AT&T Technical Journal
Vol. 64, No. 10, December 1985
Printed in U.S.A.

On Binary Differential Detection for Coherent Lightwave Communication

By J. E. MAZO*

(Manuscript received July 31, 1985)

Motivated by the communication problems caused by phase noise in those semiconductor lasers that may be used for fiber-optic data transmission, we consider heterodyned binary Differential Phase-Shift Keying (DPSK) in conjunction with high-rate (short time chip) redundancy as provided by repetition or by more complex coding techniques. In surprising contrast to repetitive coherent phase-shift keying where only a loss of a $2/\pi$ (2 db) in power is incurred in the limit of infinitely many infinitesimal time chips, we show that DPSK requires, in this limit, an infinite number of photons per bit. This is true regardless of the coding scheme used with the DPSK modulation. Next we find the bandwidth expansion that minimizes the number of received photons per bit required to hold the error rate at 10^{-9} for two situations: first for a simple repetition code, and then for a repeated (24, 12) Golay code with maximum likelihood detection. The performance of the latter is assumed to be representative of other optimally detected codes of the same rate, such as convolutional codes with Viterbi decoding. Explicit curves relating required photons per bit to the bandwidth expansion are given for B/R ratios of 0.01 to 10, where B is the laser linewidth and R is the data rate. An example of the results is that for $B/R = 0.1$ and a bandwidth expansion of 10, about 23 photons per bit are required for the repeated Golay code to perform as well as uncoded DPSK without phase noise (which requires 20 photons per bit for $Pe = 10^{-9}$). If $B/R = 0.01$ the bandwidth expansion is reduced to 2, and 12 photons per bit are required, thus outperforming the phase-stable, but uncoded, situation.

* AT&T Information Systems.

I. INTRODUCTION

Semiconductor lasers that may be used for coherent data transmission over optical fibers* have severe phase instabilities. Attempts to transmit a sine wave of frequency f_c result in outputs that are modeled as

$$\cos(2\pi f_c t + \phi + w(t)), \qquad 0 \leqslant \phi \leqslant 2\pi, \tag{1}$$

where $w(t)$ (measured in radians) is a random process representing the phase instability. The process $w(t)$ is usually taken to be a Wiener process and that then implies a Lorentzian [see eq. (4)] line shape for the power spectrum of (1). Such spectra are indeed observed and 3-db bandwidths as large as 10 to 20 MHz have been measured.† These bandwidths imply that the standard deviation of the change in $w(t)$ over a μs can be as large as 4π. Severe problems would be encountered with any conventional coherent detection scheme if one is transmitting data at ten-megabit rates (or lower) rather than gigabit rates. Nevertheless, one may wish to do precisely that, and our purpose here is to mathematically explore one very natural approach, Differentially coherent Phase-Shift Keying (DPSK) in combination with code symbols that have short transmission time. The purpose of using code chips of short duration is to mitigate the effects of phase wander between adjacent chips.

We emphasize that the scheme we are about to investigate is not the only possible one. One could use on-off keying of the optical carrier (1) with photon counting for detection. Theoretically this outperforms DPSK by 3 db, even with a stable transmitting carrier assumed for the latter. However, photon counting is not easy to implement, and practical avalanche photodiodes can introduce 20 db of loss. Thus other techniques, which involve heterodyning, are of interest, in hopes that their implementations can be closer to their own theoretical ideals. For a general survey of lightwave communications we recommend Ref. 1. In Ref. 2 a large number of modulation schemes for coherent optics are evaluated with the main purpose being that of determining the range of B/R values for which coding is not required. We choose here to examine DPSK in detail, but the general behavior of its performance with coding is expected to be representative of modulation formats that do not involve tracking the phase $w(t)$ with a phase-locked loop. The latter was one of the methods considered in Ref. 2 and was shown to be feasible only if $B/R < 0.003$.

Returning to the repetition-DPSK scheme, we note that it might be expected that in the limit of an infinite number of infinitely rapid

* In optical-fiber work, coherent transmission refers to any modulation format where an optical oscillator is required at the receiver.

† The carrier wavelength of interest is 1.55 μm, or f_c is roughly 2×10^{14} Hz.

code chips, the performance would approach something like that of full interval DPSK with a stable oscillator. One of the surprises that we have uncovered (and perhaps the major point of interest of this work) is that this is far from the truth. We show that in this limit an infinite number of incident photons per bit are required for fixed error rate.

Decreasing the number of repetitions while holding the error rate fixed also will ultimately require the photons per bit to become unbounded. This occurs when one approaches (from above) the number of repetitions required to achieve the given error rate, with phase noise being the sole impairment. Consequently, one expects that there will be an optimum bandwidth expansion. This is in fact true, and it is discussed in Section IV, while the similar problem for more sophisticated coding is treated in V.

In Section II we begin by presenting the mathematical details of the model, while the two limiting cases of large and minimum bandwidth expansion are investigated for the repetition code in Section III.

II. MODEL DETAILS

In the representation (1) we take

$$w(t) = \int_0^t n(t')dt' \tag{2}$$

and set the (two-sided) spectral density of the white noise $n(t)$ to be $N/2$. The variance $\sigma_w^2(t)$ of $w(t)$ at time t is then

$$\sigma_w^2(t) = \frac{Nt}{2}\ (\text{radians})^2. \tag{3}$$

The power spectrum of (1) can be calculated in terms of these quantities and is given by[3]

$$\frac{1}{8}G(f - f_c) + \frac{1}{8}G(f + f_c)$$

with

$$G(f) = \frac{N}{\omega^2 + (N/4)^2}, \tag{4}$$

where, as usual, $\omega = 2\pi f$. From (4) the 3-db bandwidth B of the spectrum is

$$B = N/4\pi, \tag{5}$$

and thus from (3) and (5)

$$\sigma_w^2(t) = 2\pi Bt. \tag{6}$$

Often in coherent optics one converts (1) to microwave frequencies (GHz) where conventional signal processing techniques are available. This heterodyning is accomplished by mixing (1) with a locally generated optical wave. The local oscillator also has phase instabilities that add to those of the received signal, and thus in this paper the effective bandwidth at microwave is taken to be double that at optical frequencies. Furthermore, shot noise fluctuations in photon counts during the heterodyning causes a white noise background to be added to the microwave signal. In our model we assume heterodyning to be done, and thus our received unmodulated carrier is modeled as

$$A \cos(\omega_o t + \phi + \phi(t)) + n(t), \tag{7}$$

where $\phi(t)$ is the Wiener process phase noise of variance

$$\sigma_\phi^2(t) = 4\pi Bt \tag{8}$$

and $n(t)$ is a Gaussian white noise process of spectral density $N_0/2$.

Assume, momentarily, that $\phi = 0$ and $\phi(t) = 0$ over a bit interval T, and we wish to coherently detect the modulation $\pm A$. We simply multiply by $\cos \omega_o t$, integrate the result for T seconds, and observe the sign of the output. The chance of making an error, Pe, is then

$$Pe = \frac{1}{2} \operatorname{erfc} \sqrt{\frac{E_b}{N_0}}, \tag{9}$$

where

$$\operatorname{erfc} x = 1 - \frac{2}{\sqrt{x}} \int_0^x \exp(-t^2) dt.$$

In (9), $E_b = A^2T/2$ is the energy per bit in the transmitting signal. When (7) arises, as it does in our case of interest, from heterodyning of an optical wave, E_b/N_0 is not an arbitrary parameter but is (see Ref. 1) numerically equal to the average number of photons per bit in the optical wave at the receiver. An E_b/N_0 corresponding to 18 photons per bit yields an error rate of 10^{-9} for coherent detection.

The quantity E_b/N_0 is also numerically equal to the signal-to-noise ratio (s/n) if the noise power N is measured in a bandwidth equal to $1/T$.

To motivate a later discussion, consider the coherent case further and instead of integrating the received signal over $(0, T)$ and making a decision (what we might unconventionally call soft-decision decoding), we make $n = 2m + 1$ hard decisions based on time chips of length T/n, and then use a majority vote to decide the sign of the transmitted bit. If p_c is the chip error rate, the bit error rate, $Pe(n)$, would be

$$Pe(n) = \sum_{k=m+1}^{2m+1} \binom{2m+1}{k} p_c^k(1-p_c)^{2m+1-k}. \qquad (10)$$

Since we want to keep E_b constant, the energy in each chip decreases as E_b/n. For n large, then, we have from (8) and (9)

$$p_c = \frac{1}{2} - \sqrt{\frac{\gamma}{n\pi}}, \qquad (11)$$

where

$$\gamma = E_b/N_0. \qquad (12)$$

For the binomial distribution represented by (10) and (11), we have that

$$(2m+1)p_c \approx m - \sqrt{\frac{2\gamma}{\pi}}\sqrt{m}$$

is the average number of errors; the variance of this number is

$$(2m+1)p_c(1-p_c) \approx \frac{m}{2}.$$

Since the lower limit on the sum in (10) is $(m+1)$, or $2\sqrt{\gamma/\pi}$ standard deviations above the mean, we have

$$\lim_{n\to\infty} Pe(n) = \frac{1}{\sqrt{2\pi}} \int_{2\sqrt{\frac{\gamma}{\pi}}}^{\infty} \exp(-x^2/2)dx = \frac{1}{2}\operatorname{erfc}\sqrt{\frac{2\gamma}{\pi}}. \qquad (13)$$

In (13) we see the ubiquitous $2/\pi$ penalty in s/n for using hard decisions. Equation (13) was derived for coherent detection. When we later consider repetition-DPSK to overcome phase noise, we will be concerned with a corresponding limit for differential detection. Then the fortunate limiting behavior we have just observed will not occur, because, with DPSK, the chip error rate approaches 1/2 more rapidly with n than it does in the coherent case exemplified by (11).

To make a tractable model for DPSK, we assume that the received waveform is

$$A\left(\sum_{-\infty}^{\infty} a_n g(t-nT_c)\right)\cos(\omega_o t + \theta + \phi(t))$$
$$+ n_x(t)\cos\omega_o t - n_y(t)\sin\omega_o t. \qquad (14)$$

The pulse $g(t)$ is assumed brick-wall Nyquist with $g(0) = 1$, and energy T_c. The Gaussian noise processes $n_x(t)$ and $n_y(t)$ are independent, flat spectrum, and of equal variance $\sigma^2 = N_o/T_c$. As stated earlier, the chip time, T_c, for the repetition code is related to the bit time T via $T_c =$

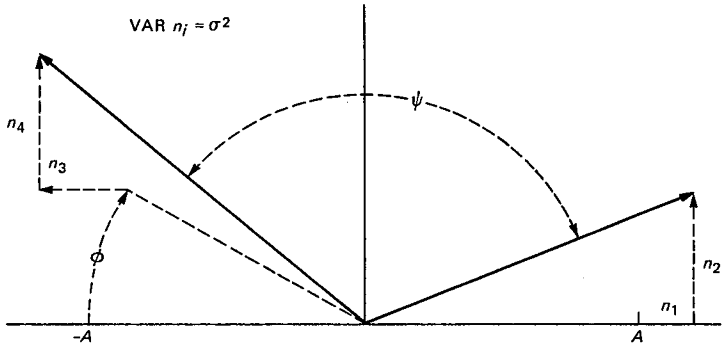

Fig. 1—Geometry of variables for differentially coherent phase-shift keying.

$T/(2m + 1)$, and $(2m + 1)$ repetitions* of the bit are differently encoded into the chips $a_n = \pm 1$.

We imagine recovering the differentially encoded chips a_n by demodulating (14) with $\cos \omega_o t$ and $-\sin \omega_o t$ and sampling the demodulation outputs at the appropriate T_c second intervals. At each sampling instant, then, we obtain a pair of real numbers, which we regard as a point in the plane. In the absence of phase and additive noise, this point would simply be $(\pm A, 0)$, relative to a coordinate system fixed by the unknown phase θ. With noise included, the geometry for two consecutive samples is equivalent to that shown in Fig. 1, drawn for the case of different consecutive a_n. All the components of the two vectors $(A, 0)$ and $(-A, 0)$ are perturbed by additive independent and identically distributed zero-mean Gaussian noise of variance σ^2. Owing to the phase noise, one of the perturbed vectors is also rotated by an angle ϕ, where ϕ itself is a zero-mean Gaussian variable of variance

$$\sigma_\phi^2 = 4\pi B T_c, \tag{15}$$

where B is the laser linewidth. If the angle ψ between the resulting vectors is less than $\pi/2$, we would declare that the two consecutive chips were the same and, in the present case, an error would be made. Of course, for $\psi > \pi/2$ we decide the chips were different.

From (15) we see that as the chip interval decreases, σ_ϕ^2 approaches zero and phase noise will make a negligible contribution to the chip error rate p_c. Note also that the additive noise variance σ^2 increases as T_c decreases.

Before we proceed, a few comments about the model are in order. An equivalent detection procedure is to delay the signal represented

* More generally, our results for the repetition code are unchanged if the bits are mapped into any two complementary patterns of n chips.

in (14) by a chip time, T_c, and sample the product of the delayed and undelayed signals. The resulting decision statistic is identical to the angle that we consider, being a representation of it as a quadratic form equal to the inner product of two noisy vectors. Secondly, a more realistic description of the signal uses a pulse $g(t)$ that has unit value over the interval T_c and is zero elsewhere. The bandwidth of the signal then is not precisely defined, but if one estimates the bandwidth of a flat front-end filter required for noise filtering to be $1/T_c$ (and ignores any intersymbol interference), then the numerical results are unchanged. Finally, we note that Salz[2] integrates the product of the signal and the delayed signal, rather than simply sampling. This complicates the analysis considerably. We have not attempted to investigate the difference in detail over the full range of B/R values, but we note the following. Salz estimates the B/R value required in order that DPSK detection suffers only 1-db degradation compared with $B/R = 0$ and finds $B/R < 0.003$ is sufficient. We calculate this precisely for our model and find $B/R < 0.002$. This suggests that the postdetection processing provided by the integrator might not be significant.

We use (10) later to calculate the bit error rate for repetitive DPSK (with an appropriate p_c). It may be objected that the use of this formula for the repetitive code is not rigorously justified, since samples used for detecting consecutive chips have one noise sample in common and thus the chip detection probabilities are not independent. This objection is easily overcome by assuming that the chips for two successive bits are interleaved in the manner *abab* $\cdots$.

III. LIMITING CASES

In this section we treat two limiting cases of repetitive DPSK. One case is concerned with a very large number of rapid repetitions. Here since the chip interval T_c becomes small, the phase noise is neglected, and for fixed E_b the chip s/n is small. The other case examines the minimum number of repetitions required to achieve a fixed error rate if phase noise were the only impairment. To approach this limit would require a large E_b provided that more than one repetition is required.

We begin with the high repetition rate limit. It is well known that the error probability for DPSK (see Ref. 4) is

$$\frac{1}{2}\exp(-\rho), \tag{16}$$

where ρ is the s/n. In terms of the parameters that apply to Fig. 1,

$$\rho = \frac{A^2}{2\sigma^2} = \frac{A^2}{2}\frac{1}{N_0(1/T_c)} = \frac{E_c}{N_0}, \tag{17}$$

where $E_c = E_b/n$ is the energy per chip. For small E_c (large n) (16) yields for the chip error probability, p_c,

$$p_c = \frac{1}{2}\left(1 - \frac{\gamma}{n}\right). \tag{18}$$

Note the difference in behavior for the differentially coherent problem [see (18)] versus the coherent one [see (11)]. In (18) the chip error rate approaches 1/2 as $1/n$, whereas in (11) the behavior is as $1/\sqrt{n}$. Thus, for DPSK the lower limit of (10) is $O(1/\sqrt{n})$ standard deviations above the mean, and for large n we have that the bit error probability $Pe(n)$ obeys

$$\lim_{n\to\infty} Pe(n) = \frac{1}{2}. \tag{19}$$

In essence, then, we have that if, in Fig. 1, only one of the vectors is noisy and ϕ is set to zero (coherent phase-shift keying), then $Pe(n)$ is small in the limit of many repetitions with constant E_b, but if both are noisy (DPSK), $Pe(n)$ limits to 1/2.

For the second limiting case, when the only perturbing influence to the transmission is the Gaussian phase noise variable ϕ, the chip error rate is

$$\begin{aligned} p_c &= \frac{2}{\sqrt{2\pi}\sigma_\phi} \sum_{k=0}^{\infty} \int_{(1+4k)\pi/2}^{(3+4k)\pi/2} \exp(-\phi^2/2\sigma_\phi^2(n))d\phi \\ &= \sum_{k=0}^{\infty}\left[\operatorname{erf}\frac{(3+4k)\pi}{2\sqrt{2}\sigma_\phi(n)} - \operatorname{erf}\frac{(1+4k)\pi}{2\sqrt{2}\sigma_\phi(n)}\right], \end{aligned} \tag{20}$$

where $\sigma_\phi^2(n)$ is given by

$$\sigma_\phi^2(n) = 4\pi BT_c = \frac{4\pi BT}{(2m+1)} = \frac{4\pi B}{nR}. \tag{21}$$

In (20), p_c is explicitly a decreasing function of the number of repetitions via the phase noise variance σ_ϕ^2. Further, if $Pe(n)$ is fixed in (10), that expression implicitly determines p_c as an increasing function of n. Requiring that both (10) and (20) determine the same value of p_c fixes the number of repetitions required to achieve the bit error probability Pe. Including additive noise in the calculations can only increase the required repetition rate for fixed Pe. Setting $Pe = 10^{-9}$, we have computed the number of repetitions, $\bar{n}$, required when phase noise is the sole impairment. The probability $\bar{p}_c$ is the chip error probability (20) for $\bar{n}$ repetitions. Both are displayed in Table I for several values of B/R. Although our main focus will not be on numerical values of p_c, it is worth emphasizing that throughout this paper values of p_c above 0.1, and even approaching 1/2, are possible for the larger values of B/R.

Table I—Minimum number of repetitions $\bar{n}$ required so that bit error rate does not exceed 10^{-9} with phase noise as sole impairment

B/R	$\bar{n}$	$\bar{p}_c$
0.01	3	2×10^{-5}
0.1	5	0.0017
1	21	0.042
10	83	0.20

The ratio signal-bandwidth/laser-linewidth equals nR/B. From Table I we see that this number is sufficiently high so that the implicit neglect of the wideband filtering on the phase noise that was made when writing the model [see (14)] seems justified for the parameters of interest.

IV. OPTIMUM REPETITION RATE

In this section we do a general investigation of DPSK detection with repetitions, including both phase noise and additive noise as impairments. Our main interest will be to determine the optimum repetition rate and the corresponding bits per photon required to achieve a bit error probability of 10^{-9}.

There are probably several useful general expressions for the chip error rate p_c when both phase noise and Gaussian noise are present. We shall work with the one given in (22), namely,

$$p_c = \frac{1}{2} - \frac{\rho \exp(-\rho)}{2} \sum_{s=0}^{\infty} \frac{(-1)^s}{(2s+1)} \cdot \left[I_s\left(\frac{\rho}{2}\right) + I_{s+1}\left(\frac{\rho}{2}\right) \right]^2 \exp\left(- \frac{(2s+1)^2 \sigma_\phi^2(n)}{2} \right), \quad (22)$$

where ρ is the chip (s/n) [see (17)], $\sigma_\phi^2(n)$ is the phase noise variance with n repetitions [see (21)], and $I_s(\rho/2)$ are the modified Bessel functions given by

$$I_s(x) = \left(\frac{x}{2}\right)^s \sum_{j=0}^{\infty} \frac{(x^2/4)^j}{j!(j+s)!}. \quad (23)$$

An expression similar to (22) for constant phase error was first derived by Blachman[5] and is given in eq. (62) of Ref. 4. Performing a simple average when this angle has Gaussian statistics yields (22). An independent derivation is given in the Appendix.

The successive terms of the sum in (22) decrease in magnitude, and hence, as for any such alternating series, the first neglected term bounds the error. We find that 15 terms in double precision (single precision on a Cray I) is enough to duplicate (16) when $\sigma_\phi^2(n) = 0$, and $\rho \leqslant 20$ ($p_c \geqslant 10^{-9}$).

We first use (22) to calculate the deterioration in p_c as B/R increases. Table II shows some results for $\rho = 20$, $n = 1$. We note that B/R = 0.002 yields about a 1-db degradation.

Next in Figs. 2 through 5 we plot as a function of the number of repetitions, n, the number of photons per bit, γ, which are required at the receiver to maintain a bit error rate of 10^{-9} for B/R values of 0.01 to 10. This is done using (10) and (22) in the following way. First n is chosen and the required p_c is determined from (10). For fixed B/R and known n, σ_ϕ in (21) is known, and (22) is then used to compute the chip s/n ρ that will achieve that p_c. Finally, $\gamma = n\rho$. These figures show quantitatively how the optimum value of n (and the correspond-

Table II—p_c vs. B/R for $\rho = 20$, $n = 1$

B/R	p_c
0	10^{-9}
0.001	6×10^{-9}
0.002	4×10^{-8}
0.01	2×10^{-4}
0.1	0.17
1	0.498841
10	0.49999...

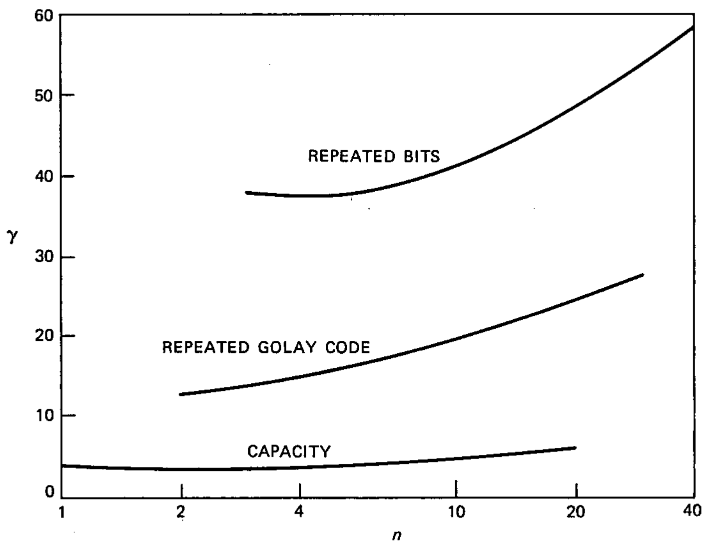

Fig. 2—Bits per photon versus bandwidth expansion for $B/R = 0.01$.

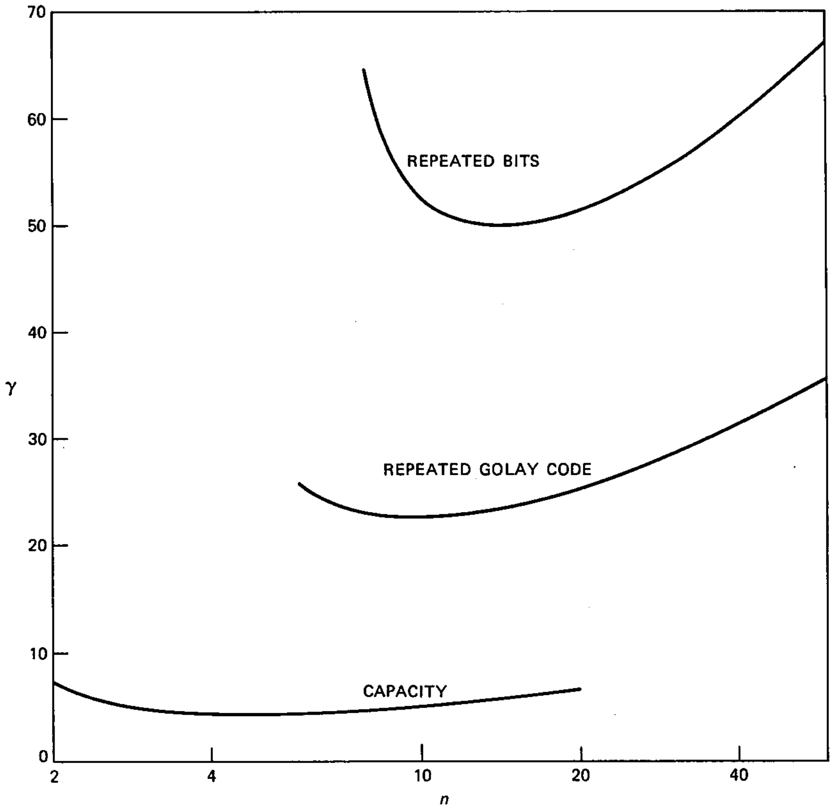

Fig. 3—Bits per photon versus bandwidth expansion for $B/R = 0.1$.

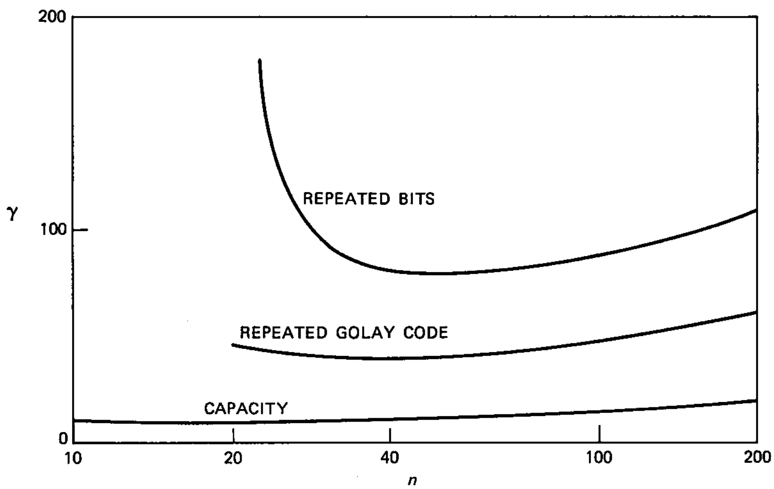

Fig. 4—Bits per photon versus bandwidth expansion for $B/R = 1$.

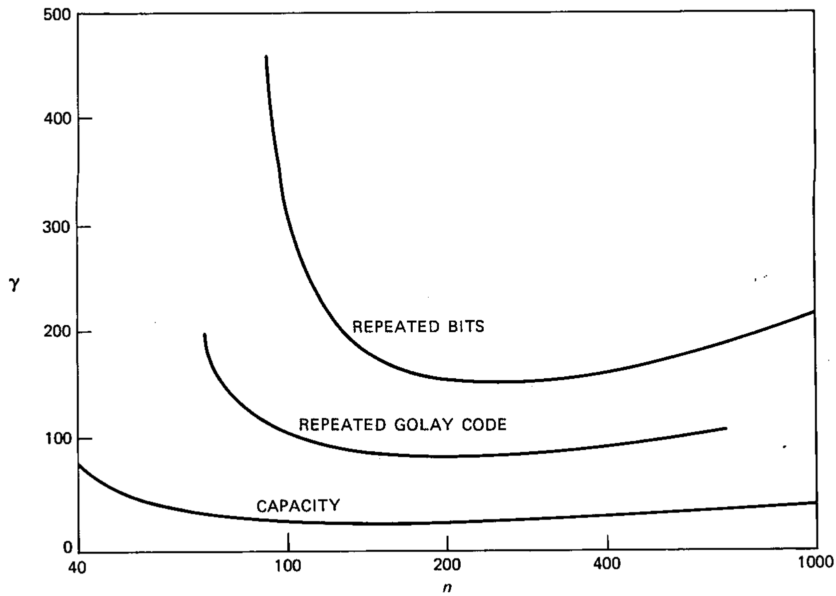

Fig. 5—Bits per photon versus bandwidth expansion for $B/R = 10$.

Table III—Optimum bandwidth expansion for repetition code ($Pe = 10^{-9}$)

B/R	n^*	p_c^*	γ^*	$\gamma^*/20$ (db)
0.01	5	4.6×10^{-4}	37	2.7 db
0.1	15	0.0255	50	4.0
1	49	0.141	79	6.0
10	239	0.313	157	8.9

ing value of γ) increase with B/R. From these figures, we derive Table III, which lists n^*, the optimum repetition rate, and γ^*, the minimum number of photons per bit required to hold the error rate at 10^{-9}. The last column in Table III compares γ^* with 20, the number of photons per bit required at this error rate for DPSK with no repetitions and a stable phase.

We note that the curves are often fairly flat around the minimum, and hence less bandwidth may be used without significantly increasing the required number of photons per bit.

V. CHANNEL CAPACITY AND GOLAY CODES FOR HETERODYNED DPSK

The repetition (or complementary) code for transmitting at rate $R = 1/T$ is but one way to use the discrete time Binary Symmetric

Table IV—Minimum and optimum bandwidth expansion for capacity and repeated Golay code

B/C	$\bar{n}$	$\bar{p}_c$ capacity	n^*	γ^*	p_c^*
0.01	<1.0001	~10^{-5}	2.2	3.192	0.126
0.1	1.7	0.0828	4.3	4.29	0.224
1	6.3	0.265	20	8.83	0.369
10	34.5	0.397	146	24	0.451
B/R	Repeated (24, 12) Golay Code $P_w = 1.67 \times 10^{-10}$				
0.01	—	—	2	12.5	0.00237
0.1	4	0.00507	10	22.8	0.0691
1	14	0.0973	36	38	0.221
10	60	0.277	194	85	0.372

Channel (BSC) that we have to work with.* Therefore, to see what is theoretically possible, we next consider the required number of photons per bit that would be required to transmit at channel capacity. If C is capacity in bits per second, we will fix B/C and plot γ versus n where the chip time is $T_c = 1/(nC)$. That is, the binary code using the chips should have a capacity $\bar{C} = 1/n$ bits per chip. Here n is the bandwidth expansion and need not be an integer.

The required chip error probability is found by equating the channel capacity for the BSC to $1/n$, that is,

$$\bar{C} \equiv 1 + p_c \log_2 p_c + (1 - p_c)\log_2(1 - p_c) = \frac{1}{n}. \tag{24}$$

The chip s/n ρ required to yield this p_c is then computed from (22), and $\gamma = n\rho$.

Curves for $B/C = 0.01$ to 10 are presented in Figs. 2 through 5, and summarized in the first half of Table IV. Once again, a feature is the existence of an optimum chip rate. The necessity of this is easily argued. The lowest value of n possible is determined by the phase noise alone, which causes p_c to increase as n decreases. Eventually, the capacity $\bar{C}$ drops below $1/n$, and this fixes the minimum $n = \bar{n}$. But to approach the pure phase noise situation, γ must increase indefinitely as $\bar{n}$ is approached from above. To see why γ must increase as n is very large, consider plotting $\bar{C}$ versus n with γ fixed. Then, as we have seen [see (18)], $p_c = (1/2) - O(1/n)$ owing to Gaussian (shot) noise. Setting $p_c = (1/2) - \epsilon$ and expanding the left side of (24) in powers of ϵ, we have, to lowest order,

$$\bar{C} = \frac{2\epsilon^2}{ln2} \text{ bits/chip.} \tag{25}$$

* In fact, interleaving (explained at the end of Section II) creates two parallel BSCs. The total capacity is the sum of the capacity of each and is the same as the capacity calculated here as if all chip errors were independent.

Since $\epsilon = O(1/n)$ and we have $O(n)$ chips per second, the capacity measured in bits per second vanishes like $1/n$. To avoid this, γ must increase.

Finally, we consider specific coding schemes that are more involved than repetition. We explicitly consider the extended binary (24, 12) Golay code. Thus for block length 24, 12 information bits are specified, yielding a rate 1/2 code. This is a linear code and any code word has 759 nearest neighbors at minimum distance $d_{\min} = 8$. The coding scheme that we consider is simply to construct low-rate code words by repeating a given Golay code word J times, make hard decisions on the chips at the receiver (assuming appropriately interleaved DPSK), and then use maximum likelihood decoding on the resulting binary code word of length $24J$. Note that the new code has $d_{\min} = 8J$ and that the code rate has decreased to $1/(2J)$.

Assuming that a word error results, on the average, in six bit errors, we set the word error rate, P_w, to be (1/6) 10^{-9}. If p_c is the chip error rate, then the union bound yields

$$P_w \leq 759 \sum_{i=4J}^{8J} \binom{8J}{i} p_c^i (1 - p_c)^{8J-i} + \binom{\text{higher}}{\text{terms}}, \tag{26}$$

where "higher terms" represents the probability of decoding into words further away than the minimum distance. These terms are neglected for the low word error probability of interest here.

We have in mind that actual attempts at coding would use repeated convolutional codes and Viterbi decoding. Our introduction of Golay codes is simply to make the analysis easier, but overall performance gains for the same repeated code rate are expected to be the same. In fact, Chase[6] finds that a (properly chosen) repeated 16-state convolutional code performs slightly better than a repeated Golay code. The effectiveness of repeating an appropriately chosen code to obtain a good low-rate code was, in fact, proposed by Chase,[6] who was concerned with codes when p_c is large, as is often the case for the present problem.* Perhaps better rate $1/(2J)$ convolutional codes exist than can be generated by repeating the symbols of a given one, but Chase shows that, at least from a minimum distance point of view, a repeated rate 1/2 convolutional code (suitably chosen) is close to optimum for code rates at least as small as 1/128.

Returning to the Golay code, note that J repeats (of any rate 1/2 code, in fact) corresponds to a bandwidth expansion of $n = 2J$.

Calculations of required photons per bit versus bandwidth expansion

* A point emphasized in Ref. 6 is that for $p_c > 0.25$ and asymptotically large block length, bounded distance algebraic decoders cannot operate ($Pe \to 1$) and maximum likelihood decoders must be used.

are done in a similar manner, as earlier. The bandwidth expansion $n = 2J$ is picked, p_c is found from (26) with $P_w = 1.67 \times 10^{-10}$, and then (22) is used to solve for ρ. The improvements obtained over simple repetition are displayed in Figs. 2 through 5, and essential features of the results are given in the second half of Table IV. In general, we see that the optimum bandwidth expansion is less than with repetition, and, of course, so is the required number of photons per bit. An important feature is that for $B/R = 0.01$ and 0.1, the required number of photons per bit is less, or comparable to, that required when no phase noise is present and no coding is used. In these cases the bandwidth expansion is relatively modest as well.

VI. CONCLUSION

To reduce the harmful effects of phase noise, coding schemes that use DPSK detection of code symbols having short duration were examined. We first considered in detail a simple repetition code and determined the optimum bandwidth expansion that minimized the number of received photons per bit required for an error probability of 10^{-9}. If $B/R = 0.01$, we found $n^* = 5$ and $\gamma^* = 37$. The corresponding numbers for a repeated Golay code were $n^* = 2$ and $\gamma^* = 12.5$. By contrast, 20 photons per bit are required for phase stable but uncoded DPSK transmission.

The performance of the repeated Golay code is intended to be typical of that obtained with other moderate coding efforts using maximum likelihood detection. In particular, it should be comparable to a repeated 16-state convolutional code (of the same overall rate) with Viterbi decoding. It is our understanding that the fastest commercially available Viterbi decoders operate at about 20 Mb/s (with 64 states).

VII. ACKNOWLEDGMENT

My interest in, and exposure to, this subset of communication problems in fiber optics is due to several enjoyable and informative discussions with Paul Henry and Jack Salz.

REFERENCES

1. Paul S. Henry, "Lightwave Primer," IEEE J. Quantum Electron., *QE-21* (December 1985).
2. J. Salz, "Coherent Lightwave Communications," AT&T Bell Lab. Tech. J., this issue.
3. J. E. Mazo and J. Salz, "Spectra of Frequency Modulation With Random Waveforms," Inform. Contr., *9*, No. 4 (August 1966), p. 419.
4. R. F. Pawula, S. O. Rice, and J. H. Roberts, "Distribution of the Phase Angle Between Two Vectors Perturbed by Gaussian Noise," IEEE Trans. Commun., *COM-30* (August 1982), pp. 1828–41.
5. N. M. Blachman, "The Effect of Phase Error on DPSK Error Probability," IEEE Trans. Commun., *COM-29* (March 1981), pp. 364–5.

6. David Chase, "Code Combining," IEEE Trans. Commun., *COM-33* (May 1985), pp. 385–93.
7. I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*, Fourth ed. New York: Academic Press, 1965.

APPENDIX

Derivation of (22)

We begin by deriving a Fourier series for the angular distribution $g(\theta_1)$ of a vector $(A, 0)$ perturbed in each component by additive, zero-mean, independent and identically distributed Gaussian noise of variance σ^2. In what follows,

$$\rho = \frac{A^2}{2\sigma^2} \tag{27}$$

and

$$I_k(z) = I_{-k}(z) = \left(\frac{z}{2}\right)^k \sum_{j=0}^{\infty} \frac{(z^2/4)^j}{j!(j+k)!}, \qquad k = 0, 1, \ldots \tag{28}$$

are the modified Bessel functions. We have

$$\exp(z \cos t) = \sum_{k=-\infty}^{\infty} I_k(z)\cos kt. \tag{29}$$

Expressing the joint density of the components of $(A + n_1, n_2)$ in polar coordinates (r, θ_1) and integrating over the r variable after applying (29), we obtain, after a variable change,

$$\begin{aligned} g(\theta_1) &= \frac{\exp(-\rho)}{2\pi} \sum_{k=-\infty}^{\infty} \cos k\theta_1 \int_0^{\infty} \exp(-x) I_k(2\sqrt{\rho x})dx \\ &= \frac{1}{2\pi} + \frac{\exp(-\rho)}{\pi} \sum_{k=1}^{\infty} \cos k\theta_1 \int_0^{\infty} \exp(-x) I_k(2\sqrt{\rho x})dx, \\ &\quad -\pi \leqslant \theta_1 \leqslant \pi. \end{aligned} \tag{30}$$

In our problem, we have a second noisy vector that is also perturbed by rotation through ϕ, ϕ being zero-mean Gaussian. The modulo 2π angle θ_2 of this vector has density $h(\theta_2)$, where

$$h(\theta_2) = E_\phi g(\theta_2 - \phi), \qquad -\pi \leqslant \theta_2 \leqslant \pi, \tag{31}$$

E_ϕ being expectation with respect to ϕ. In (31), $g(\theta_2 - \phi)$ is evaluated by the periodic extension of $g(\theta)$. We are assuming here that the two consecutive chips are identical, but the error probability is the same when two consecutive chips have opposite signs.

Since θ_1 and θ_2 are independent, the difference angle $\psi = (\theta_2 - \theta_1) \bmod 2\pi$ has density $p(\psi)$ [see eq. [6] of Ref. 4], where

$$p(\psi) = E_\phi \int_{-\pi}^{\pi} g(\theta_1)g(\theta_1 + \psi - \phi)d\theta_1. \tag{32}$$

Finally, using the symmetry $p(\psi) = p(-\psi)$, the chip error probability is

$$p_c = 2 \int_{\pi/2}^{\pi} p(\psi)d\psi. \tag{33}$$

Performing the θ_1 and then the ψ integrations gives

$$p_c = \frac{1}{2} - \frac{2}{\pi}\exp(-2\rho) \sum_{s=0}^{\infty} \frac{(-1)^s}{2s+1} \cdot \left(\int_0^{\infty} dx \exp(-x) I_{2s+1}(2\sqrt{\rho x}) \right)^2 \exp\left(-\frac{(2s+1)^2\sigma_\phi^2}{2} \right). \tag{34}$$

A final use of eq. 6.614 (1) from Ref. 7 to evaluate the integral in (34) results in (22).

AUTHOR

James E. Mazo, B.S. (Physics), 1958, The Massachusetts Institute of Technology; M.S. (Physics), 1960, and Ph.D. (Physics), 1963, Syracuse University; Research Associate, Department of Physics, University of Indiana, 1963–1964; AT&T Bell Laboratories, 1964–1985; AT&T Information Systems, 1985—. At AT&T Bell Laboratories, Mr. Mazo had been concerned with theoretical problems in data transmission. He now supervises the Data Theory group at AT&T Information Systems.