THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 55, No. 3, March 1976
Printed in U. S. A.

On Optical Data Communication via Direct Detection of Light Pulses

By J. E. MAZO and J. SALZ

(Manuscript received October 16, 1975)

A number of problems are considered relevant to understanding the performance of optical-fiber communication systems that use pulse transmission. The methods used are typically exact solutions or bounds, and we concentrate on simple examples that aid our understanding. Some of our work makes contact with previous studies, particularly by Personick and Hubbard. The major results are:

(*i*) *Presentation of an integral equation for the output density for single-pulse detection with arbitrary avalanche gain*
(*ii*) *Exact solution for the probability distribution for gains in physical avalanche diodes*
(*iii*) *Bounds on performance when intersymbol interference is present (but no avalanche gain) which suggest that an optimum-bit detector can perform, under practical conditions, only two or three dB better than a simple integrate-and-dump filter, yielding results still many dB from the quantum limit. Thus, in particular, little performance gain is to be expected from equalization techniques.*

I. INTRODUCTION AND OVERVIEW

A large part of traditional communication theory has been directed to detecting and processing electrical signals transmitted over wires, cables, or the like. While the physical realization of each of these traditional systems may have led to mathematical treatments designed to handle problems such as linear distortion or fading, which were peculiar to one, or even perhaps several, systems, the principal concern of all mathematical treatments of these time-continuous channels has been the ubiquitous additive gaussian noise. In fact, it would be fair to say that much of the structure of the mathematical treatments used has been dictated by the mathematical properties of this noise. In the absence of noise, many problems would immediately degenerate, at least theoretically, to situations of perfect detection, infinite capacity, etc.

The consideration of some promising optical communication systems seems to alter the above picture. We have in mind the transmission of information by way of light pulses propagating through an optical fiber and subsequently detected by a photodetector that converts electromagnetic energy in the fiber to electrical signals in a circuit. We immediately note certain features which this problem has in common with the traditional problems. For one thing, the fiber can delay, attenuate, or spread the transmitted pulses. For another, the electrical signal after photodetection may be corrupted by additive gaussian noise. Yet there is another fundamental impairment. The electromagnetic signal that propagates in the fiber (which acts as a wave guide) is, under practical considerations, of sufficiently weak intensity that any effective detection mechanism must be based upon the quantum nature of the electromagnetic disturbance. In other words, detection must be based upon photon counting. Here, a new element enters the problem—photon counting is subject to statistical fluctuations. In the quantum case, a signal uncorrupted by any external disturbance still carries with it its own "noise," as it were, which is not additive gaussian. This new noise manifests itself in the following way. The photon-counting process is a time-varying Poisson process whose intensity (or rate) function $\lambda(t)$ varies in direct proportion to the information-bearing pulse train, the latter being thought of in the conventional way (except it must now always be positive). Our purpose here is to explore some of the communication theory of this new situation, paying particular attention to the use of our considerations in proposed fiber-optic communication systems.

The general background of the material that we treat, namely, direct detection of photons in an optical fiber, may be found in works by Personick[1,2] and Foschini et al.[3] Direct detection refers to the processing of the electrical signal at the output of a photodetector as opposed to, say, more esoteric detection schemes based on optimum processing of the existing electromagnetic field, considered as a quantum system. In the case of binary transmission, the choice between a one or a zero is, in the systems considered here, translated into the presence or absence of a short burst of optical power (light) in the fiber. To understand this in more detail, we shall trace the passage of a single pulse through our mathematical model of the system (see Fig. 1). In the case of a one being transmitted, an electrical signal (a square pulse of duration T) turns on our "flashlight," which in this case is a laser or light-emitting diode, and electromagnetic energy is sent into the transmission medium (optical fiber). If a photodetector is placed at the end of the fiber, photons will be detected due to the electromagnetic energy present. Exactly when in time the photons

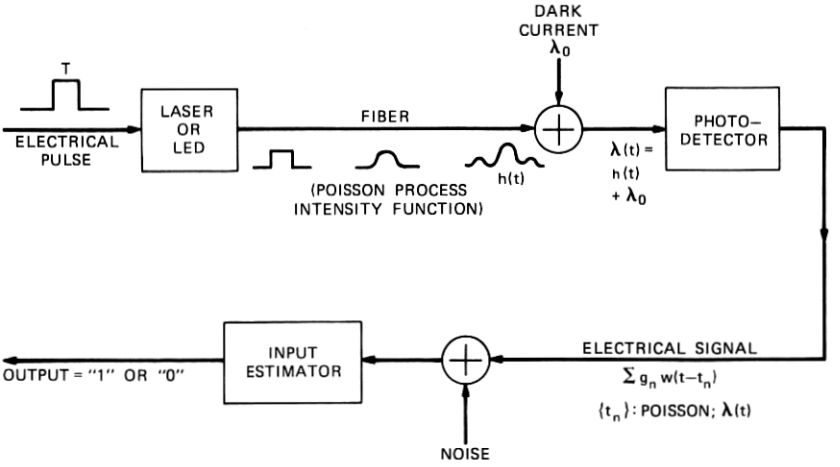

Fig. 1—Passage of a single pulse through the optical system.

register on the detector is random and is the Poisson process spoken of earlier. The probability of receiving a count between time t and $t + dt$ is given by $h(t)dt$, where, owing to effects in the fiber, $h(t)$ is a distorted and attenuated version of the transmitted pulse. The accumulation of distortion as the pulse propagates down the fiber is also sketched in Fig. 1. In practice, a background of counts also exists. This is called the dark current and is modeled by introducing a constant additive intensity function λ_0 before the detector, although some of these counts can originate in the physical detector itself. Typically, transmitted power and transmission loss are adjusted so that on the order of one or two hundred photons per pulse are, on the average, detected. The dark current contributes from about 1 to 5 percent of the counts.

To transmit a zero, we simply do not turn on the transmitting power, and the detector only registers counts resulting from the dark current.

We have been loosely speaking of the output of the photodetector as "counts." The actual electrical current at the output of this device caused by a photon is a wideband pulse $g \cdot w(t)$ (very narrow compared with T, a delta function in the limit), where g = integer-valued random variable or $g \equiv 1$, depending on whether or not an avalanche diode is used. The electrical current at the output of the photodetector is further distorted by gaussian noise whose effect is often lessened in importance when an avalanche diode is used, but not for the $g = 1$ case. In the most literal modeling of the experimental situation, the finite bandwidth of $w(t)$ prevents one from assuming that the Poisson

part of the observation is singular, i.e., can be separated out from the background gaussian noise; however, whenever we feel there are insights to be gained from the separation we shall make it.

If we take into account the facts that Personick[4] has shown superposition to hold (approximately) for optical-fiber transmission and that optical power is positive, then we may extend our single-pulse description to a model for transmission of an entire pulse train. If we transmit a sequence of on or off pulses, then the "received signal," defined as that electrical signal on which we may do processing, can be written as

$$\sum_n g_n w(t - t_n) + n(t), \tag{1}$$

where the time points $\{t_n\}$ form a Poisson process having intensity function $\lambda(t)$, with

$$\lambda(t) = \sum_n a_n h(t - nT) + \lambda_0 \tag{2}$$

and

$$\begin{aligned} h(t) \geqq 0 &= \text{distorted pulse} \\ a_n = 0, 1 &= \text{independent, equiprobable data symbols} \\ \lambda_0 \geqq 0 &= \text{dark current} \\ T &= \text{signaling interval} \\ n(t) &= \text{gaussian noise} \\ g_n &= \text{avalanche gain factors} \\ w(t) &= \text{output pulse of photodetector.} \end{aligned} \tag{3}$$

At various stages of our discussion, we may, for interests of simplicity or clarity, idealize or eliminate certain aspects of the full model given by (1), (2), and (3).

The communication theorist is interested in processing the signal (1) to estimate the a_n given in (2). If the distortion is not severe, one may simply process in an intuitive way and (assuming proper synchronization) count the number of photons detected in the appropriate T-second interval. If $g_n = 1$, this is accomplished by integrating the output for T seconds (so-called integrate-and-dump detection). However, the simplicity of this technique demands its investigation even when g_n are random. Neglecting the gaussian noise and assuming g_n are exponential random variables allow one to determine exactly the probability distribution of the output statistic and to determine error rates. This is done in Section II. In Section III we return to the $g = 1$ case to observe the effects of the random gain. In Section IV, Personick's implicit equation for the random gains g_n of actual photodetectors is studied in detail and the exact distribution of these gains

is found. Also, the use of Chernoff bounds for bounding the error rate in the general situation is discussed. Section V branches out to include a worst-case analysis of intersymbol interference [the case of appreciable spreading of $h(t)$] using integrate-and-dump detection. A particular example is also computed. Finally, in Section VI, we consider the question of replacing the integrate-and-dump detector with an optimum detector. We know that equalization can achieve considerable improvements for voiceband telephone transmission, but can we expect the same here? Using the lower bound on performance which we derive for the optimum detector and applying this to the example of Section V, we find that performance greatly surpassing that of integrate-and-dump detection cannot be expected.

II. INTEGRATE-AND-DUMP DETECTION—AVALANCHE DETECTORS

As already mentioned in the introduction, a simple way to detect the jth bit in (2) is to integrate the output of the photodetector over the jth T-second interval and compare the random variable thus obtained with a threshold F; if the output is greater than F, a one is declared (pulse present); if it is less than F, a zero is declared (pulse absent). In this section, we discuss the exact error rate for such a situation when pulse overlap in (2) can be neglected, as well as the additive noise. Further, the gains g_n are assumed to be exponentially distributed.

We shall need the moment-generating function (MGF) for the indicated random variable, but we may as well begin by giving the MGF for a general linear filter $P(t)$ rather than simply an integrator. Consider a Poisson point process having an arbitrary intensity function $\lambda(t)$ [not necessarily of the form (2)], and let the nth count be given nonnegative weight g_n, i.e., consider

$$\sum_n g_n \delta(t - t_n), \tag{4}$$

where the sequence of time points $\{t_n\}$ is Poisson with intensity function $\lambda(t)$. If (4) is linearly filtered, with $P(t)$ being the impulse response of the filter, then the output of the filter at time t, $x(t)$ can be shown by elementary calculations to have moment-generating function given by

$$M_X = E \exp [sX] = \exp \left[\int_{-\infty}^{\infty} \lambda(\tau) \{ M_g[sP(t - \tau)] - 1 \} d\tau \right], \tag{5}$$

where $M_g(s)$ is the moment-generating function of the g_n, assumed independent, and we have set $x(t) = X$. In particular, if $P(\tau) = 1$ for $0 < \tau < T$ and zero elsewhere, and if $t = T$, (5) will simplify to

$$M_X = \exp \{ \Lambda [M_g(s) - 1] \}, \tag{6}$$

where

$$\Lambda = \int_0^T \lambda(\tau)d\tau. \tag{7}$$

If the pulse $w(t)$ in (1) (assumed of unit area) is narrow enough so that end effects are negligible when doing the integration and if pulse overlap in (2) is negligible, then (6) and (7) are relevant quantities to consider in determining the error rate for integrate-and-dump detection of (1) and (2). To treat the two separate cases of a one or a zero, we need only replace Λ in (6) by Λ_i, $i = 1$ or 0, where

$$\Lambda_1 = \int_0^T h(t) + T\lambda_0 \tag{8}$$

$$\Lambda_0 = T\lambda_0. \tag{9}$$

While the gaussian noise will be neglected here, let us at least note that to include the effect of the added noise term on the integrated output, we would multiply (6) by the moment-generating function of the noise $M_n(s)$,

$$M_n(s) = \exp\left(\frac{s^2\sigma^2}{2}\right), \tag{10}$$

to obtain the MGF of the new output variable. In (10), the variance of the noise σ^2 is given by

$$\sigma^2 = \frac{N_0}{2} T \tag{11}$$

for the case of the integrator with white noise of two-sided spectral density $N_0/2$, or

$$\sigma^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} N(\omega)\,|\tilde{P}(\omega)\,|^2 d\omega, \tag{12}$$

in general, where $N(\omega)$ denotes a general noise spectrum and $\tilde{P}(\omega)$ is the Fourier transform of $P(t)$.

In the special case where the g_n in (4) are continuous variables and are exponentially distributed, i.e.,

$$p(g) = \alpha \exp(-\alpha g), \qquad g > 0, \tag{13}$$

we have

$$M_g(s) = \frac{\alpha}{\alpha - s}, \qquad s < \alpha. \tag{14}$$

At this stage, it is easier to work with the characteristic function version of (6), namely,

$$C_X(\omega) = \exp\{\Lambda[C_g(\omega) - 1]\}, \tag{15}$$

with $C(\omega)$ denoting characteristics functions now, e.g.,

$$C_X(\omega) = E \exp(i\omega X).$$

To obtain our integral equation for $p(x)$, differentiate (15) once with respect to ω, multiply by $\exp(-i\omega x)$, and integrate over x to obtain

$$xp(x) = \Lambda \int_0^x up_g(u)p(x-u)du, \tag{16}$$

where $p(x)$ denotes the density of the random variable X in (14), and $p_g(u)$ denotes the density of the nonnegative gain variable g.

For the exponential gain case (13), an exact solution to (16) can be found. Note that then the variable x has probability $\exp(-\Lambda$ of being zero (no counts) and $p(x)$ will thus contain a δ function at the origin. Introducing this explicitly by writing

$$p(x) = \exp(-\Lambda)\delta(x) + \exp(-\alpha x)f(x), \tag{17}$$

we find

$$xf(x) = (\alpha\Lambda e^{-\Lambda})x + \alpha\Lambda \int_0^x (x-w)f(w)dw, \tag{18}$$

where use has been made of (13). Differentiating (18) twice, we obtain Bessel's equation

$$x^2f'' + 2xf' - (\alpha\Lambda)xf = 0, \tag{19}$$

where f' stands for differentiation. The appropriate solution of (19) gives, finally, for the density $p(x)$ of the detection statistic

$$p(x) = e^{-\Lambda}\delta(x) + e^{-\Lambda}\sqrt{\frac{\Lambda\alpha}{x}}\, e^{-\alpha x}I_1(2\sqrt{\alpha\Lambda x}), \tag{20}$$

$I_1(\cdot)$ being the modified Bessel function.* The following may be useful in connection with (20):

$$I_1(x) \leqq \frac{\exp(x)}{\sqrt{2\pi x}}, \qquad x \geqq 0 \tag{21}$$

$$I_1(x) \sim \frac{\exp(x)}{\sqrt{2\pi x}}, \qquad x \text{ large} \tag{22}$$

$$I_1(x) \sim \frac{x}{2}, \qquad x \text{ small.} \tag{23}$$

Typically, Λ_1 is in the range of 100 to 200 for a light pulse present and

* This exact result, as well as several useful approximations to it found later in this section, were first derived by Hubbard (Ref. 5) using other techniques.

Λ_0 in the range of 5 to 10 for dark current only. The quantity $1/\alpha$, the average gain, may be 100 or 200. The average number of counts for a pulse is then Λ_1/α so, to within a factor of 2 or so, the decision threshold will be around $\Lambda_1/2\alpha$. Thus, virtually all the area of interest in (20) occurs for $x > 1/\alpha\Lambda_i$ for both $i = 1$ or 2, and (22) may be used and, to excellent accuracy,

$$p(x)dx \approx \frac{(\alpha\Lambda)^{\frac{1}{2}}}{\sqrt{4\pi}} \frac{1}{(x)^{\frac{3}{4}}} \exp\left\{-\frac{(\sqrt{x} - \sqrt{\Lambda}/\alpha)^2}{2(1/2\alpha)}\right\} dx, \qquad x > \frac{1}{\alpha\Lambda}. \tag{24}$$

Equation (24) is slightly more attractive if we write instead the density for $u = \sqrt{x}$, $p_u(u)$,

$$p_u(u)du \approx \left(\frac{\Lambda}{\alpha}\right)^{\frac{1}{2}} \frac{1}{\sqrt{u}} \cdot \frac{\exp\left\{-\frac{(u - \sqrt{\Lambda}/\alpha)^2}{2(1/2\alpha)}\right\}}{\sqrt{2\pi}(\sqrt{1}/2\alpha)} du, \qquad u > \frac{1}{\sqrt{\alpha\Lambda}}, \tag{25}$$

showing that $\sqrt{X}$ is, over a rather wide range, gaussian with mean $\sqrt{\Lambda/\alpha}$ and variance $1/2\alpha$. Note $\Lambda/\alpha = EX$, while variance of X is $2\Lambda/\alpha^2$. Also, eq. (25) should not be confused with the central limit theorem version of (24), which is obtained when one writes (for large Λ) $x = (\Lambda/\alpha) + \epsilon$ and ϵ becomes gaussian. Since, from (21), eq. (22) is an upper bound as well as being asymptotic, we have

$$p\left[x > F > \frac{\Lambda}{\alpha}\right] \leqq \left(\frac{\Lambda}{F\alpha}\right)^{\frac{1}{2}} Q(\sqrt{2\alpha F} - \sqrt{2\Lambda}), \tag{26}$$

where

$$Q(y) = \frac{1}{\sqrt{2\pi}} \int_y^\infty e^{-u^2/2} du \sim \frac{e^{-y^2/2}}{\sqrt{2\pi}y}. \tag{27}$$

Likewise, in the same spirit of approximation that indicates (26) to be an excellent approximation (in addition to being an upper bound), one may write for the lower tail

$$p\left[x < F < \frac{\Lambda}{\alpha}\right] \approx \left(\frac{\Lambda}{F\alpha}\right)^{\frac{1}{2}} Q(\sqrt{2\Lambda} - \sqrt{2\alpha F}). \tag{28}$$

Even for Λ's differing by a factor of 100, the fourth root factor in front of (26) and (28) is weak indeed. Thus, we may, to excellent approximation, find the best threshold by equating the arguments of the Q function for the two cases of error. This results in

$$\sqrt{2\Lambda_1} - \sqrt{2\alpha F} = \sqrt{2\alpha F} - \sqrt{2\Lambda_0}. \tag{29}$$

The left-hand side of (29) comes, of course, from using (28) for a pulse present (the number of counts is then expected to exceed the threshold).

Table I — Tabulation of error rate and threshold for an avalanche detector with exponentially distributed gains

Λ_0	Λ_{1s}	$F_{opt}(\alpha = 1)$	P_e [eq. (31)]	Quantum Limit
4	100	37.20	2.09×10^{-9}	1.86×10^{-44}
4	200	66.28	8.88×10^{-19}	$\sim 10^{-88}$
4	400	122.1	3.9×10^{-38}	$\sim 10^{-176}$
10	100	46.58	8.03×10^{-8}	1.86×10^{-44}
10	200	77.91	3.61×10^{-16}	$\sim 10^{-88}$
10	400	137.0	3.66×10^{-34}	$\sim 10^{-176}$

Likewise, the right member of (29) comes from using (26) for only dark current where the counts usually fall below the threshold F and an error is made only if they exceed it. We immediately obtain from (29)

$$\sqrt{F_{opt}} = \sqrt{\frac{\Lambda_1}{4\alpha}} + \sqrt{\frac{\Lambda_0}{4\alpha}}, \tag{30}$$

where, again, Λ_0 is not to be too small, for example, $\Lambda_0 \geqq 2$. In the above, we have in mind, from (8), taking $\Lambda_1 = \Lambda_{1s} + \Lambda_0$ where Λ_{1s} is due to signal alone.

For future comparisons, we should inject at this point the fact that the best detection probability one can obtain with no dark current (or no gaussian noise) is $\frac{1}{2} \exp(-\Lambda_{1s})$, often referred to as the quantum limit.

Table I displays values of the right member of (26), for the optimum F given by (30), i.e., it displays the quantity

$$\left(\frac{\Lambda_0}{F\alpha}\right)^{\frac{1}{4}} Q\left(\sqrt{\frac{\Lambda_1}{2}} - \sqrt{\frac{\Lambda_0}{2}}\right) \tag{31}$$

evaluated for several values of Λ_0 and Λ_{1s}, along with the quantum limit. Note that only αF enters the expressions, and thus the actual value of α plays no role in determining the probabilities for this problem. The fact should also be evident from the scaling properties of the problem. In real applications, $1/\alpha$ would be large so that the electronic circuitry could "see" the pulses above the gaussian noise.

Table I shows (for the parameters shown) about a 7-dB loss relative to the quantum limit, owing to the dark current, and also in part to the random nature of the gain mechanism.*

* To be perfectly clear on this point, it is really the additional (random) gain provided by the avalanche detector that allows one to formulate the physical problem as in (4) without gaussian noise. However, from a mathematical point of view, once (4) is written down, the random gains are hypothesis-insensitive, and thus would be ignored by an optimum detector.

III. INTEGRATE-AND-DUMP DETECTION—PURE POISSON CASE

We now give a brief discussion for the g = 1 case of (4), namely, the random variable X is Poisson,

$$p(X = n) = \frac{e^{-\Lambda}\Lambda^n}{n!} \qquad n = 0, 1, 2, \cdots, \tag{32}$$

$$EX = \Lambda, \qquad \text{var } X = \Lambda^2. \tag{33}$$

The purpose of the remarks will be to shed light on the degradation suffered when the g_n are random, as mentioned at the end of the last section.

If X is Poisson, then the probability that X is larger than or equal to k is

$$\sum_{n=k}^{\infty} \frac{e^{-\Lambda}\Lambda^n}{n!} = \frac{e^{-\Lambda}\Lambda^k}{k!}\left[1 + \frac{\Lambda}{k+1} + \frac{\Lambda^2}{(k+1)(k+2)} + \cdots\right]. \tag{34}$$

If, in addition, we assume $(k + 1) > \Lambda$, then a simple consequence of (34) is that

$$\left(1 + \frac{\Lambda}{k+1}\right)\frac{e^{-\Lambda}\Lambda^k}{k!} < \Pr[X \geqq k > \Lambda - 1] < \frac{1}{1 - (\Lambda/k+1)} \cdot \frac{e^{-\Lambda}\Lambda^k}{k!}. \tag{35}$$

Similarly, for the lower tail we have

$$\left(1 + \frac{k}{\Lambda}\right)\frac{e^{-\Lambda}\Lambda^k}{k!} < \Pr[x \leqq k < \Lambda] < \frac{1}{1 - (k/\Lambda)} \frac{e^{-\Lambda}\Lambda^k}{k!}. \tag{36}$$

Thus, ignoring the weak effects of the coefficient in front, the optimum threshold F for a problem such as the one described in Section II is obtained by equating probabilities such as these in (35) and (36), yielding

$$e^{-\Lambda_0}\Lambda_0^F = e^{-\Lambda_1}\Lambda_1^F \tag{37}$$

or, equivalently, the optimum threshold in this case is

$$F = \frac{\Lambda_1 - \Lambda_0}{\ln(\Lambda_1/\Lambda_0)}. \tag{38}$$

Table II displays the right-hand side of (35) for k given by the rounded-off values of (38). In particular, we see degradation ranging from 3.5 to 4 dB compared to the quantum limits given in Table I. Typically, then, detecting the presence or absence of a single pulse using random amplitudes, as a linear detector might, results in a 3- to 4-dB degradation (for the exponential case), compared with an "ideal"

Table II — Tabulation of error rate and threshold for detection with constant gain

Λ_0	Λ_{1s}	F_{opt}	P_e [eq. (35)]
4	100	30.69	1.17×10^{-17}
4	200	50.87	6.49×10^{-38}
4	400	86.67	2.18×10^{-82}
10	100	41.70	4.21×10^{-14}
10	200	65.69	9.80×10^{-32}
10	400	107.7	3.78×10^{-71}

avalanche detector, which has a large gain but whose distribution is concentrated at a delta function.

The loss due to "gain jitter" suggests a possible remedy. The physical pulse $g_n w(t - t_n)$ in the detection circuits following the avalanche diode should be clearly detectable against the background noise if g_n is sufficiently large; in particular, if it is something like the mean gain G. Suppose this is also true for pulse gains $g_n \geqq fG$, $f < 1$. Now suppose one processed the circuit output of the avalanche diode by first passing it through a pulse detector that detects pulses of height greater than fG and generates a pulse of fixed height if a pulse is detected. The output pulses of this device have fixed gain, which is beneficial, but, on the other hand, we have lost a fraction θ,

$$\theta = \frac{1}{G} \int_0^{fG} \exp(-g/G) dg, \tag{39}$$

of light intensity. Seemingly, by a simple scheme we may have still gained a dB or two in performance. Because of effects such as possible overlap of two close pulses $w(t)$ and even in the pulse shape of $w(t)$ itself, the merits of this proposal are hard to assess without further study. It does appear to be an interesting possibility for a future detailed investigation.

IV. INTEGRATE-AND-DUMP DETECTION—OTHER AVALANCHE GAIN DISTRIBUTIONS

Personick[6] has considered the physics of a class of real avalanche detectors in considerable detail and has derived the following implicit equation for their moment-generating function $M_g(s)$:*

$$s = \ln M - \frac{1}{1 - k} \ln [(1 - a)M + a], \tag{40}$$

* We shall drop the subscript on the MGF M_g of the gain variable when we refer to the particular M_g given by (40). Also, the k in this section has nothing to do with the k in (35) and (36).

where we have set

$$M \equiv M(s) = \sum_{n=1}^{\infty} e^{sn} p_n. \tag{41}$$

The parameters k and a are related to the physical properties of these photon detectors. Since (40) has never been explicitly solved for $M(s)$, we think it worthwhile to investigate the structure of $M(s)$ implied by (40) in more detail. In addition to yielding structural properties of $M(s)$, we shall find that (40) allows us to determine the p_n of (41) exactly.

To begin with, the gain G, given by $G = Eg$, is

$$G \equiv Eg = \frac{d}{ds} M(s)\Big|_{s=0}, \tag{42}$$

which, using (40), yields

$$G = \frac{1 - k}{a - k}. \tag{43}$$

From (43) we see that the restrictions

$$\begin{aligned} 0 < a \leqq 1 \\ 0 \leqq k < a \end{aligned} \tag{44}$$

are to be imposed on the parameters in (40).

When $a = 1$, (40) gives $M = e^s$, the $g = 1$ case. When $k = 0$, (40) is easily solved to give

$$M(s) = \frac{ae^s}{1 - (1 - a)e^s}, \qquad k = 0. \tag{45}$$

Equation (45) is the MGF of the discrete geometric distribution having probabilities p_n concentrated on the positive integers, where

$$p_n = \frac{a}{1 - a}(1 - a)^n, \qquad n = 1, 2, \cdots. \tag{46}$$

It is reasonable to treat the continuous version of this density, and that was done in Section II.

In the general case of (40), the variance may be calculated to give

$$\text{var } g = G^3 \left[1 - \frac{(1 - a)^2}{1 - k}\right] - G^2. \tag{47}$$

If higher moments are desired, they can be obtained recursively from (40). This can be done by expanding $M(s)$ in a power series and equating like powers in s.

In view of the discussion in Section III, one might prefer the detectors represented by (40) that have small variance. A simple in-

vestigation of (47) reveals that, for any $a < 1$, $k = 0$ uniquely gives minimum variance. Since even this minimum variance is large (equal to the mean), it may well not be a reliable guide.

Returning to the general case represented in (40), it is evident from the relation

$$M(s) = \sum_{n=1}^{\infty} e^{sn} p_n$$

that the MGF exists for all $s \leqq 0$. However, it does not exist for all positive s, and, in fact, setting $ds/dM = 0$ yields a critical value of M (call it M_c) given by

$$M_c = \frac{a}{1-a} \frac{1-k}{k} \tag{48}$$

and thus a critical value s_c of s given by

$$s_c = s(M_c) = \ln \frac{1-k}{1-a} - \frac{k}{1-k} \ln \frac{a}{k}, \tag{49}$$

beyond which $M(s)$ does not exist. Note that, if $b \neq 0$ (and $a \neq 1$), the value of the MGF at the critical s is finite. This shows that the far-tail behavior of the g variable has an exponential-like tail, with damping factor related to s_c, but in general there is a multiplicative factor, e.g., an inverse power that allows the MGF to be finite at its critical value.

If we let $s_c - s = \delta > 0$, $M_c - M(s) = \Delta > 0$, and write

$$s_c - s \cong s(M) = s(M_c - \Delta) = s(M_c) - \Delta \left.\frac{ds}{dM_g}\right|_{M_c} + \frac{1}{2} \Delta^2 \left.\frac{d^2 s}{dM_g^2}\right|_{M_c} + \cdots, \tag{50}$$

we obtain, after evaluating the second derivative in (50), that

$$\Delta \cong \sqrt{\delta} \sqrt{\frac{2}{k}} M_c \tag{51}$$

or, equivalently,

$$M \approx M_c \left[1 - \sqrt{\frac{2}{k}} \sqrt{s_c - s}\right], \tag{52}$$

thus exhibiting a square-root singularity of $M(s)$ in the neighborhood of s_c. This type of behavior is consistent with a far-tail fall-off of the "density" of the g variable being given by

$$\text{const.} \frac{\exp(-s_c g)}{g^{\frac{3}{2}}}. \tag{53}$$

Let us now proceed to the exact solution for the p_n in (41) when $M(s)$ is given by (40). We use instead $z = \exp(s)$ and write, with a slight abuse of notation,

$$M(z) = \sum_{n=1}^{\infty} z^n p_n. \tag{54}$$

Equation (40) becomes, setting $M = M(z)$ when convenient,

$$z = \frac{M}{[M(1-a)+a]^{1/(1-k)}}. \tag{55}$$

In (55) it is useful to make the substitutions

$$M(z) = \frac{a}{1-a} F[(1-a)a^{k/(1-k)}z] \tag{56a}$$

$$u = (1-a)a^{k/(1-k)}z \tag{56b}$$

$$\rho = \frac{1}{1-k} \tag{56c}$$

to obtain

$$u = \frac{F}{[1+F]^{\rho}}, \tag{57}$$

where $F(0) = 0$ and F is regarded as an implicit function of u in the neighborhood of $u = 0$. Equation (57) is a canonical form for the Lagrange inversion formula[7] for obtaining the coefficients c_j in the power series

$$F = \sum_{j=1}^{\infty} c_j u^j. \tag{58}$$

The formula yields, for the present problem,

$$c_j = \frac{1}{j!}\left\{\left(\frac{d}{dF}\right)^{j-1}(1+F)^{\rho j}\right\}_{F=0} \tag{59}$$

or

$$c_1 = 1,$$

$$c_j = \frac{\prod_{s=0}^{j-2}(j\rho - s)}{j!} = \frac{\Gamma[j/(1-k)+1]}{\Gamma(j+1)\Gamma[kj/(1-k)+2]}, \qquad j \geqq 2. \tag{60}$$

From (54) and (56), the probabilities p_j are then given by

$$p_j = \frac{a}{1-a}[(1-a)a^{k/(1-k)}]^j c_j. \tag{61}$$

For (kj) large, we have, from Stirling's asymptotic formula for the

gamma function,

$$\Gamma(z+1) \sim e^{-z} z^{z+\frac{1}{2}} \sqrt{2\pi}, \tag{62}$$

that

$$c_j \sim \frac{1}{\sqrt{2\pi}} \frac{1}{(kj)^{\frac{1}{2}}(1-k)^{j-1}} \frac{1}{(k^{k/(1-k)})^j}, \qquad \text{as} \quad kj \to \infty. \tag{63}$$

One can show that the behavior given in (63) is, via (61), in complete agreement with (49) and (53).

Remarkably, Personick reports that McIntyre,* from special-case calculations, has conjectured the exact form of (61).

Knowing the p_j does, in principle, allow the exact calculation of the output statistics of the integrate-and-dump filter. The integral equation (16), appropriately interpreted with sums, provides one such way. Instead of discussing this, however, we now turn our attention to bounding techniques. We shall make some remarks directed toward the Chernoff bound, used by Personick[6] for this type of problem.

The Chernoff bound states that, if x has MGF $M_x(s)$, then the probability that x is greater than (less than) F obeys

$$\Pr\left[x \underset{(<)}{>} F\right] \leqq \exp(-sF) M_x(s) \qquad \text{for any} \quad s \underset{(<)}{>} 0. \tag{64}$$

One makes the bound as tight as possible by minimizing the right member of (64) over s. This, of course, assumes that $M_x(s)$ is known or can be obtained explicitly as a function of s. For the general class of avalanche diodes for which Personick derives the moment-generating function, we saw that s is given explicitly as a function of M and, in fact, an explicit function of M vs s is difficult to obtain analytically. Personick gets M numerically as a function of s and then proceeds to optimize with respect to s—a rather tedious procedure. We found from our experience that a simpler approach is to eliminate s in (64) by using (40) and then to optimize over M. This optimization still has to be done numerically. Nevertheless, we could generate curves very quickly this way. We do not present these curves here, since they do not reveal more than those which Personick has already published.

For insight concerning the accuracy of the bound for present purposes, we shall apply it below to the problem of exponential gains, for which we have exact solutions available for comparison.

The function appearing in the right member of (64) is, for the exponential gain case,

$$\exp(-sF) \exp\{\Lambda[1/(1-s) - 1]\}. \tag{65}$$

---

* In addition to the cited reference of Personick, other experimental properties of avalanche photodiodes may be found in Webb, McIntyre, and Conradi (Ref. 8).

Finding the optimum s is easy in this case, and (64) then yields, for these optimum s,* $s_{opt} = 1 - \sqrt{\Lambda/F}$, and, consequently,

$$\begin{aligned} P[x > F > \Lambda] &\leqq \exp\left[-(\sqrt{\Lambda} - \sqrt{F})^2\right] \\ P[x < F < \Lambda] &\leqq \exp\left[-(\sqrt{\Lambda} - \sqrt{F})^2\right]. \end{aligned} \tag{66}$$

From the asymptotic forms of (26) and (28), we see that the Chernoff bound has given us the "right exponent."

From saddle-point considerations, this would be expected to be true in this problem for any $M_y(s)$; however, it by no means has to be true in general, where complex variable (saddle-point) techniques must be resorted to in order to decide the question.

The optimum threshold for single-bit detection that would be obtained by equating the two expressions in (64) (for different Λ's, of course) also results in (30). Table III lists the Chernoff upper bounds to the bit error rate, and these should be compared to the exact answers shown in Table I. Numerically, the Chernoff bound is off by one to two orders of magnitude in error rate due to "coefficient effects." However, even numerically this bound is judged to perform respectably. Also shown in Table III is $s_{opt} = \alpha[1 - \sqrt{\Lambda/F}]$, where the gain ($\alpha$) effect has been included. For the optimum choice of F, it turns out that the two choices of s_{opt} (due to two possible Λ's) are the negative of each other. Hence, only the positive one is shown in Table III.

If one wishes to include the effects of gaussian noise here, one multiplies the right-hand side of (64) by the appropriate MGF, namely, (10). Instead of finding the optimum s for this problem, one can use the s_{opt} that held for the problem without additive noise (any s of appropriate sign furnishes a bound). The value $\sigma^2 = 10^4$ was used in further Chernoff bound calculations for the $M_y(s)$ given in (41) and may be found in the article by Personick.[6]

V. INTERSYMBOL INTERFERENCE—INTEGRATE-AND-DUMP FILTER

We turn now to the situation where $\lambda(t)$ is given by (2), i.e., a train of interfering pulses instead of just one of them. Personick has claimed that $h(t)$ has a gaussian shape in real fibers and, hence, in practice only a few pulses would be expected to contribute intersymbol interference.

It is evident that, if the filter $P(t)$ that processes the output of the photon detector is always positive, as, for example, for an integrate-and-dump filter, the presence of intersymbol interference increases

* In setting the derivative equal to zero, one must choose the positive s that satisfies $s < 1$, since in the real-variable techniques used here, the MGF of the exponential does not exist for $s \geqq 1$.

Table III — Tabulation of Chernoff bound (CB) for error rate, exponential gain case. Also given are $s_{opt} = \alpha[1 - \sqrt{\Lambda/F}]$ for a gain $1/\alpha = 100$, and a correction $\exp[S_{opt}^2 \sigma^2/2]$ for $\sigma^2 = 10{,}000$. The latter is a correction for gaussian noise.

Λ_0	Λ_{1s}	CB [eq. (66)]	s_{opt} (gain = 100)	$\exp \frac{S_{opt}^2 \sigma^2}{2}$
4	100	5.05×10^{-8}	6.72×10^{-3}	1.25
4	200	4.16×10^{-17}	7.54×10^{-3}	1.33
4	400	2.70×10^{-36}	8.19×10^{-3}	1.40
10	100	1.49×10^{-6}	5.37×10^{-3}	1.16
10	200	1.16×10^{-14}	6.42×10^{-3}	1.23
10	400	2.01×10^{-32}	7.30×10^{-3}	1.31

the counts observed over any interval. Therefore, if a pulse is present, this intersymbol interference helps detection (helps keep output greater than the threshold) while, if the pulse is absent and no-counts is ideal, it hurts. Hence, the worst-case situation is to evaluate the probability of a one being decoded into a zero when no other pulses are present, while for the reverse error we assume all pulses are on.

Since we are still considering an integrator, i.e., $P(t) = 1$, $|t| < \tau$, we are still to use (6), but now for the two worst cases given we replace Λ in (6) by either

$$\Lambda_1 = 2\tau\lambda_0 + \int_{-\tau}^{\tau} h(t)dt$$

or (67)

$$\Lambda_0 = 2\tau\lambda_0 + \sum_{n \neq 0} \int_{-\tau}^{\tau} h(t - nT)dt.$$

Of course, we assume $\Lambda_0 < \Lambda_1$ for any reasonable operating situation. In addition to the threshold choice, we must also contend with the optimum choice of τ, half the time width of the integration. This latter step is easily handled numerically.

Many calculations may be done and, for the worst-case situation described, nothing new is involved in addition to what has already been discussed. As an illustration, we will deal explicitly with one example. We take $\lambda_0 = 0$, no avalanche gain ($g = 1$), and

$$h(t) = \frac{100}{T}\left[1 - \frac{|\tau|}{T}\right], \tag{68}$$

where T is the pulse repetition rate. Thus, there is considerable overlap from neighboring pulses, but not from others. Also, $\int h(t)dt = 100$,

Table IV — An intersymbol interference example from Section V

$\frac{\tau}{T}$	F	P_e [eq. (35)]
0.1	7	7.3×10^{-5}
0.2	15	1.5×10^{-5}
0.3	25	5.7×10^{-6}
0.4	34	1.5×10^{-5}
0.5	46	5.1×10^{-5}

so the quantum limit for single-pulse detection may be read from Table I.

Table IV gives the worst-case error rate for the above example, using formulas (35) and (38) for the Poisson case. The optimum choice of τ here is 0.3, i.e., 30 percent toward the peak of the neighboring pulse. Also, a 20-percent change in the value of τ does not change the error rate drastically. Note that we are not inferring that one should be careless in the choice of τ, because in calculating Table IV the optimum threshold (F) for each τ is assumed. Also, note the large degradation with respect to the quantum limit caused by the intersymbol interference. For the present example, the error rate averaged over all sequences cannot be much better than shown, because the worst case occurs with probability $\frac{1}{4}$, and hence $(P_e)_{\text{av}}$ cannot be more than a factor of 4 better.

VI. AN INTERSYMBOL INTERFERENCE EXAMPLE AND A LOWER BOUND ON PERFORMANCE

We present now a lower bound on performance which can be readily evaluated for the intersymbol interference problem of the last section [pulses given by (68)]. This lower bound is valid for optimum bit detection and thus sets a limit on how well *any* detector can do in coping with intersymbol interference. In particular, the bound sheds light on the performance in the present situation of suboptimum schemes such as equalization, which have found such wide application in voiceband telephone transmission.

The derivation of the lower bound proceeds along lines used by Mazo[9] to generalize Forney's lower bound for optimum bit-by-bit detection in the gaussian noise. Our approach is to assume that we are optimally detecting the kth bit in a sequence of $(N+1)$ independent bits, i.e., sequences of the form (2) of length $(N+1)$ are being considered. We suppose a_n are binary, equiprobable, and independent. Let $p_1(x \mid i)$ and $p_0(x \mid i)$ be the two probability densities of the received signal under the hypotheses $a_n = 1$ or 0, respectively,

and i denote conditioning on the ith, $i = 1, \cdots, 2^N$ sequence being transmitted. Then the probability of error for the optimum detector is (in somewhat formal notation)

$$P_e = \frac{1}{2} \int dx \min \left[\frac{1}{2^N} \sum_{i=1}^{2^N} p_1(x \mid i), \frac{1}{2^N} \sum_{j=1}^{2^N} p_0(x \mid j) \right], \tag{69}$$

which, as in Ref. 9, can be lower-bounded by

$$P_e \geqq \frac{1}{2^N} P_e \text{ (binary } i, j \text{ problem)}. \tag{70}$$

In (70), P_e (binary i, j problem) is the probability of error which would result for the simple binary problem of distinguishing between sequence i (one having $a_k = +1$) from sequence j (one having $a_k = 0$). The bound (70) holds for all such (i, j) pairs. Finally, (70) holds if the sequences of length $(N + 1)$ are shortened to $N' + 1$, with N being replaced by N' on the right side of (65).

For communication in the Poisson regime, the right member of (70) has no known evaluation as it does for the gaussian case. What is known about the binary problem is the optimum detector, which is linear. The optimum filter $P(t)$ and threshold F are known explicitly if one is deciding between equiprobable intensity functions $\lambda_a(t)$ and $\lambda_b(t)$. In fact, from the work of Bar-David,[10]

$$P(t) = \ln \frac{\lambda_a(t)}{\lambda_b(t)} \tag{71}$$

and

$$F = \int \lambda_a(t) - \int \lambda_b(t). \tag{72}$$

Thus, the set of received impulses is filtered through $P(t)$ and the resulting output variable X at the end of the observation interval is compared to the threshold F, choosing $\lambda_a(t)$ if $X > F$ and $\lambda_b(t)$ otherwise. Assuming $\lambda_a(t)$ is transmitted, the moment-generating function of X is, from (5) and (71) (recall $g = 1$ in this section),

$$\begin{aligned} M_x(s) &= \exp \left[\int \lambda_0(t) [\exp \{ s \ln [\lambda_1(t)/\lambda_0(t)] \} - 1] dt \right] \\ &= \exp \left[\int [\lambda_0^{1-s}(t) \lambda_1^s(t) - \lambda_0(t)] dt \right]. \end{aligned} \tag{73}$$

From this MGF, one can see why the right side of (70) is not known in general.

We now apply (73) to the intersymbol interference of the previous section, where $h(t)$ is given by (68). We choose $N = 2$, $\lambda_1(t)$ to cor-

respond to the pulse sequence (1, 1, 1) and $\lambda_b(t)$ to correspond to the pulse sequence (1, 0, 1). When applied to (70), we interpret the results as applying to the center bit of the sequence. We have, explicitly,*

$$\begin{aligned} \lambda_1(t) &= 1 \qquad \text{for} \quad |t| \leqq 1 \\ \lambda_0(t) &= |t|, \qquad \text{for} \quad |t| \leqq 1 \\ \lambda_1(t) &= \lambda_0(t) \qquad \text{for} \quad |t| > 1. \end{aligned} \tag{74}$$

Since, from (71), $P(t) = 0$ for $|t| > 1$, the detection interval $t \in [-1, 1]$. Using (74) in (73), the decision variable has MGF

$$M_x(s) = \exp\left[\frac{2}{2 - s} - 1\right]. \tag{75}$$

Remarkably enough, this is the moment-generating function of the random variable dealt with in Section II; in the notation of that section, it corresponds to $\Lambda = 1$, $\alpha = 2$. The density is given by (20), and the threshold is, from (72) and (74), to be set equal to unity. Putting this all together, (70) becomes

$$P_e \geqq \frac{1}{4} e^{-1} \int_1^\infty \sqrt{\frac{2}{x}}\, e^{-2x} I_1(2\sqrt{2x})\, dx. \tag{76}$$

Or, scaling (76) to reinsert the factor of 100 in front of (68),

$$P_e \geqq \frac{1}{4} e^{-100} \int_{100}^\infty \sqrt{\frac{200}{x}}\, e^{-2x} I_1(2\sqrt{200x})\, dx. \tag{77}$$

So an excellent approximation in the right-hand side of (77) may be evaluated via (26) to give

$$P_e \geqq \tfrac{1}{4}(\tfrac{1}{2})^{\frac{1}{2}} Q(\sqrt{400} - \sqrt{200}) \approx 5.06 \times 10^{-10}. \tag{78}$$

The numerical value of (78) should be compared with Table IV for performance with integrate-and-dump filter and Table I for the quantum limit. Indeed, for this case our bound shows that the optimum detector performance is still far from the quantum limit and, in fact, is roughly only 2.2 dB (comparing powers of 10) better than the integrate-and-dump filter.† The present problem seems to imply that equalization,‡ in particular, cannot be expected to approach the quantum limit bound for the type of distortion found in present optical fibers. In fact, a simple integrate-and-dump receiver with properly

---

* For the moment, we ignore the factor of 100 in (68) and also set $T = 1$. These are reintroduced only in the final numerical calculations.

† More precisely, the figure is 2.9 dB for strong signals.

‡ Some references on equalization for optical communication systems are Refs. 1 and 11.

chosen threshold compares well with a lower performance bound. The above problem ignored many practical factors, but in fact ignoring them focused even more on the pure intersymbol interference problem in the Poisson regime. It would seem that effects such as dark current and finite width of $w(t)$ would surely make the integrate-and-dump and the optimum detector perform even more equally, and it would seem too much for an equalizer to compensate for gain jitter, which is a rather nonlinear effect.

Another linear filter $P(t)$, which performs better than the integrate-and-dump, may be inferred from (71). This is discussed and evaluated in the appendix for the present problem. This new linear filter has a worst-case exponent approximately 1 dB better than the integrate-and-dump situation.

APPENDIX

A New Filter

We have already noted that (asymptotically) the integrate-and-dump filter performs within 2.9 dB of a lower bound on performance for the optimum processor for our particular example. We now show how a modified $P(t)$ can perform within 2 dB of this bound. We confine ourselves to the worst case again, for which, we recall, the best integrator had $P(t) = 1$ for $|t| \leqq 0.3$ (choosing $T = 1$). The worst case with signal present was $\lambda_1(t) = 1 - |t|$, $|t| < 1$, and $\lambda_0(t) = |t|$, $|t| < 1$, for the worst case with signal absent. Now the optimum filter

$$P(t) = \ln \frac{1 - |t|}{|t|}, \qquad |t| < 1, \tag{79}$$

which distinguishes between these two signals, is not always positive (it is negative for $|t| > \frac{1}{2}$). Therefore, if (79) were used, there could be no claim for a worst-case bound. However, we modify (79) and use

$$P(t) = \ln \frac{1 - |t|}{|t|}, \qquad |t| < \tfrac{1}{2} \tag{80}$$

instead. The filter represented by (80) is always positive, and therefore worst-case claims still obtain. The filter (80) clearly has to outperform our integrate-and-dump one, since the latter integrated only to $|t| = 0.3$, while (80) is optimum for an observation interval $|t| \leqq 0.5$. The optimum threshold for (80) is, from (72),

$$F = \int_{-\frac{1}{2}}^{\frac{1}{2}} \lambda_1(t)dt - \int_{-\frac{1}{2}}^{\frac{1}{2}} \lambda_0(t)dt = \tfrac{1}{2}. \tag{81}$$

Using the Chernoff bound for the case when $\lambda_0(t)$ is sent, we have, from

(64) and (73),

$$P_e \leqq \exp\left\{\int \lambda_0(t)[e^{sP(t)} - 1] - sF\right\}$$

$$= \exp\left\{2\int_0^{\frac{1}{2}} t^{1-s}(1-t)^s dt - \frac{1}{4} - \frac{s}{2}\right\}, \qquad s > 0, \tag{82}$$

where we have used the expression for $\lambda_0(t)$, the filter (80), and threshold (81). If we let $u = (1-t)/t$, then we may write

$$\int_0^{\frac{1}{2}} t^{1-s}(1-t)^s dt = \int_1^\infty \frac{u^s}{(1+u)^3}\, du. \tag{83}$$

Two integrations by parts give

$$\int_1^\infty \frac{u^s}{(1+u)^3}\, du = \frac{1}{8} + \frac{s}{4} + \frac{s(s-1)}{2}\int_1^\infty \frac{u^{s-2}}{1+u}\, du, \tag{84}$$

or, using (84) in (82),

$$P_e \leqq \exp\left[s(s-1)\int_1^\infty \frac{u^{s-2}}{1+u}\, du\right]. \tag{85}$$

Equation (85) makes it evident that the exponent in (82) will be negative for $0 < s < 1$. If we expand the $1/(1+u)$ part of the integrand in (80) in powers of $(1/u)$ and integrate term by term, the exponent in (85) becomes

$$s(s-1)\sum_{k=0}^{\infty} \frac{(-1)^k}{k+2-s}$$

$$= s(s-1)\sum_{\substack{k=0 \\ k \text{ even}}}^{\infty} \frac{1}{(k+2-s)(k+3-s)}. \tag{86}$$

Convergence in (86) can be improved if we write

$$\sum_{k \text{ even}} = \tfrac{1}{2}\sum_{\text{all } k} + \tfrac{1}{2}\sum_{k \text{ even}} - \tfrac{1}{2}\sum_{k \text{ odd}}$$

and use the fact that

$$\sum_{n=1}^{\infty} \frac{1}{(x+n)(x+n+1)} = \frac{1}{1+x}$$

to obtain

$$s(s-1)\left[\frac{1}{2(2-s)} + \sum_{\substack{k=0 \\ k \text{ even}}}^{\infty} \frac{1}{(k+2-s)(k+3-s)(k+4-s)}\right]. \tag{87}$$

The optimum s is easily found numerically by plotting (87); we truncated the sum after $k = 10$. We find the optimum s is about 0.6, giving a value of (87) of 0.11138. As a check on the possible accuracy

of our use of (87), we note that our technique gives 0.10696 when $s = \frac{1}{2}$, for which the exact answer can be shown to be $\pi/8 - \frac{1}{2} \approx 0.10730$. Thus, the Chernoff bound is

$$\begin{aligned} P_e &\leqq \exp(-0.111\Lambda_0) \\ \Lambda_0 &= \int_{-1}^{1} \lambda_0(t)dt, \end{aligned} \tag{88}$$

while (73) yields as a lower bound something which behaves exponentially as

$$\exp\left(-\Lambda_0\left[\frac{(\sqrt{4}-\sqrt{2})^2}{2}\right]\right) = \exp(-0.172\Lambda_0). \tag{89}$$

The exponent of (88) is 1.9 dB worse than that of (89). Concluding, we note that (80) has a logarithm singularity at $t = 0$. Including dark current in the $\lambda_i(t)$ will remove this, and will also decrease the improvement which this kind of filter provides over the integrate-and-dump filter.

REFERENCES

1. S. D. Personick, "Receiver Design for Digital Fiber Optic Communication Systems, I," B.S.T.J., *52*, No. 6 (July–August 1973), pp. 843–874.
2. S. D. Personick, "Receiver Design for Digital Fiber Optic Communication Systems, II," B.S.T.J., *52*, No. 6 (July–August 1973), pp. 875–886.
3. G. J. Foschini, R. D. Gitlin, and J. Salz, "Optimum Direct Detection for Digital Fiber-Optic Communication Systems," B.S.T.J., *54*, No. 8 (October 1975), pp. 1389–1430.
4. S. D. Personick, "Baseband Linearity and Equalization in Fiber Optic Digital Communication Systems," B.S.T.J., *52*, No. 7 (September 1973), pp. 1175–1194.
5. W. M. Hubbard, "Comparative Performance of Twin-Channel and Single-Channel Optical-Frequency Receivers," IEEE Trans. Commun., *COM-20*, No. 6 (December 1972), pp. 1079–1086.
6. S. D. Personick, "New Results on Avalanche Multiplication Statistics with Applications to Optical Detection," B.S.T.J., *50*, No. 1 (January 1971), pp. 167–190.
7. N. G. DeBruijn, *Asymptotic Methods in Analysis*, Amsterdam: North-Holland, 1961. See p. 22, 2nd edition.
8. P. P. Webb, R. J. McIntyre, and J. Conradi, "Properties of Avalanche Photodiodes," RCA Review, *35*, June 1974, pp. 234–276.
9. J. E. Mazo, "Faster-than-Nyquist Signaling," B.S.T.J., *54*, No. 8 (October 1975), pp. 1451–1462. See Section II.
10. I. Bar-David, "Communication Under the Poisson Regime," IEEE Trans. Inform. Theory, *IT-15*, No. 1 (January 1969), pp. 31–37.
11. David G. Messerschmitt, "Optimum Mean-Square Equalization for Digital Fiber Optic Systems," *International Conference on Communications Conference Record*, Vol. III, paper 43, June 1975.