KOLMOGOROV'S ε-ENTROPY OF SOME GAUSSIAN PROCESSES

KIMIO KAZI

(Received May 18, 1967)

1. Introduction and definitions

Since C. Shannon first introduced the notion of information, various definitions were presented about the amount of information of random variables in probability theory. Among them, however, A. N. Kolmogorov's definition seems efficient for analysis of stochastic processes.

Let X and Y be random variables on the probability space $(\Omega, \mathcal{B}, P)$, taking values in $\mathcal{X}$ and $\mathcal{Y}$, respectively. The amount of information contained in Y with respect to X is then

$$I(X, Y)=\iint_{\mathcal{X}\times\mathcal{Y}} P_{XY}(dxdy)\log\frac{P_{XY}(dxdy)}{P_X(dx)P_Y(dy)} \tag{1.1}$$

where $P_X(\ )$ and $P_Y(\ )$ are the probability distributions of X and Y, respectively, and $P_{XY}(\ )$ is the joint distribution of X and Y. The above integral is to be understood in the Lebesgue-Stieltjes sense. $I(X, Y)$ has the basic properties that we want to require for the definition of the "amount of information". When we put $X=Y$, we call $I(X, X)$ entropy of X and denote with $H(X)$.

$$H(X)=I(X, X)=-\int_{\mathcal{X}} P_X(dx)\log P_X(dx) \tag{1.2}$$

Hereafter we assume that $\mathcal{X}=\mathcal{Y}$ and it is a metric space with distance function ρ.

When we calculate the amount of information $I(X, Y)$ or entropy $H(X)$, we often get infinite values. (In order that $I(X, Y)$ be finite, it is necessary that P_{XY} is absolutely continuous with respect to $P_X\times P_Y$.) A. N. Kolmogorov gave a further definition of "ε-entropy of a random variable X" from a viewpoint of transmission of information, that is,

$$\mathcal{H}_\varepsilon(X)=\inf I(X, Y)\,, \tag{1.3}$$

where the infimum is taken over all Y's satisfying

$$\text{(1.4)} \qquad \mathrm{E}\,|\rho(X, Y)|^2=\int_\Omega |\rho(X(\omega), Y(\omega))|^2\,\mathrm{P}(d\omega)\leqq\varepsilon^2 .$$

It is the least amount of information contained in Y with respect to X when we use Y in place of X under the condition that the mean square error is less than ε^2. It is related to C. Shannon's "rate of generating of messages". Kolmogorov stressed the importance of evaluating the asymptotic order of the quantity $\mathscr{H}_\varepsilon(X)$ when ε is close to 0.

When $\mathscr{X}=R^n$ (n-dimensional Euclidean space with the ordinary distance), $\mathscr{H}_\varepsilon(X)$ is calculated as:

$$\text{(1.5)} \qquad \mathscr{H}_\varepsilon(X)=n\log\frac{1}{\varepsilon}+[h(X)-n\log\sqrt{2\pi e}\,]+O(\varepsilon)$$

where

$$h(X)=-\int_{R^n} p(x)\log p(x)dx$$

$p(x)$: probability density function of X in R^n

for sufficiently smooth $p(x)$.

The ε-entropy of the real-valued stochastic process $\{X(t);\ 0\leqq t\leqq T\}$ defined in the time interval $[0, T]$ is

$$\text{(1.3)}' \qquad \mathscr{H}_\varepsilon(\{X(t)\})=\inf I(\{X(t)\}, \{Y(t)\}) ,$$

where the infimum is taken over all the processes $\{Y(t);\ 0\leqq t\leqq T\}$ that satisfy the condition

$$\text{(1.4)}' \qquad \mathrm{E}\int_0^T |X(t)-Y(t)|^2dt\leqq\varepsilon^2 .$$

The ε-entropy of Gaussian random variables is calculated in exact form [7] and the ε-entropy of a Gaussian process can be evaluated rather easily. In [4] the following fact is noted without proof: let $X(t)$ be such a process that satisfies the differential equation in the symbolic sense:

$$\text{(1.6)} \qquad X^{(N)}(t)+a_1X^{(N-1)}(t)+\cdots+a_NX(t)=aB'(t) ,$$

where $a_1, a_2, \cdots, a_N$, $a\neq0$ are real constants and $B'(t)$ is the formal derivative of the Brownian motion $B(t)$, then the ε-entropy of $\{X(t);\ 0\leqq t\leqq T\}$ is given asymptotically by

$$\text{(1.7)} \qquad \mathscr{H}_\varepsilon(\{X(t)\})=C\,\varepsilon^{-2/(2N-1)}+o(\varepsilon^{-2/(2N-1)})$$

as $\varepsilon\to0$. C is a positive constant depending on T, N and $a_1, a_2, \cdots, a_N, a$. The Gaussian process $X(t)$ satisfying (1.6) is an N-ple Markov process in Doob's sense.

In this paper we shall show a similar evaluation of the ε-entropy for such an N-ple Markov Gaussian process that is represented in the form:

$$X(t)=\int_0^t \sum_{i=1}^{N} f_i(t)g_i(u)dB(u) \qquad (0\leqq t\leqq T) \tag{1.8}$$

under certain conditions. As to the general consideration of such a Gaussian process see T. Hida [5].

2. Some known results

Let $\{X(t);\ 0\leqq t\leqq T\}$ be a real-valued Gaussian process with mean $\mathrm{E}\{X(t)\}=0$ whose correlation function

$$K(t, s)=\mathrm{E}\{X(t)X(s)\} \qquad 0\leqq t,\ s\leqq T$$

is continuous in $0\leqq t,\ s\leqq T$. Since $K(t, s)$ is symmetric:

$$K(t, s)=K(s, t)$$

and positive semi-definite:

$$\int_0^T\int_0^T K(t, s)\varphi(s)\varphi(t)dsdt\geqq 0 \qquad \text{for every continuous function } \varphi(t),$$

the integral operator K with the kernel $K(t, s)$

$$(K\varphi)(t)=\int_0^T K(t, s)\varphi(s)ds$$

is a self-adjoint completely continuous operator, so there are a countable number of eigenvalues of the integral equation

$$K\varphi=\lambda\varphi \tag{2.1}$$

and they are all non-negative, and can be arranged so that $\lambda_1\geqq\lambda_2\geqq\cdots\geqq\lambda_n\to 0$ as n increases. It is verified that K is a trace operator, i.e. $\sum_n \lambda_n<+\infty$, because of continuity of $K(t, s)$.

The ε-entropy of $\{X(t);\ 0\leqq t\leqq T\}$ is then calculated exactly as follows:

$$\mathscr{H}_\varepsilon(\{X(t)\})=\frac{1}{2}\sum_n \log\left(\frac{\lambda_n}{\theta^2}\vee 1\right)^* \tag{2.2}$$

where θ^2 is a constant determined by the equation:

$$\sum_n (\lambda_n\wedge\theta^2)=\varepsilon^2\,^*.$$

Especially, if the eigenvalues of the integral equation (2.1) are evaluated

* $a\vee b=\max\{a, b\}$ $\quad a\wedge b=\min\{a, b\}$

asymptotically as

$$\lambda_n = Cn^{-k} + o(n^{-k}), \tag{2.3}$$

then we calculate $\mathcal{H}_\varepsilon(\{X(t)\})$ on the basis of (2.2):

$$\mathcal{H}_\varepsilon(\{X(t)\}) = \frac{1}{2} C^{1/(k-1)} k^{k/(k-1)} (k-1)^{-1/(k-1)} \varepsilon^{-2/(k-1)} + o(\varepsilon^{-2/(k-1)}) \tag{2.4}$$

as $\varepsilon \to 0$.

When we take the Brownian motion $B(t)$ $(0 \leqq t \leqq T)$,

$$K(t, s) = t \wedge s,$$

and

$$\lambda_n = \frac{4T^2}{\pi^2(1+2n)^2} = \frac{T^2}{\pi^2} \frac{1}{n^2} + o\left(\frac{1}{n^2}\right).$$

Applying (2.3) and (2.4), we have

$$\mathcal{H}_\varepsilon(\{B(t)\}) = \frac{2T^2}{\pi^2} \frac{1}{\varepsilon^2} + o\left(\frac{1}{\varepsilon^2}\right)$$

as $\varepsilon \to 0$.

The above results are due to Pinsker [7].

Thus, in order to calculate the ε-entropy of a Gaussian process with a continuous correlation function, we need to know the asymptotic order of the nth largest eigenvalue λ_n of the integral equation (2.1). Relying on (2.3) and (2.4), we shall evaluate eigenvalues about such a Gaussian process as (1.8).

3. Evaluation of the ε-entropy

We shall start with a process $X(t)$ which is expressed in the form:

$$X(t) = \int_0^t \sum_{i=1}^N f_i(t) g_i(u) dB(u) \qquad 0 \leqq t \leqq T. \tag{3.1}$$

Following [5]*, we assume

ASSUMPTION 1. f_i, g_i $(i=1, 2, \cdots, N)$ are differentiable in $[0, T]$ as many times as necessary for the following argument; f_i and the Wronskian $W(g_1, g_2, \cdots, g_i)$ never vanish in $[0, T]$ for every i.

* The original assumptions 1 and 2 in [5] do not contain $t=0$. See section 4 about the situation that we here include $t=0$ in assumptions 1 and 2.

ASSUMPTION 2.

$$\sum_{i=1}^{N} f_i^{(k)}(t)g_i(t)\equiv 0 \qquad (k=0, 1, \cdots, N-2) \tag{3.2}$$

and

$$\sum_{i=1}^{N} f_i^{(N-1)}(t)g_i(t)\neq 0 \tag{3.3}$$

for any $t\in[0, T]$.

Under these assumptions $X(t)$ is differentiable with respect to $L^2(\Omega)$ norm up to $N-1$ times and satisfies the stochastic differential equation of the type (1.6) where $a_1, a_2, \cdots, a_N$, a are functions of t in this case. $X(t)$ is an N-ple Markov Gaussian process in the restricted Lévy sense.

The correlation function $K(t, s)$ of $X(t)$ is written:

$$K(t, s)=\sum_{i=1}^{N} f_i(t)h_i(s) \qquad t\geqq s\,, \tag{3.4}$$

where

$$h_i(s)=\sum_{j=1}^{N} f_j(s)\int_0^s g_i(u)g_j(u)du \qquad (i=1, 2, \cdots, N)\,. \tag{3.5}$$

This correlation function is continuous in $0\leqq s,\ t\leqq T$. Putting $A_{ij}(s)=\int_0^s g_i(u)g_j(u)du$ and using relations (3.2), we get

$$h_i^{(k)}(s)=\sum_{j=1}^{N} f_j^{(k)}(s)A_{ij}(s) \qquad (i=1, 2, \cdots, N;\ k=0, 1, \cdots, N-1)\,. \tag{3.6}$$

From the assumption 1, we know $g_1, g_2, \cdots, g_N$ are linearly independent functions. We note that $f_1, f_2, \cdots, f_N$ are also linearly independent from the assumption 2, because we obtain

$$W(f_1, f_2, \cdots, f_N)\cdot W(g_1, g_2, \cdots, g_N)=\left\{\sum_{i=1}^{N} f_i^{(N-1)}(s)g_i(s)\right\}^N\neq 0$$

by writing down relations (3.2) and (3.3) in the product form of matrices and taking determinants of both sides. Taking account of non-singularity of the Gram's matrix $\{A_{ij}(s)\}_{\substack{1\leqq i\leqq N\\ 1\leqq j\leqq N}}$, we obtain from (3.6)

$$W(h_1, h_2, \cdots, h_N)=\det\{A_{ij}\}\, W(f_1, f_2, \cdots, f_N)\neq 0$$

and hence $h_1, h_2, \cdots, h_N$ are also linearly independent. Moreover, $2N$ functions $f_1, f_2, \cdots, f_N,\ h_1, h_2, \cdots, h_N$ are also linearly independent since we get

$$(3.7)\qquad \begin{aligned} &W(f_1, f_2, \cdots, f_N,\ h_1, h_2, \cdots, h_N) \\ &= \left\{ \sum_{i=1}^{N} f_i^{(N-1)}(s) g_i(s) \right\}^N W(f_1, f_2, \cdots, f_N) W(g_1, g_2, \cdots, g_N) \\ &= \left\{ \sum_{i=1}^{N} f_i^{(N-1)}(s) g_i(s) \right\}^{2N} \neq 0 \end{aligned}$$

according to (3.6). From (3.6) it also follows that

$$(3.8)\qquad \sum_{i=1}^{N} \{ f_i^{(k)}(s) h_i^{(l)}(s) - f_i^{(l)}(s) h_i^{(k)}(s) \} = 0 \qquad \text{for } 0 \leqq k+l \leqq 2N-2 .$$

For later use we compute especially the case $k=2N-1$, $l=0$ in (3.8):

$$\begin{aligned} &\sum_{i=1}^{N} \{ f_i^{(2N-1)}(s) h_i(s) - f_i(s) h_i^{(2N-1)}(s) \} \\ &= \frac{d}{ds} \sum_{i=1}^{N} \{ f_i^{(2N-2)}(s) h_i(s) - f_i(s) h_i^{(2N-2)}(s) \} \\ &\qquad\qquad - \sum_{i=1}^{N} \{ f_i^{(2N-2)}(s) h_i^{(1)}(s) - f_i^{(1)}(s) h_i^{(2N-2)}(s) \} \\ &= -\sum_i \{ f_i^{(2N-2)} h_i^{(1)} - f_i^{(1)} h_i^{(2N-2)} \} \\ &= +\sum_i \{ f_i^{(2N-3)} h_i^{(2)} - f_i^{(2)} h_i^{(2N-3)} \} \\ &= \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \end{aligned}$$

by using (3.8) repeatedly for $k+l=2N-2$, we obtain

$$= (-1)^{N-1} \sum_i \{ f_i^{(N)} h_i^{(N-1)} - f_i^{(N-1)} h_i^{(N)} \}$$

and from (3.6),

$$\begin{aligned} &= (-1)^{N-1} \sum_i \{ f_i^{(N)} \sum_j A_{ij} f_j^{(N-1)} - f_i^{(N-1)} \sum_j (f_j^{(N)} A_{ij} + f_j^{(N-1)} g_i g_j) \} \\ &= (-1)^{N-1} \sum_{i,j} \{ (-1) f_i^{(N-1)} f_j^{(N-1)} g_i g_j \} \\ &= (-1)^N \{ \sum_i f_i^{(N-1)}(s) g_i(s) \}^2 . \end{aligned}$$

Thus we have

$$(3.9)\qquad \sum_{i=1}^{N} \{ f_i^{(2N-1)}(s) h_i(s) - f_i(s) h_i^{(2N-1)}(s) \} = (-1)^N \left\{ \sum_{i=1}^{N} f_i^{(N-1)}(s) g_i(s) \right\}^2 .$$

Now we want to study the eigenvalue problem of the integral equation corresponding to (2.1) with the continuous kernel (3.4):

$$(3.10)\qquad \lambda \varphi(t) = \sum_{i=1}^{N} \{ f_i(t) A_i(t) + h_i(t) B_i(t) \} ,$$

where

$$A_i(t)=\int_0^t h_i(s)\varphi(s)ds\,,\qquad B_i(t)=\int_t^T f_i(s)\varphi(s)ds\,.$$

This eigenvalue problem (3.10) is turned into that of the classical Sturm-Liouville type of $2N$th order differential equation in the following way. Differentiating both sides of (3.10) and using the equalities (3.8) repeatedly, we get

$$(3.11)\qquad \begin{cases} \lambda\varphi(t)=\sum_i \{f_i(t)A_i(t)+h_i(t)B_i(t)\} \\ \lambda\varphi^{(1)}(t)=\sum_i \{f_i^{(1)}(t)A_i(t)+h_i^{(1)}(t)B_i(t)\} \\ \quad\cdots\cdots\cdots\cdots\cdots \\ \lambda\varphi^{(N)}(t)=\sum_i \{f_i^{(N)}(t)A_i(t)+h_i^{(N)}(t)B_i(t)\} \\ \quad\cdots\cdots\cdots\cdots\cdots \\ \lambda\varphi^{(2N-1)}(t)=\sum_i \{f_i^{(2N-1)}(t)A_i(t)+h_i^{(2N-1)}(t)B_i(t)\} \\ \lambda\varphi^{(2N)}(t)=\sum_i \{f_i^{(2N)}(t)A_i(t)+h_i^{(2N)}(t)B_i(t)\} \\ \qquad\qquad +\sum_i \{f_i^{(2N-1)}(t)h_i(t)-f_i(t)h_i^{(2N-1)}(t)\}\varphi(t)\,. \end{cases}$$

From these $2N+1$ equalities we eliminate $A_1(t), A_2(t),\cdots, A_N(t)$; $B_1(t), B_2(t),\cdots, B_N(t)$ and obtain a differential equation:

$$(3.12)\qquad \begin{vmatrix} \varphi & f_1 & \cdots & f_N & h_1 & \cdots & h_N \\ \varphi^{(1)} & f_1^{(1)} & \cdots & f_N^{(1)} & h_1^{(1)} & \cdots & h_N^{(1)} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \varphi^{(2N-1)} & f_1^{(2N-1)} & \cdots & f_N^{(2N-1)} & h_1^{(2N-1)} & \cdots & h_N^{(2N-1)} \\ \varphi^{(2N)}-\frac{\varphi}{\lambda}\sum_i \{f_i^{(2N-1)}h_i-f_ih_i^{(2N-1)}\} & f_1^{(2N)} & \cdots & f_N^{(2N)} & h_1^{(2N)} & \cdots & h_N^{(2N)} \end{vmatrix}=0\,.$$

Expanding the linear ordinary equation (3.12) by elements of the first column, we see that the coefficient of $\varphi^{(2N)}$ is $W(f_1,\cdots,f_N, h_1,\cdots,h_N)\neq 0$, as was already pointed out by (3.7), and therefore (3.12) is in fact a differential equation of $2N$th order. (3.9) allows us to write (3.12) in the following form:

$$(3.12)'\qquad L\varphi\equiv\sum_{k=0}^{2N} p_k(t)\varphi^{(k)}(t)=\frac{\varphi}{\lambda}\,,$$

where

$$p_k(t) = -\frac{W_k(f_1, \cdots, f_N, h_1, \cdots, h_N)}{W(f_1, \cdots, f_N, h_1, \cdots, h_N)} \cdot \frac{1}{\sum_i \{f_i^{(2N-1)} h_i - f_i h_i^{(2N-1)}\}}$$
$$(k=0, 1, 2, \cdots, 2N-1)$$

$$p_{2N}(t) = \frac{1}{\sum_i \{f_i^{(2N-1)} h_i - f_i h_i^{(2N-1)}\}}$$

and $W_k(f_1, \cdots, f_N, h_1, \cdots, h_N)$ is obtained from $W(f_1, \cdots, f_N, h_1, \cdots, h_N)$ by replacing the $(k+1)$th row by $f_1^{(2N)}, \cdots, f_N^{(2N)}, h_1^{(2N)}, \cdots, h_N^{(2N)}$. Note that since $f_1, \cdots, f_N, h_1, \cdots, h_N$ are linearly independent functions, they constitute a system of the fundamental solutions of the homogeneous equation:

$$L\varphi = 0 .$$

As to the boundary conditions: since

$$A_i(0) = B_i(T) = 0 \qquad (i=1, 2, \cdots, N)$$
$$h_i^{(k)}(0) = 0 \qquad (i=1, 2, \cdots, N;\ k=0, 1, \cdots, N-1),$$

which is obvious from (3.6) and (3.10), φ should satisfy (3.11) at $t=0$ and T:

$$\begin{cases} \varphi(0) = \varphi^{(1)}(0) = \cdots = \varphi^{(N-1)}(0) = 0 \\ \lambda\varphi^{(N)}(0) = \sum_i h_i^{(N)}(0) B_i(0) \\ \lambda\varphi^{(N+1)}(0) = \sum_i h_i^{(N+1)}(0) B_i(0) \\ \quad \cdots\cdots\cdots\cdots \\ \lambda\varphi^{(2N-1)}(0) = \sum_i h_i^{(2N-1)}(0) B_i(0) \end{cases}$$

and

$$\begin{cases} \lambda\varphi(T) = \sum_i f_i(T) A_i(T) \\ \lambda\varphi^{(1)}(T) = \sum_i f_i^{(1)}(T) A_i(T) \\ \quad \cdots\cdots\cdots\cdots \\ \lambda\varphi^{(2N-1)}(T) = \sum_i f_i^{(2N-1)}(T) A_i(T) . \end{cases}$$

From the above relations we have boundary conditions at $t=0$ and $t=T$ of the equation (3.12) or (3.12)':

(3.13)
$$\begin{cases} \Delta_k(\varphi)\equiv\varphi^{(k)}(0)=0 \qquad (k=0,1,\cdots,N-1)\,, \\ \Delta_k(\varphi)\equiv\begin{vmatrix} \varphi^{(k)}(T) & f_1^{(k)}(T) & \cdots & f_N^{(k)}(T) \\ \varphi^{(N-1)}(T) & f_1^{(N-1)}(T) & \cdots & f_N^{(N-1)}(T) \\ \vdots & \vdots & & \vdots \\ \varphi^{(1)}(T) & f_1^{(1)}(T) & \cdots & f_N^{(1)}(T) \\ \varphi(T) & f_1(T) & \cdots & f_N(T) \end{vmatrix}=0\,. \\ \qquad\qquad (k=N,N+1,\cdots,2N-1) \end{cases}$$

As is well known, the eigenvalue problem of the integral equation (3.10) is equivalent to that of the Sturm-Liouville type of differential equation (3.12)′ under the boundary conditions (3.13). Hence we evaluate the eigenvalues of (3.12)′ under (3.13) instead of (3.10). The eigenvalues of this system are the roots of the following equation with variable λ:

(3.14)
$$\begin{vmatrix} \Delta_0(\chi_1) & \Delta_0(\chi_2) & \cdots & \Delta_0(\chi_{2N}) \\ \Delta_1(\chi_1) & \Delta_1(\chi_2) & \cdots & \Delta_1(\chi_{2N}) \\ \vdots & \vdots & & \vdots \\ \Delta_{2N-1}(\chi_1) & \Delta_{2N-1}(\chi_2) & \cdots & \Delta_{2N-1}(\chi_{2N}) \end{vmatrix}=0$$

where $\chi_1=\chi_1(t\,;\lambda)$, $\chi_2=\chi_2(t\,;\lambda),\cdots,\chi_{2N}=\chi_{2N}(t\,;\lambda)$ are the fundamental solutions of the equation (3.12)′.

Apart from the general formulation of the problem we consider here a rather easy case.

Example 1. (N-ple Markov Gaussian process in Doob's sense)

Consider a stochastic differential equation (1.6):

$$X^{(N)}(t)+a_1X^{(N-1)}(t)+\cdots+a_{N-1}X^{(1)}(t)+a_NX(t)=aB'(t)$$

and suppose that the equation:

(3.15)
$$\nu^N+a_1\nu^{N-1}+\cdots+a_{N-1}\nu+a_N=0$$

has simple roots $\nu_1,\nu_2,\cdots,\nu_N$. Then the solution of (1.6) which satisfies the initial condition:

$$X(0)=X^{(1)}(0)=\cdots=X^{(N-1)}(0)=0$$

is

(3.16)
$$X(t)=\frac{a}{\delta}\int_0^t\sum_{i=1}^N\tilde{\delta}_{Ni}e^{\nu_i(t-u)}dB(u)\,^*,$$

where $\tilde{\delta}_{Ni}$ is the cofactor of ν_i^{N-1} in the determinant:

$$\delta=\begin{vmatrix}1 & 1 & \cdots & 1\\ \nu_1 & \nu_2 & \cdots & \nu_N\\ \cdot & \cdot & & \cdot\\ \cdot & \cdot & & \cdot\\ \cdot & \cdot & & \cdot\\ \nu_1^{N-1} & \nu_2^{N-1} & \cdots & \nu_N^{N-1}\end{vmatrix}.$$

Although $X(t)$ is in fact a real-valued process in spite of its apparent form (3.16), we make use of its complex-valued expression for convenience of calculation. Put

$$f_i(t)=\frac{a}{\delta}\tilde{\delta}_{Ni}e^{\nu_i t}\,,\qquad g_i(t)=e^{-\nu_i t}\qquad (i=1,2,\cdots,N)\,.$$

They satisfy the assumptions 1 and 2.

$$h_i(t)=\frac{a}{\delta}\sum_{j=1}^N\frac{\tilde{\delta}_{Nj}}{\nu_i+\nu_j}(e^{\nu_j t}-e^{-\nu_i t})\,.$$

$$\sum_{i=1}^N\{f_i^{(2N-1)}(t)h_i(t)-f_i(t)h_i^{(2N-1)}(t)\}=(-1)^N\left\{\sum_{i=1}^N f_i^{(N-1)}(t)g_i(t)\right\}^2$$
$$=(-1)^N a^2\,.$$

If we substitute them in the equation (3.12), an easy matrix calculation shows that the fundamental solutions of $L\varphi=0$ are $e^{\nu_1 t}, e^{\nu_2 t},\cdots, e^{\nu_N t}, e^{-\nu_1 t}, e^{-\nu_2 t},\cdots, e^{-\nu_N t}$ and the equation (3.12)′ becomes

(3.17)
$$\frac{(-1)^N}{a^2}(D^2-\nu_1^2)(D^2-\nu_2^2)\cdots(D^2-\nu_N^2)\varphi=\frac{\varphi}{\lambda}\,,$$

where D is a differential operator: $D\varphi=\frac{d\varphi}{dt}$. The differential equation (3.17) with constant coefficients is solved by the usual method. We are interested in the case of large $\frac{1}{\lambda}$, for we intend to treat the small eigenvalues λ. Then for positive large $\frac{1}{\lambda}$ the algebraic equation:

* The stationary solution of (1.6) is
$$X(t)=\frac{a}{\delta}\int_{-\infty}^t\sum_{i=1}^N\tilde{\delta}_{Ni}e^{\nu_i(t-u)}dB(u)\,.$$
See Doob [3].

$$\frac{(-1)^N}{a^2}(z^2-\nu_1^2)(z^2-\nu_2^2)\cdots(z^2-\nu_N^2)=\frac{1}{\lambda} \tag{3.18}$$

has always just two pure imaginary roots whether N is even or odd. Let $\pm\mu_1$ be the pure imaginary roots and $\pm\mu_2, \pm\mu_3, \cdots, \pm\mu_N$ be other roots of (3.18). (Choose $\mu_1, \mu_2, \cdots, \mu_N$ so that the imaginary part of μ_1 and the real parts of $\mu_2, \cdots, \mu_N$ may be positive.) For large $\frac{1}{\lambda}$,

$$|\mu_j| \asymp \lambda^{-1/2N} \qquad (j=1, 2, \cdots, N)\,* \tag{3.19}$$

$$\text{Real part of } \mu_j \asymp \lambda^{-1/2N} \qquad (j=2, 3, \cdots, N).$$

The fundamental solutions of (3.17) are

$$\chi_{\pm j}(t)=e^{\pm\mu_j t} \qquad (j=1, 2, \cdots, N).$$

In order to evaluate asymptotically the nth smallest eigenvalue $\frac{1}{\lambda_n}$ of (3.17) under the boundary conditions (3.13), we solve the equation (3.14) picking up terms of the highest order and ignoring terms of lower order for large $\frac{1}{\lambda}$ under (3.19). We substitute in (3.14)

$$\Delta_k(\chi_{\pm j})=\begin{cases} (\pm\mu_j)^k & (k=0, 1, \cdots, N-1) \\ W(f_1, \cdots, f_N)\,|_{t=T}\,(\pm\mu_j)^k e^{\pm\mu_j T}+o\left(\frac{1}{\sqrt[2N]{\lambda^k}}\exp\left\{\pm\frac{1}{\sqrt[2N]{\lambda}}\right\}\right) & \\ & (k=N, N+1, \cdots, 2N-1) \end{cases}$$

$$(W(f_1, \cdots, f_N)\,|_{t=T} \text{ does not vanish.})$$

and we have

$$\begin{vmatrix} 1 & 1 & 0 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 \\ \mu_1 & -\mu_1 & 0 & 0 & \cdots & 0 & -\mu_2 & -\mu_3 & \cdots & -\mu_N \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \mu_1^{N-1} & (-\mu_1)^{N-1} & 0 & 0 & \cdots & 0 & (-\mu_2)^{N-1} & (-\mu_3)^{N-1} & \cdots & (-\mu_N)^{N-1} \\ \mu_1^N e^{\mu_1 T} & (-\mu_1)^N e^{-\mu_1 T} & 1 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ \mu_1^{N+1} e^{\mu_1 T} & (-\mu_1)^{N+1} e^{-\mu_1 T} & \mu_2 & \mu_3 & \cdots & \mu_N & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \mu_1^{2N-1} e^{\mu_1 T} & (-\mu_1)^{2N-1} e^{-\mu_1 T} & \mu_2^{N-1} & \mu_3^{N-1} & \cdots & \mu_N^{N-1} & 0 & 0 & \cdots & 0 \end{vmatrix}=0.$$

* $f(\lambda)\asymp g(\lambda)$ means $f(\lambda)=O(g(\lambda))$ and $g(\lambda)=O(f(\lambda))$.

After some manipulation we obtain the relation

$$(\mu_1-\mu_2)^2(\mu_1-\mu_3)^2\cdots(\mu_1-\mu_N)^2e^{\mu_1 T} \\ +(-1)^{N-1}(\mu_1+\mu_2)^2(\mu_1+\mu_3)^2\cdots(\mu_1+\mu_N)^2e^{-\mu_1 T}=0$$

i.e.

$$e^{2\mu_1 T}=(-1)^N\left\{\frac{(\mu_1+\mu_2)(\mu_1+\mu_3)\cdots(\mu_1+\mu_N)}{(\mu_1-\mu_2)(\mu_1-\mu_3)\cdots(\mu_1-\mu_N)}\right\}^2. \tag{3.20}$$

Taking account of (3.19), we see that argument of the right-hand side is close to a constant for large $\frac{1}{\lambda}$, and therefore μ_1 should satisfy

$$2T|\mu_1|=2n\pi+O(1)$$

for sufficiently large n (integer). Since multiplicity of each eigenvalue — dimension of the space of solutions corresponding to a eigenvalue — of the equation (3.10) is at most $2N$, we can conclude from the above asymptotical equality and (3.19) that for sufficiently large n

$$\lambda_n \asymp \pi^{-2N}T^{2N}n^{-2N}. \tag{3.21}$$

Hence, from (2.3) and (2.4), we obtain (1.7).

THEOREM 1. *The ε-entropy of an N-ple Markov Gaussian process in Doob's sense is given by* (1.7).

Let us return to the previous formulation. We prepare two lemmas to evaluate eigenvalues of (3.12) or (3.12)'.

LEMMA 1. *The differential operator L in* (3.12)' *is written in the form*

$$L\varphi=\sum_{k=0}^{N}(-1)^k\frac{d^k}{dt^k}\left(q_k(t)\frac{d^k\varphi}{dt^k}\right) \tag{3.22}$$

by taking appropriate functions $q_0(t), q_1(t), \cdots, q_N(t)$.

PROOF. From the equality:

$$\sum_{k=0}^{2N}p_k(t)\frac{d^k\varphi}{dt^k}=\sum_{k=0}^{N}(-1)^k\frac{d^k}{dt^k}\left(q_k(t)\frac{d^k\varphi}{dt^k}\right)$$

we have

$$(3.23)\qquad p_i=\begin{cases}\sum_{k=(i+1)/2}^{i}(-1)^k\binom{k}{2k-i}q_k^{(2k-i)} & (i=0,1,\cdots,N-1)\\ \sum_{k=(i+1)/2}^{N}(-1)^k\binom{k}{2k-i}q_k^{(2k-i)} & (i=N,N+1,\cdots,2N).\end{cases}$$

In order that there exist $q_0, q_1,\cdots, q_N$ that satisfy (3.23), it is necessary and sufficient that (3.24) holds.

$$(3.24)\qquad p_l=\sum_{k=l}^{2N}(-1)^k\binom{k}{l}p_k^{(k-l)}\qquad (l=0,1,\cdots,2N).$$

It is obtained from (3.23) by eliminating $q_0, q_1,\cdots, q_N$. In that course of calculation we must use the following relations of binomial coefficients:

$$\begin{cases}\sum_{k=i}^{2i}(-1)^k\binom{k}{l}\binom{i}{2i-k}=0 & (0\leqq l\leqq i-1)\\ \sum_{k=l}^{2i}(-1)^k\binom{k}{l}\binom{i}{2i-k}=\binom{i}{2i-l} & (i\leqq l\leqq 2i).\end{cases}$$

On the other hand, as the integral operator K in (2.1) is a self-adjoint operator, the differential operator L in (3.12)′, the inverse operator of K, is also a self-adjoint operator, and (3.24) is nothing but a relation obtained from $L\varphi=L^*\varphi$:

$$\sum_{k=0}^{2N}p_k\frac{d^k\varphi}{dt^k}=\sum_{k=0}^{2N}(-1)^k\frac{d^k}{dt^k}(p_k\varphi).$$

Therefore it is possible to write $L\varphi$ in the form (3.22). q's are determined from (3.23) successively. Especially we note

$$(3.25)\qquad q_N(t)=(-1)^N p_{2N}(t)=\frac{1}{\{\sum_i f_i^{(N-1)}(t)g_i(t)\}^2}.$$

It is well known in the theory of integral equation as to (2.1) that the greatest eigenvalue λ_1 is given by

$$(3.26)\qquad \frac{1}{\lambda_1}=\min\,(K\varphi,\varphi)=\min\int_0^T\int_0^T K(t,s)\varphi(s)\varphi(t)ds\,dt$$

where the minimum is taken over the class of functions φ that are continuous and $\|K\varphi\|^2=\int_0^T\left[\int_0^T K(t,s)\varphi(s)ds\right]^2dt=1$. The corresponding normalized eigenfunction φ_1 is the function that attains the minimum of the right-hand side of (3.26). Then nth greatest eigenvalues λ_n of (2.1) is determined inductively by

(3.27) $$\frac{1}{\lambda_n}=\min\,(K\varphi,\varphi)$$

where the minimum is taken over the class of functions φ that are continuous and

$$\|K\varphi\|^2=1 \quad\text{and}\quad (\varphi,\varphi_j)=\int_0^T \varphi(t)\varphi_j(t)dt=0 \qquad (j=1,2,\cdots,n-1)\,.$$

(φ_j is the jth normalized eigenfunction corresponding to λ_j.)

The nth eigenfunction is the function that attains the minimum of the right-hand side of (3.27).

Now we can translate the above remark about K into the expression about L; i.e.

(3.28) $$\frac{1}{\lambda_1}=\min\,(L\varphi,\varphi)=\min\int_0^T L(\varphi)\varphi dt$$

where the minimum is taken over the class of functions that have continuous derivatives up to the $2N$th order and satisfy the boundary condition (3.13) and $\|\varphi\|^2=1$. λ_n is determined inductively by

(3.29) $$\frac{1}{\lambda_n}=\min\,(L\varphi,\varphi)$$

where the minimum is taken over the class of functions that have continuous derivatives up to the $2N$th order and satisfy the boundary condition (3.13) and such that $\|\varphi\|^2=1$ and $(\varphi,\varphi_j)=0$ for the jth eigenfunction φ_j $(j=1,2,\cdots,n-1)$.

It is inconvenient that we need to know the preceding eigenfunctions in order to minimize the energy functional in (3.28) or (3.29):

(3.30) $$D[\varphi]=(L\varphi,\varphi)=\int_0^T\sum_{k=0}^N(-1)^k\varphi\frac{d^k}{dt^k}\left(q_k\frac{d^k\varphi}{dt^k}\right)dt$$

under those conditions. But we have a fundamental method which shows the basic property of eigenvalues that the eigenvalue is given an extremal value.

LEMMA 2. (Courant) *Let C^{2N} be the class of functions that have continuous derivatives up to the $2N$th order in the interval* $[0,T]$. *For arbitrary piecewise continuous functions $\psi_1,\psi_2,\cdots,\psi_n$, put*

$$d(\psi_1,\psi_2,\cdots,\psi_{n-1})=\inf\left\{\frac{D[\varphi]}{\|\varphi\|^2}\,;\quad \begin{array}{l}(\varphi,\psi_j)=0 \quad (j=1,2,\cdots,n-1)\\ \varphi\in C^{2N} \textit{ satisfies the boundary}\\ \textit{conditions } (3.13)\end{array}\right\}.$$

Then the nth eigenvalue of the differential equation $L\varphi=\frac{1}{\lambda}\varphi$ *under the boundary conditions* (3.13) *is given by*

$$\frac{1}{\lambda_n}=\max\left\{d(\phi_1, \phi_2,\cdots, \phi_{n-1})\,;\ \begin{array}{l}\phi_1, \phi_2,\cdots, \phi_{n-1} \textit{ are piecewise}\\ \textit{continuous function}\end{array}\right\}$$

and the maximum of the right-hand side is attained by taking $\phi_1=\varphi_1, \cdots, \phi_{n-1}=\varphi_{n-1}$. ($\varphi_1,\cdots,\varphi_{n-1}$ *are the first n eigenfunctions.*)

PROOF. The proof is carried out just in the same way as found in Courant and Hilbert [2] where the lemma is proved for the second order differential equation.

(3.30) is calculated as

$$(3.31)\quad \begin{aligned} D[\varphi]&=\sum_{k=1}^{N}(-1)^k\sum_{l=0}^{k-1}(-1)^l\varphi^{(l)}(t)\,(q_k(t)\varphi^{(k)}(t))^{(k-l-1)}\big|_{t=T}\\ &\quad+\int_0^T\sum_{k=0}^{N}q_k(t)\left(\frac{d^k\varphi}{dt^k}\right)^2 dt\\ &=(\text{the term of the bilinear form of } \varphi^{(k)}(T)\\ &\qquad\qquad (k=0, 1,\cdots, N-1))\\ &\quad+\int_0^T\sum_{k=0}^{N}q_k(t)\left(\frac{d^k\varphi}{dt^k}\right)^2 dt\end{aligned}$$

where integration by parts is applied and the boundary conditions (3.13) are taken into account. Since $q_0(t), q_1(t),\cdots, q_N(t)$ are continuous in $[0, T]$ and $q_N(t)$ is positive in $[0, T]$, we choose strictly positive constants $\tilde{m}$ and $\underset{\sim}{m}$ such that

$$|q_k(t)|\leqq\tilde{m}\qquad (k=0, 1,\cdots, N) \text{ and } q_N(t)\geqq\underset{\sim}{m} \text{ in } [0, T].$$

We define differential operator $\tilde{L}$ by replacing $q_k(t)$ in L by $\tilde{m}$:

$$(3.32)\qquad \tilde{L}\varphi=\sum_{k=0}^{N}(-1)^k\tilde{m}\frac{d^{2k}\varphi}{dt^{2k}}$$

and the boundary conditions:

$$(3.33)_1\qquad \varphi(0)=\varphi^{(1)}(0)=\cdots=\varphi^{(N-1)}(0)=0$$

$$(3.33)_2\qquad \varphi^{(k)}(T)=a_{k0}\varphi(T)+a_{k1}\varphi^{(1)}(T)+\cdots+a_{k\,N-1}\varphi^{(N-1)}(T)\qquad (k=N, N+1,\cdots, 2N-1)$$

where a_{kj}'s are constants chosen appropriately so that the first term of (3.31) may coincide, as a bilinear form, with that obtained from $\int_0^T\tilde{L}(\varphi)\varphi\,dt$ by integration by parts and by applying $(3.33)_1$ and $(3.33)_2$.

As is easily seen from lemma 2, for the nth eigenvalues $\frac{1}{\lambda_n}$, $\frac{1}{\tilde{\lambda}_n}$ of the equation $L\varphi=\frac{\varphi}{\lambda}$ under the boundary conditions (3.13) and the equation $\tilde{L}\varphi=\frac{\varphi}{\tilde{\lambda}}$ under the boundary conditions $(3.33)_1$ and $(3.33)_2$ respectively, the following relation holds; $\frac{1}{\lambda_n}\leqq\frac{1}{\tilde{\lambda}_n}$. On the other hand, eigenvalues $\frac{1}{\tilde{\lambda}_n}$'s are evaluated for sufficiently large n directly in the same way as in example 1:

$$\lambda_n\geqq\tilde{\lambda}_n\asymp T^{2N}n^{-2N}\,.$$

We also get

$$\lambda_n\leqq\underset{\sim}{\lambda}_n\asymp T^{2N}n^{-2N}\,,$$

where $\underset{\sim}{\lambda}_n$ is the nth eigenvalue of the equation:

$$\underset{\sim}{L}\varphi=\sum_{k=0}^{N-1}(-1)^k(-\tilde{m})\frac{d^{2k}\varphi}{dt^{2k}}+(-1)^N\underset{\sim}{m}\frac{d^{2N}\varphi}{dt^{2N}}$$

under the boundary conditions of the same type as $(3.33)_1$ and $(3.33)_2$. Hence we have the final evaluation for λ_n.

$$\lambda_n\asymp T^{2N}n^{-2N}\,.$$

THEOREM 2. *Under the assumptions* 1 *and* 2, *the ε-entropy of the Gaussian process*

$$X(t)=\int_0^t\sum_{i=1}^N f_i(t)g_i(u)dB(u)\qquad 0\leqq t\leqq T$$

is given by (1.7).

Example 2. The coefficient of the term of principal order in (1.7) is calculated easily for a Gaussian process that is expressed as

$$X(t)=\int_0^t g(u)dB(u)\,.$$

In this case we have by the same notation as in section 2,

$$L\varphi=-\frac{d}{dt}\left(\frac{1}{g(t)^2}\frac{d\varphi}{dt}\right)=\frac{\varphi}{\lambda}$$

under the boundary condition:

$$\varphi(0)=0\,,\qquad\varphi^{(1)}(T)=0\,.$$

If we introduce new variables x, y by the Liouville transformation

$$x=\int_0^t |g(u)|\,du\,, \qquad y=\frac{\varphi(t)}{\sqrt{|g(t)|}}\,,$$

the above problem is transformed into a boundary value problem in $[0,a]$:

$$\begin{cases} -\dfrac{d^2y}{dx^2}+z(x)y=\dfrac{y}{\lambda} \\ y(0)=0\,, \qquad y^{(1)}(a)=y^* \end{cases}$$

where $a=\int_0^T |g(u)|\,du$, $z=\sqrt{|g|}\dfrac{d^2}{dt^2}\sqrt{|g|}$ and y^* is a constant. By the method of lemma 2, we obtain the asymptotic evaluation

$$\begin{aligned}\frac{1}{\lambda_n}&=\frac{n^2\pi^2}{a^2}+O(1)\\ &=n^2\pi^2\left[\int_0^T |g(u)|\,du\right]^{-2}+O(1)\,,\end{aligned}$$

that is,

$$\lambda_n=\frac{1}{\pi^2}\left[\int_0^T |g(u)|\,du\right]^2 n^{-2}+o(n^{-2})\,.$$

Therefore we get an evaluation of the ε-entropy from (2.3) and (2.4)

$$\mathcal{H}_\varepsilon(\{X(t)\})=\frac{2}{\pi^2}\left[\int_0^T |g(u)|\,du\right]^2\frac{1}{\varepsilon^2}+o\left(\frac{1}{\varepsilon^2}\right)$$

as $\varepsilon\to 0$.

4. ε-entropy of $M(t)$ process

It is not essential that we have included $t=0$ in the assumptions 1 and 2. If we rewrite the assumptions 1 and 2 excluding $t=0$, we have generally an eigenvalue problem for the "singular" Sturm-Liouville problem. But also in this case, we may be able to trace the argument in section 2 with a slight modification taking notice of behaviors at $t=0$ of f's, h's, g's and etc. and reexamining the boundary conditions at $t=0$.

As an example of the singular Sturm-Liouville problem, we shall take Lévy's $M(t)$ process. Y. Baba [1] evaluated the ε-entropy of $M_3(t)$, obtaining an exact solution of $L\varphi=\dfrac{\varphi}{\lambda}$ and making use of properties of the Bessel function.

Lévy's $M_N(t)$ process is defined for $t\in[0, T]$ by

(4.1) $$M_N(t)=\int_{S_N(t)} X(A)d\sigma(A)$$

where $X(A)$ is a "Brownian motion with a parameter space R^N *" and $S_N(t)$ is the sphere with center O and radius t, and $d\sigma$ is the uniform measure on $S_N(t)$ with total measure 1.

If $N=2p+1$ (odd number), $M_{2p+1}(t)$ is expressed as

(4.2) $$M_{2p+1}(t)=\int_0^t P_{2p+1}\left(\frac{u}{t}\right)dB(u)\,,$$

where

$$P_{2p+1}(u)=\frac{2p}{\sqrt{\pi}}\sqrt{I_{2p}}\int_u^1(1-x^2)^{p-1}dx$$
$$=\text{polynomial of degree } 2p-1,$$
$$I_{2p}=\int_0^{\pi/2}\sin^{2p}\theta\, d\theta\,.$$

$M_{2p+1}(t)$ is a $(p+1)$-ple Markov Gaussian process in Lévy's restricted sense. Its covariance function $K_{2p+1}(t,s)$ $(0\leqq s\leqq t\leqq T)$ was calculated by Lévy [6].

$$K_{2p+1}(t,s)=\frac{s}{2}-\frac{1}{2}\sum_{l=1}^{p}a_l\frac{s^{2l}}{t^{2l-1}}\,,$$

where

$$a_l=\frac{1}{4l-1}\prod_{k\neq l}\frac{k(2k-1)}{(k-l)(2l+2k-1)}\,.$$

We shall sketch an outline of the calculation to evaluate the eigenvalues of the integral equation (2.1) for K_{2p+1}, taking up the case $p=2$ as an example. For brevity of notation, let the time interval considered be [0,1].

(4.3) $$K_5(t,s)=\frac{s}{2}-\frac{s^2}{5t}+\frac{s^4}{70t^3}\qquad(0\leqq s\leqq t\leqq 1)\,.$$

It does not satisfy the assumptions 1 and 2 at the point $t=0$. But, since it is continuous at 0 as $t\geqq s\to 0$, there are countably many eigenvalues λ converging to zero for the integral equation (4.4) defined

* $X(A,\omega)$ $(A\in R^N)$ is called a "Brownian motion with a parameter space R^N", if
i) $X(A,\omega)$ is a Gaussian random variable,
ii) $X(O)=0$ for the origin $O\in R^N$,
$E\{X(A)\}=0$, $E\,|X(A)-X(B)|^2=|A-B|$, where $|A-B|$ is the Euclidean distance between A and B in R^N.

by the symmetric, positive semi-definite and completely continuous kernel $K_5(t, s)$:

$$\int_0^t \left(\frac{s}{2}-\frac{s^2}{5t}+\frac{s^4}{70t^3}\right)\varphi(s)\,ds+\int_t^1\left(\frac{t}{2}-\frac{t^2}{5s}+\frac{t^4}{70s^3}\right)\varphi(s)ds=\lambda\varphi(t)\,. \tag{4.4}$$

Differentiating both sides of (4.4) in the same way as in section 3, we obtain the differential equation

$$-\frac{t^4}{12}\varphi^{(6)}-t^3\varphi^{(5)}-2t^2\varphi^{(4)}+2t\varphi^{(3)}-\frac{\varphi}{\lambda}=0\,,$$

which is equivalent to

$$L(\varphi)\equiv-\frac{d^3}{dt^3}\left(\frac{t^4}{12}\frac{d^3\varphi}{dt^3}\right)+\frac{d^2}{dt^2}\left(t^2\frac{d^2\varphi}{dt^2}\right)-\frac{d}{dt}\left(2\frac{d\varphi}{dt}\right)=\frac{\varphi}{\lambda}\,. \tag{4.5}$$

The boundary conditions at $t=1$ are given quite similarly to (3.13):

$$(4.6)_1 \qquad \Delta_k(\varphi)=\begin{vmatrix} \varphi^{(k)}(1) & f_1^{(k)}(1) & f_2^{(k)}(1) & f_3^{(k)}(1) \\ \varphi^{(2)}(1) & f_1^{(2)}(1) & f_2^{(2)}(1) & f_3^{(2)}(1) \\ \varphi^{(1)}(1) & f_1^{(1)}(1) & f_2^{(1)}(1) & f_3^{(1)}(1) \\ \varphi(1) & f_1(1) & f_2(1) & f_3(1) \end{vmatrix}=0 \qquad (k=3, 4, 5)\,,$$

where $f_1(t)=1$, $f_2(t)=\frac{1}{t}$, $f_3(t)=\frac{1}{t^3}$.

As to the boundary conditions at $t=0$ we have from (4.4),

$$(4.6)_2 \qquad \begin{aligned} &\Delta_0(\varphi)=\varphi(0)=0 \\ &\Delta_k(\varphi)=\varphi^{(k)}(0) \quad \text{is finite} \qquad (k=1, 2)\,. \end{aligned}$$

The equation (4.5) is in fact "singular", for $q_3(t)=\frac{t^4}{12}$ vanishes at the boundary $t=0$. The eigenvalues of (4.5) under the boundary conditions $(4.6)_1$ and $(4.6)_2$ are evaluated from a variational method based on lemma 2. The nth eigenvalue $\frac{1}{\lambda_n}$ of (4.5) is obtained from

$$\frac{1}{\lambda_n}=\max_{\phi_1,\cdots,\phi_{n-1}}\inf\left\{\frac{D[\varphi]}{\|\varphi\|^2};\ \begin{array}{l}\varphi\in C^6 \text{ and satisfies the boundary} \\ \text{conditions } (4.6)_1 \text{ and } (4.6)_2, \\ (\varphi, \phi_j)=0 \quad (j=1, 2, \cdots, n-1)\end{array}\right\} \tag{4.7}$$

where the maximum is taken over all piecewise continuous functions $\phi_1, \phi_2, \cdots, \phi_{n-1}$ and $D[\varphi]=\int_0^1 L(\varphi)\varphi\,dt$. Integrating by parts and substituting $(4.6)_1$ and $(4.6)_2$, we have

$$(4.7)' \quad \frac{1}{\lambda_n}=\max_{\phi_1,\cdots,\phi_{n-1}} \inf\left\{\frac{F(\varphi(1))+\int_0^1 G(\varphi(t))dt}{\|\varphi\|^2};\quad \begin{matrix}\varphi\in C^6 \text{ and } (\varphi,\phi_j)=0\\ (j=1,2,\cdots,n-1)\end{matrix}\right\}$$

from (4.7), where $F(\varphi(t))$ is a bilinear form of $\varphi(t)$, $\varphi^{(1)}(t)$, $\varphi^{(2)}(t)$ and

$$G(\varphi(t))=\frac{t^4}{12}\left(\frac{d^3\varphi}{dt^3}\right)^2+t^2\left(\frac{d^2\varphi}{dt^2}\right)^2+2\left(\frac{d\varphi}{dt}\right)^2.$$

(i) Asymptotic evaluation of $\frac{1}{\lambda_n}$ from above

(4.7)′ is smaller than

$$(4.8) \quad \frac{1}{\tilde{\lambda}_n}=\max_{\phi_1,\cdots,\phi_{n-1}} \inf\left\{\frac{1}{\|\varphi\|^2}\left[F(\varphi(1))+\int_0^1\left\{\left(\frac{d^3\varphi}{dt^3}\right)^2+\left(\frac{d^2\varphi}{dt^2}\right)^2+2\left(\frac{d\varphi}{dt}\right)^2\right\}dt\right]\right.$$
$$\left.;\varphi\in C^6 \text{ and } (\varphi,\phi_j)=0 \quad (j=1,2,\cdots,n-1)\right\}.$$

As (4.8) is equivalent to the eigenvalue problem of the differential equation of type (3.32) under $(3.33)_1$ and $(3.33)_2$, we easily obtain

$$(4.9) \qquad \frac{1}{\lambda_n}\leqq\frac{1}{\tilde{\lambda}_n}\asymp n^6.$$

(ii) Asymptotic evaluation of $\frac{1}{\lambda_n}$ from below

First we give a lemma which holds for usual variational eigenvalue problems.

LEMMA 3. *Fix a certain $t_0\in(0,1)$ and consider the following variational eigenvalue problems:*

$$(4.10) \qquad \frac{\int_0^{t_0} G(\varphi(t))dt}{\int_0^{t_0}\varphi(t)^2dt},$$

$$(4.11) \qquad \frac{F(\varphi(1))+\int_{t_0}^1 G(\varphi(t))dt}{\int_{t_0}^1\varphi(t)^2dt},$$

and denote by $V_1\left(\frac{1}{\lambda}\right)$ and $V_2\left(\frac{1}{\lambda}\right)$ the number of eigenvalues not more than $\frac{1}{\lambda}$ of (4.10) and (4.11) respectively. Then the number $V\left(\frac{1}{\lambda}\right)$ of *eigenvalues less than* $\frac{1}{\lambda}$ *of the variational eigenvalue problem:*

$$(4.12) \qquad \frac{F(\varphi(1))+\int_0^1 G(\varphi(t))dt}{\int_0^1 \varphi(t)^2 dt},$$

which appears in (4.7)′, *is not more than the sum of* $V_1\left(\frac{1}{\lambda}\right)$ *and* $V_2\left(\frac{1}{\lambda}\right)$:

$$V\left(\frac{1}{\lambda}\right) \leqq V_1\left(\frac{1}{\lambda}\right)+V_2\left(\frac{1}{\lambda}\right).$$

PROOF. Put $V_1\left(\frac{1}{\lambda}\right)=m$, $V_2\left(\frac{1}{\lambda}\right)=n$. Since each sequence of eigenvalues forms an increasing and divergent sequence, it is sufficient to show

$$\frac{1}{\lambda_{m+n+1}} > \frac{1}{\lambda},$$

where $\frac{1}{\lambda_l}$ is the lth eigenvalue of (4.12). We have

$$\begin{aligned}\frac{1}{\lambda_{m+n+1}} &= \max_{\substack{\psi_1,\cdots,\psi_m,\\ \psi_{m+1},\cdots,\psi_{m+n}}} \inf\left\{(4.12);\ \begin{matrix}\varphi\in C^6[0,1] \text{ and } (\varphi,\psi_j)=0\\ (j=1,2,\cdots,m+n)\end{matrix}\right\}\\ &\geqq \max_{\substack{\xi_1^{(1)},\cdots,\xi_m^{(1)},\\ \xi_1^{(2)},\cdots,\xi_n^{(2)}}} \inf\left\{(4.12);\ \begin{matrix}\varphi\in C^6[0,1] \text{ and } (\varphi,\xi_j^{(1)})=0\ (j=1,\cdots,m)\\ (\varphi,\xi_k^{(2)})=0\ (k=1,\cdots,n)\end{matrix}\right\}\end{aligned}$$

where $\xi_1^{(1)},\cdots,\xi_m^{(1)}$ run through the class of piecewise continuous functions that vanish outside $[0, t_0]$ and $\xi_1^{(2)},\cdots,\xi_n^{(2)}$ run through the class of piecewise continuous functions that vanish outside $[t_0, 1]$. The above is

$$\begin{aligned}&\geqq \max_{\substack{\xi_1^{(1)},\cdots,\xi_m^{(1)},\\ \xi_1^{(2)},\cdots,\xi_n^{(2)}}} \min\Bigg[\inf\left\{(4.10);\ \begin{matrix}\varphi\in C^6[0,t_0] \text{ and } \int_0^{t_0}\varphi\xi_j^{(1)}dt=0\\ (j=1,2,\cdots,m)\end{matrix}\right\},\\ &\qquad\qquad \inf\left\{(4.11);\ \begin{matrix}\varphi\in C^6[t_0,1] \text{ and } \int_{t_0}^{1}\varphi\xi_k^{(2)}dt=0\\ (k=1,2,\cdots,n)\end{matrix}\right\}\Bigg]\\ &= \max_{\substack{\xi_1^{(1)},\cdots,\xi_m^{(1)},\\ \xi_1^{(2)},\cdots,\xi_n^{(2)}}} \min\{d^{(1)}(\xi_1^{(1)},\cdots,\xi_m^{(1)}),\ d^{(2)}(\xi_1^{(2)},\cdots,\xi_n^{(2)})\}\\ &\geqq \max_{\xi_1^{(1)},\cdots,\xi_m^{(1)}} \min\{d^{(1)}(\xi_1^{(1)},\cdots,\xi_m^{(1)}),\ d^{(2)}(\varphi_1^{(2)},\cdots,\varphi_n^{(2)})\}\end{aligned}$$

$$=\min\{\max_{\xi_1^{(1)},\cdots,\xi_m^{(1)}} d^{(1)}(\xi_1^{(1)},\cdots,\xi_m^{(1)}),\, d^{(2)}(\varphi_1^{(2)},\cdots,\varphi_n^{(2)})\}$$

$$=\min\{d^{(1)}(\varphi_1^{(1)},\cdots,\varphi_m^{(1)}),\, d^{(2)}(\varphi_1^{(2)},\cdots,\varphi_n^{(2)})\}$$

$$=\min\left\{\frac{1}{\lambda_{m+1}^{(1)}},\, \frac{1}{\lambda_{n+1}^{(2)}}\right\}>\frac{1}{\lambda}$$

where $d^{(1)}(\xi_1^{(1)},\cdots,\xi_m^{(1)})$ and $d^{(2)}(\xi_1^{(2)},\cdots,\xi_n^{(2)})$ are the functionals defined in lemma 2 to determine eigenvalues of (4.10) and (4.11), and $\varphi_1^{(1)},\cdots,\varphi_m^{(1)}$ are the first m eigenfunctions of (4.10) and $\varphi_1^{(2)},\cdots,\varphi_n^{(2)}$ are the first n eigenfunctions of (4.11).

Take an arbitrary large $\frac{1}{\lambda}$ and fix it. We choose a positive integer k such that $\frac{1}{\lambda}<\frac{2^k}{c}$, where $c=\max\{|K_5(t,s)|;\ 0\leqq t, s\leqq 1\}$. First of all, from the inequality

$$\left|\int_0^{1/2^k}\varphi(t)\left(\int_0^{1/2^k}K_5(t,s)\varphi(s)ds\right)dt\right|$$
$$\leqq c\left\{\int_0^{1/2^k}|\varphi(t)|\,dt\right\}^2\leqq\frac{c}{2^k}\int_0^{1/2^k}|\varphi(t)|^2dt\,,$$

it can be concluded that the integral equation

$$\int_0^{1/2^k}K_5(t,s)\varphi(s)ds=\lambda\varphi(t)\qquad\left(0\leqq t\leqq\frac{1}{2^k}\right)\tag{4.13}$$

has no eigenvalues larger than $\frac{c}{2^k}$. Now we divide the time interval $[0,1]$ into $k+1$ subintervals $I_k=\left[0,\frac{1}{2^k}\right]$, $I_{k-1}=\left[\frac{1}{2^k},\frac{1}{2^{k-1}}\right],\cdots,I_l=\left[\frac{1}{2^{l+1}},\frac{1}{2^l}\right],\cdots,I_0=\left[\frac{1}{2},1\right]$. Since $G(\varphi(t))\geqq\frac{1}{12\cdot 2^{4(l+1)}}\left(\frac{d^3\varphi}{dt^3}\right)^2$ in the subinterval I_l, the conclusion of lemma 3 remains true if $G(\varphi(t))$ is substituted by $\frac{1}{12\cdot 2^{4(l+1)}}\left(\frac{d^3\varphi}{dt^3}\right)^2$ in some of those subintervals. If we translate each variational problem into the eigenvalue problem of differential equation, we have $k+1$ equations:

	Differential equation	Boundary conditions
(eq. k)	$L\varphi=\frac{1}{\lambda}\varphi$ on I_k	$(4.6)_2$ at $t=0$ and we substitute 1 in $(4.6)_1$ by $\frac{1}{2^k}$ at $t=\frac{1}{2^k}$,

(This is equivalent to the problem (4.13).)

(eq. $k-1$) $\quad \dfrac{1}{12\cdot 2^{4k}}\dfrac{d^6\varphi}{dt^6}=\dfrac{1}{\lambda}\varphi$ on I_{k-1} $\quad$ boundary conditions of type $(3.33)_2$ at $t=\dfrac{1}{2^k}$ and $\varphi^{(3)}\left(\dfrac{1}{2^{k-1}}\right)=\varphi^{(4)}\left(\dfrac{1}{2^{k-1}}\right)=\varphi^{(5)}\left(\dfrac{1}{2^{k-1}}\right)=0$

(eq. l) $\quad \dfrac{1}{12\cdot 2^{4(l+1)}}\dfrac{d^6\varphi}{dt^6}=\dfrac{1}{\lambda}\varphi$ on I_l $\quad \varphi\left(\dfrac{1}{2^{l+1}}\right)=\varphi^{(1)}\left(\dfrac{1}{2^{l+1}}\right)=\varphi^{(2)}\left(\dfrac{1}{2^{l+1}}\right)=0$

$$\varphi^{(3)}\left(\frac{1}{2^{l}}\right)=\varphi^{(4)}\left(\frac{1}{2^{l}}\right)=\varphi^{(5)}\left(\frac{1}{2^{l}}\right)=0$$

$$l=1, 2, 3, \cdots, k-2$$

(eq. 0) $\quad \dfrac{1}{12\cdot 2^{4}}\dfrac{d^6\varphi}{dt^6}=\dfrac{1}{\lambda}\varphi$ on I_0 $\quad \varphi\left(\dfrac{1}{2}\right)=\varphi^{(1)}\left(\dfrac{1}{2}\right)=\varphi^{(2)}\left(\dfrac{1}{2}\right)=0$ and boundary conditions of type $(3.33)_2$ at $t=1$.

Let $V_l\left(\dfrac{1}{\lambda}\right)$ be the number of eigenvalues less than $\dfrac{1}{\lambda}$ of (eq. l). Then, according to lemma 3,

$$V\left(\frac{1}{\lambda}\right)\leqq \sum_{l=0}^{k} V_l\left(\frac{1}{\lambda}\right).$$

$V_l\left(\dfrac{1}{\lambda}\right)$ $(l=0, 1, \cdots, k-1)$ is evaluated from above after the example 1 by

$$2T\,|\mu_1|=2n\pi+O(1),$$

where

$$|\mu_1|=\left\{\frac{12\cdot 2^{4(l+1)}}{\lambda}\right\}^{1/6} \quad \text{and} \quad T=\frac{1}{2^{l+1}}.$$

Thus we have

$$V_l\left(\frac{1}{\lambda}\right)\leqq\frac{6}{\pi}\frac{1}{2^{l+1}}\left\{\frac{12\cdot 2^{4(l+1)}}{\lambda}\right\}^{1/6}$$
$$=\frac{6\cdot 12^{1/6}}{\pi}\,2^{-(l+1)/3}\,\frac{1}{\lambda^{1/6}},$$

taking into consideration the multiplicity of eigenvalues. On the other hand $V_k\left(\dfrac{1}{\lambda}\right)=0$, as is noted beforehand.

Since

$$V\left(\frac{1}{\lambda}\right) \leqq \frac{6\cdot 12^{1/6}2^{-1/3}}{\pi\lambda^{1/6}} \sum_{l=0}^{k-1} 2^{-l/3} = \frac{6\cdot 12^{1/6}2^{-1/3}}{\pi\lambda^{1/6}} \frac{1-2^{-k/3}}{1-2^{-1/3}}$$
$$\leqq \text{const.}\ \frac{1}{\lambda^{1/6}}\,,$$

we have an evaluation from below for the nth eigenvalue $\frac{1}{\lambda_n}$ of (4.5) under $(4.6)_1$ and $(4.6)_2$ whose order is n^6 at most.

Combining (i) and (ii), we have

$$\mathscr{H}_\varepsilon\{M_5(t)\} \asymp \varepsilon^{-2/5}\,.$$

For $M_7(t)$ and $M_9(t)$ the following differential equations are obtained:

$$\left(\frac{t^6}{360}\varphi^{(4)}\right)^{(4)} - \left(\frac{t^4}{10}\varphi^{(3)}\right)^{(3)} + (t^2\varphi^{(2)})^{(2)} - (2\varphi^{(1)})^{(1)} = \frac{\varphi}{\lambda}\,,$$

$$-\left(\frac{t^8}{20160}\varphi^{(5)}\right)^{(5)} + \left(\frac{t^6}{252}\varphi^{(4)}\right)^{(4)} - \left(\frac{3}{28}t^4\varphi^{(3)}\right)^{(3)} + (t^2\varphi^{(2)})^{(2)} - (2\varphi^{(1)})^{(1)} = \frac{\varphi}{\lambda}\,.$$

We can ascertain $\lambda_n \asymp n^{-8}$ and $\lambda_n \asymp n^{-10}$, respectively for them just like as the above. Thus we have the following evaluations of ε-entropy:

$$\mathscr{H}_\varepsilon\{M_7(t)\} \asymp \varepsilon^{-2/7}$$
$$\mathscr{H}_\varepsilon\{M_9(t)\} \asymp \varepsilon^{-2/9}\,.$$

It is troublesome to write down the differential equation, corresponding to an arbitrary $M_{2p+1}(t)$, in the general form of (4.5). But if we do not get weary of those calculation, it seems that we can obtain an evaluation:

$$\mathscr{H}_\varepsilon\{M_{2p+1}(t)\} \asymp \varepsilon^{-2/(2p+1)}\,.$$

Acknowledgement

The author wishes to express his hearty thanks to N. Ikeda and Y. Baba who turned his attention to ε-entropy and furnished him with materials about the topics, and to T. Hida who read through the manuscript to give him some useful advices.

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] Y. Baba, "On ε-entropy of Gaussian processes," (in Japanese), *Jour. of Kobe Univ. of Commerce*, 18 (1966), 422-430.

[2] R. Courant and D. Hilbert, *Methoden der mathematischen Physik I*, Springer, 1931.

[3] J. L. Doob, "The elementary Gaussian processes," *Ann. Math. Statist.*, 15 (1944), 229-282.

[4] И. М. Гельфанд, А. Н. Колмогоров, А. М. Яглом, "Количество информации и знтропия для непрерывных распределений," Доклад на третьем Всесоюзном математическом съезде, 1956.

[5] T. Hida, "Canonical representations of Gaussian processes and their applications," *Memoirs of the College of Sci., Kyoto Univ.*, Series A, 13, Math. No. 1 (1960), 109-155.

[6] P. Lévy, "A special problem of Brownian motion, and a general theory of Gaussian random functions," *Third Berkeley Symposium*, (1956), 133-175.

[7] М. С. Пинскер, "Гауссовские Источники," Провлемы Передачи Информации, 14 (1963), 59-100.