N 68- 29528

M-ary POISSON DETECTION
AND OPTICAL COMMUNICATIONS

by

Sherman Karp

Electronics Research Center
Cambridge, Mass.

and

Robert M. Gagliardi

University of Southern California
Los Angeles, Calif.

M-ary POISSON DETECTION

AND

OPTICAL COMMUNICATIONS

By Sherman Karp
Electronics Research Center
Cambridge, Mass.

and

Robert M. Gagliardi
University of Southern California
Los Angeles, Calif.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

M-ary POISSON DETECTION AND OPTICAL COMMUNICATIONS

By Sherman Karp
Electronics Research Center

and

Robert M. Gagliardi
University of Southern California

SUMMARY

This report presents an investigation of the problem of maximum likelihood detection of one of M Poisson processes in a background of additive Poisson noise. When the observables correspond to counts of emitted photoelectrons, the problem models a discrete version of a coherent M-ary optical communication system using photon counters in the presence of background radiation. Consideration is given to an average distance and a detection probability criterion. The advantages of an M-ary-pulsed intensity set (Poisson intensities wholly concentrated in a single counting interval) are demonstrated. The performance of such intensity sets is exhibited in terms of error probabilities, pulse widths, signal-to-noise ratio, and channel capacity. Behavior as a function of number of intensities M is also discussed. By appropriate conversion these latter results may be used for determining power requirements in an optical pulse position modulation system.

I. INTRODUCTION

The application of detection theory to the field of optical communications has been a subject of increasing interest. Since the output of a photodetecting surface is often modeled as sequences of electron "counts", and since optical photoelectrons have been readily accepted as obeying Poisson statistics, the analysis problem is basically one of signal detection involving Poisson processes. The problem was first formulated in this context by Reiffen and Sherman [ref. 1], and further contributions were made by Abends [ref. 2], Kailath [ref. 3], and Helstrom [ref. 4]. In this report the general problem of M-ary detection based upon observations of events described by a time-discrete Poisson process is investigated.

Though the formulation of the problem is of a general nature, the principal application is to optical communications, and the practical limits of such a system will govern the mathematical assumptions imposed. Consideration is given to the divergence criterion for detection and to a criterion of maximization of probability of detection, both readily accepted as suitable design objectives. The intensity set yielding optimal performance based on special cases of these criteria is shown to be a special type of orthogonal intensity set, composed of M disjoint intensities wholly concentrated in a counting interval. Previously, the superiority of this type of signal set in binary detection had been shown by Abends [ref. 2], based on a signal-to-noise ratio criterion, and by Kailath [ref. 3], based on distance criteria. This report represents an extension of these results to M-ary Poisson detection.

The formulation of the problem follows that of Reiffen and Sherman [ref. 1] . The occurrence of events over an observed interval ΔT is said to obey a Poisson process if the probability of exactly k (an integer) occurring is given by

$$P(k) = \frac{(n\Delta T)^k}{k!} \, e^{-n\Delta T} \; . \tag{1}$$

The parameter n is the average rate of occurrence and is called the intensity of the process. The average number of events occurring is then $n\Delta T$ and is often called the level of the process. If the events occur over a sequence of ΔT intervals in which the intensity may vary from one interval to the next, but is constant over each interval, we have a discrete time-varying Poisson process. In photodetection each event corresponds to the emission of an electron, which occurs upon arrival of a photon, each photon having a fixed energy. The level is therefore proportional to the average energy received per interval, while the intensity n is proportional to the average power (see section V). Thus, constraints upon level and intensity in Poisson processes are equivalent to energy and power constraints on the incident radiation.

In optical PCM communications, information is transmitted, as shown in Figure 1-a, by sending an optical signal intensity modulated with one of a set of possible intensities. The modulated signal is corrupted by fixed intensity background radiation during reception, resulting in a process, the intensity of which is the sum of each intensity. The output of the photodetecting surface at the receiver is then a time-varying Poisson process of electron counts having the received intensity. In an M-ary system the transmitter selects one of a set of M intensities for the optical process. The receiver, after photodetection, counts the number of electrons in each of M intervals ΔT seconds long and attempts to maximum likelihood detect which of M intensities is controlling the observed process. It is assumed that ΔT is suitably shorter than the inverse bandwidth of the intensities so that the intensity remains approximately constant over ΔT. In addition we assume the counting interval is exactly known at the receiver by a perfect synchronization link. Thus, the above system can be modeled by the block diagram in Figure 1-b. The input signal corresponds to a discrete Poisson process, while the interference appears as additive Poisson noise. (Recall that the sum of independent Poisson processes is itself a Poisson process having an intensity equal to the sum of the intensities.) The assumption of additive Poisson noise implies that the optical bandwidth is in general wider than the receiver bandwidth. This is referred to as "classical" counting. To obtain deviations from "classical" counting takes a considerable effort and requires very special circumstances not normally encountered. The model is idealized since other sources of interference, such as thermal noise and dark currents, are neglected. With this model the M-ary Poisson detection problem can be formulated as follows. Let a sequence of events obeying a discrete Poisson process occur over a sequence of M disjoint intervals ΔT where $M \, \Delta T = T$ and the count over each interval is independent of all others. Let the observed process be controlled by one of M possible intensity vectors $\underline{n}_q + \underline{n}_o$, q=1, 2, ... M, where
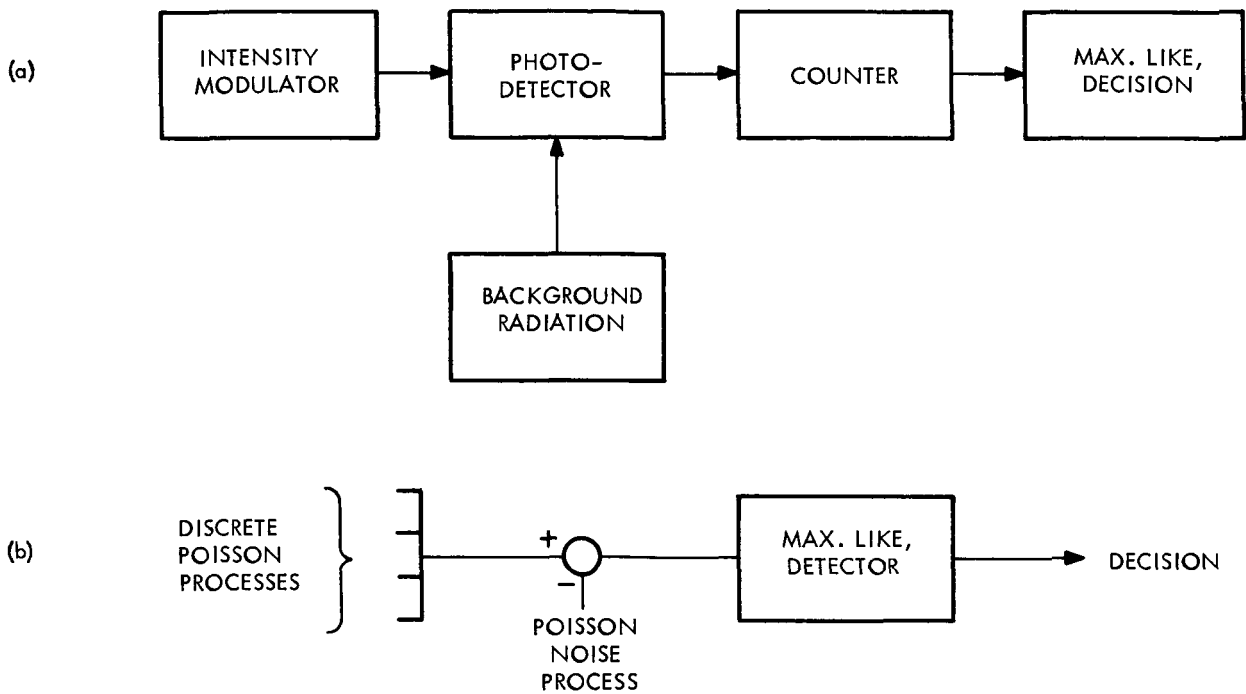
Figure 1. -- Optimum receiver

$$\underline{n}_q = \left\{ n_{q1}, \ n_{q2}, \ \ldots n_{qM} \right\} \tag{2}$$

$$\underline{n}_o = \left\{ n_o, \ n_o, \ n_o, \ \ldots n_o \right\}$$

$$n_{qi}, \ n_o, \ \geq 0 \quad .$$

The non-negative n_{qi} is thus the intensity of $\underline{n}_q$ during the ith interval. Under a fixed energy constraint for each signal, it is required that

$$\sum_{i=1}^{M} n_{qi} \Delta T = N, \text{ for all } q. \tag{3}$$

The intensity vector $\underline{n}_o$ represents constant intensity background noise superimposed upon the desired intensity. Let the corresponding number of events occurring in the ith interval be k_i. The problem then is to determine which of the possible intensity vectors $\underline{n}_q$ is controlling the received Poisson process by observing the sequence of independent counts $\underline{k} = \left\{ k_1, \ k_2, \ k_3, \ldots k_M \right\}$. Under a maximum likelihood detection criterion and a priori equilikely intensities, it is well known [ref. 1] that the optimal test is to form the likelihood functions:

3

$$\Lambda_q(\underline{k}) = \sum_{i=1}^{M} a_{qi} k_i \tag{4}$$

where

$$a_{qi} = \log\left[(n_{qi} + n_o) \Delta T\right] \tag{5}$$

and select $n_q + n_o$ as the true intensity if no other likelihood function exceeds $\Lambda_q(\underline{k})$. If a likelihood draw occurs (more than one $\Lambda_q(k)$ is maximum), it is known that any randomized choice amongst the maxima is equivalent. In the following a purely random selection in the case of likelihood draws is used. Equation (4) can be interpreted as a crosscorrelation of $\underline{k}$ with the a_{qi}'s, an operation readily performed by a digital crosscorrelator [ref. 1].

II. THE DIVERGENCE OF THE DETECTION TEST

The divergence between two intensities $\underline{n}_j$ and $\underline{n}_q$ of the above test is defined as

$$D_{jq} = E_{k/j}(\Lambda_{jq}) - E_{k/q}(\Lambda_{jq}) \qquad (6)$$

where

$$\Lambda_{jq} = \Lambda_j(\underline{k}) - \Lambda_q(\underline{k})$$

and $E_{k/j}(\Lambda)$ is the conditional average of Λ with respect to $\underline{k}$ given the intensity $\underline{n}_j$. Abends [ref. 2] has shown that for M = 2 (binary detection) and the condition $\underline{n}_2 = 0$, the divergence normalized by the variance of Λ is maximized by a "pulsed" type of intensity, where the level of the process is wholly concentrated in a single-count interval. Kialath [ref. 3] has extended this result by showing that, under a total energy constraint, other suitable forms of "distance" are maximized by similar "pulsed" intensities. We extend these notions here to the M-ary case and the equal energy constraint of Eq. (3).

The average divergence of an M-ary intensity set $\{\underline{n}_q\}$ will be defined as

$$\overline{D} = \frac{1}{M^2} \sum_j \sum_q D_{jq} \;\; . \qquad (7)$$

Since $E_{k/j}(k_i) = (n_{ji} + n_o)\Delta T$, the average divergence becomes

$$\overline{D} = \frac{\Delta T}{M^2} \sum_j \sum_q \sum_i (n_{ji} - n_{qi})\left\{\log\left[(n_{ji}+n_o)\Delta T\right]\right.$$

$$\left. - \log\left[(n_{qi}+n_o)\Delta T\right]\right\} = \frac{2\Delta T}{M^2} \sum_i \left[M \sum_j n_{ji}\log\left[1 + \frac{n_{ji}\Delta T}{K}\right]\right.$$

$$\left. - \sum_{j \neq q} \sum n_{ji}\log\left[\left(1 + \frac{n_{qi}\Delta T}{K}\right)\right]\right] \qquad (8)$$

where $K = n_o\Delta T$. The non-negativeness of the n_{ji} and n_o allows one to write:

$$\overline{D} \le \frac{2\Delta T}{M} \sum_i \sum_j n_{ji} \log \left[1 + \frac{n_{ji}\Delta T}{K} \right]$$

$$\le \frac{2}{M} \left[\max_{i,j} \log\left(1 + \frac{n_{ji}\Delta T}{K}\right) \right] \sum_j \sum_i n_{ji} \Delta T$$

$$= 2N \log \left(1 + \frac{N}{K}\right) \tag{9}$$

as an upper bound under the constrain of Eq. (3). However, the first equality holds if the second term in Eq. (8) is zero, requiring n_{ji}, $j \ne q$ to be zero for all i at which n_{qi} is non-zero; that is, the intensities must be mutually disjoint. The second equality in Eq. (9) holds if $n_{ji} = N$ for one i and $n_{ji} = 0$ for all other i. Thus, the upper bound for $\overline{D}$ occurs if the intensities of the set are disjoint and wholly concentrated in single counting interval. This is satisfied with the set

$$\underline{N}_q = \left\{ N/\Delta T \, \delta_{iq} \right\}, \quad q = 1, 2, \ldots M \tag{10}$$

where δ_{iq} is the Kronecher delta. The above represents an M-ary "pulsed" intensity set with each intensity occupying one of M intervals. It is significant to note that any disjoint intensity set, no matter how many intervals are used, yields the bound of the first inequality of Eq. (9), but only the pulsed intensity set of Eq. (10) yields the second bound. Thus, of all disjoint intensity sets, only the pulsed set maximizes $\overline{D}$, which immediately implies only M intervals are required for maximization with M intensities. Lastly, it may be noted that with an average energy constraint over all intensities:

$$\frac{1}{M} \sum_j \sum_i n_{ji}\Delta T = N \tag{11}$$

instead of Eq. (3), $n_{ji} \le MN/\Delta T$ results and Eq. (9) becomes

$$\overline{D} \le 2N \log \left[1 + \frac{MN}{K} \right] \tag{12}$$

which exceeds that previously derived. Furthermore, the upper bound in Eq. (12) occurs when M - 1 intensities are zero everywhere, and one intensity is a pulsed intensity having a value $MN/\Delta T$.

6

III. DETECTION PROBABILITY

The optimality of the M-ary pulsed intensity set has been shown, based on a divergence criterion. This section will show that in certain cases this superiority also extends over a criterion based on maximization of the detection probability. First an expression for the detection probability for a general intensity set $\{n_q\}$ is required. Usually this is obtained by first writing the conditional probability density of $\Lambda_q(k)$, and then integrating over regions of correct decisioning. However, $\Lambda_q(k)$ in Eq. (4) is a weighted sum of independent Poisson variates which in general is $\underline{not}$ a Poisson variable. Rather, the true density involves an M-fold convolution of modified Poisson densities, yielding a result that is difficult to integrate. Instead an alternative expression for the detection probability is used. It was derived in the appendix and has the form:

$$P_D = \frac{e^{-N}}{M} \sum_{R^M} \max_q \left\{ \Psi(q, \underline{j}) \right\} \tag{13}$$

where N is the intensity energy constraint of Eq. (3), R^M is the space of all M-dimensional vectors $\underline{j}$ having non-negative integer components, and

$$\Psi(q, j) = \prod_{i=1}^{M} \frac{\left[(n_{qi} + n_o)\Delta T \right]^{j_i}}{j_i!} \exp\left[-Mn_o\Delta T \right] . \tag{14}$$

The derivation of Eq. (13) follows an analogous procedure used in Gaussian channels (see Viterbi $\{$ref. 5$\}$), but is somewhat complicated by the fact that likelihood draws occur with non-zero probability.

Determination of the intensity set $\{n_q\}$ for which P_D is maximum is of interest. This has been obtained for two particular cases of interest:

Case I — M = 2 and symmetric intensity sets: -

Let M = 2 and consider the set of all possible symmetric intensity sets; i.e., if $n_1 = \{a, b\}$, then $n_2 = \{b, a\}$. For this case it is easy to show that for any intensity set of this type, the vectors $\underline{j}$ for which $\Psi(1, \underline{j}) \leq \Psi(2, \underline{j})$ when $a > b$ is simply the set $\underline{j} = \{j_1, j_2\}$ such that $j_1 \leq j_2$. Using the constraint of Eq. (3) and letting $n_1 = \{a, N-a\}$ and $n_2 = \{N-a, a\}$, $\frac{N}{2} < a \leq N$, the detection probability is:

$$P_D = \frac{e^{-(N+K)}}{2} \left\{ \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{j_1-1} \frac{(a+K)^{j_1}}{j_1!} \cdot \frac{(N-a+K)^{j_2}}{j_2!} \right.$$

$$\left. + \sum_{j_1=0}^{\infty} \sum_{j_2=j_1}^{\infty} \frac{(N-a+K)^{j_1}}{j_1!} \cdot \frac{(a+K)^{j_2}}{j_2!} \right\} \qquad (15)$$

where again $K = n_o \Delta T$. Differentiating with respect to "a" yields

$$\frac{dP_D}{da} = \frac{e^{-(N+K)}}{2} \left\{ \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{j_1-1} \left[\frac{A^{j_1-1}}{(j_1-1)!} \cdot \frac{B^{j_2}}{j_2!} \right.\right.$$

$$\left. - \frac{A^{j_1}}{j_1!} \cdot \frac{B^{j_2-1}}{(j_2-1)!} \right]$$

$$+ \sum_{j_1=0}^{\infty} \sum_{j_2=j_1}^{\infty} \left[\frac{B^{j_1}}{j_1!} \cdot \frac{A^{j_2-1}}{(j_2-1)!} \right.$$

$$\left.\left. - \frac{A^{j_2}}{j_2!} \cdot \frac{B^{j_1-1}}{(j_1-1)!} \right] \right\} =$$

$$\frac{e^{-(N+K)}}{2} \left\{ 1 + \sum_{j_1=1}^{\infty} \sum_{j_2=j_1-1}^{j_1} \frac{A^{j_1}B^{j_2} + B^{j_1}A^{j_2}}{j_1! \, j_2!} \right\}$$

where $A = (a + K)$ and $B = (N - a + K)$. Since A and B are positive, the above substantiates P_D as a monotone increasing function of a. Therefore P_D is maximized with "a" having its maximum value, $a = N$, corresponding to the pulsed intensity set of (10) with $M = 2$.

8

Case II -- Any M, N/K → 0

The limit above implies a high background noise level situation. We observe here that

$$P_D = \frac{e^{-N}}{M} \sum_{R^M} \max_q \Psi(q,\underline{j})$$

$$= \frac{e^{-N}}{M} \sum_{R^M} C \max_q \left\{ \exp \left[\sum_{i=1}^{M} j_i \ln \left[1 + \frac{n_{qi}\Delta T}{K} \right] \right] \right\} \quad (16)$$

$$\xrightarrow[\frac{N}{K} \to o]{} \frac{e^{-N}}{M} \sum_{R^M} C \max_q \left\{ \exp \sum_{i=1}^{M} j_i n_{qi} \Delta T/K \right\}$$

where $C = e^{-MK} \prod_i (1/j_1!)$ and the limit follows since $n_{qi}\Delta T/K \leq N/K \to 0$.

Now, with the constraint of Eq. (3):

$$\sum_i j_i n_{qi} \Delta T/K \leq j_{max} N/K \quad (17)$$

where

$$j_{max} = \max_i \left\{ j_i \right\}.$$

Thus:

$$P_D \leq \frac{e^{-N}}{M} \sum_{R^M} C \exp \left\{ j_{max} N/K \right\}, \quad \frac{N}{K} \to 0$$

The upper bound occurs when

$$\max_q \left\{ \exp \sum_{i=1}^{M} j_i n_{qi} \Delta T/K \right\} = \exp \left\{ j_{max} N/K \right\} \quad (18)$$

which clearly is true for the pulsed intensity set of Eq. (10), signifying asymptotic optimality for any M.

To determine optimal intensity sets (either global or local) in the general case, using Eq. (13), still remains an elusive task. It has been conjectured by many (e.g., see refs. [1] and [3]) that the pulsed intensity set is, in fact, the true optimal set, but to the authors' knowledge a rigorous proof has not been shown.

IV. ERROR PROBABILITIES WITH PULSED INTENSITY SETS

In this section the performance of the pulsed intensity set in M-ary detection is investigated by evaluating the error probability $P_E = 1 - P_D$. This can be obtained by using Eq. (13), but the computation can be more conveniently handled by noting that for the pulsed intensity set of Eq. (10) $\{\Lambda_q\}$ of Eq. (4) constitutes a set of independent, Poisson random variables. The variable Λ_q has level $(N + K)$ if the qth intensity is sent, and has level K otherwise $(K = n\Delta T)$. Recall, if the qth intensity is sent, a correct decision will be made with probability $1/r + 1$ if Λ_q equals r other Λ's and exceeds the remaining $M - 1 - r$. Therefore, when all possibilities are considered, the conditional detection probability is

$$P_{D/q} = \frac{e^{-(N + MK)}}{M} + \sum_{r=0}^{M-1} \sum_{x=1}^{\infty} \left[\frac{(N + K)^x}{x!} e^{-(N + K)} \right]$$

$$\left\{ \sum_{t=0}^{x-1} \frac{K^t}{t!} e^{-K} \right\}^{M - 1 - r} \left[\frac{K^x}{x!} e^{-K} \right]^r$$

$$\left[\frac{(M - 1)!}{r! \, (M - 1 - r)! \, (r + 1)} \right] . \tag{19}$$

The right side is independent of q and thus represents the average detection probability. By applying the identity:

$$\sum_{r=0}^{M-1} \frac{(M - 1)!}{(r + 1)! \, (M - 1 - r)!} A^{M-1-r} B^r = \frac{A^{M-1}}{M\left(\frac{B}{A}\right)} \left[\left(1 + \frac{B}{A}\right)^M - 1 \right],$$

one can rewrite the error probability as

$$P_E(N, K, M) = 1 - P_D = 1 - \frac{e^{-(N+MK)}}{M} - \sum_{x=1}^{\infty} \left[\frac{(N + K)^x \, e^{-(N+K)}}{x!} \right]$$

$$\left\{ \sum_{t=0}^{x-1} \frac{K^t e^{-K}}{t!} \right\}^{M-1} \cdot \frac{1}{Ma} \left[(1 + a)^M - 1 \right] \tag{20}$$

where

$$a = \frac{K}{x! \left[\sum_{t=0}^{x-1} \frac{K^t}{t!} \right]}$$

The parameter P_E (N, K, M) has been digitally computed by the authors for various values of N, K, and M. An exemplary plot is shown in Figure 2 in which P_E (N, 3, M) has been plotted for various M as a function of N. It is of paramount significance to note that P_E depends on both signal energy N and noise energy K in the counting interval, and not simply on their ratio. This fact is emphasized in Figure 3 in which P_E (N, K, 2) is plotted as a function of K for two fixed ratios N/K. This dependence on both signal and noise energies distinguishes the Poisson detection problem from the analogous coherent gaussian channel problem. Note that the interfering noise energy K depends only upon the background energy in the interval ΔT, which is the width of the transmitted intensity pulse. The prime advantage of optical systems is precisely their ability to remove the effect of background noise by making ΔT small, and has been emphasized in the literature (ref. 6). * This fact can be illustrated graphically, using Eq. 20, by considering a binary Poisson channel (M=2) sending information at a rate 1/T bits/sec. The effect of the parameter ΔT is indicated by plotting $P_E(N, n_o T \Delta T/T, 2)$ as a function of $\Delta T/T$, for fixed intensity energy N and background noise energy per bit interval $n_o T$. This is shown in Figure 4. The results indicate the continuous improvement obtained by decreasing the "duty cycle" $\Delta T/T$, the ultimate limit corresponding to $\Delta T = 0$. The improvement, of course, is made at the expense of information bandwidth and peak power (both inversely proportional to ΔT). Surprisingly, the improvement is quite small at low values of N, and the increase in bandwidth may not be worth the improvement obtained in error probability.

A quantity of particular interest to communication engineers is the detected "signal-to-noise ratio". This is often denoted [ref. 7] as the ratio of the square of average electron count with no noise to the variance of the count when noise is present. For Poisson counting statistics with pulsed intensities this becomes $(S/N) = N^2/(N+K)$. The behavior of P_E of Eq. (20) as a function of K for fixed (S/N) is illustrated in Figure 5 for a binary system. The results again indicate the ambiguity in using S/N as a design criterion. The asymptotes show the wide functional variation of P_E as K increases from zero.

As illustrated in Figure 2, the error probabilities increase as M increases. However, the use of a single set of curves to compare various M-ary systems is misleading. An M-ary system with ΔT second counting intervals transmits $\log_2 M$ bits of information in $M \Delta T$ seconds. It therefore communicates at a rate

$$R = \frac{\log_2 M}{M \Delta T} \quad \text{bit/sec.} \tag{21}$$

---

*Also Karp, S., and Gagliardi, R. : A Low Duty Cycle Optical Communication System. Presented at Eastcon Convention, Washington, D. C., October 1967.
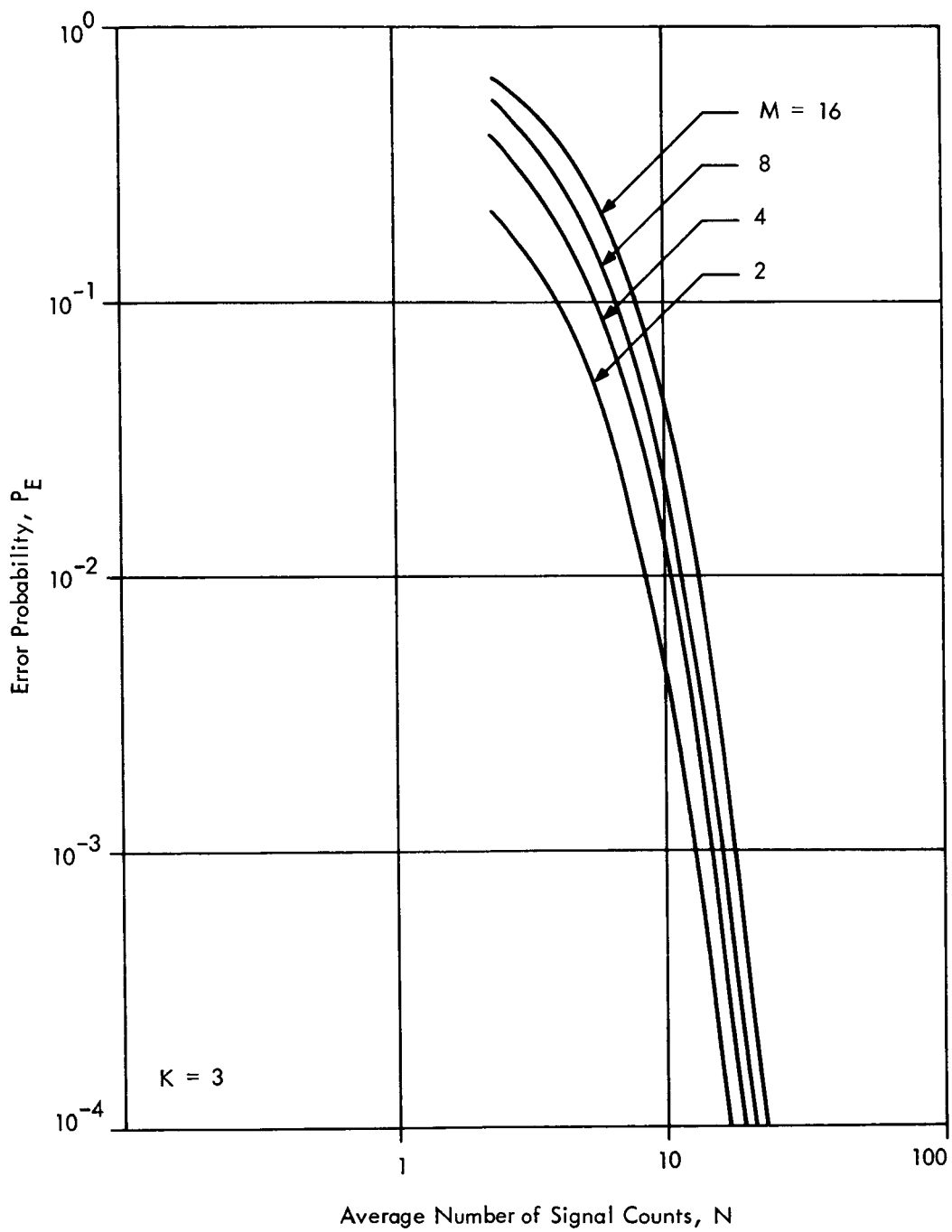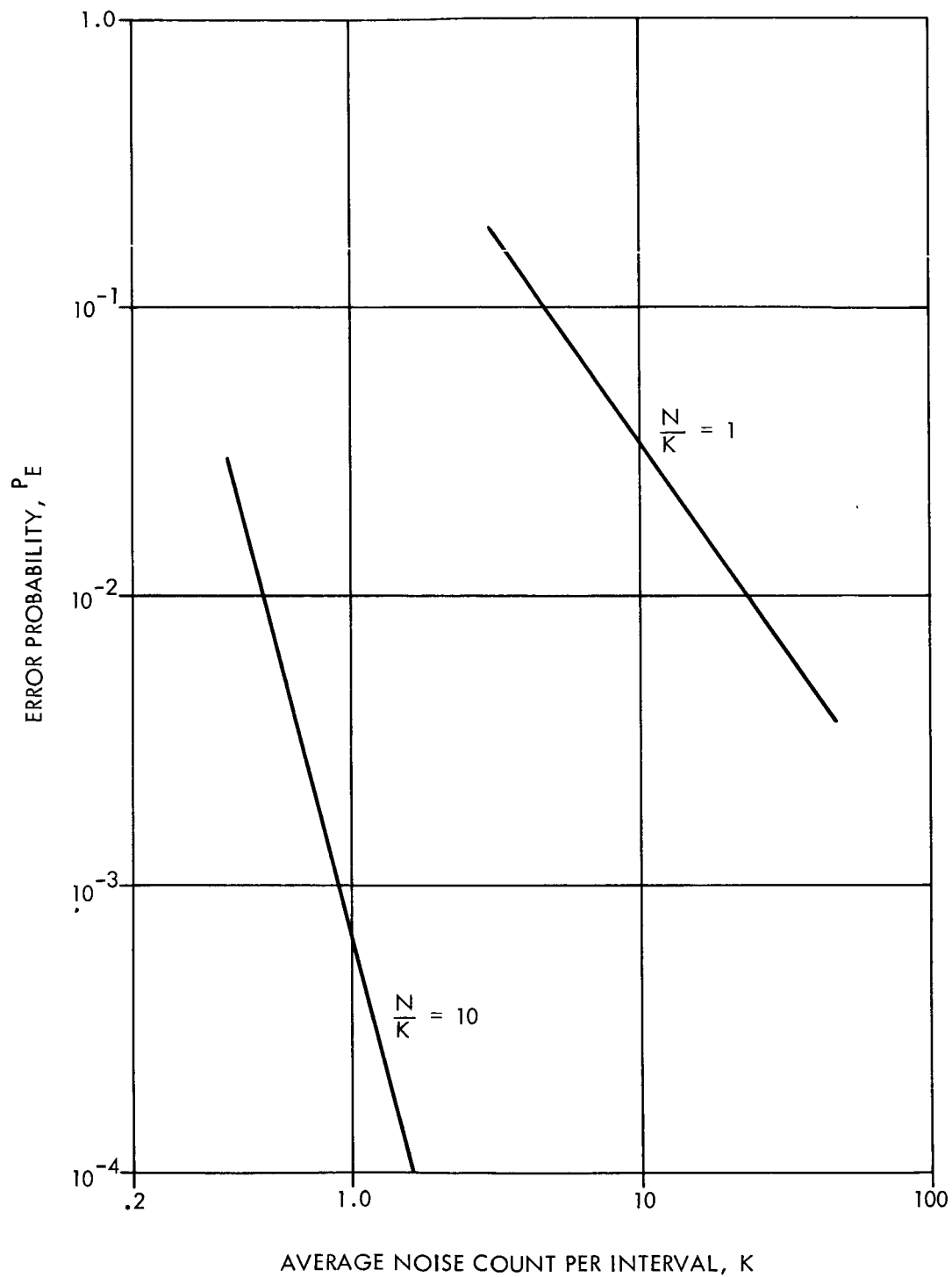
Figure 2. -- Error probability versus M

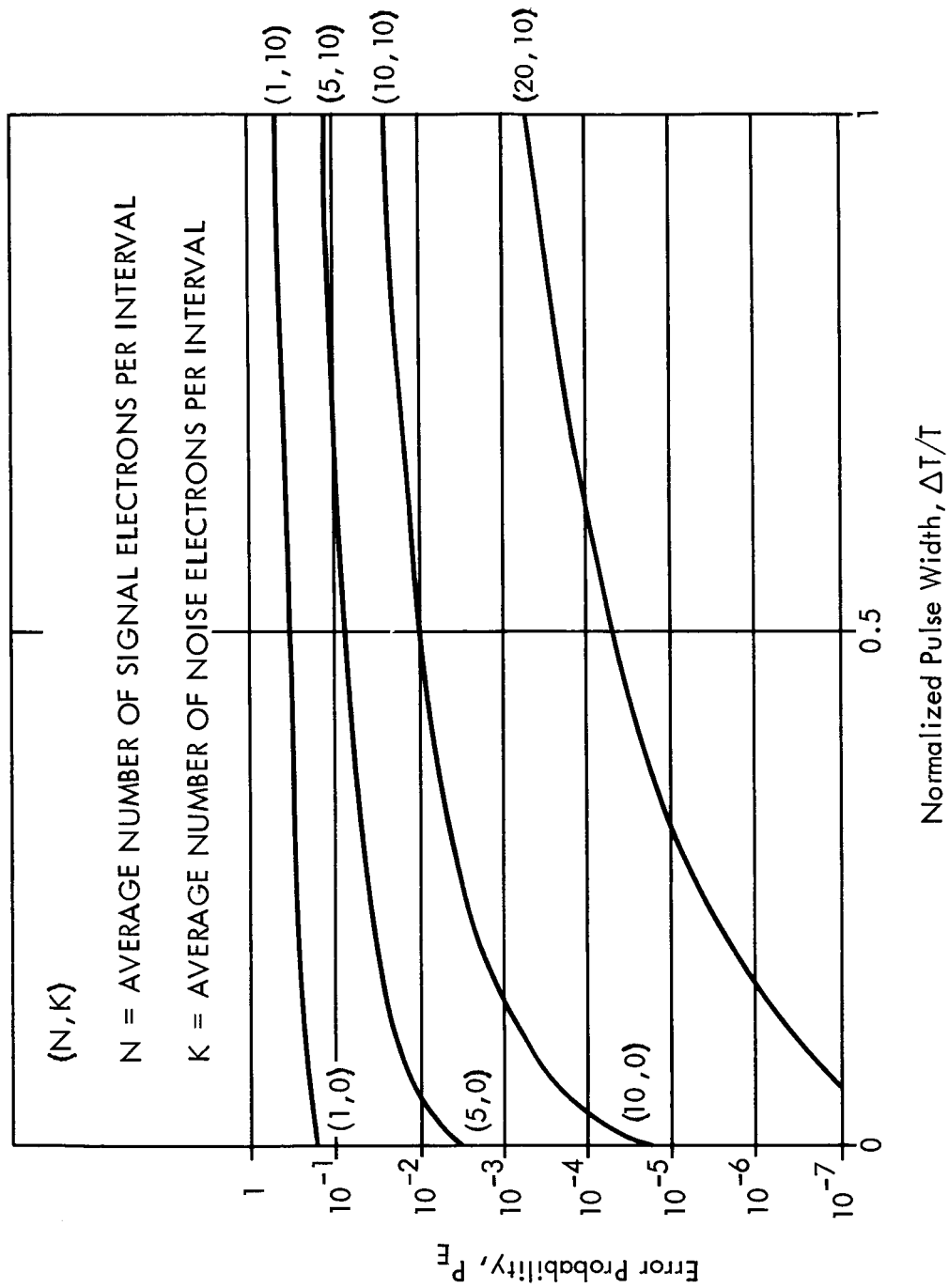Figure 3. -- Comparison of error probabilities for fixed ratio of average signal to average noise counts

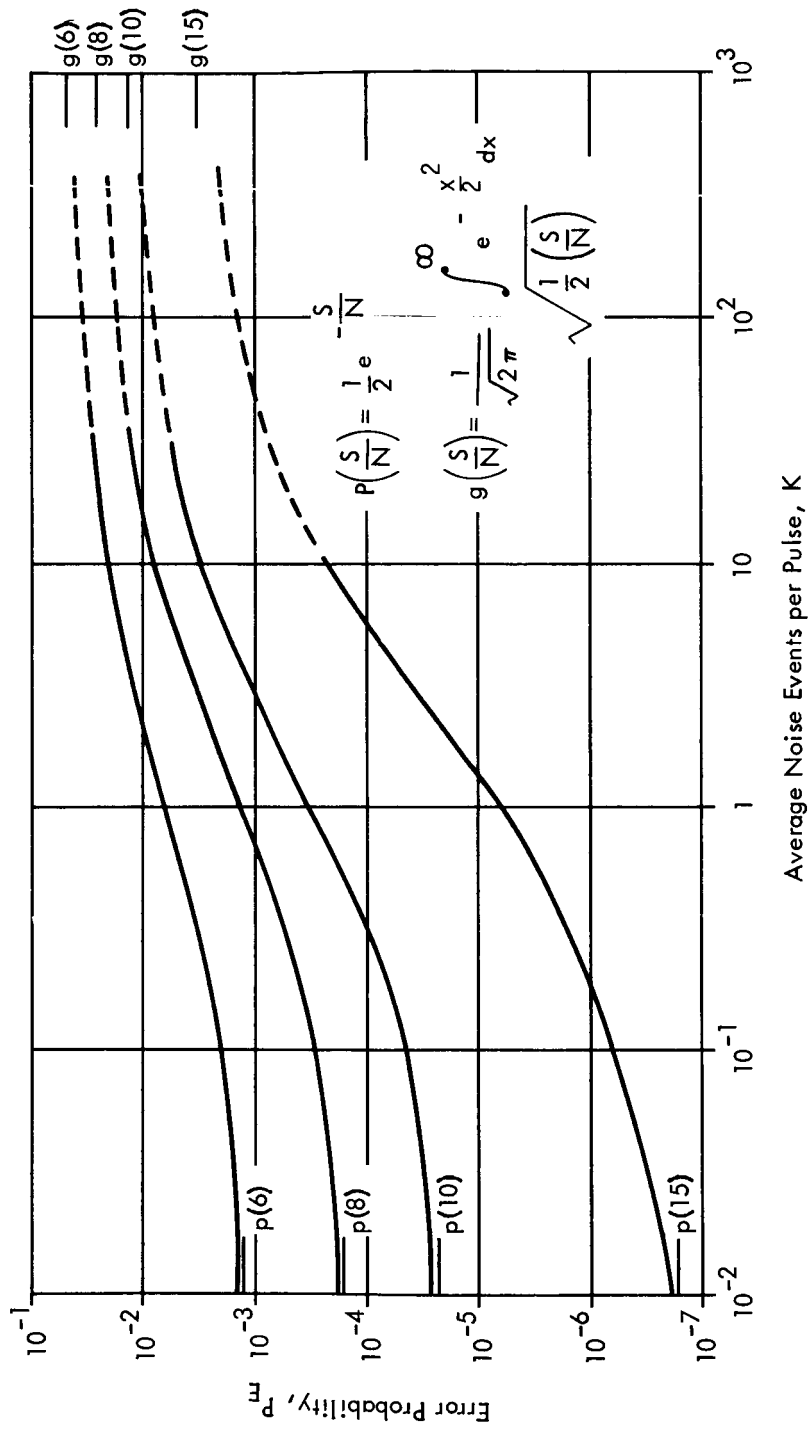Figure 4. -- Error probability versus normalized pulse width

g(6)
g(8)
g(10)
g(15)

$$P\left(\frac{S}{N}\right) = \frac{1}{2} e^{-\frac{S}{N}}$$

$$g\left(\frac{S}{N}\right) = \frac{1}{\sqrt{2\pi}} \int^{\infty}_{\sqrt{\frac{1}{2}\left(\frac{S}{N}\right)}} e^{-\frac{x^2}{2}} dx$$

p(6)
p(8)
p(10)
p(15)

Error Probability, P_E

Average Noise Events per Pulse, K

Figure 5. -- Probability of error for fixed signal-to-noise ratio versus noise background

If the transmission rate is normalized for each M, ΔT must be readjusted to maintain a fixed rate $R = R_o$. The effective noise level per counting interval is then

$$n_o \Delta T = n_o \log_2 M / MR_o = \left(\frac{n_o}{2R_o}\right) \frac{2\log_2 M}{M} = 2K_o \log_2 M / M \quad (22)$$

where K_o is the noise energy in an interval $1/2R_o$. Thus, for a comparison of different M systems, each with fixed information rates, one should compare the parameter $P_E(N, 2K_o \log M / M, M)$ for each N. If this adjustment is made using Eq. (20), the curves of Figure 6 are generated, with $K_o = 1$.

The curve corresponding to $M = \infty$ is also shown, and is determined by taking the limit of $P_E(N, 2K_o \log M / M, M)$ as $M \to \infty$. This can be obtained by replacing K in Eq. (20) by $K' = 2K_o \log M / M$ and noting:

$$\lim_{M \to \infty} e^{-(N+MK')} \to 0$$

$$\lim_{M \to \infty} \frac{(1+a)^M - 1}{Ma} \to 1 \text{ for } x \geq 1$$

$$\lim_{M \to \infty} \left[\sum_{t=o}^{x-1} \frac{(K')^t}{t!} e^{-K'}\right]^{M-1} \longrightarrow \begin{cases} 0 \text{ for } x = 1 \\ 1 \text{ for } x > 1 \end{cases}.$$

Using the above results in

$$\lim_{M \to \infty} P_E(N, 2K_o \log M / M, M) \to 1 - \sum_{x=2}^{\infty} \frac{(N)^x e^{-N}}{x!}$$

$$= 1 - (1 - Ne^{-N} - e^{-N}) = (1 + N)e^{-N} \quad (23)$$

which is plotted as $M = \infty$ in Figure 6. It is noteworthy that Eq. (23) is precisely the probability of an event count of zero or one occurring in a noiseless counting interval of signal energy N. This has the following interpretation. As $M \to \infty$ the number of intervals becomes infinite, but the normalized noise energy per interval, $K' = 2K_o \log M / M$, approaches zero. The probability that more than one event will occur in any one of M - 1 independent non-signalling intervals having noise energy K' is given by
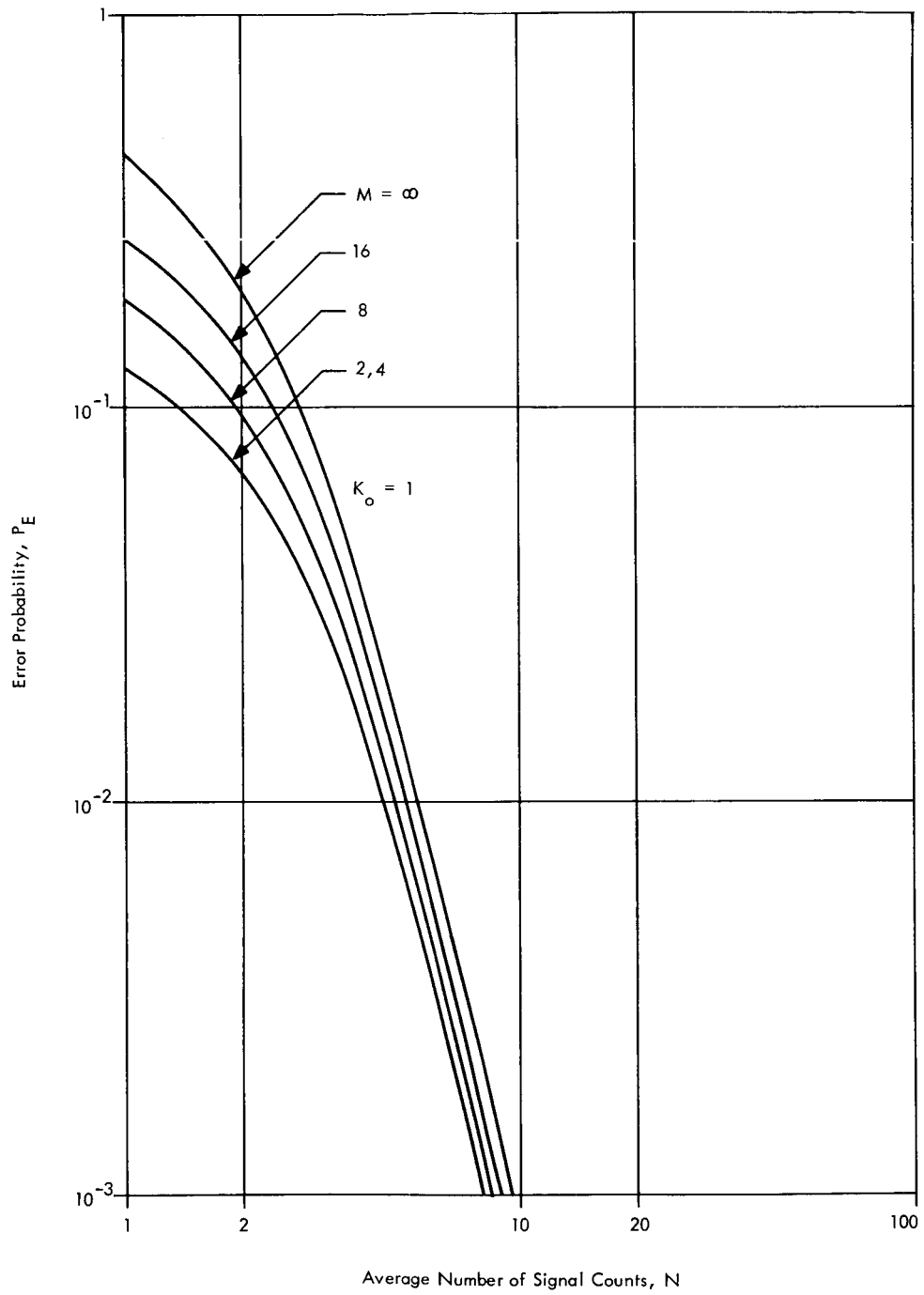
$$1 - \left(1 + K'e^{-K'}\right)^{M-1}.$$

Figure 6. -- Normalized error probability versus M

This approaches zero as $M \to \infty$, indicating that counts of zero or one will occur in every such interval with probability one. Furthermore, there will be an infinite number of intervals with a zero count and with a one count. Therefore, as $M \to \infty$, an error will occur (with probability approaching one) whenever the signalling interval has a count of zero or one, and an error will never occur when the latter interval has a count greater than one; hence, Eq. (23).

It is also interesting to note in Figure 6 that best system operation, in terms of minimal error probability, does not always correspond to $M \to \infty$. In fact, it can be shown that best M operation depends strongly on the amount of background noise K_o. For example, if $K_o \approx 0$, it is easy to show [using Equation (20)] that for M finite $P_E(N, K', M) \approx (M - 1)e^{-N}/M$, which is monotonically decreasing with M and always less than the $M = \infty$ value of Eq. (23). Thus, with negligible background noise, system operation improves with decreasing M and is best for $M = 2$. Physically, this means the noise reduction advantages due to decreasing ΔT as M increases does not offset the increasing errors due to the larger numbers of likelihood draws that will occur. (Recall a random choice is made in the event of draws.) For large amounts of background noise, however, the converse is true, and $M = \infty$ does yield minimal error probability.

It should be emphasized that a fixed energy constraint was imposed on the signal intensity, and therefore the time average power $P_o = N/T = NR/\log M$ actually approaches zero as $M \to \infty$. If the average power level of the source had been fixed at some level P_o, then N must be replaced by $P_o \log M/R$ in the previous equations, and it can be seen that $P_E \to 0$ as $M \to \infty$ for any $P_o > 0$. This result may be compared to a similar result for an additive gaussian channel (ref. 8) in which zero error probability occurred only if P_o satisfied a condition dependent on the rate R.

The P_E results above are useful for determining the channel capacity (maximum information rate) of an M-ary pulsed intensity set. Assume a transmitter sends one of a group of M-pulsed intensities every T second, with each pulse having width $\Delta T = T/M$. If the transmitter operates at a fixed rate R_o, then again $T = \log M/R_o$ as given by Eq. (21). The channel can now be represented as a symmetric channel in which each of the M equal likely intensities is converted to itself with probability $1 - P_E$, and is converted to any of the other intensities with equal probability $P_E/M-1$. The channel capacity for this type of system is known to be

$$C = \frac{\log M + P_E \log \frac{P_E}{M - 1} + (1 - P_E) \log (1 - P_E)}{\log M/R_o} \qquad (24)$$

where $P_E = P_E(N, n_o \log M/MR_o, M)$. Again the signal intensity energy N, the background noise power n_o, and the rate R_o to be held fixed are considered. Then, as $M \to \infty$, P_E approaches the limit in Eq. (23), while the channel capacity has the limit

$$C \to [1 - (1 + N)e^{-N}] R_o \qquad (25)$$

for N finite. The above indicates information transfer can be forced to approach any desired rate with a finite signal energy by using an increasingly larger number of intensities and adjusting R_o at the transmitter. However, each level is transmitted with a non-zero error probability, and the information bandwidth and peak power become infinite. Again introduction of a transmitter power constraint, instead of an energy constraint, will yield operation at a capacity R_o with a zero error probability, as $M \rightarrow \infty$.

V. CONCLUSIONS

This report has investigated the M-ary Poisson detection problem, defined as maximum likelihood detection of one of a set of M discrete Poisson processes in the presence of an additive discrete Poisson noise process. The problem model represents a discrete version of an optical communication system in which the observables represent counts of photoelectrons, the signals represents intensity modulated CW optical sources, and the noise represents background radiation received within the optical bandwidth. The photoelectron count can then be modeled as a time-varying Poisson process, the average rate of which is proportional to the sum of the intensities of the modulated source and the background radiation. In practical operation the intensity of the optical source is a continuous process, but the analyses may be put on a discrete basis by partitioning the signalling intervals into subintervals over which the intensity is taken to be constant. The above Poisson model is examined, and the advantages of a pulsed type of intensity set is demonstrated. The latter corresponds to an optical system using pulse position modulation in which information is transmitted by a burst or pulse of optical energy located in one of a set of pulse positions. The performance of such a system, in terms of pulse width and numbers of pulse positions, is presented herein. The results of this report basically represent theoretical limits to which an optical link can approach, since the deleterious effects of receiver (thermal) noise have been neglected. This latter assumption becomes valid, for example, when photomultipliers are used in detection, and the background radiation collected at the receiver is the predominant source of noise.

The analyses and performance results are in terms of N and K, the average electron counts due to signal and noise, respectively. However, these results may be easily converted to average power requirements using the relations

$$N = \eta P_s M/hfB$$

$$K = \eta P_n/hfB$$

where P_s and P_n are the average signal and background noise power, h = 6.62 x 10^{-34}, η is the photodetector efficiency (including photomultiplication), f is the optical frequency of the CW source, and B = $1/\Delta T$. The average power P_s and P_n can be further converted to transmitted power by introducing space losses and receiver optics (e.g., ref. 9, Ch. 1, 2). Exact synchronization has been assumed here between transmitter and receiver at all times. In addition to receiver thermal noise, the analysis has excluded the effects of photomultiplier statistics, saturation, and dark currents.

REFERENCES

1. Reiffen, B., and Sherman, H.: An Optimum Demodulator for Poisson Processes. Proc. IEEE, Vol. 51, p. 1319, October 1963.

2. Abend, K.: Optimum Photon Detection. IEEE Transactions in Information Theory, Vol. 12, p. 64, January 1966.

3. Kailath, T.: The Divergence and Bhattacharyya Distance Measures in Signal Selection. IEEE Transactions on Communication Technology, Vol. Com-15, p. 52, February 1967.

4. Helstrom, C.: Quantum Limitations on the Determination of Coherent and Incoherent Signals. IEEE Transactions on Information Theory, Vol. IT-11, p. 482, October 1965.

5. Viterbi, A.: Principles of Coherent Communication. McGraw Hill Book Co., p. 234, 1966.

6. Ross, M.: Pulse Interval Modulation Laser Communications. Proc. Eastcon Convention, Washington, D. C., October 1967.

7. Pratt, W.: Binary Detection in Optical Polarization Modulation Communication Channel. IEEE Trans. on Communication Technology, Vol. Com-14, p. 664, October 1966.

8. Viterbi, A.: ibid., p. 226.

9. Ross, M.: Laser Receivers. J. Wiley and Sons, New York, 1966.

APPENDIX

This appendix presents the derivation of Eq. (13). The average probability of correctly determining the true intensity in M-ary transmission is

$$P_D = \frac{1}{M} \sum_{q=1}^{M} P(D/q) \qquad (A-1)$$

where $P(D/q)$ is the probability of correct detection, given that $n_q + n_o$ is the true intensity. Now the conditional probability of the occurrence of an observed vector $\underline{k} = \underline{j} = \{ j_1, j_2, j_3, \ldots j_M \}$, given $\underline{n}_q + \underline{n}_o$, is

$$P(\underline{k} = \underline{j}/q) = \prod_{i=1}^{M} \frac{[(n_{qi}+n_o)\Delta T]^{j_i}}{j_i!} e^{-(Mn_o)\Delta T} e^{-N}$$

$$\overset{\Delta}{=} \Psi(q, \underline{j}) e^{-N} \qquad (A-2)$$

where N is the energy constraint given in Eq. (3). The conditional detection probability $P(D/q)$ is then obtained by summing over the set of all $\underline{j}$ such that a correct decision is made. A correct decision will occur, when the qth intensity is used, if Λ_q is selected as being the largest. If no other Λ_t exceeds Λ_q but r of the Λ_t's equals Λ_q, the receiver will be correct with a probability of $1/r + 1$, assuming a purely random selection is made when likelihood equalities occur. Now $\underline{j}$ is an M-dimensional vector with non-negative integer components, and the space of all such vectors is denoted as R^M. The conditional detection probability $P(D/q)$ can therefore be written by summing over all $\underline{j} \in R^M$ leading to a correct decision. Thus,

$$P(D/q) = \sum_{r=0}^{M} \frac{1}{r+1} \sum_{J_{qr}} \Psi(q, \underline{j}) e^{-N} \qquad (A-3)$$

where J_{qr} is the set of $\underline{j} \in R^M$ such that no other Λ_t exceeds Λ_q and r other Λ_t's equal Λ_q. If we let I_q denote the r dimensional index set corresponding to these r Λ_t's, for simplicity, J_{qr} can be denoted symbolically as

$$J_{qr} = \left\{ \underline{j} \subset R^M : \Lambda_q = \max_k \Lambda_k = \Lambda_t, t \subset I_q \right\} . \qquad (A-4)$$

Substituting Eq. (A-3) into Eq. (A-1) yields a general expression for the detection probability:

$$P_D = \frac{e^{-N}}{M} \sum_{q=1}^{M} \sum_{r=0}^{M-1} \sum_{J_{qr}} \frac{1}{r+1} \Psi(q, \underline{j}) \qquad (A-5)$$

Now by examining carefully the set J_{qr}, the expression above can be simplified. Making use of the monotonicity of the exponential function, one can write:

$$J_{qr} = \left\{ \underline{j} \subset R^M: \exp(\Lambda_q) = \exp(\max_k \Lambda_k) = \exp \Lambda_t, \, t \subset I_q \right\}$$

$$= \left\{ \underline{j} \subset R^M: \prod_{i=1}^{M} [(n_{qi} + n_o)\Delta T]^{j_i} = \max_k \prod_{i=1}^{M} [(n_{ki} + n_o)\Delta T]^{j_i} \right.$$

$$= \left. \prod_{i=1}^{M} [(n_{ti} + n_o)\Delta T]^{j_i}, \, t \subset I_q \right\}$$

$$= \left\{ \underline{j} \subset R^M: \Psi(q, \underline{j}) = \max_k \Psi(k, \underline{j}) = \Psi(t, j), \, t \subset I_q \right\} . \quad (A-6)$$

Thus J_{qr} can be alternatively defined as the set of j for which $\Psi(q, j)$ is one of $r + 1$ maximum $\Psi(k, \underline{j})$ functions. This means every $\underline{j}$ in J_{qr} also belongs to r other sets $J_{tr}, \, t \subset I_q$, or correspondingly, a point $\underline{j}$ in $J_{tr}, \, t \subset I_q$, exists such that $\Psi(t, \underline{j}) = \Psi(q, \underline{j})$. Note that the set of subspaces $\{J_{qr}\}$ are disjoint for different r but not for different q. With these facts consider the summation.

$$\sum_{q=1}^{M} \sum_{J_{qr}} \frac{\Psi(q, \underline{j})}{r+1} \qquad (A-7)$$

for fixed r. For any term of the sum, say $\Psi(q_o, j_o)/r+1$, there exists r other terms having the same value, one for each point j_o of $J_{tr}, \, t \subset I_{qo}$. The total contribution to the sum above from this set of $r + 1$ terms is then

$$(r + 1) \left[\frac{\Psi(q_o, j_o)}{r + 1} \right] = \Psi(q_o, j_o) = \max_q \Psi(q, j_o) , \qquad (A-8)$$

the last equation following since $j_o \in J_{qr}$. Thus, overlapping points in the summation of Eq. (A-7) contribute a total amount given by Eq. (A-3). It therefore follows that

24

$$\sum_{q=1}^{M} \sum_{J_{qr}} \frac{\Psi(q,\underline{j})}{r+1} = \sum_{\underset{q}{\cup} J_{qr}} \max_{q} \Psi(q,\underline{j}) \qquad (A-9)$$

where $\underset{q}{\cup} J_{qr}$ is the union over of the subsets $\{J_{qr}\}$. Inverting the order of summation in Eq. (A-5) and using Eq. (A-9) allows one to rewrite:

$$P_D = \frac{e^{-N}}{M} \sum_{r=0}^{M-1} \sum_{\underset{q}{\cup} J_{qr}} \max_{q} \Psi(q,\underline{j})$$

$$= \frac{e^{-N}}{M} \sum_{R^M} \max_{q} \Psi(q,\underline{j}) \qquad (A-10)$$

where we have employed the fact that $\left\{ \underset{q}{\cup} J_{qr} \right\}$ are disjoint subspaces, and the sum over all r spans the whole space R^M. Equation (A-10) is Eq. (13) of the report.

"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."

— NATIONAL AERONAUTICS AND SPACE ACT OF 1958

NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

TECHNICAL REPORTS: Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

TECHNICAL NOTES: Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

TECHNICAL MEMORANDUMS: Information receiving limited distribution because of preliminary data, security classification, or other reasons.

CONTRACTOR REPORTS: Scientific and technical information generated under a NASA contract or grant and considered an important contribution to existing knowledge.

TECHNICAL TRANSLATIONS: Information published in a foreign language considered to merit NASA distribution in English.

SPECIAL PUBLICATIONS: Information derived from or of value to NASA activities. Publications include conference proceedings, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

TECHNOLOGY UTILIZATION PUBLICATIONS: Information on technology used by NASA that may be of particular interest in commercial and other non-aerospace applications. Publications include Tech Briefs, Technology Utilization Reports and Notes, and Technology Surveys.

Details on the availability of these publications may be obtained from:

SCIENTIFIC AND TECHNICAL INFORMATION DIVISION

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
Washington, D.C. 20546