

# Coding for Optical Channels

L. D. Baumert  
Arizona State University

R. J. McEliece  
University of Illinois

H. Rumsey, Jr.  
Communications Systems Research Section

*In a recent paper Pierce considered the problem of optical communication from a novel viewpoint, and concluded that performance will likely be limited by issues of coding complexity rather than by thermal noise. This paper reviews the model proposed by Pierce and presents some results on the analysis and design of codes for this application.*

## I. Introduction

In a recent paper Pierce (Ref. 1) considered the problem of optical communication from a novel point of view. He showed that for optical frequencies and low temperatures, the maximum signaling rate will be determined by coding complexity issues rather than by thermal noise. He exhibited one simple coding scheme (referred to below as a type 1 code), and challenged future workers to go further. This paper is a response to that challenge.

In Section II, we will describe the model Pierce arrived at as an appropriate description of the optical communication problem. It will be seen that Pierce's model is the familiar Z-channel. A complication arises because in Pierce's model the transmission of a "1" (which corresponds physically to the transmission of photons) is more costly than the transmission of a "0" (no photons). After some fairly simple analysis, we conclude that an appropriate figure of merit for a binary code  $\{x_1, \dots, x_M\}$  which is to be used on Pierce's channel is

$$Q = \left[ \frac{M \log M}{\sum_{i=1}^M w(x_i)} \right] d_a \quad (1)$$

where  $w(x_i)$  is the Hamming weight of the codeword  $x_i$ , and  $d_a$  is the minimum *asymmetric distance* of the code. (The term is defined precisely in Section II).

Calling the first term on the right side of Eq. (1)  $R_0$ , in Section III we show that  $R_0$  is largest for Pierce's type 1 code. In fact, we show that  $R_0 \leq [(n+1)/n] \log(n+1)$ , with equality only for type 1 codes. On the other hand, Pierce's codes all have  $d_a = 1$ , and in Section IV, we exhibit several codes with larger  $d_a$  which are in a certain sense superior to Pierce's codes.

Finally, in Section V, we prove the existence of a sequence of codes for which  $Q$  grows *linearly* with the block length  $n$ . (For Pierce's codes  $Q$  grows only logarithmically.)

## II. The Channel Model

The channel model arrived at by Pierce (Ref. 1) can be described as follows. At the transmitting end there is a light source and a shutter, and at the receiving end there is a photon counter. To transmit a binary sequence  $\mathbf{x} = (x_1, \dots, x_n)$  over this channel in  $T$  seconds, we divide the time interval into  $n$  equal segments of duration  $T/n$ ; we close the shutter during the  $i$ -th time interval if  $x_i = 0$ , and open it if  $x_i = 1$ . The receiver's estimate of  $x_i$  is 0 if no photons strike the photon counter during the appropriate interval, and 1 if one or more do. Assuming that the expected number of photons emitted by the light source during an interval of length  $T/n$  is  $\lambda$ , the probability that a transmitted 1 will be detected as a 0 is  $e^{-\lambda}$ , because of the Poisson nature of photon emissions. Of course if no photons are transmitted, none will be detected, and so this channel is the Z-channel depicted in Fig. 1. Furthermore it is assumed that there is a unit energy "cost" associated with the transmission of each photon. The basic coding problem here is to study the trade-off between the transmission rate measured in nats per photon) and the decoded bit error probability.

Suppose we wish to communicate over this channel using a binary code  $\{\mathbf{x}_1, \dots, \mathbf{x}_M\}$  of length  $n$  with  $M$  code words. The rate of the code is  $\log M$  nats<sup>1</sup> per code word. On the other hand the expected number of photons required by the  $i$ -th code word is  $\lambda w(\mathbf{x}_i)$ , where  $w(\mathbf{x}_i)$  denotes the Hamming weight of  $\mathbf{x}_i$ . Hence (provided each code word is transmitted with probability  $M^{-1}$ ) the average number of photons required per code word is

$$\frac{\lambda \sum_{i=1}^M w(\mathbf{x}_i)}{M}$$

Hence the transmission rate, measured in *nats per photon* is

$$R = \frac{1}{\lambda} \left[ \frac{M \log M}{\sum_{i=1}^M w(\mathbf{x}_i)} \right] \quad (2)$$

The quantity in brackets in Eq. (2) depends only on the code, and we call it the *asymmetric rate* of the code:

$$R_0 = \frac{M \log M}{\sum_{i=1}^M w(\mathbf{x}_i)} \quad (3)$$

<sup>1</sup>Throughout, all logarithms will be natural.

We have now defined the transmission rate of a code for our channel model. We now need a measure of the code's error correcting ability. To this end we are led to define the *asymmetric distance* between two binary  $n$ -tuples  $\mathbf{x} = (x_1, \dots, x_n)$  and  $\mathbf{y} = (y_1, \dots, y_n)$ . Let  $r$  denote the number of coordinates where  $x_i = 1$  and  $y_i = 0$ , and  $s =$  number where  $x_i = 0$  and  $y_i = 1$ . Then we define

$$d_a(\mathbf{x}, \mathbf{y}) = \max(r, s) \quad (4)$$

This distance plays a role for asymmetric errors analogous to that played by the Hamming distance  $d_H(\mathbf{x}, \mathbf{y}) = r + s$  for symmetric errors. The main result is the following.

**Theorem 1:** If the minimum asymmetric distance between distinct code words is  $d_a$ , then the code is capable of correcting any pattern of  $d_a - 1$  or fewer asymmetric errors. (N. B. An asymmetric error is an error of the type  $1 \rightarrow 0$ . The symmetry of the definition in Eq. (4) implies immediately the curious fact that any code capable of correcting  $d_a - 1$  " $1 \rightarrow 0$ " errors will also correct  $d_a - 1$  " $0 \rightarrow 1$ " errors.)

**Proof:** We begin by considering an example:

$$\begin{array}{l} \mathbf{x} = 1111110000000 \\ \mathbf{y} = 0000001111111 \end{array}$$

$\underbrace{\hspace{10em}}_r \quad \underbrace{\hspace{10em}}_s$

If  $\mathbf{x}$  is transmitted over the Z-channel of Fig. 1, can it be mistaken at the receiver as  $\mathbf{y}$ ? Clearly not, unless *each* of the 1's in  $\mathbf{x}$  is received as 0, since the presence of a 1 in the first  $r$  received components would immediately rule  $\mathbf{y}$  out. (The transition  $0 \rightarrow 1$  is impossible.) Hence, the smallest number of errors that could possibly cause confusion is  $r$ . If then  $\mathbf{x}$  is received with these  $r$  errors as  $\mathbf{x}' = 0000000000000$ ,  $\mathbf{x}'$  differs from  $\mathbf{x}$  in  $r$  positions and from  $\mathbf{y}$  in  $s$  positions. If the decoder picks the vector for which the number of disagreements is smallest, an error is possible only if  $r \geq s$  (and certain only if  $r > s$ ). The conclusion is that if  $\mathbf{x}$  is transmitted, an error is possible only if at least  $r$  errors occur, and  $r \geq s$ . Similarly if  $\mathbf{y}$  is transmitted, an error is possible only if at least  $s$  errors occur, and  $s \geq r$ . This shows that the code consisting of only the two code words  $\{\mathbf{x}, \mathbf{y}\}$  can correct any pattern of up to  $\max(r, s) - 1 = d_a(\mathbf{x}, \mathbf{y}) - 1$  asymmetric errors. Finally if the code has  $M$  code words  $\{\mathbf{x}_1, \dots, \mathbf{x}_M\}$ , and  $d_a = \min \{d_a(\mathbf{x}_i, \mathbf{x}_j) : i \neq j\}$ , the above argument shows that no pattern of  $d_a - 1$  or fewer asymmetric errors can possibly cause one transmitted code word to be mistaken for another.

Now consider the decoded error probability  $P_E$  of the code  $\{x_1, \dots, x_M\}$ , when it is used on the channel of Fig. 1.  $P_E$  is given by a complicated expression of the form

$$P_E = \frac{1}{M} \sum_{i=1}^M \sum_{z \in B_i} e^{-\lambda w(z)} (1 - e^{-\lambda})^{n-w(z)} \quad (5)$$

where  $B_i$  is the set of error patterns which causes decoder error, when  $x_i$  is transmitted. Theorem 1 implies that  $w(z) \geq d_a$  for all

$$z \in \bigcup_{i=1}^M B_i$$

and so in the limit as  $\lambda \rightarrow \infty$ , the sum on the right side of Eq. (5) is dominated by the terms of the form  $e^{-\lambda d_a}$ , i.e.,

$$\lim_{\lambda \rightarrow \infty} \frac{1}{\lambda} \log P_E = -d_a \quad (6)$$

where  $d_a$  is the code's asymmetric minimum distance. If now we define the parameter  $\gamma$  to be the number of photons per nat required by the code (this is a sort of normalized energy budget, analogous to the bit signal-to-noise ratio on the more familiar Gaussian channel), we get from Eq. (2) that  $\lambda = R_0 \gamma$ , and Eq. (6) becomes

$$\lim_{\gamma \rightarrow \infty} \frac{1}{\gamma} \log P_E = -R_0 d_a \quad (7)$$

Hence we are led to define the following quantity  $Q$  for any binary code, which is a measure of its asymptotic effectiveness when used on our photon counting channel:

$$Q = R_0 d_a \quad (8)$$

In summary: the bigger the code's "Q," the better we expect it to be.<sup>2</sup> In the next section we will show that  $R_0 \leq [(n+1)/n] \log(n+1)$  for any code of length  $n+1$ . In Section IV, we give some examples of binary codes with fairly large "Q." Finally in Section V we demonstrate that as a function of  $n$ , the code length, the best possible  $Q$  grows linearly with  $n$ .

<sup>2</sup>For the Gaussian channel, the corresponding number is  $R \cdot d_H$ , where  $R$  is the ordinary dimensionless rate of the code, and  $d_H$  is its minimum Hamming distance.

### III. An Upper Bound on $R_0$

In view of the definition in Eq. (8), it is clearly important to know how large the asymmetric rate  $R_0$  (see Eq. 3) can be. Theorem 2, below, shows that  $R_0$  can be no larger than  $(1+n^{-1}) \log(n+1)$ .

Thus let  $C = \{x_1, x_2, \dots, x_M\}$  be any binary code of length  $n$ . Define

$$R_0(C) = \frac{M \log M}{\sum_{i=1}^M w(x_i)}$$

**Theorem 2:**  $R_0(C) \leq [(n+1)/n] \log(n+1)$ , with equality if and only if  $C$  consists of the  $n+1$  words of weight  $\leq 1$ .

**Proof:** Let  $y_1, y_2, \dots$  be an ordering of the  $2^n$  binary vectors according to increasing weight:  $0 = w(y_1) \leq w(y_2) \leq \dots \leq w(y_{2^n}) = n$ . If  $C = \{y_i; i \in I\}$  is any code with  $|I| = M$  code words, then clearly the code  $C' = \{y_i; i \leq M\}$  must satisfy  $R_0(C') \geq R_0(C)$ . Thus, for the remainder of the proof we focus our attention on the codes  $C_M = \{y_1, y_2, \dots, y_M\}$  for  $M = 1, 2, \dots, 2^n$ . If we let  $\rho_M = R_0(C_M)$ , the assertion of the theorem is that  $\rho_M$ , as a function of  $M$ , is maximized for  $M = n+1$ .

Let  $\bar{M}$  be a value of  $M$  that maximizes  $\rho_M$ . Our first result is that if we define, for each  $k = 0, 1, \dots, n$ ,

$$M_k = \sum_{j=0}^k \binom{n}{j} \quad (9)$$

we must have  $\bar{M} \in \{M_0, M_1, \dots, M_n\}$ . To see this suppose that  $M_k < \bar{M} < M_{k+1}$ , let  $x = \bar{M} - M_k$ , and observe that  $\rho_M$  is given by

$$\rho_M = f(x) = \frac{(M_k + x) \log(M_k + x)}{w_k + (k+1)x} \quad (10)$$

where

$$w_k = \sum_{j=0}^k j \binom{n}{j}$$

Suppose  $\rho_M \geq \max(\rho_{M_k}, \rho_{M_{k+1}})$ . Then the function  $f(x)$ , viewed as a continuous function of the real number  $x$ , would have a maximum somewhere in the interval  $[0, \binom{n}{k+1}]$ , i.e.,  $f'(x) = 0$  and  $f''(x) \leq 0$ . But from Eq. (10) one easily sees that

$$f'(x) = \frac{1}{w_k + (k+1)x} \left\{ 1 + [w_k - (k+1)M_k] \frac{\log(M_k + x)}{w_k + (k+1)x} \right\} \quad (11)$$

From Eq. (11) it follows that  $f'(x) = 0$  can only occur if the equation

$$[(k+1)M_k - w_k] \frac{\log(M_k + x)}{w_k + (k+1)x} = 1 \quad (12)$$

is satisfied. One easily verifies, however, that if Eq. (12) is satisfied, then  $f''(x) = (M_k + x)^{-1} [w_k + (k+1)x]^{-1} > 0$ ; hence,  $f(x)$  has no maximum for  $x > 0$ . Thus  $\rho_M < \max(\rho_{M_k}, \rho_{M_{k+1}})$ , and we have shown that the largest value of  $\rho_M$  occurs for  $M \in \{M_0, M_1, \dots, M_n\}$ .

It remains to show that the maximum of  $\rho(C_{M_k})$  occurs at  $k = 1$ . To do this, define

$$T_k = M_k \log M_k \quad (13)$$

and let  $k \geq 1$  be an index that maximizes the function  $\rho(C_{M_k}) = T_k/w_k$ . Then in particular  $T_k/w_k \geq T_{k-1}/w_{k-1}$ ; substituting  $w_{k-1} = w_k - k \binom{n}{k}$  into this inequality we obtain

$$w_k(T_k - T_{k-1}) \geq k \binom{n}{k} T_k \quad (14)$$

But  $T_k - T_{k-1} = M_k \log M_k - M_{k-1} \log M_{k-1} = [\binom{n}{k} + M_{k-1}] \log M_k - M_k \log M_{k-1} = \binom{n}{k} \log M_k + M_{k-1} \log(M_k/M_{k-1})$ . Using the elementary inequality  $\log x \leq x - 1$ , we thus obtain  $T_k - T_{k-1} \leq \binom{n}{k} (1 + \log M_k)$ . Substituting this into Eq. (14) we obtain  $1 + \log M_k \geq k T_k/w_k$ . But since  $k$  is presumed to maximize the ratio  $T_k/w_k$ , it follows that  $T_k/w_k \geq T_1/w_1 = [(n+1)/n] \log(n+1) > \log n$ . Hence, from Eq. (14) follows

$$1 + \log M_k < k \log n \quad (15)$$

Equation (15) is a strong necessary condition on the optimizing parameter  $k$ ; it cannot be satisfied unless  $k = 1$  or  $2$ , or  $k = 3$  and  $n \leq 2$ , as we shall now see.

By a well-known inequality (Ref. 2), we have

$$\begin{aligned} \log M_k &\leq k \log \frac{n}{k} + (n-k) \log \frac{n}{n-k} \\ &= k \log n - k \log k + (n-k) \log \left( 1 + \frac{k}{n-k} \right) \end{aligned} \quad (16)$$

Again using the inequality  $\log x \leq x - 1$ , we get from Eq. (16)

$$\log M_k < k \log n - k \log k + k \quad (17)$$

It follows from Eq. (17) that  $1 + \log M_k < k \log n$ , provided  $k < k \log k - 1$ . This is true for all  $k \geq 4$ , and hence Eq. (15) is *not* true if  $k \geq 4$ .

If  $k = 3$  we compute directly that  $M_k = (n^3 + 5n + 6)/6$ , and hence Eq. (15) is false for  $k = 3$  and all  $n \geq 3$ .

Thus we have shown that for  $n \geq 3$ , the optimizing value for  $k$  must be  $k = 1$  or  $k = 2$ . (The verification that Theorem 1 is true for  $n = 1$  or  $2$  is trivial). We now conclude our proof of Theorem 1 by showing that  $R_0(C_{M_2}) < R_0(C_{M_1})$ . Since  $M_2 = (n^2 + n + 2)/2$ ,  $w_2 = n^2$ ,  $M_1 = n + 1$ ,  $w_1 = n$ , this is equivalent to

$$\frac{n^2 + n + 2}{2} \log \frac{n^2 + n + 2}{2} < n(n+1) \log(n+1) \quad (18)$$

For  $n \geq 1$ ,  $n^2 + n + 2 \leq (n+1)^2$ , and so the left side of Eq. (18) is upper bounded by

$$\begin{aligned} \frac{n^2 + n + 2}{2} \log \frac{(n+1)^2}{2} &= (n^2 + n + 2) \log(n+1) \\ &\quad - \frac{\log 2}{2} (n^2 + n + 2) \end{aligned}$$

This is less than the right side of Eq. (15) since

$$\begin{aligned} \frac{n^2 + n + 2}{2} \log(n+1) - \frac{\log 2}{2} (n^2 + n + 2) &- n(n+1) \log(n+1) \\ &= 2 \log(n+1) - \frac{\log 2}{2} (n^2 + n + 2) \end{aligned}$$

and  $\log(n+1)/(n^2 + n + 2) \leq (\log 2)/4$  for all  $n \geq 2$ . This completes the proof of Theorem 1.

## IV. Examples

**Example 1:** In the previous section we saw that the code of length  $n$  with 1 word of weight zero and  $n$  of weight 1 (hereafter called a *type 1* code) has the largest possible  $R_0$  among codes of length  $n$ . Clearly  $d_a = 1$  for these codes, and so from Eq. (8), the coding gain  $Q$  is given by

$$Q = \frac{n+1}{n} \log(n+1) \quad (19)$$

Surprisingly, it is quite difficult to find codes of length  $n$  for which  $Q$  is larger than this, for small values of  $n$ .

**Example 2:** Consider the extended (24, 12) Golay code. The average weight of its code words is 12 (this follows from the general theorem that any linear code of length  $n$  whose generator matrix has no zero columns has average weight  $n/2$ ). Hence from Eq. (3)  $R_0 = \log 2^{12}/12 = \log 2$ . It is well-known that the minimum Hamming distance of this code is 8, i.e., if  $r$  and  $s$  are as in Eq. (4),  $r + s \geq 8$ ; and hence  $d_a \geq 4$  for this code. In fact one can also show that  $d_a = 4$  for the Golay code, and so

$$Q = 4 \log 2 = 2.7726$$

Note that a type 1 code with  $n = 12$  has  $Q = (13/12) \log(13) = 2.7787$ , which exceeds that of the Golay code. The type 1 code with  $n = 24$  has  $Q = 3.3105$ , so that whatever desirable properties the Golay code may have in ordinary circumstances are certainly lost in the present application.

**Example 3:** Let  $C$  be the (128, 64) extended BCH code, with  $d_H = 22$ , hence  $d_a \geq 11$  (actually  $d_a = 11$ ). Then as above the average weight is 64 and so

$$Q = 11 \log 2 = 7.6246$$

This is better than the type 1 code with  $n = 128$ , which has only  $Q = 4.8978$ . Indeed, one needs  $n \geq 2041$  in order to exceed  $Q = 11 \log 2$  with a type 1 code.

## V. The Existence of Codes With Large $Q$

In the last example of Section IV, we saw that it is possible for the best code of length  $n$  to have a value of  $Q$  which is larger than  $(1 + 1/n) \log(n+1)$ . In fact, if we denote the largest possible  $Q$  for a code of length  $n$  by  $Q_n$ , we can show that  $Q_n$  grows *linearly* with  $n$ . Specifically, we shall show in this section that

$$0.052 \leq \liminf_{n \rightarrow \infty} \frac{Q_n}{n} \leq \limsup_{n \rightarrow \infty} \frac{Q_n}{n} \leq 1.39 \quad (20)$$

(Note: The implied logarithms in Eq. (20) are natural logarithms.)

First we derive the upper bound in Eq. (20). From Eq. (4) we know that the asymmetric distance  $d_a$  of a given code is less than or equal to the Hamming distance,  $d_H$ , and so by Eqs. (3) and (8) we have

$$Q \leq \frac{\log M}{\frac{1}{M} \sum_{i=1}^M w(x_i)} \cdot d_H \quad (21)$$

for any code  $\{x_1, x_2, \dots, x_M\}$  of length  $n$ . Let us view this code as an  $M \times n$  binary array, and let  $s_k$  denote the number of ones in the  $k$ -th column of this array. We now compute the sum  $\sum d(x_i, x_j)$  over all distinct pairs  $i < j$  in two ways. On one hand it is  $\geq \binom{M}{2} d_H$ , since  $d(x_i, x_j) \geq d_H$  for all  $i \neq j$ . On the other hand, a pair of entries  $(x_{ik}, x_{jk})$  in column  $k$  contribute 1 to the sum if and only if  $x_{ik} \neq x_{jk}$ . Thus

$$\begin{aligned} \binom{M}{2} d_H &\leq \sum_{i < j} d(x_i, x_j) = \sum_{k=1}^n s_k (M - s_k) \\ &= M \sum_{k=1}^n s_k - \sum_{k=1}^n s_k^2 \quad (22) \end{aligned}$$

Now

$$\sum_{k=1}^n s_k = \sum_{i=1}^M w(x_i)$$

and by Schwarz's inequality

$$\sum_{k=1}^n s_k^2 \geq \frac{1}{n} (\sum s_k)^2$$

Hence, if we denote

$$\frac{1}{nM} \sum_{i=1}^M w(x_i)$$

by  $\omega$ , Eq. (22) yields

$$\omega(1 - \omega) \geq \frac{M-1}{2M} \cdot \frac{d_H}{n} \quad (23)$$

Denoting the ratio  $d_H/n$  in Eq. (23) by  $\delta$ , and using the fact that  $(M-1)/M \leq 1$ , Eq. (23) gives

$$\omega \geq \frac{1 - \sqrt{1 - 2\delta}}{2} \quad (24)$$

Using Eq. (24) in Eq. (21), we obtain

$$Q \leq \frac{\log M}{n} \cdot \frac{2}{1 - \sqrt{1 - 2\delta}} \cdot \delta n \quad (25)$$

In Eq. (25) the term  $(\log M)/n$  is the rate of the code. If we denote by  $R(n, d)$  the rate of the largest code of length  $n$  and minimum Hamming distance  $d$ , and for  $0 \leq \delta \leq 1$

$$R(\delta) = \sup_{n \rightarrow \infty} \lim r(n, d_n)$$

where the "sup" is over all sequences  $(d_n)$  such that  $d_n/n \rightarrow \delta$ , it follows from Eq. (25) that

$$\limsup_{n \rightarrow \infty} \frac{Q_n}{n} \leq \sup_{\delta} \frac{2R(\delta)}{1 - \sqrt{1 - 2\delta}} \cdot \delta \quad (26)$$

Now the function  $R(\delta)$  is not precisely known, but using the well-known Plotkin bound

$$\begin{aligned} R(\delta) &\leq 1 - 2\delta & 0 \leq \delta \leq \frac{1}{2} \\ &= 0 & \delta \geq \frac{1}{2} \end{aligned}$$

we find that (let  $x = \sqrt{1 - 2\delta}$ )

$$\limsup_{n \rightarrow \infty} \frac{Q_n}{n} \leq \sup_{0 \leq x \leq 1} (x^2 + x^3) = 2 \quad (27)$$

Finally we make a small correction in Eq. (27). The rate of a code,  $R = (1/n)\log M$ , is usually defined for base 2 logarithms, and so our bound in Eq. (27) is a bound using base 2. Recalling that our assertions in Eq. (20) are base  $e$ , we must replace

$2 = \log_2(4)$  with  $\log_e(4) = 1.386$ . Hence we have, finally, one-half of Eq. (20), viz,

$$\limsup_{n \rightarrow \infty} \frac{Q_n}{n} \leq 1.386$$

It remains to prove the lower bound of Eq. (20). To do this we consider the class of *constant weight* codes. A code  $C = \{x_1, x_2, \dots, x_M\}$  is said to be a constant weight code of weight  $w$  if all code words have weight  $w$ . For such a code the formula for  $Q$  simplifies:

$$Q = \frac{\log M}{w} \cdot d_a \quad (28)$$

Or since we know that  $d_a \geq d_H/2$ ,

$$Q \geq \frac{\log M}{2w} d_H \quad (29)$$

Now according to the Gilbert bound for constant weight codes for any  $\delta$  and  $\alpha$  satisfying

$$0 \leq \delta \leq \frac{1}{2}$$

$$\frac{1 - \sqrt{1 - 2\delta}}{2} \leq \alpha \leq \frac{1}{2}$$

there exists a sequence of constant weight codes with (length, minimum Hamming distance, weight) =  $(n, d_n, w_n)$  such that  $d_n/n \rightarrow \delta$ ,  $w_n/n \rightarrow \alpha$  and with rates at least  $R(\delta, \alpha) = H(\alpha) - 2H(\delta/2\alpha) - (1 - \alpha)H[\delta/2(1 - \alpha)]$ , where  $H_2(x) = -x \log x - (1 - x) \log(1 - x)$  is the entropy function. Using this result, we see from Eq. (29) that

$$\liminf_{n \rightarrow \infty} \frac{Q_n}{n} \geq \frac{R(\delta, \alpha)}{\alpha} \cdot \frac{\delta}{2} \quad (30)$$

for all choices of  $\alpha$  and  $\delta$ . The maximum of the function on the right side of Eq. (30) can be found numerically. It is 0.052, and occurs at  $\delta = 0.06$ ,  $\alpha = 0.264$ . Hence

$$\liminf_{n \rightarrow \infty} \frac{Q_n}{n} \geq 0.052$$

as asserted in Eq. (20).

Of course the bounds in Eq. (20) are very far apart and it would be desirable to improve them. One obvious weakness in our technique is that we nowhere deal directly with the asymmetric distance  $d_a$ , but use instead the weak bounds

$d_H/2 \leq d_a \leq d_H$ . It would be highly desirable to develop techniques for constructing and analyzing codes with good asymmetric distance properties, for both theoretical and practical reasons.

## References

1. Pierce, J. R., "Optical Channels: Practical Limits With Photon Counting," *IEEE Trans. Comm.*, COM-26 (1978), in press.
2. Peterson, W. W., *Error-Correcting Codes*, MIT Press, 1961, Appendix A.

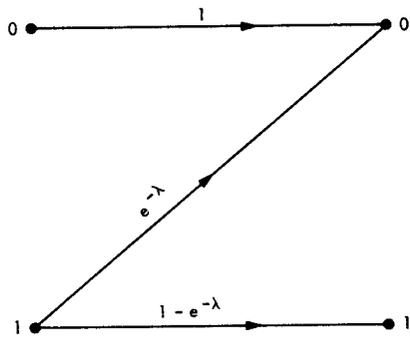


Fig. 1. The Z-channel