MA: MIT, 1963.
[9] P. R. Chevillat, "Fast sequential decoding and a new complete decoding algorithm," Ph.D. dissertation, Dept. Elec. Eng., Illinois Institute of Technology, Chicago, IL, May 1976.
[10] K. Zigangirov, "Some sequential decoding procedures," *Problemy Peredachi Informatsii*, vol. 2, pp. 13–25, 1966.
[11] F. Jelinek, "A fast sequential decoding algorithm using a stack," *IBM J. of Res. and Dev.*, vol. 13, pp. 675–685, Nov. 1969.
[12] J. Geist, "Algorithmic aspects of sequential decoding," Ph.D. dissertation, Dept. Elec. Eng., Univ. Notre Dame, Notre Dame, IN, Aug. 1970.
[13] W. W. Peterson and E. J. Weldon, *Error-Correcting Codes*, 2nd Ed. Cambridge, MA: MIT, 1972.
[14] P. R. Chevillat and D. J. Costello, "Distance and computation in sequential decoding," *IEEE Trans. Commun.*, vol. COM-24, pp. 440–447, Apr. 1976.
[15] R. Johannesson, "Robustly optimal rate one-half binary convolutional codes," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 464–468, July 1975.
[16] D. J. Costello, "Free distance bounds for convolutional codes," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 356–365, May 1974.
[17] J. M. Wozencraft, Private Communication, Oct. 1974.
[18] J. L. Massey and D. J. Costello, "Nonsystematic convolutional codes for sequential decoding in space applications," *IEEE Trans. Commun. Tech.*, vol. COM-19, pp. 806–813, Oct. 1971.
[19] L. R. Bahl and F. Jelinek, "Rate 1/2 convolutional codes with complementary generators," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 718–727, Nov. 1971.
[20] K. J. Larsen, "Short convolutional codes with maximal free distance for rates 1/2, 1/3, and 1/4," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 371–372, May 1973.
[21] E. Paaske, "Short binary convolutional codes with maximal free distance for rates 2/3 and 3/4," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 683–689, Sept. 1974.
[22] D. Haccoun and M. J. Ferguson, "Generalized stack algorithms for decoding convolutional codes," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 638–651, Nov. 1975.
[23] R. Johannesson, "On the computational problem with sequential decoding," presented at the IEEE International Sym. on Inform. Theory, Ronneby, Sweden, June 21–24, 1976.

The Gaussian Wire-Tap Channel

S. K. LEUNG-YAN-CHEONG, MEMBER, IEEE, AND MARTIN E. HELLMAN, MEMBER, IEEE

***Abstract*—Wyner's results for discrete memoryless wire-tap channels are extended to the Gaussian wire-tap channel. It is shown that the secrecy capacity C_s is the difference between the capacities of the main and wire-tap channels. It is further shown that $Rd = C_s$ is the upper boundary of the achievable rate-equivocation region.**

I. INTRODUCTION

IN A RECENT insightful paper [1] Wyner introduced the wire-tap channel shown in Fig. 1. It is a form of degraded broadcast channel [2], with the novel difference that one information rate is to be maximized and the other minimized. The object is to maximize the rate of *reliable* communication from the source to the legitimate receiver, subject to the constraint that the wire-tapper learns as little as possible about the source output. The wire-tapper knows the encoding scheme used at the transmitter and the decoding scheme used by the legitimate receiver, and is kept ignorant solely by the greater noise present in his received signal. Thus while the objective is the same as in cryptography, the technique used to achieve privacy is very different.

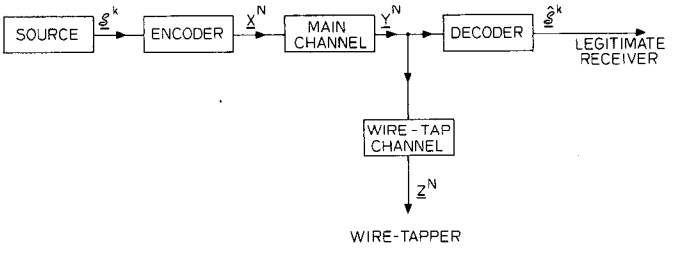

Fig. 1. General wire-tap channel.

The source is stationary and ergodic, and has a finite alphabet. The first k source outputs s^k are encoded into an N-vector x^N which is input to the main channel. The legitimate receiver makes an estimate $\hat{s}^k$ of s^k based on the output y^N of the main channel, incurring a block error rate

$$P_e = \Pr(s^k \neq \hat{s}^k). \tag{1}$$

y^N is also the input to the wire-tap channel and the wire-tapper has an average residual uncertainty $H(S^k|Z^N)$ after observing the output z^N of the wire-tap channel. Of course it does not change the problem if z^N is the output of a single channel with input x^N, which is statistically equivalent to the cascade of the main and wire-tap channels, since dependencies between z^N and y^N are im-

Manuscript received June 4, 1976; revised November 9, 1977. This work was supported in part by the National Science Foundation under Grant ENG-10173, in part by the United States Air Force Office of Scientific Research under Contract F44620-73-C-0065, and in part by the Joint Service Electronics Program under Contract N00014-75-C-0601.

S. K. Leung-Yan-Cheong was with the Department of Electrical Engineering, Stanford, CA. He is now with the Department of Electrical Engineering, Massachusetts Institute of Technology, Cambridge, MA.
M. E. Hellman is with the Department of Electrical Engineering, Stanford University, Stanford, CA.

0018-9448/78/0700-451$00.75

material. We define the fractional equivocation of the wire-tapper to be

$$\Delta = H(\mathcal{S}^k | Z^N)/H(\mathcal{S}^k) \tag{2}$$

and the rate of transmission to be

$$R = H(\mathcal{S}^k)/N. \tag{3}$$

We shall say that the pair (R^*, d^*) is achievable if for all $\epsilon > 0$ there exists an encoder-decoder pair such that

$$R \geqslant R^* - \epsilon, \quad \Delta \geqslant d^* - \epsilon, \text{ and } P_e \leqslant \epsilon. \tag{4}$$

Our definitions are slightly different from Wyner's original definitions. For example, Wyner defines $\Delta = H(\mathcal{S}|Z)/k$. (We will drop superscripts when the context permits.) The new definitions have the advantage that the achievable (R,d) region depends only on the channels and not on the source.

Wyner has determined the achievable (R,d) region when both channels are discrete memoryless channels. He shows that in most cases there is a secrecy capacity $C_s > 0$ such that $(R,d) = (C_s, 1)$ is achievable. By operating at rates below C_s, it is possible to ensure that the wire-tapper is essentially no better informed about s after observing z than he was before.

A particularly simple example results when both the main and wire-tap channels are binary symmetric channels (BSC) with crossover probabilities of 0 and p respectively, and the source is binary symmetric. (Then $H(\mathcal{S}^k) = k$, and our definition is equivalent to Wyner's.) Wyner shows that

$$R \leqslant 1 \tag{5}$$

$$d \leqslant 1 \tag{6}$$

$$Rd \leqslant h(p) \tag{7}$$

defines the set of achievable points. As noted by Wyner, this region is not convex. Surprisingly, however, $Rd = c$ as in (7) corresponds to a time-sharing curve as established in the following lemma.

Lemma 1: Let $R_1 d_1 = R_2 d_2 = c$, a constant. Assume $R_1 > R_2$ and hence $d_1 < d_2$. If the points (R_1, d_1) and (R_2, d_2) are achievable, then by time-sharing any point (R,d) with $R_2 \leqslant R \leqslant R_1$, $d_1 \leqslant d \leqslant d_2$, and $Rd = c$ is achievable.

Proof: Consider a block of N channel uses. Assume that for αN transmissions we operate at (R_1, d_1) and for $(1-\alpha)N$ transmissions we operate at (R_2, d_2). Then the effective equivocation is

$$d = \frac{\alpha N R_1 d_1 + (1-\alpha) N R_2 d_2}{\alpha N R_1 + (1-\alpha) N R_2}. \tag{8}$$

The effective transmission rate is

$$R = [\alpha N R_1 + (1-\alpha) N R_2]/N \tag{9}$$

so that

$$Rd = \alpha R_1 d_1 + (1-\alpha) R_2 d_2. \tag{10}$$

We see that time-sharing averages the product $R_i d_i$, so if $R_1 d_1 = R_2 d_2 = c$, then $Rd = c$. Q.E.D.

This lemma will be of importance in establishing the achievable (R,d) region for the Gaussian wire-tap channel, shown in Fig. 2. As before the source is a stationary ergodic finite alphabet source. The noise vectors $\boldsymbol{n}_1^N$ and $\boldsymbol{n}_2^N$ are independent and have components that are i.i.d. $\mathcal{N}(0, \sigma_1^2)$ and $\mathcal{N}(0, \sigma_2^2)$, respectively. The channel is power limited in that

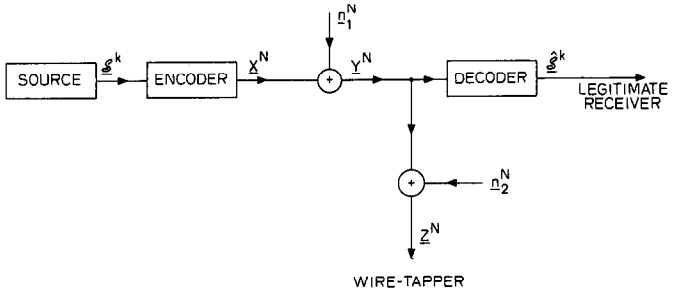

Fig. 2. Gaussian wire-tap channel.

$$(1/N) \sum_{i=1}^{N} E(X_i^2) \leqslant P. \tag{11}$$

In the following two sections we utilize Wyner's framework to completely characterize the achievable (R,d) region for this Gaussian wire-tap channel. Letting

$$C_M = 1/2 \log (1 + P/\sigma_1^2) \tag{12}$$

and

$$C_{MW} = 1/2 \log (1 + P/(\sigma_1^2 + \sigma_2^2)) \tag{13}$$

denote the capacities of the main and overall wire-tap channels, respectively, we shall show that the secrecy capacity is given by

$$C_s = C_M - C_{MW} \tag{14}$$

and that in general we have the following result.

Theorem 1: For the Gaussian wire-tap channel, the set $\mathcal{R}$ of all achievable (R,d) pairs is defined by

$$R \leqslant C_M \tag{15}$$

$$d \leqslant 1 \tag{16}$$

$$Rd \leqslant C_s. \tag{17}$$

In the next section, we establish the achievability of this region by showing that the extreme points of $\mathcal{R}$

$$(R_1, d_1) = (C_M, C_s/C_M) \tag{18}$$

and

$$(R_2, d_2) = (C_s, 1) \tag{19}$$

are both achievable. We then invoke the time sharing argument, and say that since $R_1 d_1 = R_2 d_2 = C_s$ the entire region $\mathcal{R}$ defined in Theorem 1 is achievable.

In Section III we establish the converse, that any point outside $\mathcal{R}$ is not achievable.

II. Direct Half of Theorem 1

As noted above we need only establish that the two points (R_1, d_1) and (R_2, d_2) defined by (18) and (19) are achievable, since time-sharing then implies the achievability of the entire region $\mathcal{R}$ of Theorem 1.

The point (R_1, d_1) is trivially achieved by coding as if the wire-tapper was absent. Usual source and channel coding arguments show that it is possible for R to be arbitrarily close to $C_M = R_1$ and P_e to be arbitrarily close to 0. But the information gained by the wire-tapper is limited by the capacity of his channel so that

$$\Delta = H(\mathcal{S}^k|\boldsymbol{Z}^N)/H(\mathcal{S}^k) \geq (H(\mathcal{S}^k) - NC_{MW})/H(\mathcal{S}^k) = 1 - (C_{MW}/R). \quad (20)$$

As R approaches C_M, this lower bound on Δ approaches $C_s/C_M = d_1$. Thus the point (R_1, d_1) is achievable.

We will establish the achievability of $(R_2, d_2) = (C_s, 1)$ by proving a somewhat stronger result, similar to that of Hellman and Carleial [3]. If $C_s = C_M/2$, Theorem 1 states that, by cutting our rate in half, we can completely foil the wire-tapper. Instead, we will show that it is possible to send two independent messages reliably, each at a rate near $C_s = C_M/2$, and each totally protected from the wire-tapper on an individual basis. The penalty is that, if the wire-tapper learns one message through other means, he can then also determine the other message. In general, if $C_s \geq C_M/L$, we will show that L independent messages can be simultaneously and reliably communicated to the legitimate receiver, each at a rate near C_M/L, and each totally protected on an individual basis. However, if the wire-tapper learns any one message, he may be able to determine all of the others. By using random noise for all but the first message, we can obtain the direct half of Theorem 1 as a special case of Theorem 2.

Theorem 2: Let $\mathfrak{u}^m$ be a sequence of m outputs from a finite-alphabet stationary ergodic source with per letter entropy $H(\mathcal{U})$, and let $\mathfrak{s}^p$ be any p consecutive components of $\mathfrak{u}^m$. Then provided

$$R_t \equiv H(\mathcal{U}^m)/N < C_M \quad (21)$$

$$R_s \equiv H(\mathcal{S}^p)/N < C_s, \quad (22)$$

it is possible, by choosing N large enough, to communicate $\mathfrak{u}^m$ to the legitimate receiver reliably in N channel uses and yet to ensure that

$$\Delta_s \equiv H(\mathcal{S}^p|\boldsymbol{Z}^N)/H(\mathcal{S}^p) \quad (23)$$

is arbitrarily close to 1.

Further, if $\{\mathfrak{s}_i^p\}_{i=1}^L$ are L such consecutive p-tuples of $\mathfrak{u}^m$, it is possible to ensure that

$$\Delta_{si} \equiv H(\mathcal{S}_i^p|\boldsymbol{Z}^N)/H(\mathcal{S}_i^p) \quad (24)$$

is arbitrarily close to 1 for $1 \leq i \leq L$, with L fixed as $N \to \infty$.

Remarks:

1) An alternative notation would be to use $\mathcal{S}^m$ in place of $\mathcal{U}^m$, but then superscripts would be necessary to distinguish between the entire ergodic source output and its projection, $\mathcal{S}^p$, to be kept secret. It is hoped that this remark will remove any confusion caused by our choice of notation.

2) Until now, the entire sequence $\mathfrak{u}^m$ was to be protected so that $m = p$, $R_s = R_t$, and $\Delta_s = \Delta_t$. Now the distinction between $\mathfrak{u}^m$ and its projection $\mathfrak{s}^p$ requires us to distinguish between the total rate R_t and the secrecy rate R_s.

3) For memoryless sources the p outputs in $\mathcal{S}^p$ need not be consecutive, and more general p-dimensional projections are possible. See [3] for a generalization to linear projections.

Theorem 2 will be established by proving a sequence of lemmas. First, since the source is ergodic, we have:

Lemma 2: It is possible to source code $\mathfrak{u}^m$ into a binary n-vector $\boldsymbol{u}^n$ such that each of the $\mathfrak{s}_i^p$ are determined reliably (i.e., as $N \to \infty$ the probability of error tends to 0) by k consecutive components $\boldsymbol{s}_i^k$ of $\boldsymbol{u}^n$ and

$$n = N(C_M - 2\epsilon) \quad (25)$$

$$k = N(C_s - \epsilon) \quad (26)$$

with $\epsilon > 0$.

Remarks: $\mathfrak{u}^m$ denotes the entire ergodic source output, and $\mathfrak{s}^p$ denotes a p-dimensional projection thereof; $\boldsymbol{u}^n$ denotes the binary source-coded version of $\mathfrak{u}^m$, and $\boldsymbol{s}^k$ denotes a k-dimensional projection thereof. Further $\boldsymbol{s}^k$ is a binary source-coded version of $\mathfrak{s}^p$.

Proof: From (21) and (22), we can define

$$\epsilon = \min\{(C_M - R_t)/3, (C_s - R_s)/2\} > 0. \quad (27)$$

In proving that R_s and R_t can be made to approach C_s and C_M, while Δ_s is kept arbitrarily close to 1, we can redefine R_s and R_t so that

$$(C_M - R_t)/3 = (C_s - R_s)/2 = \epsilon \quad (28)$$

where ϵ is given by (27) since excess rate can be discarded.

The noiseless source coding theorem for ergodic sources [4, theorem 3.5.3] then implies that (25) and (26) can be satisfied. There is a minor problem in ensuring that $\boldsymbol{s}^k$ consists of k consecutive bits of $\boldsymbol{u}^n$, but this is easily overcome.

If the $\{\mathfrak{s}_i^p\}$ are disjoint, we clearly can code in subblocks while satisfying (26) and the condition that $\boldsymbol{s}_i$ be consecutive bits of $\boldsymbol{u}$. Even if the $\{\mathfrak{s}_i^p\}$ are not disjoint, we can still satisfy these conditions. For example, if $\mathfrak{s}_1^p$ constitutes the first 3/4 of $\mathfrak{u}^m$ and $\mathfrak{s}_2^p$ constitutes the last 3/4 of $\mathfrak{u}^m$, we can code $\mathfrak{u}^m$ in four equal subblocks to obtain $\boldsymbol{u}^n$. The union bound guarantees that the overall coding from $\mathfrak{u}$ to $\boldsymbol{u}$ is reliable since each of the four subcodings is reliable. Q.E.D.

We will henceforth deal with only one of the $\boldsymbol{s}_i$ (or $\mathfrak{s}_i$) that we shall denote as $\boldsymbol{s}$ (or $\mathfrak{s}$). We shall show that, over a suitable ensemble of codes, $\boldsymbol{u}$ can be communicated reliably to the receiver and Δ_s kept arbitrarily near 1, with probability that approaches 1 as $N \to \infty$. Use of the union bound then allows us to state that, with probability approaching 1, all L of the Δ_{si} can be kept near 1. Now define an ensemble of channel codes as follows. Each code in the ensemble has 2^n codewords with blocklength N,

$$C = \{\boldsymbol{X}^1, \boldsymbol{X}^2, \cdots, \boldsymbol{X}^{2^n}\}. \quad (29)$$

Each component of each codeword is an i.i.d. random variable with a $\mathcal{N}(0, P-\alpha)$ distribution, where $\alpha>0$ is chosen so that

$$C_M(\alpha) \equiv 1/2 \log\left(1+(P-\alpha)/\sigma_1^2\right) > C_M - \epsilon \quad (30)$$

and

$$C_{MW}(\alpha) \equiv 1/2 \log\left(1+(P-\alpha)/\left(\sigma_1^2+\sigma_2^2\right)\right) > C_{MW} - \epsilon. \quad (31)$$

Since $n = N(C_M - 2\epsilon)$, the normal coding theorem for Gaussian channels [4, theorem 7.4.2] states that $\boldsymbol{u}^n$ is reliably transmitted to the receiver by almost all codes in the ensemble as $N\to\infty$. And as $N\to\infty$ almost all codes in the ensemble satisfy the power constraint (11), so almost all codes satisfy both conditions as $N\to\infty$.

All that remains is to show that $\Delta_s \doteq 1$ for almost all codes in the ensemble.

Lemma 3:

$$\Delta_s \geq [H(\boldsymbol{U}) - H(\boldsymbol{U}|\boldsymbol{S},\boldsymbol{Z}) - I(\boldsymbol{U};\boldsymbol{Z})]/NC_s. \quad (32)$$

Proof: Since $\boldsymbol{s}$ is a deterministic function of $\boldsymbol{s}$,

$$\Delta_s = H(\mathcal{S}|\boldsymbol{Z})/H(\mathcal{S}) \quad (33)$$

$$\geq H(\boldsymbol{S}|\boldsymbol{Z})/H(\mathcal{S}), \quad (34)$$

and from (22)

$$H(\mathcal{S}) \leq NC_s. \quad (35)$$

We complete the proof by showing that

$$H(\boldsymbol{S}|\boldsymbol{Z}) = H(\boldsymbol{U}) - H(\boldsymbol{U}|\boldsymbol{S},\boldsymbol{Z}) - I(\boldsymbol{U};\boldsymbol{Z}). \quad (36)$$

By definition

$$H(\boldsymbol{U}|\boldsymbol{Z}) = H(\boldsymbol{U}) - I(\boldsymbol{U};\boldsymbol{Z}) \quad (37)$$

and, since $\boldsymbol{s}$ is a function of $\boldsymbol{u}$,

$$H(\boldsymbol{U}|\boldsymbol{Z}) = H(\boldsymbol{U},\boldsymbol{S}|\boldsymbol{Z}) = H(\boldsymbol{S}|\boldsymbol{Z}) + H(\boldsymbol{U}|\boldsymbol{S},\boldsymbol{Z}). \quad (38)$$

Q.E.D.

We now proceed to bound the three terms in (32).

Lemma 4: There exists a sequence of source codes of increasing blocklength such that

$$H(\boldsymbol{U}) \geq NC_M(1-\epsilon'-\delta) \quad (39)$$

where ϵ' stands for any term which tends to 0 as $\epsilon\to 0$, and δ stands for any term which tends to 0 as $N\to\infty$ with $\epsilon>0$ fixed.

Proof: From (21), (27), and (28)

$$H(\mathcal{U}) = NR_t = N(C_M - 3\epsilon) = NC_M(1-\epsilon'). \quad (40)$$

Since $\boldsymbol{u}$ is a deterministic function of u,

$$H(\boldsymbol{U}) = H(\mathcal{U}) - H(\mathcal{U}|\boldsymbol{U}). \quad (41)$$

Using the noiseless source coding theorem for ergodic sources [4, theorem 3.5.3] and Fano's inequality [4, theorem 4.3.1], we get

$$H(\mathcal{U}|\boldsymbol{U}) \leq 1 + \delta N, \quad (42)$$

so that

$$H(\boldsymbol{U}) \geq NC_M(1-\epsilon') - 1 - \delta N = NC_M(1-\epsilon'-\delta). \quad (43)$$

(Note that the two δ's are not equal.) Q.E.D.

We now bound the second term in (32).

Lemma 5: As $N\to\infty$ almost all codes in the ensemble obey

$$H(\boldsymbol{U}|\boldsymbol{S},\boldsymbol{Z}) \leq \delta N. \quad (44)$$

Proof: Since $\boldsymbol{s}$ is a k-dimensional projection of the n-vector $\boldsymbol{u}$, given $\boldsymbol{s}$ there are only $2^{n-k} = 2^{N(C_{MW}-\epsilon)}$ $\boldsymbol{u}$'s to be decided among on the basis of $\boldsymbol{z}$. But the codewords associated with each of these $\boldsymbol{u}$'s were chosen according to the capacity-achieving distribution, and from (31) we know the error probability given $\boldsymbol{s}$ and $\boldsymbol{z}$ tends to 0 as $N\to\infty$ for almost all codes. Use of Fano's inequality completes the proof.

Finally, the data processing theorem and the definition of C_{MW} yield a bound on the third term in (32):

$$I(\boldsymbol{U};\boldsymbol{Z}) \leq NC_{MW}. \quad (45)$$

Combining (45) and the preceding three lemmas we find that for almost all codes

$$\Delta_s \geq [NC_M(1-\epsilon'-\delta) - \delta N - NC_{MW}]/NC_s$$
$$= NC_s(1-\epsilon'-\delta)/NC_s$$
$$= 1-\epsilon'-\delta. \quad (46)$$

Then letting $N\to\infty$ with fixed $\epsilon>0$ we find that

$$\lim_{N\to\infty} \Delta_s \geq 1-\epsilon' \quad (47)$$

and

$$\lim_{\epsilon\to 0} \lim_{N\to\infty} \Delta_s = 1 \quad (48)$$

for almost all codes.

This completes the proof that $(R_2, d_2) = (C_s, 1)$ is achievable. An intuitive partial interpretation of the proof is as follows. Suppose the wire-tapper could determine $\boldsymbol{s}$ from $\boldsymbol{z}$. The residual rate of $\boldsymbol{u}$ is then below the capacity C_{MW} of his channel, and the code is designed so that then the wire-tapper could reliably learn the rest of $\boldsymbol{u}$. But then the wire-tapper would be gaining information at an overall rate above his channel's capacity, which is impossible. Therefore, the initial assumption (that the wire-tapper could determine $\boldsymbol{s}$) is wrong.

III. Converse Theorem

In this section we prove the converse part of Theorem 1, that any point (R,d) outside $\mathcal{R}$ is not achievable. That $R \leq C_M$ and $d \leq 1$ is self-evident from the definitions. Our real task is to show that

$$Rd \leq C_s \quad (17)$$

must hold if P_e is arbitrarily close to 0. (Note that in this section we are dealing solely with $\boldsymbol{s}$, and not at all with the $\boldsymbol{u}$ of the last section. We can therefore use R in place of R_s and Δ in place of Δ_s without ambiguity. The formulation of the last section led to a stronger forward theorem, but would yield a weaker converse if used here.) The proof of the following theorem is therefore the goal of this section.

Theorem 3: With R, Δ, and P_e defined as in (1), (2), and (3)

$$R\left[\Delta-\frac{kP_e\log(v)+h(P_e)}{H(\mathcal{S}^k)}\right]\leqslant C_s \tag{49}$$

where v is the size of the source alphabet and C_s is defined by (14).

If instead the per digit error rate

$$p_e\equiv 1/k\sum_{i=1}^{k}\Pr(s_i\neq\hat{s}_i) \tag{50}$$

is used, (49) becomes

$$R\left[\Delta-\frac{k[h(p_e)+p_e\log(v-1)]}{H(\mathcal{S}^k)}\right]\leqslant C_s. \tag{51}$$

Thus the use of this more lenient measure of reliability would not expand the region $\mathcal{R}$.

The proof of this theorem will be established through a sequence of lemmas.

Lemma 6:

$$R\left[\Delta-\frac{kP_e\log(v)+h(P_e)}{RN}\right]\leqslant\frac{I(\boldsymbol{X}^N;\boldsymbol{Y}^N|\boldsymbol{Z}^N)}{N} \tag{52}$$

and

$$R\left[\Delta-\frac{k[h(p_e)+p_e\log(v-1)]}{RN}\right]\leqslant\frac{I(\boldsymbol{X};\boldsymbol{Y}|\boldsymbol{Z})}{N}. \tag{53}$$

Proof: First note that, through use of the data processing theorem [4, theorem 4.3.3] and Fano's inequality [4, theorem 4.3.1],

$$H(\mathcal{S}|\boldsymbol{Z},\boldsymbol{Y})\leqslant H(\mathcal{S}|\boldsymbol{Y})\leqslant H(\mathcal{S}|\hat{\mathcal{S}})\leqslant h(P_e)+kP_e\log(v). \tag{54}$$

Then $RN\Delta=H(\mathcal{S}|\boldsymbol{Z})$ from the definitions (2), (3) of R and Δ. Using (54), we obtain

$$RN\Delta\leqslant H(\mathcal{S}|\boldsymbol{Z})-H(\mathcal{S}|\boldsymbol{Z},\boldsymbol{Y})+h(P_e)+kP_e\log(v)=I(\mathcal{S};\boldsymbol{Y}|\boldsymbol{Z})+h(P_e)+kP_e\log(v). \tag{55}$$

Since $\mathcal{S}$, $\boldsymbol{X}$, $\boldsymbol{Y}$, $\boldsymbol{Z}$ form a Markov chain, the data processing theorem implies

$$I(\mathcal{S};\boldsymbol{Y}|\boldsymbol{Z})\leqslant I(\boldsymbol{X};\boldsymbol{Y}|\boldsymbol{Z}), \tag{56}$$

so

$$RN\Delta\leqslant I(\boldsymbol{X};\boldsymbol{Y}|\boldsymbol{Z})+h(P_e)+kP_e\log(v) \tag{57}$$

which with minor algebra establishes (52). Equation (53) is established in exactly the same manner using the per digit error rate version of Fano's inequality [4, theorem 4.3.2].

Lemma 7:

$$I(\boldsymbol{X};\boldsymbol{Y}|\boldsymbol{Z})=\frac{N}{2}\log\left(\frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2}\right)-[H(\boldsymbol{Z})-H(\boldsymbol{Y})]. \tag{58}$$

Proof: Although the entropy of a continuous random variable is lacking in physical significance, if we define

$$H(A)=-\int p(a)\log[p(a)]\,da \tag{59}$$

it is known that differences in entropy are still physically meaningful as mutual informations: e.g., $H(A)-H(A|B)=I(A;B)$ (see [5] for a full development). We may thus write

$$I(\boldsymbol{X};\boldsymbol{Y}|\boldsymbol{Z})=H(\boldsymbol{X}|\boldsymbol{Z})-H(\boldsymbol{X}|\boldsymbol{Y},\boldsymbol{Z})=H(\boldsymbol{X}|\boldsymbol{Z})-H(\boldsymbol{X}|\boldsymbol{Y}) \tag{60}$$

since $\boldsymbol{X}$ is conditionally independent of $\boldsymbol{Z}$ given $\boldsymbol{Y}$. Using

$$H(A,B)=H(A)+H(B|A)=H(B)+H(A|B), \tag{61}$$

we can recast (60) as

$$\begin{aligned}I(\boldsymbol{X};\boldsymbol{Y}|\boldsymbol{Z})&=[H(\boldsymbol{X})+H(\boldsymbol{Z}|\boldsymbol{X})-H(\boldsymbol{Z})]-[H(\boldsymbol{X})+H(\boldsymbol{Y}|\boldsymbol{X})-H(\boldsymbol{Y})]\\&=H(\boldsymbol{Z}|\boldsymbol{X})-H(\boldsymbol{Y}|\boldsymbol{X})-[H(\boldsymbol{Z})-H(\boldsymbol{Y})].\end{aligned} \tag{62}$$

Because the channel is memoryless,

$$H(\boldsymbol{Y}|\boldsymbol{X})=\sum_{i=1}^{N}H(Y_i|X_i)=(N/2)\log(2\pi e\sigma_1^2) \tag{63}$$

where the last expression comes from integration as in (59) [4, p. 32]. Similarly

$$H(\boldsymbol{Z}|\boldsymbol{X})=(N/2)\log[2\pi e(\sigma_1^2+\sigma_2^2)]. \tag{64}$$

Substituting (63) and (64) into (62) yields (58). Q.E.D.

Lemma 8: Define

$$g(P)=1/2\log(2\pi eP),\qquad P>0, \tag{65}$$

$$g^{-1}(\alpha)=(1/2\pi e)e^{2\alpha}, \tag{66}$$

$$A(v)=g[\sigma_2^2+g^{-1}(v)]-v. \tag{67}$$

Then $A(v)$ is decreasing in v.

Proof:

$$A(v)=1/2\log\left[2\pi e\left(\sigma_2^2+\frac{1}{2\pi e}e^{2v}\right)\right]-v. \tag{68}$$

Differentiating (68) yields

$$\frac{dA}{dv}=[e^{2v}/(2\pi e\sigma_2^2+e^{2v})]-1\leqslant 0. \tag{69}$$

Lemma 9:

$$H(\boldsymbol{Y})\leqslant Ng(P+\sigma_1^2)=(N/2)\log[2\pi e(P+\sigma_1^2)]. \tag{70}$$

Proof: We know that

$$I(\boldsymbol{X};\boldsymbol{Y})=H(\boldsymbol{Y})-H(\boldsymbol{Y}|\boldsymbol{X})\leqslant NC_M \tag{71}$$

or

$$H(\boldsymbol{Y})\leqslant(N/2)\log[(P+\sigma_1^2)/\sigma_1^2]+(N/2)\log(2\pi e\sigma_1^2)=(N/2)\log[2\pi e(P+\sigma_1^2)]. \tag{72}$$

Lemma 10: If

$$H(\boldsymbol{Y})=Nv, \tag{73}$$

then

$$H(\boldsymbol{Z})-H(\boldsymbol{Y})\geqslant NA(v)=Ng[\sigma_2^2+g^{-1}(v)]-Nv. \tag{74}$$

Proof: See Shannon [6, theorem 15], Blachman [7], and Bergmans [8].

Combining Lemmas 8; 9, and 10 we see that

$$H(Z)-H(Y)\geq NA\left[g\left(P+\sigma_1^2\right)\right]$$
$$=Ng\left[\sigma_2^2+g^{-1}g\left(P+\sigma_1^2\right)\right]-Ng\left(P+\sigma_1^2\right)$$
$$=(N/2)\log\left(\frac{P+\sigma_1^2+\sigma_2^2}{P+\sigma_1^2}\right). \quad (75)$$

Using (75) with Lemma 7 yields

$$I(X;Y|Z)\leq\frac{N}{2}\log\left(\frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2}\right)-\frac{N}{2}\log\left(\frac{P+\sigma_1^2+\sigma_2^2}{P+\sigma_1^2}\right)$$
$$=N(C_M-C_{MW})=NC_s, \quad (76)$$

which together with Lemma 6 completes the proof of Theorem 3.

IV. Discussion

It is interesting that the secrecy capacity $C_s=C_M-C_{MW}$ completely characterizes the achievable (R,d) region of a Gaussian wire-tap channel, just as in the case of binary symmetric channels. Motivated by this observation, Leung [9] has shown that this is true whenever both the main channel and the cascade of the main and wire-tap channels are symmetric [4, p. 94]. (Strictly speaking, Leung only shows this for discrete memoryless channels.) Wyner's results [1], although derived for discrete memoryless channels, can also be combined with Lemmas 8, 9, and 10 to yield Theorem 1.

In the power limited region, when $P\ll\sigma^2$,

$$C_M\doteq P/\left(\sigma_1^2 2\ln 2\right), \quad (77)$$

$$C_{MW}\doteq P/\left[\left(\sigma_1^2+\sigma_2^2\right)2\ln 2\right], \quad (78)$$

and

$$C_s/C_M\doteq\sigma_2^2/\left(\sigma_1^2+\sigma_2^2\right). \quad (79)$$

In the bandwidth limited region, when $P\gg\sigma^2$,

$$C_M\doteq 1/2\log\left(P/\sigma_1^2\right), \quad (80)$$

$$C_{MW}\doteq 1/2\log\left[P/\left(\sigma_1^2+\sigma_2^2\right)\right], \quad (81)$$

so that

$$C_s\doteq 1/2\log\left[\left(\sigma_1^2+\sigma_2^2\right)/\sigma_1^2\right] \quad (82)$$

and

$$C_s/C_M\doteq=0. \quad (83)$$

For a fixed bandwidth C_s is increasing in P, but there is a finite limit on C_s no matter how large we make P. Our results are therefore of most use on power limited channels. Of course, if the main channel is bandwidth limited $(P/\sigma_1^2\gg 1)$ and the wire-tap channel is power limited $(P/(\sigma_1^2+\sigma_2^2)\ll 1)$, then C_s/C_M will be even closer to 1. It is really only the wire-tap channel that must be power limited.

There is a potential problem if the SNR's on the channels are somewhat uncertain. Then the system may operate several dB below the actual capacity of the main channel, but if the wire-tap channel's SNR is only several dB below the main channel's nominal SNR, then secrecy is lost. In spite of this problem there may be practical applications for these results. If, for example, the wire-tapper is listening to unintentional electromagnetic radiation from a terminal or computer, his SNR may be tens of dB down from that of the "main channel." Such a wire-tap channel allows almost no reduction in rate of information flow to be coupled with high uncertainty on the part of the wire-tapper.

We are currently investigating the wire-tap channel with feedback and have shown that even when the main channel is inferior to the wire-tapper's channel it is possible to transmit reliably and securely [10]. This would obviously eliminate many of the above problems.

Acknowledgment

The authors wish to thank Prof. Thomas Cover and Dr. Aaron Wyner for several valuable discussions on this problem.

References

[1] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, pp. 1355–1387, Oct. 1975.
[2] P. P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 197–207, Mar. 1973.
[3] M. E. Hellman and A. B. Carleial, "A note on Wyner's wire-tap channel," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 387–390, May 1977.
[4] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.
[5] R. Ash, *Information Theory*. New York: Interscience, 1965.
[6] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 623–656, 1948.
[7] N. M. Blachman, "The convolution inequality for entropy powers," *IEEE Trans. Inform. Theory*, vol. IT-11, pp. 267–271, Apr. 1965.
[8] P. P. Bergmans, "A simple converse for broadcast channels with additive white Gaussian noise," *IEEE Trans. Inform Theory*, vol. IT-20, pp. 279–280, Mar. 1974.
[9] S. K. Leung-Yan-Cheong, "On a special class of wire-tap channels," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 625–627, Sept. 1977.
[10] S. K. Leung-Yan-Cheong, *Multi-User and Wire-tap Channels Including Feedback*, Ph.D. thesis, Stanford University, Stanford, CA, 1976.