0000	1.5850	0.7925	−1.01	2.9648

Table 5. Cut-off rate per modulation bit and modulation bit energy-to-noise ratios at which the efficiency factor σ_2' is obtained for three-phase modulation [a]

R_0 per modulation bit	$(E_0)_4/N_0$	$(E_0)_3/N_0$	σ_2'	σ_2' in dB
0.0072	0.0100	0.0100	1.000	0
0.0703	0.1000	0.0994	1.005	0.02
0.2002	0.3000	0.2955	1.015	0.07
0.5481	1.000	0.9547	1.048	0.20
0.6523	1.300	1.227	1.060	0.25
0.7094	1.500	1.406	1.067	0.28
0.7579	1.700	1.583	1.074	0.31
0.9299	3.000	2.707	1.108	0.45
0.9999	9.577	8.252	1.161	0.65
0.8751	2.403	2.196	1.095	0.39 (0.41)
0.9735	3.988	3.546	1.125	0.51 (0.53)
0.9976	6.416	5.593	1.147	0.60 (0.61)
0.99980	8.900	7.683	1.158	0.64 (0.65)
0.99999885	14.04	12.007	1.169	0.68 (0.68)
$\to 1$	$\to\infty$	$\to\infty$	1.189	0.75 (0.75)

[a] Numbers in parentheses are the values of σ_2 at the same $(E_0)_4/N_0$.