THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 52, No. 5, May–June, 1973
Printed in U.S.A.

Utilization of Optical-Frequency Carriers for Low- and Moderate-Bandwidth Channels

By W. M. HUBBARD

(Manuscript received November 10, 1972)

Recent advances in solid-state optical-frequency sources and detectors and low-loss optical fibers make feasible the consideration of optical communication systems for low- and moderate-bandwidth channels (a few kHz to, say, 100 MHz). This paper explores the use of optical-frequency carrier systems for transmission over such channels. Analog intensity modulation, pulse position modulation, delta modulation, and pulse code modulation are considered. This paper is intended to be tutorial in nature.

I. INTRODUCTION

Since the advent of the laser, communication engineers have been intrigued by the promise of fantastic bandwidth capability in optical-frequency communication systems. As a result, attention has been focused on high-capacity, high-bandwidth considerations. Recent advances in component fabrication—for example, light-emitting diodes (LED's), junction lasers, avalanche photodiode detectors, and low-loss optical fibers—have made it feasible to consider the use of optical-frequency carrier systems for moderate- and even low-bandwidth channels.

Fundamental and practical differences between optical-frequency channels and radio-frequency channels* necessitate a reevaluation of concepts acquired from experience with the latter. To this end, in the following sections we consider four potentially attractive forms of modulation of optical-frequency signals and derive results for the required average received signal power in terms of system requirements and parameters. In Section II we consider a system using analog inten-

* We use the term "optical frequency" to mean frequencies roughly in the range 10 to 1000 THz and the term "radio frequency" to indicate frequencies below roughly 3 THz.

sity modulation (IM) of the light source. In some respects such a system is analogous to a baseband system. In Section III we consider pulse position modulation which, because of the nature of optical-frequency sources and detectors, is particularly attractive. In Section IV we consider binary PCM and also delta modulation which we treat as a special case of a binary PCM channel. It is not within the scope of this paper to identify and analyze optimal receivers for these types of modulation. The approach adopted here is rather to analyze the performance of receivers which can be realized and which hold hope of providing an economically attractive approach to transmission of low- and moderate-bandwidth channels.

The four types of systems to be discussed in the following sections are considered in terms of their applicability to the problem of transmitting a comparatively narrow information bandwidth b over a channel with noise (signal) bandwidth $\mathcal{B}$. The features which these systems have in common will be discussed in this section.

Two types of signal sources are considered in this discussion–lasers and light-emitting diodes. For the purposes of this discussion, the difference in these two sources is that the laser possesses a substantial degree of temporal and spatial coherence, while the LED does not. The effect of this incoherence is to give rise to an additional type of noise (later referred to as beat noise) in an LED system. We shall see in the following calculations, however, that this noise is usually negligible in systems of interest. This is because the beat noise is proportional to the ratio $\mathcal{B}/(WJ)$ where W is the spectral width of the LED and J is the number of spatial modes of the signal viewed by the receiver. For typical GaAs LED's, $W \approx 20 \times 10^{12}$ Hz and J is an integer which depends on the details of the channel between transmitter and receiver and which is usually very large.

The receiver in all cases is assumed to begin with an avalanche photodiode with quantum efficiency η and avalanche current gain G. This photodiode is followed by a baseband amplifier which presents a load resistance R to the photodiode and which has a noise figure F_t. Only direct detection receivers are considered in this treatment, since recently developed avalanche photodetectors make heterodyne and homodyne methods look quite unattractive in view of the difficulties encountered in phase-front matching in such systems.*

II. ANALOG INTENSITY MODULATION

The simplest form of modulation is analog intensity modulation. Both light-emitting diodes and double-heterostructure junction lasers

* See appendix for elaboration on this point.

have output power versus bias current characteristics which are sufficiently linear over a reasonable range that they can be modulated directly by modulating their bias currents. Modulation depths of up to about 85 percent can be achieved with suitable light-emitting diodes with very small harmonic distortion. It should be noted that the optical power, not amplitude, is proportional to the drive signal but, since the photodetector is a square-law device, its output current is proportional to the received power. Thus, in many respects, an intensity-modulated optical system can be regarded as equivalent to a baseband system with a transducer (the light-emitting diode or laser) which converts electrons into photons and a subsequent transducer (the photodetector) which converts photons back into electrons.

For a sinusoidally modulated carrier with modulation index m, the mean-square signal current in the photodetector output is given by

$$\langle i_s^2 \rangle = \frac{1}{2}\left[\eta \frac{e}{h\nu} G m p_o\right]^2.$$

When a coherent source is used, the mean-square noise current in the photodetector output is the resultant of the five noise currents described below.

The most important (in most applications) of the noise currents is the quantum noise with its mean-square given by

$$\langle i_Q^2 \rangle = 2e \frac{e}{h\nu} \eta p_o G^2 F_d b = N_Q G^2 F_d$$

where in this and the following equations e is the electronic charge, $h\nu$ is the energy per photon, η is the quantum efficiency of the photodiode, p_o is the average received optical power, b is the bandwidth of the information source (which in the case of analog intensity modulation is equal to the bandwidth of the channel), G is the avalanche gain of the photodetector, and F_d is a noise figure associated with the random nature of the avalanche process. F_d is, in general, a function of G which for silicon is well approximated[1,2] by $F_d = \sqrt{G}$. N_Q is the value of the quantum noise in the absence of avalanche gain.

The next most important noise source is the thermal-noise current with mean-square value

$$\langle i_T^2 \rangle = \frac{4kT}{R_{eq}} b F_t = N_T$$

where kT is Boltzmann's constant times the absolute temperature, R_{eq} is an equivalent load resistance, and F_t is the noise figure of the (baseband) amplifier.

The dark-current noise can be rendered negligible by suitable choice of photodetector. At the present time, this generally dictates that the photodetector be made of silicon. There are actually two kinds of dark-current noise. The first, which will be referred to simply as dark current in the following, consists of electrons (and/or holes) which are thermally liberated in the pn junction and which experience the avalanche gain G. The mean-square value of this current is given by:

$$\langle i_D^2 \rangle = 2eI_d' G^2 F_d b = G^2 F_d N_D$$

when I_d' is the primary detector dark current. The other "dark current," which will henceforth be referred to as leakage current, bypasses the drift region and experiences no avalanche gain. The mean-square value of this current is therefore given by:

$$\langle i_L^2 \rangle = 2eI_L b = N_L$$

where I_L is the leakage current.

Finally, if there is incoherent background radiation with average power p_G incident on the detector, there will be an additional noise current given by:

$$\langle i_G^2 \rangle = 2e \frac{e}{h\nu} \eta p_G G^2 F_d b = G^2 F_d N_G.$$

(This assumes that the background radiation is at about the same wavelength as the signal. This is justified since other wavelengths could be effectively removed by filters.)

Since $\langle i_G^2 \rangle$ and $\langle i_D^2 \rangle$ have the same form, we can simply write

$$I_d = I_d' + \frac{e}{h\nu} \eta p_G$$

and lump both of these terms into an effective dark current. This is done in the following calculations.

When an incoherent source such as a light-emitting diode is used, there is an additional noise term due to the beats between spectral components within the spectral width of the source. This phenomenon gives rise to a noise current with variance

$$\langle i_B^2 \rangle = 2 \left(\frac{e}{h\nu} G\eta p_o \right)^2 \frac{b}{JW} \left(1 - \frac{1}{2} \frac{b}{W} \right) = G^2 N_B$$

where W is the spectral width of the source and J is the number of spatial modes of the source which are viewed by the receiver.[3] (In most cases the ratio b/JW renders this term negligible.) The factor $(1 - b/(2W))$ is always very nearly unity in cases of interest.

In each case, $\langle i_X^2 \rangle$ represents the mean-square value of the corresponding noise current after avalanche gain and N_X represents the value it would have in the absence of avalanche gain.

Thus, the signal-to-noise ratio is given by

$$\text{SNR} = \frac{\frac{1}{2}\left(\eta \frac{e}{h\nu} Gmp_o\right)^2}{\langle i_Q^2 \rangle + \langle i_T^2 \rangle + \langle i_D^2 \rangle + \langle i_L^2 \rangle + \langle i_B^2 \rangle}. \tag{1}$$

It is instructive to consider the behavior of (1) for some particular cases. For our examples we use the following numerical values throughout this paper:

$$\lambda = 0.85\ \mu\text{m} \qquad R = 10^3\ \text{ohms}$$
$$\eta = 0.5 \qquad b = 4\ \text{kHz}$$
$$m = 0.85.$$

Figure 1 shows the signal-to-noise ratio (expressed in dB) computed from (1) for $I_d = 10^{-9}$ A, $I_L = 10^{-8}$ A, $WJ = 10^{15}$ Hz, and $G = 10$. It must be emphasized that these values are chosen for illustrative purposes only and are not meant to be typical of a real receiver. This choice of parameters allows us to consider the form of the contribution of each noise source to the resulting SNR. The curves labeled Q, T, D, L, and B are the ratio of the mean signal power to the mean quantum-noise power, mean thermal-noise power, mean dark-current-noise power, mean leakage-current-noise power, and mean beat-noise power, respectively. We observe that for this comparatively low value of G the thermal noise is dominant over a large range of incident light power p_o. Then for -30 dBm $< p_o < -15$ dBm the quantum noise dominates the picture. Finally, for larger p_o, the beat noise clamps a ceiling on the SNR. The dark current and leakage current are unimportant.

Figure 2 shows the same curves for the same illustrative parameters as Fig. 1 except that G is now taken to be 100. Two differences between Fig. 1 and Fig. 2 are immediately apparent. First, the increased gain has caused the quantum-limited region to extend to lower values of p_o and, second, the dark-current noise is more important relative to the thermal noise.

In Figs. 1 and 2, F_d has been taken to be given by $G^{\frac{1}{2}}$. Thus, in Fig. 2, for example, $F_d = 10$ and one must be cautious about referring to the behavior as "quantum limited" in the "quantum-limited region." It is "quantum limited" only in the sense that the quantum noise term dominates the other noise terms, but the actual results are 10 dB poorer than could be achieved if the avalanche gain process were noise free. Figure 3 shows the same curves but for the values $I_d = 10^{-10}$ A,

$I_L = 10^{-9}$ A, which are typical of good, but available, silicon photodiodes;[4] $WJ = 10^{17}$ Hz which is typical of GaAs luminescent diodes and multimode optical fibers; and $G = 17$. We see that in this case beat

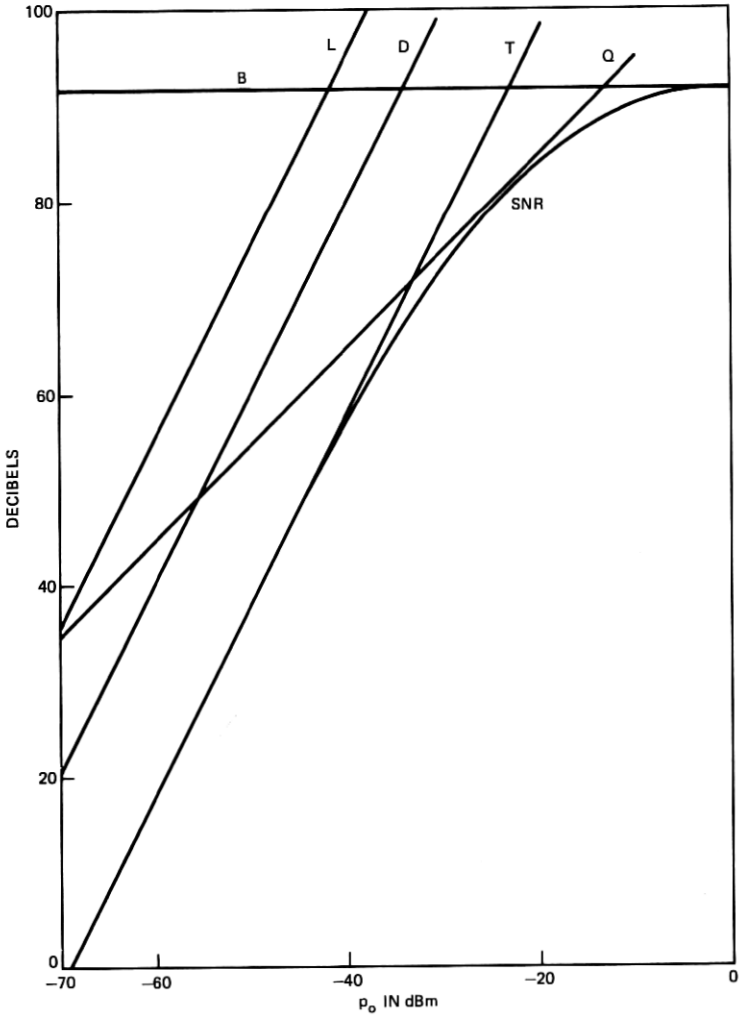

Fig. 1—Signal-to-noise ratio, SNR, and ratio of mean signal power to each component of the mean noise power for the illustrative parameters: $\lambda = 0.85\ \mu$m, $\eta = 0.5$, $M = 0.85$, $R = 1000$ ohms, $I_d = 10^{-9}$ A, $I_L = 10^{-8}$ A, $WJ = 10^{15}$ Hz, $b = 4000$ Hz, $G = 10$, for an IM channel. B = ratio of mean signal power to beat-noise power, L = ratio of mean signal power to leakage-current-noise power, D = ratio of mean signal power to dark-current-noise power, T = ratio of mean signal power to thermal-noise power, Q = ratio of mean signal power to quantum-noise power.

noise, leakage noise, and dark-current noise are unimportant and the quantum-excess noise controls above $p_o = -30$ dBm with thermal noise controlling below.

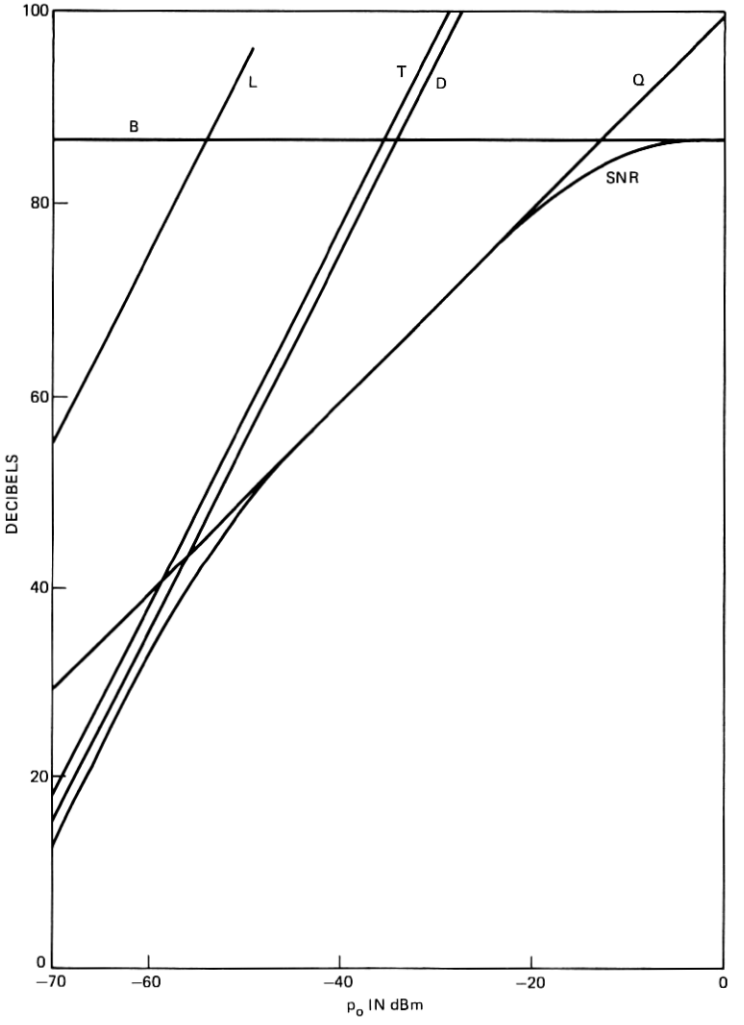

Fig. 2—Signal-to-noise ratio, SNR, and ratio of mean signal power to each component of the mean noise power for the illustrative parameters: $\lambda = 0.85$ μm, $\eta = 0.5$, $M = 0.85$, $R = 1000$ ohms, $I_d = 10^{-9}$ A, $I_L = 10^{-8}$ A, $WJ = 10^{15}$ Hz, $b = 4000$ Hz, $G = 100$, for an IM channel. B = ratio of mean signal power to beat-noise power, L = ratio of mean signal power to leakage-current-noise power, D = ratio of mean signal power to dark-current-noise power, T = ratio of mean signal power to thermal-noise power, Q = ratio of mean signal power to quantum-noise power.

It is now well known[1,2,5] that the excess noise figure F_d of an avalanche photodiode increases (in most cases) with increasing gain. In particular, for silicon photodiodes, F_d is well approximated by $G^{\frac{1}{2}}$.

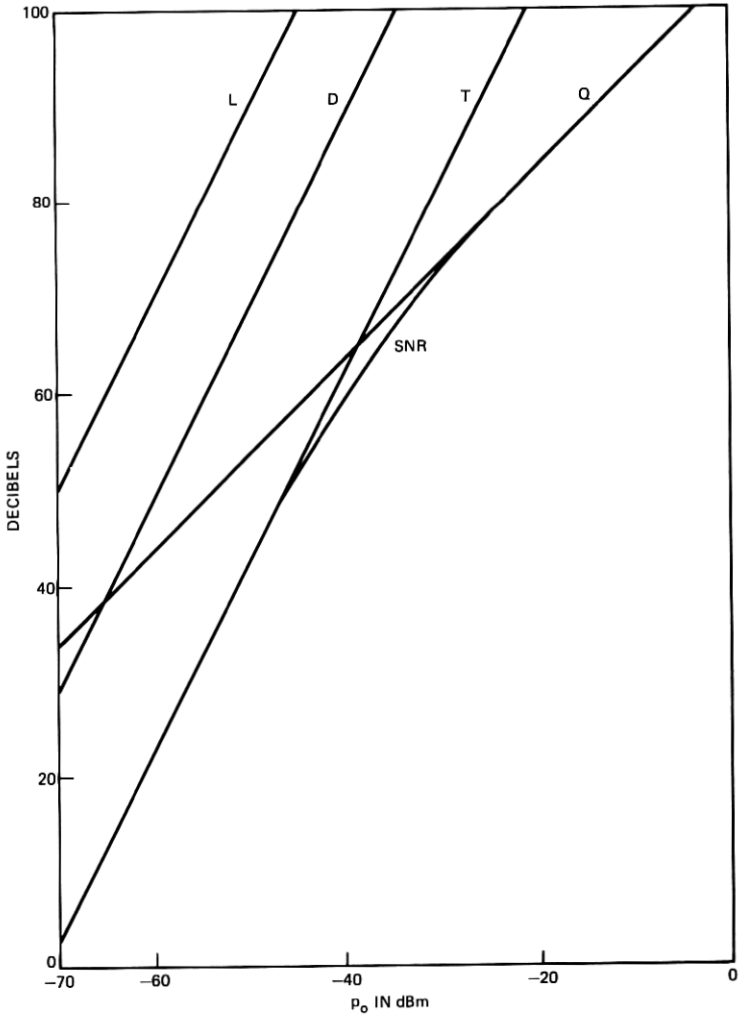

Fig. 3—Signal-to-noise ratio, SNR, and ratio of mean signal power to each component of the mean noise power for the typical parameters: $\lambda = 0.85\ \mu$m, $\eta = 0.5$, $M = 0.85$, $R = 1000$ ohms, $I_d = 10^{-10}$ A, $I_L = 10^{-9}$ A, $WJ = 10^{17}$ Hz, $b = 4000$ Hz, $G = 17$, for an IM channel. B = ratio of mean signal power to beat-noise power, L = ratio of mean signal power to leakage-current-noise power, D = ratio of mean signal power to dark-current-noise power, T = ratio of mean signal power to thermal-noise power, Q = ratio of mean signal power to quantum-noise power.

When a form $F_d = G^r$ is assumed, SNR as a function of G has a maximum given by

$$\mathrm{SNR}_{\max} = \frac{\langle i_S^2 \rangle G^{-2}}{\left(\frac{2\gamma}{r}\right)^{r/(2+r)} \beta^{2/(2+r)} \left(1 + \frac{r}{2}\right) + N_B}$$

for

$$G = G_{\mathrm{opt}} = \left(\frac{2\gamma}{r\beta}\right)^{1/(2+r)}$$

where

$$\gamma = \langle i_T^2 \rangle + \langle i_L^2 \rangle = N_T + N_L,$$

the variance of the gain-independent noise, and

$$\beta = G^{-(1+r)}[\langle i_Q^2 \rangle + \langle i_D^2 \rangle] = N_Q + N_D,$$

the variance of the gain-dependent noise *before* the gain process. For $r = \frac{1}{2}$ (silicon) this reduces to

$$G_{\mathrm{opt}} = \left[\frac{4(N_T + N_L)}{N_Q + N_D}\right]^{2/5} \tag{2}$$

$$\mathrm{SNR}_{\max} = \frac{\frac{2}{5}\left[\eta \frac{e}{h\nu} m p_o\right]^2}{(4\gamma)^{1/5}\beta^{4/5} + N_B}. \tag{3}$$

It is interesting to note in passing that the condition for G to be optimum is that

$$\beta G_{\mathrm{opt}}^{2+r} = \frac{2}{r}\gamma.$$

But the left-hand side of this equation is the total mean-square current due to gain-dependent noise (excluding beat noise), while the right-hand side is $2/r$ times the total mean-square current due to gain-independent noise. For $r = 0.5$, for example, the gain is optimum when the gain-dependent noise exceeds the gain-independent noise by 6 dB.

This result is illustrated in Fig. 4 for the same parameters used in Figs. 1 and 2 except that in Fig. 4 $G = G_{\mathrm{opt}}$, which is a function of p_o according to (2). First we observe that G_{opt} varies from 132 to 1 over the range of p_o plotted in these figures. These are values which are readily achievable with existing photodiodes. We see that the thermal noise (leakage noise remains negligible) is just a constant 6 dB below the sum of the quantum- and dark-current noises as the condition for optimum gain dictates.

Figure 5 illustrates the optimum gain behavior for the typical device

parameters used in Fig. 3. Note that here the optimum gain becomes rather large below about $p_o = -60$ dB and might be difficult or impossible to realize in practice.

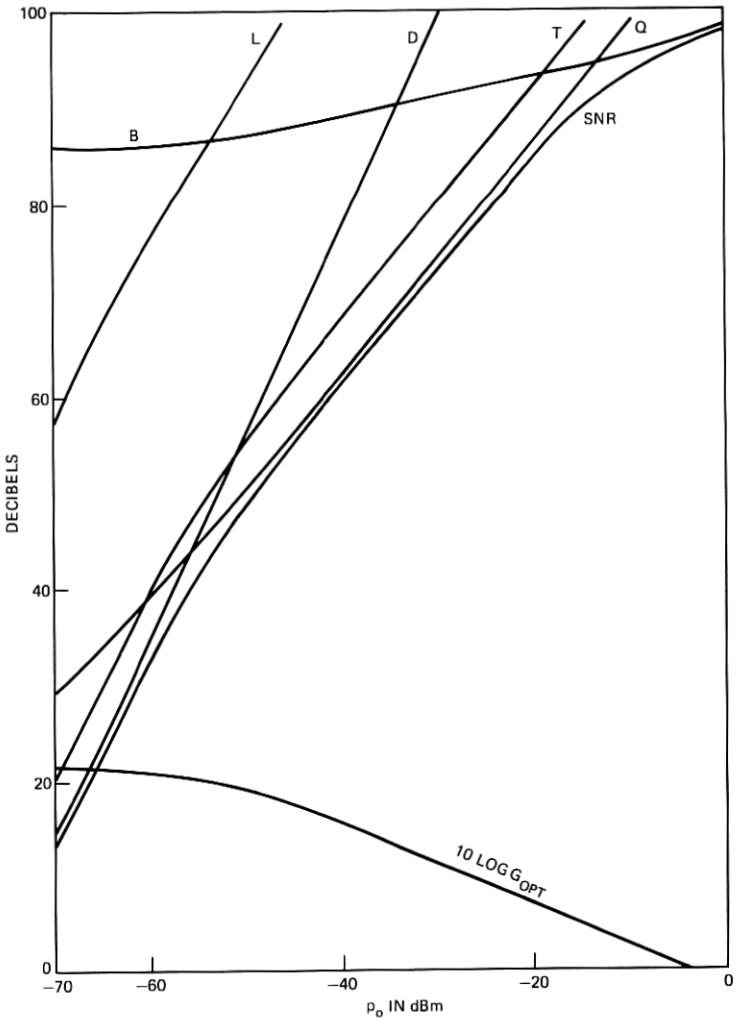

Fig. 4.—Signal-to-noise ratio, SNR, and ratio of mean signal power to each component of the mean noise power for the illustrative parameters: $\lambda = 0.85$ μm, $\eta = 0.5$, $M = 0.85$, $R = 1000$ ohms, $I_d = 10^{-9}$ A, $I_L = 10^{-8}$ A, $WJ = 10^{15}$ Hz, $b = 4000$ Hz, $G = G_{opt}$, for an IM channel. B = ratio of mean signal power to beat-noise power, L = ratio of mean signal power to leakage-current-noise power, D = ratio of mean signal power to dark-current-noise power, T = ratio of mean signal power to thermal-noise power, Q = ratio of mean signal power to quantum-noise power.

The value of $R = 10^3$ ohms is used consistently in the numerical examples throughout this paper. The detector c is usually presumed to dictate the maximum value of load resistance R through the relationship

$$R < \frac{1}{4c\mathfrak{B}}$$

where $\mathfrak{B}$ is the bandwidth of the signal; but in practice it is often beneficial to use a much larger value of R than this and equalize the resultant signal distortion later on in the receiver. It is valid to object to the use of so low a resistance for a 4-kHz channel [the value is more appropriate to the other types of systems (without equalization) to be considered in the following sections]. However, (3) shows that at optimum avalanche gain when, as is usually the case, the leakage current is negligible, SNR $\propto R^{1/5}$, so very little is gained by going to larger load resistors except in the region where G_{opt} is so large that it is difficult to achieve. Here the fact that $G_{opt} \propto R^{-2/5}$ may be important.

It is useful to present the results of (3) graphically in a form suitable for system design. This can be done in a very general and simple manner, when the leakage current is negligible, by defining the quantities

$$\Psi = 10 \log \left\{ \frac{\mathrm{SNR}_{max} b}{m^2} \right\}$$

$$x = 10 \log \left\{ \frac{\eta p_o}{h\nu} \right\}$$

where, in the argument of the first logarithm, b is taken to be dimensionless, i.e., it is interpreted as the ratio of bandwidth in Hz to 1 Hz. Ψ vs x is plotted in Fig. 6 for $R/F_t = 10^3$ ohms. Figure 6 can be used as a computational aid as follows. Suppose one needs to design a system with an SNR of 70 dB, a bandwidth of 4 kHz. Suppose further that a modulation index of 0.85 is possible with available devices. Then the value of Ψ which characterizes such a system is 107.4 which, Fig. 6 tells us, can be achieved if $x = 121$. Now $x = 121$ means that $\eta p_o/h\nu = 1.26 \times 10^{12}$; for $\eta = 0.5$ and $\lambda = 0.85$ this gives $p_o = 5.88 \times 10^{-7}$ W which corresponds to -32.5 dBm. If a value of R/F_t which differs from 1000 ohms is desired, Ψ can be modified according to $\Delta\Psi = 2 \log [(R/F_t)/1000]$.

Figure 7 can be used to determine the value of G required to achieve the result computed from Fig. 6. If a value of R/F_t which differs from 1000 ohms was used, recall that $G_{opt} \propto R^{-2/5}$.

III. PULSE POSITION MODULATION

Considerable improvement in noise immunity can be achieved by properly exploiting the wide available bandwidth of optical systems.

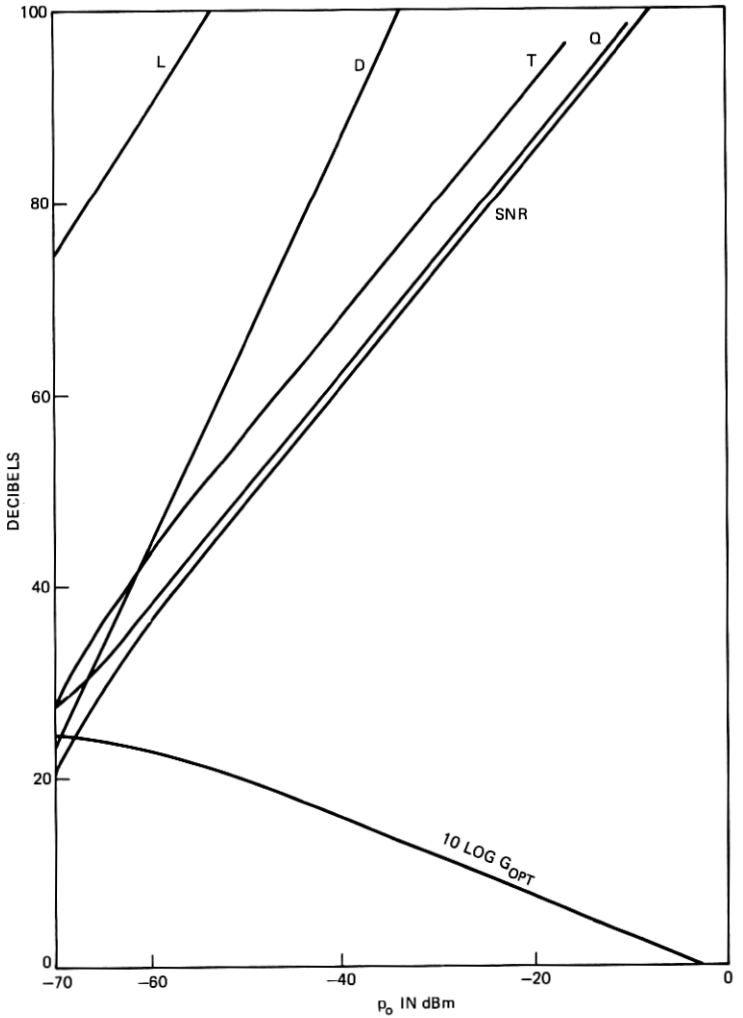

Fig. 5—Signal-to-noise ratio, SNR, and ratio of mean signal power to each component of the mean noise power for the typical parameters: $\lambda = 0.85$ μm, $\eta = 0.5$, $M = 0.85$, $R = 1000$ ohms, $I_d = 10^{-10}$ A, $I_L = 10^{-9}$ A, $WJ = 10^{17}$ Hz, $b = 4000$ Hz, $G = G_{opt}$, for an IM channel. B = ratio of mean signal power to beat-noise power, L = ratio of mean signal power to leakage-current-noise power, D = ratio of mean signal power to dark-current-noise power, T = ratio of mean signal power to thermal-noise power, Q = ratio of mean signal power to quantum-noise power.

Pulse position modulation (PPM) offers an attractive method of accomplishing this end. Figure 8 shows a block diagram of the system to be analyzed in this section.

The pulse-position modulation signal is encoded by sampling the message signal periodically at times nT (where n is an integer and T is the sampling interval or time slot duration). The value v_n of the nth sample is transmitted during the nth time slot by sending a short pulse

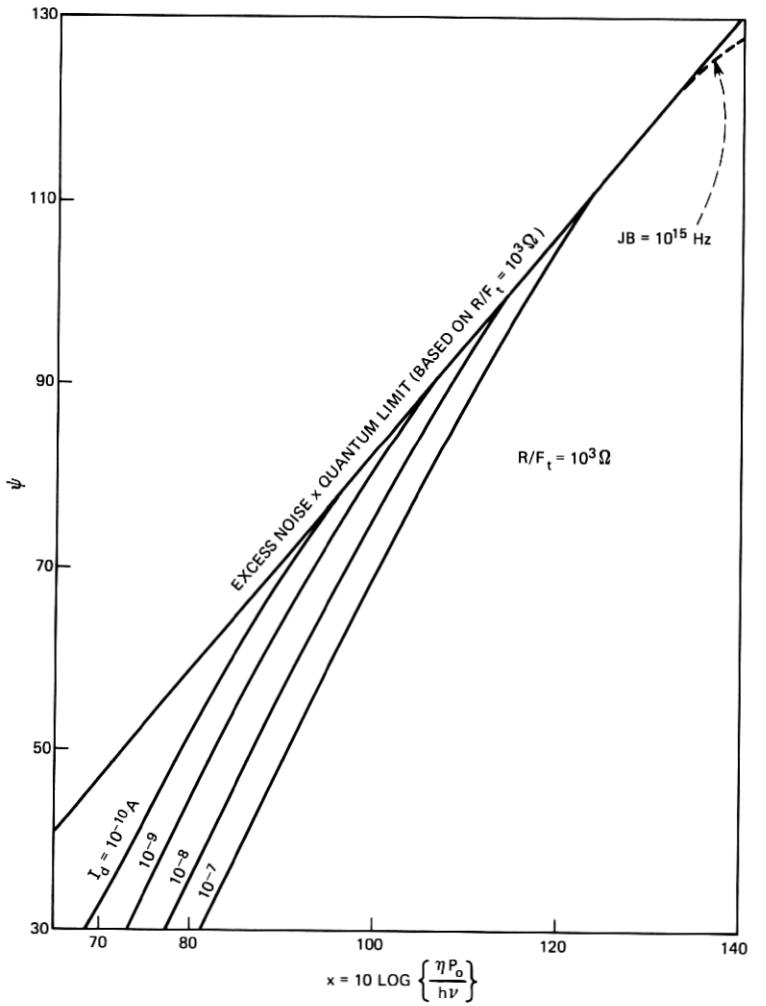

Fig. 6—$\psi = 10 \log \left\{\frac{\mathrm{SNR}_{\max} b}{M^2}\right\}$ vs $x = 10 \log \left\{\frac{\eta P_o}{h\nu}\right\}$ for $R/F_T = 1000$ ohms.

of optical energy at a time which is shifted from the center of the nth time slot by an amount proportional to v_n.

At the receiver, the values of v_n are recovered by measuring the time interval between the center of the time slot and the time at which the amplified output current from the photodetector crosses a threshold. This system is described in some detail in Ref. 6.

Consider a PPM signal consisting of a sequence of light pulses whose

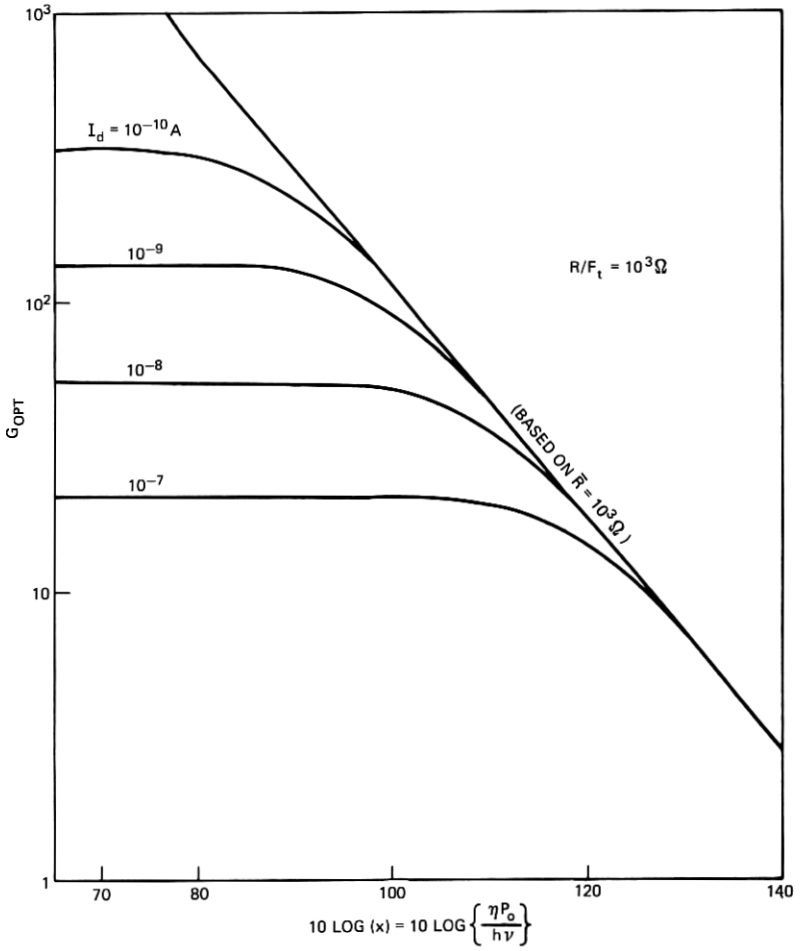

Fig. 7—Optimum gain vs $x = 10 \log \left\{\frac{\eta P_o}{h\nu}\right\}$ for $R/F_T = 1000$ ohms.

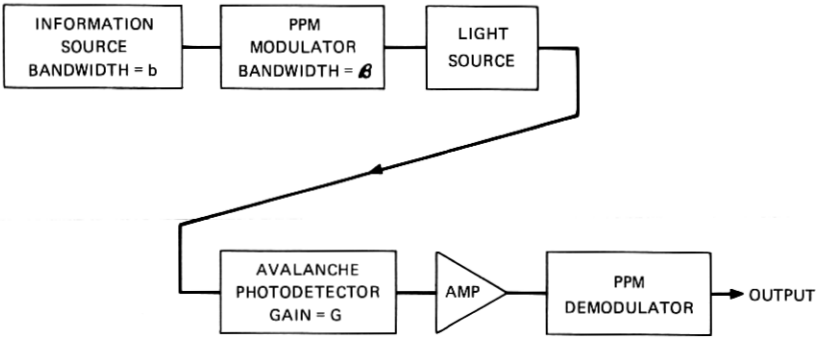

Fig. 8—Block diagram of a PPM channel.

power $P(t)$ is of the form

$$P(t) = \frac{1}{2}\left(1 + \cos\left(\pi \frac{t}{\mathcal{T}}\right)\right) P_m \qquad -\mathcal{T} < t < \mathcal{T} \tag{4}$$

where P_m is the peak power achieved by the pulse and $2\mathcal{T}$ is the total pulse duration. The average power P_o in this signal is related to P_m by

$$P_o = \frac{\mathcal{T}}{T} P_m = \frac{1}{2\kappa} P_m \tag{5}$$

where $\kappa \equiv T/2\mathcal{T}$. The detected current pulse in the receiver (neglecting noise) is given by

$$i(t) = \frac{1}{2}\left(1 + \cos\left(\pi \frac{t}{\mathcal{T}}\right)\right) i_m$$

where $i_m = \eta(e/h\nu)Gp_m = \eta(e/h\nu)G(T/\mathcal{T})p_o$ is the peak current, $p_m = AP_m$, $p_o = AP_o$, A = attenuation between transmitter and receiver.

Noise affects the SNR of a PPM signal in two ways. First, it can perturb the time of the threshold-crossing of the received signal and thereby effectively shift the position of the pulse. This is the predominant effect when the bandwidth expansion is small. Second, the noise can cause the received current to exceed the threshold in the absence of the signal pulse, thereby triggering a "false alarm" in the circuit.

First, we consider the perturbation of the time of threshold-crossing due to the noise. Assume that the threshold current level is one-half of

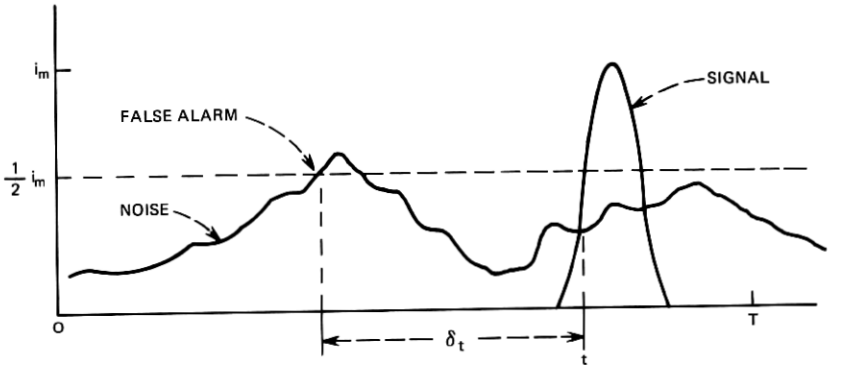

Fig. 9—Illustration of a "false alarm" or threshold violation.

the peak current level,* i.e., $\frac{1}{2}i_m$. The perturbation τ in the position of the pulse due to a noise current i_n is†

$$\tau = i_n \left. \frac{1}{\frac{di(t)}{dt}} \right|_{t=\mathcal{T}/2}$$

$$= \frac{2}{\pi} \frac{i_n}{i_m} \mathcal{T}. \tag{6}$$

The expectation value of τ^2 is then

$$\langle \tau^2 \rangle = \frac{4}{\pi^2} \frac{\langle i_n^2 \rangle}{i_m^2} \mathcal{T}^2$$

$$= \frac{1}{\pi^2} \frac{\langle i_n^2 \rangle \mathcal{T}^2}{\left(\eta \frac{e}{h\nu} G\kappa p_o \right)^2} \tag{7}$$

where $\langle i_n^2 \rangle$ is the expected value of the square of the noise current.

Now consider the contribution to the noise due to "false alarms" generated when the noise current exceeds the threshold value $\frac{1}{2}i_m$. Consider a time slot $(0, T)$ as illustrated in Fig. 9. Let $t \in (0, T)$ be the time at which the signal pulse crosses the threshold. Assume that a false alarm occurs in this time slot. It must occur (with uniform probability density) on the interval $(0, t)$, since the receiver, having sensed a pulse at t, is disabled on the interval (t, T). The mean-square value of the error in t is therefore

$$\langle \delta t^2 \rangle = \int_0^t p(\delta t)(\delta t)^2 d(\delta t) = \frac{t^2}{3}.$$

* In practice, a slightly different threshold may be optimum due to the details of the pulse shape.

† The remainder of this paragraph follows the derivation on pages 256–257 of Ref. 6.

Since we have no statistical information on t, we make an approximation which is clearly very conservative: We assume that t always has its largest possible value, $T - 2\mathcal{T}$. This gives

$$\langle \delta t^2 \rangle = \frac{(T - 2\mathcal{T})^2}{3} = \tfrac{4}{3}\mathcal{T}^2(\kappa - 1)^2. \tag{8}$$

Define Π as the probability of a fase alarm occurring during a time interval $T - 2\mathcal{T}$. The mean-square value of the error in t due to false alarms is then $\Pi\langle \delta t^2 \rangle$.

Let $\pm\theta$ be the limits of the allowable variation of the pulse position. The baseband signal-to-noise ratio *at full load,** SNR, is determined as follows. The output signal power due to a sinusoidal input signal which swings the pulse position by an amount $\pm\theta$ about its mean is proportional to $\frac{1}{2}\theta^2$, the output noise power due to perturbation of the threshold crossings of the signal is proportional to $\langle \tau^2 \rangle$, and the output noise power due to false alarms is proportional to $\Pi\langle \delta t^2 \rangle$. Thus we can write

$$\text{SNR} = \frac{\frac{1}{2}\theta^2}{\langle \tau^2 \rangle + \Pi\langle \delta t^2 \rangle}. \tag{9}$$

Since T is the duration of time slot, θ is constrained by the requirement

$$2(\theta + \mathcal{T}) \leqq T.$$

Choosing equality in the above expression gives the best possible SNR; substitution of this along with (7) and (8) into (9) gives

$$\text{SNR} = \frac{\frac{\pi^2}{2}(\kappa - 1)^2 \left(\eta \frac{e}{h\nu} \kappa p_o\right)^2 G^2}{\langle i_n^2 \rangle + \frac{4\pi^2}{3}(\kappa - 1)\Pi \left(\eta \frac{e}{h\nu} \kappa p_o\right)^2 G^2}. \tag{10}$$

The next steps are to evaluate $\langle i_n^2 \rangle$ and Π. We begin by evaluating $\langle i_n^2 \rangle$. The noise currents which make up $\langle i_n^2 \rangle$ for the PPM systems are the same as those which made up $\langle i_n^2 \rangle$ for the analog system except that the noise bandwidth $\mathcal{B}$ is not equal to the signal bandwidth b; and the signal-power-dependent noises are evaluated not at p_o but rather at $p_m/2$ since this is the expected value of the signal when the threshold crossing is to occur.

In the remainder of this section, the reciprocal pulse width, $1/\mathcal{T}$, and the noise bandwidth, $\mathcal{B}$, are assumed to be equal. From (5) one sees that the threshold level $\frac{1}{2}p_m$ is related to the average power by

$$\tfrac{1}{2}p_m = \kappa p_o.$$

* SNR is the ratio of mean signal power when the signal is a sinusoid of maximum allowable amplitude to the mean noise power.

Therefore, the mean-square values of the noise currents given in Section II are appropriate for PPM signals with the substitution of κp_o for p_o and $\mathcal{B}$ for b. Thus,

$$\langle i_n^2 \rangle = \langle i_Q^2 \rangle + \langle i_T^2 \rangle + \langle i_D^2 \rangle + \langle i_L^2 \rangle + \langle i_B^2 \rangle \tag{11}$$

where

$$\langle i_Q^2 \rangle = 2e \frac{e}{h\nu} \eta \kappa p_o G^2 F_d \mathcal{B}, \quad \text{etc.}$$

Now we turn to the problem of evaluating Π, the probability of a threshold violation on the interval $(0, T - 2\mathcal{T})$. We assume for the purposes of this calculation that the noise current,* i_n, during the interval when no signal pulse is present, can be treated as a Gaussian random process strictly bandlimited to the interval $(0, \mathcal{B})$. S. O. Rice[7] computes the probability of such a signal passing a particular value I_1, with positive slope, on the interval Δt to be

$$\frac{1}{\sqrt{3}} \exp - (I_1^2/2\langle \mathrm{i}_n^2 \rangle) \mathcal{B} \Delta t.$$

Thus the probability of the noise alone crossing the threshold $(\frac{1}{2} i_m)$ during the time $T - 2\mathcal{T}$ is

$$\begin{aligned} \Pi &= \frac{1}{\sqrt{3}} \exp \left[(\tfrac{1}{2} i_m)^2 / 2\langle \mathrm{i}_n^2 \rangle \right] \mathcal{B}(T - 2\mathcal{T}) \\ &= \frac{2}{\sqrt{3}} (\kappa - 1) \exp \left(- i_m^2 / 8\langle \mathrm{i}_n^2 \rangle \right). \end{aligned} \tag{12}$$

Now the noise current i_n, which is important for threshold violation, is not the same as the noise current i_n characterized by (11) because there is no signal during the interval between pulses (when threshold violations can occur) and two of the terms which contribute to i_n, namely, the quantum-noise current i_Q and the beat-noise current i_B, are correspondingly absent. Therefore

$$\langle \mathrm{i}_n^2 \rangle = \langle i_D^2 \rangle + \langle i_T^2 \rangle + \langle i_L^2 \rangle.$$

It will turn out that in many cases of interest $\langle \mathrm{i}_n^2 \rangle \ll \langle i_n^2 \rangle$. This result, which has no classical radio-frequency analog, allows considerably more bandwidth-for-signal-power trade in optical systems than in radio-frequency systems.

It is convenient to write (12) as

$$\Pi = \frac{2}{\sqrt{3}} (\kappa - 1) e^{-\mathrm{XNR}/2} \tag{13}$$

* Note the distinction between i_n and i_n of the preceding paragraphs.

where

$$\text{XNR} \equiv \frac{i_m^2}{4\langle i_n^2 \rangle}. \tag{14}$$

We can now substitute (13) into (10) to obtain

$$\text{SNR} = \frac{\frac{\pi^2}{2}(\kappa - 1)^2 \left(\eta \frac{e}{h\nu} \kappa p_o\right)^2 G^2}{\langle i_n^2 \rangle + \frac{8\pi^2}{3\sqrt{3}}(\kappa - 1)^3 e^{-\text{XNR}/2}}. \tag{15}$$

Equation (15) is plotted in Fig. 10, for the same parameters used in Fig. 2, with $\kappa = 250$. A certain similarity in the *relative* positions of the corresponding curves is evident in the two figures, but two differences are also immediately apparent. First, the curves in Fig. 10 are translated (horizontally) to smaller values of p_o and (vertically) to larger values of SNR; second, a threshold is introduced (by the threshold violation term) below which the SNR degrades extremely rapidly. In fact, this threshold term goes from negligible to dominant over about a 1-dB change in p_o.

As in the case of analog IM, there is an optimum value of G in PPM systems. It can be found by differentiating (15) with respect to G and solving for the value of G, which renders this derivative equal to zero. Unfortunately, the resulting expression for G_{opt} is quite complicated. The fact that the threshold effect sets in so rapidly, however, can be exploited to simplify the determination of G_{opt}. Over the range on which the threshold effect is negligible we neglect it, and, as before, obtain G_{opt} as given by (2) [but with the noise terms redefined as described in connection with (11)]; over the range on which the threshold effect is dominant, one readily obtains

$$G_{\text{opt}} = \left[\frac{4(N_L + N_T)}{N_D}\right]^{2/5}.$$

Figure 11 illustrates (15) for optimum gain for the same set of parameters used in Fig. 4. Figure 12 presents these results for a typical set of parameters.

It is interesting to compare (15) for SNR with the result obtained in Section II for the signal-to-noise ratio in an intensity modulated system with modulation index m. In order to do this, we first observe that, in order to properly sample a signal of bandwidth b, the sampling rate must be (at least) $2b$. This gives the relationships

$$T = \frac{1}{2b}, \qquad \kappa = \frac{T}{2\mathcal{T}} = \frac{\mathcal{B}}{4b}. \tag{16}$$

If we substitute $4bK$ for $\mathcal{B}$ in the noise terms in (15), we find that, above threshold, the expression for SNR in a PPM system is formally identical to that in an analog IM system (1) except that we make the

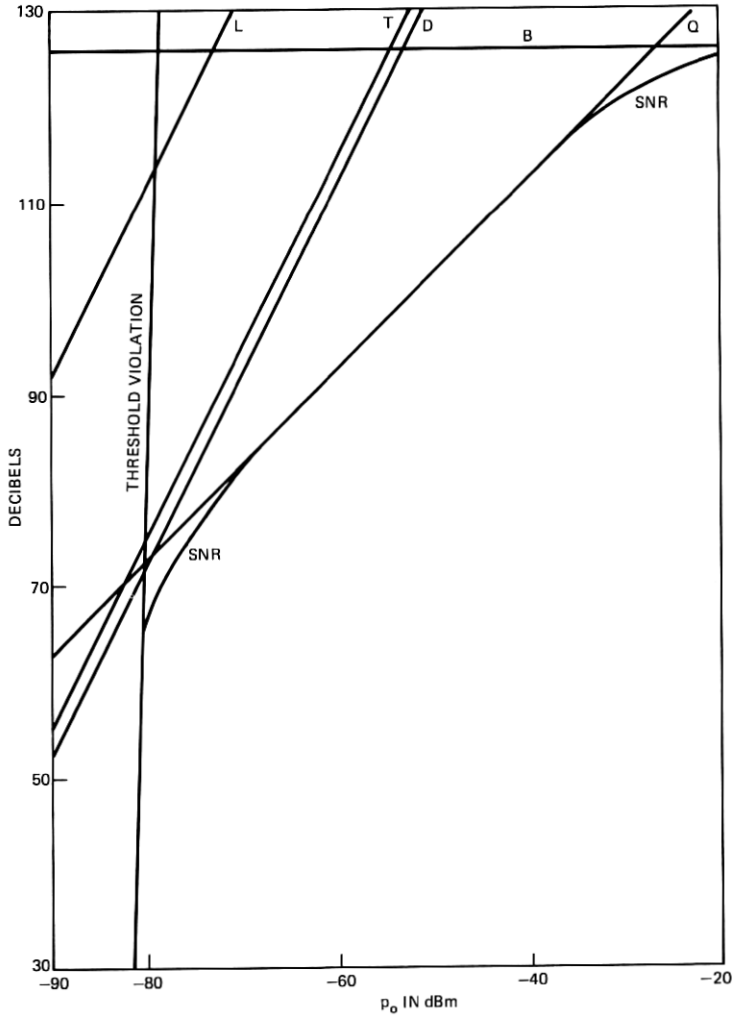

Fig. 10—Signal-to-noise ratio, SNR, and ratio of mean signal power to each component of the mean noise power for the illustrative parameters: $\lambda = 0.85\ \mu$m, $\eta = 0.5$, $M = 0.85$, $R = 1000$ ohms, $I_d = 10^{-9}$ A, $I_L = 10^{-8}$ A, $WJ = 10^{15}$ Hz, $b = 4000$ Hz, $G = 100$, for a PPM channel. B = ratio of mean signal power to beat-noise power, L = ratio of mean signal power to leakage-current-noise power, D = ratio of mean signal power to dark-current-noise power, T = ratio of mean signal power to thermal-noise power, Q = ratio of mean signal power to quantum-noise power, X = ratio of mean signal power to noise power due to threshold violations.

replacements:

$$p_o \rightarrow \kappa p_o, \qquad m \rightarrow \frac{\pi}{2}\frac{\kappa - 1}{\sqrt{\kappa}}.$$

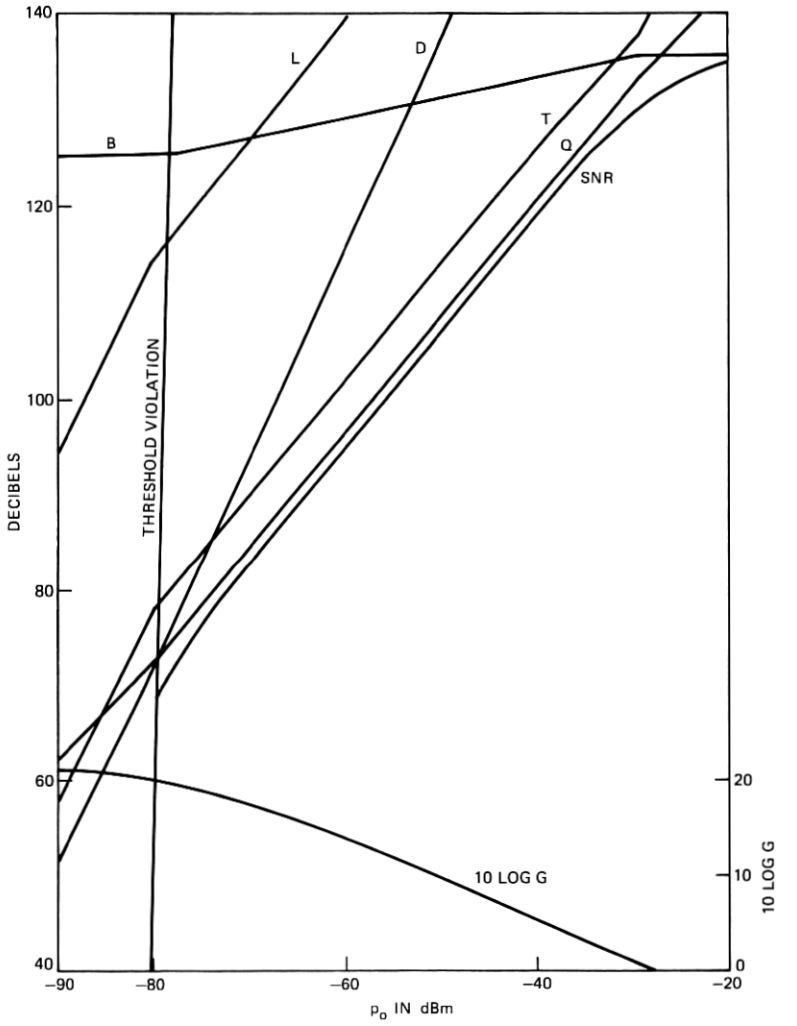

Fig. 11.—Signal-to-noise ratio, SNR, and ratio of mean signal power to each component of the mean noise power for the illustrative parameters: $\lambda = 0.85\ \mu m$, $\eta = 0.5$, $M = 0.85$, $R = 1000$ ohms, $I_d = 10^{-9}$ A, $I_L = 10^{-8}$ A, $WJ = 10^{15}$ Hz, $b = 4000$ Hz, $G = G_{opt}$, for a PPM channel. B = ratio of mean signal power to beat-noise power, L = ratio of mean signal power to leakage-current-noise power, D = ratio of mean signal power to dark-current-noise power, T = ratio of mean signal power to thermal-noise power, Q = ratio of mean signal power to quantum-noise power, X = ratio of mean signal power to noise power due to threshold violations.

From this we see that the PPM system yields considerable improvement over the IM system. For PPM, the average signal power is effectively increased by a factor κ and the modulation index (which is less

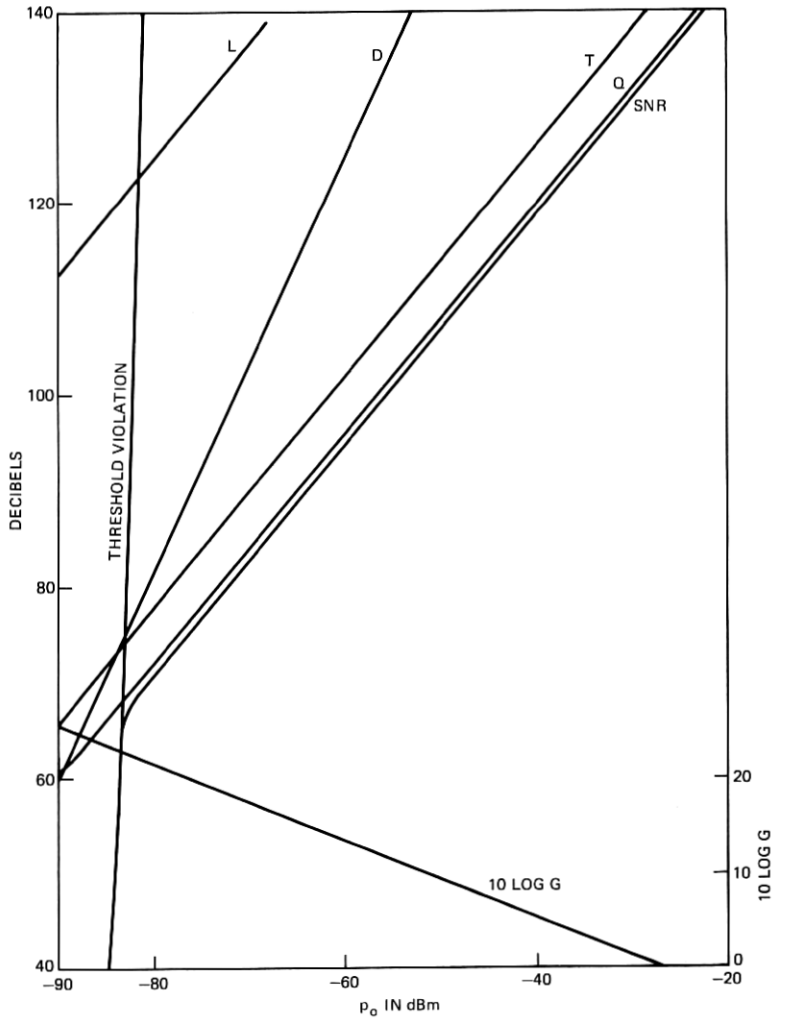

Fig. 12—Signal-to-noise ratio, SNR, and ratio of mean signal power to each component of the mean noise power for the typical parameters: $\lambda = 0.85$ μm, $\eta = 0.5$, $M = 0.85$, $R = 1000$ ohms, $I_d = 10^{-10}$ A, $I_L = 10^{-9}$ A, $WJ = 10^{17}$ Hz, $b = 4000$ Hz, $G = G_{opt}$, for a PPM channel. B = ratio of mean signal power to beat-noise power, L = ratio of mean signal power to leakage-current-noise power, D = ratio of mean signal power to dark-current-noise power, T = ratio of mean signal power to thermal-noise power, Q = ratio of mean signal power to quantum-noise power, X = ratio of mean signal power to noise power due to threshold violations.

than 1 for an IM system) is replaced by $(\pi/2)(\kappa - 1)/\sqrt{\kappa}$ which can be substantially larger than 1.

In practical applications, bandwidth expansion factors of over a thousand are sometimes possible before threshold violations become important.

3.1 *Power Available*

It is tempting to hypothesize that the average light power obtainable from a given diode is proportional to the average thermal power which can be dissipated in the device without causing catastrophic failure. The thermal power dissipated by the device, P_I [for the signal given by (4)], is

$$P_I = \tfrac{1}{4} I_m^2 \frac{\mathfrak{R}}{T} \int_{-T}^{T} \left(1 + \cos \pi \frac{t}{T}\right)^2 dt$$

$$= \frac{3}{8\kappa} \mathfrak{R} I_m^2 \tag{17}$$

where $\mathfrak{R}$ is the effective resistance of the device. This gives

$$P_o = \mu \sqrt{\frac{2}{3} \frac{P_I}{\kappa \mathfrak{R}}} \tag{18}$$

where μ is a constant of proportionality such that $P_m = \mu I_m$. Thus for P_I and $\mathfrak{R}$ fixed, P_o (and hence p_o) varies as $\kappa^{-\frac{1}{2}}$.

Unfortunately, the behavior of real LED's and injection lasers is not this simple. First, the heat capacity of some LED's is so small that if a step function change occurs in the diode current, the diode temperature reaches its new steady-state value very quickly. For example, some diodes have such small heat capacity that burnout occurred whenever pulse duration exceeded a few microseconds (independent of duty cycle) if the peak pulse current exceeds the tolerable dc value.[8] Second, even when the pulse is short enough to avoid this problem, the concept of constant power dissipation is not exactly correct. For example, the peak current may be limited by saturation effects.

We observe experimentally for diodes of the type described in Ref. 8 that, for a pulse repetition rate of 8 kHz, the maximum peak pulse power achievable with pulses of 0.1 to 0.25 μs duration is about 5 dB less than that predicted by the constant power dissipation model. Nevertheless, the model is useful as a qualitative guide to diode behavior in pulsed operation.

IV. DIGITAL BINARY PULSE CODE MODULATION

A second method of trading bandwidth for noise immunity is the use of pulse code modulation (PCM). This also has the advantage of being

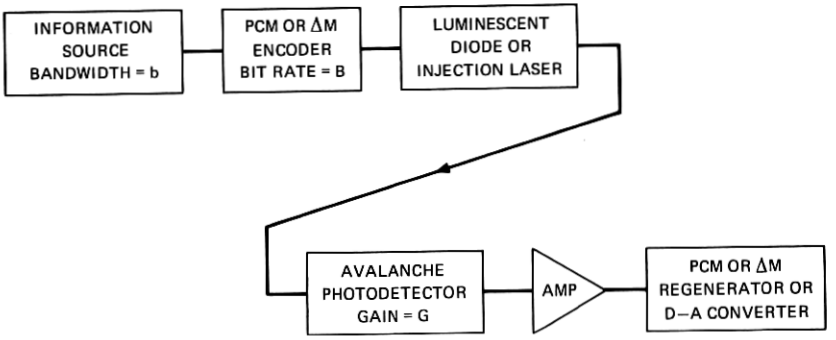

Fig. 13—Block diagram of a single-channel PCM optical communication system.

readily compatible with digital data transmission. Analysis of a digital PCM channel is somewhat different from that of an analog channel in that the parameter used to characterize a PCM channel is not SNR but rather the error probability P_e.

Memoryless binary optical digital communication systems can be divided into two broad classes–single-channel systems in which the information is coded such that a pulse of energy represents a "1" and no pulse represents a "0"; and twin-channel systems in which energy is transmitted for both information states, but the signal is modulated in such a manner that an appropriate device in the receiver routes the signal energy into one of two channels when a "1" is transmitted and into the other when a "0" is transmitted. It has been shown,[9] however, that unlike the classical radio-frequency case, the twin-channel receiver offers little, if any, advantage in an optical system. In fact, if the transmitter is average-power limited, a single-channel receiver has at least a 1.5-dB advantage in noise immunity over a twin-channel receiver; if the transmitter is peak-power limited, the single-channel receiver suffers, at worst, a 1.5-dB disadvantage. Since the single-channel system is considerably easier to implement than the twin-channel system, and since the twin-channel system offers no significant advantages, we confine our treatment to a single-channel system. Figure 13 is a block diagram of such a system.

It has previously been shown[9,10] that if we assume that the avalanche current gain G is deterministic,* the probability that the receiver

* By this we mean that if m primary electrons are liberated, exactly Gm electrons will be delivered to the load. This artificial constraint will be relaxed in the next section.

mistakes a "0" for a "1" is given by

$$P(1|0) = \tfrac{1}{2} \sum_{n=0}^{\infty} p_o(n) \operatorname{erfc} \left\{ \frac{x_t - nG}{\sqrt{2\langle x_T^2 \rangle}} \right\} \tag{19}$$

and the probability that it mistakes a "1" for a "0" is

$$P(0|1) = \tfrac{1}{2} \sum_{n=0}^{\infty} p_1(n) \operatorname{erfc} \left\{ \frac{nG - x_t}{\sqrt{2\langle x_T^2 \rangle}} \right\} \tag{20}$$

where $\langle x_T^2 \rangle = \langle i_T^2 \rangle / (eB)^2$ is the mean-square thermal noise current expressed as the mean-square average number of electrons flowing during a time slot due to thermal noise and x_t is the decision threshold also expressed in terms of the number of electrons per time slot. This normalization will turn out to be very convenient in that it will allow us to present the results in a form which is independent of bit rate. And where

$$p_i(n) = \frac{m_i^n}{n!} e^{-m_i}, \qquad i = 0, 1,$$

B = bit rate,

m_0 = mean number of primary electrons liberated when a "0" is transmitted,

m_1 = mean number of primary electrons liberated when a "1" is transmitted,

$\operatorname{erfc}(\cdot)$ is the complement of the error function.

These equations are derived under the assumption that the thermal noise is Gaussian and that the statistics of the primary electrons liberated in the photodetector due to signal photons, background-illumination photons, and dark current are independent Poisson processes. Then m_0 represents the sum of the means of the background illumination and dark current processes and $m_1 = m_0 + m_s$ where m_s is the mean number of photoelectrons liberated due to the signal. The majority of this section is devoted to the problem of determining the required value of m_s in order to achieve a specified error probability. The required average optical power is, of course, just

$$p_0 = \frac{1}{2} \frac{h\nu m_s}{\eta} B \tag{21}$$

where the factor $\frac{1}{2}$ comes from the assumption that 0's and 1's are equally probable.

The threshold value x_t is, of course, chosen to minimize the total error probability. To a very good approximation, this is achieved when

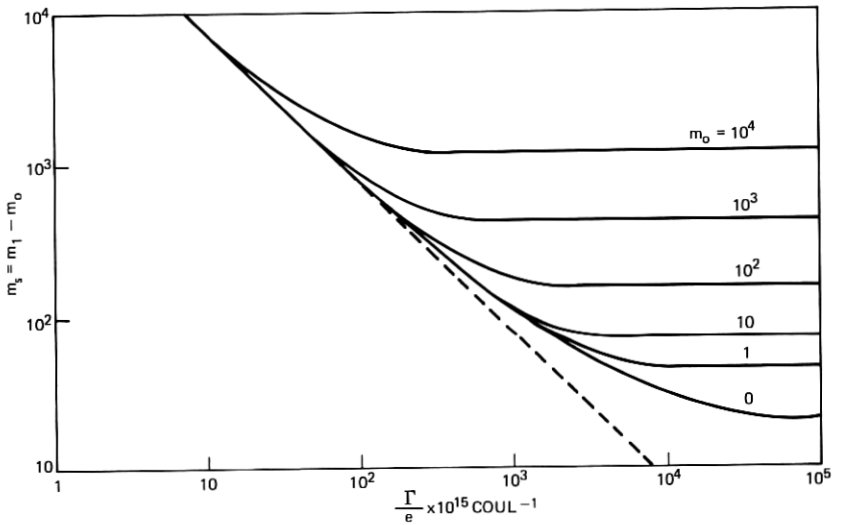

Fig. 14—Number of "signal photoelectrons" for a 10^{-9} error probability from the deterministic gain model vs $\frac{\Gamma}{e} \times 10^{15}$ Coul^{-1}. (Note $\frac{\Gamma}{e} \times 10^{15}$ Coul$^{-1} \approx G$ for reasonable system parameters.)

$P(0|1) = P(1|0)$. Thus, in practice, one can compute x_t from (19) setting $P(1|0)$ equal to the required error probability P_e; and then, knowing x_t, compute the required value of m_1 from (20). In the remainder of this section, it is assumed that the noise bandwidth $\mathcal{B}$ is equal to the bit rate B.

Figure 14 illustrates the results of the calculation described in the preceding paragraph. It is expedient to introduce the dimensionless parameter $\Gamma \equiv G/\langle x_T^2 \rangle^{\frac{1}{2}}$. Examination of this figure reveals that, for small values of Γ, the required signal power is inversely proportional to Γ. However, as Γ is increased, a limiting value, set by the dark current, is soon reached. This sets a maximum value of useful gain (for a given $\langle x_T^2 \rangle$), dependent only on m_0, beyond which no further significant improvement can be achieved. Note that this is true even here for the deterministic gain model with no excess noise factor of the sort to be discussed in Section 4.1.

4.1 *Gain-Dependent Excess Diode Noise Factor*

The gain-dependent excess diode noise factor* F_d played a very important role in the behavior of analog IM and PPM systems. There is

* Recall that F_d results from the random nature of the avalanche gain process.

no reason to believe that it has any less significant role in a digital system. The statistics of the avalanche gain process are very difficult to analyze except in two limiting cases, namely, when the ionization probabilities of holes and electrons in the avalanche region are equal, and when only one carrier contributes to the avalanche process.[1,9,11] Unfortunately, neither case applies to silicon and germanium photodetectors, the two most promising candidates.

Recently S. D. Personick[12] has obtained a rigorous upper bound to the error rate which is applicable to the general case (arbitrary ionization-probability ratio). The result of Personick's calculation is in the form of an integral equation, however, which must be computed numerically. In this section, we derive an approximate relationship between m_s and G which is in excellent agreement* with Personick's result.

This gain-dependent noise figure will have two effects on system behavior: (*i*) it will establish an optimum value of gain in the sense that m_s will have a minimum as a function of G, everything else held constant, and (*ii*) it will cause a larger number of signal photoelectrons to be required, for a given value of G, to satisfy a given error-probability requirement.

The expression we seek is derived as follows. We approximate the Poisson probability density function which describes the primary electron emission by a Gaussian probability density function with the same mean and variance. This approximation, which is discussed in detail in Ref. 13, turns out to be valid for most cases of interest. We also assume, without justification, that the statistics of the output of the photodetector are still Gaussian with variance equal to F_d times the variance of the primary electron distribution† where F_d is the excess detector noise factor used in Sections II and III.

With this assumption, eqs. (19) and (20) can be reduced to:

$$P(1|0) = \tfrac{1}{2} \operatorname{erfc} \left\{ \frac{x_t - m_o G}{[2(\langle x_T^2 \rangle + m_o F_d G^2)]^{\frac{1}{2}}} \right\} \tag{22}$$

$$P(1|0) = \tfrac{1}{2} \operatorname{erfc} \left\{ \frac{m_1 G - x_t}{[2(\langle x_T^2 \rangle + m_1 F_d G^2)]^{\frac{1}{2}}} \right\}. \tag{23}$$

We choose x_t in such a way that

$$P(1|0) = P(0|1) = P_e$$

* The results of this approximation are typically less than Personick's upper bound by about 1.0 ± 0.5 dB.

† It is not difficult to show rigorously that this is the correct value of the variance.

where P_e is the error probability. Defining a quantity Q by the relationship

$$P_e = \tfrac{1}{2}\,\mathrm{erfc}\left(\frac{Q}{\sqrt{2}}\right)$$

enables one to write

$$\frac{x_t - m_o G}{(\langle x_T^2\rangle + m_o F_d G^2)^{\frac{1}{2}}} = Q = \frac{m_1 G - x_t}{(\langle x_T^2\rangle + m_1 F_d G^2)^{\frac{1}{2}}}.$$

Eliminating x_t between these equations and setting $F_d = G^{\frac{1}{2}}$ gives

$$m_s = m_1 - m_o = 2Q\left[\frac{\langle x_T^2\rangle}{G^2} + m_o G^{\frac{1}{2}}\right]^{\frac{1}{2}} + Q^2 G^{\frac{1}{2}}. \tag{24}$$

It is clear from (24) that m_s has a minimum in G.

It was previously stated that $\langle x_T^2\rangle$ is independent of bit rate. This comes about as follows. From its definition and that of $\langle i_T^2\rangle$,

$$\langle x_T^2\rangle = \frac{4\mathrm{k}T}{e^2 B}\frac{F_t}{R},$$

but R is inversely proportional to bit rate. In fact, for a well-designed detector,* one can write

$$R = \frac{1}{4Bc}$$

where c is the capacitance of the photodetector. Then

$$\langle x_T^2\rangle = \frac{16\mathrm{k}TF_t c}{e^2}.$$

A value of $e\langle x_T^2\rangle^{\frac{1}{2}} = 4(\mathrm{k}TF_t c)^{\frac{1}{2}} = 10^{-15}$ Coul implies a value of $F_t c = 15$ pF which is typical of good avalanche photodetectors. This value is used in the example illustrated in Fig. 15. In this figure, we observe that for small values of avalanche gain, where the system performance is thermal-noise limited, the behavior is identical to that of the deterministic gain case. As the gain is increased, however, m_s reaches a broad minimum at $G = G_{\mathrm{opt}}$ and then increases slowly as the gain is further increased.

With proper exclusion of background illumination, typical values of dark and leakage currents are 10^{-10} and 10^{-9} A, respectively. The background count, m_o, is given by

$$m_o = \frac{I_b}{eB} + \frac{I_L}{eBG} = \frac{10^9}{1.6B} + \frac{10^{10}}{1.6BG}.$$

From Fig. 16, we observe that $G_{\mathrm{opt}} \approx 100$ over a wide range of condi-

* Note added in proof: Recent work by S. D. Personick shows that some advantage may be obtained by using larger R and post-detection equalization.

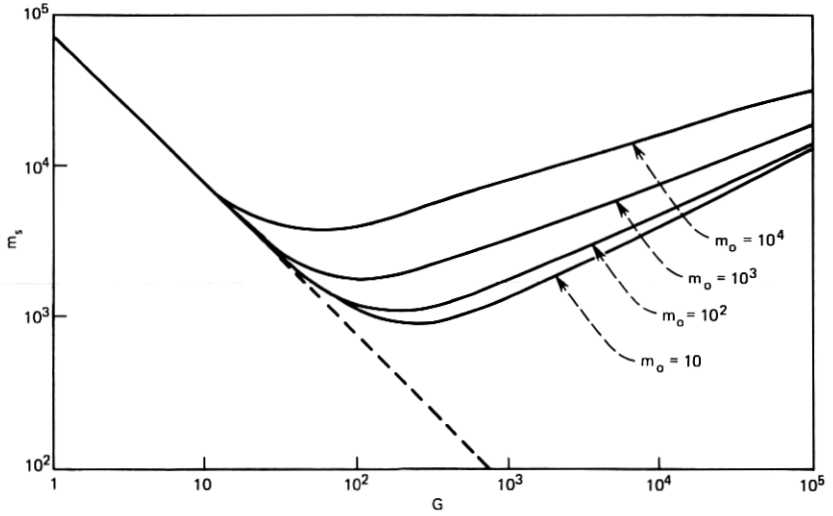

Fig. 15—Number of "signal photoelectrons" for a 10^{-9}-error probability from the gain-dependent noise model ($F_d = G^{\frac{1}{2}}$) vs avalanche gain. $e\langle x_T^2\rangle^{\frac{1}{2}} = 10^{-15}$ Coul.

tions and therefore the dark-current term usually dominates the leakage-current term. Typical values of m_o would therefore be from about 600 at $B = 1$ Mb/s down to about 0.6 at $B = 1$ Gb/s.

Figure 17 shows the value of m_s required for $P_e = 10^{-9}$ ($Q = 6.00$) for $e\langle x_T^2\rangle^{\frac{1}{2}} = 10^{-15}$ Coul [from (24)] for $G = 100$ and for $G = G_{\text{opt}}$. The value of G_{opt} is also shown in Fig. 17. Two important results are apparent from Fig. 17: m_o is not important until it exceeds about 100, and using $G = 100$ instead of $G = G_{\text{opt}}$ costs no more than about 1 dB. This last result is important because G_{opt} is so large over much of the region of interest that it would be difficult to obtain.

We now turn our attention to the use of nonadaptive delta modulation (ΔM) for transmitting analog signals. Delta modulation is a form of digital modulation which allows a trade-off between bandwidth and both terminal cost and signal power.

Noise in ΔM systems has been studied in detail by several authors.[14–17] We present here a sketch of how one might estimate the bit-rate requirements for a ΔM system. Suppose that the frequency band of the information source extends from 0 to b and that the step size of the coder is s. The maximum slope of a sine wave with amplitude A and frequency f is $2\pi A f$, while the maximum slope of the quantized signal with step size s and sampling rate B is sB. The limiting condition for slope overload is then

$$sB = 2\pi A f. \tag{25}$$

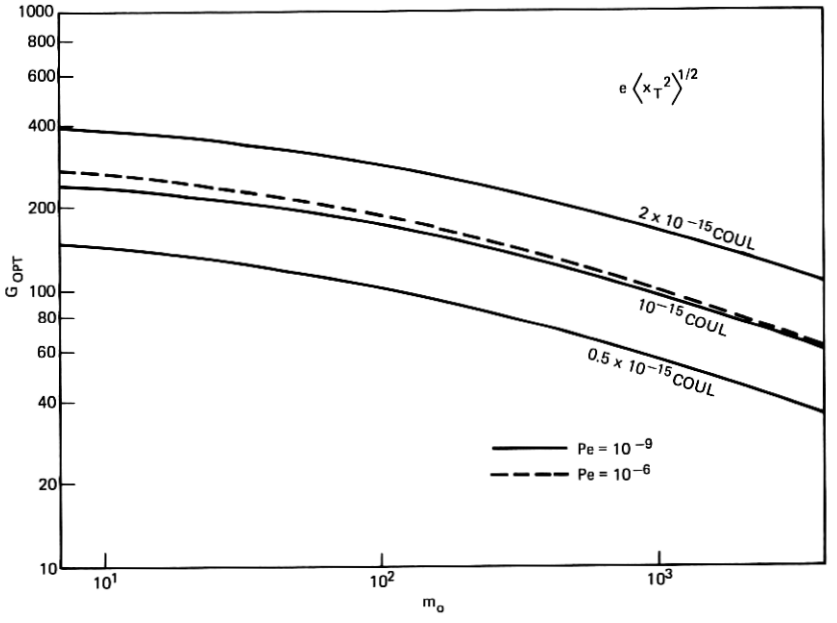

Fig. 16—Optimum gain ($F_d = G^{\frac{1}{2}}$) vs primary background count, m_0.

In order to compute the mean-square error in the quantized signal we assume, following Van de Weg,[14] that there is no correlation between samples and that the difference between the source signal and the quantized signal is uniformly distributed on the interval $(-s, s)$. We obtain:

$$\langle \delta s^2 \rangle = \frac{1}{2s}\int_{-s}^{s} x^2 dx = \tfrac{1}{3}s^2 = \text{mean quantizing noise power.} \quad (26)$$

The spectrum of the noise is quite complicated, but for our immediate purpose it is sufficient to assume that this noise is spread more or less uniformly over a bandwidth B so that the fraction b/B of the quantizing noise power falls into the information band b. (The remainder of the noise can then be eliminated by a low-pass filter of bandwidth b.) Assuming that quantizing noise is the only significant noise source, the signal-to-noise ratio, SNR, is then given for a sinusoidal signal of amplitude A by:

$$\text{SNR} = \frac{\frac{1}{2}A^2}{\frac{1}{3}s^2\frac{b}{B}} = \frac{3}{2}\frac{B}{b}\left(\frac{A}{s}\right)^2 = \frac{3}{8\pi^2}\frac{B^3}{bf^2} \quad (27)$$

where f is the highest frequency we are required to transmit with the

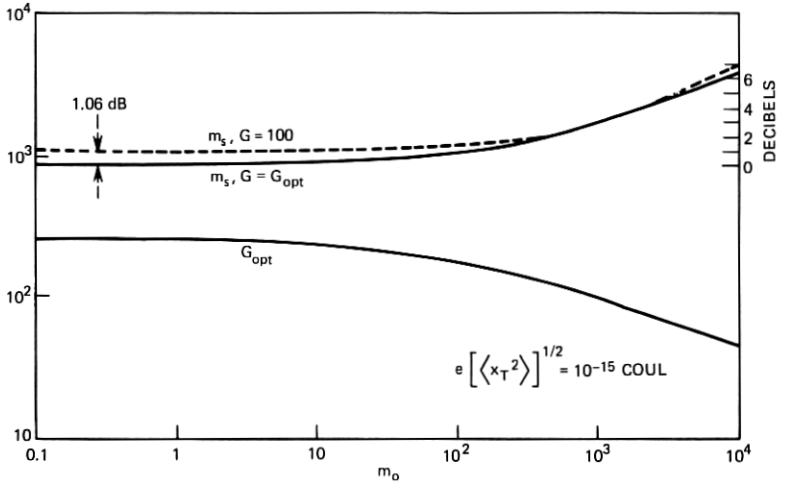

Fig. 17—Number of "signal photoelectrons," m_s, vs background count, m_o, for $P_e = 10^{-9}$.

specified value of SNR.* For a voice channel, we take (as conservative values) $b = 4 \times 10^3$ Hz and $f = 2 \times 10^3$ Hz. Then

$$\mathcal{B} = \left[\frac{8\pi^2}{3} b f^2 (\mathrm{SNR})\right]^{\frac{1}{3}} = 7500(\mathrm{SNR})^{\frac{1}{3}} \text{ Hz}. \tag{28}$$

For SNR = 70 dB, this gives $B = 1.6$ MHz. Van de Weg's calculation[14] takes into account the correlation between samples (which we have neglected) and the details of the quantizing noise spectrum. His result (for $B/b \geqq 4$) replaces the factor $8\pi^2/3 = 26.3$ in (28) by the factor 25.0 (and leaves it otherwise unaffected). Thus, his result is essentially identical with the one derived above. Experimental work by Laane and Murphy[17] indicates that a value of $B = 1.5$ MHz is adequate for transmitting a single voice channel by ΔM; we use this value in the following calculation.

For a 1.5-Mb/s rate, $I_D = 10^{-9}$ and $I_L = 10^{-10}$ gives

$$m_o = 417 + 4170/\mathrm{G} \approx 460.$$

From Fig. 17, the required value of m_s is 1450 (which is 2.2 dB poorer

* A ΔM system is limited by slope overload as indicated by (25). The limiting condition is set not by amplitude or by frequency alone, but by their product Af. We define the SNR for the ΔM system in terms of the ratio of full load power in a sinusoidal signal at some frequency f to the mean quantizing noise power. At higher frequencies, the available SNR degrades at a rate of 6 dB per octave. One could, of course, choose f to be the highest frequency in the information source bandwidth, b. For voice signals, however, this turns out to be an unreasonable constraint.[18]

than the case of no dark current and optimum gain, and 1.1 dB poorer than the case of no dark current and $G = 100$). The corresponding value of p_o is (for $\lambda = 0.85$ μm, $\eta = 0.5$) 0.51×10^{-9} W $= -63$ dBm. This is based on a tolerable error probability of 10^{-9}, which is probably much better than necessary. However, the error probability varies at the rate of about one order of magnitude for every 0.5 dB change in optical signal power. Therefore it makes little difference what value of error probability is chosen for this calculation.

It has been stated in the preceding sections that dark current and leakage current are generally negligible with good silicon photodiodes. The case of 1.5-Mb/s PCM is just on the border line of being dark-current limited with the numbers used in this example. The values of I_D and I_L used in this example must not, however, be regarded as ultimate performance. Indeed, photodetectors with $I_D \leqq 10^{-11}$ A have been built.[4]

V. CONCLUSIONS

Because light-emitting diodes and diode lasers can be directly modulated, analog intensity modulation is the simplest form of modulation to implement. Considerable improvement in noise immunity can be obtained, however, by judiciously exploiting the wide available bandwidth in optical systems.

Thus, pulse position modulation is particularly attractive because the square-law nature of the detector makes the "bunching" of the optical power beneficial and because the signal-power-dependent nature of the noise makes very large bandwidth-expansion factors feasible. Improvement of over 40 dB (relative to intensity modulation) is theoretically possible with pulse position modulation. Improvement of about 30 dB for a single high-quality 4-kHz voice channel appears to be realizable with existing light-emitting diodes. Use of delta modulation affords a theoretical improvement in noise immunity of about 25 dB relative to analog intensity modulation.

Optical carriers appear attractive for pulse code modulation even at low bit rates. At a bit rate of 6 Mb/s, only about -58 dBm of signal power is required for 10^{-9} error rate.

APPENDIX

Comparison of Practical Direct Detection Receivers With Homodyne Receivers and With Ideal Reception for Binary PCM Channels

It is instructive to compare the performance of a direct detection receiver of the sort described in Section IV with a homodyne receiver

operating on a similar signal. Consider a signal in which a "1" is represented by a pulse with peak power P_s and a "0" is represented by the absence of a pulse. Let P_{LO} be the local oscillator power. The detected photocurrent (with no avalanche gain) is then

$$i_s = \eta \frac{e}{h\nu} [2\sqrt{p_m P_{LO}} + P_{LO} + p_m] \tag{29}$$

at the peak of the pulse when a "1" is transmitted and is $\eta(e/h\nu)P_{LO}$ when a "0" is transmitted. In practical operation, $P_{LO} \gg p_m$ and the last term in (29) is negligible. The quantity $\eta(e/h\nu)P_{LO}$ is just a dc shift and can be neglected in the following calculation.

The mean-square noise current is given by

$$\langle i_n^2 \rangle = 2e \frac{e}{h\nu} \eta(P_{LO} + p_m)\mathcal{B} + 2eI_d \mathcal{B} + \frac{4kTF_t\mathcal{B}}{R}.$$

But once again $P_{LO} \gg p_m$ and in any reasonable receiver one also has $\eta(e/h\nu)P_{LO} \gg I_d$, so $\langle i_n^2 \rangle$ becomes

$$\langle i_n^2 \rangle = \left[2e\eta \frac{e}{h\nu} P_{LO} + \frac{4kTF_t}{R} \right] \mathcal{B}.$$

For large P_{LO}, this (Poisson) noise can be regarded as Gaussian. The optimum decision threshold will be near $(\frac{1}{2})i_s$. Therefore, the probability of error is well approximated by

$$P_e = \frac{1}{\sqrt{2\pi\langle i_s^2 \rangle}} \int_{\eta(e/h\nu)\sqrt{P_{LO}p_m}}^{\infty} \exp\left\{ -\frac{x^2}{2\langle i_n^2 \rangle} \right\} dx$$

$$= \tfrac{1}{2} \operatorname{erfc} \left\{ \frac{1}{\sqrt{2}} \sqrt{\frac{\eta p_m}{2h\nu\mathcal{B} + \dfrac{4kTF_t h\nu\mathcal{B}}{eR\eta \dfrac{e}{h\nu} P_{LO}}}} \right\}.$$

In order to fully exploit the advantages of homodyne detection, one must require

$$P_{LO} \gg \frac{2kTF_t}{e \dfrac{e}{h\nu} \eta R}.$$

When this condition obtains, one has

$$P_e = \tfrac{1}{2} \operatorname{erfc} \left\{ \frac{1}{\sqrt{2}} \sqrt{\frac{\eta p_m}{2h\nu\mathcal{B}}} \right\}.$$

Assuming a square pulse of duration $1/\mathcal{B}$ gives

$$m_s = \frac{\eta p_m}{h\nu \mathcal{B}}$$

and

$$P_e = \tfrac{1}{2} \operatorname{erfc} \left\{ \frac{1}{\sqrt{2}} \sqrt{\frac{m_s}{2}} \right\}. \tag{30}$$

We model an "ideal receiver" by a device which unerringly distinguishes between the case when no photoelectrons were liberated and the case when one or more were liberated. Since the photoelectrons are Poisson distributed, the probability that none were liberated when the expected number was m_s is just

$$e^{-m_s}$$

and for the ideal receiver this is twice the probability of error. For a $P_e = 10^{-9}$ this gives:

Ideal receiver:	$m_s = 20$
Homodyne receiver:	$m_s = 72.$

Thus the homodyne receiver is 5.5 dB poorer than the ideal. From Section IV we see that a practical direct detection receiver requires $m_s \approx 1000$ which places it 17 dB worse than the ideal receiver and about 12 dB worse than the homodyne receiver. Of course, 12 dB is not insignificant; but homodyne (or heterodyne) detection requires both a coherent source and precise phase-front matching between the signal and the local oscillator. With LED's this is impossible, with existing diode lasers it is at best extremely difficult. Even if adequate phase-front matching could be achieved, phase-lock for the homodyne receiver would be extremely difficult—if at all possible; heterodyne detection would reduce the advantage to 9 dB.

Direct detection without avalanche gain requires $m_s = 7.2 + 10^4$ (for $e\langle x_T^2 \rangle^{\frac{1}{2}} = 10^{-15}$ Coul) which is almost 36 dB worse than the ideal receiver. Thus, in the example of Section IV the avalanche gain ($G = 100$) gives almost 19 dB improvement.

These relative performance numbers are based on $P_e = 10^{-9}$ but over the range $10^{-10} < P_e < 10^{-4}$ the relative performance of the ideal detector, the homodyne detector, and the avalanche photodetector varies by less than 1 dB while the performance of direct detector without avalanche gain relative to the ideal detector varies by no more than 2 dB on this range.

REFERENCES

1. McIntyre, R. J., "Multiplication Noise in Uniform Avalanche Diodes," IEEE Trans. Elec. Dev., *ED-13*, No. 1 (January 1966), pp. 164–168.
2. Anderson, L. K., DiDomenico, M., Jr., and Fisher M. B., "High-Speed Photodetectors for Microwave Demodulation of Light," in *Advances in Microwaves*, ed. L. Young, vol. 5, New York: Academic Press, 1970.
3. Davenport, Jr., W. B., and Root, W. L., *An Introduction to the Theory of Random Signals and Noise*, New York: McGraw-Hill, 1958.
4. Melchior, H., Fischer, M. B., and Arams, F. R., "Photodectors for Optical Communication Systems," IEEE, *58*, No. 10 (October 1970), pp. 1466–1486.
5. Melchior, H., and Lynch, W. T., "Signal and Noise Response of High-Speed Germanium Avalanche Photodiodes," IEEE Trans. Elec. Dev., *ED-13*, No. 12 (December 1966), pp. 829–838.
6. Schwartz, M., Bennett, W. R., and Stein, S., *Communication Systems and Techniques*, New York: McGraw-Hill.
7. Rice S. O., "Mathematical Analysis of Random Noise," B.S.T.J., *23*, No. 3 (July 1944), pp. 282–332.
8. Burrus, C. A., and Dawson, R. W., "Small-Area High-Current Density GaAs Electroluminescent Diodes and a Method of Operation for Improved Degradation Characteristics," Appl. Phys. Lett., *17*, No. 3 (August 1970), pp. 97–99.
9. Hubbard, W. M., "Comparative Performance of Twin-Channel and Single-Channel Optical-Frequency Receivers," IEEE Trans. Commun. COM–20, No. 6, Dec. 1972, pp. 1079–1086.
10. Karp, S., and Gagliardi, R. M., "The Design of a Pulse-Position Modulated Optical Communication System," IEEE Trans. Commun. Tech., *Com-17*, No. 6 (December 1969), pp. 670–676.
11. Personick, S. D., "New Results on Avalanche Multiplication Statistics with Applications to Optical Detection," B.S.T.J., *50*, No. 1 (January 1971), pp. 167–190.
12. Personick, S. D., "Statistics of a General Class of Avalanche Detectors with Application to Optical Communication," B.S.T.J., *50*, No. 10 (December 1971), pp. 3075–3095.
13. Hubbard, W. M., "The Approximation of a Poisson Distribution by a Gaussian Distribution," Proc. IEEE, *58*, No. 9 (September 1970), pp. 1374–1375.
14. Van de Weg, H., "Quantizing Noise of a Single Integration Delta Modulation System with an N-Digit Code," Philips Res. Rep., *8*, 1953, pp. 367–385.
15. O'Neal, J. B., Jr., "Delta Modulation Quantizing Noise Analytical and Computer Simulation Results for Gaussian and Television Input Signals," B.S.T.J., *45*, No. 1 (January 1966), pp. 117–142.
16. Iwersen, J. E., "Calculated Quantizing Noise of Single-Integration Delta-Modulation Coders," B.S.T.J., *48*, No. 7 (September 1969), pp. 2359–2389.
17. Laane, R. R., and Murphy, B. T., "Delta Modulation Codec for Telephone Transmission and Switching Applications," B.S.T.J., *49*, No. 6 (July 1970), pp. 1013–1032.
18. deJager, F., "Delta Modulation, A Method of PCM Transmission Using 1-Unit Code," Philips Res. Rep., *7*, 1952, pp. 442–446.