The Annals of Mathematical Statistics
1969, Vol. 40, No. 3, 870–904

PRODUCT ENTROPY OF GAUSSIAN DISTRIBUTIONS[1]

BY EDWARD C. POSNER, EUGENE R. RODEMICH AND HOWARD RUMSEY, JR.

Jet Propulsion Laboratory, California Institute of Technology

0. Summary. This paper studies the product epsilon entropy of mean-continous Gaussian processes. That is, a given mean-continuous Gaussian process on the unit interval is expanded into its Karhünen expansion. Along the kth eigenfunction axis, a partition by intervals of length ϵ_k is made, and the entropy of the resulting discrete distribution is noted. The infimum of the sum over k of these entropies subject to the constraint that $\sum \epsilon_k^2 \leqq \epsilon^2$ is the product epsilon entropy of the process. It is shown that the best partition to take along each eigenfunction axis is the one in which 0 is the midpoint of an interval in the partition. Furthermore, the product epsilon entropy is finite if and only if $\sum \lambda_k \log \lambda_k^{-1}$ is finite, where λ_k is the kth eigenvalue of the process. When the above series is finite, the values of ϵ_k which achieve the product entropy are found. Asymptotic expressions for the product epsilon entropy are derived in some special cases. The problem arises in the theory of data compression, which studies the efficient representation of random data with prescribed accuracy.

1. Introduction. This paper is motivated by the problem of data compression, the efficient representation of data for the purpose of information transmission. We shall consider the case in which the data to be represented consists of a sample function from a Gaussian process $X(t)$ on the unit interval which is mean-continuous; i.e. $E[x(s) - x(t)]^2 \to 0$ as $s \to t$, for all t. Our basic problem is how to transmit (over a noiseless channel) information as to which sample function of X occurred. We assume that the recipient of the transmitted data has full knowledge of the statistics of the process. In particular he knows the Karhünen expansion [1] of the process; namely

$$X(t) = \sum_{k=1}^{\infty} \lambda_k^{\frac{1}{2}} y_k \phi_k(t), \tag{1}$$

where the y_k are mutually independent unit normal random variables (they determine which sample function of the processes occurred); the $\phi_k(t)$ are the (orthonormal) eigenfunctions of the process; they are known *a priori* as are the λ_k, which are the eigenvalues of the process, and are non-negative. We note that the series in (1) converges with probability 1. If

$$R(s, t) = E(X(s)X(t)) \tag{2}$$

is the covariance function of the process, then $R(s, t)$ is continuous, by the mean

Received 28 December 1967; revised 22 November 1968.

[1] This paper presents the results of one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract No. NAS 7-100, sponsored by the National Aeronautics and Space Administration.

continuity of $X(t)$, and

$$R(s, t) = \sum \lambda_k \phi_k(s) \phi_k(t), \tag{3}$$

the convergence being uniform on the unit square. The functions ϕ_k are continuous and satisfy the integral equation

$$\lambda_k \phi_k(s) = \int_0^1 R(s, t) \phi_k(t) \, dt, \tag{4}$$

where the λ_k are non-negative and are the eigenvalues of this integral equation. It follows that

$$\sum \lambda_k = \int_0^1 R(s, s) \, ds < \infty. \tag{5}$$

In the special case when all but a finite number of the λ_k are zero, the process X is just a finite dimensional Gaussian distribution. The interesting cases, from the point of product entropy, turn out to be the one-dimensional processes and the infinite-dimensional processes.

In the data compression problem we wish to represent the sample functions of the known process X. By equation (1) we can fully describe a sample function $X(t)$ by specifying the values of the y_k which occur in (1). Loosely speaking, we shall call y_k the projection of the process along the kth coordinate axis.

Our final assumption concerning the nature of our problem is the requirement that the information which is transmitted must be adequate to locate the sample function in some set of L_2-diameter at most ϵ. The data compression procedure we propose is as follows. Observe $X(t)$ and compute its projections, y_k, along the coordinate axes. Quantize the kth coordinate axis into intervals of diameter at most ϵ_k. For each k, transmit the index of the interval which actually occurred. If the ϵ_k satisfy

$$\sum \epsilon_k^2 \leqq \epsilon^2, \tag{6}$$

then, with probability 1, when the intervals of uncertainty are known, the original sample function is known to within a set which is a hyper-rectangle of diameter at most ϵ.

Our main concern in this paper is to study the entropy of the above procedure. We observe that this entropy does not depend on the eigenfunctions ϕ_k of the process, but only on the eigenvalues λ_k. This is because any two mean-continuous Gaussian processes with the same λ_k possess measure-preserving isometries between the Hilbert spaces generated by their ϕ_k. It follows that assumptions about stationarity, band-limiting, etc. are relevant only insofar as they help estimate the eigenvalues λ_k.

Further discussion of data compression as well as a definition of epsilon entropy for mean-continuous stochastic processes are found in [2]. The entropy defined in [2] is bounded from above by the product epsilon entropy considered here; for it uses partitions by arbitrary measurable sets of diameter at most ϵ, instead of by hyper-rectangles of diameter at most ϵ. In [3] it is shown that the epsilon entropy of a mean-continuous Gaussian process on the unit interval is always finite. It turns out, however, that product epsilon entropy is finite if and only if $\sum \lambda_k \log \lambda_k^{-1}$ converges.

We shall now describe the organization of the rest of this paper. Section 2 treats the 1-dimensional case. We show that the best ϵ-partition (the ϵ-partition with least entropy) is that partition by intervals of length ϵ which contains the interval $[-\epsilon/2, \epsilon/2]$. This is the longest section of the paper. We treat the cases of large and small ϵ separately. Techniques of analytic function theory are necessary.

In Section 3 we show that the product epsilon entropy $J_\epsilon(X)$ of a mean-continuous Gaussian process on the unit interval is finite if and only if the "entropy of the eigenvalues" $\sum \lambda_k \log \lambda_k^{-1}$ is finite. In the case in which $J_\epsilon(X)$ is finite, we give a product partition whose entropy equals $J_\epsilon(X)$.

Section 4 gives an asymptotic form for $J_\epsilon(X)$ when the eigenvalues satisfy a relation of the form $\lambda_k \sim Bk^{-p}$. In particular, for the Weiner process, $J_\epsilon(X) \sim C/\epsilon^2$ as $\epsilon \to 0$, where C is a constant between 6 and 7.

Section 5 considers a general lower bound $L_\epsilon(X)$ for $J_\epsilon(X)$. We show that if

$$\sum_{k=n}^{\infty} \lambda_k = O(n\lambda_n)$$

then the ratio J_ϵ/L_ϵ remains bounded as ϵ tends to 0; and if

$$\sum_{k=n}^{\infty} \lambda_k = o(n\lambda_n),$$

then $J_\epsilon \sim L_\epsilon$ as $\epsilon \to 0$.

This last result implies that, when $\sum_n^\infty \lambda_k = o(n\lambda_n)$, product ϵ-entropy is asymptotically as good as ϵ-entropy for small ϵ. As an application of our techniques we show that for a stationary band-limited Gaussian process on the unit interval, with well-behaved spectrum,

$$J_\epsilon(X) \sim (\log^2 \epsilon^{-1})(2 \log\log \epsilon^{-1})^{-1}.$$

2. The one-dimensional normal distribution. In this section we consider a normal random variable of mean 0 on the line. We show that the ϵ-partition of the line with least entropy is the "centered" partition consisting of non-overlapping intervals of length ϵ, and containing the interval $[-\epsilon/2, \epsilon/2]$.

We need a series of six lemmas to prove this result, which is Theorem 1. The first lemma shows that we need only consider portions consisting of non-overlapping intervals of length ϵ. Lemmas 2–3 show that the centered partition is best (has smallest entropy) if $\epsilon \geqq 3$. Lemmas 4–6 are devoted to showing that the centered partition is best when $\epsilon \leqq \pi$.

We begin by defining the entropy of a countable partition U of the real line under a probability measure: Let the probabilities of the sets of U be denoted by p_i. Then the entropy $H(U)$ of the partition U is the (Shannon) entropy of the discrete probability distribution $\{p_i\}$, that is

$$H(U) = \sum_i p_i \log p_i^{-1}. \tag{7}$$

The term "epsilon entropy" in the following lemma refers to the definition of

[2]: the epsilon entropy $H_\epsilon(X)$ of a separable metric space X with a probability distribution on the Borel sets is the infimum of the entropies of all partitions of the space by measurable sets of diameters at most ϵ.

For conciseness, the statement of the lemma neglects the behavior of the partition on sets of probability zero. More precisely, the sets of positive probability in an optimal partition can be intervals of length ϵ with sets of probability zero omitted.

LEMMA 1. *Let X be the real line with a probability distribution μ on the Borel sets of X such that μ has a density $p(x)$ which achieves its maximum value at 0, is monotonic on $(0, \infty)$, and even $(p(-x) = p(x))$. Then the ϵ-entropy $H_\epsilon(X)$ of X is attained only by a partition which consists of consecutive intervals of length ϵ (or one which agrees with such a partition on the interval supporting μ if this interval is finite).*

PROOF. Let U be an ϵ-partition of X (partition of X by measurable sets of diameters at most ϵ) which does not consist of consecutive intervals of length ϵ. We shall show that U can be modified to get another ϵ-partition of X of smaller entropy.

We can assume that the sets U_j of U are intervals. For let U_k be a set of maximum probability which is not an interval. If we modify U by replacing U_k by the closed interval which it spans, removing the adjoined points from the other sets, the entropy of U is decreased. This follows from the concavity of the function $p \log p^{-1}$. Furthermore, if any interval of length ϵ contains two of the sets U_j, these can be combined into a single set, which decreases the entropy of U. Thus we can assume that the partition has the property that as j ranges from $-\infty$ to ∞, U_{j+1} is the interval to the right of U_j, 0 lies in U_0, and the length of $U_j \cup U_{j+1}$ is greater than ϵ.

If not all of the intervals of U have length ϵ, we can suppose by symmetry that there is a first $j = j_0$ such that U_j has length less than ϵ, and such that U_{j_0} intersects $(0, \infty)$. Let

$$U_{j_0} = (a, b), \qquad U_{j_0+1} = (b, c),$$

where the assignment of the end points is immaterial. By our assumptions, $b < a + \epsilon < c$. If we replace these two sets by

$$U'_{j_0} = (a, a + \epsilon), \qquad U'_{j_0+1} = (a + \epsilon, c),$$

then the monotonicity and symmetry of $p(x)$ implies that

$$\mu(U'_{j_0}) \geqq \max [\mu(U_{j_0}), \mu(U_{j_0+1})], \tag{8}$$

with equality only if $p(x)$ is zero to the right of b, or $p(x)$ is constant from a to c. The new partition U' has smaller entropy unless we have equality in (8).

If $p(x)$ is zero beyond b, this shows that the lemma is true for the part of U which intersects $(0, \infty)$. Similarly the rest of U also has the stated property.

If $p(x)$ is constant and positive on (a, c) and $c = b + \epsilon$, then $H(U') = H(U)$. The above procedure can be applied to the intervals U'_{j_0+1}, U_{j_0+2} of U', for U'_{j_0+1}

has length less than ϵ. We get a partition with smaller entropy unless $p(x)$ was constant over U_{j_0}, U_{j_0+1}, U_{j_0+2}. If this procedure is applied repeatedly, eventually a pair of intervals will be encountered on which $p(x)$ is not constant. Thus $H(U) > H_\epsilon(X)$. Lemma 1 is proved.

We remark that the hypothesis of unimodality of the distribution is essential for the conclusion of Lemma 1. (The distribution need not be symmetric, however. This assumption was used to simplify the treatment of a partition in which the interval containing zero has length less than ϵ.) In the problem at hand, Lemma 1 implies that, for Gaussian distributions, the epsilon entropy is attained only for a partition by consecutive intervals of length ϵ. We are thus led to the following definition:

DEFINITION. Let X be the real line with the probability distribution of a normal random variable with mean zero and variance 1: Let $h(\epsilon, \alpha)$ be the entropy of the partition of X by intervals of length ϵ centered at the points $\epsilon(k - \alpha)$, $= 0, \pm 1, \pm 2, \cdots$; $h(\epsilon, 0)$ is denoted by $h(\epsilon)$, the entropy of the centered ϵ partition of X.

Lemmas 3 and 6 below show that for any $\epsilon > 0$ we have $h(\epsilon, \alpha) \geqq h(\epsilon)$, with equality if and only if α is an integer. But we first need to define two functions and state some of their properties. Let $P(\epsilon, z)$ be the probability of the interval of length ϵ centered at ϵz, so that

$$P(\epsilon, z) = \int_{(z-\frac{1}{2})\epsilon}^{(z+\frac{1}{2})\epsilon} \phi(y)\, dy = \int_{(z-\frac{1}{2})\epsilon}^{(z+\frac{1}{2})\epsilon} e^{-y^2/2} (2\pi)^{-\frac{1}{2}}\, dy, \tag{9}$$

where ϕ is the normal density function.

Since $P(\epsilon, z) \sim ((z - \frac{1}{2})\epsilon)^{-1} \phi((z - \frac{1}{2})\epsilon)$ for large z, all the series which we encounter will converge absolutely; we need make no further mention of convergence.

Define

$$F(z) = F(\epsilon, z) = \log [P(\epsilon, z - \tfrac{1}{2})/P(\epsilon, z + \tfrac{1}{2})].$$

We observe that $P(\epsilon, z)$ is an even function of z, and $F(z)$ an odd function. Some of the properties of $F(\epsilon z)$ which will be needed are given by the following lemma

LEMMA 2.

(i) $F''(z) > 0, \quad$ *for* $\quad z > 0$;

(ii) $\epsilon^2 > F'(z) \geqq F'(0) = 2\epsilon[\phi(0) - \phi(\epsilon)]/P(\epsilon, \frac{1}{2})$,

(iii) $\epsilon^2(z - \frac{1}{2}) < F(z) < \epsilon^2 z, \quad$ *for* $\quad z > 0$;

(iv) $0 < \partial F(\epsilon, z)/\partial \epsilon < 2F(\epsilon, z)/\epsilon, \quad$ *for* $\quad z > 0$.

PROOF. The definition of $F(z)$ can be rewritten as

$$F(z) = \log [J(z - \tfrac{1}{2})/J(z + \tfrac{1}{2})],$$

where

$$J(z) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\frac{1}{2}\epsilon^2(y+z)^2}\, dy.$$

Then

$$F''(z) = J_2(z - \tfrac{1}{2}) - J_2(z + \tfrac{1}{2}),$$

where

$$J_2(z) = [J(z)J''(z) - J'(z)^2]/J(z)^2.$$

$J_2(z)$ is an even function. It will be shown that $J_2(z)$ is strictly decreasing for $z > 0$. This implies that $F''(z) > 0$ for $z > 0$.

Consider a fixed positive value z_1 of z, and let $J_2(z_1) = \lambda$. We consider the function

$$M(z) = J(z)J''(z) - J'(z)^2 - \lambda J(z)^2.$$

This function can be written as a double integral:

$$\begin{aligned} M(z) &= \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\frac{1}{2}\epsilon^2(x+z)^2}\,dx \int_{-\frac{1}{2}}^{\frac{1}{2}} [\epsilon^4(y+z)^2 - \epsilon^2]e^{-\frac{1}{2}\epsilon^2(y+z)^2}\,dy \\ &\quad - \epsilon^4 \int_{-\frac{1}{2}}^{\frac{1}{2}} (x+z)e^{-\frac{1}{2}\epsilon^2(x+z)^2}\,dz \int_{-\frac{1}{2}}^{\frac{1}{2}} (y+z)e^{-\frac{1}{2}\epsilon^2(y+z)^2}\,dy \\ &\quad - \lambda \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\frac{1}{2}\epsilon^2(x+z)^2}\,dx \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\frac{1}{2}\epsilon^2(y+z)^2}\,dy \\ &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} [\epsilon^4(y-x)(y+z) - \epsilon^2 - \lambda] \exp[-\tfrac{1}{2}\epsilon^2(x+z)^2 \\ &\quad - \tfrac{1}{2}\epsilon^2(y+z)^2]\,dx\,dy. \end{aligned}$$

Now make the substitution $x = u + v$, $y = u - v$. We get

$$M(z) = 2\int\int_S [2\epsilon^4 v(v-u-z) - \epsilon^2 - \lambda] \exp[-\epsilon^2(u^2 + v^2 + 2uz + z^2)]\,du\,dv,$$

where S is the square with vertices $(u, v) = (\pm\frac{1}{2}, 0)$ and $(0, \pm\frac{1}{2})$. Drop the terms which are odd in v, for they have integral zero. Combining the contributions of positive and negative u,

$$M(z) = 8e^{-\epsilon^2 z^2} \int_0^{\frac{1}{2}} \cosh(2\epsilon^2 uz)K(u)\,du,$$

where $$K(u) = e^{-\epsilon^2 u^2} \int_0^{\frac{1}{2}-u} (2\epsilon^4 v^2 - \epsilon^2 - \lambda)e^{-\epsilon^2 v^2}\,dv.$$

We know $M(z_1)$ is zero. Hence $K(u)$ changes sign on $(0, \frac{1}{2})$, and the function $2\epsilon^4 v^2 - \epsilon^2 - \lambda$ must also change sign. Clearly it can only only change from negative to positive as v increases. Thus $K(u)$ is first positive on an interval adjacent to 0, then negative on the rest of $(0, \frac{1}{2})$. In the expression

$$M'(z_1) = 8e^{-\epsilon^2 z_1^2} \int_0^{\frac{1}{2}} 2\epsilon^2 u \sinh(2\epsilon^2 u z_1)K(u)\,du,$$

$K(u)$ is multiplied by a more rapidly increasing function than $\cosh(2\epsilon^2 u z_1)$. Hence $M'(z_1) < 0$. Since

$$J_2'(z_1) = M'(z_1)/J(z_1)^2,$$

and z_1 was an arbitrary positive number, $J_2'(z) < 0$ for $z > 0$. As mentioned above, this proves (i).

The value of $F'(0)$ given in (ii) follows by direct differentiation of the defining formulas for $F(z)$. From the asymptotic form of $P(\epsilon, z)$, it follows easily that

$$F'(z) \to \epsilon^2, \qquad \text{as} \quad z \to \infty.$$

Thus (ii) follows from (i) for $z \geqq 0$, and is also true for $z < 0$ since $F'(z)$ is an even function.

It follows from (ii) that $F(z) - \epsilon^2 z$ is strictly decreasing function. Its value is zero at $z = 0$, and

$$F(z) - \epsilon^2 z \longrightarrow -\tfrac{1}{2}\epsilon^2, \qquad \text{as} \quad z \longrightarrow \infty,$$

from the asymptotic formula for $P(\epsilon, z)$. Hence (iii) is true.

The inequalities of (iv) again follow from finding the signs of certain double integrals. Going back to the expression for $F(\epsilon, z)$ in terms of the function $J(z)$, we have

$$\partial F(\epsilon, z)/\partial\epsilon = M_1(z)/J(z + \tfrac{1}{2})J(z - \tfrac{1}{2}),$$

where

$$\begin{aligned} M_1(z) &= J(z + \tfrac{1}{2})(\partial J(z - \tfrac{1}{2})/\partial\epsilon) - J(z - \tfrac{1}{2})(\partial J(z + \tfrac{1}{2})/\partial\epsilon) \\ &= -\epsilon \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\frac{1}{2}\epsilon^2(x+z+\frac{1}{2})^2}\,dx \int_{-\frac{1}{2}}^{\frac{1}{2}} (y + z - \tfrac{1}{2})^2 e^{-\frac{1}{2}\epsilon^2(y+z-\frac{1}{2})^2}\,dy \\ &\quad + \epsilon \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\frac{1}{2}\epsilon^2(y+z-\frac{1}{2})^2}\,dy \int_{-\frac{1}{2}}^{\frac{1}{2}} (x + z + \tfrac{1}{2})^2 e^{-\frac{1}{2}\epsilon^2(x+z+\frac{1}{2})^2}\,dx \\ &= \epsilon \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} [(x + z + \tfrac{1}{2})^2 - (y + z - \tfrac{1}{2})^2] \\ &\quad \cdot \exp[-\tfrac{1}{2}\epsilon^2(x + z + \tfrac{1}{2})^2 - \tfrac{1}{2}\epsilon^2(y + z - \tfrac{1}{2})^2]\,dx\,dy. \end{aligned}$$

Again we make the substitution $x = u + v$, $y = u - v$. We get

$$\begin{aligned} M_1(z) = 2\epsilon \int\int_S &[4uv + 4vz + 2u + 2z] \\ &\cdot \exp[-\epsilon^2(u^2 + v^2 + 2uz + v + z^2 + \tfrac{1}{4})]\,du\,dv. \end{aligned}$$

Combining the contributions of positive and negative v,

$$\begin{aligned} M_1(z) &= 8\epsilon \int_0^{\frac{1}{2}} e^{-\epsilon^2(v^2+\frac{1}{4})}[\cosh \epsilon^2 v - 2v \sinh \epsilon^2 v][\textstyle\int_{-\frac{1}{2}+v}^{\frac{1}{2}-v} (u + z)e^{-\epsilon^2(u+z)^2}\,du]\,dv \\ &= 8\epsilon^{-1} \int_0^{\frac{1}{2}} e^{-\epsilon^2(2v^2-v+z^2+\frac{1}{4})}[\cosh \epsilon^2 v - 2v \sinh \epsilon^2 v] \sinh[\epsilon^2 z(1 - 2v)]\,dv \\ &> 0, \end{aligned}$$

for $z > 0$. This proves the first inequality of (iv).

Now we put $F(z)$ in the form

$$F(\epsilon, z) = -\int_{z-\frac{1}{2}}^{z+\frac{1}{2}} J'(\zeta)/J(\zeta)\,d\zeta = \epsilon^2 \int_{z-\frac{1}{2}}^{z+\frac{1}{2}} J_1(\zeta)/J(\zeta)\,d\zeta$$

where
$$J_1(z) = \int_{-\frac{1}{2}}^{\frac{1}{2}} (y + z)e^{-\frac{1}{2}\epsilon^2(y+z)^2}\,dy\,.$$

Then
$$2\epsilon^{-1}F(z) - \partial F(z)/\partial\epsilon = \epsilon^2 \int_{z-\frac{1}{2}}^{z+\frac{1}{2}} M_2(\zeta)/J(\zeta)^2\,d\zeta,$$

where
$$M_2(z) = J_1(z)(\partial J(z)/\partial\epsilon) - J(z)(\partial J_1(z)/\partial\epsilon).$$

It will be shown that $M_2(z)$ is an odd function which is positive for $z > 0$. Since $J(z)$ is even, this implies that the above integral is positive for $z > 0$, proving the last inequality of (iv).

We have

$$\begin{aligned} M_2(z) &= -\epsilon \int_{-\frac{1}{2}}^{\frac{1}{2}} (x + z)e^{-\frac{1}{2}\epsilon^2(x+z)^2}\,dx \int_{-\frac{1}{2}}^{\frac{1}{2}} (y + z)^2 e^{-\frac{1}{2}\epsilon^2(y+z)^2}\,dy \\ &\quad + \epsilon \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\frac{1}{2}\epsilon^2(x+z)^2}\,dx \int_{-\frac{1}{2}}^{\frac{1}{2}} (y + z)^3 e^{-\frac{1}{2}\epsilon^2(y+z)^2}\,dy \\ &= \epsilon \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} (y - x)(y + z)^2 \exp\left[-\tfrac{1}{2}\epsilon^2(x + z)^2 - \tfrac{1}{2}\epsilon^2(y + z)^2\right] dx\,dy \\ &= 4\epsilon \int\int_S -v(u - v + z)^2 \exp\left[-\epsilon^2(u + z)^2 - \epsilon^2 v^2\right] du\,dv \\ &= 8\epsilon \int\int_S v^2(u + z) \exp\left[-\epsilon^2(u + z)^2 - \epsilon^2 v^2\right] du\,dv. \end{aligned}$$

Integrating over u first,

$$\begin{aligned} M_2(z) &= 16\epsilon \int_0^{\frac{1}{2}} v^2 e^{-\epsilon^2 v^2}\left[\int_{-\frac{1}{2}+v}^{\frac{1}{2}-v} (u + z)e^{-\epsilon^2(u+z)^2}\,du\right] dv \\ &= 16\epsilon^{-1} \int_0^{\frac{1}{2}} v^2 \exp\left[-\epsilon^2 v^2 - \epsilon^2 z^2 - \epsilon^2(\tfrac{1}{2} - v)^2\right] \sinh\left[\epsilon^2 z(1 - 2v)\right] dv. \end{aligned}$$

This is clearly an odd function of z, positive if $z > 0$.

This completes the proof of Lemma 2.

The next lemma proves Theorem 1 for large ϵ. However, the difficult case is the case of small ϵ.

Lemma 3. *If* $\epsilon \geqq 3$, $h(\epsilon, \alpha)$ *assumes its minimum value only when* α *is an integer.*

Proof. Note that $h(\epsilon, \alpha)$ is an even periodic function of period 1. Thus it is sufficient to show that $\partial h(\epsilon, \alpha)/\partial\alpha > 0$ for $0 < \alpha < \frac{1}{2}$.

The proof that $\partial h/\partial\alpha > 0$, $\epsilon \geqq 3$ and $0 < \alpha < \frac{1}{2}$ consists of two parts: first we prove the result for an interval of the form $0 < \alpha \leqq \alpha_0$, then for an interval $\alpha_0 \leqq \alpha < \frac{1}{2}$. In either interval, we have

$$h(\epsilon, \alpha) = \sum_{k=-\infty}^{\infty} P(\epsilon, k - \alpha) \log\left[P(\epsilon, k - \alpha)\right]^{-1}.$$

So

$$\partial h/\partial\alpha = \epsilon \sum_{-\infty}^{\infty} \left[\phi((k + \tfrac{1}{2} - \alpha)\epsilon) - \phi((k - \tfrac{1}{2} - \alpha)\epsilon)\right] \log P(\epsilon, k - \alpha)$$

$$\partial h/\partial\alpha = \epsilon \sum_{-\infty}^{\infty} \phi((k + \tfrac{1}{2} - \alpha)\epsilon)F(k + \tfrac{1}{2} - \alpha), \tag{10}$$

$$\begin{aligned} \partial h/\partial\alpha = \epsilon \sum_{k=0}^{\infty} \{&\phi[(k + \tfrac{1}{2} - \alpha)\epsilon]F(k + \tfrac{1}{2} - \alpha) \\ &- \phi[(k + \tfrac{1}{2} + \alpha)\epsilon]F(k + \tfrac{1}{2} + \alpha)\}. \end{aligned} \tag{11}$$

The kth term in this series is $\phi[k + \frac{1}{2} + \alpha)\epsilon]G_k(\alpha)$, where

$$G_k(\alpha) = e^{\alpha\epsilon^2(2k+1)}F(k + \tfrac{1}{2} - \alpha) - F(k + \tfrac{1}{2} + \alpha).$$

By Lemma 2, (ii), $F(k + \frac{1}{2} + \alpha) < F(k + \frac{1}{2} - \alpha) + 2\alpha\epsilon^2$; hence

$$G_k(\alpha) > \left[e^{\alpha\epsilon^2(2k+1)} - 1\right]F(k + \tfrac{1}{2} - \alpha) - 2\alpha\epsilon^2. \tag{12}$$

This expression is clearly an increasing function of $k \geqq 0$, for $\alpha > 0$. Thus the terms in (11) are all positive if

$$(e^{\alpha\epsilon^2} - 1)F(\tfrac{1}{2} - \alpha) - 2\alpha\epsilon^2 > 0,$$

or

$$\left[(e^{\alpha\epsilon^2} - 1)(\alpha\epsilon^2)^{-1}\right] F(\tfrac{1}{2} - \alpha) > 2. \tag{13}$$

The two factors on the left are both increasing functions of ϵ for $0 < \alpha < \frac{1}{2}$, using (iv) of Lemma 2. Thus $\partial h/\partial\alpha > 0$ for $\epsilon \leqq 3$ if $(e^{9\alpha} - 1)(9\alpha)^{-1}F(3, \frac{1}{2} - \alpha) > 2$. By (i) of Lemma 2, $F(3, \frac{1}{2} - \alpha) > (\frac{1}{2} - \alpha)F'(3, 0)$. Hence it is sufficient to show that

$$(e^{9\alpha} - 1)(9\alpha)^{-1}(\tfrac{1}{2} - \alpha) > 2/F'(3, 0). \tag{14}$$

It is easily verified that the function on the left in (14) has one relative maximum on $(0, \infty)$, and no other stationary points. Thus the set of positive α for which (14) is true is an interval. The left side takes the values $\frac{1}{2}$ at $\alpha = 0$, and $(e^4 - 1)/72 = 0.74 \cdots$ at $\alpha = 4/9$, both of which are greater than

$$2/F'(3, 0) = \tfrac{1}{3}(2\pi)^{\frac{1}{2}} \cdot (1 - e^{-9/2})^{-1} \int_0^3 \phi(x)\, dx = 0.42 \cdots .$$

Hence

$$\partial h/\partial\alpha > 0, \qquad \text{for} \quad \epsilon \geqq 3, \qquad 0 < \alpha \leqq 4/9. \tag{15}$$

We now turn to the case in which α is close to $\frac{1}{2}$. It is convenient to rearrange the series for $\partial h/\partial\alpha$ (equation (10)) as follows:

$$\begin{aligned} \partial h/\partial\alpha &= \epsilon \textstyle\sum_{-\infty}^{\infty} \phi((k + \frac{1}{2} - \alpha)\epsilon)F(\epsilon, k + \frac{1}{2} - \alpha) \\ &= \epsilon \textstyle\sum_{-\infty}^{\infty} \phi((k + \beta)\epsilon)F(\epsilon, k + \beta), \end{aligned} \tag{10'}$$

where we have set $\beta = \frac{1}{2} - \alpha$; so $0 < \beta < \frac{1}{2}$. Continuing with (10′), we have

$$\begin{aligned} \partial h/\partial\alpha = {} & \epsilon\phi(\epsilon\beta)F(\epsilon, \beta) \\ & + \epsilon \textstyle\sum_{k=1}^{\infty} \{\phi((k + \beta)\epsilon)F(k + \beta) - \phi((k - \beta)\epsilon)F(k - \beta)\}. \end{aligned}$$

Thus,

$$\begin{aligned} & (\epsilon\phi(\epsilon\beta))^{-1}(\partial h/\partial\alpha) \\ & \quad = F(\epsilon, \beta) - \textstyle\sum_1^{\infty} e^{-\epsilon^2k^2/2}\{e^{\epsilon^2k\beta}F(k - \beta) - e^{-\epsilon^2k\beta}F(k + \beta)\}. \end{aligned} \tag{16}$$

Now by (ii) and (iii) of Lemma 2,

$$\begin{aligned} e^{+\epsilon^2k\beta}F(k - \beta) - e^{-\epsilon^2k\beta}F(k + \beta) &< (e^{\epsilon^2k\beta} - e^{-\epsilon^2k\beta})F(k) \\ &< 2\epsilon^2k \sinh \epsilon^2k\beta. \end{aligned}$$

Returning to equation (16), we find that

$$\begin{aligned} (\epsilon\phi(\epsilon\beta))^{-1}(\partial h/\partial\alpha) &> F(\epsilon, \beta) - \textstyle\sum_1^{\infty} e^{-\epsilon^2k^2/2}2\epsilon^2k \sinh \epsilon^2k\beta \\ &= F(\epsilon, \beta) - \textstyle\sum_1^{\infty} A_k \qquad \text{say}. \end{aligned} \tag{17}$$

We have

$$A_{k+1}/A_k = e^{-\epsilon^2(k+\frac{1}{2})}(k + 1)k^{-1}(\sinh \epsilon^2(k + 1)\beta)/(\sinh \epsilon^2k\beta).$$

Since the function ctnh z is decreasing for $z > 0$,

$$\begin{aligned}(\sinh \epsilon^2(k+1)\beta)/(\sinh \epsilon^2 k\beta) &= \cosh \epsilon^2\beta + \sinh \epsilon^2\beta \operatorname{ctnh} \epsilon^2 k\beta \\ &\leqq \cosh \epsilon^2\beta + \sinh \epsilon^2\beta \operatorname{ctnh} \epsilon^2\beta \\ &= 2\cosh \epsilon^2\beta \\ &\leqq 2e^{\epsilon^2\beta},\end{aligned}$$

for $k \geqq 1$. For $0 < \beta < \frac{1}{2}$, this last function is less than $2e^{\frac{1}{2}\epsilon^2}$, and

$$A_{k+1}/A_k < 2(k+1)k^{-1}e^{-\epsilon^2 k} \leqq 4e^{-\epsilon^2} \leqq 4e^{-9},$$

for $\epsilon \geqq 3$. Therefore $\sum_1^\infty A_k \leqq A_1/(1-4e^{-9})$. Combining this with inequality (17), we see that $\partial h/\partial\alpha > 0$ whenever

$$(18) \qquad F(\epsilon, \beta) > 2\epsilon^2 e^{-\epsilon^2/2}(1-4e^{-9})^{-1} \sinh \epsilon^2\beta.$$

A simple calculation shows that the right-hand side of the above inequlity is a decreasing function of ϵ for fixed $0 < \beta < \frac{1}{2}$. We have already shown that the left hand side is an increasing function of ϵ ((iv) of Lemma 2). So again, it is only necessary to consider the case $\epsilon = 3$.

By Lemma 2, (ii),

$$F(\epsilon, \beta) \geqq \beta F'(0) = 2\epsilon\beta(\phi(0) - \phi(\epsilon))/P(\epsilon, \tfrac{1}{2}).$$

Combining this with (18), we see that $\partial h/\partial\alpha > 0$ for $\epsilon \geqq 3$ if

$$6\beta(\phi(0) - \phi(3))/P(3, \tfrac{1}{2}) > 18e^{-9/2}(1-4e^{-9})^{-1} \sinh 9\beta,$$

or

$$(19) \quad (\sinh 9\beta)(9\beta)^{-1} < (e^{9/2}-1)(1-4e^{-9})[27(2\pi)^{\frac{1}{2}} \textstyle\int_0^3 \phi(x)\,dx]^{-1} = 2.63\cdots.$$

This is true for $\beta \leqq \frac{1}{4}$. Thus

$$(20) \qquad \partial h/\partial\alpha > 0, \qquad \text{for } \epsilon \geqq 3, \qquad \tfrac{1}{4} \leqq \alpha < \tfrac{1}{2}.$$

By combining this result with (15) we have finally proved Lemma 3.

To complete the proof of Theorem 1, we shall have to study the function $h(\epsilon, \alpha)$ very carefully. This is because for small ϵ

$$(\partial/\partial\alpha)h(\epsilon, \alpha) = O(e^{-2\pi^2/\epsilon^2}),$$

so that h is very flat as $\epsilon \to 0$.

The rapid convergence of the series for $h(\epsilon, \alpha)$ ensures that it is C^∞. From the periodicity of the function in α, it follows that it is the sum of a convergent Fourier series:

$$(21) \qquad h(\epsilon, \alpha) = \tfrac{1}{2}C_0(\epsilon) + \textstyle\sum_{n=1}^\infty C_n(\epsilon)\cos(2n\pi\alpha),$$

where

$$\begin{aligned}C_n(\epsilon) &= 2\textstyle\int_{-\frac{1}{2}}^{\frac{1}{2}} h(\epsilon, \alpha)\cos(2n\pi\alpha)\,d\alpha \\ &= 2\textstyle\int_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{k=-\infty}^{\infty} P(\epsilon, k-\alpha)\log[P(\epsilon, k-\alpha)]^{-1}\cos(2n\pi\alpha)\,d\alpha.\end{aligned}$$

We interchange the order of integration and summation here; after the substitution $k - \alpha = x$, we have

(22) $$C_n(\epsilon) = 2\int_{-\infty}^{\infty} P(\epsilon, x) \log [P(\epsilon, x)]^{-1} \cos (2n\pi x)\, dx.$$

To get useful inequalities for these coefficients we need to investigate the properties of $P(\epsilon, z)$ as an entire function of the complex variable z.

Define

(23) $$Q(\epsilon, z) = (2\pi)^{\frac{1}{2}}\epsilon^{-1}e^{\frac{1}{2}z^2/\epsilon^2}P(\epsilon, z/\epsilon^2), \qquad \text{so that} \quad Q(\epsilon, z) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-zy-\frac{1}{2}\epsilon^2y^2}\, dy,$$

which shows that $Q(\epsilon, z)$ is an even entire function of z of exponential type. Hence it can be expressed in terms of the canonical product of its zeros $\pm\zeta_1, \pm\zeta_2, \cdots$, as

(24) $$Q(\epsilon, z) = Q(\epsilon, 0)\prod_{k=1}^{\infty} (1 - z^2/\zeta_k^2).$$

Thus, information about the zeros of $Q(\epsilon, z)$ would be quite useful, and the next lemma furnishes the required information.

LEMMA 4. *The zeros* $\{\pm\zeta_k\}$ *of* $Q(\epsilon, z)$ *are all distinct and are on the imaginary axis for* $0 < \epsilon < \epsilon_0 = 4.309 \cdots$. *Furthermore, under the appropriate indexing, we have*

$$2\pi k < \zeta_k/i < 2\pi(k+1), \qquad k = 1, 2, \cdots.$$

PROOF. As $\epsilon \to 0$, $Q(\epsilon, z) \to (2/z)\sinh(z/2)$, with zeros at $\pm 2\pi i, \pm 4\pi i, \cdots$. Integrating by parts twice in (23), we get

(25) $$(1 + \epsilon^2/z^2)Q(\epsilon, z) = e^{-1/8\epsilon^2}[(2/z)\sinh(z/2) + (\epsilon^2/z^2)\cosh(z/2)] + (2\epsilon^4/z^2)\int_0^{\frac{1}{2}} y^2 \cosh yz e^{-\frac{1}{2}\epsilon^2y^2}\, dy.$$

The integral in (25) is bounded by $\frac{1}{8}e^{\frac{1}{2}|\mathrm{Re}\, z|}$. Hence for small ϵ, and $|z|$ not small, the last term on the right in (25) is small compared to the sum of the first two, except in small neighborhoods of the zeros of

$$(2/z)\sinh(z/2) + (\epsilon^2/z^2)\cosh(z/2),$$

or equivalently, except in small neighborhoods (for ϵ small) of the zeros of $(2/z)\sinh(z/2)$. Hence, as $\epsilon \to 0$, the zeros of $Q(\epsilon, z)$ either converge to the zeros of $(2/z)\sinh(z/2)$, or else approach infinity. It follows that for ϵ sufficiently small, the ζ_k are all distinct, and can be indexed so that $\zeta_k \to 2\pi ik$ as $\epsilon \to 0$.

For any ϵ, if z is sufficiently large, the first term on the right in (25) is dominant in the sense described above. Thus when ϵ varies, no zeros can appear or disappear at infinity.

The set of zeros of $Q(\epsilon, z)$ must be symmetric with respect to the imaginary axis as well as to the real axis. Since this set varies continuously with ϵ, it is easily seen that the ζ_k must lie on the imaginary axis up to the first value ϵ_0 of ϵ at which two zeros coincide.

Let primes denote differentiation with respect to z. Then

$$Q'(\epsilon, z) = 2\int_0^{\frac{1}{2}} y \sinh zy e^{-\frac{1}{2}\epsilon^2y^2}\, dy,$$

and, by integration by parts,

$$Q(\epsilon, z) = (2/z)e^{-\frac{1}{8}\epsilon^2} \sinh (z/2) + (2\epsilon^2/z) \int_0^{\frac{1}{2}} y \sinh zye^{-\frac{1}{2}\epsilon^2 y^2}\, dy.$$

Comparing these two equations, we see that for $\epsilon > 0$

$$Q'(\epsilon, \zeta_k) = -2\epsilon^{-2}e^{-\frac{1}{8}\epsilon^2} \sinh (\zeta_k/2). \tag{26}$$

Hence a zero of $Q(\epsilon, z)$ is not simple if and only if it is one of the points $\pm 2n\pi i$, $n = 1, 2, \cdots$; and ϵ_0 is the first positive value of ϵ at which one of the numbers

$$q_n(\epsilon) = \int_0^{\frac{1}{2}} \cos (2\pi ny)e^{-\frac{1}{2}\epsilon^2 y^2}\, dy, \qquad n = 1, 2, \cdots, \tag{27}$$

is zero.

If we differentiate (27) with respect to ϵ, then integrate by parts, a differential equation satisfied by $q_n(\epsilon)$ is obtained:

$$\epsilon q_n'(\epsilon) + (1 - 4\pi^2 n^2/\epsilon^2)q_n(\epsilon) = \tfrac{1}{2}(-1)^n e^{-\frac{1}{8}\epsilon^2}.$$

The constant of integration in the solution of this equation is fixed by the condition

$$q_n(\epsilon) \sim (2\epsilon)^{-1}(2\pi)^{\frac{1}{2}}, \qquad \text{as} \quad \epsilon \to \infty.$$

Thus we have

$$q_n(\epsilon) = (2\epsilon)^{-1}e^{-2\pi^2 n^2/\epsilon^2}\{(2\pi)^{\frac{1}{2}} - (-1)^n \int_\epsilon^\infty e^{-\frac{1}{8}v^2 + 2\pi^2 n^2/v^2}\, dv\}.$$

We see that $q_n(\epsilon)$ has no positive zeros if n is odd, while for n even there is a unique positive zero at

$$\int_\epsilon^\infty e^{-\frac{1}{8}v^2 + 2\pi^2 n^2/v^2}\, dv = (2\pi)^{\frac{1}{2}}.$$

This zero is clearly an increasing function of n. Hence ϵ_0 is the zero of $q_2(\epsilon)$, so that

$$\int_{\epsilon_0}^\infty e^{-\frac{1}{8}v^2 + 8\pi^2/v^2}\, dv = (2\pi)^{\frac{1}{2}}.$$

By numerical integration it is found that $\epsilon_0 = 4.309 \cdots$.

Differentiating with respect to ϵ, we have

$$(\partial/\partial\epsilon)Q(\epsilon, z) = -2\epsilon \int_0^{\frac{1}{2}} y^2 \cosh zye^{-\frac{1}{2}\epsilon^2 y^2}\, dy.$$

Comparing with equation (25), we have

$$(\partial/\partial\epsilon)Q(\epsilon, z)|_{z=\zeta_k} = \epsilon^{-3}e^{-\frac{1}{8}\epsilon^2}[2\zeta_k \sinh (\zeta_k/2) + \epsilon^2 \cosh (\zeta_k/2)]. \tag{28}$$

The differential equation satisfied by ζ_k as a function of ϵ is

$$0 = (d/d\epsilon)Q(\epsilon, \zeta_k) = (\partial/\partial\epsilon)Q(\epsilon, z)|_{z=\zeta_k} + Q'(\epsilon, \zeta_k)(d\zeta_k/d\epsilon).$$

Using (26) and (28), we get

$$\epsilon(d\zeta_k/d\epsilon) = \zeta_k + \tfrac{1}{2}\epsilon^2 \operatorname{ctnh} (\zeta_k/2).$$

Put

$$\zeta_k = 2\pi ki + i\delta.$$

Then δ is real for $0 < \epsilon < \epsilon_0$, and $\delta \to 0$ as $\epsilon \to 0$. This $\delta(\epsilon)$ satisfies the differential equation

$$\epsilon(d\delta/d\epsilon) = 2\pi k + \delta - \tfrac{1}{2}\epsilon^2 \operatorname{ctn}(\delta/2). \tag{29}$$

Since $\delta(0+)$ must be zero, δ cannot enter the interval $(-\pi/2, 0)$, for at all points of this interval the right-hand side of (29) is positive. Hence we see that $0 < \delta < 2\pi$ for $0 < \epsilon < \epsilon_0$, which verifies (30):

$$2\pi k < \zeta_k/i < 2\pi(k+1), \qquad k = 1, 2, \cdots. \tag{30}$$

This completes the proof of Lemma 4.

Next, Lemma 4 will be applied to get estimates for the Fourier coefficients $C_n(\epsilon)$ of $h(\epsilon, \alpha)$. This is the content of Lemma 5.

LEMMA 5. *If* $0 < \epsilon < \epsilon_0$,

$$C_1(\epsilon) \leqq -2e^{-2\pi^2/\epsilon^2}[1 - P(\epsilon, 0)], \tag{31}$$

and, for $n \geqq 2$,

$$|C_n(\epsilon)| \leqq e^{-2\pi^2 n^2/\epsilon^2}[2 + 4P(\epsilon, 0)] + 2\epsilon n^{-1}(2\pi)^{-\frac{1}{2}} \textstyle\sum_{k=1}^{n-1} \exp[-2\pi^2(2nk - k^2)\epsilon^{-2}]. \tag{32}$$

PROOF. Integrating by parts in (22), we get

$$C_n(\epsilon) = -(n\pi)^{-1} \int_{-\infty}^{\infty} \sin(2n\pi x)[\log[P(\epsilon, x)]^{-1} - 1](d/dx)P(\epsilon, x)\, dx.$$

The derivative of $P(\epsilon, x)$ can be expressed as

$$\begin{aligned}(d/dx)P(\epsilon, x) &= \epsilon[\phi((x + \tfrac{1}{2})\epsilon) - \phi((x - \tfrac{1}{2})\epsilon)] \\ &= \epsilon(2\pi)^{-\frac{1}{2}}\{\exp[-\tfrac{1}{2}\epsilon^2(x + \tfrac{1}{2})^2] - \exp[-\tfrac{1}{2}\epsilon^2(x - \tfrac{1}{2})^2]\}.\end{aligned}$$

Using the exponential formula for the sine function, we obtain

$$C_n(\epsilon) = i\epsilon(2n\pi(2\pi)^{\frac{1}{2}})^{-1} \textstyle\sum_{r,s=\pm 1} rs \int_{-\infty}^{\infty} \exp[2\pi i r n x - \tfrac{1}{2}\epsilon^2(x + s/2)^2] \cdot [\log[P(\epsilon, x)]^{-1} - 1]\, dx. \tag{33}$$

The function $\log[1/P(\epsilon, x)]$ is analytic throughout the complex plane except at the points $x = \pm\zeta_k/\epsilon^2$, $k = 1, 2, \cdots$. Take the branch of this function which is real on the real axis and with cuts along horizontal lines from the poles to $+\infty$. In (33), move the path of integration to the line $\operatorname{Im} x = 2\pi n r/\epsilon^2$. The result is the integral along this line, together with a contribution from each pole of $\log P^{-1}$ that has been crossed, due to the jump in the logarithm across the cut. We obtain

$$\begin{aligned}C_n(\epsilon) = i\epsilon(2n\pi(2\pi)^{\frac{1}{2}})^{-1} \textstyle\sum_{r,s=\pm 1} rs \{ &-2\pi i r \textstyle\sum_{k<n} \int_{r\zeta_k/\epsilon^2}^{\infty} \\ &\exp[2\pi i r n x - \tfrac{1}{2}\epsilon^2(x + s/2)^2]\, dx + \int_{-\infty+2\pi n r i/\epsilon^2}^{+\infty+2\pi n r i/\epsilon^2} \exp[2\pi i r n x - \tfrac{1}{2}\epsilon^2(x \\ &+ s/2)^2][\log[P(\epsilon, x)]^{-1} - 1]\, dx\},\end{aligned} \tag{34}$$

by (30).

The integrals from the poles to ∞ in (34) can be combined after the substitution $x = -\frac{1}{2}s + ry$. In the integral along the line $\operatorname{Im} x = 2\pi nr/\epsilon^2$, replace x by $(2\pi nri/\epsilon^2) + u$. The result is

$$C_n(\epsilon) = -2\epsilon(-1)^n(n(2\pi)^{\frac{1}{2}})^{-1} \sum_{k=1}^{n-1} \int_{\zeta_k/\epsilon^2-\frac{1}{2}}^{\zeta_k/\epsilon^2+\frac{1}{2}} e^{2\pi iny-\frac{1}{2}\epsilon^2y^2}\, dy$$
$$+ i\epsilon(-1)^n(2n\pi(2\pi)^{\frac{1}{2}})^{-1}e^{-2\pi^2n^2/\epsilon^2} \sum_{s=\pm1} s \int_{-\infty}^{\infty} e^{-\frac{1}{2}\epsilon^2(u+s/2)^2}$$
$$\cdot \log\,[P(\epsilon, u - 2\pi in/\epsilon^2)/P(\epsilon, u + 2\pi in/\epsilon^2)]\, du.$$

In the second integral, express P in terms of Q by (23). We find that

$$\begin{aligned}\log\,[P(\epsilon, u - 2\pi in/\epsilon^2)/P(\epsilon, u + 2\pi in/\epsilon^2)] &= 4\pi inu - 2i \arg Q(\epsilon, \epsilon^2u + 2\pi in)\\ &= 2\pi i(n - 1 + 2nu) - 2i \arg\,[(-1)^{n-1}Q(\epsilon, \epsilon^2u + 2\pi in)],\end{aligned}$$

where the branch of the last argument is zero at $u = 0$. Hence

$$\begin{aligned}C_n(\epsilon) = \epsilon(-1)^n(n\pi(2\pi)^{\frac{1}{2}})^{-1}e^{-2\pi^2n^2/\epsilon^2} &\sum_{s=\pm1} s \int_{-\infty}^{\infty} e^{-\frac{1}{2}\epsilon^2(u+s/2)^2}\\ &\cdot \arg\,[(-1)^{n-1}Q(\epsilon, \epsilon^2u + 2\pi in)]\, du\\ &+ 2(-1)^n e^{-2\pi^2n^2/\epsilon^2} - 2\epsilon(-1)^n(n(2\pi)^{\frac{1}{2}})^{-1} \sum_{k<n} I_{nk}\end{aligned} \tag{35}$$

where

$$I_{nk} = \int_{\zeta_k/\epsilon^2-\frac{1}{2}}^{\zeta_k/\epsilon^2+\frac{1}{2}} e^{2\pi iny-\frac{1}{2}\epsilon^2y^2}\, dy. \tag{36}$$

First we consider

$$C_1(\epsilon) = e^{-2\pi^2/\epsilon^2}[-2 - \epsilon(\pi(2\pi)^{\frac{1}{2}})^{-1} \sum_{s=\pm1} s \int_{-\infty}^{\infty} e^{-\frac{1}{2}\epsilon^2(u+s/2)^2} \cdot \arg Q(\epsilon, \epsilon^2u + 2\pi i)\, du]. \tag{37}$$

In the integral, Q is evaluated on a line which lies below the first cut in the upper half plane. Hence, the argument is an odd function of u, and the sum of integrals is just J_1, say, where

$$J_1 = 2\int_0^{\infty} [e^{-\frac{1}{2}\epsilon^2(u+\frac{1}{2})^2} - e^{-\frac{1}{2}\epsilon^2(u-\frac{1}{2})^2}] \arg Q(\epsilon, \epsilon^2u + 2\pi i)\, du. \tag{38}$$

The function $Q(\epsilon, z + 2\pi i)$ is positive at $z = 0$, while for z real and positive

$$\operatorname{Im} Q(\epsilon, z + 2\pi i) = 2\int_0^{\frac{1}{2}} \sin 2\pi y \sinh zy e^{-\frac{1}{2}\epsilon^2y^2}\, dy > 0,$$

since the integrand is positive. Thus the argument in (38) lies between 0 and π, and

$$\begin{aligned}|J_1| &< 2\pi \int_0^{\infty} [e^{-\frac{1}{2}\epsilon^2(u-\frac{1}{2})^2} - e^{-\frac{1}{2}\epsilon^2(u+\frac{1}{2})^2}]\, du\\ &= 2\pi \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\frac{1}{2}\epsilon^2u^2}\, du = 2\pi(2\pi)^{\frac{1}{2}}\epsilon^{-1}P(\epsilon, 0).\end{aligned}$$

It follows from (37) that (31) is true.

To estimate $C_n(\epsilon)$ for $n \geqq 2$, we need to first estimate

$$J_n = \int_{-\infty}^{\infty} [e^{-\frac{1}{2}\epsilon^2(u+\frac{1}{2})^2} - e^{-\frac{1}{2}\epsilon^2(u-\frac{1}{2})^2}] \arg\,[(-1)^{n-1}Q(\epsilon, \epsilon^2u + 2\pi in)]\, du. \tag{39}$$

From the product expansion (24),

$$\arg[(-1)^{n-1}Q(\epsilon, x + 2\pi in)] = \sum_{k<n} \arg[(2\pi in + x)^2/\zeta_k^2 - 1] + \sum_{k\geqq n} \arg[1 - (2\pi in + x)^2/\zeta_k^2],$$

where the branch of each argument is zero at $x = 0$. Any linear factor of Q has argument which varies by $\pi/2$ on each of the half lines $0 < x < \infty$ and $-\infty < x < 0$.

Thus, putting $\eta_k = \operatorname{Im} \zeta_k$,

$$|\arg (-1)^{n-1}Q(\epsilon, x + 2\pi in)| \leqq n\pi + \sum_{k>n} \tan^{-1}[4\pi n|x|/(\eta_k^2 - 4\pi^2n^2 + x^2)]$$

Note that, since $\eta_k > 2\pi k$ by (30),

$$\eta_k^2 - 4\pi^2n^2 > 4\pi^2(k - n)^2,$$

and

$$\begin{aligned}\sum_{k>n} \tan^{-1}[4\pi n|x|/(\eta_k^2 - 4\pi^2n^2 + x^2)] &\leqq \sum_{j=1}^{\infty} \tan^{-1}[4\pi n|x|/(4\pi^2j^2 + x^2)] \\ &\leqq \sum_{j=1}^{\infty} 4\pi n|x|/(4\pi^2j^2 + x^2) \\ &\leqq \int_0^{\infty} 4\pi n|x|/(4\pi^2v^2 + x^2)\,dv = n\pi.\end{aligned}$$

Thus the argument in (39) is bounded in absolute value by $2n\pi$, and

$$\begin{aligned}|J_n| &\leqq 2n\pi \int_{-\infty}^{\infty} |e^{-\frac{1}{2}\epsilon^2(u+\frac{1}{2})^2} - e^{-\frac{1}{2}\epsilon^2(u-\frac{1}{2})^2}|\,du \\ &= 4n\pi \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\frac{1}{2}\epsilon^2u^2}\,du = 4n\pi(2\pi)^{\frac{1}{2}}\epsilon^{-1}P(\epsilon, 0).\end{aligned}$$

Now from (35)

$$|C_n(\epsilon)| \leqq e^{-2\pi^2n^2/\epsilon^2}[4P(\epsilon, 0) + 2] + 2\epsilon(n(2\pi)^{\frac{1}{2}})^{-1}\sum_{k<n}|I_{nk}|. \tag{40}$$

In formula (36) for I_{nk}, put $y = (\zeta_k/\epsilon^2) + v$:

$$\begin{aligned}|I_{nk}| &\leqq \int_{-\frac{1}{2}}^{\frac{1}{2}} \exp[\operatorname{Re}(2\pi iny - \tfrac{1}{2}\epsilon^2y^2)]\,dv \\ &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \exp[-\tfrac{1}{2}\epsilon^2v^2 + \tfrac{1}{2}\eta_k^2\epsilon^{-2} - 2\pi n\eta_k\epsilon^{-2}]\,dv \\ &\leqq \exp[\tfrac{1}{2}\eta_k^2\epsilon^{-2} - 2\pi n\eta_k\epsilon^{-2}],\end{aligned}$$

where $\zeta_k = i\eta_k$. By (30), $\eta_k > 2\pi k$. Hence

$$|I_{nk}| \leqq \exp[-2\pi^2(2nk - k^2)/\epsilon^2],$$

which, with (40), yields (32). Lemma 5 is proved.

LEMMA 6. *For $0 < \epsilon \leqq \pi$, $h(\epsilon, \alpha) > h(\epsilon)$ when α is not an integer.*

PROOF. From the Fourier series for $h(\epsilon, \alpha)$,

$$\begin{aligned}h(\epsilon, \alpha) - h(\epsilon) &= \sum_{n=1}^{\infty} C_n(\epsilon)[\cos 2n\pi\alpha - 1] \\ &= (1 - \cos 2\pi\alpha)[-C_1(\epsilon) - \sum_{n=2}^{\infty} C_n(\epsilon)(1 - \cos 2n\pi\alpha)/(1 - \cos 2\pi\alpha)].\end{aligned}$$

It is enough to show that this difference is positive for $0 < \alpha \leqq \frac{1}{2}$. To do this, we will show that

$$-C_1(\epsilon) > \sum_{n=2}^{\infty} C_n(\epsilon)(1 - \cos 2n\pi\alpha)/(1 - \cos 2\pi\alpha).$$

Since

$$(1 - \cos 2n\pi\alpha)/(1 - \cos 2\pi\alpha) = (\sin n\pi\alpha/\sin \pi\alpha)^2 \leqq n^2,$$

it is sufficient to show

$$(41) \qquad -C_1(\epsilon) > \sum_{n=2}^{\infty} n^2 |C_n(\epsilon)|.$$

Using (32), we see that

$$\sum_{n=2}^{\infty} n^2 |C_n(\epsilon)| \leqq [2 + 4P(\epsilon, 0)] \sum_{n=2}^{\infty} e^{-2\pi^2 n^2/\epsilon^2} n^2 + 2\epsilon(2\pi)^{-\frac{1}{2}} \sum_{k=1}^{\infty} \sum_{n=k+1}^{\infty} n \exp[-2\pi^2(2nk - k^2)\epsilon^{-2}].$$

In estimating these sums, we use the fact that when the ratio of successive terms in a series decreases, the sum is less than the sum of the geometric series with the same first two terms. Thus

$$\begin{aligned}\sum_{n=k+1}^{\infty} n \exp[-2\pi^2(2nk - k^2)/\epsilon^2] &\leqq (k + 1)e^{-2\pi^2/\epsilon^2(k^2+2k)} \\ &\quad \cdot [1 - (k + 2)(k + 1)^{-1} \\ &\quad \cdot e^{-4\pi^2 k/\epsilon^2}]^{-1} \\ &\leqq (k + 1)e^{-2\pi^2/\epsilon^2}(k^2 + 2k) \\ &\quad [1 - \tfrac{3}{2}e^{-4\pi^2/\epsilon^2}]^{-1}, \\ \sum_{k=1}^{\infty} \sum_{n=k+1}^{\infty} n \exp[-2\pi^2(2nk - k^2)/\epsilon^2] &\leqq (1 - \tfrac{3}{2}e^{-4\pi^2/\epsilon^2})^{-1} 2e^{-6\pi^2/\epsilon^2} \\ &\quad \cdot [1 - \tfrac{3}{2}e^{-10\pi^2/\epsilon^2}]^{-1}\end{aligned}$$

and

$$\sum_{n=2}^{\infty} n^2 e^{-2\pi^2 n^2/\epsilon^2} \leqq 4e^{-8\pi^2/\epsilon^2}/[1 - \tfrac{9}{4}e^{-10\pi^2/\epsilon^2}].$$

Applying these estimates and the inequality (31), we see that (41) is true provided

$$(42) \qquad 2[1 - P(\epsilon, 0)] > 4e^{-6\pi^2/\epsilon^2}[1 - \tfrac{9}{4}e^{-10\pi^2/\epsilon^2}]^{-1}[2 + 4P(\epsilon, 0)] + 4\epsilon(2\pi)^{-\frac{1}{2}}e^{-4\pi^2/\epsilon^2}[(1 - \tfrac{3}{2}e^{-4\pi^2/\epsilon^2})(1 - \tfrac{3}{2}e^{-10\pi^2/\epsilon^2})]^{-1}.$$

The left side of inequality (42) decreases as ϵ increases, while the right side increases. Hence if it is true for any value ϵ_1 of ϵ, it is true for $0 < \epsilon < \epsilon_1$, so that (41) is true for $0 < \epsilon < \min(\epsilon_1, \epsilon_0)$. On the other hand, (42) is easily verified for $\epsilon = \pi$, using the computed value $P(\pi, 0) = 0.884$. Hence, Lemma 6 is proved.

We can now state and prove Theorem 1.

THEOREM 1. *The ϵ-entropy $H_\epsilon(X)$ of the real line X under a one-dimensional Gaussian distribution with mean 0, variance σ^2, is $h(\epsilon/\sigma)$. The only ϵ-partition of the line with this entropy is the partition into consecutive intervals of length ϵ with one interval centered at zero.*

PROOF. We can assume $\sigma = 1$, since the general case follows by a change of scale. By Lemma 1, the only ϵ-partitions whose entropy can be the ϵ-entropy of the space are those which subdivide the line into intervals of length ϵ. We run through all these partitions by taking the partition into ϵ-intervals with one interval centered at $-\epsilon\alpha$, $0 \leqq \alpha < 1$. These partitions have entropies $h(\epsilon, \alpha)$, so that

$$H_\epsilon(X) = \inf_{0 \leqq \alpha < 1} h(\epsilon, \alpha).$$

By Lemmas 3 and 6, for each positive ϵ, this infimum is assumed only at $\alpha = 0$, which proves Theorem 1.

The final lemma of this section lists some properties of the function $h(\epsilon)$. These properties are interesting in themselves, and they are also needed at various points throughout the remainder of this paper.

LEMMA 7. *For* $0 < \epsilon < \infty$, $h'(\epsilon) < 0$ *and* $(h'(\epsilon)/\epsilon)' > 0$. *The function* $h'(\epsilon)/\epsilon$ *various monotonically from* $-\infty$ *to* 0 *for* ϵ *on* $(0, \infty)$. *Also, the following asymptotic formulas hold*:

as $\epsilon \to 0$,

$$h(\epsilon) \sim \log(1/\epsilon), \qquad h'(\epsilon) \sim -1/\epsilon; \tag{43}$$

as $\epsilon \to \infty$,

$$h(\epsilon) \sim \epsilon(2(2\pi)^{\frac{1}{2}})^{-1}e^{-\epsilon^2/8}, \qquad h'(\epsilon) \sim -\epsilon^2(8(2\pi)^{\frac{1}{2}})^{-1}e^{-\epsilon^2/8}. \tag{44}$$

PROOF. From the definition of $P(\epsilon, z)$,

$$\begin{aligned}(\partial/\partial\epsilon)P(\epsilon, z) &= (\partial/\partial\epsilon)\textstyle\int_{(z-\frac{1}{2})\epsilon}^{(z+\frac{1}{2})\epsilon} \phi(t)\,dt \\ &= (z + \tfrac{1}{2})\phi[\epsilon(z + \tfrac{1}{2})] - (z - \tfrac{1}{2})\phi[\epsilon(z - \tfrac{1}{2})].\end{aligned}$$

Hence

$$\begin{aligned}h'(\epsilon) &= (\partial/\partial\epsilon)\textstyle\sum_{k=-\infty}^{\infty} P(\epsilon, k)\log[P(\epsilon, k)]^{-1} \\ &= \textstyle\sum_{k=-\infty}^{\infty} \{(k + \tfrac{1}{2})\phi[\epsilon(k + \tfrac{1}{2})] - (k - \tfrac{1}{2})\phi[\epsilon(k - \tfrac{1}{2})]\}[\log[P(\epsilon, k)]^{-1} - 1].\end{aligned}$$

Rearranging terms, we obtain

$$h'(\epsilon) = \textstyle\sum_{k=-\infty}^{\infty} (k + \tfrac{1}{2})\phi[\epsilon(k + \tfrac{1}{2})]\log[P(\epsilon, k + 1)/P(\epsilon, k)].$$

The terms of this sum are unchanged when $k \to -k - 1$. Hence

$$h'(\epsilon) = -2\textstyle\sum_{k=0}^{\infty} (k + \tfrac{1}{2})\phi[\epsilon(k + \tfrac{1}{2})]\log[P(\epsilon, k)/P(\epsilon, k + 1)]. \tag{45}$$

This formula shows that $h'(\epsilon) < 0$, the first assertion of Lemma 7.

Now define θ_k, $k \geqq 0$, by

$$\log[P(\epsilon, k)/P(\epsilon, k + 1)] = \epsilon^2(k + \tfrac{1}{2} - \theta_k). \tag{46}$$

In the notation of Lemma 2,

$$F(\epsilon, k + \tfrac{1}{2}) = \epsilon^2(k + \tfrac{1}{2} - \theta_k). \tag{47}$$

By part (iii) of Lemma 2,

$$0 < \theta_k < \tfrac{1}{2}. \tag{48}$$

By the last inequality of (iv), Lemma 2, $F(\epsilon, k + \frac{1}{2})/\epsilon^2$ is a decreasing function of ϵ, with a negative derivative. Hence

$$d\theta_k/d\epsilon > 0. \tag{49}$$

Expressing the right side of (45) in terms of θ_k, we find

$$h'(\epsilon)/\epsilon = -2 \sum_{k=0}^{\infty} \phi[\epsilon(k + \tfrac{1}{2})]\epsilon(k + \tfrac{1}{2})(k + \tfrac{1}{2} - \theta_k), \tag{50}$$

and

$$[h'(\epsilon)/\epsilon]' = 2 \sum_{k=0}^{\infty} \phi[\epsilon(k + \tfrac{1}{2})][\epsilon^2(k + \tfrac{1}{2})^2 - 1](k + \tfrac{1}{2} - \theta_k)(k + \tfrac{1}{2}) + 2 \sum_{k=0}^{\infty} \phi[\epsilon(k + \tfrac{1}{2})]\epsilon(k + \tfrac{1}{2})\, d\theta_k/d\epsilon. \tag{51}$$

The terms in the second series are positive by (49). If $\epsilon \geqq 2$, then by (48) the terms in the first series are all non-negative, and $(h'/\epsilon)'$ is positive. If $\epsilon < 2$, and if k_1 is the integral part of $1/\epsilon - \frac{1}{2}$, the terms of the first series are negative up to k_1, positive for $k > k_1$. Thus we have

$$[h'(\epsilon)/\epsilon]' > 2 \sum_{k=0}^{k_1} \phi[\epsilon(k + \tfrac{1}{2})][\epsilon^2(k + \tfrac{1}{2})^2 - 1](k + \tfrac{1}{2})^2 + 2 \sum_{k=k_1+1}^{\infty} \phi[\epsilon(k + \tfrac{1}{2})][\epsilon^2(k + \tfrac{1}{2})^2 - 1]k(k + \tfrac{1}{2}),$$

and

$$[h'(\epsilon)/\epsilon]' > 2\epsilon^{-2}(2\pi)^{-\frac{1}{2}}\{\sum_{k=0}^{k_1} m[\epsilon(k + \tfrac{1}{2})] + \tfrac{2}{3} \sum_{k=k_1+1}^{\infty} m[\epsilon(k + \tfrac{1}{2})]\}, \tag{52}$$

where
$$m(x) = (x^4 - x^2)e^{-\frac{1}{2}x^2}.$$

This function $m(x)$ is negative on $(0, 1)$, assuming its minimum value $m(x_1) \cong -0.198$ at $x = x_1 = [(5 - (17)^{\frac{1}{2}})/2]^{\frac{1}{2}}$. The following argument shows that any closed interval of length 2 or more on the positive real axis must contain at least $k_1 + 1$ points of the sequence $\epsilon(k + \frac{1}{2})$, when $\epsilon \leqq 2$: If the interval contains exactly n points, then $(n + 1)\epsilon > 2$, $n > 2/\epsilon - 1$, and $n \geqq [2/\epsilon]$. For $\epsilon \leqq 2$, $2/\epsilon \geqq 1/\epsilon + \frac{1}{2}$, hence $n \geqq [1/\epsilon - \frac{1}{2}] + 1 = k_1 + 1$. In particular, there are at least $k_1 + 1$ terms in the second sum in (52) which have values greater than $(\frac{3}{2})|m(x_1)|$, since the interval (x_2, x_3) on which $m(x) > (\frac{3}{2})|m(x_1)|$ has $x_2 \cong 1.19$, $x_3 \cong 3.51$, and has length greater than 2.

Thus $(h'/\epsilon)'$ is positive for any $\epsilon > 0$. This proves the second assertion of Lemma 7.

We now know that $h'(\epsilon)/\epsilon$ is an increasing function of ϵ. Its range will follow from the asymptotic formulas to be derived.

First we consider the behavior of $h(\epsilon)$ as $\epsilon \to 0$. The asymptotic form of $h(\epsilon)$ is a general property of the epsilon entropy of continuous one-dimensional distributions [4]. However, the form of $h(\epsilon)$ when $\epsilon \to 0$, as well as when $\epsilon \to \infty$, follows directly from the forms given for the derivative. To show the second part of (43), note that when (50) is multiplied by ϵ^2, the sum on the right tends to an integral as $\epsilon \to 0$:

$$\epsilon h'(\epsilon) \to -2 \int_0^{\infty} (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}x^2} x^2\, dx = -1.$$

This proves (43).

Again, as $\epsilon \to \infty$, the contribution of the terms after the first in the series (50) is $O(\epsilon e^{-9\epsilon^2/8})$. Thus

$$h'(\epsilon) = -\epsilon^2\phi(\epsilon/2)(\tfrac{1}{2} - \theta_0) + O(\epsilon^2 e^{-9/8\epsilon^2}).$$

But as $\epsilon \to \infty$, $P(\epsilon, 0) \to 1$, and

$$P(\epsilon, 1) = (2\pi)^{-\frac{1}{2}} \int_{\frac{1}{2}\epsilon}^{\frac{3}{2}\epsilon} e^{-\frac{1}{2}x^2}\, dx \sim 2e^{-\frac{1}{8}\epsilon^2}/\epsilon(2\pi)^{\frac{1}{2}}.$$

From (46),

$$\epsilon^2(\tfrac{1}{2} - \theta_0) \sim \tfrac{1}{8}\epsilon^2.$$

Thus, as $\epsilon \to \infty$,

$$-\epsilon^2\phi(\epsilon/2)(\tfrac{1}{2} - \theta_0) \sim -\tfrac{1}{8}\epsilon^2\phi(\epsilon/2)$$
$$\sim -\tfrac{1}{8}(2\pi)^{-\frac{1}{2}}\epsilon^2 e^{-\epsilon^2/8}.$$

Since the contribution of the remaining terms in the series is the smaller value $O(\epsilon e^{-9\epsilon^2/8})$, the above expression is asymptotically the value of $h'(\epsilon)$, and the second half of (44) holds. Lemma 7 is proved.

Now that we have gotten "preliminaries" about the one-dimensional Gaussian distribution out of the way, we can begin to study the case of arbitrary mean-continuous Gaussian processes on the unit interval.

3. The product epsilon entropy function $J_\epsilon(X)$. In this section we define the product ϵ-entropy $J_\epsilon(X)$ of a mean-continuous Gaussian process X on the unit interval. The main results are contained in Theorem 2. We find a necessary and sufficient condition for $J_\epsilon(X)$ to be finite. In the case when $J_\epsilon(X)$ is finite we construct a product ϵ-partition with entropy equal to $J_\epsilon(X)$.

In order to define the product ϵ-entropy function $J_\epsilon(x)$ we first consider the class π_ϵ' of all *product ϵ-partitions* of $L_2[0,1]$. A product ϵ-partition of $L_2[0, 1]$ is the Cartesian product of ϵ_k-partitions of the kth coordinate axis in the Karhünen expansion of the process, where $\sum \epsilon_k^2 \leqq \epsilon^2$. Thus product ϵ-partitions consist of hyper-cubes of diameter at most ϵ. Next define π_ϵ to be that subclass of π_ϵ' consisting of partitions in which a countable collection of the sets have a union of probability 1. I.e., a product partition in π_ϵ contains a denumerable partition of a subset of X of probability 1. By the *entropy* of the product partition we mean the entropy of this denumerable partition.

The product epsilon entropy is defined as

$$J_\epsilon(X) = \infty, \qquad \text{if} \quad \pi_\epsilon \quad \text{is empty},$$
$$J_\epsilon(X) = \inf_{U \varepsilon \pi_\epsilon} H(U), \qquad \text{if} \quad \pi_\epsilon \quad \text{is not empty}.$$

The entropy $H(U)$ is defined as in (7) over the sets of U of positive probability. It turns out that π_ϵ is empty if the series (64) diverges, and otherwise $J_\epsilon(X)$ is finite.

Our first lemma shows how to compute the entropy of a product partition in terms of the entropies of its one-dimensional partitions.

LEMMA 8. *Let the probability space X be the product of a sequence of probability spaces $X_1, X_2, \cdots$, with product measure. If U_k is a partition of X_k, $k = 1, 2, \cdots$, and U the product partition of X, then*

$$H(U) = \sum_{k=1}^{\infty} H(U_k).$$

This is to be interpreted to mean that if the union of countably many sets of U does not have probability 1, then $H(U)$ is infinite.

PROOF. It follows from [2], Lemma 3, that

$$H(U) \leqq \sum_{k=1}^{\infty} H(U_k), \tag{53}$$

even with the special interpretation on $H(U)$. Hence, we need to show that inequality cannot occur. The above-cited reference shows that we can assume that the union of countably many sets of U has probability 1.

For any positive integer n consider the finite product

$$X^{(n)} = X_1 \times X_2 \times \cdots \times X_n .$$

The partition $U^{(n)} = U_1 \times U_2 \times \cdots \times U_n$ of $X^{(n)}$ has the same entropy as the partition $U^{(n)} \times X_{n+1} \times \cdots$ of X, of which U is a refinement. Thus $H(U^{(n)}) \leqq H(U)$. If the sets of U_k have probabilities p_{kj}, $j = 1, 2, \cdots$, then the sets of $U^{(n)}$ have probabilities $p_{1j_1} p_{2j_2} \cdots p_{nj_n}$, $j_1, \cdots, j_n = 1, 2, \cdots$, and

$$H(U^{(n)}) = \sum_{j_1, \cdots, j_n=1}^{\infty} p_{1j_1} \cdots p_{nj_n} \log [p_{1j_1} \cdots p_{nj_n}]^{-1}.$$

Breaking up the logarithm into a sum,

$$\begin{aligned} H(U^{(n)}) &= \sum_{m=1}^{n} \sum_{j_1, \cdots, j_n=1}^{\infty} p_{1j_1} \cdots p_{nj_n} \log (p_{mj_m})^{-1} \\ &= \sum_{m=1}^{\infty} \sum_{j_m=1}^{\infty} p_{mj_m} \log (p_{mj_m})^{-1} \\ &= \sum_{m=1}^{n} H(U_m). \end{aligned}$$

Hence

$$H(U) \geqq \sum_{m=1}^{n} H(U_m),$$

and, letting $n \to \infty$,

$$H(U) \geqq \sum_{m=1}^{\infty} H(U_m).$$

This inequality together with (53) proves Lemma 8.

The next two lemmas taken together show that, for a mean-continuous Gaussian process on the unit interval, either π_ϵ is empty for all $\epsilon > 0$, or else π_ϵ contains a partition of finite entropy for all $\epsilon > 0$.

LEMMA 9. *Let $X(t)$ be a mean -continuous Gaussian process on the unit interval. Let U be a product partition of $L_2[0, 1]$ obtained as the product of partitions U_k of the coordinate axes by intervals of lengths ϵ_k. Then the following three conditions are equivalent:*

(a) *The union of countably many sets of U has probability* 1;

(b) *U contains a set of positive probability*;

(c) *With probability* 1, *all but a finite number of components of an element of* $L_2\,[0,1]$ *lie in the unique interval containing zero in the partition of that coordinate.*

If the partitions U_k *are centered, these conditions are also equivalent to*

(d) $\sum_{k=1}^{\infty} \lambda_k^{\frac{1}{2}} \epsilon_k^{-1} \phi(\epsilon_k/2\lambda_k^{\frac{1}{2}}) < \infty$, *where* ϕ *is the unit normal density function and* $\{\lambda_k\}$ *are the eigenvalues of the process.*

PROOF. Clearly (c) implies (a), since the collection of sets of U with the property (c) is denumerable, and (a) implies (b). The equivalence of these conditions will follow if we show that (b) implies (c).

Assume (b) is true. Let the set V of positive probability be the product of the intervals $\{V_k\}$ on the coordinate axes, with $p_k = \mu(V_k)$. Then

$$\mu(V) = \prod_{k=1}^{\infty} p_k > 0. \tag{54}$$

First we note that this product can be positive only if $p_k \to 1$ as $k \to \infty$. In particular, for k sufficiently large ($k \geqq K$), $p_k > \frac{1}{2}$, which implies that V_k has the origin of the kth coordinate axis in its interior. Another consequence of (54) is that

$$\lim_{m\to\infty} \prod_{k=m}^{\infty} p_k = 1. \tag{55}$$

For $m \geqq K$, consider those sets of U which are formed by using V_k for $k \geqq m$. The union of these sets has probability $\prod_{k=m}^{\infty} p_k$, and is a subset of the set characterized in (c). Hence, by (55), (c) is true.

Now let the U_k be centered. The set in U which is the product of the center intervals of the U_k has at least as large a probability as many other set, so that (b) is equivalent to (54), where p_k is the probability of the center interval of U_k. Then

$$1 - p_k = 2\lambda_k^{-\frac{1}{2}} \int_{\epsilon_k/2}^{\infty} \phi(t/\lambda_k^{\frac{1}{2}})\,dt \sim 2\lambda_k^{\frac{1}{2}} \epsilon_k^{-1} \phi(\epsilon_k/2\lambda_k^{\frac{1}{2}}), \quad \text{as } k \to \infty. \tag{56}$$

Condition (b) is equivalent to

$$\sum_{k=1}^{\infty} (1 - p_k) < \infty, \tag{57}$$

which is equivalent to

$$\sum_{k=1}^{\infty} \lambda_k^{\frac{1}{2}} \epsilon_k^{-1} \phi(\epsilon_k/2\lambda_k^{\frac{1}{2}}) < \infty, \tag{58}$$

since the convergence of either of these series implies that $\epsilon_k/\lambda_k^{\frac{1}{2}} \to \infty$. Therefore (56) is true. Thus (b) and (d) are equivalent. This completes the proof of Lemma 9.

LEMMA 10. *For* $k = 1, 2 \cdots$, *let* U_k *be a given* ϵ_k'*-partition of the* kth *coordinate axis. Let* $\sum \epsilon_k'^2$ *converge and let a countable subpartition of the product partition* $U = \prod_k U_k$ *cover a set of probability* 1 *in* $L_2\,[0,1]$. *Then for every* $\epsilon > 0$ *there exist* ϵ_k*-partitions* V_k *of the* kth *coordinate axis such that*

$$\epsilon^2 = \sum \epsilon_k^2 \qquad \textit{and} \qquad \sum H(V_k) < \infty.$$

PROOF. We can assume that U_k is a centered partition, and that every interval in U_k is of length ϵ_k'. By Lemma 9, if a countable subpartition of U covers a set of probability 1, then

$$\sum_k \lambda_k^{\frac{1}{2}}(\epsilon_k')^{-1}e^{-\epsilon_k'^2/8\lambda_k} < \infty. \tag{59}$$

On the other hand, by (44), we will have $\sum H(V_k) < \infty$ if both $\epsilon_k/\lambda_k^{\frac{1}{2}} \to \infty$ and

$$\sum \epsilon_k \lambda_k^{-\frac{1}{2}} e^{-\epsilon_k^2/8\lambda_k} < \infty. \tag{60}$$

Now if (59) holds with ϵ_k', we proceed as follows. Let K be so large that for $k > K$, $\epsilon_k'/\lambda_k^{\frac{1}{2}} \geqq 10$ and

$$\sum_{k>K} \epsilon_k'^2 \leqq \epsilon^2/4. \tag{61}$$

The function $t^2e^{-t^2/8}$ is less than $1/2^{\frac{1}{2}}$ for $t \geqq 10$; hence, for such t,

$$2^{\frac{1}{2}}te^{-t^2/4} < t^{-1}e^{-t^2/8},$$

and for $k > K$,

$$2^{\frac{1}{2}}\epsilon_k'\lambda_k^{-\frac{1}{2}}e^{-\epsilon_k'^2/4\lambda_k} < \lambda_k^{\frac{1}{2}}(\epsilon_k')^{-1}e^{-\epsilon_k'^2/8\lambda_k}. \tag{62}$$

Thus if we put $\epsilon_k = 2^{\frac{1}{2}}\epsilon_k'$ for $k > K$ and assign to $\epsilon_1, \cdots, \epsilon_K$ any positive values such that $\sum \epsilon_k^2 = \epsilon^2$, (60) is satisfied, for the terms in that series are bounded for $k > K$ by the terms in the series (59). This proves Lemma 10.

The next lemma is the last before Theorem 2.

LEMMA 11. *For a mean-continuous Gaussian process on* [0, 1] *with eigenvalues* $\lambda_n = \sigma_n^2$, $n = 1, 2, \cdots$, *the product* ϵ*-entropy is given by*

$$J_\epsilon(X) = \inf_{\sum\epsilon_k^2=\epsilon^2} \sum_{k=1}^{\infty} h(\epsilon_k/\sigma_k), \tag{63}$$

where $h(x)$ *is the function defined on page* 874.

PROOF. With each product ϵ-partition of X we can associate a sequence $\{\epsilon_k\}$ such that the partition of the kth component space X_k is an ϵ_k-partition, and $\sum \epsilon_k^2 = \epsilon^2$. For given $\{\epsilon_k\}$, the minimum possible entropy of the partition of X_k is $h(\epsilon_k/\sigma_k)$, by Theorem 1. Hence (63) follows from Lemma 8. Lemma 11 is proved.

Equation (63) reduces the problem of finding an optimal product ϵ-partition to the problem of selecting an optimal set $\{\epsilon_k\}$ of "quantizations" for the coordinate axes. The next theorem solves this problem and gives at the same time a necessary and sufficient condition for $J_\epsilon(X)$ to be finite.

THEOREM 2. *The product* ϵ*-entropy* $J_\epsilon(X)$ *of a mean-continuous Gaussian process on* [0, 1] *with eigenvalues* $\{\lambda_k\}$ *is finite if and only if*

$$\sum \lambda_k \log \lambda_k^{-1} < \infty. \tag{64}$$

If this condition is satisfied, the equations

$$h'(\delta_k) = -A\lambda_k\delta_k, \qquad k = 1, 2, \cdots, \tag{65}$$

have a unique solution $\{\delta_k\}$, *with* A *such that*

$$\sum \lambda_k\delta_k^2 = \epsilon^2. \tag{66}$$

Then

$$J_\epsilon(X) = \textstyle\sum_{k=1}^{\infty} h(\delta_k). \tag{67}$$

On the other hand, if (64) *is violated,* (65) *and* (66) *have no solution. The condition* (64) *is also the condition that there be a countable subpartition of some product epsilon partition covering a set of probability* 1.

PROOF. Set $\sigma_k = \lambda_k^{\frac{1}{2}}$. We want to minimize $J(\epsilon_1, \epsilon_2, \cdots) = \sum h(\epsilon_k/\sigma_k)$ subject to condition $\sum \epsilon_k^2 = \epsilon^2$. Equations (65) are the conditions for a minimum, by the method of Lagrange multipliers, if $\delta_k = \epsilon_k/\sigma_k$. To avoid justifying the use of this method in an infinite-dimensional space, we will consider finite dimensional subspaces of X.

First we show that (65) and (66) have a (unique) solution for any $\epsilon > 0$ if and only if (64) is satisfied. According to Lemma 7, for any $A > 0$ there is a unique solution $\{\delta_k\}$ of (65); each δ_k is a monotonic decreasing function of A, and

$$\lim_{A\to 0} \delta_k = \infty, \qquad \lim_{A\to\infty} \delta_k = 0.$$

For a given value of A, $A\lambda_k \to 0$ as $k \to \infty$. Hence for k sufficiently large, δ_k is so large that we can conclude from Lemma 7 that

$$h'(\delta_k) = -C_k\delta_k^2 e^{-\frac{1}{8}\delta_k^2},$$

where

$$\tfrac{1}{16}(2\pi)^{-\frac{1}{2}} < C_k < 1.$$

Then we have

$$C_k\delta_k e^{-\frac{1}{8}\delta_k^2} = A\lambda_k,$$

which implies

$$\delta_k^2 \sim 8 \log (1/\lambda_k). \tag{68}$$

We see that the series (66) is finite if and only if (64) is satisfied. If (64) holds, then the monotone dependence of δ_k on A shows that the series in (66) is a strictly decreasing function of A, taking all positive values as A ranges over all positive values. Therefore, (65) and (66) have a unique solution.

Notice also that the existence of a solution of (65) and (66) implies that $J_\epsilon(x)$ is finite, for if we put $\epsilon_k = \sigma_k\delta_k$, then

$$\sum \epsilon_k^2 = \epsilon^2 \qquad \text{and} \qquad J_\epsilon(X) \leqq \sum h(\epsilon_k/\sigma_k) = \sum h(\delta_k).$$

This series converges, for by Lemma 7,

$$h(\delta_k) \sim -\delta_k h'(\delta_k) = A\delta_k^2\lambda_k.$$

Now let $X^{(n)}$ be the product of the first n coordinate spaces. By Lemma 11,

$$J_\epsilon(X^{(n)}) = \inf_{\sum_1^n \epsilon_k^2 = \epsilon^2} \textstyle\sum_{k=1}^{n} h(\epsilon_k/\sigma_k).$$

This sum is a continuous function over the positive 2^n-tant of the n-sphere

$\sum \epsilon_k^2 = \epsilon^2$, approaching infinity at the boundaries. Hence the infimum is assumed at some interior point, and we have there

$$h'(\epsilon_k/\sigma_k)/\sigma_k = -A^{(n)}\epsilon_k, \qquad k = 1, \cdots, n,$$

where $A^{(n)}$ is a positive constant. Let $\epsilon_k = \delta_k^{(n)}\sigma_n$ be a solution of this system of equations, which lies on the m-sphere. Then

$$h'(\delta_k^{(n)})/\delta_k^{(n)} = -A^{(n)}\lambda_k, \qquad k = 1, \cdots, n, \tag{69}$$

and

$$\sum_{k=1}^{n} \lambda_k (\delta_k^{(n)})^2 = \epsilon^2. \tag{70}$$

For any value of $A^{(n)}$, the solutions of (69) are unique, by Lemma 7. Furthermore, as $A^{(n)}$ varies from 0 to ∞, each $\delta_k^{(n)}$ varies monotonically from ∞ to 0. Thus there is a unique value of $A^{(n)}$ at which (70) is satisfied. We have

$$J_\epsilon(X) \geqq J_\epsilon(X^{(n)}) = \sum_{k=1}^{n} h(\delta_k^{(n)}). \tag{71}$$

This bounding can be done for any n. In particular, for the numbers $A^{(n+1)}$ and $\{\delta_k^{(n+1)}\}$, $\delta_1^{(n+1)}, \cdots, \delta_n^{(n+1)}$ are solutions of (69) with $A^{(n)}$ replaced by $A^{(n+1)}$, and

$$\sum_{k=1}^{n} \lambda_k (\delta_k^{(n+1)})^2 \leqq \epsilon^2.$$

It follows that $A^{(n+1)} \geqq A^{(n)}$. Define

$$\bar{A} = \lim_{n\to\infty} A^{(n)}.$$

$\bar{A}$ is either a positive real number or ∞.

First suppose $\bar{A} = \infty$. Then as $n \to \infty$, $A^{(n)}\lambda_1 \to \infty$ and $\delta_1^{(n)} \to 0$. From (71),

$$J_\epsilon(X) \geqq h(\delta_1^{(n)}) \to \infty,$$

so $J_\epsilon(X) = \infty$. It follows from above that in this case (64) is violated.

Now let $\bar{A}$ be finite, and let $\{\bar{\delta}_k\}$ be the solution of (65) when $A = \bar{A}$. Since $A^{(n)} \leqq \bar{A}$,

$$\sum_{k=1}^{n} \lambda_k \bar{\delta}_k^2 \leqq \sum_{k=1}^{n} \lambda_k \delta_k^{(n)2} = \epsilon^2,$$

hence

$$\sum_{k=1}^{\infty} \lambda_k \bar{\delta}_k^2 \leqq \epsilon^2.$$

This shows that there is a value A^* of A for which the solution of (65) satisfies (66), and $A^* \leqq \bar{A}$. Denoting this solution by $\{\delta_k^*\}$, we have

$$\sum_{k=1}^{n} \lambda_k \delta_k^{*2} \leqq \epsilon^2;$$

Hence $A^* \geqq A^{(n)}$ for all n. It follows that $\bar{A} = A^*$.

For each k, we have $\delta_k^{(n)} \to \bar{\delta}_k$ as $n \to \infty$. From (71), if $m \leqq n$,

$$J_\epsilon(X) \geqq \sum_{k=1}^{m} h(\delta_k^{(n)}).$$

Let $n \to \infty$; then $m \to \infty$, and we obtain

$$J_\epsilon(X) \geqq \sum_{k=1}^\infty h(\bar{\delta}_k).$$

On the other hand, we have seen above that this series is the entropy of an ϵ-product partition of X. Therefore equality holds, and (67) is true. The last assertion of the theorem follows from Lemmas 9 and 10. This completes the proof of Theorem 2.

COROLLARY. *$J_\epsilon(X)$ is a continuous function of ϵ.*

PROOF. This is a consequence of the formulas of Theorem 2. Namely, the asymptotic formula (68) is uniform over any interval $0 < A_1 \leqq A \leqq A_2 < \infty$. Thus the series in (66) and (67) are uniformly convergent. It follows that these series are continuous functions of A. Since ϵ, given by (66), is a strictly decreasing function of A, A and $J_\epsilon(X)$ are continuous functions of ϵ. This proves the corollary.

We remark that when the λ_k are written in non-increasing order, condition (64) is easily shown to be equivalent to the condition

$$\sum \lambda_k \log k < \infty.$$

Also note that (64) is the entropy of the distribution $\{\lambda_k\}$, provided the λ_k are normalized so that $\sum \lambda_k = 1$. The occurrence of the entropy of the eigenvalues in this way appears to be fortuitous, although quite intriguing.

4. Some special processes. In this section, we shall consider a class of Gaussian processes whose product ϵ-entropies can be estimated for small ϵ by Theorem 2.

We begin with some general remarks on product ϵ-entropy. Let X be a finite-dimensional mean-continuous Gaussian process on [0, 1]. That is, X has only a finite number of non-zero eigenvalues, $\lambda_1, \cdots, \lambda_n$ say. It is an easy consequence of Theorem 2 and Lemma 7 that

$$J_\epsilon(X) \sim n \log \epsilon^{-1}$$

as $\epsilon \to 0$. For this reason the interesting processes to now consider, from the point of view of product ϵ-entropy, are the infinite dimensional ones.

The first thing we observe about an infinite dimensional process X is that, as $\epsilon \to 0$, its product ϵ-entropy must increase faster than any positive multiple of $\log \epsilon^{-1}$. To verify this let $X^{(n)}$ be the finite dimensional process obtained from X by setting $\lambda_k = 0$ for $k > n$. Then as $\epsilon \to 0$

$$J_\epsilon(X) \geqq J_\epsilon(X^{(n)}) \sim n \log \epsilon^{-1}.$$

Since n was arbitrary this proves our assertion.

In the final section of this paper we shall develop some techniques which are more generally applicable than Theorem 2. For the present, however, we shall content ourselves with the consideration of mean-continuous Gaussian processes on [0, 1] whose eigenvalues satisfy a relation of the form

$$\lambda_k \sim Bk^{-p} \qquad \text{as} \quad k \to \infty,$$

where $B > 0$ and $p > 1$ are constants. Special cases of these processes arise as solutions of the stochastic differential equation

$$d^nX/dt^n + a_{n-1}\cdot d^{n-1}X/dt^{n-1} + \cdots + a_0X = b_m\cdot d^mN/dt^m + \cdots + b_0N$$

where $N(t)$ is white Gaussian noise of spectral density $\frac{1}{2}$ and the a's and b's are constants with $b_m \neq 0$ and $n > m$. For these processes $R(s, t) = E(X(s)X(t))$ can be found as well as the λ_k. However, for our purposes it is enough to know that $\lambda_k \sim Bk^{-p}$, where $B > 0$ and $p = 2(n - m)$. This is true for stationary processes by [6], and (apparently) is also true for non-stationary processes. The most important special case is the Weiner process, for which $dX/dt = N$, $R(s, t) = \min(s, t)$ and

$$\lambda_k = 1/\pi^2(k - \tfrac{1}{2})^2, \qquad k = 1, 2, \cdots.$$

The main result of this section is the following theorem which gives an asymptotic formula for $J_\epsilon(X)$ as $\epsilon \to 0$.

THEOREM 3. *Let X be a mean-continuous Gaussian process on the unit interval with eigenvalues $\{\lambda_n\}$ such that*

$$\lambda_n \sim Bn^{-p},$$

$B > 0$, $p > 1$. *Then, as* $\epsilon \to 0$,

$$J_\epsilon(X) \sim \epsilon^{-2/(p-1)}(2B/(p-1))^{1/(p-1)}\{\textstyle\int_0^\infty [-h'(x)/x]^{1-1/p}x\,dx\}^{p/(p-1)}. \tag{72}$$

PROOF. According to Lemma 7, the equation $-h'(\epsilon)/\epsilon = x$ has an inverse $\epsilon = r(x)$, where $r(x)$ is a monotonic function with

$$\begin{aligned} r(x) &\sim x^{-\frac{1}{2}} && \text{as } x \to \infty, \\ r(x) &\sim [8 \log (1/x)]^{\frac{1}{2}} && \text{as } x \to 0. \end{aligned} \tag{73}$$

In terms of this function, the solution of (65) is

$$\delta_k = r(A\lambda_k),$$

and (66) becomes

$$\epsilon^2 = A^{-1} \textstyle\sum_{k=1}^\infty A\lambda_k r(A\lambda_k)^2. \tag{74}$$

We want to get an symptotic formula for this function of A as $A \to \infty$. Thus we are led to consider sums of the form

$$\textstyle\sum_{k=1}^\infty f(A\lambda_k). \tag{75}$$

Suppose that $f(x)$ is a continuous function on $(0, \infty)$ with

$$\begin{aligned} f(x) &= O(x \log x^{-1}), && x \to 0, \\ &= O(\log x), && x \to \infty. \end{aligned} \tag{76}$$

We break up the sum (75) into three parts. Let η be a small positive number, fixed for now. Since

$$\sum_{A\lambda_k \leqq \eta} A\lambda_k \log (A\lambda_k)^{-1} \sim \sum_{n^p > AB/\eta} ABn^{-p} \log (n^p/AB) \sim (AB)^{1/p}(p-1)^{-1}\eta^{1-1/p} \log \eta^{-1},$$

as $A \to \infty$, we have

$$\sum_{A\lambda_k \leqq \eta} f(A\lambda_k) = O(A^{1/p}\eta^{1-1/p} \log \eta^{-1}). \tag{77}$$

By hypothesis, there is a constant B_1 such that $\lambda_n < B_1 n^{-p}$ for all n. Hence, if M is any positive number,

$$\sum_{A\lambda_k \geqq M} (A\lambda_k)^{1/2p} \leqq \sum_{Mn^p < AB_1} (AB_1)^{1/2p} n^{-\frac{1}{2}} \sim 2(AB_1)^{1/p} M^{-1/2p}$$

as $A \to \infty$. From (76), $f(x) = O(x^{1/2p})$ for large x. Hence

$$\sum_{A\lambda_k \geqq M} f(A\lambda_k) = O(A^{1/p} M^{-1/2p}). \tag{78}$$

Now for $0 < a < b < \infty$, consider $\nu(a, b)$, the number of values of k for which $A\lambda_k$ lies on (a, b). The condition $a < A\lambda_k < b$ is equivalent to $a' < ABk^{-p} < b'$, where $a' \to a$ and $b' \to b$ as $A \to \infty$. The inequality which k must satisfy is

$$(AB/b')^{1/p} < k < (AB/a')^{1/p}.$$

Hence

$$\nu(a, b) \sim (AB)^{1/p}(a^{-1/p} - b^{-1/p}).$$

It follows that

$$\sum_{\eta < A\lambda_k < M} f(A\lambda_k) \sim (AB)^{1/p} \int_\eta^M f(x)\, d(-x^{-1/p}).$$

Comparing this result with (77) and (78), we have

$$\sum_{k=1}^{\infty} f(A\lambda_k) \sim p^{-1}(AB)^{1/p} \int_0^\infty f(x) x^{-1-1/p}\, dx. \tag{79}$$

If we take $f(x) = xr(x)^2$, (76) is satisfied. Hence from (74)

$$\epsilon^2 \sim (AB)^{1/p}(Ap)^{-1} \int_0^\infty r(x)^2 x^{-1/p}\, dx.$$

Put $x = -h'(t)/t$, and integrate by parts. The integrated term vanishes at the limits, and we get

$$\epsilon^2 \sim 2(AB)^{1/p}[A(p-1)]^{-1} \int_0^\infty [-h'(t)/t]^{1-1/p} t\, dt. \tag{80}$$

From (67),

$$J_\epsilon(X) = \sum_{k=1}^{\infty} h(r(A\lambda_k)).$$

We apply (79), with $f(x) = h(r(x))$; this function f satisfies (76), so

$$J_\epsilon(X) \sim p^{-1}(AB)^{1/p} \int_0^\infty h(r(x)) x^{-1-1/p}\, dx.$$

After the substitution $x = -h'(t)/t$ and integration by parts, this becomes

$$J_\epsilon(X) \sim (AB)^{1/p} \int_0^\infty [-h'(t)/t]^{1-1/p} t\, dt.$$

Eliminating A by (80), we obtain (72). This proves Theorem 3.

COROLLARY. *For the Weiner process on* [0, 1], *we have*

$$J_\epsilon(x) \sim C/\epsilon^2$$

as $\epsilon \to 0$, *with*

$$C = 2\pi^{-2}\{\int_0^\infty [-xh'(x)]^{\frac{1}{2}}\, dx\}^2 = 6.711 \cdots .$$

PROOF. We apply Theorem 3 with $B = 1/\pi^2$, $p = 2$, and evaluate the integral numerically to prove this corollary.

In [3], the ϵ-entropy $H_\epsilon(X)$ of the Weiner process was considered, where $H_\epsilon(X)$, as in Section 2 of this paper, is the infimum of the entropies of all countable partitions of sets of probability 1 in L_2 [0, 1] by measurable sets of diameters at most ϵ. Thus, $H_\epsilon(X) \leqq J_\epsilon(X)$. However, it was shown in [3] that for the Weiner process

$$\frac{17}{32\epsilon^2} \lesssim H_\epsilon(X) \lesssim \epsilon^{-2}$$

(the notation $U \lesssim V$ means lim sup $(U/V) \leqq 1$). Thus, for the Weiner process,

$$\liminf_{\epsilon\to 0} J_\epsilon(X)/H_\epsilon(X) \geqq 6.711 \cdots .$$

This means that for small ϵ the optimal product ϵ-partition requires at least 6.7+ times as many bits, on the average, to transmit the outcome of the process as does the optimal ϵ-partition.

5. The order of magnitude of $J_\epsilon(X)$. In this final section, a useful lower bound $L_\epsilon(X)$ for $J_\epsilon(X)$ is considered. Conditions on the eigenvalues λ_k are given, which guarantee that $J_\epsilon(X) = O(L_\epsilon(X))$, or even $J_\epsilon(X) \sim L_\epsilon(X)$. Since $L_\epsilon(X)$ is a lower bound for the epsilon entropy $H_\epsilon(X)$, these results imply that $H_\epsilon(X)$ is of the same order as, or even asymptotically equal to, $J_\epsilon(X)$, so that not much is lost by the restriction to product partitions in these special cases. Finally, these results are applied to a stationary band limited Gaussian process on the unit interval to obtain a simple asymptotic expression for $J_\epsilon(X)$ in that case.

The lower bound $L_\epsilon(X)$ derived in [3] for the ϵ-entropy $H_\epsilon(X)$ of a Gaussian process X is as follows: Assume $\epsilon^2 < \sum \lambda_k$. Define the number $b = b(\epsilon)$ by

(81) $$\epsilon^2 = \sum \lambda_k/(1 + b\lambda_k).$$

Then

(82) $$L_\epsilon(X) = \tfrac{1}{2} \sum \log\ (1 + b\lambda_k).$$

Since $L_\epsilon(X) \leqq H_\epsilon(x)$ and $H_\epsilon(X) \leqq J_\epsilon(X)$, $L_\epsilon(X)$ also provides a lower bound for $J_\epsilon(X)$.

The next lemma gives a lower bound for $L_\epsilon(X)$, which is actually the bound we shall be using instead of $L_\epsilon(X)$ itself.

LEMMA 12. *Let X be a mean-continuous Gaussian process on* [0, 1] *with eigenvalues* $\lambda_1 \geqq \lambda_2 \geqq \cdots$. *Define* $\lambda(x)$, $x \geqq 1$, *as the function such that* $\lambda(n) = \lambda_n$,

$n = 1, 2, \cdots$, *and such that* $x\lambda(x)$ *is linear on each interval* $(n, n+1)$. *For* $\epsilon^2 < \lambda_1$, *define the function* $y = y(\epsilon)$ *to be the smallest root on* $(1, \infty)$ *of the equation* $y\lambda(y) = \epsilon^2$. *Then*

$$L_\epsilon(X) \geqq \int_\epsilon^{\lambda_1^{\frac{1}{2}}} [y(t) - 1] t^{-1}\, dt + O(1) \qquad \text{as} \quad \epsilon \to 0. \tag{83}$$

Proof. Let $N = N(\epsilon)$ be the integral part of $y(\epsilon)$. Consider

$$S(\epsilon) = \tfrac{1}{2} \textstyle\sum_{k=1}^{N} \log [\lambda(y)(1 + b\lambda_k)/\lambda_k],$$

where $b = b(\epsilon)$ is determined by Equation (81). From that equation,

$$\textstyle\sum_{k=1}^{N} \lambda_k/(1 + b\lambda_k) \leqq \epsilon^2. \tag{84}$$

The problem of minimizing $S(\epsilon)$ for any set of λ_k satisfying (84) is easily solved. We find that

$$S(\epsilon) \geqq \tfrac{1}{2}N \log [N\lambda(y)/\epsilon^2] = \tfrac{1}{2}N \log (N/y) \geqq \tfrac{1}{2}N \log [N/(N+1)] \geqq -\tfrac{1}{2}.$$

Hence

$$L_\epsilon(X) \geqq \tfrac{1}{2} \textstyle\sum_{k=1}^{N} \log (1 + b\lambda_k) \geqq \tfrac{1}{2} \sum_{k=1}^{N} \log (\lambda_k/\lambda(y)) - \tfrac{1}{2}. \tag{85}$$

For $k \leqq x \leqq k+1$, since $x\lambda(x)$ is linear and $\lambda_k \geqq \lambda_{k+1}$, it follows that $\lambda(x) \leqq (1 + 1/k)\lambda_k$. Hence

$$\int_1^y \log [\lambda(x)/\lambda(y)]\, dx \leqq \textstyle\sum_{k=1}^{N} \log [\lambda_k/\lambda(y)] + O(\log y),$$

and, combining this inequality with (85),

$$L_\epsilon(X) \geqq \tfrac{1}{2} \int_1^y \log [\lambda(x)/\lambda(y)]\, dx + O(\log y). \tag{86}$$

If we define

$$\epsilon(x) = [x\lambda(x)]^{\frac{1}{2}},$$

then, by integration by parts, we find

$$\begin{aligned} \int_1^y \log [\lambda(x)/\lambda(y)]\, dx &= -\log [\lambda_1/\lambda(y)] + \int_1^y x\, d \log [1/\lambda(x)] \\ &= -\log [\lambda_1/\lambda(y)] + y - 1 + 2\int_1^y x\, d \log [\epsilon(x)]^{-1}. \end{aligned} \tag{87}$$

In terms of the variable $t = \epsilon(x)$,

$$\int_1^y x\, d \log [\epsilon(x)]^{-1} = \int_\epsilon^{\lambda_1^{\frac{1}{2}}} y(t) t^{-1}\, dt + R, \tag{88}$$

where R accounts for the ranges of values of x which have been lost.

We shall show that $R \geqq 0$. The function $t = \epsilon(x)$, $1 < x < y$, can have the same value for several values of x. The interval which is the range of $\epsilon(x)$ for $1 < x < y$ can be broken up into a finite number of intervals such as $a < \epsilon(x) < b$. each of which is the 1-1 image of each of m intervals, $I_1, I_2, \cdots, I_m$, (in order of increasing x), contained in $(1, y)$. Let the inverse function from (a, b) to I_j be $y_j(t)$. For $m \geqq 2$, the last $2[m/2]$ of these intervals contribute to R. The functions $y_j(t)$ are monotonic with alternate senses. If $a \geqq \epsilon$, $y_m(t)$ is a decreasing function, and

$$\int_{I_1 \cup I_2 \cup \cdots \cup I_m} x \, d \log [\epsilon(x)]^{-1} = \int_a^b \sum_{j=1}^m (-1)^{m-j} y_j(t) t^{-1} \, dt.$$

Since $y_m(t) \geqq y_{m-1}(t) \geqq \cdots$,

$$\begin{aligned} \sum_{j=1}^m (-1)^{m-j} y_j(t) \geqq y_1(t) &= y(t), \qquad m \text{ odd}, \\ &\geqq 0, \qquad m \text{ even}. \end{aligned}$$

Thus these intervals give a positive contribution to R. All of R is thus accounted for, because the definition of $y = y(\epsilon)$ implies that $\epsilon(x) > \epsilon$ for $x < y$. Hence $R \geqq 0$. Thus, using (86) through (88), we conclude that

$$(89) \quad L_\epsilon(X) \geqq \int_\epsilon^{\lambda_1^{1/2}} y(t) t^{-1} \, dt + \tfrac{1}{2}[y - 1 + \log [\lambda_1/\lambda(y)]] + O(\log y).$$

The terms after the integral in (89) are

$$\tfrac{1}{2}y - \tfrac{1}{2} - \log \epsilon^{-1} + O(\log y) \geqq - \log \epsilon^{-1} + O(1),$$

since y is bounded away from zero as $\epsilon \to 0$. Hence (83) is true. This proves Lemma 12.

The next lemma estimates the number $A = A(\epsilon)$ given by (65) and (66) in terms of the function $y(\epsilon)$ of the preceding lemma. To make these estimates, certain restrictions must be put on the eigenvalues λ_k; these restrictions imply that the influence of the eigenvalues "far out" is not "too large."

LEMMA 13. *Let $A = A(\epsilon)$ be the number in the solution of* (65) *and* (66). *If the Gaussian process X has an infinite number of positive eigenvalues, and*

$$\sum_{k=n}^\infty \lambda_k = O(n\lambda_n)$$

when the eigenvalues are arranged in non-increasing order, then (65) *and* (66) *have a solution, and $A\epsilon^2 = O(y(\epsilon))$ as $\epsilon \to 0$. If the stronger condition*

$$(90) \qquad \sum_{k=n}^\infty \lambda_k = o(n\lambda_n)$$

holds, then $A\epsilon^2 \sim y(\epsilon)$.

PROOF. Under the first hypothesis,

$$(91) \qquad \sum_{k=n}^\infty \lambda_k \leqq Dn\lambda_n$$

for all n, where D is a positive constant. Using the inequality $\log x \leqq x - 1$, we have for $L > N' \geqq 1$

$$\begin{aligned} &\sum_{k=N'+1}^L \lambda_k \log (\lambda_{N'}/\lambda_k) \\ &\quad = \sum_{k=N'+1}^L \lambda_k \sum_{m=N'+1}^k \log (\lambda_{m-1}/\lambda_m) \leqq \sum_{m=N'+1}^L (\lambda_{m-1}/\lambda_m - 1) \sum_{k=m}^L \lambda_k \\ &\quad \leqq \sum_{m=N'+1}^L (\lambda_{m-1}/\lambda_m - 1) Dm\lambda_m = D[(N' + 1)\lambda_{N'} + \sum_{m=N'+1}^{L-1} \lambda_m - L\lambda_L] \\ &\quad \leqq D(N' + 1)\lambda_{N'} + D^2(N' + 1)\lambda_{N'+1}. \end{aligned}$$

Hence, letting $L \to \infty$,

$$(92) \qquad \sum_{k=N'+1}^\infty \lambda_k \log (\lambda_{N'}/\lambda_k) \leqq (D\lambda_{N'} + D^2\lambda_{N'+1})(N' + 1).$$

This shows that $\sum \lambda_k \log \lambda_k^{-1}$ converges. By Theorem 2, $J_\epsilon(X)$ is finite and (65) and (66) have a solution.

Using the function $r(x)$ defined in the proof of Theorem 3, we have, as in (74),

$$A\epsilon^2 = \sum_{k=1}^{\infty} A\lambda_k r(A\lambda_k)^2. \tag{93}$$

Let $\delta < 1$ be a positive number, and $N' = N'(\epsilon)$ the positive integer for which

$$A\lambda_{N'} > \delta \geqq A\lambda_{N'+1}. \tag{94}$$

Also let the first condition of the hypotheses hold. We have by (73), for $x < \delta$,

$$xr(x)^2 = O(x \log (e/x)).$$

Hence

$$\sum_{k=N'+1}^{\infty} A\lambda_k r(A\lambda_k)^2 = O(\sum_{k=N'+1}^{\infty} A\lambda_k \log (e/A\lambda_k)). \tag{95}$$

The inequality (92) was shown by using (91) only for $n > N'$. Hence we can replace $\lambda_{N'}$ by $\epsilon\delta/A$ there, and, by (94),

$$\sum_{k=N'+1}^{\infty} A\lambda_k \log (e\delta/A\lambda_k) \leqq (eD + D^2)\delta(N' + 1).$$

Breaking up the logarithm into two parts, and applying (91) and (94),

$$\begin{aligned}\sum_{k=N'+1}^{\infty} A\lambda_k \log (e/A\lambda_k) &\leqq A \log \delta^{-1} \sum_{k=N'+1}^{\infty} \lambda_k + (eD + D^2)\delta(N' + 1) \\ &\leqq A \log \delta^{-1} \cdot D(N' + 1)\delta/A + (eD + D^2)\delta(N' + 1) \\ &= O(N'\delta \log (e/\delta)).\end{aligned}$$

Thus, from (95),

$$A\epsilon^2 = \sum_{k=1}^{N'} A\lambda_k r(A\lambda_k)^2 + O(N'\delta \log (e/\delta)). \tag{96}$$

The function $xr(x)^2$ approaches 1 as $x \to \infty$. If

$$B = \sup_{0<x<\infty} xr(x)^2,$$

$$A\epsilon^2 \leqq N'B + O(N'\delta \log (e/\delta)) = O(N'). \tag{97}$$

Multiplying by $\lambda_{N'}$, we have, by (94),

$$\epsilon^2 = O(N'\lambda_{N'}/\delta).$$

Let $N_1 = N_1(\epsilon)$ be the smallest integer not less than $y(\epsilon)$, so that $N_1\lambda_{N_1} \leqq \epsilon^2$; then there is a constant C_1 such that

$$N_1\lambda_{N_1} \leqq C_1 N'\lambda_{N'} \tag{98}$$

for ϵ sufficiently small. This inequality implies that there is a constant C_2 such that

$$N' \leqq C_2 N_1. \tag{99}$$

To show this, we can assume that $N' > N_1$. Then by (91),

$$\begin{aligned} D^2 N_1 \lambda_{N_1} &\geqq D \textstyle\sum_{k=N_1}^{N'} \lambda_k \geqq \sum_{k=N_1}^{N'} k^{-1} \sum_{l=k}^{N'} \lambda_l \\ &\geqq \lambda_{N'} \textstyle\sum_{k=N_1}^{N'} (N' + 1 - k)/k > \lambda_{N'} \int_{N_1}^{N'} (N' - x)/x \, dx \\ &= N' \lambda_{N'} (\log (N'/N_1) - 1 + N_1/N'). \end{aligned}$$

From (94) and (98), we must have

$$\log (N'/N_1) - 1 + N_1/N' \leqq C_1 D^2.$$

Thus if we define $x = F(t)$ to be the solution of $\log x - 1 + 1/x = t$, $x > 1$, (99) is true, with $C_2 = F(C_1 D^2)$.

Now by (97) and (99),

$$A\epsilon^2 = O(N') = O(N_1) = O(y(\epsilon)),$$

which proves the first part of the lemma.

To prove the second part of the lemma, let M be a large positive number. If

$$A\lambda_{N''+1} \leqq M < A\lambda_{N''}$$

then by (94)

$$\lambda_{N''+1} \leqq M\delta^{-1}\lambda_{N'},$$

and, since $N'' + 1 \leqq N' + 1 \leqq 2N'$,

$$(N'' + 1)\lambda_{N''+1} \leqq 2M\delta^{-1}N'\lambda_{N'}.$$

As above,

$$N' \leqq F(2MD^2/\delta)(N'' + 1).$$

Let

$$g_M = \max_{x \geqq M} |xr(x)^2 - 1|,$$

so that $\lim_{M\to\infty} g_M = 0$. Then

$$\begin{aligned} &|\textstyle\sum_{k=1}^{N'} A\lambda_k r(A\lambda_k)^2 - N'| \\ &\qquad \leqq N'' g_M + (N' - N'')(1 + B) \qquad (100) \\ &\qquad \leqq \{1 + [1 - 1/F(2MD^2/\delta)]N'\}(1 + B) + g_M N'. \end{aligned}$$

Under hypothesis (90), the inequality (91) is valid for arbitrarily small D if n is sufficiently large. For all inequalities developed above in this proof, we need (91) only for n such that $\lambda_n \leqq M/A$. Hence D can be arbitrarily small, since $A \to \infty$ as $\epsilon \to 0$. Since

$$\lim_{D\to 0} F(CD^2) = 1,$$

we see from (100) that

$$|\textstyle\sum_{k=1}^{N'} A\lambda_k r(A\lambda_k)^2 - N'| \leqq g_M N' + o(N').$$

Applying this to (96),

$$|A\epsilon^2 - N'| \leqq g_M N' + O(N'\delta \log (e/\delta)),$$

which can be made less than any multiple of N' by taking δ small, M large. Hence

$$A\epsilon^2 = N'[1 + o(1)]. \tag{101}$$

In particular, if $C_3 > 1$, we ultimately have

$$N' + 1 \leqq C_3 A\epsilon^2,$$

and, by (94) and the definition of $N = N(\epsilon)$,

$$(N' + 1)\lambda_{N'+1} \leqq C_3 \delta N \lambda_N .$$

This equation is another equation of the type (98). It therefore implies

$$N \leqq F(C_3\delta D^2)(N' + 1).$$

Thus we have from (99)

$$N'/F(C_1 D^2) - 1 \leqq N \leqq F(C_3\delta D^2)(N' + 1).$$

The values of the function F here can be made arbitrarily close to 1 by making D small. Hence $N \sim N'$. Combining this result with (101),

$$A\epsilon^2 \sim N' \sim N \sim y(\epsilon).$$

This completes the proof of Lemma 13.

We now are ready to proceed with the main result of this section.

THEOREM 4. *Let X be a mean-continuous Gaussian process on the unit interval with infinitely many non-zero eigenvalues $\{\lambda_n\}$ arranged in non-increasing order. If*

$$\textstyle\sum_{k=n}^{\infty} \lambda_k = O(n\lambda_n),$$

then, as $\epsilon \to 0$, we have

$$J_\epsilon(X) = O(L_\epsilon(X)).$$

If the stronger condition

$$\textstyle\sum_{k=n}^{\infty} \lambda_k = o(n\lambda_n) \tag{102}$$

holds, we have $J_\epsilon(X) \sim L_\epsilon(X)$.

PROOF. Denoting the solution of (65) and (66) by $\{\delta_k(\epsilon)\}$, $A(\epsilon)$, we have

$$\begin{aligned} h(\delta_k(\epsilon)) &= \textstyle\int_\epsilon^\infty -h'(\delta_k(t))(d\delta_k(t)/dt)\, dt \\ &= \textstyle\int_\epsilon^\infty A(t)\lambda_k\delta_k(t)(d\delta_k/dt)\, dt. \end{aligned}$$

Since the integrands are non-negative, we can interchange the order of summation and integration in (67). We find

$$J_\epsilon(X) = \textstyle\int_\epsilon^\infty A(t) \sum_{k=1}^{\infty} \lambda_k\delta_k(t)(d\delta_k(t)/dt)\, dt = \int_\epsilon^\infty tA(t)\, dt,$$

since $\sum \lambda_k \delta_k^2 = t^2$. Thus

$$J_\epsilon(X) = \int_\epsilon^{\lambda_1^{\frac{1}{2}}} tA(t)\,dt + O(1), \tag{103}$$

since $J_\epsilon(X)$ is finite.

When infinitely many eigenvalues are positive, $y(t) \to \infty$ as $t \to 0$. Hence, by Lemma 13, under the first hypothesis

$$tA(t) = O(y(t)/t) = O([y(t) - 1]/t),$$

and

$$\int_\epsilon^{\lambda_1^{\frac{1}{2}}} tA(t)\,dt = O\left(\int_\epsilon^{\lambda_1^{\frac{1}{2}}} [y(t) - 1]t^{-1}\,dt\right).$$

Since the second integral approaches infinity as $\epsilon \to 0$, the terms $O(1)$ in (103) and (83) are negligible, and by Lemma 12

$$J_\epsilon(X) = O(L_\epsilon(X)).$$

Similarly, under the stronger hypothesis (102),

$$J_\epsilon(X) \sim \int_\epsilon^{\lambda_1^{\frac{1}{2}}} [y(t) - 1]t^{-1}\,dt \lesssim L_\epsilon(X).$$

Asymptotic equality must hold here because $L_\epsilon(X) \leqq H_\epsilon(X) \leqq J_\epsilon(X)$. This completes the proof of Theorem 4.

An important consequence of Theorem 4 is the next result, which has been proved within Theorem 4.

THEOREM 5. *Let X be a mean-continuous Gaussian process on the unit interval with infinitely non-zero eigenvalues $\{\lambda_n\}$ arranged in non-increasing order. If*

$$\sum_{k=n}^{\infty} \lambda_k = O(n\lambda_n),$$

then

$$J_\epsilon(X) = O\left(\int_\epsilon^{\lambda_1^{\frac{1}{2}}} y(t)t^{-1}\,dt\right).$$

If the stronger condition

$$\sum_{k=n}^{\infty} \lambda_k = o(n\lambda_n)$$

holds, then

$$J_\epsilon(X) \sim \int_\epsilon^{\lambda_1^{\frac{1}{2}}} y(t)t^{-1}\,dt.$$

Note that Theorem 5 applies in the case of Theorem 3, but gives less precise information.

Since $J_\epsilon(X) \geqq H_\epsilon(X) \geqq L_\epsilon(X)$, Theorem 4 can be thought of as a condition for

$$J_\epsilon(X) = O(H_\epsilon(X)), \quad \text{or} \quad J_\epsilon(X) \sim H_\epsilon(X). \tag{104}$$

In the former case, X can be transmitted by product partitions with a number of bits not worse than the optimal system by more than a constant multiple. For processes with the stronger property (102), the product partition system is asymptotically as good as the best possible system as $\epsilon \to 0$. It can moreover be shown that $J_\epsilon(X)$ can be finite and yet not $O(H_\epsilon(X))$.

We close this paper with an application of Theorem 5 to band-limited processes. That is, let X be a mean-continuous stationary Gaussian process on the real line whose covariance function

$$\rho(\tau) = R(s, s + \tau)$$

has Fourier Transform $dS(f)$ with support in some finite interval. Suppose $dS(f) = a(f)\,df$ with $a(f)$ continuous. Then when X is restricted to the unit interval, it is known ([7], Lemma 2) that

$$\lambda_n \sim n^{-1}(Cn)^{-2n}$$

for some constant C. It is then easily seen that

$$y(\epsilon) \sim \log \epsilon^{-1}/\log \log \epsilon^{-1}.$$

Theorem 5 now implies that

$$J_\epsilon(X) \sim \int_\epsilon^{\lambda_1^{\frac{1}{2}}} [\log (1/t)/\log \log (1/t)]t^{-1}\, dt,$$

so that

$$J_\epsilon(X) \sim \tfrac{1}{2}[\log (1/\epsilon)]^2/[\log \log (1/\epsilon)]. \tag{105}$$

Equation (105) shows that band-limited processes are not much more random than finite-dimensional distributions, since $J_\epsilon(X)$ does not increase much more rapidly than a constant times log $(1/\epsilon)$. This is to be expected, since the sample functions are analytic with probability 1, and not very "random."

REFERENCES

[1] LOÈVE, MICHEL (1955). *Probability Theory—Foundations, Random Sequences*. Van Nostrand, New York.

[2] POSNER, EDWARD C., RODEMICH, EUGENE R., and RUMSEY, HOWARD, JR. (1967). Epsilon entropy of stochastic processes. *Ann. Math. Statist.* **38** 1000–1020.

[3] POSNER, EDWARD C., RODEMICH, EUGENE R., and RUMSEY, HOWARD, JR. (1969). Epsilon entropy of Gaussian processes. *Ann. Math. Statist.* To appear.

[4] POSNER, EDWARD C. and RODEMICH, EUGENE R. (1969). Epsilon entropy and data compression. In preparation.

[5] GARSIA, ADRIANO M., POSNER, EDWARD C. and RODEMICH, EUGENE R. (1967). Some properties of the measures on function spaces induced by Gaussian processes. *J. Combinational Theor.* **2** 150–161.

[6] WIDOM, HAROLD (1964). Asymptotic behavior of eigenvalues of certain integral operators. *Arch. Rat. Mech.* **17** 215–229.

[7] WIDOM, HAROLD (1963). Asymptotic behavior of eigenvalues of certain integral equations. *Trans. Amer. Math. Soc.* **109** 278–295.