FILE COPY

(4)

AD-A207 388

Capacity of Discrete Time Gaussian Channel With and Without Feedback, II*

Dedicated to Professor Shōzo Koshi on his 60th Birthday

Shunsuke Ihara**
Department of Mathematics
Kochi University
Kochi 780
Japan

Kenjiro Yanagi
Department of Mathematics
Yamaguchi University
Yamaguchi 753
Japan

LISS-26

May, 1988

DTIC
ELECTE
APR 28 1989
S H D

*To appear, *Japan J. of Applied Mathematics.*

**Partially supported by University of North Carolina (ONR Contract N00014-84-C-0212).

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

89 4 28 145

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for Public Release: Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Department of Statistics 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code) University of North Carolina Chapel Hill, North Carolina 27514	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Office of Naval Research 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-84-C-0212
8c. ADDRESS (City, State and ZIP Code) Statistics & Probability Program Arlington, VA 22217	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO.: NR PROJECT NO.: 042 TASK NO.: 269 WORK UNIT NO.: SRO 105
11. TITLE (Include Security Classification) Capacity of Discrete Time Gaussian Channel ..	
12. PERSONAL AUTHOR(S) S. Ihara and K. Yanagi	
13a. TYPE OF REPORT TECHNICAL 13b. TIME COVERED FROM _____ TO _____	14. DATE OF REPORT (Yr., Mo., Day) May 1988 15. PAGE COUNT 26
16. SUPPLEMENTARY NOTATION	

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	Gaussian channel with feedback, information capacity

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The problem considered in this paper is whether the capacity of a discrete time Gaussian channel is increased by feedback or not. It is well known that the capacity of a white Gaussian channel under the average power constraint is not changed by feedback. We give some conditions under which the capacity of a discrete time Gaussian channel is increased by feedback. It is also shown that there exists a non-white Gaussian channel whose capacity is not increased by feedback.

11. TITLE CONT.: With and Without Feedback; II

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED XX SAME AS RPT. XX DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION	
22a. NAME OF RESPONSIBLE INDIVIDUAL C.R. Baker	22b. TELEPHONE NUMBER (Include Area Code) (919) 962-2189	22c. OFFICE SYMBOL

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

Abstract —— The problem considered in this paper is whether the capacity of a discrete time Gaussian channel is increased by feedback or not. It is well known that the capacity of a white Gaussian channel under the average power constraint is not changed by feedback. We give some conditions under which the capacity of a discrete time Gaussian channel is increased by feedback. It is also shown that there exists a non-white Gaussian channel whose capacity is not increased by feedback.

Key words: Gaussian channel with feedback, information capacity

Accession For
NTIS GRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

§ 1. Introduction

The problem considered in this paper is whether the capacity of a discrete time Gaussian channel (GC) is increased by feedback or not.

It is known that the capacity of a white GC is not increased by feedback (see [6] for continuous time case and Theorem 2 of this paper for discrete time case). On the other hand, some examples of non-white GC's have been presented to show that the capacity can be increased by feedback (see [4] for continuous time case and [3], [5] for discrete time case).

The main aim of this paper is to give conditions under which the capacity of a discrete time GC is increased by feedback. The GC to be considered is presented by

$$Y_n = X_n + Z_n, \quad n \in \mathbb{N} \equiv \{1,\cdots,N\}, \tag{1}$$

where X_n, Y_n and Z_n represent the channel input, the output and the noise, respectively, at time n, and N is the terminal time. The noise $Z = \{Z_n;\ n \in \mathbb{N}\}$ is a Gaussian process such that Z_n, $n \in \mathbb{N}$, are linearly independent. We assume that an average power constraint

$$\sum_{n=1}^{N} E[X_n^2] \le \rho^2, \tag{2}$$

is imposed on the input $X = \{X_n\}$, where $\rho > 0$ is a given constant. Let θ be a message, to be transmitted, which is a random variable or a stochastic process independent of the noise Z. Denote by $\mathcal{F}(\theta)$ and $\mathcal{F}_m(Y)$ the σ-fields generated by θ and $\{Y_n;\ n \le m\}$,

respectively. The GC (1) is *with feedback* if X_n is a function of the message and the channel output up to $n - 1$, namely

(a.1) X_n is $\mathcal{F}(\theta) \vee \mathcal{F}_{n-1}(Y)$ measurable.

The GC is *without feedback* if X_n is a function of the message, namely

(a.2) X_n is $\mathcal{F}(\theta)$ measurable.

In each case we denote by $\mathcal{A}_F(\rho^2)$ and $\mathcal{A}_0(\rho^2)$ the classes of all admissible pairs of a message and an input:

$$\mathcal{A}_F(\rho^2) = \{(\theta,X);\ \theta \text{ and } X \text{ satisfy (2) and (a.1)}\};$$

$$\mathcal{A}_0(\rho^2) = \{(\theta,X);\ \theta \text{ and } X \text{ satisfy (2) and (a.2)}\}.$$

We denote by $I(\theta,Y)$ the mutual information between the message θ and the output $Y = \{Y_n;\ n \in \mathbb{N}\}$. Under the constraint (2), the capacities $C_F(\rho^2)$ and $C_0(\rho^2)$ of the GC (1) with feedback and without feedback, respectively, are defined by

$$C_F(\rho^2) = \sup\ \{I(\theta,Y);\ (\theta,X) \in \mathcal{A}_F(\rho^2)\}$$

and

$$C_0(\rho^2) = \sup\ \{I(\theta,Y);\ (\theta,X) \in \mathcal{A}_0(\rho^2)\}.$$

Note that if the GC is without feedback we can identify the message with the input and that

$$C_0(\rho^2) = \sup\ \{I(X,Y);\ X \in \mathcal{B}_0(\rho^2)\},$$

where $\mathcal{B}_0(\rho^2) = \{X;\ X \text{ satisfies (2) and is independent of } Z\}$.

If the GC is without feedback, a formula for the capacity

$C_0(\rho^2)$ is known (see Theorem 1). If $Z = \{Z_n\}$ is a white noise, namely if Z_n, $n \in \mathbb{N}$, are mutually independent, the GC is called a *white* Gaussian channel (WGC). Since it is known that the capacity of the WGC is not increased by feedback, we are concerned with the case of the non-white GC.

Let $\Sigma = \left(\sigma_{mn} \right)_{m,n\in\mathbb{N}}$ be the covariance matrix of Z. We say that the Gaussian process $Z = \{Z_n;\ n \in \mathbb{N}\}$ is *blockwise white*, if there exist $m_k \times m_k$ matrices $\Sigma^{(k)}$, $k = 1,\cdots,K$ $(1 < K < N$ and $\sum_{k=1}^{K} m_k = N)$ such that

$$\Sigma \sim \begin{pmatrix} \Sigma^{(1)} & & & 0 \\ & \Sigma^{(2)} & & \\ & & \ddots & \\ 0 & & & \Sigma^{(K)} \end{pmatrix} \tag{3}$$

The relation $\sim$ is defined as follows. For $N\times N$ matrices $A = \left(a_{mn} \right)_{m,n\in\mathbb{N}}$ and $B = \left(b_{mn} \right)_{m,n\in\mathbb{N}}$, we denote $A \sim B$ if there exists a permutation π on $\mathbb{N} = \{1,\cdots,N\}$ such that $b_{mn} = a_{\pi(m)\pi(n)}$, $m, n \in \mathbb{N}$. We note that if Z is blockwise white then $Z^{(k)} = \{Z_{\pi(n)};\ N_{k-1} < n \le N_k\}$, $k = 1,\cdots,K$, are mutually independent, where $N_k = m_1 + \cdots + m_k$. The GC is said to be *blockwise white* if the noise Z is blockwise white. If Z is not blockwise white, Z is called a *completely non-white noise* and the corresponding GC is also called a *completely non-white* GC.

It will be shown that the capacity of the GC is increased by feedback if the GC is completely non-white or the GC is blockwise white but not white and the power ρ^2 is greater than a constant ρ_0^2 specified later (Theorem 4). It is conjectured that if the GC is blockwise white and $\rho^2 < \rho_0^2$ the capacity of the GC with feedback is

the same as the capacity of the same channel without feedback. We can give an example of a blockwise white GC whose capacity is not increased by feedback (Theorem 5).

It is known that the capacity of the GC with feedback is never more than twice the capacity of GC without feedback [7]. We have given in Part I [5] an example of a GC whose capacity is almost doubled by feedback.

§ 2. Preliminaries

This section is devoted to summarizing previously known results on the capacity of GC, which will be used later.

The mutual information $I(\xi,\eta)$ between random variables ξ and η is defined by

$$I(\xi,\eta) = \int \log \frac{d\mu_{\xi\eta}}{d\mu_\xi \times \mu_\eta}\, d\mu_{\xi\eta},$$

if the joint probability distribution $\mu_{\xi\eta}$ of ξ and η is absolutely continuous with respect to the product measure $\mu_\xi \times \mu_\eta$ of the probability distributions μ_ξ and μ_η of ξ and η, respectively, where $d\xi_{\xi\eta}/d\mu_\xi \times \mu_\eta$ is the Radon-Nikodym derivative; otherwise $I(\xi,\eta)$ is infinite. The differential entropy $h(\xi)$ of an m-dimensional random variable ξ with a continuous distribution is defined by

$$h(\xi) = -\int_{R^m} p_\xi(x) \log p_\xi(x)\, dx,$$

if the integral exists, where $p_\xi(x)$ is the probability density function of ξ. If $h(\xi,\eta)$ $(= h((\xi,\eta)))$ and $h(\eta)$ exist, the quantity $h(\xi|\eta)$ defined by

$$h(\xi|\eta) = h(\xi,\eta) - h(\eta)$$

is called the conditional entropy of ξ given η.

The following properties of the mutual information and the differential entropy are used later.

LEMMA 1. *Let* θ, $X = \{X_n\}$ *and* $Y = \{Y_n\}$ *be a message, an input and the corresponding output of the GC* (1), *respectively.*

(i)
$$I(\theta,Y) = \sum_{n=1}^{N} \{h(Y_n|Y_1,\cdots,Y_{n-1}) - h(Y_n|\theta,Y_1,\cdots,Y_{n-1})\}$$
$$= \sum_{n=1}^{N} \{h(Y_n|Y_1,\cdots,Y_{n-1}) - h(Z_n|Z_1,\cdots,Z_{n-1})\}, \qquad (4)$$

where $h(Y_n|Y_1,\cdots,Y_{n-1}) = h(Y_1)$ *if* $n = 1$.

(ii) *Denote by* τ_n^2 *the variance of the conditional distribution of* Y_n *given* $(Y_1,\cdots,Y_{n-1})$. *Then*

$$h(Y_n|Y_1,\cdots,Y_{n-1}) \le \frac{1}{2} \log (2\pi e \tau_n^2). \qquad (5)$$

The equality holds in (5) *if and only if the conditional distribution is Gaussian.*

(iii) *If* $(X,Z) = \{X_n,Z_n;\ n \in \mathbb{N}\}$ *is Gaussian, then*

$$I(\theta, Y) = \frac{1}{2} \{\log |\Gamma_Y| - \log |\Sigma|\},$$

where Γ_Y *is the covariance matrix of* $Y = \{Y_n\}$ *and* $|\cdot|$ *denotes the determinant.*

The capacity $C_0(\rho^2)$ of the GC (1) without feedback is known to be determined by the eigenvalues $\{\lambda_n^2;\ n \in \mathbb{N}\}$ of the covariance matrix Σ of the noise $Z = \{Z_n\}$.

THEOREM 1 (see, e.g. [1]). *The capacity* $C_0(\rho^2)$ *of the GC* (1) *without feedback subject to* (2) *is given by*

$$C_0(\rho^2) = \frac{1}{2} \sum_{n=1}^{N} \log \max (1,\ a^2\lambda_n^{-2}),$$

where $a \equiv a(\rho) > 0$ *is a constant uniquely determined by*

$$\sum_{n=1}^{N} \max (0,\ a^2 - \lambda_n^2) = \rho^2. \tag{6}$$

There exists an orthogonal matrix $P = \left(p_{mn} \right)_{m,n\in\mathbb{N}}$ such that ${}^tP\Sigma P = \Lambda$, where $\Lambda = \begin{pmatrix} \lambda_1^2 & & 0 \\ & \ddots & \\ 0 & & \lambda_N^2 \end{pmatrix}$ and ${}^tP$ $(= P^{-1})$ is the transposed matrix of P. A Gaussian process $\zeta = \{\zeta_n;\ n \in \mathbb{N}\}$ defined by

$$\zeta = ZP \quad (\zeta_n = \sum_{m=1}^{N} p_{mn}Z_m, \quad n \in \mathbb{N}) \tag{7}$$

is a white noise. Indeed, the covariance matrix of ζ is the

diagonal matrix Λ. Define constants $\tau_n \equiv \tau_n(\rho) \geq 0$, $n \in \mathbb{N}$, by

$$\tau_n^2 = \max\,(0,\ a^2 - \lambda_n^2). \tag{8}$$

Let $\xi = \{\xi_n;\ n \in \mathbb{N}\}$ be a zero mean Gaussian process independent of ζ, such that $E[\xi_m \xi_n] = \delta_{mn}\tau_n^2$, and $\eta = \{\eta_n;\ n \in \mathbb{N}\}$ be defined by

$$\eta_n = \xi_n + \zeta_n, \quad n \in \mathbb{N}.$$

We define Gaussian processes $X^0 = \{X_n^0;\ n \in \mathbb{N}\}$ and $Y^0 = \{Y_n^0;\ n \in \mathbb{N}\}$ by

$$X^0 = \xi^t P \quad (X_n^0 = \sum_{m=1}^{N} p_{nm}\xi_m) \tag{9}$$

and

$$Y_n^0 = X_n^0 + Z_n, \quad n \in \mathbb{N}.$$

Then X^0 is independent of Z and $\sum_{n=1}^{N} E[(X_n^0)^2] = \sum_{n=1}^{N} E[\xi_n^2] = \rho^2$. In the case of without feedback, it is known that X^0 is the optimal input signal in the sense that the capacity is attained by transmitting X^0:

$$C_0(\rho^2) = \frac{1}{2}\sum_{n=1}^{N} \log \max\,(1, a^2\lambda_n^{-2}) = I(\xi,\eta) = I(X^0,Y^0). \tag{10}$$

We put $\gamma_{mn} \equiv \gamma_{mn}(\rho) = E[X_m^0 X_n^0]$ and denote by $\Gamma = \left(\gamma_{mn}\right)_{m,n\in\mathbb{N}}$ the covariance matrix of $X^0 = \{X_n^0\}_{n\in\mathbb{N}}$. Since $a = a(\rho)$ and $\tau_n = \tau_n(\rho)$ are monotone non-decreasing in ρ, $\gamma_{mn} = \sum_{m=1}^{N} p_{mn}^2\tau_m^2$ is also non-decreasing in ρ. We assume, for a moment, that the noise Z is blockwise white but not white. Then there exists a permutation π on $\mathbb{N}$ such that $Z^{(k)} = \{Z_{\pi(n)};\ N_{k-1} < n \leq N_k\}$, $k = 1,\cdots,K$, $(0 = N_0$

$< N_1 < \cdots < N_K = N$ and $1 < K < N$) are mutually independent. Here we may assume that

$$\begin{cases} m_k = 1, & k \le L, \\ m_k \ge 2 \text{ and } Z^{(k)} \text{ is completely non-white}, & k > L, \end{cases} \tag{11}$$

where $m_k = N_k - N_{k-1}$ and $0 \le L < N$. In this case we can easily show that

$$p_{\pi(m)\pi(n)} = 0 \quad \text{if} \quad N_{k-1} < m \le N_k,\ N_{\ell-1} < n \le N_\ell \ (k \ne \ell)$$

and $\left(p_{\pi(m)\pi(n)} \right)_{N_{k-1}<m\le N_k}$ is an orthogonal matrix. Hence we have

$$\sum_{m=N_{k-1}+1}^{N_k} \gamma_{\pi(m)\pi(m)} = \sum_{m,n=N_{k-1}+1}^{N_k} p^2_{\pi(m)\pi(n)} \tau^2_{\pi(n)}$$

$$= \sum_{n=N_{k-1}+1}^{N_k} \tau^2_{\pi(n)}. \tag{12}$$

We define a constant $\rho_0 \ge 0$ by

$$\rho_0 = \inf \{\rho;\ a^2(\rho) \ge \min_{n>N_L} \lambda^2_{\pi(n)}\}. \tag{13}$$

From (8) and (12) we see that $\rho > \rho_0$ if and only if

$$\sum_{n=N_L+1}^{N} \gamma_{\pi(n)\pi(n)}(\rho) > 0.$$

Thus we have obtained the following

LEMMA 2. *Let* $Z = \{Z_n\}$ *be a non-white and blockwise white noise satisfying* (11). *Then, if* $\rho > \rho_0$, *there exists an integer*

k *such that* $m_k = N_k - N_{k-1} \geq 2$ *and*

$$\sum_{n=N_{k-1}+1}^{N_k} \gamma_{\pi(n)\pi(n)}(\rho) > 0 \tag{14}$$

When the GC (1) is white, the following theorem is known.

THEOREM 2. *If the GC* (1) *is white, then the capacity under the constraint* (2) *is not increased by feedback*:

$$C_F(\rho^2) = C_0(\rho^2) = \frac{1}{2} \sum_{n=1}^{N} \log \max (1, a^2\lambda_n^{-2}),$$

where $\lambda_n^2 = E[Z_n^2]$ *and* a *is given by* (6).

For the completeness we shall give the proof in Appendix.

We do not have explicit formulae for the capacity with feedback except for the case of WGC. However it has been known that the capacity is attained in a Gaussian scheme. We consider the following conditions.

(b.0) A message $\theta = \{\theta_n;\ n \in \mathbb{N}\}$ is a white Gaussian noise with $E[\theta_n] = 0$ and $E[\theta_n^2] = 1$.

(b.1) An input $X = \{X_n;\ n \in \mathbb{N}\}$ is of the form

$$X_n = a_n\theta_n + \sum_{k=1}^{n-1} a_k\ (\theta_k - E[\theta_k | \mathcal{F}_{n-1}(Y)]),$$

where a_k, $k = 1,\cdots,n$, are constants and $E[\theta_k | \mathcal{F}_{n-1}(Y)]$ denotes the conditional expectation.

THEOREM 3 [2]. *It holds that*

$$C_F(\rho^2) = \sup \{I(\theta,Y);\ (\theta,X) \in \mathcal{G}(\rho^2)\},$$

where $\mathcal{G}(\rho^2) = \{(\theta,X) \in \mathcal{A}_F(\rho^2);\ \theta$ and X satisfy (b.0) and (b.1)$\}$.

§ 3. Main Results

In this section we shall give the statements of our main results. The proof of them will be given in § 4.

First, we give conditions for the increase of the capacity by feedback.

THEOREM 4. *If*

(i) *the noise* Z *of the GC* (1) *is completely non-white,*

or

(ii) *the noise* Z *is blockwise white but not white and* $\rho > \rho_0$, *where* ρ_0 *is the constant given by* (13),

then the capacity of the GC subject to (2) *is increased by feedback:*

$$C_0(\rho^2) \lneqq C_F(\rho^2). \tag{15}$$

By (10) the capacity $C_0(\rho^2)$ of the GC without feedback is achieved by transmitting the signal $X^0 = \{X_n^0\}$ of (9). To prove Theorem 4, using the signal X^0 and feedback, we shall construct a

coding scheme which enables us to transmit mutual information strictly greater than $C_0(\rho^2)$.

If the dimension N of the GC (1) is equal to 2, a non-white Gaussian noise $Z = (Z_1, Z_2)$ is completely non-white. Therefore, as a consequence of Theorem 2 and Theorem 4, we have the following

COROLLARY. *The capacity of a 2-dimensional GC is increased by feedback if and only if the GC is not white.*

It is conjectured that if $\rho < \rho_0$ the capacity of a blockwise white GC is not increased by feedback. Although we have not succeeded in proving this for the general case, we can show that there exists an N (≥ 3) dimensional blockwise white GC whose capacity is not increased by feedback. We consider a 3-dimensional non-white GC

$$Y_n = X_n + Z_n, \quad n = 1,2,3, \tag{16}$$

to get an example of a GC whose capacity is not increased by feedback. We denote by λ_n^2, $n = 1,2,3$, the eigenvalues of the covariance matrix $\Sigma = \left(\sigma_{mn}\right)_{m,n=1,2,3}$ of the noise $Z = (Z_1, Z_2, Z_3)$. We assume that Z_3 is independent of (Z_1, Z_2) so that

$$\sigma_{13} = \sigma_{23} = 0. \tag{17}$$

We may take λ_1^2 and λ_2^2 as the eigenvalues of the matrix $\left(\sigma_{mn}\right)_{m,n=1,2}$. Then $\lambda_3^2 = \sigma_{33}$. We assume that

$$\lambda_1^2 \geq \lambda_2^2 \geq \rho^2 + \lambda_3^2 > \lambda_3^2 \tag{18}$$

and

$$\frac{\sigma_{11} + \sigma_{22} + \rho^2 + |\sigma_{12}| \sqrt{1 + \sigma_{11}^{-1}\rho^2}}{\Delta} \le \frac{1}{\rho^2 + \sigma_{33}} \tag{19}$$

are satisfied, where $\Delta = \sigma_{11}\sigma_{22} - \sigma_{12}^2$. Note that if σ_{11} and σ_{22} are large enough, and σ_{12} is small enough, then λ_1^2 and λ_2^2 are large, and (18) and (19) hold.

THEOREM 5. *Assume that the conditions* (17) - (19) *are satisfied. Then the capacity of the* 3-*dimensional non-white GC* (16) *subject to* (2) *is not increased by feedback.*

In order to prove Theorem 5 we need to use a result from the capacity of a 2-dimensional GC

$$Y_n = X_n + Z_n, \quad n = 1,2, \tag{20}$$

where $Z = (Z_1, Z_2)$ is a 2-dimensional zero mean Gaussian random variable with covariance matrix $\Sigma = \left(\sigma_{mn} \right)_{m,n=1,2}$ satisfying $\Delta = \sigma_{11}\sigma_{22} - \sigma_{12}^2 > 0$. We denote by $C_F(\rho_1^2, \rho_2^2)$ the capacity of the GC (20) with feedback under the constraint

$$E[X_n^2] \le \rho_n^2, \quad n = 1,2, \quad (\rho_n \ge 0). \tag{21}$$

We can show the following result.

THEOREM 6. *Under the constraint* (21), *the capacity of the* 2-*dimensional GC* (20) *with feedback is given by*

$$C_F(\rho_1^2,\rho_2^2)$$

$$= \frac{1}{2}\log\left[1 + \frac{\sigma_{11}\rho_2^2 + \sigma_{22}\rho_1^2 + \rho_1^2\rho_2^2 + 2\rho_1\rho_2|\sigma_{12}|\sqrt{1+\sigma_{11}^{-1}\rho_1^2}}{\Delta}\right]. \tag{22}$$

Moreover it holds that

$$C_F(\rho^2) = \sup_{\rho_1,\rho_2} C_F(\rho_1^2,\rho_2^2), \tag{23}$$

where the supremum is taken for all $\rho_1 \geq 0$ *and* $\rho_2 \geq 0$ *such that* $\rho_1^2 + \rho_2^2 = \rho^2$.

§ 4. Proof of Theorems

We recall that the capacity $C_0(\rho^2)$ of the GC (1) without feedback is attained by transmitting the signal $X^0 = \{X_n^0\}$ given by (9), and the covariance of X^0 is denoted by $\gamma_{mn} = \gamma_{mn}(\rho) = E[X_m^0 X_n^0]$.

In order to prove Theorem 4 we prepare a lemma.

LEMMA 3. *Assume that* (i) *or* (ii) *of Theorem* 4 *hold. Then there exist* m *and* n $(n \neq m)$ *such that* $\gamma_{mn} \neq 0$.

PROOF. At first we assume (i). Let us suppose that $\gamma_{mn} = 0$

for any $m \neq n$. Let $\mathbb{L}$ be the subset of $\mathbb{N}$ such that $X_n^0 \neq 0$, $n \in \mathbb{L}$ and $X_n^0 = 0$, $n \in \mathbb{M} = \mathbb{N} \setminus \mathbb{L}$. Since $X^0 = \xi^t P$ and $\xi = X^0 P$ we have

$$\mathcal{M}[X_n^0;\ n \in \mathbb{N}] = \mathcal{M}[X_n^0;\ n \in \mathbb{L}] = \mathcal{M}[\xi_n;\ \xi_n \neq 0],$$

where $\mathcal{M}[X_n;\ n \in \mathbb{N}]$ denotes the set of all linear combinations of X_n, $n \in \mathbb{N}$. Therefore we may assume that

$$\xi_n \neq 0, \quad n \in \mathbb{L}, \quad \text{and} \quad \xi_n = 0, \quad n \in \mathbb{M}. \tag{24}$$

It follows from (24) that

$$p_{nm} = 0 \quad \text{if} \quad n \in \mathbb{M} \ \text{and} \ m \in \mathbb{L}. \tag{25}$$

Using (7) and (25) we have

$$Z_n = \sum_{m=1}^{N} p_{nm}\zeta_m = \sum_{m \in \mathbb{M}} p_{nm}\zeta_m \in \mathcal{M}[\zeta_m;\ m \in \mathbb{M}], \quad n \in \mathbb{M},$$

and $\mathcal{M}[Z_n;\ n \in \mathbb{M}] \subset \mathcal{M}[\zeta_n;\ n \in \mathbb{M}]$. Since Z_n, $n \in \mathbb{N}$, are linearly independent, it should be held that

$$\mathcal{M}[Z_n;\ n \in \mathbb{M}] = \mathcal{M}[\zeta_n;\ n \in \mathbb{M}].$$

Therefore we see that

$$\zeta_n = \sum_{m=1}^{N} p_{mn}Z_m = \sum_{m \in \mathbb{M}} p_{mn}Z_m, \quad n \in \mathbb{M},$$

and that

$$p_{mn} = 0 \quad \text{if} \quad m \in \mathbb{L} \ \text{and} \ n \in \mathbb{M}.$$

Thus, if $n \in \mathbb{L}$, $Z_n = \sum_{\ell \in \mathbb{L}} p_{n\ell}\zeta_\ell$ is independent of ζ_m, $m \in \mathbb{M}$, and

consequently independent of Z_m, $m \in \mathbf{M}$. Since the noise Z is completely non-white, this means that $\mathbf{L} = \mathbf{N}$. Therefore, $\xi_n \neq 0$ and $E[\xi_n^2] = \tau_n^2 = a^2 - \lambda_n^2 > 0$ for all $n \in \mathbf{N}$. Thus, for any $m \neq n$, we have

$$0 = \gamma_{mn} = \sum_{k=1}^{N} p_{mk}p_{nk}\tau_k^2 = \sum_{k=1}^{N} p_{mk}p_{nk}(a^2 - \lambda_k^2)$$

$$= a^2 \sum_{k=1}^{N} p_{mk}p_{nk} - \sum_{k=1}^{N} p_{mk}p_{nk}\lambda_k^2 = - \sum_{k=1}^{N} p_{mk}p_{nk}\lambda_k^2$$

$$= - \sigma_{mn}.$$

But this does not occur unless Z is a white noise, and contradicts the assumption (i). Therefore we conclude that there exist m and n $(m \neq n)$ such that $\gamma_{mn} \neq 0$. Secondly we assume (ii). In this case, by Lemma 2, we know that there exists k such that $Z^{(k)} = \{Z_{\pi(n)};\ N_{k-1} < n \leq N_k\}$ is completely non-white, $N_k - N_{k-1} \geq 2$ and (14) holds. Therefore, as we have shown above, there exist $N_{k-1} < m < n \leq N_k$ such that $\gamma_{\pi(m)\pi(n)} \neq 0$. Q.E.D.

We are now in a position to prove Theorem 4.

PROOF OF THEOREM 4. Let $X^0 = \{X_n^0\}$ be the optimal input signal given by (9). By Lemma 3, there exist k and ℓ $(1 \leq k < \ell \leq N)$ such that $\gamma_{k\ell} = E[X_k^0 X_\ell^0] \neq 0$. For simplicity, we put

$$\alpha = \frac{\gamma_{k\ell}}{\gamma_{kk} + \sigma_{kk}}, \qquad \beta = \frac{\gamma_{k\ell}^2}{\gamma_{kk} + \sigma_{kk}}.$$

We define a message $\theta = \{\theta_n\}_{n \in \mathbf{N}}$ by

$$\theta_n = X_n^0, \quad n \neq \ell, \qquad \theta_\ell = X_\ell^0 + X_0,$$

where X_0 is a zero mean Gaussian random variable with variance β which is independent of X^0 and Z. Then we define an input signal $X = \{X_n\}_{n\in\mathbb{N}}$ by

$$X_n = \theta_n, \quad n \neq \ell, \quad X_\ell = \theta_\ell - \alpha Y_k,$$

where $Y = \{Y_n\}$ is the corresponding output:

$$\begin{cases} Y_n = X_n + Z_n = \theta_n + Z_n, & n \neq \ell, \\ Y_\ell = X_\ell + Z_\ell = (\theta_\ell - \alpha Y_k) + Z_\ell. \end{cases}$$

From the definition, $E[X_n^2] = E[\theta_n^2] = \gamma_{nn}$, $n \neq \ell$, and

$$\begin{aligned} E[X_\ell^2] &= E[(\theta_\ell - \alpha(\theta_k + Z_k))^2] \\ &= E[X_0^2] + E[(X_\ell^0 - \alpha X_k^0)^2] + \alpha^2\, E[Z_k^2] \\ &= \beta + \gamma_{\ell\ell} - 2\alpha\gamma_{k\ell} + \alpha^2(\gamma_{kk} + \sigma_{kk}) = \gamma_{\ell\ell}. \end{aligned}$$

Hence

$$\sum_{n=1}^{N} E[X_n^2] = \sum_{n=1}^{N} \gamma_{nn} = \rho^2.$$

Thus $(\theta, X) \in \mathcal{A}_F(\rho^2)$ and consequently

$$I(\theta, Y) \leq C_F(\rho^2). \tag{26}$$

We define $\tilde{Y} = \{\tilde{Y}_n\}$ by

$$\tilde{Y}_n = \theta_n + Z_n, \quad n \in \mathbb{N}.$$

Then it is clear that $\mathcal{M}[\tilde{Y}_n;\ n \in \mathbb{N}] = \mathcal{M}[Y_n;\ n \in \mathbb{N}]$ and

$$I(\theta,\tilde{Y}) = I(\theta,Y). \tag{27}$$

By Lemma 1, $I(\theta,\tilde{Y})$ is given by

$$I(\theta,\tilde{Y}) = \frac{1}{2}\{\log|\tilde{Q}| - \log|\Sigma|\}, \tag{28}$$

where $\tilde{Q} = \left(\tilde{q}_{mn}\right)_{m,n\in\mathbb{N}}$ is the covariance matrix of $\tilde{Y}$. We denote by $Q^0 = \left(q^0_{mn}\right)_{m,n\in\mathbb{N}}$ the covariance matrix of the output $Y^0 = \{Y^0_n\}$ given by $Y^0_n = X^0_n + Z_n$. Then $\tilde{q}_{mn} = E[\tilde{Y}_m\tilde{Y}_n] = E[Y^0_m Y^0_n] = q^0_{mn}$ for any $(m,n) \neq (\ell,\ell)$ and $\tilde{q}_{\ell\ell} = E[(\tilde{Y}_\ell)^2] = E[(Y^0_\ell)^2] + E[X^2_0] = q^0_{\ell\ell} + \beta$. Therefore

$$|\tilde{Q}| = |Q^0| + \beta\Delta^0_{\ell\ell} > |Q^0|, \tag{29}$$

where $\Delta^0_{\ell\ell}$ is (ℓ,ℓ) cofactor of Q^0. It follows from (28), (29) and (10) that

$$I(\theta,\tilde{Y}) > \frac{1}{2}\{\log|Q^0| - \log|\Sigma|\} = I(X^0,Y^0) = C_0(\rho^2). \tag{30}$$

Thus we obtain the desired inequality (15) from (26), (27) and (30).

Q.E.D.

PROOF OF THEOREM 5. By (18) and Theorem 1 we know that the capacity $C_0(\rho^2)$ without feedback is given by

$$C_0(\rho^2) = \frac{1}{2}\log\left(1 + \rho^2\lambda_3^{-2}\right). \tag{31}$$

We proceed to calculate the capacity $C_F(\rho^2)$ of the GC (16) with feedback. Let (θ,X) be any admissible pair of a message and an input signal, namely θ and $X = (X_1,X_2,X_3)$ satisfy (2) and (a.1).

We put $\rho_n^2 = E[X_n^2]$, $n = 1,2,3$. We may assume that $\rho_1^2 + \rho_2^2 + \rho_3^2 = \rho^2$. Using Lemma 1 we have

$$I(\theta,Y) = I(\theta,(Y_1,Y_2)) + h(Y_3|Y_1,Y_2) - h(Y_3|\theta,Y_1,Y_2)$$

$$= I(\theta,(Y_1,Y_2)) + h(Y_3|Y_1,Y_2) - \frac{1}{2}\log(2\pi e\lambda_3^2). \tag{32}$$

Since Z_3 is independent of (Y_1,Y_2), Z_3 is independent of X_3 and it follows from (ii) of Lemma 1 that

$$h(Y_3|Y_1,Y_2) \le \frac{1}{2}\log\{2\pi e(\rho_3^2 + \lambda_3^2)\}. \tag{33}$$

Clearly $I(\theta,(Y_1,Y_2))$ is upper bounded by $C_F(\rho_1^2,\rho_2^2)$. We put $\gamma^2 = \rho_1^2 + \rho_2^2 = \rho^2 - \rho_3^2$, then $\rho_1^2, \rho_2^2, 2\rho_1\rho_2 \le \gamma^2 \le \rho^2$. Using (22) we get the following inequality:

$$I(\theta,(Y_1,Y_2)) \le C_F(\rho_1^2,\rho_2^2)$$

$$\le \frac{1}{2}\log\left[1 + \left(\frac{\sigma_{11} + \sigma_{22} + \rho^2 + |\sigma_{12}|\sqrt{1 + \sigma_{11}^{-1}\rho^2}}{\Delta}\right)\gamma^2\right]. \tag{34}$$

It follows from (34) and (19) that

$$I(\theta,(Y_1,Y_2)) \le \frac{1}{2}\log\left(1 + \frac{\gamma^2}{\rho^2 + \lambda_3^2}\right)$$

$$\le \frac{1}{2}\log\left(1 + \frac{\gamma^2}{\rho_3^3 + \lambda_3^2}\right) = \frac{1}{2}\log\left(\frac{\rho^2 + \lambda_3^2}{\rho_3^3 + \lambda_3^2}\right). \tag{35}$$

Combine (31) - (33) and (35) to get

$$I(\theta,Y) \le \frac{1}{2}\log\left(\frac{\rho_3^2+\lambda_3^2}{\lambda_3^2}\right) + \frac{1}{2}\log\left(\frac{\rho^2+\lambda_3^2}{\rho_3^2+\lambda_3^2}\right)$$

$$= \frac{1}{2}\log\left(1+\rho^2\lambda_3^{-2}\right) = C_0(\rho^2),$$

for any admissible pair (θ,X). This means that $C_F(\rho^2) \le C_0(\rho^2)$. Since $C_0(\rho^2) \le C_F(\rho^2)$ by definition, $C_F(\rho^2)$ is equal to $C_0(\rho^2)$.

Q.E.D.

PROOF OF THEOREM 6. We consider the following coding scheme

$$\begin{cases} X_1 = \rho_1\theta_1 \\ X_2 = \alpha\theta_2 + \beta(\theta_1 - \hat{\theta}_1), \end{cases} \tag{36}$$

where θ_1 and θ_2 are mutually independent zero mean Gaussian random variables with variance one, $\hat{\theta}_1 = E[\theta_1|\mathcal{F}(Y_1)]$, and α and β are constants such that

$$\alpha^2 + \beta^2 E[|\theta_1 - \hat{\theta}_1|^2] = \rho_2^2. \tag{37}$$

In [2] it is shown that

$$C_F(\rho_1^2,\rho_2^2) = \sup_{\alpha,\beta} I(\theta,Y),$$

where $\theta = (\theta_1,\theta_2)$, $Y = (Y_1,Y_2)$ and the supremum is taken for all constants α and β satisfying (37). For simplicity we put $\tau = (\sigma_{11} + \rho_1^2)^{1/2}$. For the coding scheme (36), we know that

$$E[|\theta_1 - \hat{\theta}_1|^2] = \sigma_{11}(\rho_1^2 + \sigma_{11})^{-1} = \sigma_{11}\tau^{-2}$$

Denote by $\hat{Z}_2 = E[Z_2|\mathcal{F}(Y_1)]$ and $\tilde{Z}_2 = E[Z_2|\mathcal{F}(Z_1)]$. Then $\hat{Z}_2 =$

$\sigma_{12}\tau^{-2}Y_1$, $\tilde{Z}_2 = \sigma_{12}\sigma_{11}^{-1}Z_1$,

$$E[|Z_2 - \hat{Z}_2|^2] = \sigma_{22} - \sigma_{12}^2\tau^{-2}$$

and

$$E[|Z_2 - \tilde{Z}_2|^2] = \Delta\sigma_{11}^{-1}.$$

By (4) we have

$$I(\theta,Y) = I(\theta,Y_1) + h(Y_2|Y_1) - h(Y_2|Y_1,\theta). \tag{38}$$

We can show that

$$I(\theta,Y_1) = I(\theta_1,Y_1) = \frac{1}{2}\log\left(\frac{\rho_1^2 + \sigma_{11}}{\sigma_{11}}\right) = \frac{1}{2}\log\frac{\tau^2}{\sigma_{11}}, \tag{39}$$

$$h(Y_2|Y_1) = \frac{1}{2}\log\{2\pi e\, E[|\alpha\theta_2 + \beta\,(\theta_1 - \hat{\theta}_1) + (Z_2 - \hat{Z}_2)|^2]\} \tag{40}$$

and

$$h(Y_2|Y_1,\theta) = h(Z_2|Y_1,\theta) = \frac{1}{2}\log\{2\pi e\, E[|Z_2 - \tilde{Z}_2|^2]\}$$

$$= \frac{1}{2}\log(2\pi e\Delta\sigma_{11}^{-1}). \tag{41}$$

In order to get the capacity $C_F(\rho_1^2,\rho_2^2)$, we need to maximize

$$\varphi(\alpha,\beta) = E[|\alpha\theta_2 + \beta\,(\theta_1 - \hat{\theta}_1) + (Z_2 - \hat{Z}_2)|^2],$$

which can be written as

$$\varphi(\alpha,\beta) = E[|\alpha\theta_2 + \beta\,(\theta_1 - \hat{\theta}_1)|^2]$$

$$+ 2\beta\, E[(\theta_1 - \hat{\theta}_1)(Z_2 - \hat{Z}_2)] + E[|Z_2 - \hat{Z}_2|^2]$$

$$= \rho_2^2 - 2\beta\, E[\theta_1\hat{Z}_2] + E[|Z_2 - \hat{Z}_2|^2]$$

$$= \rho_2^2 - 2\beta\rho_1\sigma_{12}\tau^{-2} + \sigma_{22} - \sigma_{12}^2\tau^{-2}.$$

Since

$$\rho_2^2 = \alpha^2 + \beta^2 \, E[|\theta_1 - \hat{\theta}_1|^2] = \alpha^2 + \beta^2\sigma_{11}\tau^{-2},$$

β can run on the interval

$$- \rho_2\sigma_{11}^{-1/2}\tau \le \beta \le \rho_2\sigma_{11}^{-1/2}\tau.$$

Thus, if $\sigma_{12} > 0$ (resp. $\sigma_{12} < 0$), $\varphi(\alpha,\beta)$ is maximized when $\beta = - \rho_2\sigma_{11}^{-1/2}\tau$ (resp. $\beta = \rho_2\sigma_{11}^{-1/2}\tau$) and $\alpha = 0$, and the maximum value is

$$\rho_2^2 + \sigma_{22} - \sigma_{12}^2\tau^{-2} + 2\rho_1\rho_2|\sigma_{12}|\sigma_{11}^{-1/2}\tau^{-1}.$$

Combining this with (38) - (41), we get (22). Eq. (23) is clear from the definition of $C_F(\rho^2)$. Q.E.D.

Appendix

PROOF OF THEOREM 2. For each $(\rho_1^2,\cdots,\rho_N^2)$ such that $\Sigma_{n=1}^N \rho_n^2 = \rho^2$, we denote by $C_F(\rho_1^2,\cdots,\rho_N^2)$ (resp. $C_0(\rho_1^2,\cdots,\rho_N^2)$) the capacity of the GC with (resp. without) feedback under a constraint

$$E[X_n^2] \le \rho_n^2, \quad N = 1,\cdots,N. \tag{42}$$

It is known that

$$C_0(\rho_1^2,\cdots,\rho_N^2) = \frac{1}{2}\sum_{n=1}^{N} \log\left(1 + \rho_n^2\lambda_n^{-2}\right).$$

In order to prove Theorem 2 it suffices to show

$$C_0(\rho_1^2,\cdots,\rho_N^2) = C_F(\rho_1^2,\cdots,\rho_N^2). \tag{43}$$

For any pair (θ,X) satisfying (a.1) and (42) we have

$$h(Z_n|Z_1,\cdots,Z_{n-1}) = h(Z_n) = \frac{1}{2}\log(2\pi e\lambda_n^2)$$

and

$$h(Y_n|Y_1,\cdots,Y_{n-1}) \le h(Y_n) \le \frac{1}{2}\log[2\pi e(\rho_n^2 + \lambda_n^2)].$$

Therefore, using Lemma 1, we have

$$I(\theta,Y) \le \frac{1}{2}\sum_{n=1}^{N} \log\left(1 + \rho_n^2\lambda_n^{-2}\right),$$

and consequently

$$C_0(\rho_1^2,\cdots,\rho_N^2) \le C_F(\rho_1^2,\cdots,\rho_N^2)$$

$$\le \frac{1}{2}\sum_{n=1}^{N} \log\left(1 + \rho_n^2\lambda_n^{-2}\right) = C_0(\rho_1^2,\cdots,\rho_N^2).$$

Thus we get (43) and the proof is completed. Q.E.D.

Acknowledgement. The authors would like to thank Professor C.R.Baker for a number of helpful discussions.

References

[1] R.G. Gallager, Information Theory and Reliable Communication. Wiley, New York, 1968.

[2] S. Ihara, On the capacity of the discrete time Gaussian channel with feedback. Trans. Eighth Prague Conf. on Inform. Theory, Stat. Decision Fun., Random Processes, vol. C, pp. 175-186, Czecho. Acad. Sci., 1979.

[3] S. Ihara, Stochastic Processes and Entropy. Iwanami Shoten, Tokyo, 1984 (in Japanese).

[4] S. Ihara, Capacity of mismatched Gaussian channels with and without feedback. submitted, 1986.

[5] S. Ihara, Capacity of discrete time Gaussian channel with and without feedback, I. Mem. Fac. Sci. Kochi Univ. (Math.), 9 (1988), 21-36.

[6] T.T. Kadota, M. Zakai and J. Ziv, Mutual information of white Gaussian channels with and without feedback. Trans. Inform. Theory, IT-17 (1971), 368-371.

[7] S. Pombra and T.M. Cover, Feedback increases capacity of Gaussian channels by at most half a bit. presented at 1986 Int. Symp. on Inform. Theory, Ann Arbor, Michigan, 1986.