PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 17, Fasc. 1 (1997), pp. 179–196

ON CAPACITIES OF QUANTUM CHANNELS

BY

MASANORI OHYA, DÉNES PETZ* AND NOBORU WATANABE (TOKYO)

Abstract. Capacities of quantum mechanical channels are defined in terms of mutual information quantities. Geometry of the relative entropy is used to express capacity as a divergence radius. The symmetric quantum spin 1/2 channel and the attenuation channel of Boson fields are discussed as examples.

1. Introduction. A discrete communication system — as modeled by Shannon — is capable of transmitting successively symbols of a finite input alphabet $\{x_1, x_2, \ldots, x_m\}$. In the stochastic approach to the communication model it is assumed that the input symbols show up with certain probability. Let p_{ji} be the probability that a symbol x_i is sent over the channel and the output symbol y_j appears at the destination. The joint distribution p_{ji} yields marginal distributions $(p_1, p_2, \ldots, p_m)$ and $(q_1, q_2, \ldots, q_k)$ on the set of input symbols and output symbols, respectively. Shannon introduced the mutual information

$$I = \sum_{i,j} p_{ji} \log \frac{p_{ji}}{p_i q_j} \tag{1.1}$$

in order to measure the amount of information going through the channel.

The interest in quantum communication channels arose in the late 1960'. The scheme of a quantum communication system is not different from a classical one, however, zero point fluctuation (noise) cannot be avoided in quantum systems. Important recent devices for communication are based on optical fiber which is a quantum object. Hence we may assume that the actual signal transmission is over a quantum mechanical medium which is described in the usual Hilbert space formalism of quantum theory. Coding, actual signal transmission and decoding (or measurement) are the main components of the communication chain. The splitting of the communication chain into these three parts C–T–M is somewhat arbitrary. The parts can be investigated individually and their capacity can be defined by means of mutual information.

* On leave from the Technical University Budapest and from the Mathematical Institute of the Hungarian Academy of Sciences. The author was partially supported by JSPS and by OTKA T016924.

Our approach to capacity is based on quantum mutual information which is defined in terms of relative entropy (or informational divergence). Therefore, relative entropy is the basic tool in the paper. The capacity is not compared with performance bounds of classical coding (as in [5] and [13]) because we are mainly interested in the purely quantum part of the channel. Section 2 contains some generalities of quantum communication channels, mutual information and relative entropy. Kholevo's bound is also discussed and we show that it is rarely achievable. In Section 3 our capacities are introduced and the quantum mechanical counterpart of Csiszár's information geometry is used to realize the pure quantum capacity as the divergence radius of the range. The toy example of symmetric quantum spin 1/2 channel is used to demonstrate our ideas. We verify that the pure quantum capacity of this channeling transformation is the same as the performance bound from coding (see [5]). Section 4 treats an infinite-dimensional example, the attenuation channel of Boson fields. The channeling transformation is treated in the abstract Weyl algebra setting as well as in Fock representation. It is proved that the capacity of the attenuation channel is infinite, however, the transmission of arbitrarily much information requires infinite energy.

2. Generalities of quantum mechanical channels. To each input symbol x_i there corresponds a signal state φ_i of the quantum communication system, φ_i functions as the codeword of x_i. The signal states φ_i are mostly pure but they can be non-orthogonal. However, we do not make any assumption on them at this level of generality. The channel state is a convex combination

$$\varphi = \sum_i p_i \varphi_i \tag{2.1}$$

whose coefficients are the corresponding probabilities, p_i is the probability that the letter x_i should be transmitted over the channel.

In the mathematical sense the *quantum channeling transformation* Λ^* is an affine transformation of the state space of the input quantum system into the state space of the output quantum system. (The notation Λ^* is used here because very often Λ^* is the dual mapping of a linear transformation of observables.) At the output some sort of detection scheme retrieves the transmitted information. To each output symbol y_j there corresponds a non-negative observable A_j, that is a self-adjoint operator A_j on the Hilbert space $\mathscr{H}$, such that $\sum_j A_j = I$. (Some people speak about *effects*, or call (A_j) a *generalized measurement*). In terms of the quantum states the transition probabilities are $(\Lambda^*\varphi_i)(A_j)$ and the probability that x_i was sent and y_j is read is

$$p_{ji} = p_i(\Lambda^*\varphi_i)(A_j). \tag{2.2}$$

On the basis of these joint probability distributions the classical mutual information (1.1) is given. Kholevo's theorem provides a fundamental bound for the mutual information in terms of the quantum von Neumann entropy. Before stating

Kholevo's result, we review the simplest entropy quantities used in quantum information theory, for details see [10].

The relative entropy of two states is defined (following Umegaki [14], Lindblad [8] and Araki [1]) as

$$S(\varphi_1, \varphi_2) = \mathrm{Tr}\, D_1 (\log D_1 - \log D_2), \tag{2.3}$$

where D_1 and D_2 are the corresponding density operators. (This formula extends the Kullback–Leibler information measure.) We shall use a kind of algebraic language and view the states as linear functionals on (operator) algebras. The basic property of relative entropy is its monotonicity under channeling transformation. More precisely, if α: $\mathscr{A} \to \mathscr{B}$ is a unitial (completely) positive mapping between the algebras $\mathscr{A}$ and $\mathscr{B}$, that is, the dual α^* is a channeling transformation from the state space of $\mathscr{B}$ into that of $\mathscr{A}$, then

$$S(\varphi_1 \circ \alpha, \varphi_2 \circ \alpha) \leqslant S(\varphi_1, \varphi_2). \tag{2.4}$$

Relative entropy (or *information gain*) is the fundamental information quantity, many other information quantities are expressed by it. For example, the von Neumann entropy is

$$S(\varphi) = -\mathrm{Tr}(D \log D) = \sup \{\sum_j \lambda_j S(\varphi_j, \varphi) \colon \sum_j \lambda_j \varphi_j = \varphi,\ \lambda_j \geqslant 0\}. \tag{2.5}$$

Let α: $\mathscr{A} \to \mathscr{B}$ be positive unitial mapping and φ be a state of $\mathscr{B}$. So φ is an initial state of the channel α^*. The *quantum mutual entropy* is defined after [9] as

$$I(\varphi; \alpha) = \sup \{\sum_j \lambda_j S(\varphi_j \circ \alpha, \varphi \circ \alpha) \colon \sum_j \lambda_j \varphi_j = \varphi\}, \tag{2.6}$$

where the least upper bound is over all orthogonal extremal decompositions. (One checks easily that this formula reduces to (1.1) when $\mathscr{A} = \boldsymbol{C}^k$ and $\mathscr{B} = \boldsymbol{C}^m$ since in this case the orthogonal extremal decomposition is unique and $(p_1, p_2, \ldots, p_m) = \sum_i p_i \delta_i$.)

THEOREM 2.1. *With the above notation the inequality*

$$I = \sum_{i,j} p_{ji} \log \frac{p_{ji}}{p_i q_j} \leqslant S(\Lambda^*(\varphi)) - \sum_i p_i S(\Lambda^*(\varphi_i))$$

holds true.

Kholevo [6] proved this inequality in 1973 when the concept of quantum relative entropy was not well understood yet. Kholevo's upper bound is

$$S(\Lambda^* \varphi) - \sum_i p_i S(\Lambda^* \varphi_i) = \sum_i p_i S(\Lambda^* \varphi_i, \Lambda^* \varphi). \tag{2.7}$$

Let $\boldsymbol{C}^m$, $B(\mathscr{H})$, $B(\mathscr{K})$ and $\boldsymbol{C}^k$ be operator algebras and consider the mappings

$$\alpha \colon \boldsymbol{C}^k \to B(\mathscr{K}), \quad (c_1, c_2, \ldots, c_k) \mapsto \sum_j c_j A_j,$$

$$\Lambda\colon B(\mathcal{H}) \to B(\mathcal{K}),$$

$$\beta\colon B(\mathcal{H}) \to C^m, \qquad H \mapsto (\varphi_1(H), \varphi_2(H), \ldots, \varphi_m(H)).$$

So the duals of the positive unitial mappings α, Λ and β correspond to the measurement, quantum state transmission and coding procedures, respectively. In this terminology the upper bound is the mutual information $I(\mu, \beta\circ\Lambda)$ of the quantum channel α with input quantum input state μ (where $\mu((c_1, c_2, \ldots, c_m)) = \sum_i p_i c_i$) and the classical mutual information I becomes the quantum mutual information $I(\mu, \beta\circ\Lambda\circ\alpha)$ of a composite quantum channel. So Kholevo's theorem is read as

$$I(\mu, \beta\circ\Lambda\circ\alpha) \leqslant I(\mu, \beta\circ\Lambda) \tag{2.8}$$

which follows from the monotonicity of the quantum mutual information. For the details see [10], in particular pp. 139–140. It is noteworthy that these ideas work in the continuous case as well as it was observed also in [15].

Yuen and Ozawa [15] propose to call Theorem 1 the *fundamental theorem of quantum communication*. The theorem bounds the performance of the detecting scheme. We see that in most cases the bound cannot be achieved. Namely, the bound may be achieved in the only case when the output states $\Lambda^*(\varphi_i)$ have commuting densities.

PROPOSITION 2.2. *If the states $\Lambda^*(\varphi_i)$, $1 \leqslant i \leqslant m$, do not commute, then*

$$I = \sum_{i,j} p_{ji} \log \frac{p_{ji}}{p_i q_j} < S(\Lambda^*(\varphi)) - \sum_i p_i S(\Lambda^*(\varphi_i))$$

is a strict inequality.

In the terminology of Chapter 8 of [10] the equality in Kholevo's theorem means that the measurement channel α is sufficient for the states $\Lambda^*(\varphi_i)$, $1 \leqslant i \leqslant m$, and the sufficiency has several characterizations, for example, the existence of states ω_j of the output quantum system such that

$$\sum_j \varphi_i(A_j)\,\omega_j = \varphi_i \qquad \text{for every } i.$$

In particular, if the bound is achieved, the states $\Lambda^*(\varphi_i)$ have to commute. Suppose that the state $\Lambda^*(\varphi_i)$ has a density D_i and let $D = \sum_i p_i D_i$ (which is the density of $\Lambda^*(\varphi)$). Then the generalized measurement $A_i = D_i D^{-1}$ achieves Kholevo's bound. (The operator $D_i D^{-1}$ is well defined even if D is not invertible, because the kernel of D is contained in that of D_i.) The technicalities of the detailed proof depend very much on the level of generality. For finite dimension one has to investigate the equality case in the Jensen inequality and this was carried out by Kholevo [6]. We consider now the infinite-dimensional case but under the restrictive assumption of faithfulness. Since Λ^* does not play any role, we skip it from the notation. The proof of the next theorem uses the idea of the paper [12] and the proof presented here is a bit sketchy. (The interested reader may consult [12] for the more detailed justification of the steps.)

THEOREM 2.3. *Assume that $\varphi = \sum_i p_i \varphi_i$ and there is a sequence (α_n) of generalized measurements such that*

$$\sum_i p_i S(\alpha_n^*(\varphi_i), \alpha_n^*(\varphi)) \to \sum_i p_i S(\varphi_i, \varphi).$$

If the limit is finite and the states φ_i are faithful, then the family (φ_i) must commute.

Proof. Since $S(\alpha_n^*(\varphi_i), \alpha_n^*(\varphi)) \leqslant S(\varphi_i, \varphi)$, the assumption implies that

$$S(\alpha_n^*(\varphi_i), \alpha_n^*(\varphi)) \to S(\varphi_i, \varphi) \quad \text{for every } i.$$

Let D_i, D be the statistical operators of φ_i, φ, respectively. In the sequel we shall use the relative modular operator technique and we work on the Hilbert space $\mathscr{K}$ of Hilbert–Schmidt operators. There exists a positive operator Δ_i such that $\{AD_i^{1/2}\colon A \in B(\mathscr{H})\}$ is a core for $\Delta_i^{1/2}$, and

$$\|\Delta_i^{1/2} A D_i^{1/2}\|^2 = \varphi(AA^*) \quad (A \in B(\mathscr{H})).$$

In fact, the relative modular operator Δ_i is the extension of the linear operator $AD_i^{1/2} \mapsto D^{1/2}A$ defined on a dense linear subspace of $\mathscr{K}$. In terms of the relative modular operator, we have

$$S(\varphi_i, \varphi) = -\langle D_i^{1/2}, (\log\Delta) D_i^{1/2}\rangle = \int_0^\infty \langle D_i^{1/2}, (\Delta_i + t)^{-1} D_i^{1/2}\rangle - (1+t)^{-1}\, dt. \tag{2.9}$$

Similarly,

$$S(\alpha^*(\varphi_i), \alpha^*(\varphi)) = \int_0^\infty \langle d_{n,i}^{1/2}, (\delta_{n,i} + t)^{-1} d_{n,i}^{1/2}\rangle - (1+t)^{-1}\, dt,$$

where the probability vector $d_{n,i}^{1/2}$ (d_n) corresponds to the state $\alpha_n^* \varphi_i$ $(\alpha_n^* \varphi)$ and

$$\delta_{n,i} = d_n / d_{n,i}.$$

(Note that $\delta_{n,i}$ is the relative modular operator of $\alpha_n^* \varphi$ with respect to $\alpha_n^*(\varphi_i)$ but, due to the simple situation coming from finite dimension and commutativity, we may just regard it as a vector.) Since

$$\langle d_{n,i}^{1/2}, (\delta_{n,i} + t)^{-1} d_{n,i}^{1/2}\rangle \leqslant \langle D_i^{1/2}, (\Delta_i + t)^{-1} D_i^{1/2}\rangle \quad \text{for every } t > 0 \text{ and } i,$$

our assumption implies that

$$\langle d_{n,i}^{1/2}, (\delta_{n,i} + t)^{-1} d_{n,i}^{1/2}\rangle \to \langle D_i^{1/2}, (\Delta_i + t)^{-1} D_i^{1/2}\rangle \quad \text{as } n \to \infty.$$

From this we infer that

$$\alpha_n\big((t + \delta_{n,i})^{-1} d_{n,i}^{1/2}\big) \to (t + \Delta_i)^{-1} D_i^{1/2}. \tag{2.10}$$

Now we consider the function $F_i(z) = \Delta_i^z D_i^{1/2}$. We know that $D_i^{1/2}$ is in the domain of $\Delta_i^{1/2}$. The range of $\Delta_i^{1/2}$ contains the operators $D^{1/2}A$ for a bounded A. Since $\|D^{-1/2} D_i^{1/2}\| \leqslant p_i^{-1/2}$, we may choose $A = D^{-1/2} D_i^{1/2}$, and infer that $D_i^{1/2}$ is in the domain of $\Delta_i^{-1/2}$. As a consequence, the function $F_i(z)$ is analytic on the strip $\{z \in \boldsymbol{C}\colon\ -1/2 < \mathrm{Re}\, z < 1/2\}$. We should not gather so much with $f_{n,i}(z) = \delta_{n,i}^z d_{n,i}^{1/2}$, because it is analytic on the whole complex plain.

Our next aim is to show that

$$V_{n,i}\big(f_{n,i}(z)\big) \to F_i(z) \tag{2.11}$$

if $-1/2 < \mathrm{Re}\, z < 1/2$ for the contraction $V_{n,i}$ defined by

$$V_{n,i}(a_n d_{n,i}^{1/2}) = \alpha_n(a_n) D_i^{1/2}.$$

Since we have an analytic function at our disposal, it suffices to prove (2.11) for $0 < s < 1/2$ in place of z. For $0 < s < 1/2$ we obtain

$$\begin{aligned} V_{n,i}\big(f_{n,i}(s)\big) &= \frac{\sin s\pi}{\pi} \int_0^\infty t^{s-1}\, V_{n,i}\big(\delta_{n,i}(t+\delta_{n,i})^{-1}\, d_{n,i}^{1/2}\big)\, dt \\ &= \frac{\sin s\pi}{\pi} \int_0^\infty t^{s-1} \Delta_i (t+\Delta_i)^{-1} D_i^{1/2}\, dt = F_i(s). \end{aligned}$$

So we may consider in (2.11) a pure imaginary $z = \mathrm{i}t$:

$$V_{n,i}\big(F_n(\mathrm{i}t)\big) = V_{n,i}(\delta_{n,i}^{\mathrm{i}t} d_{n,i}^{1/2}) \to \Delta_i^{\mathrm{i}t} D_i^{1/2} = D^{\mathrm{i}t} D_i^{(-\mathrm{i}t+1)/2}$$

and

$$\alpha_n(\delta_{n,i}^{\mathrm{i}t}) D_i^{1/2} \to (D^{\mathrm{i}t} D_i^{-\mathrm{i}t}) D_i^{1/2}$$

in the Hilbert–Schmidt norm. The strong operator convergence follows from the Hilbert–Schmidt norm convergence and we arrive at

$$\alpha_n(\delta_{n,i}^{\mathrm{i}t}) \to (D^{\mathrm{i}t} D_i^{-\mathrm{i}t}) \ \text{(strongly)}. \tag{2.12}$$

In particular, $D^{\mathrm{i}t} D_i^{-\mathrm{i}t}$ is a unitary group for fixed i, and D and D_i must commute. ∎

3. Capacity of channels. Let $\mathscr{H}$ and $\mathscr{K}$ be the input and output Hilbert spaces of a quantum communication system. The *channeling transformation* $\Lambda^*\colon \Sigma(\mathscr{H}) \to \Sigma(\mathscr{K})$ sends density operators acting on $\mathscr{H}$ into those acting on $\mathscr{K}$. A *pseudo-quantum code* is a probability distribution on $\Sigma(\mathscr{H})$ with finite support. So $\{(p_i), (\varphi_i)\}$ is a pseudo-quantum code if (p_i) is a probability vector and φ_i are states of $B(\mathscr{H})$. The quantum states φ_i are sent over the quantum mechanical media, for example, optical fiber, and yield the output quantum states $\Lambda^*\varphi_i$. The performance of coding and transmission is measured by the *mutual information*

$$I\big((p_i), (\varphi_i), \Lambda^*\big) = \sum_i p_i S(\Lambda^*\varphi_i, \Lambda^*\varphi). \tag{3.1}$$

Taking the supremum over certain classes of pseudo-quantum codes, we obtain various capacities of the channel. Here we consider one subclass of pseudo--quantum codes. A *quantum code* is defined by the additional requirement that $\{\varphi_i\}$ is a set of pairwise orthogonal pure states. Correspondingly, we arrive at two alternative concepts of capacity:

(3.2) $$C_{\mathrm{pq}}(\Lambda^*) = \sup\{I((p_i), (\varphi_i), \Lambda): ((p_i), (\varphi_i)) \text{ is a pseudo-quantum code}\}$$

and

(3.3) $$C_{\mathrm{q}}(\Lambda^*) = \sup\{I((p_i), (\varphi_i), \Lambda^*): ((p_i), (\varphi_i)) \text{ is a quantum code}\}.$$

We can write $C_{\mathrm{q}}(\Lambda^*)$ in a slightly different form by using the notation (2.6):

(3.4) $$C_{\mathrm{q}}(\Lambda^*) = \sup\{I(\varphi, \Lambda^*): \varphi \text{ is an input state}\}.$$

The capacity C_{q} may be viewed as the characteristic of the purely quantum mechanical signal transmission.

It follows from the definition that

$$C_{\mathrm{q}}(\Lambda^*) \leqslant C_{\mathrm{pq}}(\Lambda^*)$$

holds for every channel.

EXAMPLE 3.1. Let Λ^* be a channel on the 2×2 density matrices such that

$$\Lambda^*: \begin{pmatrix} a & b \\ \bar{b} & c \end{pmatrix} \mapsto \begin{pmatrix} a & 0 \\ 0 & c \end{pmatrix}.$$

Consider the input density matrix

$$D_\lambda = \frac{1}{2}\begin{pmatrix} 1 & 1-2\lambda \\ 1-2\lambda & 1 \end{pmatrix} \quad (0 < \lambda < 1).$$

For $\lambda \neq 1/2$ the orthogonal extremal decomposition is unique; in fact,

$$D_\lambda = \frac{\lambda}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \frac{1-\lambda}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

and we have

$$I(D_\lambda, \Lambda^*) = 0 \quad \text{for } \lambda \neq 1/2.$$

However, $I(D_{1/2}, \Lambda^*) = \log 2$. Since $C_{\mathrm{q}}(\Lambda^*) \leqslant C_{\mathrm{pq}}(\Lambda^*) \leqslant \log 2$, we conclude that $C_{\mathrm{q}}(\Lambda^*) = C_{\mathrm{pq}}(\Lambda^*) = \log 2$.

The example shows that the quantity $I(\varphi, \Lambda^*)$ may be discontinuous at φ when φ has some degeneracy in the spectrum.

In order to estimate the quantum mutual information, we introduce the concept of *divergence center*. Let $\{\omega_i\colon i \in I\}$ be a family of states and $R > 0$. We say that the state ω is a *divergence center for* $\{\omega_i\colon i \in I\}$ *with* *radius* $\leqslant R$ if

$$S(\omega_i, \omega) \leqslant R \quad \text{for every } i \in I.$$

In the following discussion about the geometry of relative entropy (or divergence as it is called in information theory) the ideas of [3] can be recognized very well.

LEMMA 3.2. *Let* $((p_i), (\varphi_i))$ *be a pseudo-quantum code for the channel* Λ^* *and* ω *be a divergence center with radius* $\leqslant R$ *for* $\{\Lambda^*\varphi_i\}$. *Then*

$$I\big((p_i), (\varphi_i), \Lambda^*\big) \leqslant R.$$

Proof. We assume that the states $\Lambda^*\varphi_i$, $\Lambda^*\varphi = \sum_i p_i \Lambda^*\varphi_i$ and ω have finite entropy and their densities are denoted by D_i, D and D', respectively. We have

$$-S(\Lambda^*\varphi_i) - \operatorname{Tr} D_i \log D' \leqslant R,$$

and hence

$$(3.5) \qquad \sum_i p_i S(\Lambda^*\varphi_i, \Lambda^*\varphi) = -\sum_i p_i S(\Lambda^*\varphi_i) - \operatorname{Tr} D \log D$$
$$\leqslant R - \operatorname{Tr} D(\log D - \log D') = R - S(\Lambda^*\varphi, \omega).$$

The extra assumption we made holds always in finite dimension. When the entropies are not finite but the relative entropies are so, one has to use more sophisticated methods for the proof. It is quite clear that inequality (3.5) is close to equality if $S(\Lambda^*\varphi_i, \omega)$ is about R and $\sum_i p_i \Lambda^*\varphi_i$ is about ω. ∎

Let $\{\omega_i\colon i \in I\}$ be a family of states. We say that the state ω is an *exact divergence center with radius* R if

$$R = \inf_{\varphi} \sup_{i} \{S(\omega_i, \varphi)\}$$

and ω is a minimizer for the right-hand side. (When R is finite, then there exists a minimizer, because $\varphi \mapsto \sup\{S(\omega_i, \varphi)\colon i \in I\}$ is lower semicontinuous with compact level sets; cf. Proposition 5.27 in [10].)

LEMMA 3.3. *Let* ψ_0, ψ_1 *and* ω *be states of* $B(\mathscr{K})$ *such that the Hilbert space* $\mathscr{K}$ *is finite dimensional and set*

$$\psi_\lambda = (1-\lambda)\psi_0 + \lambda\psi_1 \quad (0 \leqslant \lambda \leqslant 1).$$

If $S(\psi_0, \omega)$ *and* $S(\psi_1, \omega)$ *are finite and*

$$S(\psi_\lambda, \omega) \geqslant S(\psi_1, \omega) \quad (0 \leqslant \lambda \leqslant 1),$$

then

$$S(\psi_1, \omega) + S(\psi_0, \psi_1) \leqslant S(\psi_0, \omega).$$

Proof. Let the densities of ψ_λ and ω be D_λ and D, respectively. Due to the assumption $S(\psi_\lambda, \omega) < +\infty$, the kernel of D is smaller than that of D_λ. The function $f(\lambda) = S(\varphi_\lambda, \omega)$ is convex on $[0, 1]$ and $f(\lambda) \geqslant f(1)$ (cf. Proposition 3.1 in [10]). It follows that $f'(1) \leqslant 0$. Hence we have

$$\begin{aligned} f'(1) &= \mathrm{Tr}(D_1 - D_0)(I + \log D_1) - \mathrm{Tr}(D_1 - D_0)\log D \\ &= S(\psi_1, \omega) - S(\psi_0, \omega) + S(\psi_0, \psi_1) \leqslant 0. \end{aligned}$$

This is the inequality we had to obtain.

We note that in the differentiation of the function $f(\lambda)$ the well-known formula

$$\frac{\partial}{\partial t} \mathrm{Tr}\, F(A + tB)|_{t=0} = \mathrm{Tr}\big(F'(A)B\big)$$

can be used. ■

Lemma 3.4. *Let $\{\omega_i \colon i \in I\}$ be a finite set of states of $B(\mathscr{K})$ such that the Hilbert space $\mathscr{K}$ is finite dimensional. Then the exact divergence center is unique and it is in the convex hull on the states ω_i.*

Proof. Let K be the (closed) convex hull of the states $\omega_1, \omega_2, \ldots, \omega_n$ and let ω be an arbitrary state such that $S(\omega_i, \omega) < +\infty$. There is a unique state $\omega \in \mathscr{K}$ such that $S(\omega', \omega)$ is minimal (where ω' runs over K), see Theorem 5.25 in [10]. Then

$$S\big(\lambda\omega_i + (1-\lambda)\omega', \omega\big) \geqslant S(\omega', \omega) \quad \text{for every } 0 \leqslant \lambda \leqslant 1 \text{ and } 1 \leqslant i \leqslant n.$$

It follows from the previous lemma that

$$S(\omega_i, \omega) \geqslant S(\omega_i, \omega').$$

Hence the divergence center of ω_i's must be in K. The uniqueness of the exact divergence center follows from the fact that the relative entropy functional is strictly convex in the second variable. ■

Theorem 3.5. *Let $\Lambda^* \colon \Sigma(\mathscr{H}) \to \Sigma(\mathscr{K})$ be a channel with finite-dimensional $\mathscr{K}$. Then the capacity $C_{\mathrm{pq}}(\Lambda^*)$ is the divergence radius of the range of Λ^*.*

Proof. Let $\big((p_i), (\varphi_i)\big)$ be a pseudo-quantum code. Then $I\big((p_i), (\varphi_i), \Lambda^*\big)$ is at most the divergence radius of $\{\Lambda^*\varphi_i\}$ (according to Lemma 3.2), which is obviously majorized by the divergence radius of the range of Λ^*. Therefore, the capacity does not exceed the divergence radius of the range.

To prove the converse inequality we assume that the exact divergence radius of $\Lambda^*\big(\Sigma(\mathscr{H})\big)$ is larger than $t \in \boldsymbol{R}$. Then we can find $\varphi_1, \varphi_2, \ldots, \varphi_n \in \Sigma(\mathscr{H})$ such that the exact divergence radius R of $\Lambda^*(\varphi_1), \ldots, \Lambda^*(\varphi_n)$ is larger than t. Lemma 3.4 states that the divergence center ω of $\Lambda^*(\varphi_1), \ldots, \Lambda^*(\varphi_n)$ lies in their convex hull K. By possible reordering of the states φ_i we can achieve that

$$S(\Lambda^*(\varphi_i), \omega)\begin{cases} = R & \text{if } 1 \leqslant i \leqslant k, \\ < R & \text{if } i < k \leqslant n. \end{cases}$$

Let K' be the convex hull of $\Lambda^*(\varphi_1), \ldots, \Lambda^*(\varphi_n)$. We claim that $\omega \in K'$; we choose $\omega' \in K'$ such that $S(\omega', \omega)$ is minimal (ω' is running over K'). Then

$$S(\Lambda^*\varphi_i, \varepsilon\omega' + (1-\varepsilon)\omega) < R$$

for every $1 \leqslant i \leqslant k$ and $0 < \varepsilon < 1$, due to Lemma 3.3. However,

$$S(\Lambda^*\varphi_i, \varepsilon\omega' + (1-\varepsilon)\omega) < R$$

for $k \leqslant i \leqslant n$ and for a small ε by a continuity argument. In this way, we conclude that there exists a probability distribution $(p_1, p_2, \ldots, p_k)$ such that

$$\sum_{i=1}^{k} p_i \Lambda^* \varphi_i = \omega, \qquad S(\Lambda^*\varphi_i, \omega) = R.$$

Consider now the pseudo-quantum code $((p_i), (\varphi_i))$ such that

$$\sum_{i=1}^{k} p_i S\big(\Lambda^*\varphi_i, \Lambda^*\big(\sum_{j=1}^{k} p_j \varphi_j\big)\big) = \sum_{i=1}^{k} p_i S(\Lambda^*\varphi_i, \omega) = R.$$

So we have found a pseudo-quantum code which has quantum mutual information larger than t. The channel capacity must exceed the entropy radius of the range. ■

Up to now our discussion has concerned the capacities of coding and transmission, which are bounds for the performance of quantum coding and quantum transmission. After a measurement is performed, the quantum channel becomes classical and Shannon's theory applied. The *total capacity* (or *classical capacity*) of a quantum channel Λ^* is

$$C_{\mathrm{cl}}(\Lambda^*) = \sup\{I((p_i), (\varphi_i), \gamma^* \circ \Lambda^*)\}, \tag{3.6}$$

where the supremum is taken over both all pseudo-quantum codes (p_i), (φ_i) and all measurements γ^*. Due to the monotonicity of the mutual information we have

$$C_{\mathrm{cl}}(\Lambda^*) \leqslant C_{\mathrm{pq}}(\Lambda^*). \tag{3.7}$$

EXAMPLE 3.6. Consider the Stokes parametrization of 2×2 density matrices:

$$D_x = \tfrac{1}{2}(I + x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3),$$

where σ_1, σ_2, σ_3 are the Pauli matrices and $(x_1, x_2, x_3) \in \mathbf{R}^3$ with $x_1^2 + x_2^2 + x_3^2 \leqslant 1$. For a positive semidefinite (3×3)-matrix A the application Γ^*: $D_x \mapsto D_{Ax}$ gives a channeling transformation when $\|A\| \leqslant 1$. This channel was introduced in [5] under the name of *symmetric binary quantum channel.* We want to compute the capacities of Γ^*. Since a unitary conjugation does not obviously change capacity, we may assume that A is diagonal with eigenvalues $1 \geqslant \lambda_1 \geqslant \lambda_2 \geqslant \lambda_3 \geqslant 0$. The range of Γ^* is visualized as an ellipsoid with (Euclidean) diameter $2\lambda_1$. It is not difficult to see that the trace state τ is the exact divergence center of the segment connected the states $(I \pm \lambda_1\sigma_1)/2$, and hence τ must be the divergence center of the whole range. The divergence radius is

$$S\left(\frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}+\frac{\lambda}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \tau\right)$$
$$= \log 2 - S\left(\frac{1}{2}\begin{pmatrix}1+\lambda & 0\\ 0 & 1-\lambda\end{pmatrix}\right) = \log 2 - \eta\left((1+\lambda)/2\right) - \eta\left((1-\lambda)/2\right).$$

This gives the capacity $C_{\mathrm{pq}}(\Gamma^*)$ according to Theorem 3.5. Inequality (3.7) states that the capacity $C_{\mathrm{q}}(\Gamma^*)$ cannot exceed this value. On the other hand,

$$I(\tau, \Gamma^*) = \log 2 - \eta\left((1+\lambda)/2\right) - \eta\left((1-\lambda)/2\right),$$

and we have $C_{\mathrm{pq}}(\Gamma^*) = C_{\mathrm{q}}(\Gamma^*)$. ∎

Shannon's communication theory is largely of asymptotic character, the message length N is supposed to be very large. So we consider the N-fold tensor product of the input and output Hilbert spaces $\mathscr{H}$ and $\mathscr{K}$,

$$\mathscr{H}_N = \bigotimes_{i=1}^{N} \mathscr{H}, \qquad \mathscr{K}_N = \bigotimes_{i=1}^{N} \mathscr{K}.$$

Note that

$$B(\mathscr{H}_N) = \bigotimes_{i=1}^{N} B(\mathscr{H}), \qquad B(\mathscr{K}_N) = \bigotimes_{i=1}^{N} B(\mathscr{K}).$$

The *(multi-) channeling transformation* is a mapping

$$\Lambda_N^*\colon\ \Sigma(\mathscr{H}_N) \to \Sigma(\mathscr{K}_N).$$

The main example is the *memoryless channel*, which is the tensor product of the same *single site channels*:

$$\Lambda_N^* = \Gamma^* \otimes \ldots \otimes \Gamma^* \quad (N\text{-fold}).$$

The sequences $C_{\mathrm{pq}}(\Lambda_N^*)$ and $C_{\mathrm{q}}(\Lambda_N^*)$ of capacities are defined as above for a single channel.

For a memoryless channel the sequences $C_{\mathrm{pq}}(\Lambda_N^*)$ and $C_{\mathrm{q}}(\Lambda_N^*)$ are superadditive. Indeed, if $((p_i), (\varphi_i))$ and $((q_j), (\psi_j))$ are (pseudo-) quantum codes of order N and M, then $((p_i, q_j), (\varphi_i \otimes \psi_j))$ is a (pseudo-) quantum code of order $N+M$ and

$$(3.8) \qquad I\left((p_i, q_j), (\varphi_i \otimes \psi_j), \Lambda_{N+M}^*\right) = I\left((p_i), (\varphi_i), \Lambda_N^*\right) + I\left((q_j), (\psi_j), \Lambda_N^*\right)$$

follows from the additivity of relative entropy under taking tensor product.

One can check that if the initial codes are (pseudo-) quantum, then the product code is (pseudo-) quantum as well. After taking the supremum, the additivity (3.8) yields the superadditivity of the sequences $C_{\mathrm{pq}}(\Lambda_N^*)$ and $C_{\mathrm{q}}(\Lambda_N^*)$. So the following limits exist and they are well known to coincide with the suprema:

$$(3.9)\quad C_{\mathrm{pq}}^{\infty} = \lim \frac{1}{N} C_{\mathrm{pq}}(\Lambda_N^*), \qquad C_{\mathrm{q}}^{\infty} = \lim \frac{1}{N} C_{\mathrm{q}}(\Lambda_N^*), \qquad C_{\mathrm{cl}}^{\infty} = \lim \frac{1}{N} C_{\mathrm{cl}}(\Lambda_N^*).$$

(For multiple channels with some memory effect, one may take the limsup in (3.9) to get a good concept of capacity per single use.) We have

$$(3.10)\qquad C_{\mathrm{cl}}^{\infty}(\Lambda_N^*) \leqslant C_{\mathrm{pq}}^{\infty}(\Lambda_N^*) \leqslant C_{\mathrm{q}}^{\infty}(\Lambda_N^*)$$

for the capacities per single use.

EXAMPLE 3.7. In the case of the memoryless symmetric binary channel we have

$$C_{\mathrm{pq}}^{\infty}(\Gamma^*) = C_{\mathrm{q}}^{\infty}(\Gamma^*) = \log 2 - \eta((1+\lambda)/2) - \eta((1-\lambda)/2),$$

that is the capacity of the single channel coincides with the capacity per single use for the multiple channel.

The proof consists in checking that the trace state remains the divergence center of certain states in the range. Since $\tau = (\varphi_1 + \varphi_2)/2$ for certain output states φ_1, φ_2 such that $S(\varphi_i, \tau)$ is the capacity, we have

$$\tau \otimes \ldots \otimes \tau = 2^{-n} \sum \varphi_{i(1)} \otimes \ldots \otimes \varphi_{i(n)}.$$

Due to symmetry, the trace state is the divergence center, the exact divergence radius is n times $S(\varphi_i, \tau)$ according to the additivity of the relative entropy. This implies that the entropy C_{pq}^{∞} equals the single site one. The argument for C_{q}^{∞} is similar to the single site case.

The work [5] deals with $C_{\mathrm{cl}}(\Lambda^*)$ in detail, and, among other things, a coding theorem relates C_{cl}^{∞} to the code rate of a sequence of pseudo-quantum codes and measurements with asymptotically vanishing average error probability. The picture looks rather similar to Shannon's coding theorem. (Note that in [5] our capacity C_{pq} was called *pseudo-capacity* because the authors were interested in the classical capacity.)

The relations among C_{pq}, C_{q} and C_{cl} form an important problem, worthy of study. For a noiseless channel, $C_{\mathrm{cl}} = \log n$ was obtained in [5], where n is the dimension of the output Hilbert space (actually identical to the input one). Since the trace state is the exact divergence center of all density matrices, we have $C_{\mathrm{pq}} = \log n$ and also $C_{\mathrm{q}} = \log n$. We expect that $C_{\mathrm{pq}} < C_{\mathrm{cl}}$ for "truly quantum mechanical channels" but $C_{\mathrm{cl}}^{\infty} = C_{\mathrm{pq}}^{\infty} = C_{\mathrm{q}}^{\infty}$ must hold for a large class of memoryless channels.

In the case of the binary symmetric channel, all the three capacities coincide as computed in Example 3.7 and in [5].

4. The attenuation channel. First we discuss the attenuation channel in the context of the Weyl algebra in a representation free way. It will turn out somewhat later that what we are describing is identical to the attenuation channel defined in terms of the bosonic Fock space in [9].

Let σ be a non-degenerate symplectic form on a linear space $\mathscr{H}$. Typically, $\mathscr{H}$ is a complex Hilbert space and $\sigma(f, g) = -\mathrm{Im}\langle f, g\rangle$. The Weyl algebra $\mathrm{CCR}(\mathscr{H})$ is generated by unitaries $\{W(f): f \in \mathscr{H}\}$ satisfying the Weyl form of the canonical commutation relation:

$$W(f)\,W(g) = e^{\mathrm{i}\sigma(f,g)}\,W(f+g) \qquad (f, g \in \mathscr{H}). \tag{4.1}$$

Since the linear hull of the unitaries $W(f)$ is dense in $\mathrm{CCR}(\mathscr{H})$, any state is determined uniquely by its values taken on the Weyl unitaries. The most important state of the Weyl algebra is the Fock state which is given as

$$\varphi(W(f)) = \exp\{-\|f\|^2/2\} \qquad (f \in \mathscr{H}). \tag{4.2}$$

The GNS Hilbert space corresponding to the Fock state is called the *(bosonic) Fock space* $\Gamma(\mathscr{H})$ and the cyclic vector Φ is said to be a *vacuum*. The states

$$\varphi_f(\cdot) = \varphi(W(f)^* \cdot W(f)) \tag{4.3}$$

are called *coherent states* and they are induced by the coherent vectors

$$\pi_{\mathrm{F}}(W(f))\Phi = \Phi_f$$

in the Fock representation π_{F}. We have

$$\begin{aligned} \langle \Phi_f, \Phi_g\rangle &= \varphi(W(f)^*\,W(g)) = \exp\{-\tfrac{1}{2}\|g-f\|^2\}\exp\{-\mathrm{i}\sigma(f, g)\} \\ &= \exp\{-\tfrac{1}{2}(\|f\|^2+\|g\|^2)+\langle f, g\rangle\}, \end{aligned} \tag{4.4}$$

and

$$\begin{aligned} \varphi_f(W(g)) &= \exp\{-\tfrac{1}{2}\|g\|^2 - 2\mathrm{i}\sigma(f, g)\} \\ &= \exp\{-\tfrac{1}{2}\|g\|^2 + 2\mathrm{i}\,\mathrm{Im}\langle f, g\rangle\} \qquad (f, g \in \mathscr{H}). \end{aligned} \tag{4.5}$$

The field operators are obtained as the generators of the unitary groups $t \mapsto \pi_{\mathrm{F}}(W(tf))$ in the Fock representation. In other words, $B(f)$ is an unbounded self-adjoint operator on $\Gamma(\mathscr{H})$ such that

$$B(f) = -\mathrm{i}\frac{d}{dt}\pi_{\mathrm{F}}(W(tf))|_{t=0}$$

with an appropriate domain. The creation and annihilation operators are defined as

$$a^*(f) = \tfrac{1}{2}(B(\mathrm{i}f) - \mathrm{i}B(f)), \qquad a(f) = \tfrac{1}{2}(B(\mathrm{i}f) + \mathrm{i}B(f)).$$

The positive self-adjoint operator $N(f) = a^*(f)a(f)$ has spectrum $\boldsymbol{Z}^+$ and it is called the *particle number operator* (for the "f-mode").

Let T be a symplectic transformation of $\mathscr{H}$ to $\mathscr{H} \oplus \mathscr{K}$, i.e., $\sigma(f, g) = \sigma(Tf, Tg)$. Then there is a homomorphism

$$\alpha_T\colon \mathrm{CCR}(\mathscr{H}) \to \mathrm{CCR}(\mathscr{H} \oplus \mathscr{K})$$

such that

$$\alpha_T(W(f)) = W(Tf). \tag{4.6}$$

We may regard the Weyl algebra $\mathrm{CCR}(\mathscr{H} \oplus \mathscr{K})$ as $\mathrm{CCR}(\mathscr{H}) \otimes \mathrm{CCR}(\mathscr{K})$ and, given a state ψ on $\mathrm{CCR}(\mathscr{K})$, a channeling transformation arises as

$$(\Lambda^*\omega)(A) = (\omega \otimes \psi)(\alpha_T(A)), \tag{4.7}$$

where the input state ω is an arbitrary state of $\mathrm{CCR}(\mathscr{H})$ and $A \in \mathrm{CCR}(\mathscr{H})$. (In the language of optical communication, ψ is called a *noise state.*) To see a concrete example discussed in [9], we choose $\mathscr{H} = \mathscr{K}$, $\psi = \varphi$ and

$$S(\xi) = a\xi \oplus b\xi. \tag{4.8}$$

If $|a|^2 + |b|^2 = 1$ holds for the numbers a and b, this S is an isometry and a symplectic transformation, and we arrive at the channeling transformation

$$(\Lambda^*\omega)\, W(g) = \omega(W(ag)) \exp\{-\tfrac{1}{2}\|bg\|^2\} \qquad (g \in \mathscr{H}). \tag{4.9}$$

In order to have an alternative description of Λ^* in terms of density operators acting on $\Gamma(\mathscr{H})$ we introduce the linear operator $V\colon \Gamma(\mathscr{H}) \to \Gamma(\mathscr{H}) \otimes \Gamma(\mathscr{H})$ defined by

$$V\pi_{\mathrm{F}}(A)\Phi = \pi_{\mathrm{F}}(\alpha_T(A))\Phi \otimes \Phi.$$

We have

$$V\pi_{\mathrm{F}}(W(f))\Phi = (\pi_{\mathrm{F}}(W(af)) \otimes \pi_{\mathrm{F}}(W(bf)))\,\Phi \otimes \Phi,$$

and hence

$$V\Phi_f = \Phi_{af} \otimes \Phi_{bf}. \tag{4.10}$$

LEMMA 4.1. *Let ω be a state of* $\mathrm{CCR}(\mathscr{H})$ *which has density D in the Fock representation. Then the output state $\Lambda^*\omega$ of the attenuation channel has density* $\mathrm{Tr}_2\, VDV^*$ *in the Fock representation.*

Proof. Since we work only in the Fock representation, we skip π_{F} in the formulas. First we show that

$$V^*(W(f) \otimes I)\, V = W(af) \exp\{-\tfrac{1}{2}\|bf\|^2\} \tag{4.11}$$

for every $f \in \mathscr{H}$. (This can be done by computing the quadratic form of both operators on coherent vectors.) Now we proceed as follows:

$$\operatorname{Tr}(\operatorname{Tr}_2 VDV^*)W(f) = \operatorname{Tr} VDV^*(W(f)\otimes I) = \operatorname{Tr} DV^*(W(f)\otimes I)V$$
$$= \operatorname{Tr} DW(af)\exp\{-\tfrac{1}{2}\|bf\|^2\} = \omega(W(af))\exp\{-\tfrac{1}{2}\|bf\|^2\},$$

which is nothing else but $(\Lambda^*\omega)(W(f))$ due to (4.9). ■

The lemma states that Λ^* is really the same (attenuation) channel discussed in [9] or [10], p. 305.

We note that Λ is a so-called quasi-free completely positive mapping of $\mathrm{CCR}(\mathscr{H})$ given as

(4.12) $$\Lambda(W(f)) = W(af)\exp\{-\tfrac{1}{2}\|bf\|^2\}$$

(cf. [4] or Chapter 8 of [11]).

PROPOSITION 4.2. *If ψ is a regular state of* $\mathrm{CCR}(\mathscr{H})$, *that is $t\mapsto\psi(W(tf))$ is a continuous function on $\mathbf{R}$ for every $f\in\mathscr{H}$, then $(\Lambda^*)^n(\psi)\to\varphi$ pointwise. (φ denotes the Fock state.)*

Proof. It is enough to look at the formula

$$((\Lambda^*)^n\psi)(W(f)) = \varphi(W(a^nf))\exp\{-\tfrac{1}{2}\sum_{k=0}^{n-1}\|a^kbf\|^2\}$$

and the statement is concluded. ■

It is worth noting that the singular state

$$\tau(W(f)) = \begin{cases} 0 & \text{if } f\neq 0,\\ 1 & \text{if } f=0\end{cases}$$

is an invariant state of $\mathrm{CCR}(\mathscr{H})$. On the other hand, the proposition applies to states with density operator in the Fock representation. Therefore, we have

COROLLARY 4.3. *Λ^* regarded as a channel of $B(\Gamma(\mathscr{H}))$ has a unique invariant state, the Fock state, and correspondingly Λ is ergodic.*

Λ is not only ergodic but it is *completely dissipative* in the sense that

(4.13) $$\Lambda(A^*A) = \Lambda(A^*)\Lambda(A)$$

may happen only in the trivial case when A is a multiple of the identity. The authors are grateful to M. Fannes and A. Verbeure for this information (private communication). In fact,

(4.14) $$\Lambda = (\mathrm{id}\otimes\omega)\circ\alpha_S,$$

where α_S is given by (4.6) and (4.8), and $\omega(W(f)) = \exp\{-\|bf\|^2\}$ is a quasi-free state. Here $\mathrm{id}\otimes\omega$ is just a conditional expectation which leaves invariant a separating product state.

LEMMA 4.4. *Let Λ^* be the attenuation channel. Then*

$$\sup I((p_i),(\varphi_i),\Lambda^*) = \log n$$

when the supremum is taken over all pseudo-quantum codes $((p_i)_{i=1}^n, (\varphi_{f(i)})_{i=1}^n)$ *applying n coherent states.*

Proof. We know that $\Lambda^*\varphi_f = \varphi_{af}$, so the output $\{\Lambda^*\varphi_{f(1)}, \ldots, \Lambda^*\varphi_{f(n)}\}$ consists of n pure states. The corresponding vectors of $\Gamma(\mathscr{H})$ span a Hilbert space of dimension $k \leqslant n$. Since the trace state on that Hilbert space is a divergence center with radius $\leqslant \log k \leqslant \log n$, $\log n$ is always a bound for the mutual information according to Lemma 3.2.

In order to show that the bound $\log n$ is really achieved we choose the vectors $f(k)$ such that

$$f(k) = \lambda kf \quad (1 \leqslant k \leqslant n),$$

where $f \in \mathscr{H}$ is a fixed non-zero vector. Then in the limit $\lambda \to \infty$ the states $\varphi_{f(k)}$ become orthogonal, since

$$|\langle \Phi_{\lambda kf} \mid \Phi_{\lambda mf} \rangle|^2 = \exp\{-\lambda^2 (k-m)^2 \|f\|^2/2\} \to 0 \tag{4.15}$$

whenever $k \neq m$. In the limit $\lambda \to \infty$ the trace state (of a subspace) becomes the exact divergence center and we have

$$\lim_{\lambda \to \infty} I((1/n), (\varphi_{f(i)}), \Lambda^*) = \log n.$$

This proves the lemma. ■

The next theorem follows directly from the previous lemma.

THEOREM 4.5. *The capacity* C_{pq} *of the attenuation channel is infinite.*

Some remarks are in order. Since the argument of the proof of Lemma 4.4 works for any quasi-free channel, we can conclude $C_{\mathrm{pq}} = \infty$ also in that more general case. Another remark concerns the classical capacity C_{cl}. Since the states $\varphi_{f(n)}$ used in the proof of Lemma 4.4 commute in the limit $\lambda \to \infty$, the total capacity C_{cl} is infinite as well. $C_{\mathrm{cl}} = \infty$ follows also from the proof of the next theorem.

THEOREM 4.6. *The capacity* C_{q} *of the attenuation channel is infinite.*

Proof. We follow the strategy of the proof of the previous theorem, but we use the number states in place of the coherent ones. The attenuation channel sends the number state $|n\rangle\langle n|$ into the binomial mixture of the number states

$$|0\rangle\langle 0| = \Phi,\ |1\rangle\langle 1|, \ldots, |n\rangle\langle n|.$$

Hence the commuting family of convex combination of number states is invariant under the attenuation channel, and the channel restricted to those states is classical with obviously infinite capacity. Since C_{q} (as well as C_{cl}) cannot have a smaller value, the claim follows. ■

Let us make some comments on the previous results. The theorems mean that arbitrarily large amount of information can go through the attenuation channel, however the theorems do not say anything about the price for it. The expectation value of the number of particles needed in the pseudo-quantum code of Lemma 4.4 tends to infinity. Indeed,

$$\sum_i \frac{1}{n} \varphi_{f(i)}(N) = \frac{1}{n} \sum_{i=1}^{n} \|f(i)\|^2 = \lambda(n+1)(2n+1)\|f\|^2/6,$$

which increases rapidly with n (here N denotes the number operator). Hence the good question is to ask for the capacity of the attenuation channel when some energy constraint is posed:

$$C(E_0) = \sup\{I((p_i), (\varphi_i), \Lambda^*)\colon \sum_i p_i \varphi_i(N) \leqslant E_0\}. \tag{4.16}$$

(To be more precise, we have posed a bound on the average energy, different constraints are also possible, cf. [2].) Since $\Lambda(N) = a^2 N$ for the number operator N, we have

$$C(E_0) = \sup\{\sum_i p_i S(\varphi_i, \sum_j p_j \varphi_j)\colon \sum_i p_i \varphi_i(N) \leqslant a^{-2} E_0\}. \tag{4.17}$$

The solution of this problem is the same as that of

$$\sup\{S(\psi)\colon \psi(N) = a^2 E_0\}$$

and the well-known maximizer of this problem is a so-called *Gibbs state*. Therefore, we have

$$C(E_0) \leqslant a^2 E_0 + \log(a^2 E_0 + 1). \tag{4.18}$$

This value can be realized as a classical capacity if the number states can be output states of the attenuation channel.

REFERENCES

[1] H. Araki, *Relative entropy for states of von Neumann algebras*, Publ. Res. Inst. Math. Sci., Kyoto Univ., 11 (1976), pp. 809–833.

[2] C. M. Caves and P. D. Drummond, *Quantum limits on bosonic communication rates*, Rev. Modern Phys. 66 (1994), pp. 481–537.

[3] I. Csiszår, *I-divergence geometry of probability distributions and minimization problems*, Ann. Probab. 3 (1975), pp. 146–158.

[4] B. Demoen, P. Vanheuverzwijn and A. Verbeure, *Completely positive maps on the CCR-algebra*, Lett. Math. Phys. 2 (1977), pp. 161–166.

[5] A. Fujiwara and H. Nagaoka, *Capacity of memoryless quantum communication channels*, Math. Eng. Tech. Rep. 94–22, University of Tokyo, 1994.

[6] A. S. Kholevo, *Some estimates for the amount of information transmittable by a quantum communication channel*, Problemy Peredachi Informatsii 9 (1973), pp. 3–11.

[7] — *Capacity of a quantum communication channel*, Problems Inform. Transmission 15 (1979), pp. 247–253.

[8] G. Lindblad, *Completely positive maps and entropy inequalities*, Comm. Math. Phys. 40 (1975), pp. 147–151.

[9] M. Ohya, *On compound state and mutual information in quantum information theory*, IEEE Trans. Inform. Theory 29, pp. 770–777.

[10] — and D. Petz, *Quantum Entropy and Its Use*, Springer, 1993.

[11] D. Petz, *The Algebra of the Canonical Commutation Relation*, Leuven University Press, 1990.

[12] — *Discrimination between states of a quantum system by observations*, J. Funct. Anal. 120 (1994), pp. 82–97.

[13] B. Schumacher, *Quantum coding*, Phys. Rev. A (1995), pp. 2738–2747.

[14] H. Umegaki, *Conditional expectations in an operator algebra IV (entropy and information)*, Kodai Math. Sem. Rep. 14 (1962), pp. 59–85.

[15] H. P. Yuen and M. Ozawa, *Ultimate information carrying limit of quantum systems*, Phys. Rev. Lett. 70 (1993), pp. 363–366.

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

Received on 9.5.1996