Japan J. Appl. Math., 3 (1986), 197–206

A New Treatment of Communication Processes with Gaussian Channels

(Dedicated to Professor H. Umegaki on his 60th birthday)

Masanori OHYA and Noboru WATANABE

Department of Information Sciences,
Science University of Tokyo,
Noda City, Chiba 278, Japan

Received May 17, 1985
Revised September 28, 1985

In order to discuss communication processes consistently for a Gaussian input with a Gaussian channel on an infinite dimensional Hilbert space, we introduce the entropy functional of an input source and the mutual entropy functional for a Gaussian channel and show a fundamental inequality for communication processes.

Key words: communication theory, quantum probability, Gauss measure, entropy

Introduction

A Gaussian measure on a Hilbert space is studied and applied to communication processes by several authors [1, 4, 15, 26]. Particularly Baker [1] introduced a mutual information for a Gaussian channel based on a work by Gelfand-Yaglom [2]. In a usual communication theory, one takes the differential entropy as the definition of the entropy (information) carried by an input source. However, for an input Gaussian measure, we understand by a simple consideration that the differential entropy for an input source is not compatible with the mutual information mentioned above in Shannon's communication theory, so that the differential entropy is not good at discussing the Gaussian communication process. The main purpose of this paper is to introduce two functionals, say the entropy functional and the mutual entropy functional, for an input Gaussian source and a Gaussian channel, and prove a fundamental inequality for the communication process. Our formulation of these entropy functionals are based on a formulation of quantum mechanical information theory given in [9].

§ 1. Gaussian Measures on a Hilbert Space

Let $\mathscr{H}$ be a real separable Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and $\mathscr{B}$ be the Borel σ-field of $\mathscr{H}$. μ is a Borel probability measure on $\mathscr{B}$ satisfying

$$\int_{\mathscr{H}} \|x\|^2 d\mu(x) < \infty .$$

Further, we define the mean vector $m_\mu \in \mathscr{H}$ and the covariance operator R_μ of μ such as

$$\langle x_1, m_\mu \rangle = \int_{\mathscr{H}} \langle x_1, y \rangle \mu(dy)$$

$$\langle x_1, R_\mu x_2 \rangle = \int_{\mathscr{H}} \langle x_1, y - m_\mu \rangle \langle y - m_\mu, x_2 \rangle \mu(dy)$$

for any $x_1, x_2, y \in \mathscr{H}$. We denote the set of all bounded linear operators on $\mathscr{H}$ by $B(\mathscr{H})$ and denote the set of all positive self-adjoint trace class operators on $\mathscr{H}$ by $T(\mathscr{H})_+$ ($\equiv \{\rho \in B(\mathscr{H}), \rho \geqq 0, \rho^* = \rho, \operatorname{tr} \rho < \infty\}$). A Gaussian measure μ in $\mathscr{H}$ is a Borel measure in $\mathscr{H}$ such that for each $x \in \mathscr{H}$, there exist real numbers m_x and σ_x (>0) satisfying

$$\mu\{y \in \mathscr{H}; \langle y, x \rangle \leqq a\} = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi\sigma_x}} e^{-(t-m_x)^2/2\sigma_x} dt .$$

Then the characteristic function of μ is given by

$$\hat{\mu}(x) = \exp\left\{ i\langle x, m_x \rangle - \frac{1}{2} \langle x, R_\mu x \rangle \right\},$$

where R_μ is an element of $T(\mathscr{H})_+$. It is known [7] that a Gaussian measure μ with a mean vector 0 one-to-one corresponds to a covariance operator of μ. The notation $\mu = [m, R]$ means that μ is a Gaussian measure on $\mathscr{H}$ with a mean vector m and a covariance operator R. $\mu \gg \nu$ means that ν is absolutely continuous with respect to μ. Furthermore we denote (1) $\mu \sim \nu$ if μ is equivalent to ν, that is $\mu \gg \nu$ and $\nu \gg \mu$; (2) $\mu \perp \nu$ if μ is singular to ν. The relation $\mu \sim \nu$ or $\mu \perp \nu$ is satisfied for any pair of Gaussian measures μ and ν ([18]).

Before closing this section, we remember the relative entropy of two probability measures μ and ν. This entropy (Kullback-Leibler information) is defined by [2, 6]

$$S(\nu \mid \mu) = \int_{\mathscr{H}} \frac{d\nu}{d\mu} \log \frac{d\nu}{d\mu} d\mu \qquad \text{when} \quad \mu \gg \nu ,$$

$$= \infty \qquad \text{when} \quad \mu \not\gg \nu ,$$

where $d\nu/d\mu$ is the Radon-Nikodym derivative of ν w.r.t. μ.

§ 2. Gaussian Channel

Let $(\mathscr{H}_1, \mathscr{B}_1)$ be an input space, $(\mathscr{H}_2, \mathscr{B}_2)$ be an output space and $P_G^{(k)}$ be the set of all Gaussian probability measures on $(\mathscr{H}_k, \mathscr{B}_k)$ ($k = 1, 2$). We call mapping λ: $\mathscr{H}_1 \times \mathscr{B}_2 \to [0, 1]$ a Gaussian channel from the input space to the output space if λ satisfies the following conditions:

(1) $\lambda(x, \cdot) \in P_G^{(2)}$ for each fixed $x \in \mathscr{H}_1$,

(2) $\lambda(\cdot, Q)$ is a measurable function on $(\mathscr{H}_1, \mathscr{B}_1)$ for each fixed $Q \in \mathscr{B}_2$.

Let $\mu_1 \in P_G^{(1)}$ be a Gaussian measure of the input space and $\mu_0 \in P_G^{(2)}$ be a Gaussian measure indicating a noise of the channel. Then, the Gaussian measure $\mu_2 \in P_G^{(2)}$ obtained in the output system can be expressed by the channel λ such as

$$\mu_2(Q)=\int_{\mathscr{H}_1} \lambda(x, Q)\mu_1(dx) \tag{2.1}$$

$$\lambda(x, Q) \equiv \mu_0(Q^x),$$

$$Q^x \equiv \{y \in \mathscr{H}_2;\ Ax+y \in Q\}, \qquad x \in \mathscr{H}_1, \quad Q \in \mathscr{B}_2,$$

where A is a linear transformation from $\mathscr{H}_1$ to $\mathscr{H}_2$.

$$P_G^{(1)} \ni \mu_1 \longrightarrow \boxed{\text{channel}} \longrightarrow \mu_2 \in P_G^{(2)}$$

$$\uparrow \mu_0 : \text{noise}$$

The compound measure μ_{12} derived from the input measure μ_1 and the output measure μ_2 is given by

$$\mu_{12}(Q_1 \times Q_2)=\int_{Q_1} \lambda(x, Q_2)\mu_1(dx), \tag{2.2}$$

for any $Q_1 \in \mathscr{B}_1$ and $Q_2 \in \mathscr{B}_2$. Then, the mutual entropy (information) w.r.t. μ_1 and λ is defined by the Kullback-Leibler information such as

$$I(\mu_1; \lambda)=S(\mu_{12} \,|\, \mu_1 \otimes \mu_2). \tag{2.3}$$

§ 3. A Model

For simplicity, we put $\mathscr{H}_1=\mathscr{H}_2=\boldsymbol{R}^2$ in this section. Let two Gaussian measures μ_1 and μ_0 be given by $\mu_1=[0, R_1]$, $\mu_0=[0, R_0]$ with

$$R_1=\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, \qquad R_0=\begin{pmatrix} \frac{1}{48} & 0 \\ 0 & \frac{1}{48} \end{pmatrix}.$$

(Remark that we have many other choices of the covariance operators R_1 and R_0.) We take the linear transformation A used in (2.1) as

$$A=\begin{pmatrix} \sqrt{\frac{23}{24}} & 0 \\ 0 & \sqrt{\frac{23}{24}} \end{pmatrix}.$$

Under these settings, the covariance operator R_2 of the output measure μ_2 becomes

$$R_2=\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}.$$

From Proposition 2 of [1], we obtain an operator $V \equiv R_1^{1/2} A^* R_2^{-1/2}$ as

$$V=\begin{pmatrix} \sqrt{\frac{23}{24}} & 0 \\ 0 & \sqrt{\frac{23}{24}} \end{pmatrix}.$$

by which the compound measure μ_{12} is determined. Further, the mutual entropy w.r.t. μ_1 and the channel λ determined by μ_0 is calculated as

$$I(\mu_1; \lambda) = S(\mu_{12} \mid \mu_1 \otimes \mu_2) = \log 24 .$$

Now, if we use the differential entropy as the definition of entropy (information) for an input Gaussian measure (this definition is often used in literature dealing with communication theory), then

$$\begin{aligned} S(\mu_1) &= -\int_{R^2} \frac{d\mu}{dm} \log \frac{d\mu}{dm} dm \\ &= \log(\pi e), \end{aligned}$$

where m is a Lebesgue measure of $\boldsymbol{R}^2$. Therefore, the mutual information $I(\mu_1; \lambda)$ becomes larger than $S(\mu_1)$, which is a contradiction to the usual Shannon's theory [2, 6, 14, 16]. Thus, the differential entropy is not suitable for the definition of entropy for an input Gaussian measure.

By the way, if we take

$$S(\mu_1) = \sup \left\{ -\sum_{A \in \tilde{A}} \mu_1(A) \log \mu_1(A);\ \tilde{A} \in \mathscr{P}(\mathscr{B}_2) \right\}$$

as the definition of entropy for an input Gaussian measure as a straight extension of the Shannon entropy for a discrete probability distribution (where $\mathscr{P}(\mathscr{B}_2)$ is the set of all finite partitions of $\mathscr{B}_2$), then it is easily seen that $S(\mu_1)$ is infinite. In this case, the mutual entropy is smaller than $S(\mu_1)$, but it is difficult to comprehend the physical meaning of the fact that every input Gaussian measure carries infinite information. Moreover, it is impossible to distinguish an input Gaussian state from other Gaussian states only by using the entropy $S(\mu_1)$, as it is always infinite. Therefore we had better find some other expressions (quantities) characterizing a Gaussian state and a Gaussian channel so that we can discuss the Gaussian communication process consistently.

§4. A New Formulation of the Mutual Entropy

In order to discuss a dynamical change of states in quantum systems, a quantum mechanical channel is useful and is studied in various aspects [8, 9, 11, 21, 22, 23]. This quantum mechanical channel is generally defined as follows:

A mapping Λ^* from $T(\mathcal{H}_1)_{+,1}$ to $T(\mathcal{H}_2)_{+,1}$ (where $T(\mathcal{H}_k)_{+,1} \equiv \{\rho \in T(\mathcal{H}_k)_+;\ \mathrm{tr}\,\rho = 1\}$ $(k = 1, 2)$) is said to be a channel if its dual map Λ from $B(\mathcal{H}_2)$ to $B(\mathcal{H}_1)$ satisfies the following three conditions: (i) *Λ is completely positive (i.e., $\sum_{k,j=1}^{n} A_k^* \Lambda(B_k^* B_j) A_j$ is positive for any $A_k \in B(\mathcal{H}_1)$, $B_j \in B(\mathcal{H}_2)$ and for any $n \in N$),* (ii) *$\Lambda I_2 = I_1$ (where I_k is the identity operator on $\mathcal{H}_k$ $(k = 1, 2)$) and* (iii) *Λ is normal (i.e., $\Lambda(A_n) \uparrow \Lambda(A)$ for any $\{A_n\} \subset B(\mathcal{H}_2)$ with $A_n \uparrow A$).*

A typical example of a channel is the conditional expectation of a set of observables to its subalgebras which plays quite an important role in quantum probability theory [21, 22, 23].

When an input state is given by a density operator $\rho \in T(\mathcal{H}_1)_{+,1}$, von Neumann introduced [25] the entropy of the input state ρ such as

$$S(\rho) = -\mathrm{tr}\,\rho \log \rho\,. \tag{4.1}$$

We now denote a Schatten decomposition [17] of ρ such as

$$\rho = \sum_k \lambda_k E_k\,, \tag{4.2}$$

where E_k is the projection from $\mathcal{H}_1$ to the one-dimensional subspace of $\mathcal{H}_1$ generated by an eigenvector x_k associated to the eigenvalue λ_k, that is, $E_k = |x_k\rangle\langle x_k|$ in the Dirac notation. In (4.2), the eigenvalue of multiplicity n is repeated precisely n times. Note that this decomposition is not unique unless every eigenvalue is nondegenerate. Then a compound state expressing the correlation existing between an initial state ρ and the final state $\Lambda^*\rho$ is defined on the tensor product Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$ such as

$$\sigma_E = \sum_k \lambda_k E_k \otimes \Lambda^* E_k\,, \tag{4.3}$$

where we use the index E because this compound state depends on the choice of $E = \{E_k\}$. This compound state is introduced in [9, 10] and plays an essential role in studying the dynamics of state change [10, 12] and quantum information theory discussed in [3, 5, 9, 19]. In particular, the mutual entropy (information) with respect to the input state ρ and the communication channel Λ^* is given by [9]

$$I(\rho; \Lambda^*) = \sup_E S(\sigma_E \,|\, \sigma_0)\,, \tag{4.4}$$

where σ_0 is a trivial compound state

$$\sigma_0 = \rho \otimes \Lambda^* \rho \tag{4.5}$$

and $S(\sigma_E \,|\, \sigma_0)$ is the relative entropy [23] of σ_E w.r.t. σ_0

$$S(\sigma_E \,|\, \sigma_0) = \operatorname{tr} \sigma_E (\log \sigma_E - \log \sigma_0) \,. \tag{4.6}$$

Let $P^{(k)}_{G,1}$ be the set $\{\mu = [0, R] \in P^{(k)}_G;\ \operatorname{tr} R = 1\}$ $(k = 1, 2)$. We assume that $A^*A = (1 - \operatorname{tr} R_0) I_1$ holds for the covariance operator R_0 of μ_0. We define a transformation Γ^* from $P^{(1)}_{G,1}$ to $P^{(2)}_{G,1}$ associated with the Gaussian channel λ:

$$(\Gamma^* \mu_1)(Q) = \int_{\mathscr{H}_1} \lambda(x, Q) \mu_1(dx) \tag{4.7}$$

for any $\mu_1 \in P^{(1)}_{G,1}$ and any $Q \in \mathscr{B}_2$. (4.7) can be expressed as

$$\Gamma^*(\mu_1) = [0, A \rho_1 A^* + R_0] \tag{4.8}$$

for any $\mu_1 = [0, \rho_1] \in P^{(1)}_{G,1}$. There exists a bijection Ξ^*_k from $P^{(k)}_{G,1}$ to $T(\mathscr{H}_k)_{+,1}$ given by

$$\operatorname{tr} \Xi^*_k(\mu_k) A_k = \int_{\mathscr{H}_k} \langle \xi, A_k \xi \rangle \mu_k(d\xi) \tag{4.9}$$

for any $A_k \in B(\mathscr{H}_k)$ and any $\mu_k \in P^{(k)}_{G,1}$ $(k = 1, 2)$. We further define a map from $T(\mathscr{H}_1)_{+,1}$ to $T(\mathscr{H}_2)_{+,1}$ such as

$$\Lambda^* \rho_1 = \Xi^*_2 \circ \Gamma^* \circ (\Xi^*_1)^{-1} \rho_1 \tag{4.10}$$

for any $\rho_1 \in T(\mathscr{H}_1)_{+,1}$. (4.10) can be expressed as

$$\Lambda^* \rho_1 = A \rho_1 A^* + R_0 \tag{4.11}$$

for any $\rho_1 \in T(\mathscr{H}_1)_{+,1}$.

$$\begin{array}{ccc} P^{(1)}_{G,1} & \xrightarrow{\ \Gamma^*\ } & P^{(2)}_{G,1} \\ \Big\downarrow \Xi^*_1 & & \Big\downarrow \Xi^*_2 \\ T(\mathscr{H}_1)_{+,1} & \xrightarrow{\ \Lambda^*\ } & T(\mathscr{H}_2)_{+,1} \end{array}$$

Λ is the dual map of Λ^* from $B(\mathscr{H}_2)$ to $B(\mathscr{H}_1)$; that is,

$$\operatorname{tr} \rho_1 \Lambda(A_2) = \operatorname{tr} \Lambda^*(\rho_1) A_2 \tag{4.12}$$

for any $A_2 \in B(\mathscr{H}_2)$ and any $\rho_1 \in T(\mathscr{H}_1)_{+,1}$. Therefore Λ is written as

$$\Lambda(Q) = A^* Q A + (\operatorname{tr} R_0 Q) I_1 \tag{4.13}$$

for any $Q \in B(\mathscr{H}_2)$.

THEOREM 1. *Λ^* is a quantum mechanical channel from $T(\mathscr{H}_1)_{+,1}$ to $T(\mathscr{H}_2)_{+,1}$.*

Proof. From the definition of quantum mechanical channel, we have to show the following three properties of the map Λ from $B(\mathscr{H}_2)$ to $B(\mathscr{H}_1)$ given in (4.12): (1) Completely positivity of Λ; (2) $\Lambda(I_2) = I_1$; (3) Normality of Λ.

Proof of (1). For any $\{Q_i\}_{i=1}^n \subset B(\mathscr{H}_1)$ and any $\{R_j\}_{j=1}^n \subset B(\mathscr{H}_2)$ with $n \in N$, we have

$$\sum_{i,j=1}^n Q_i^* \Lambda(R_i^* R_j) Q_j = \sum_{i,j=1}^n Q_i^*(A^* R_i^* R_j A + (\operatorname{tr} R_0 R_i^* R_j) I_1) Q_j$$
$$= \sum_{i,j=1}^n Q_i^* A^* R_i^* R_j A Q_j + \sum_{i,j=1}^n (\operatorname{tr} R_0 R_i^* R_j) Q_i^* Q_j$$
$$= \left(\sum_{i=1}^n R_i A Q_i\right)^* \left(\sum_{j=1}^n R_j A Q_j\right) + \sum_{i,j=1}^n \sum_k \langle x_k, R_0 R_i^* R_j x_k \rangle Q_i^* Q_j .$$

Let $\{y_m\}$ be any CONS in $\mathscr{H}_1$ and put $C = \sum_{i=1}^n R_i A Q_i$. Then the above equality is identical to

$$C^*C + \sum_{i,j=1}^n \sum_{k,m} \langle x_k, R_j R_0^{1/2} y_m \rangle \langle y_m, R_0^{1/2} R_i^* x_k \rangle Q_i^* Q_j$$
$$= C^*C + \sum_{k,m} \left(\sum_{i=1}^n \langle y_m, R_0^{1/2} R_i^* x_k \rangle Q_i^* \right) \left(\sum_{j=1}^n \langle x_k, R_j R_0^{1/2} y_m \rangle Q_j \right)$$
$$= C^*C + \sum_{k,m} \left(\sum_{i=1}^n \overline{\langle x_k, R_i R_0^{1/2} y_m \rangle} Q_i^* \right) \left(\sum_{j=1}^n \langle x_k, R_j R_0^{1/2} y_m \rangle Q_j \right)$$
$$= C^*C + \sum_{k,m} \left(\sum_{i=1}^n (\langle x_k, R_i R_0^{1/2} y_m \rangle Q_i)^* \right) \left(\sum_{j=1}^n \langle x_k, R_j R_0^{1/2} y_m \rangle Q_j \right)$$
$$= C^*C + \sum_{k,m} \left(\sum_{i=1}^n \langle x_k, R_i R_0^{1/2} y_m \rangle Q_i \right)^* \left(\sum_{j=1}^n \langle x_k, R_j R_0^{1/2} y_m \rangle Q_j \right) \geqq 0 .$$

This inequality holds for any $n \in N$, so that Λ is a completely positive map from $B(\mathscr{H}_2)$ to $B(\mathscr{H}_1)$.

Proof of (2).

$$\Lambda(I_2) = A^* I_2 A + (\operatorname{tr} R_0 I_2) I_1 = A^*A + (\operatorname{tr} R_0) I_1 = (1 - \operatorname{tr} R_0) I_1 + (\operatorname{tr} R_0) I_1 = I_1 .$$

Proof of (3). For any increasing net $\{B_\alpha\}$ ($\subset B(\mathscr{H}_2)$) ultrastrongly converging to $B \in B(\mathscr{H}_2)$ and for any sequence $\{x_n\}$ ($\subset \mathscr{H}_1$) satisfying $\sum_n \|x_n\|^2 < \infty$, we have

$$\sum_n \|(\Lambda(B) - \Lambda(B_\alpha)) x_n\|^2 = \sum_n \|\Lambda(B - B_\alpha) x_n\|^2$$
$$= \sum_n \|\{A^*(B - B_\alpha)A + (\operatorname{tr} R_0(B - B_\alpha)) I_1\} x_n\|^2$$
$$\leqq \sum_n \|A^*(B - B_\alpha) A x_n\|^2 + \sum_n |\operatorname{tr} R_0(B - B_\alpha)|^2 \|x_n\|^2$$
$$\leqq \|A^*\|^2 \sum_n \|(B - B_\alpha) A x_n\|^2 + |\operatorname{tr} R_0(B - B_\alpha)|^2 \sum_n \|x_n\|^2 .$$

Since $\sum_n \|Ax_n\|^2 \leqq \|A\|^2 \sum_n \|x_n\|^2 < \infty$ is satisfied and R_0 is a trace class operator, then $\Lambda(B_\alpha)$ ultrastrongly converges to $\Lambda(B)$. Q.E.D.

In order to introduce new functionals for a consistent treatment of the Gaussian communication process, we first prove a theorem for the Gaussian measure with the covariance operator σ_E. Let us define two probability measures $\bar{\mu}_1$ and $\bar{\mu}_2$ on $\mathscr{H}_1$ and $\mathscr{H}_2$ respectively such as

$$\bar{\mu}_k(A)=\int_A \|\xi\|^2 d\mu_k(\xi)\,,$$

for any $A\in\mathscr{B}_k$ and any $\mu_k\in P_{G,1}^{(k)}$.

THEOREM 2. *The Gaussian measure* $\mu=[0, \sigma_E]$ *is a compound state (measure) derived from the input measure* $\mu_1=[0, \Xi_1^*(\mu_1)]$ *on* $\mathscr{H}_1$ *and the output measure* $\mu_2=[0, \Lambda^*\circ\Xi_1^*(\mu_1)]$ *on* $\mathscr{H}_2$ *in the sense that*

$$\bar{\mu}_1(A)=\bar{\mu}(A\otimes\mathscr{H}_2),\ \textit{for any subspace } A \textit{ in } \mathscr{B}_1\,,$$

$$\bar{\mu}_2(B)=\bar{\mu}(\mathscr{H}_1\otimes B),\ \textit{for any subspace } B \textit{ in } \mathscr{B}_2\,.$$

Proof. From (4.9), we have

$$\bar{\mu}_k(A)=\operatorname{tr}\,\Xi_k^*(\mu_k)P_A\,,$$

where P_A is the projection operator from $\mathscr{H}_k$ on the subspace A in $\mathscr{B}_k$.
Hence

$$\begin{aligned}\bar{\mu}(A\otimes\mathscr{H}_2)&=\operatorname{tr}\sigma_E P_A\otimes I_2\\ &=\operatorname{tr}\rho P_A\\ &=\operatorname{tr}\,\Xi_1^*(\mu_1)P_A\\ &=\bar{\mu}_1(A)\,.\end{aligned}$$

Similarly,

$$\bar{\mu}_2(B)=\bar{\mu}(\mathscr{H}_1\otimes B),\ \text{for any subspace } B \text{ in } \mathscr{B}_2\,. \qquad \text{Q.E.D.}$$

From (4.4), we define a functional (say the mutual entropy functional) with respect to the input Gaussian measure μ_1 and the Gaussian channel λ as

$$\begin{aligned}\tilde{I}(\mu_1;\lambda)&=\sup_E S(\sigma_E\,|\,\sigma_0)\\ &=\sup_E \operatorname{tr}\sigma_E(\log\sigma_E-\log\sigma_0)\,,\end{aligned}\tag{4.14}$$

where σ_E is a compound state given by (4.3) for a density operator $\Xi_1^*(\mu_1)$ and σ_0 is $\Xi_1^*(\mu_1)\otimes\Lambda^*\circ\Xi_1^*(\mu_1)$. Another functional (say the entropy functional) of the input Gaussian measure $\mu_1=[0, \Xi_1^*(\mu_1)]$ expressing certain "information" of μ_1 is given by

$$\tilde{S}(\mu_1)=-\operatorname{tr}\,\Xi_1^*(\mu_1)\log\Xi_1^*(\mu_1)\,.$$

By the following result, we understand that the above functionals play the same role as the entropy and the mutual entropy in Shannon's communication theory.

THEOREM 3. *For any $\mu_1 \in P_{G,1}^{(1)}$ and for some Gaussian channel λ, we obtain*

$$0 \leqq \tilde{I}(\mu_1; \lambda) \leqq \tilde{S}(\mu_1).$$

Proof. This theorem has been essentially proved in ref. [9]. However, we here sketch the proof for completeness of this paper. After some calculation, we obtain

$$\begin{aligned} S(\sigma_E \mid \sigma_0) &\leqq \sum_n \lambda_n S(E_n \mid \Xi_1^*(\mu_1)) \\ &= \sum_n \lambda_n (\operatorname{tr} E_n \log E_n - \operatorname{tr} E_n \log \Xi_1^*(\mu_1)) \\ &= -\operatorname{tr} \sum_n \lambda_n E_n \log \Xi_1^*(\mu_1) = -\operatorname{tr} \Xi_1^*(\mu_1) \log \Xi_1^*(\mu_1) \\ &= \tilde{S}(\mu_1). \end{aligned}$$

where $E = \{E_n\}$ of the Schatten decomposition $\Xi_1^*(\mu_1) = \sum_n \lambda_n E_n$. Taking the supremum over E, we get $0 \leqq \tilde{I}(\mu_1; \lambda) \leqq \tilde{S}(\mu_1)$. Q.E.D.

For the model discussed in § 3, we calculate the entropy functional $\tilde{S}(\mu_1)$ and the mutual entropy functional $\tilde{I}(\mu_1; \lambda)$ concretely:

$$\tilde{I}(\mu_1; \lambda) = \log 2 = \tilde{S}(\mu_1).$$

Consequently, the difficulty appearing in a model of § 3 is resolved in our formulation. Besides mathematical formulation, our functionals classify the Gaussian inputs and might be useful to analyse the Gaussian communication process in detail, upon which we are still working.

Acknowledgment. The authors wish to express their gratitude to Professor H. Umegaki for his valuable advice and constant encouragement. They also thank the referee for his useful comments.

References

[1] C. R. Baker, Capacity of the Gaussian channel without feedback. Inform. and Control, **37** (1978), 70–89.

[2] I. M. Gelfand and A. M. Yaglom, Calculation of the amount of information about a random function contained in another such function. Amer. Math. Soc. Transl., **12** (1959), 199–246.

[3] C. R. Helstrom, J. W. S. Liu and J. P. Gordon, Quantum mechanical communication theory. Proc. IEEE, **58** (1970), 1578–1598.

[4] S. Ihara, On the capacity of the discrete time Gaussian channel with feedback. Trans. Eighth Prague Conf., Vol. C, Czechoslovak Acad. Sci., 1979, 175–186.

[5] R. S. Ingarden, Quantum information theory, Rep. Math. Phys., **10** (1976), 43–73.

[6] S. Kullback and R. A. Leibler, On information and sufficiency, Ann. Math. Statistics, **21** (1953), 79–86.

[7] H. H. Kuo, Gaussian Measures in Banach Spaces. Springer, Berlin, 1975.

[8] M. Ohya, Quantum ergodic channels in operator algebras. J. Math. Anal. Appl., **84** (1981), 318–328.

[9] M. Ohya, On compound state and mutual information in quantum information theory. IEEE Trans. Inform. Theory, **29** (1983), 770–774.

[10] M. Ohya, Note on quantum probability. Lett. Nuovo Cimento, **38** (1983), 402–404.
[11] M. Ohya, Entropy transmission in C^*-dynamical systems. J. Math. Anal. Appl., **100** (1984), 222–235.
[12] M. Ohya, State change and entropies in quantum dynamical systems. Lecture Note in Math., **1136**, Springer, 1985, 397–408.
[13] M. Ohya and N. Watanabe, Construction and analysis of a mathematical model in quantum communication processes. IECE of Japan, J67-A, No. 6 (1984), 548–552.
[14] M. S. Pinsker, Information and Information Stability of Random Variable and Processes. Holden-Day, Inc., 1964.
[15] C. R. Rao and V. S. Varadarajan, Discrimination of Gaussian processes. Sankhya, Ser. A, **25** (1963), 303–330.
[16] C. Shannon, A mathematical theory of communication. Bell System Tech. J., **27** (1948), 379–423, 623–656.
[17] R. Schatten, Norm Ideals of Completely Continuous Operators. Springer-Verlag, Berlin/New York, 1960.
[18] A. V. Skorohod, Integration in Hilbert Space. Springer-Verlag, Berlin, New York, 1974.
[19] H. Takahashi, Information theory of quantum mechanical channels. Advances in Communication Systems. Vol. 1. Academic Press, 1966, 227–310.
[20] M. Takesaki, Theory of Operator Algebra I, Springer-Verlag, New York, 1981.
[21] H. Umegaki, Conditional expectation in an operator algebra. Tôhoku Math. J., **6** (1954), 177–181.
[22] H. Umegaki, Conditional expectation in an operator algebra, III. Kōdai Math. Sem. Rep., **11** (1959), 51–64.
[23] H. Umegaki, Conditional expectation in an operator algebra, IV (entropy and information). Kōdai Math. Sem. Rep., **14** (1962), 59–85.
[24] H. Umegaki, General treatment of alphabet-message space and integral representation of entropy. Kōdai Math. Sem. Rep., **16** (1964), 18–26.
[25] J. von Neumann, Die Mathematischen Grundlagen der Quantenmechanik, Springer, Berlin, 1932.
[26] K. Yanagi, On some properties of Gaussian channels. J. Math. Anal. Appl., **88** (1982), 364–377.