Bounds on Communication with Polyphase Coding

By A. D. WYNER

(Manuscript received December 23, 1965)

The theoretical capabilities of a "polyphase" coding-modulation scheme with additive white Gaussian noise are studied. The channel capacity of this system is found and the error exponent estimated. Bounds are also found on $R_o(\rho_{\max})$, *the maximum (asymptotic) rate for which polyphase codes can be found with maximum correlation between code words* $\rho_{\max}$.

I. DEFINITIONS AND PRELIMINARIES

We shall consider the following ("polyphase") coding-modulation system (schematized in Fig. 1):

Every T seconds the message source emits one of M equally likely messages. The information rate is $R = 1/T \ln M$ nats per second. Corresponding to the ith message ($i = 1, 2, \cdots, M$) the coder emits an n-vector $\mathbf{x}_i = (x_{i1}, x_{i2}, \cdots, x_{in})$, where

$$-\pi \leqq x_{ik} \leqq \pi, \qquad k = 1, 2, \cdots, n, \tag{1}$$

and where the integer n will be specified later. The time interval $[0, T]$, during which this information must be transmitted, is divided into n equal subintervals of length T/n. During the kth of these subintervals, the modulated signal is

$$s_i(t) = \sqrt{2S} \cos (\omega_c t + x_{ik}), \qquad (k - 1) \frac{T}{n} \leqq t < \frac{kT}{n}, \quad k = 1,2, \cdots, n. \tag{2}$$

Thus, we have employed phase modulation with carrier frequency ω_c radians per second and average power S.

We assume that the noise is additive, white, and Gaussian with one-sided spectral density N_o. The receiver must examine the received signal $y(t)$, the sum of $s_i(t)$ and the noise, and determine which of the M messages was actually transmitted. It is well known that (since all

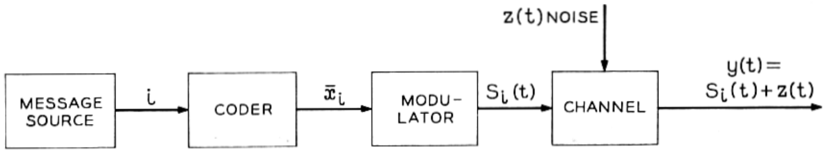

Fig. 1 — Polyphase coding-modulation system.

signals are equally likely to be transmitted, and have equal energy ST) the optimal decoder (which minimizes the average error probability) selects that signal $s_i(t)$ which maximizes ρ_i, the (normalized) correlation between $s_i(t)$ and $y(t)$:

$$\rho_i = \frac{1}{ST}\int_0^T s_i(t)y(t)\,dt. \tag{3}$$

Let us remark at this point that the correct operation of the decoder depends on its exact knowledge of the possible transmitted signals, so that in particular all delays and distortions to which the signals are subjected in transmission must be known exactly by the receiver. This is a so-called "coherent" receiver.

We let P_{ei} equal the probability that the decoder output is incorrect given that message i was transmitted, so that the average error probability is

$$P_e = \frac{1}{M}\sum_{i=1}^{M} P_{ei}\,. \tag{4}$$

Now, the same channel is to be used by a number of users simultaneously, each at a different carrier frequency. Let W cycles per second be the separation of carrier frequencies between adjacent users (W will be taken as the "bandwidth"). Then the carrier frequency for the αth user (α an integer) is $\omega_c = \alpha 2\pi W$ radians per second. Further, we shall set $n = WT$ (let us say that T is such that WT is an integer), where n is the number of subintervals defined previously. With ω_c and n so chosen and the signals constructed as in (2), it is easy to show that the signals of the αth and βth ($\alpha \neq \beta$) users are orthogonal on the interval [0, T]. Hence, the presence of the signal due to the βth user does not affect the correlator in the decoder of the αth user.

Let us say that the transmission rate R and the bandwidth W are held fixed, and let T, the duration of the signals (hence $n = WT$), become large. Every T seconds the message source will produce one of $M = e^{RT}$ equally likely messages to which the coder must assign an n-vector. The channel capacity C is the maximum rate for which we may make P_e vanishing small for T sufficiently large. Formally, for any $R < C$ and $\varepsilon > 0$, there is a T sufficiently large so that the transmitter may transmit one of $M = e^{RT}$ messages with $P_e < \varepsilon$. (This will necessitate a set of $M = e^{RT}$ n-vectors stored in the coder.) The channel capacity C of this coding-modulation scheme is found in Section III.

Let us consider again the decoding scheme. Making use of the fact that the ρ_i of (3) are normally distributed random variables, it is possible to write an expression for the error probability P_e* which depends only on the signal energy to noise ratio ST/N_o and the matrix of normalized inner products among signals

$$\rho_{ij} = \frac{1}{ST}\int_0^T s_i(t)s_j(t)dt, \qquad i,j = 1,2, \cdots, M. \tag{5}$$

From (2) we obtain

$$\rho_{ij} = \frac{1}{n}\sum_{k=1}^{n} \cos(x_{ik} - x_{jk}), \qquad i,j = 1,2, \cdots, n. \tag{6}$$

It is known† that the error probability P_e (as given in (4)) using the optimal decoder may be bounded by

$$P_e \leqq f(\max_{i \neq j} \rho_{ij}),$$

where $f(x)$ is an increasing function of x. Accordingly, a reasonable procedure for designing good coding systems would be to try to make $\rho_{\max} = \max_{i \neq j} \rho_{ij}$ as small as possible. Alternately we pose the problem as follows:

> With W, T, $\rho_{\max}$ held fixed, what is the largest rate for which we can design codes with parameters W, T, $\rho_{\max}$? (7)

Let us observe that from (6)

$$\begin{aligned} \rho_{ij} &= 1 - \frac{1}{n}\sum_{k=1}^{n} [1 - \cos(x_{ik} - x_{jk})] \\ &= 1 - \frac{1}{n}\sum_{k} 2\sin^2\frac{(x_{ik} - x_{jk})}{2} = 1 - \frac{d^2(\mathbf{x}_i, \mathbf{x}_j)}{2n} \end{aligned} \tag{8}$$

* Ref. 1, (2.11).
† Ref. 1, (4.7) and Ref. 2, p. 498.

where the "distance" $d(\mathbf{x}_i, \mathbf{x}_j)$ is defined by

$$d^2(\mathbf{x}_i, \mathbf{x}_j) = \sum_{k=1}^{n} \left[2 \sin \frac{(x_{ik} - x_{jk})}{2} \right]^2. \tag{9}$$

Thus, a code with maximum $\rho_{ij} = \rho_{\max}$, has minimum $d^2(\mathbf{x}_i, \mathbf{x}_j)/2n = (1 - \rho_{\max})$. In the light of the above, we shall reformulate the problem as follows:

Let $\mathcal{A}_n$ be the space of real n-vectors $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ (where $n = WT$) which satisfy

$$-\pi \leqq x_k \leqq +\pi, \qquad k = 1, 2, \cdots, n. \tag{10}$$

Let $\mathbf{x} = (x_1, \cdots, x_n)$ and $\mathbf{y} = (y_1, \cdots, y_n) \, \varepsilon \mathcal{A}_n$, and define the *distance* between $\mathbf{x}$ and $\mathbf{y}$ as

$$d(\mathbf{x},\mathbf{y}) = \left[\sum_{k=1}^{n} \left(2 \sin \frac{(x_k - y_k)}{2} \right)^2 \right]^{\frac{1}{2}}. \tag{11}$$

It will be shown in Section IV that $d(\mathbf{x},\mathbf{y})$ is, in fact, a metric. A *code* is a set of M members of $\mathcal{A}_n$, $\{\mathbf{x}_i = (x_{i1}, x_{i2}, \cdots, x_{in})\}_{i=1}^{M}$. The *transmission rate* is $\hat{R} = 1/n \ln M$ nats per *symbol*. The transmission rate in nats per *second* is $R = (1/T) \ln M = W\hat{R}$. We will define $M(n,d)$ as the maximum number of code vectors in an n-dimensional code with minimum distance between pairs of code words d. Then $\hat{R}(n,d) = (1/n) \ln M(n,d)$, and $R(n,d) = (1/T) \ln M(n,d)$ are the corresponding transmission rates. A problem equivalent to that of (7) is the determination of $\hat{R}(n,d)$. In Section IV we shall let n (and hence T) become large while the ratio $\beta = d^2/2n$ is held fixed (corresponding to a fixed $\rho_{\max}$) and estimate $\hat{R}(\beta) = \lim_{n \to \infty} \hat{R}(n, \sqrt{2n\beta})$ by upper and lower bounds. Since $\beta = 1 - \rho_{\max}$, $\hat{R}(1 - \rho_{\max})$ is the (asymptotic) maximum rate for polyphase coding with $\max_{i \neq j} \rho_{ij} = \rho_{\max}$.

II. SUMMARY AND DISCUSSION OF RESULTS

The channel capacity is shown in Section III to be

$$C = W \left[-\int_0^\infty \hat{f}(\rho) \ln \frac{\hat{f}(\rho)}{\rho} \, d\rho + \ln 2 \frac{A}{e} \right], \tag{12}$$

where

$$A = S/N_o W \tag{12a}$$

is the signal-to-noise ratio, and

$$\hat{f}(\rho) = 2\rho A e^{-A(1+\rho^2)} I_0(2\rho A), \tag{12b}$$

and $I_\nu(x)$ is the modified Bessel function of νth order. Another formula for C is (93). Approximate formulas for C for large and small values of the signal-to-noise ratio A are obtained in Appendix A. For large values of A,

$$C = \frac{W}{2} \ln \left(\frac{4\pi}{e} A \right) + \varepsilon_1(A), \tag{13}$$

where $\varepsilon_1(A) \to 0$ as $A \to \infty$. For values of A close to zero

$$C = W[A + 0(A^2)]. \tag{14}$$

The capacity C is plotted versus the signal-to-noise ratio A in Fig. 2. Estimates of the optimal achievable error probability are obtained in Appendix D.

The upper and lower bounds on $\hat{R}(\beta)$ are expressed in terms of the function $C_0(\xi)$ which is defined as follows. Let ξ be chosen

$$0 < \xi \leqq 1,$$

then define $\lambda(\xi)$ as the (unique) solution of

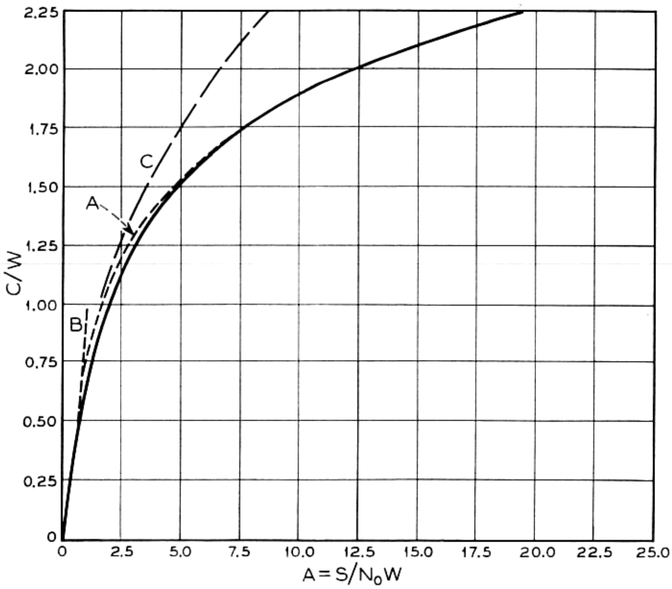

Fig. 2 — The channel capacity C vs the signal-to-noise ratio $A = S/N_oW$ — (12) (solid line). (Curves A and B are the approximations to the capacity C for large and small values of the signal-to-noise ratio A, respectively — (13) and (14). Curve C is $W \ln (1 + A)$, the capacity of a channel with bandwidth W and no restriction on the modulating scheme.

$$\xi = \left[1 - \frac{I_1(2\lambda(\xi))}{I_0(2\lambda(\xi))}\right]. \tag{15}$$

The existence (and uniqueness) of the solution to (15) is established in Appendix B. A graph of $\lambda(\xi)$ versus ξ is shown in Fig. 3. The function $C_0(\xi)$ is then defined as

$$C_0(\xi) = -\ln I_0(\lambda(\xi)) + (1 - \xi)\lambda(\xi). \tag{16}$$

A graph of $C_0(\xi)$ versus ξ is shown in Fig. 4. Our bounds on $\hat{R}(\beta)$, which are obtained in Section IV (and plotted in Fig. 5) are

$$C_0(\beta) \leqq \hat{R}(\beta) \leqq C_0(\gamma^2\beta), \tag{17}$$

where $C_0(\xi)$ is defined in (1) and

$$\gamma^2 = \frac{1}{\beta}(1 - \sqrt{1 - \beta}). \tag{18}$$

The lower bound is of the same type as the Gilbert bound for binary coding, and the upper bound makes use of the Blichfeldt density method.[3] Let us remark that the upper and lower bounds of (14) agree when $\beta = 1$, yielding $\hat{R}(\beta) = 0$ for $\beta \geqq 1$. When β is small it is shown in Appendix E that

$$\tfrac{1}{2}\ln\frac{1}{\pi e\beta} + \varepsilon_2(\beta) = \hat{R}(\beta) = \tfrac{1}{2}\ln\frac{2}{\pi e\beta} + \varepsilon_2(\beta), \tag{19}$$

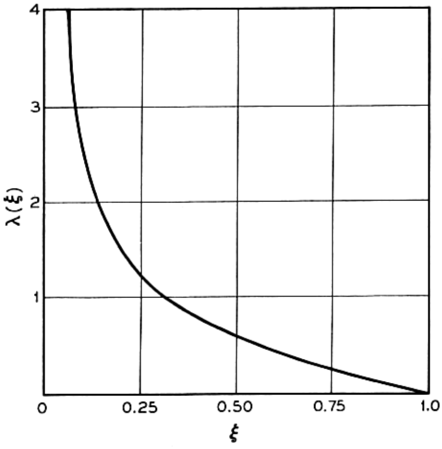

Fig. 3 — The function $\lambda(\xi)$ vs ξ — (15).

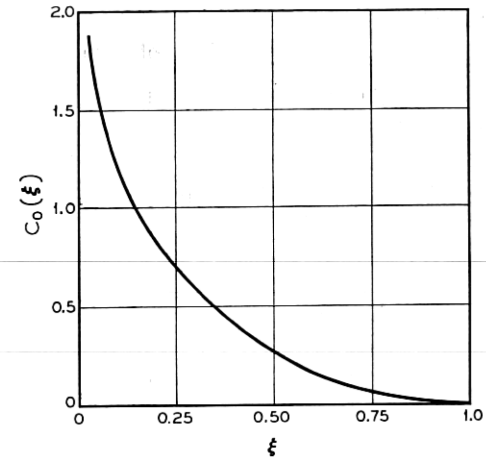

Fig. 4 — The function $C_0(\xi)$ vs ξ — (16).

where ε_1, $\varepsilon_2 \to 0$ as $\beta \to 0$. Thus, for sufficiently small β, $\hat{R}(\beta)$ is within $\frac{1}{2}$ ln 2 of $\frac{1}{2}$ ln $(2/\pi e\beta)$.

In terms of the modulation scheme discussed in Section I it is more revealing to rewrite inequalities (17) in terms of $\rho_{\max}$ the maximum correlation between pairs of signals. Let $R_0(\rho_{\max}, W, T) = R_0(\rho_{\max}, T)$ be the maximum rate (in nats per second) attainable for the polyphase

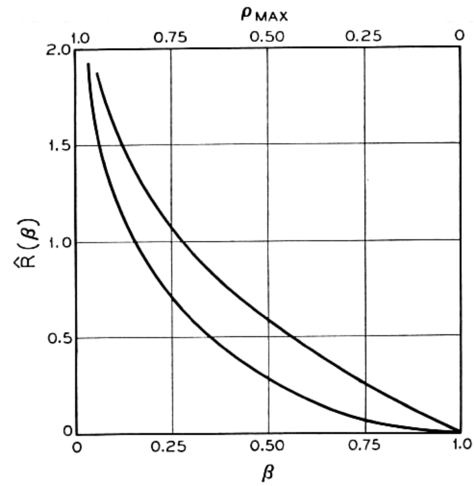

Fig. 5 — The upper and lower bounds $\hat{R}$ (β) vs β and $\rho_{\max} - 1 = \beta$ — (17). $\hat{R}$ (β) lies in the shaded region.

modulation scheme of Section I with parameters $\rho_{\max}$, W, and T. Let $R_o(\rho_{\max}) = \lim_{T\to\infty} R_o(\rho_{\max}, T)$. In the light of comment following (9), $R_o(\rho_{\max}) = R[(1 - \rho_{\max})]$. The upper and lower bounds on $R_o(\rho_{\max})$ are plotted versus $\rho_{\max}$ in Fig. 5.

Appendix F contains a comparison of the capabilities of this polyphase system and another important modulation system.

III. CHANNEL CAPACITY

The signals $s_i(t)$, $i = 1, 2, \cdots, M$, are of the form

$$s_i(t) = \sqrt{2S} \cos(\alpha 2\pi Wt + x_{ik}), \qquad (k-1)\frac{T}{n} \leqq t < \frac{kT}{n}, \quad k = 1,2, \cdots, n, \tag{20a}$$

where $n = WT$ and

$$-\pi \leqq x_{ik} \leqq \pi, \qquad k = 1, 2, \cdots, n. \tag{20b}$$

Alternately, we may write

$$s_i(t) = x_{ik}^{(1)} \cos \alpha 2\pi Wt + x_{ik}^{(2)} \sin \alpha 2\pi Wt, \qquad (k-1)\frac{T}{n} \leqq t < \frac{kT}{n}, \quad k = 1,2, \cdots, n, \tag{21a}$$

where

$$x_{ik}^{(1)} = \sqrt{2S} \cos x_{ik}, \qquad x_{ik}^{(2)} = \sqrt{2S} \sin x_{ik}. \tag{21b}$$

The noise function $z(t)$ is a sample from a white Gaussian noise process with one sided spectral density N_o (so that the covariance is $R(\tau) = (N_o/2)\delta(\tau)$). The received signal is $y(t) = s_i(t) + z(t)$, where $s_i(t)$ is one of the M signals. The optimal decoder computes

$$\rho_i = \frac{1}{ST}\int_0^T s_i(t)y(t)\,dt,$$

and decodes $y(t)$ as that $s_i(t)$ with largest ρ_i. If $y(t)$ is the received signal, let $y^*(t)$ be

$$y^*(t) = y_k^{(1)} \cos \alpha 2\pi Wt + y_k^{(2)} \sin \alpha 2\pi Wt, \qquad (k-1)\frac{T}{n} \leqq t < \frac{kT}{n}, \quad k = 1,2, \cdots, n, \tag{22a}$$

where

$$y_k^{(1)} = 2W \int_{(k-1)(T/n)}^{kT/n} y(t) \cos \alpha 2\pi W t \, dt, \tag{22b}$$

and

$$y_k^{(2)} = 2W \int_{(k-1)(T/n)}^{kT/n} y(t) \sin \alpha 2\pi W t \, dt. \tag{22c}$$

We may think of $y^*(t)$ as the projection of $y(t)$ onto the space of allowable signals. It follows by direct computation that

$$\frac{1}{ST} \int_0^T s_i(t) y^*(t) dt,$$

the correlation of $y^*(t)$ and the ith signal $s_i(t)$, equals ρ_i. Thus, without loss of generality, we may consider the received signal to be $y^*(t)$. From (21) and (22), it suffices to consider the noise to be

$$z^*(t) = y^*(t) - s_i(t) = z_k^{(1)} \cos \alpha 2\pi W t + z_k^{(2)} \sin \alpha 2\pi W t$$
$$(k-1)\frac{T}{n} \leqq t < \frac{kT}{n}, \qquad k = 1,2, \cdots, n, \tag{23a}$$

where

$$z_k^{(1)} = y_k^{(1)} - x_{ik}^{(1)} \quad \text{and} \quad z_k^{(2)} = y_k^{(2)} - x_{ik}^{(2)},$$
$$k = 1, 2, \cdots, n. \tag{23b}$$

From (23b), (22b), (21b), and (20a) we may write

$$\begin{aligned} z_k^{(1)} &= 2W \int_{(k-1)(T/n)}^{kT/n} (y(t) - s_i(t)) \cos \alpha 2\pi W t \, dt \\ &= 2W \int_{(k-1)(T/n)}^{kT/n} z(t) \cos \alpha 2\pi W t \, dt, \qquad k = 1,2, \cdots, n, \end{aligned} \tag{24}$$

so that $z_k^{(1)}$ is a normally distributed random variable with mean zero and variance

$$\begin{aligned} E(z_k^{(1)2}) = 4W^2 \int_{(k-1)(T/n)}^{kT/n} \\ \cdot \int_{(k-1)(T/n)}^{kT/n} \cos \alpha 2\pi W t \, (\cos \alpha 2\pi W \tau) \overline{z(t) z(\tau)} \, dt \, d\tau, \end{aligned} \tag{25}$$

where the over-bar denotes expectation. Since $\overline{z(t)z(\tau)} = R(t - \tau) = (N_o/2)\delta(t - \tau)$, the variance of $z_k^{(1)}$ is $N_o W$. Similarly for $z_k^{(2)}$. Further, $E(z_k^{(1)} z_k^{(2)}) = 0$, and

$$E(z_{k_1}^{(i)}, z_{k_2}^{(j)}) = 0 \qquad (i, j = 1, 2) \quad \text{if} \quad k_1 \neq k_2 .$$

Thus, these random variables are independent.

We conclude from the above that our channel is equivalent to the following *time-discrete memoryless* channel. Every $T/n = 1/W$ seconds, the channel input is a real number $X \varepsilon [-\pi, \pi]$. The output is a pair of numbers Y_1 and Y_2 given by

$$Y_1 = X_1 + Z_1 , \qquad Y_2 = X_2 + Z_2 \tag{26a}$$

where

$$X_1 = \sqrt{2S} \cos X, \qquad X_2 = \sqrt{2S} \sin X, \tag{26b}$$

and Z_1, Z_2 are independent normally distributed random variables with mean zero and variance $N = N_0 W$. Consequently, known results for determining capacity may be used.

If an input probability distribution is specified, the mutual information of the input and the output is

$$I(Y_1 , Y_2 ; X) = H(Y_1 , Y_2) - H(Y_1 , Y_2 \mid X), \tag{27}$$

where $H(Y_1 , Y_2)$ is the joint uncertainty of Y_1 and Y_2 and

$$H(Y_1 , Y_2 \mid X)$$

is the conditional uncertainty of Y_1, Y_2 given X. The channel capacity C, in nats per *second* is

$$C = W[\max I(Y_1 , Y_2 ; X)], \tag{28}$$

where the maximization is performed over all possible input distributions. We proceed to find C.

Say $X = x$, and let $x_1 = \sqrt{2S} \cos x$, $x_2 = \sqrt{2S} \sin x$, then

$$H(Y_1 , Y_2 \mid X = x)$$
$$= \int_{-\infty}^{+\infty} \cdot \int_{-\infty}^{+\infty} dy_1 dy_2 g(y_1 - x_1 , y_2 - x_2) \ln g(y_1 - x_1 , y_2 - x_2), \tag{29a}$$

where

$$g(z_1 , z_2) = \frac{1}{2\pi N} \exp [-(z_1^2 + z_2^2)/2N] \tag{29b}$$

is the joint probability density of Z_1, Z_2. After changing the variables of integration and integrating (29a), we obtain,

$$H(Y_1 , Y_2 \mid X = x) = \ln 2\pi e N, \tag{30}$$

independent of x. Thus,

$$H(Y_1Y_2 \mid X) = \ln 2\pi eN, \tag{31}$$

independent of the input distribution.

Thus, to find C, we must maximize $H(Y_1, Y_2)$. Say $p_0(x)$ is the probability density of the input X, and $p_{12}(y_1, y_2)$ the resulting joint probability density of the output pair (Y_1, Y_2). If we characterize the output pair by polar coordinates $(\mathfrak{R}, \Phi)$, then the corresponding density for $\mathfrak{R}, \Phi$ is

$$f_{12}(r,\varphi) = rp_{12}(r \cos \varphi, r \sin \varphi), \qquad r \geqq 0, \qquad -\pi \leqq \varphi \leqq \pi, \tag{32}$$

where r is the Jacobian of the transformation. Hence,

$$\begin{aligned} H(Y_1Y_2) &= -\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p_{12}(y_1, y_2) \ln p_{12}(y_1, y_2)dy_1dy_2 \\ &= -\int_{r=0}^{\infty}\int_{-\pi}^{\pi} p_{12}(r \cos \varphi, r \sin \varphi) \\ &\qquad \cdot \ln [p_{12}(r \cos \varphi, r \sin \varphi)]r\, dr\, d\varphi \\ &= -\int_{0}^{\infty}\int_{-\pi}^{\pi} f_{12}(r,\varphi) \ln \frac{f_{12}(r,\varphi)}{r}\, dr\, d\varphi \\ &= -\int_{0}^{\infty}\int_{-\pi}^{\pi} f_{12}(r,\varphi) \ln f_{12}(r,\varphi)dr\, d\varphi \\ &\qquad + \int_{0}^{\infty}\int_{-\pi}^{\pi} f_{12}(r,\varphi) \ln r\, dr\, d\varphi \\ &= H(\mathfrak{R},\Phi) + \int_{0}^{\infty} f_1(r) \ln r\, dr, \end{aligned} \tag{33}$$

where $H(\mathfrak{R}, \Phi)$ is the joint uncertainty of $\mathfrak{R}, \Phi$, and $f_1(r)$ is the marginal density of $\mathfrak{R}$. Now

$$H(\mathfrak{R}, \Phi) \leqq H(\mathfrak{R}) + H(\Phi), \tag{34}$$

(where $H(\mathfrak{R})$, $H(\Phi)$ are the uncertainties of $\mathfrak{R}, \Phi$, respectively) with equality if and only if $\mathfrak{R}, \Phi$ are independent, and

$$H(\Phi) \leqq \ln 2\pi, \tag{35}$$

with equality if and only if Φ is uniformly distributed on the interval $[-\pi,\pi]$. Hence, from (33), (34), and (35),

$$H(Y_1, Y_2) \leqq H(\mathfrak{R}) + \ln 2\pi + \int_{0}^{\infty} f_1(r) \ln r\, dr. \tag{36}$$

We shall now find $f_1(r)$, the density of $\mathfrak{R}$, and show that it is independent of the input density $p_0(x)$. To begin with, let us say that $X = x$. Then the joint density of (Y_1, Y_2), given that $X = x$ is

$$\begin{aligned} p_{12}(y_1, y_2 \mid X = x) \\ = \frac{1}{2\pi N} \exp \{-[(y_1 - \sqrt{2S} \cos x)^2 + (y_2 - \sqrt{2S} \sin x)^2]/2N\}. \end{aligned} \tag{37}$$

The joint density of Y_1, Y_2 or the corresponding joint density of R, Φ is obtained from (37) by averaging over x:

$$\begin{aligned} f_{12}(r,\varphi) &= rp_{12}(r \cos \varphi, r \sin \varphi) \\ &= r \int_{-\pi}^{\pi} p_0(x) p_{12}(r \cos \varphi, r \sin \varphi \mid X = x) dx \\ &= r \int_{-\pi}^{\pi} p_0(x) dx \frac{1}{2\pi N} \\ &\quad \cdot \exp \left\{ -\frac{1}{2N} [r \cos \varphi - \sqrt{2S} \cos x)^2 \right. \\ &\quad \left. + (r \sin \varphi - \sqrt{2S} \sin x)^2] \right\} \\ &= \frac{1}{2\pi N} re^{-(r^2+2S)/2N} \int_{-\pi}^{\pi} p_0(x) \\ &\quad \cdot \exp \left(\frac{r\sqrt{2S}}{N} \cos (x - \varphi) \right) dx. \end{aligned} \tag{38}$$

Now, the marginal density for R is obtained by integrating φ out of (38)

$$\begin{aligned} f_1(r) &= \int_{-\pi}^{\pi} f_{12}(r,\varphi) d\varphi \\ &= \frac{re^{-(r^2+2S)/2N}}{2\pi N} \int_{-\pi}^{\pi} d\varphi \int_{-\pi}^{\pi} dx\, p_0(x) \exp \left(\frac{r\sqrt{2S} \cos (x - \varphi)}{N} \right). \end{aligned} \tag{39}$$

Interchanging the order of integration, we get

$$\begin{aligned} f_1(r) &= \frac{re^{-(r^2+2S)/2N}}{2\pi N} \int_{-\pi}^{\pi} p_0(x) dx \int_{-\pi}^{\pi} d\varphi \exp \frac{r\sqrt{2S}}{N} \cos (\varphi - x) \\ &= \frac{re^{-(r^2+2S)/2N}}{N} I_0 \left(\frac{\sqrt{2S} r}{N} \right), \end{aligned} \tag{40}$$

independent of $p_0(x)$.* We conclude from (40) and (36) that

$$\max_{p_0(x)} H(Y_1, Y_2) \leqq -\int_0^\infty f_1(r) \ln \frac{f_1(r)}{r} dr + \ln 2\pi, \tag{41}$$

where $f_1(r)$ is given by (40).

Let us now say that the input distribution is $p_0(x) = 1/2\pi$. Then from (38)

$$f_{12}(r,\varphi) = f_1(r) \frac{1}{2\pi}, \tag{42}$$

so that $\mathcal{R}$, Φ are independent with the marginal density of Φ, $f_2(\varphi) = 1/2\pi$. Thus, in this case, the equalities in (34) and (35) and hence in (36) hold yielding

$$H(Y_1, Y_2) = -\int_0^\infty f_1(r) \ln \frac{f_1(r)}{r} dr + \ln (2\pi), \tag{43}$$

so that (41) is satisfied with equality. From (28), (30), (41), and (43), the channel capacity C is given by

$$\frac{C}{W} = -\int_0^\infty f_1(r) \ln \frac{f_1(r)}{r} dr - \ln eN. \tag{44}$$

If we set $\rho = r/\sqrt{2S}$ and $A = S/N = S/N_oW$, the "signal-to-noise ratio", we obtain

$$\frac{C}{W} = -\int_0^\infty \hat{f}(\rho) \ln \frac{\hat{f}(\rho)}{\rho} d\rho + \ln \frac{2A}{e}, \tag{45a}$$

where

$$\hat{f}(\rho) = 2A\rho e^{-A(\rho^2+1)} I_0(2A\rho). \tag{45b}$$

IV. BOUNDS ON $\hat{R}(\beta)$

4.1 *Upper Bound on* $\hat{R}(\beta)$

We need the following two lemmas:

---

* We shall make frequent use of the formula

$$I_0(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{x \cos \theta} d\theta,$$

which can be found in Ref. 4, p. 79.

Lemma 1: Let $g_1, g_2, \cdots, g_p$ be real numbers. Then

$$\sum_{k=1}^{p} g_k^2 \geqq \frac{1}{p} (\Sigma g_k)^2. \tag{46}$$

Proof: From the Schwarz inequality

$$\left(\sum_{k=1}^{p} 1 \cdot g_k\right)^2 \leqq \left(\sum_{k=1}^{p} 1^2\right)\left(\sum_{k=1}^{p} g_k^2\right) = p \sum_{k=1}^{p} g_k^2. \tag{47}$$

Lemma 2: Let $\{\mathbf{x}_i\}_{i=1}^m$ be a set of m n-vectors from $\mathfrak{A}_n$ with minimum distance d between pairs of vectors. The distance is given by (11). Let $\mathbf{y}$ be an arbitrary vector in $\mathfrak{A}_n$, and denote by d_i the distance $d(\mathbf{x}_i, \mathbf{y})$. Then

$$\left(\sum_{i=1}^{m} \frac{d_i^2}{n}\right) - 4m \left(\sum_{i=1}^{m} \frac{d_i^2}{n}\right) + 2(m)(m-1) \frac{d^2}{n} \leqq 0. \tag{48}$$

Proof: Let us define a mapping of $\mathfrak{A}_n$ into E_{2n}, Euclidean $2n$-space, as follows. If $\mathbf{x} = (x_1, x_2, \cdots, x_n) \varepsilon \mathfrak{A}_n$, then the corresponding $2n$-vector is $\mathbf{x}' = (u_1, v_1, u_2, v_2, \cdots, u_n, v_n)$ where

$$u_k = \cos x_k, \qquad v_k = \sin x_k, \qquad k = 1, 2, \cdots, n. \tag{49}$$

Then letting $\mathbf{x}_1, \mathbf{x}_2 \varepsilon \mathfrak{A}_n$, and letting $\mathbf{x}_1' = (u_{11}, v_{11}, u_{12}, v_{12}, \cdots, u_{1n}, v_{1n})$ and $\mathbf{x}_2' = (u_{21}, v_{21}, u_{22}, v_{22}, \cdots, u_{2n}, v_{2n})$ be the corresponding members of E_{2n}, the distance between $\mathbf{x}_1$ and $\mathbf{x}_2$ is

$$\begin{aligned} d^2(\mathbf{x}_1, \mathbf{x}_2) &= \sum_{k=1}^{n} \left[2 \sin \frac{(x_{1k} - x_{2k})}{2}\right]^2 \\ &= \sum_{k=1}^{n} \{(u_{1k} - u_{2k})^2 + (v_{1k} - v_{2k})^2\}. \end{aligned} \tag{50}$$

To see this we need only observe that if the $x_{1k}, x_{2k}, k = 1, 2, \cdots, n$ are considered as arc lengths on a unit circle with center at the origin, then (u_{1k}, v_{1k}) and (u_{2k}, v_{2k}) are the Cartesian coordinates of x_{1k}, and x_{2k}, respectively, (see Fig. 6). The quantity $2 \sin [(x_{1k} - x_{2k})/2]$ is then the Euclidean distance between (u_{1k}, v_{1k}) and (u_{2k}, v_{2k}). Hence, $d(\mathbf{x}_1, \mathbf{x}_2)$ is the Euclidean distance between $\mathbf{x}_1'$ and $\mathbf{x}_2'$. This also provides a justification for calling $d(\mathbf{x},\mathbf{y})$ a metric. We are now in a position to prove the lemma.

Without loss of generality we may take $\mathbf{y} = (0, 0, \cdots, 0)$ so that $\mathbf{y}' = (1, 0, 1, 0, \cdots, 1, 0)$. Let $\mathbf{x}_i' = (u_{i1}, v_{i1}, u_{i2}, v_{i2}, \cdots, u_{in}, v_{in})$ then

$$d_i^2 = d^2(\mathbf{x}_i, \mathbf{y}) = \sum_{k=1}^{n} \{(1 - u_{ik})^2 + v_{ik}^2\}.$$

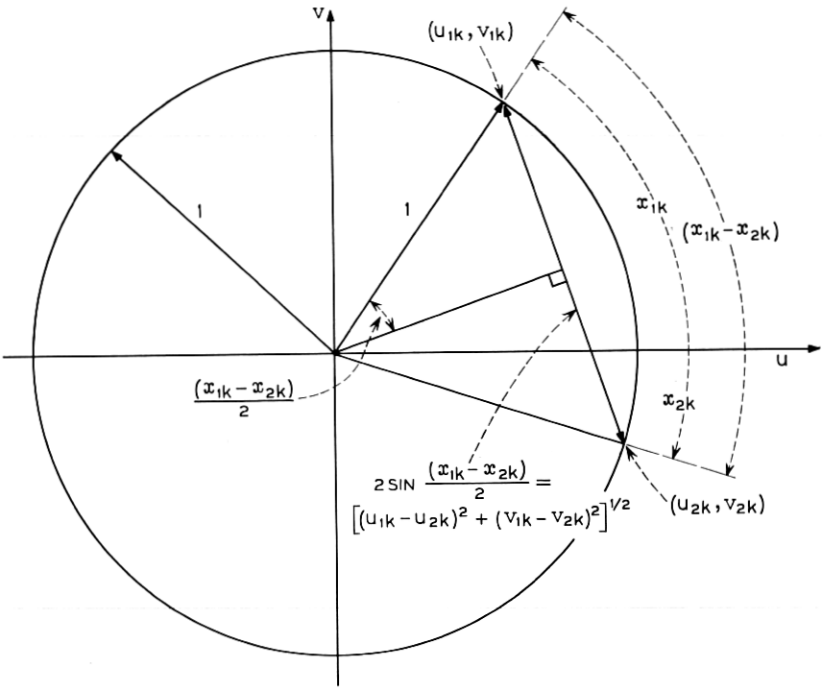

Fig. 6 — Proof of Lemma 2.

Since $d(\mathbf{x}_i, \mathbf{x}_j) \geqq d$,

$$\begin{aligned}\binom{m}{2} d^2 \leqq \sum_{i<j} d^2(\mathbf{x}_i, \mathbf{x}_j) &= \sum_{i<j} \sum_{k=1}^{n} \{(u_{ik} - u_{jk})^2 + (v_{ik} - v_{jk})^2\} \\ &= \sum_k \left\{ m \sum_{i=1}^m u_{ik}^2 - \left(\sum_{i=1}^m u_{ik} \right)^2 \right. \\ &\qquad \left. + m \sum_{i=1}^m v_{ik}^2 - \left(\sum_{i=1}^m v_{ik} \right)^2 \right\} \\ &= \sum_k \left\{ m \left(\sum_i (1 - u_{ik})^2 \right) \right. \\ &\qquad - \left(m - \sum_i u_{ik} \right)^2 + m \sum_i v_{ik}^2 \\ &\qquad \left. - \left(\sum_i v_{ik} \right)^2 \right\} \\ &= m \sum_i \sum_k \{(1 - u_{ik})^2 + v_{ik}^2\}\end{aligned} \tag{51}$$

$$- \sum_k \left(\sum_i (1 - u_{ik}) \right)^2$$

$$- \sum_k \left(\sum_i v_{ik} \right)^2.$$

$$\leqq m \sum_i d_i^2 - \sum_{k=1}^{n} \left(\sum_i (1 - u_{ik}) \right)^2.$$

From Lemma 1, (51) becomes

$$\binom{m}{2} d^2 \leqq m \sum_i d_i^2 - \frac{1}{n} \left(\sum_{k=1}^{n} \sum_{i=1}^{m} (1 - u_{ik}) \right)^2. \tag{52}$$

Now, since $u_{ik}^2 + v_{ik}^2 = 1$, we have

$$\begin{aligned} d_i^2 &= \sum_k \{(1 - u_{ik})^2 + v_{ik}^2\} = \sum_k \{1 - 2u_{ik} + u_{ik}^2 + v_{ik}^2\} \\ &= 2 \sum_k (1 - u_{ik}). \end{aligned} \tag{53}$$

Substituting (53) into (52) yields

$$\binom{m}{2} d^2 \leqq m \sum_i d_i^2 - \frac{1}{4n} \left(\sum_{i=1}^{n} d_i^2 \right)^2. \tag{54}$$

The lemma follows on multiplying both sides of (54) by $4/n$.

Derivation of the Bound:

If $\mathbf{z} \ \varepsilon \ \mathcal{A}_n$ let us define the "sphere" $S(\mathbf{z},\rho)$ as

$$S(\mathbf{z},\rho) = \{\mathbf{x} \ \varepsilon \ \mathcal{A}_n: \ d(\mathbf{x},\mathbf{z}) < \rho\}. \tag{55}$$

Since the distance d defined on $\mathcal{A}_n$ is a metric, it follows that if a code $\{\mathbf{x}_i\}_{i=1}^{M}$ has minimum distance (as defined by d), then the spheres $S(\mathbf{x}_i, d/2)$ are disjoint.

Consider the maximum size n-dimensional code with minimum distance d and $M(n,d)$ code words $\{\mathbf{x}_i\}_{i=1}^{M}$. Consider the spheres $S(\mathbf{x}_i, \gamma d)$ about each code word, where

$$\gamma^2 = \frac{1}{\beta} (1 - \sqrt{1 - \beta}) \tag{56a}$$

$$\beta = d^2/2n. \tag{56b}$$

Note that since $\gamma > \frac{1}{2} (0 \leqq \beta \leqq 1)$,* these spheres are not necessarily

---

* This follows immediately when we write $\gamma^2 = 1/1 + (1 - \beta)^{1/2}$, so that γ increases from $1/\sqrt{2}$ to 1 as β increases from 0 to 1.

disjoint. To each point in the sphere at distance r from the center assign a density $\sigma(r) = \gamma^2 d^2 - r^2$. Then the "mass" of each sphere is

$$\mu = \int_{r<\gamma d} (\gamma^2 d^2 - r^2) dV, \tag{57}$$

where the integration in (57) is performed with respect to the Euclidean measure, assigned to $\mathcal{A}_n$ in the obvious way.

In general, a vector $\mathbf{y} \varepsilon \mathcal{A}_n$ will belong to the spheres about m code words say $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_m$. We assign to $\mathbf{y}$, a density equal to the sum of the densities contributed by each sphere, i.e.,

$$\sigma_{\mathbf{y}} = \sum_{i=1}^{m} \sigma(d_i) = m\gamma^2 d^2 - \sum_{i=1}^{m} d_i^2, \tag{58}$$

where $d_i = d(\mathbf{y},\mathbf{x}_i)$. If $\mathbf{y}$ belongs to no sphere $\sigma_{\mathbf{y}} = 0$. Thus, we have

$$\text{mass of } \mathcal{A}_n = \int_{\mathbf{y}\epsilon\mathcal{A}_n} \sigma_{\mathbf{y}} dV = M(n,d) \cdot \mu. \tag{59}$$

We will bound $M(n,d)$ by finding an upper bound on the mass of $\mathcal{A}_n$.

Letting $s = s_{\mathbf{y}} = \sigma_{\mathbf{y}}/n$, (58) becomes

$$\frac{\sum d_i^2}{n} = \frac{m\gamma^2 d^2}{n} - \frac{\sigma_{\bar{y}}}{n} = 2m\gamma^2\beta - s, \tag{60}$$

where $\beta = d^2/2n$. Substituting (60) into (48) we get

$$(2m\gamma^2\beta - s)^2 - 4m(2m\gamma^2\beta - s) + 4(m)(m-1)\beta \leqq 0. \tag{61}$$

Rewriting (61)

$$0 \leqq s^2 \leqq m\{4\beta - 2m\beta(2\gamma^4\beta - 4\gamma^2 + 2) - 4s(1 - \gamma^2\beta)\}. \tag{62}$$

With γ chosen by (56), $2\gamma^4\beta - 4\gamma^2 + 2 = 0$ and $1 - \gamma^2\beta > 0$, so that (62) can only be satisfied if

$$s = \frac{\sigma}{n} \leqq \beta/(1 - \gamma^2\beta) \overset{\Delta}{=} K(\beta). \tag{63}$$

Hence, from (63) and (59) we have

$$M(n,d) = \frac{1}{\mu} \int_{\mathcal{A}_n} \sigma_{\bar{y}} dV \leqq \frac{K(\beta)n}{\mu} (\text{Volume of } \mathcal{A}_n). \tag{64}$$

Now from (57)

$$\mu = \int_{r<\gamma d/2} (\gamma^2 d^2 - r^2) dV > \int_{r<\sqrt{\gamma^2 d^2 - 1}} dV = V_n(\sqrt{\gamma^2 d^2 - 1}) \tag{65}$$

where $V_n(r)$ is the volume of the sphere in $\mathcal{A}_n$ $S(\mathbf{z},r)$, which is independent of $\mathbf{z}$ (due to the symmetry of $\mathcal{A}_n$). Thus, (64) becomes

$$M(n, \sqrt{2\beta n}) \leqq \frac{nK(\beta)(2\pi)^n}{V_n(\sqrt{\gamma^2 d^2 - 1})}.$$

The asymptotic rate $\hat{R}(\beta)$ satisfies

$$\begin{aligned}\hat{R}(\beta) &= \operatorname*{limit}_{n\to\infty} \frac{1}{n} \ln M(n, \sqrt{2\beta n}) \\ &\leqq \lim_{n\to\infty} \frac{1}{n} \ln \frac{nK(\beta)(2\pi)^n}{V_n(\sqrt{2\gamma^2\beta n - 1})} \triangleq \hat{R}_U(\beta).\end{aligned} \tag{66}$$

Applying the result of Appendix C we have $\hat{R}(\beta) \leqq C_0(\gamma^2\beta)$ establishing the upper bound.

4.2 *Lower Bound on* $R(\beta)$

Again let us consider a maximum size n-dimensional code with minimum distance d and $M(n,d)$ code words. About each of the code words $\mathbf{x}_i (i = 1, 2, \cdots, M)$ consider the spheres $S_n(\mathbf{x}_i, d)$. We claim that the union of these spheres $\bigcup_{i=1}^{M} S_n(\mathbf{x}_i, d)$ covers the entire space $\mathcal{A}_n$. This follows from the fact that if $\mathbf{x}_0 \in \mathcal{A}_n$ is in no $S_n(\mathbf{x}_i, d)$, then $d(\mathbf{x}_0, \mathbf{x}_i) \geqq d$, $i = 1, 2, \cdots, M$, so that $\mathbf{x}_0$ may be added to the code destroying the maximality. If $V_n(d)$ is the volume of $S_n(\mathbf{x}_i, d)$ (independent of $\mathbf{x}_i$), then

$$M \cdot V_n(d) \geqq \text{volume of } \mathcal{A}_n = (2\pi)^n. \tag{67}$$

Thus, our lower bound is

$$M(n,d) \geqq \frac{(2\pi)^n}{V_n(d)}. \tag{68}$$

The asymptotic rate $\hat{R}(\beta)$ satisfies

$$\hat{R}(\beta) = \operatorname*{limit}_{n\to\infty} \frac{1}{n} \ln M(n, \sqrt{2\beta n}) \geqq \operatorname*{limit}_{n\to\infty} \frac{1}{n} \ln \frac{(2\pi)^n}{V_n(\sqrt{2\beta n})} \triangleq \hat{R}_L(\beta). \tag{69}$$

Again applying the result of Appendix C, we have $R(\beta) \geqq C_0(\beta)$ establishing the lower bound.

APPENDIX A

Asymptotic Estimates of the Channel Capacity

The channel capacity C is given by (12) as

$$\frac{C}{W} = -\int_0^\infty \hat{f}(\rho) \ln \frac{\hat{f}(\rho)}{\rho} \, d\rho + \ln \frac{2A}{e} \tag{70}$$

where

$$\hat{f}(\rho) = 2A\rho e^{-A(1+\rho^2)} I_0(2A\rho). \tag{71}$$

In this appendix we obtain estimates of C for large and small signal-to-noise ratio A.

A.1 *Large A:* We show here that

$$\frac{C}{W} = \tfrac{1}{2} \ln \frac{4\pi A}{e} + \varepsilon_1(A), \tag{72}$$

where $\varepsilon_1(A) \to 0$ as $A \to \infty$. To prove this we will show that for large A nearly all the contribution to the integral in (70) is for ρ in the neighborhood of unity. Part (*i*) is an estimate of this contribution. Part (*ii*) shows that the remaining contribution vanishes as $A \to \infty$.

(*i*) We shall show that if $\delta = A^{-\frac{1}{4}}$,

$$T(A) \triangleq -\int_{1-\delta}^{1+\delta} \hat{f}(\rho) \ln \frac{\hat{f}(\rho)}{\rho} \, d\rho \to \tfrac{1}{2} \ln \frac{\pi e}{A}, \tag{73}$$

as $A \to \infty$.

Using the asymptotic formula for $I_0(x)$ for large argument*

$$I_0(x) = \frac{e^x}{\sqrt{2\pi x}} \left[1 + 0\left(\frac{1}{x}\right) \right], \tag{74}$$

we obtain from (71) (for large A)

$$f(\rho) = \sqrt{\frac{A}{\pi}} \, \rho^{\frac{1}{2}} e^{-A(\rho-1)^2} \left[1 + 0\left(\frac{1}{A}\right) \right] \qquad 1 - \delta \leqq \rho \leqq 1 + \delta. \tag{75}$$

Substituting into (73) yields (after a change of variable)

$$T(A) = \left[-1 + 0\left(\frac{1}{A}\right) \right] [B_1 + B_2 + B_3] \tag{76}$$

* Ref. 5, p. 86.

where

$$B_1 = \int_{-\delta}^{\delta} \sqrt{\frac{A}{\pi}}\, e^{-Ax^2}(1+x)^{\frac{1}{2}} \ln \sqrt{\frac{A}{\pi}}\, e^{-Ax^2}\, dx \tag{76a}$$

$$B_2 = \int_{-\delta}^{\delta} \sqrt{\frac{A}{\pi}}\, e^{-Ax^2}(1+x)^{\frac{1}{2}} \ln (1+x)^{\frac{1}{2}}\, dx \tag{76b}$$

$$B_3 = \int_{-\delta}^{\delta} \sqrt{\frac{A}{\pi}}\, e^{-Ax^2}(1+x)^{\frac{1}{2}} 0\left(\frac{1}{A}\right) dx. \tag{76c}$$

Noting that the range of integration is $-\delta \leqq x \leqq \delta$ we can write

$$B_1 = K_1 \int_{-\delta}^{\delta} \sqrt{\frac{A}{\pi}}\, e^{-Ax^2} \ln \sqrt{\frac{A}{\pi}}\, e^{-Ax^2}\, dx \tag{77a}$$

$$B_2 = K_2 \int_{-\delta}^{\delta} \sqrt{\frac{A}{\pi}}\, e^{-Ax^2}\, dx \leqq K_2 \tag{77b}$$

$$B_3 = K_3 0\left(\frac{1}{A}\right) \int_{-\delta}^{\delta} \sqrt{\frac{A}{\pi}}\, e^{-Ax^2}\, dx \leqq K_3 0\left(\frac{1}{A}\right) \tag{77c}$$

where $(1-\delta)^{\frac{1}{2}} \leqq K_1, K_3 \leqq (1+\delta)^{\frac{1}{2}}$, $|K_2| \leqq (1+\delta)^{\frac{1}{2}} \ln (1+\delta)^{\frac{1}{2}}$ and $\delta = 1/A^{\frac{1}{4}}$. From (77b) and (77c) we see immediately that $B_2, B_3 \rightarrow 0$ as $A \rightarrow \infty$ so that we need consider only B_1. From (77a) (letting $y = \sqrt{2A}x$) and setting $\delta = A^{-\frac{1}{4}}$, we have

$$B_1 = \tfrac{1}{2} K_1 \ln \frac{A}{\pi} \int_{-\sqrt{2}A^{\frac{1}{4}}}^{\sqrt{2}A^{\frac{1}{4}}} \frac{1}{\sqrt{2\pi}}\, e^{-y^2/2}\, dy - \frac{K_1}{2} \int_{-\sqrt{2}A^{\frac{1}{4}}}^{\sqrt{2}A^{\frac{1}{4}}} \frac{y^2}{\sqrt{2\pi}}\, e^{-y^2/2}\, dy. \tag{78}$$

Since both integrals in (78) and K_1 tend to unity as $A \rightarrow \infty$, we have $B_1 \rightarrow \frac{1}{2} \ln (A/\pi e)$ as $A \rightarrow \infty$. Applying these results to (76) yields

$$\lim_{A\to\infty} T(A) = \lim_{A\to\infty} \left(-1 + 0\left(\frac{1}{A}\right)\right)\left(\tfrac{1}{2} \ln \frac{A}{\pi e} + B_2 + B_3\right) = \tfrac{1}{2} \ln \frac{\pi e}{A}$$

which is (73).

(*ii*) Here we shall show that with $\delta = A^{-\frac{1}{4}}$ as in (*i*) above,

$$\eta(A) \triangleq \int_{\substack{\rho \leqq 1-\delta \\ \rho \geqq 1+\delta}} \hat{f}(\rho) \ln \frac{\hat{f}(\rho)}{\rho}\, d\rho \rightarrow 0, \quad \text{as} \quad A \rightarrow \infty. \tag{79}$$

To do this we write

$$\eta(A) = \int_0^{\alpha} \hat{f}(\rho) \ln \frac{\hat{f}(\rho)}{\rho}\, d\rho + \int_{\alpha}^{1-\delta} f(\rho) \ln \frac{\hat{f}(\rho)}{\rho}\, d + \int_{1+\delta}^{\infty} f(\rho) \ln \frac{\hat{f}(\rho)}{\rho}\, d\rho = C_1 + C_2 + C_3, \tag{80}$$

where $\alpha(0 < \alpha < \frac{1}{2})$ is arbitrary. We will show that for arbitrary $\varepsilon > 0$, we can choose A sufficiently large so that $\eta(A) < \varepsilon$. Let us consider each of the integrals C_1, C_2 and C_3 of (80) in turn.

C_1 : For $0 \leqq \rho \leqq \alpha$ we may write

$$\frac{\hat{f}(\rho)}{\rho} \leqq 2Ae^{-A}I_0(2A\alpha), \tag{81}$$

since $I_0(x)$ is an increasing function of x. Making use of the asymptotic formula for $I_0(x)$ (74) we obtain from (81)

$$\left|\frac{\hat{f}(\rho)}{\rho}\right| < \sqrt{\frac{A}{\alpha\pi}}\, e^{-A(1-2\alpha)} \left(1 + 0\left(\frac{1}{A}\right)\right) \rightarrow 0, \quad \text{as} \quad A \rightarrow \infty,$$

since $\alpha < \frac{1}{2}$. Thus, with A sufficiently large,

$$\left|\frac{\hat{f}(\rho)}{\rho} \log \frac{\hat{f}(\rho)}{\rho}\right| \leqq \frac{2\varepsilon}{3\alpha^2}, \qquad 0 \leqq \rho \leqq \alpha$$

and

$$|C_1| \leqq \int_0^\alpha \frac{2\varepsilon}{3\alpha^2}\, \rho\, d\rho = \frac{\varepsilon}{3}. \tag{82}$$

C_2 : Again using the asymptotic formula for $I_0(x)$ (74) we may write, for $\alpha \leqq \rho \leqq 1 - \delta$,

$$\begin{aligned} \frac{\hat{f}(\rho)}{\rho} &= \sqrt{\frac{A}{\pi\alpha}}\, e^{-A(1-\rho)^2} \left(1 + 0\left(\frac{1}{A}\right)\right) \leqq \sqrt{\frac{A}{\pi\alpha}}\, e^{-A\delta^2} \left(1 + 0\left(\frac{1}{A}\right)\right) \\ &= \sqrt{\frac{A}{\pi\alpha}}\, e^{-A^{\frac{1}{3}}} \left(1 + 0\left(\frac{1}{A}\right)\right) \rightarrow 0 \quad \text{as} \quad A \rightarrow \infty. \end{aligned} \tag{83}$$

Thus, with A sufficiently large

$$\left|\frac{\hat{f}(\rho)}{\rho} \ln \frac{\hat{f}(\rho)}{\rho}\right| \leqq \tfrac{2}{3}\varepsilon,$$

from which

$$|C_2| \leqq \tfrac{2}{3}\varepsilon \int_\alpha^{1-\delta} \rho\, d\rho \leqq \frac{\varepsilon}{3}. \tag{84}$$

C_3 : As above, we may write for $\rho \geqq 1 + \delta$,

$$\frac{\hat{f}(\rho)}{\rho} = \sqrt{\frac{A}{\pi\rho}}\, e^{-A(\rho-1)^2} \left(1 + 0\left(\frac{1}{A}\right)\right). \tag{85}$$

Substituting (85) into the defining integral for C_3 (80), and making change of variable $y = (\rho - 1)$, we obtain

$$\begin{aligned} C_3 = \int_\delta^\infty \sqrt{\frac{A}{\pi}}\,(1+y)^{\frac{1}{2}}\,e^{-Ay^2}\ln\sqrt{\frac{A}{\pi}}\,dy \\ + \int_\delta^\infty \sqrt{\frac{A}{\pi}}\,(1+y)^{\frac{1}{2}}\,e^{-Ay^2}\ln\,(1+y)^{-\frac{1}{2}}\,dy \\ + \int_\delta^\infty \sqrt{\frac{A}{\pi}}\,(1+y)^{\frac{1}{2}}\,e^{-Ay^2}\ln e^{-Ay^2}\,dy \\ + \int_\delta^\infty \sqrt{\frac{A}{\pi}}\,(1+y)^{\frac{1}{2}}\,e^{-Ay^2}\ln\left(1+0\left(\frac{1}{A}\right)\right)dy. \end{aligned} \tag{86}$$

Since for $y \geqq \delta$, $(1+y)^{\frac{1}{2}} \leqq 2e^{y^2}$ and $|\,(1+y)^{\frac{1}{2}}\ln\,(1+y)^{-\frac{1}{2}}\,| \leqq e^{y^2}$, we have from (86)

$$\begin{aligned} |C_3| \leqq \int_\delta^\infty \sqrt{\frac{A}{\pi}}\,e^{-(A-1)y^2}\left[\tfrac{1}{2}\ln\frac{A}{\pi} + 2 + 0\left(\frac{1}{A}\right)\right]dy \\ + \int_\delta^\infty \sqrt{\frac{A}{\pi}}\,e^{-(A-1)y^2}(Ay^2)\,dy. \end{aligned} \tag{87}$$

Using the well known asymptotic formula for the cumulative error function, and the fact that $\delta = A^{-\frac{1}{4}}$, it is readily shown that for large A,

$$\int_\delta^\infty \sqrt{\frac{A}{\pi}}\,e^{-(A-1)y^2}\,dy \approx \frac{1}{\sqrt{2\pi}A^{\frac{1}{4}}}\,e^{-A^{\frac{1}{2}}}, \tag{88a}$$

and

$$\int_\delta^\infty \sqrt{\frac{A}{\pi}}\,e^{-(A-1)y^2}\,(Ay^2)\,dy \approx \frac{A^{\frac{1}{4}}}{\sqrt{2\pi}}\,e^{-A^{\frac{1}{2}}}. \tag{88b}$$

Equations (88a and b) tell us that with A sufficiently large $|\,C_3\,| \leqq \varepsilon/3$.

Taking the above results together yields

$$|\,\eta(A)\,| \leqq |\,C_1\,| + |\,C_2\,| + |\,C_3\,| \leqq \varepsilon,$$

with A sufficiently large.

(*iii*) The final step is to substitute (73) and (79) into (70) and obtain

$$\frac{C}{W} \to \tfrac{1}{2}\ln\frac{\pi e}{A} + \ln\frac{2A}{e} = \tfrac{1}{2}\ln\frac{4\pi}{e}A, \quad \text{as} \quad A \to \infty,$$

which is what is to be proved, (72).

A.2 *Small A:* We show here that

$$\frac{C}{W} = A(1 + 0(A)), \tag{89}$$

as $A \to 0$.

Substituting (71) into (70) yields, after a bit of straightforward manipulation,

$$\begin{aligned}\frac{C}{W} = A - 1 + 2A^2e^{-A} \int_0^\infty \rho^3 e^{-A\rho^2} I_0(2\rho A)d\rho \\ - 2Ae^{-A} \int_0^\infty \rho e^{-A\rho^2} I_0(2\rho A) \ln I_0(2\rho A)d\rho.\end{aligned} \tag{90}$$

If we change the variable of integration to $x = 2\rho A$, we obtain from (90)

$$\begin{aligned}\frac{C}{W} = A - + \frac{e^{-A}}{8A^2} \int_0^\infty x^3 I_0(x)e^{-x^2/4A}dx \\ - \frac{e^{-A}}{2A} \int_0^\infty xI_0(x)e^{-x^2/4A} \ln I_0(x)dx.\end{aligned} \tag{91}$$

Now the first integral of (90) is known* and is

$$\int_0^\infty x^3 I_0(x)e^{-x^2/4A}\, dx = 8A^2(1 + A)e^A, \tag{92}$$

so that

$$\frac{C}{W} = 2A - \frac{e^A}{2A} \int_0^\infty xI_0(x)e^{-x^2/4A} \ln I_0(x)dx = 2A - \frac{e^{-A}}{2A} D. \tag{93}$$

We can estimate the integral D for small A, by noting that most of the contribution is for small x. Making use of the asymptotic formula for $I_0(x)$, for small x

$$I_0(x) = 1 + \frac{x^2}{4} + 0(x^4), \tag{94}$$

we have

$$D = \int_0^\infty \left[\frac{x^3}{4} + 0(x^5)\right] e^{-x^2/4A}\, dx. \tag{95}$$

* Ref. 6, p. 198, (4a).

Since

$$\int_0^\infty \frac{x^3}{4} e^{-x^2/4A}\, dx = 2A^2,$$

and

$$\int_0^\infty x^5 e^{-x^2/4A}\, dx = 64A^3\Gamma(3) = 0(A^3),$$

we have

$$D = 2A^2(1 + 0(A)). \tag{96}$$

From (96) and (93) we get

$$\frac{C}{W} = 2A - Ae^{-A}(1 + 0(A)) = A(1 + 0(A))$$

which is what was to be proved (89).

APPENDIX B

The Function $\lambda(\xi)$

In this appendix, we show that for ξ satisfying

$$0 < \xi \leqq 1, \tag{97}$$

there exists a unique $\lambda(\xi)$ which satisfies

$$\xi = 1 - \frac{I_1(\lambda(\xi))}{I_0(\lambda(\xi))}. \tag{98}$$

If we define the function $\xi(\lambda)$ by

$$\xi(\lambda) = 1 - \frac{I_1(\lambda)}{I_0(\lambda)}, \qquad 0 \leqq \lambda < \infty, \tag{99}$$

it will suffice to show that

(*i*) $\xi(\lambda)$ is strictly monotone decreasing,
(*ii*) $\xi(0) = 1$,
(*iii*) $\lim_{\lambda \to \infty} \xi(\lambda) = 0$.

If (*i*), (*ii*), and (*iii*) are true, $\xi(\lambda)$ is a one-to-one mapping of the half line $[0, \infty)$ onto the interval $(0,1]$.

(i) Making use of the fact that $I_0'(\lambda) = I_1(\lambda)$ we can write

$$\frac{d\xi(\lambda)}{d\lambda} = \frac{-I_0(\lambda)I_1'(\lambda) + I_1^2(\lambda)}{(I_0(\lambda))^2}. \tag{100}$$

Since*

$$I_0(\lambda) = \frac{1}{\pi}\int_0^\pi e^{\lambda \cos \varphi}\, d\varphi,$$

we have

$$I_1(\lambda) = I_0'(\lambda) = \frac{1}{\pi}\int_0^\pi \cos \varphi e^{\lambda \cos \varphi}\, d\varphi,$$

and

$$I_1'(\lambda) = \frac{1}{\pi}\int_0^\pi \cos^2 \varphi e^{\lambda \cos \varphi}\, d\varphi.$$

Thus (100) becomes

$$\frac{d\xi(\lambda)}{d\lambda} = \frac{-\frac{1}{\pi^2}\int_0^\pi e^{\lambda \cos \varphi}\, d\varphi \int_0^\pi \cos^2 \varphi e^{\lambda \cos \varphi}\, d\varphi + \frac{1}{\pi^2}\left(\int_0^\pi \cos \varphi e^{\lambda \cos \varphi}\, d\varphi\right)^2}{[I_0(\lambda)]^2}. \tag{101}$$

By the Schwarz inequality

$$\left(\int_0^\pi \cos \varphi e^{\lambda \cos \varphi}\, d\varphi\right)^2 < \left(\int_0^\pi \cos^2 \varphi e^{\lambda \cos \varphi}\, d\varphi\right)\left(\int_0^\pi e^{\lambda \cos \varphi}\, d\varphi\right),$$

(the strict inequality holding). Hence $\frac{d\xi(\lambda)}{d\lambda} < 0$ and (i) follows.

$$(ii)\ \xi(0) = 1 - \frac{I_1(0)}{I_0(0)} = 1 - \frac{0}{1} = 1.$$

(iii) We make use of the asymptotic formula for $I_0(x)$ and $I_1(x)$ for large x†

$$\begin{aligned} I_0(x) &= \frac{e^x}{\sqrt{2\pi x}}\left[1 + \frac{1}{8x} + 0\left(\frac{1}{x^2}\right)\right] \\ I_1(x) &= \frac{e^x}{\sqrt{2\pi x}}\left[1 - \frac{3}{8x} + 0\left(\frac{1}{x^2}\right)\right]. \end{aligned} \tag{102}$$

Substitution of (102) into (99) yields (iii) immediately.

* Ref. 4, p. 76.
† Ref. 4, p. 86.

Let us remark here that since $I_0(x)$ and $I_1(x)$ are even functions of x, if $\lambda(\xi) = a \geqq 0$ is the unique nonnegative solution to (98), then $\lambda(\xi) = -a$ is the unique nonpositive solution to (98).

APPENDIX C

Completion of Asymptotic Estimates of $R(\beta)$

We have defined $V_n(r)$ as the volume of the sphere $S_n(\mathbf{z},r) = \{x \in \mathfrak{A}_n : d(\mathbf{x},\mathbf{z}) < r\}$. Due to the symmetry of $\mathfrak{A}_n$, $V_n(r)$ is independent of $\mathbf{z}$. Thus, we shall take $V_n(r)$ as the volume of

$$S(\bar{0},\mathbf{r}) = \left\{\mathbf{x} = (x_1, x_2, \cdots, x_n) \in \mathfrak{A}_n:\ d^2(\bar{0},\mathbf{x}) = \sum_{k=1}^{n} \left(2 \sin \frac{x_k}{2}\right)^2 < r^2\right\}.$$

In this appendix, we evaluate

$$\lim_{n\to\infty} \frac{1}{n} \ln \frac{(2\pi)^n}{V_n(\sqrt{an})} \triangleq E_a . \tag{103}$$

We shall find E_a by solving an equivalent probability problem: Let $X_1, X_2, \cdots$ be a sequence of independent random variables uniformly distributed on the interval $[-\pi,\pi]$. Let

$$Y_n = \sum_{k=1}^{n} \left(2 \sin \frac{X_k}{2}\right)^2 .$$

It is clear that

$$\Pr[Y_n < r^2] = \frac{V_n(r)}{(2\pi)^n}, \tag{104}$$

hence,

$$-\lim_{n\to\infty} (1/n) \ln \Pr[Y_n < an] = E_a . \tag{105}$$

We now make use of

Chernoff's Theorem:[7] *Let $Z_1, Z_2, \cdots$ be a sequence of independent identically distributed random variables with moment generating function $E[e^{Z_k t}] = M(t)$. Let*

$$P_n = \Pr\left[\sum_{k=1}^{n} Z_i \leqq an\right],$$

where $a \leqq E(Z_k)$. Then

$$\lim_{n\to\infty} \frac{1}{n} \ln P_n = \ln m,$$

where $m = \min_{t \leqq 0} e^{-at} M(t)$.

If we set

$$Z_k = \left[2 \sin \frac{X_k}{2}\right],$$

where X_k is the above random variable, then

$$Y_n = \sum_{k=1}^{n} Z_k .$$

Thus, from (105) and Chernoff's Theorem, $E_a = -\ln m$.

The moment generating function of Z_k is

$$\begin{aligned} M(t) = E[e^{Z_k t}] &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp \left\{ \left(2 \sin \frac{x}{2}\right)^2 t \right\} dx \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{(2t - 2t \cos x)}\, dx = e^{2t} I_0(2t). \end{aligned} \tag{106}$$

Hence,

$$m = \min_{t \leqq 0} e^{(2-a)t} I_0(2t). \tag{107}$$

To find the minimum, set the derivative of (107) equal to zero:

$$0 = e^{(2-a)t}[(2 - a)I_0(2t) + 2I_1(2t)],$$

so that the t which minimizes (107) satisfies

$$\frac{a}{2} = 1 - \frac{I_1(2t)}{I_0(2t)} . \tag{108}$$

The solution to (108), for $t \leqq 0$, is $2t = -\lambda(a/2)$, where $\lambda(\xi)$ is defined by (12). (See the remark at the conclusion of Appendix B.) Hence, from (107)

$$m = \exp\left[-\left(1 - \frac{a}{2}\right)\lambda\left(\frac{a}{2}\right)\right] I_0\left(\lambda\left(\frac{a}{2}\right)\right), \tag{109}$$

so that

$$\begin{aligned} E_a = -\ln m &= -\ln I_0\left(\lambda\left(\frac{a}{2}\right)\right) + \left(1 - \frac{a}{2}\right)\lambda\left(\frac{a}{2}\right) \\ &= C_0\left(\frac{a}{2}\right) \end{aligned} \tag{110}$$

where $C_0(\xi)$ is defined by (16).

Applying (110) to (66) yields $\hat{R}_U(\beta) = C_0(\gamma^2\beta)$, and applying (110) to (69) yields $\hat{R}_L(\beta) = C_0(\beta)$.

APPENDIX D

Exponential Error Bounds

It is known that for any time-discrete (amplitude continuous) memoryless channel the smallest attainable error probability $P_e^*(n,\hat{R})$ for an n-dimensional code with $e^{n\hat{R}}$ code words may be written

$$P_e^*(n,\hat{R}) = \exp\,[-nE(\hat{R}) + o(n)], \tag{111}$$

where $E(\hat{R}) > 0$ when $\hat{R} < \hat{C}$ (the channel capacity in nats per symbol). Although $E(\hat{R})$ is not always known exactly it can be estimated by upper and lower bounds. The best known lower bound on $E(\hat{R})$ is given in Gallager (Ref. 8, Theorem 10) and the best known upper bound on $E(\hat{R})$ by Shannon, Gallager, and Berlekamp.[9]

Let $P(y \mid x)$ be the channel transition probability density. We assume that any n-sequence of input symbols is an allowable channel input — i.e., no "input constraint". The bounds of Refs. 8 and 9 can then be stated as follows:

For any $\rho \geqq 0$ and input probability density $f(x)$, let us define

$$E(\rho,f) = E_0(\rho,f(x)) - \rho R_o(\rho,f(x)), \tag{112}$$

where

$$E_0(\rho,f(x)) = -\ln \int_y dy \left[\int_x dx\, f(x) P(y \mid x)^{1/1+\rho} \right]^{1+\rho} \tag{112a}$$

and

$$R_o(\rho,f(x)) = \frac{\partial}{\partial\rho} E_0(\rho,f(x)). \tag{112b}$$

With $\rho \geqq 0$ specified let $f_\rho(x)$ be that input density which maximizes $E(\rho.f(x))$. It is shown in Ref. 8 that with ρ fixed $f_\rho(x)$ is the unique density which satisfies

$$\int_y dy\, P(y \mid x)^{1/1+\rho} \alpha_\rho^{\rho}(y) \geqq \int_y \alpha_\rho(y)^{1+\rho} dy, \qquad \text{all } x, \tag{113}$$

with equality if $f_\rho(x) \neq 0$ (all x) where

$$\alpha_\rho(y) = \int_x f_\rho(x) P(y \mid x)^{1/1+\rho} dx. \tag{113a}$$

It can be shown that with $\rho = 0$, $f_0(x)$ is that input density which achieves capacity $\hat{C}$, and $R_o(0,f_0) = \hat{C}$. In most channels of interest $R_o(\rho,f_\rho)$ decreases from $\hat{C}$ to 0 as ρ increases from 0 to ∞.

We define the rate $\hat{R}$ parametrically in terms of ρ by

$$\hat{R} = \hat{R}(\rho) = R_o(\rho,f_\rho(x)). \tag{114}$$

Then for $0 \leqq \rho \leqq 1$, which corresponds to $\hat{R}_o(1,f_1(x)) \leqq \hat{R} \leqq \hat{C}$, the exponent is known exactly:

$$E(\hat{R}) = E(\rho,f_\rho) = E_0(\rho,f_\rho) - \rho\hat{R} \tag{115}$$

where E, E_0, and f_ρ are defined by (112). For $\rho \geqq 1$, which corresponds to $\hat{R} \leqq \hat{R}_o(1,f_1)$, the ("sphere-packing"), upper bound on $E(\hat{R})$ is

$$E(\hat{R}) \leqq E_0(\rho,f_\rho), \tag{116}$$

and the ("random-coding") lower bound is for $0 \leqq \hat{R} \leqq \hat{R}_o(1,f_1)$

$$\mathrm{E}(\hat{R}) \geqq E_0(1,f_1) - \hat{R}. \tag{117}$$

This estimate of $E(\hat{R})$ may be improved for low rates $\hat{R}$. It is shown in Ref. 9 that if $E^*(\hat{R})$ is an upper bound on $E(\hat{R})$ which is sharper than the sphere-packing bound (116) for low rates $\hat{R}$ (such a bound can always be found), and if $E^*(\hat{R})$ and the sphere-packing bound are plotted versus $\hat{R}$, then their common tangent is also an upper bound on $E(\hat{R})$.

The lower bound may be sharpened for low rates $\hat{R}$ as follows. For $\rho \geqq 1$, and input density $g(x)$, define

$$E_x(\rho,g) = E_{0x}(\rho,g) - \rho R_{0x}(\rho,g), \tag{118}$$

where

$$E_{0x}(\rho,g) = -\rho \ln \int_x g(x)dx \int_{x'} g(x')dx' \left[\int_y P(y \mid x)^{\frac{1}{2}} P(y \mid x')^{\frac{1}{2}} dy\right]^{1/\rho} \tag{118a}$$

and

$$R_{0x}(\rho,g) = \frac{\partial}{\partial\rho} E_{0x}(\rho,g). \tag{118b}$$

Then for any fixed $g(x)$ and with $\hat{R}$ again given parametrically in terms of ρ by

$$\hat{R} = R_{0x}(\rho,g), \tag{119}$$

the ("expurgated") lower bound is

$$E(R) \geqq E_x(\rho,g). \tag{120}$$

We shall now apply these results to our channel using the time-discrete model defined before and after (26). Here the input is a number $X \,\varepsilon\, [-\pi,\pi]$, and the output is a pair of real numbers (Y_1, Y_2). If $X = x$ is the input, then the conditional transition probability density is the two-dimensional

$$P(y_1, y_2 \mid x) = \frac{1}{2\pi N} \exp \{-[(y_1 - \sqrt{2S} \cos x)^2 + (y_2 - \sqrt{2S} \sin x)^2]/2N\}$$

or in polar coordinates

$$\begin{aligned} P(r,\varphi \mid x) &= rP(r \cos \varphi, r \sin \varphi \mid x) \\ &= \frac{r}{2\pi N} e^{-S/N} e^{-r^2/2N} \exp\left(-r \frac{\sqrt{2S'}}{N} \cos (\varphi - x)\right). \end{aligned} \tag{121}$$

It may be verified by substitution into (113), that the input density $f(x) = 1/2\pi$ maximizes $E(\rho,f(x))$ for all $\rho \geqq 0$. Further a direct substitution of (121) into (112a) yields after a straightforward computation

$$E_0(\rho,f\rho) = -\ln 2Ae^{-A} \int_0^\infty ve^{-Av^2} \left[I_0\left(\frac{2Av}{1+\rho}\right)\right]^{1+\rho} dv, \tag{122}$$

where $A = S/N$, the signal-to-noise ratio. The rate, $\hat{R}$, can be gotten by differentiating (122) with respect to ρ. This yields

$$\begin{aligned} \hat{R}(\rho) &= \frac{\partial}{\partial\rho} E_0(\rho,f_\rho) \\ &= \left[-\int_0^\infty ve^{-Av^2} I_0\left(\frac{2vA}{1+\rho}\right)^{1+\rho} \ln I_0\left(\frac{2vA}{1+\rho}\right) dv \right. \\ &\quad \left. + \frac{2A}{(1+\rho)} \int_0^\infty v^2 e^{-Av^2} I_0\left(\frac{2Av}{1+\rho}\right)^{\rho} I_1\left(\frac{2Av}{1+\rho}\right) dv\right] \Big/ \\ &\quad \int_0^\infty ve^{-Av^2} I_0\left(\frac{2Av}{1+\rho}\right)^{1+\rho} dv. \end{aligned} \tag{123}$$

After some manipulation one can show that $\hat{R}(\rho)\mid_{\rho=0} = \hat{C}$, the channel capacity as given by (12). The estimate of the exponent $E(\hat{R})$ of (115), (116), and (117) is plotted versus $\hat{R}$ in Fig. 7 for signal-to-noise ratio

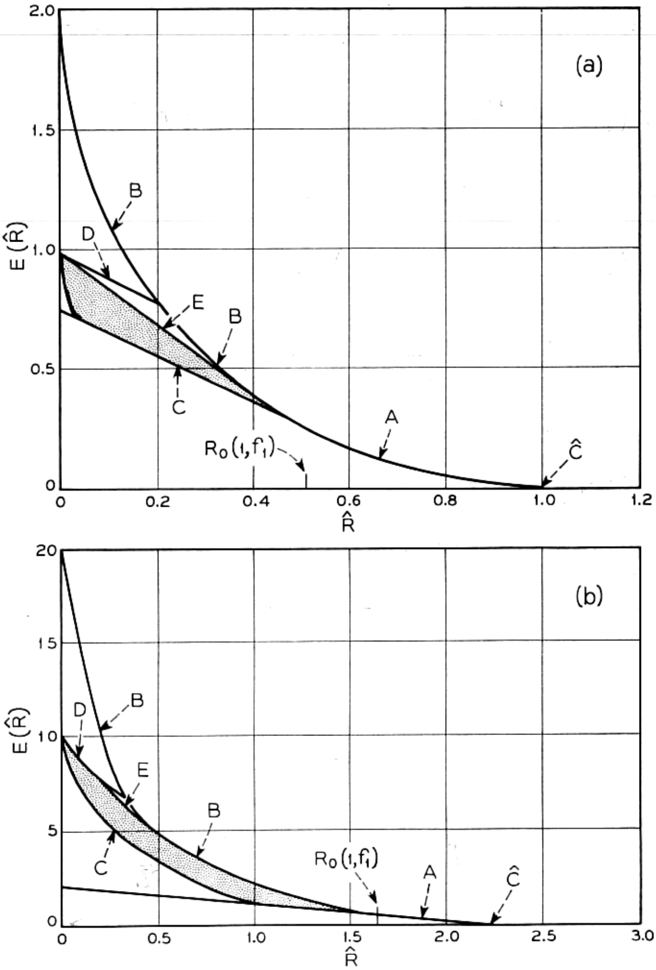

Fig. 7 — Upper and lower bounds on the error exponent $E\ (\hat{R})$ vs. $\hat{R}$ for signal-to-noise ratios of (a) $A = 2$, (b) $A = 20$. (Curve A is the exponent $E\ (\hat{R})$ in the range where it is known exactly (115). Curve B is the "sphere-packing" upper bound on $E\ (\hat{R})$ (116). Curve C is the "random coding" lower bound on $E\ (\hat{R})$ (117). Curve D is $E^*\ (\hat{R})$, the low rate upper bound (133). Curve E is the common tangent to $E^*\ (\hat{R})$ and the sphere-packing bound. $E\ (\hat{R})$ lies in the shaded region.

$A = 2{,}20$. The sphere-packing upper bound and the random-coding lower bound diverge for small rates $\hat{R}$. We shall improve this situation by computing the low-rate expurgated lower bound on $E(\hat{R})$ (120). If we again choose the input density to be $g(x) = 1/2\pi$, $-\pi \leqq x \leqq \pi$ we have from (121)

$$\int_y P(y \mid x)^{\frac{1}{2}} P(y \mid x')^{\frac{1}{2}} dy = \int_0^\infty dr \frac{r}{2\pi N} \exp(-S/N_e - r^2/2N)$$

$$\cdot \int_{-\pi}^{\pi} \exp\left(-\frac{r}{N}\sqrt{\frac{S}{2}}[\cos(\varphi - x) + \cos(\varphi - x')]\right) d\varphi$$

$$= \int_0^\infty dr \frac{r}{2\pi N} \exp(-S/N - r^2/2N)$$

$$\cdot \int_{-\pi}^{\pi} \exp\left(-\frac{\sqrt{S}}{N} Br \cos(\varphi - a)\right) d\varphi$$

$$= \frac{\exp(-S/N)}{N} \int_0^\infty r \exp(-r^2/2N) I_0\left(\frac{\sqrt{S}}{N} Br\right) dr,$$

where

$$B = \sqrt{1 + \cos(x - x')} \quad \text{and} \quad a = \tan^{-1}\left[\frac{\sin x + \sin x'}{\cos x + \cos x'}\right].$$

This integral is tabulated [Ref. 6, p. 198, #5] so that we have

$$\int_y P(y \mid x)^{\frac{1}{2}} P(y \mid x')^{\frac{1}{2}} dy = \exp(-(S/2N)[1 - \cos(x - x')]). \quad (124)$$

Substituting (124) into (118a) yields

$$\begin{aligned} E_{0x}(\rho,g) &= -\rho \ln \int_{-\pi}^{\pi} \frac{dx}{2\pi} \int_{-\pi}^{\pi} \frac{dx'}{2\pi} \\ &\quad \cdot \exp[-S/2N\rho - (S/2N\rho)\cos(x - x')] \\ &= -\rho \ln \int_{-\pi}^{\pi} \frac{dx}{2\pi} \exp(-S/2N\rho) I_0\left(\frac{S}{2N\rho}\right) \\ &= -\rho \ln[\exp(-A/2\rho) I_0(A/2\rho)], \qquad (\rho \geqq 1) \end{aligned} \quad (125)$$

where $A = S/N$. The rate $\hat{R}$ is given parametrically in terms of ρ by

$$\hat{R} = \hat{R}(\rho) = \frac{\partial E_{0x}}{\partial \rho}(\rho,g) = \frac{A}{2\rho}\frac{I_1(A/2\rho)}{I_0(A/2\rho)} - \ln I_0(A/2\rho) \qquad (\rho \geqq 1). \quad (126)$$

Let us note that as $\rho \to \infty$ $\hat{R}(\rho) \to 0$. The expurgated bound is given by (120), (125), and (126). It is easy to show that as $\rho \to \infty$, $(\hat{R} \to 0)$ the lower bound $E_x(\rho,g) \to A/2$.

We shall now obtain a sharper upper bound for low rates $E^*(\hat{R})$ which will in fact have $E^*(0) = A/2$, establishing that $E(0) = A/2$.

Let us denote by $\rho_n(M)$, the smallest maximum (normalized) correlation obtainable for an n-dimensional polyphase code with M code words. Paralleling arguments of Shannon (Ref. 10, pp. 647–648) it is not hard to show that the error probability for a code with $M = \exp(n\hat{R})$ code words satisfies

$P_e \geqq \frac{1}{2}\text{Pr}$ [error in a code with two code words with energy ST and correlation $\rho_n(M/2)$ in white Gaussian noise with spectral density N_o].

The right member of this inequality is known [Ref. 11, (38)], and is equal to

$$\tfrac{1}{2}\Phi\left(-\sqrt{\frac{ST}{N_0}\left(1-\rho_n\left(\frac{M}{2}\right)\right)}\right), \tag{127}$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-u^2/2}du$$

is the cumulative error function.

We now bound $E(\hat{R})$ by finding a bound on $\rho_n(M/2)$. Since for large M, a code with $M/2$ code words has about the same rate as one with M code words, it will suffice to bound $\rho_n(M)$. With $M = \exp(n\hat{R})$ and n large, we have from (17) (since $\beta = (1 - \rho)$)

$$\hat{R} \leqq C_0(1 - \sqrt{\rho_n(M)}). \tag{128}$$

If we define $\hat{R}$ parametrically by

$$\hat{R} = \hat{R}(\sigma) = C_0(\sigma), \qquad 0 \leqq \sigma \leqq 1, \tag{129}$$

we have from (129)

$$\sigma \geqq 1 - \sqrt{\rho_n(M)} \tag{130}$$

or

$$[1 - \rho_n(M)] \leqq \sigma(2 - \sigma). \tag{131}$$

Substituting (131) into (127) yields

$$P_e \leqq \tfrac{1}{2}\Phi\left(-\sqrt{\frac{ST}{N_o}\,\sigma(2-\sigma)}\right). \tag{132}$$

Making use of the well-known asymptotic formula for the cumulative error function $\Phi(-x) \approx (1/\sqrt{2\pi}x)\exp(-x^2/2)$ (large x), we obtain from (132) for large T (and therefore large $n = WT$), the upper bound on the error exponent

$$E(\hat{R}) = -\lim_{n\to\infty} \frac{1}{n} \ln P_e \leqq \frac{A}{2} \sigma(2 - \sigma) \triangleq E^*(R) \tag{133}$$

where $A = S/N_oW = S/N$. When $R = 0, \sigma = 1$, so that $E(0) = A/2$. The expurgated bound and the bound of (133) are plotted in Fig. 7. The upper bound is, of course, sharpened by drawing the common tangent of $E^*(\hat{R})$ and the sphere-packing bound.

APPENDIX E

Asymptotic Estimates of $C_0(\xi)$

In this appendix we obtain estimates of $C_0(\xi)$ as $\xi \to 0$ and $\xi \to 1$.

E.1 Small ξ: We show here that

$$C_0(\xi) = \tfrac{1}{2} \ln \frac{\pi}{e\xi} + \varepsilon_1(\xi) \tag{134}$$

where $\varepsilon_1(\xi) \to 0$ as $\xi \to \infty$.

From proposition (*iii*) in Appendix C, we know that as $\xi \to 0$, $\lambda(\xi) \to \infty$. Again making use of the asymptotic formula for $I_0(x)$ and $I_1(x)$ for large x (102), we obtain by substitution into (15),

$$\xi = \frac{1}{2\lambda}\left[1 + 0\left(\frac{1}{\lambda}\right)\right]. \tag{135}$$

or

$$\lambda = \frac{1}{2\xi} + 0(1) = \frac{1}{2\xi} + k + \varepsilon(\xi), \tag{136}$$

where $\varepsilon(\xi) \to 0$ as $\xi \to 0$ and k is a constant. Substitution of (136) into (16), and another application of the asymptotic formula for $I_0(x)$ yields (134).

E.2 Large ξ: We show here that

$$C_0(\xi) = (1 - \xi)^2\{1 + 0[(1 - \xi)^2]\} \tag{137}$$

as $\xi \to 1$. As above our first task is to estimate $\lambda(\xi)$ when ξ is near unity or $\lambda(\xi)$ is near zero. We need the asymptotic formulas for $I_0(x)$ and $I_1(x)$ for x near zero (Ref. 4, p. 77):

$$I_0(x) = 1 + \frac{x^2}{4} + 0(x^4),$$
$$I_1(x) = \frac{x}{2} + \frac{x^3}{16} + 0(x^5). \tag{138}$$

Substituting (138) into (15) yields

$$\xi = 1 - \frac{I_1(\lambda)}{I_0(\lambda)} = 1 - \frac{\lambda}{2} + 0(\lambda^3). \tag{139}$$

Setting $\hat{\xi} = 1 - \xi$ we have,

$$\hat{\xi} = \frac{\lambda}{2} + 0(\lambda^3). \tag{140}$$

We show that

$$\lambda = 2\hat{\xi} + 0(\hat{\xi}^3) = 2(1 - \xi) + 0((1 - \xi)^3). \tag{141}$$

Equation (141) follows on setting $x = \lambda - 2\xi$ and observing [from (140)] that

$$\frac{x}{\hat{\xi}^3} = \frac{0(\lambda^3)}{\left[\frac{\lambda}{2} + 0(\lambda)^3\right]^3} \to k,$$

as $\lambda \to 0$ or $\hat{\xi} \to 0$ $(\xi \to 1)$. Substitution of (141) into (16) and another application of the asymptotic formula for $I_0(x)$, yields (137).

APPENDIX F

Comparison of Modulation Schemes

In this appendix, we shall describe an amplitude modulation scheme and compare its performance with that of the phase modulation scheme studied in this paper.

Referring to (21a) we see that our phase modulated signal may be written (during the kth subinterval)

$$s_i(t) = x_{ik}^{(1)} \sin \alpha 2\pi Wt + x_{ki}^{(2)} \cos \alpha 2\pi Wt, \tag{142}$$

where from (21b)

$$[x_{ik}^{(1)}]^2 + [x_{ij}^{(2)}]^2 = 2S, \qquad k = 1, 2, \cdots, n. \tag{143}$$

Consider an amplitude modulation (AM) scheme in which the signals $s_i(t)$ are given by (142) but with (143) replaced by the "mean square" constraint

$$\sum_{k=1}^{n} \{[x_{ik}^{(1)}]^2 + [x_{ik}^{(2)}]^2\} \leqq 2Sn = 2WST. \tag{144}$$

The resulting signals $s_i(t)$ are then amplitude modulated signals with carrier frequency $\alpha 2\pi W$ radians per second and average power

$$\frac{1}{T}\int_0^T s_i^2(t)dt = \frac{1}{T}\sum_{k=1}^{n}\int_{(k-1)T/n}^{kT/n} s_i^2(t)dt \leqq S. \tag{145}$$

Thus, in this case the signals are constrained to have average power not exceeding S. It is clear that, as for the phase modulation, the signals of the αth and βth users of the channel are orthogonal so that we may again take the bandwidth (i.e., difference in carrier frequencies of adjacent users) to be W cps. Further, it follows from the analysis in Section III that this channel is mathematically equivalent to the time-discrete channel Gaussian channel considered by Shannon.[10,12] This channel accepts real numbers at a rate of $2W$ per second and adds to each number an independent Gaussian variate with mean zero and variance N_0W. Messages are encoded in blocks (vectors) of $2WT$ real numbers (which take T seconds to transmit), each $2WT$-vector having the sum of the squares of the coordinates not exceeding $2WST$. Shannon has found the capacity of this channel to be (in nats per second)

$$W \ln\left(1 + \frac{S}{N_0W}\right) = W \ln(1 + A), \tag{146}$$

where $A = S/N_0W$, the signal to noise ratio. Equation (146) is plotted in Fig. 2 so that it may be compared to the capacity of the polyphase system. Note that for small A, $\ln(1 + A) \approx A$ so that from (14) the

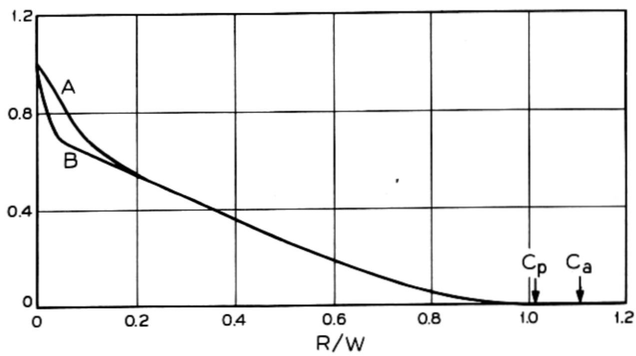

Fig. 8 — Lower bound on the exponents $E_a(R)$ (curve A) and $E_p(R)$ curve B). C_p and C_a are the capacities of the polyphase and AM systems, respectively. The signal-to-noise ratio A=2.

capacities of this AM scheme and the polyphase scheme are nearly the same.

Further, letting $P_{ep}^*(T,R)$ and $P_{ea}^*(T,R)$ be the smallest attainable error probability for a code with parameter T and rate R nats per sec. for the polyphase and AM systems, respectively, we can write

$$P_{ep}^* = \exp\left[-TE_p(R) + 0(T)\right]$$

$$P_{ea}^* = \exp\left[-TE_a(R) + 0(T)\right].$$

The exponent $E_p(R)$ is estimated in Appendix D and may be written

$$E_p(R) = WE(R/W),$$

where $E(\hat{R})$ is defined by (111). The exponent $E_a(R)$ is estimated in Refs. 10, 8, 9, 13. The exponents are compared in Fig. 8 for $A = 2$ by plotting their known lower bounds. It is also possible to show that $E_a(0) = E_p(0) = AW/2$ for all A.

REFERENCES

1. Viterbi, A. J., Systematic Coding for the Continuous Gaussian Channel, Ph.D. Dissertation, University of Southern California, August, 1962.
2. Balakrishnan, A. V., A Contribution to the Sphere-Packing Problem of Communication Theory, J. Math. Anal. Appl., *3*, 1961, pp. 485–506.
3. Blichfeldt, H. F., The Minimum Value of Quadratic Forms, and the Closest Packing of Spheres, Math Ann., *101*, 1929, pp. 605–608.
4. Watson, G. N., *The Theory of Bessel Functions*, Cambridge Univ. Press, London, 1958.
5. Erdélyi, A., et al., *Higher Transcendental Functions*, Vol. 2, McGraw-Hill Book Company, New York, 1953.
6. Grobner, W. and Hofreiter, N., *Integraltafel* (*Vol. II*), Springer-Verlag, Vienna, 1958.
7. Chernoff, H., A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations, Ann. Math. Statist., *23*, 1952, pp. 493–507.
8. Gallager, R. G., A Simple Derivation of the Coding Theorem and Some Applications, IEEE Trans. Inform. Theor., *IT-11*, January 1965, pp. 3–18.
9. Shannon, C. E., Gallager, R. G., and Berlekamp, E. R., forthcoming paper to be published in Inform. and Control.
10. Shannon, C. E., Probability of Error for Optimal Codes in the Gaussian Channel, B.S.T.J., *38*, May, 1959, pp. 611–656.
11. Nuttal, A. H., Error Probabilities for Equicorrelated M-ary Signals Under Phase-Coherent and Phase-Incoherent Reception, IEEE Trans. Inform Theor., *IT-8*, July, 1962, pp. 305–314.
12. Shannon, C. E., Communication in the Presence of Noise, Proc. IRE, *37*, Jan., 1949, pp. 10–21.
13. Wyner, A. D., An Inproved Error Bound for Gaussian Channels, B.S.T.J., *43*, Nov., 1964, pp. 3070–3075.