A Real Variable Lemma and the Continuity of Paths of Some Gaussian Processes

A. M. GARSIA, E. RODEMICH & H. RUMSEY JR.

Communicated by M. ROSENBLATT

Introduction. The now standard method introduced by Doob of constructing a separable and measurable model for a mean continuous stochastic process, starting from a given consistent system of joint distributions, although somewhat arbitrary, is perhaps unavoidable in the general case.

However, for Gaussian processes one would hope for a more direct approach. A procedure, frequently used by communications engineers in the Gaussian case, is to expand the paths in terms of the eigenfunctions of the covariance kernel. The resulting expression, which is usually referred to as the Karhunen–Loève expansion and was apparently introduced quite early by M. Kač, is a very natural tool to use.

The purpose of this paper is to show that, at least in the cases when the paths are known to be almost surely continuous this procedure can indeed be used to produce the desired models. In fact, it will also be shown that the best possible estimates for the modulus of continuity of the sample paths can be directly obtained from a study of the partial sums of this expansion.

Our main tool here is a real variable lemma whose significance will, perhaps, transcend the applications that have led us to its discovery. This is a lemma which, roughly speaking, states that the finiteness of a certain integral involving a given function has a bearing on the modulus of continuity of such a function.

Results of similar nature have appeared in Fourier analysis and partial differential equations, but as far as we know, no result of the generality of our Lemma 1.1 has appeared in the literature before.

As we shall see, the power of our lemma lies in the fact that it provides a step which enables us to pass from global estimates, often readily available in a probabilistic setting, to local estimates which usually appear to be of a more elusive nature.

1. The Lemma. Here and in the following, $\Psi(u)$ will denote a non-negative even function on $(-\infty, +\infty)$ and $p(u)$ a non-negative even function on $[-1, 1]$.

Indiana University Mathematics Journal, Vol. 20, No. 6 (1970).

We shall assume that $p(0) = 0$ and $\Psi(\infty) = \infty$. Furthermore, $p(u)$ is supposed to be continuous while $p(u)$ and $\Psi(u)$ are to be both non decreasing for $u \geqq 0$.

For $u \geqq \Psi(0)$ we shall set

$$\Psi^{-1}(u) = \sup \{v: \Psi(v) \leqq u\},$$

while for $0 \leqq u \leqq p(1)$ we set

$$p^{-1}(u) = \max \{v: p(v) \leqq u\}.$$

We can now state our basic estimate.

Lemma 1.1. Let $f(x)$ be defined and continuous on $[0, 1]$ *and suppose that*

1.1 $$\int_0^1 \int_0^1 \Psi\left(\frac{f(x) - f(y)}{p(x - y)}\right) dx\, dy \leqq B < \infty$$

then for all $s, t \in [0, 1]$ we have

1.2 $$|f(s) - f(t)| \leqq 8 \int_0^{|s-t|} \Psi^{-1}\left(\frac{4B}{u^2}\right) dp(u).$$

Proof. We shall first prove the inequality in 1.2 for $|f(1) - f(0)|$. To this end we let

$$I(t) = \int_0^1 \Psi\left(\frac{f(s) - f(t)}{p(s - t)}\right) ds.$$

From 1.1 it then follows that for some $t_0 \in (0, 1)$

$$I(t_0) \leqq B.$$

This given, we choose a sequence $\{t_n\}$ converging to zero and satisfying

$$t_0 > t_1 > t_2 > \cdots,$$

by induction, as follows. Given t_{n-1}, define

1.3 $$d_{n-1} = p^{-1}[\tfrac{1}{2}p(t_{n-1})],$$

and choose

1.4 $$t_n \leqq d_{n-1}$$

in such a way that

1.5 $$I(t_n) \leqq \frac{2B}{d_{n-1}}$$

and in addition

1.6 $$\Psi\left(\frac{f(t_n) - f(t_{n-1})}{p(t_n - t_{n-1})}\right) \leqq \frac{2I(t_{n-1})}{d_{n-1}}.$$

It is always possible to do this, since each of the last two inequalities can be violated only on a set of t_n's of measure less than $1/2\ d_{n-1}$.

Note that since $d_n \leqq d_{n-1}$, for $n \geqq 1$, 1.5 implies

$$I(t_n) \leqq \frac{2B}{d_n}.$$

This inequality is trivially true also for $n = 0$. Hence, for $n \geqq 1$, from 1.6 we derive

1.7 $$|f(t_n) - f(t_{n-1})| \leqq p(t_{n-1} - t_n)\Psi^{-1}\left(\frac{2I(t_{n-1})}{d_{n-1}}\right)$$
$$\leqq p(t_{n-1} - t_n)\Psi^{-1}\left(\frac{4B}{d_{n-1}^2}\right).$$

By 1.3

$$p(t_{n-1} - t_n) \leqq p(t_{n-1}) = 2p(d_{n-1}),$$

1.3 and 1.4 then give

$$p(d_n) = \tfrac{1}{2}p(t_n) \leqq \tfrac{1}{2}p(d_{n-1}),$$

hence

$$p(t_{n-1} - t_n) \leqq 4[p(d_{n-1}) - p(d_n)].$$

Combining with 1.7 we obtain

$$|f(t_0) - f(0)| \leqq 4 \sum_{n=1}^{\infty} [p(d_{n-1}) - p(d_n)]\Psi^{-1}\left(\frac{4B}{d_{n-1}^2}\right)$$
$$\leqq 4 \sum_{n=1}^{\infty} \int_{d_n}^{d_{n-1}} \Psi^{-1}\left(\frac{4B}{u^2}\right) dp(u)$$
$$\leqq 4 \int_0^1 \Psi^{-1}\left(\frac{4B}{u^2}\right) dp(u).$$

Working in a similar fashion with $f(1 - t)$ instead of $f(t)$ we can obtain the same bound for

$$|f(t_0) - f(1)|.$$

Hence

1.8 $$|f(1) - f(0)| \leqq 8 \int_0^1 \Psi^{-1}\left(\frac{4B}{u^2}\right) dp(u).$$

Plainly our lemma is vacuous if the above integral is divergent. Thus here and after we shall assume that it is finite.

This given, to prove 1.2 for general s and t we set

$$\bar{f}(t') = f(s + t'(t - s)), \quad \text{for} \quad 0 \leqq t' \leqq 1$$
$$\bar{p}(u) = p(u\,|s - t|).$$

Upon restricting the range of integration in 1.1 and carrying out a change of variables we get

$$\int_0^1 \int_0^1 \Psi\left(\frac{\bar{f}(s') - \bar{f}(t')}{\bar{p}(s' - t')}\right) ds'\, dt' \leqq \frac{B}{|s - t|^2}.$$

So by 1.8 we deduce

$$|f(s) - f(t)| = |\bar{f}(1) - \bar{f}(0)| \leqq 8 \int_0^1 \Psi^{-1}\left(\frac{4B}{u^2\, |s - t|^2}\right) dp(u\, |s - t|).$$

However, this is none other than 1.2 after a further change of variables.

2. Applications. Let $R(s, t)$ be a continuous positive definite, symmetric kernel on $[0, 1] \times [0, 1]$. From Mercer's Theorem, $R(s, t)$ has an eigenfunction expansion

2.1 $$R(s, t) = \sum_{n=1}^{\infty} \lambda_n \varphi_n(s) \varphi_n(t)$$

which converges uniformly in the square $[0, 1] \times [0, 1]$. The φ_n's form an orthonormal system of continuous functions and for each $n \geqq 1$ we have

2.2 $$\lambda_n \varphi_n(s) = \int_0^1 R(s, t) \varphi_n(t)\, dt.$$

Further, the λ_n's are non-negative and

$$\sum_{n=1}^{\infty} \lambda_n = \int_0^1 R(s, s)\, ds.$$

It is well known, that for each such kernel $R(s, t)$ there is a real, mean continuous, separable and measurable Gaussian process $X_t(\omega)$ on $[0, 1]$ such that for all $t, s \in [0, 1]$

2.3 $$E(X_t X_s) = R(s, t).$$

Perhaps the most natural way to produce such a process is to write formally

2.4 $$X_t(\omega) = \sum_{n=1}^{\infty} (\lambda_n)^{1/2} \varphi_n(t)\, \theta_n(\omega)$$

where $\{\theta_n(\omega)\}$ is a sequence of independent standard (*i.e.* mean zero, variance one) Gaussian variables. This is the expansion we referred to in the introduction.

Starting from 2.2 one easily verifies, at least when only a finite number of the λ_n's are different from zero, that the expression in 2.4 defines a process with the desired properties.

However, since in general the series in 2.4, without a deeper analysis, can only be shown to converge almost everywhere in t with probability one, this expansion is not used to show the existence of separable models. This, notwithstanding the fact that N. Wiener used precisely this approach in his famous construction of a model for the Brownian motion process.

We shall show here that Wiener's approach can be carried through in a much greater generality. Indeed, for the record, it is perhaps worthwhile mentioning that it was precisely our desire to find a substitute for Bernstein's inequality for trigonometric polynomials (which is a crucial tool in Wiener's proof) that led us to our lemma 1.1.

Let us get on now with our arguments. For convenience we set

$$\Delta R(s, t) = R(s, s) + R(t, t) - 2R(s, t) = \sum_{n=1}^{\infty} \lambda_n(\varphi_n(s) - \varphi_n(t))^2.$$

Then, define for $|u| \leqq 1$

$$p(u) = \max_{|s-t| \leq u} (\Delta R(s, t))^{1/2}.$$

We can now state our result in the following fashion

Theorem 2.1 Suppose, that

2.5 $$\int_0^1 \left(\log \frac{1}{u}\right)^{1/2} dp(u) < \infty,$$

Then, with probability one the series in 2.4 converges uniformly. Furthermore, there is a random variable $B(\omega)$, with finite expectation, such that, for all $n \geqq 1$, the partial sums

$$X_t^{(n)}(\omega) = \sum_{\nu=1}^{n} (\lambda_\nu)^{1/2} \varphi_\nu(t) \theta_\nu(\omega)$$

satisfy

2.6 $$|X_t^{(n)}(\omega) - X_s^{(n)}(\omega)| \leqq 16 \int_0^{|s-t|} \left(\log \frac{4B(\omega)}{u^2}\right)^{1/2} dp(u).$$

Proof. We set

2.7 $$B(\omega) = \sup_n \int_0^1 \int_0^1 \exp\left[\frac{(X_s^{(n)}(\omega) - X_t^{(n)}(\omega))^2}{4p^2(s-t)}\right] ds\, dt,$$

and Lemma 1.1 immediately gives us 2.6.

This given, for each ω where

2.8
$$\text{a) } B(\omega) < \infty,$$
$$\text{b) } \sum_{\nu=1}^{\infty} \lambda_\nu (\theta_\nu(\omega))^2 < \infty,$$

the partial sums

$$X_t^{(n)}(\omega) = \sum_{\nu=1}^{n} (\lambda_\nu)^{1/2} \varphi_\nu(t) \theta_\nu(\omega)$$

are, as functions of t, both equicontinuous and convergent in the square mean, thus uniformly convergent.

In view of the relation

$$E\left(\sum_{\nu=1}^{\infty} \lambda_\nu \theta_\nu^2(\omega)\right) = \sum_{\nu=1}^{\infty} \lambda_\nu < \infty,$$

the condition in 2.8 b) is satisfied with probability one. Thus to complete the proof of the theorem we need only show that $B(\omega)$ has a finite expectation.

To this end note that, for fixed s and t, the random variables

$$P_n(s, t; \omega) = \exp\left[\frac{(X_s^{(n)}(\omega) - X_t^{(n)}(\omega))^2}{8p^2(s-t)}\right],$$

by the convexity of e^{u^2} and the definition of $X_s^{(n)}(\omega)$, form a submartingale. Therefore it follows that

2.9 $$E(\max_{m \leqq n} P_m^2(s, t; \omega)) \leqq 4E(P_n^2(s, t; \omega)).$$

On the other hand, since for each n

$$\frac{X_s^{(n)}(\omega) - X_t^{(n)}(\omega)}{p(s-t)}$$

is Gaussian, has mean zero and variance less than or equal to one, we must have

2.10 $$E(P_n^2(s, t; \omega)) \leqq (2)^{1/2}.$$

Combining 2.9 with 2.10 and integrating with respect to s and t from zero to one, Fubini's Theorem gives

2.11 $$E\left(\int_0^1 \int_0^1 \max_{m \leqq n} P_m^2(s, t; \omega)\, ds\, dt\right) \leqq 4(2)^{1/2}$$

Going back to 2.7 and using the monotone convergence theorem in 2.11 we immediately deduce not only that $B(\omega)$ is integrable but that we must have as well

$$E(B(\omega)) \leqq 4(2)^{1/2}.$$

This completes our proof.

When the only thing we know about a process is its mean continuity, our methods can still be used to obtain results and estimates that are the best that can be obtained without further information.

Indeed, it is well-known that a formal expansion of the form

$$X_t(\omega) = \sum_{\nu=1}^{\infty} (\lambda_\nu)^{1/2} \varphi_\nu(t) \theta_\nu(\omega)$$

can always be obtained for any process whose covariance kernel

$$R(s, t) = E(X_s X_t)$$

has the Mercer's expansion

$$R(s, t) = \sum_{\nu=1}^{\infty} \lambda_\nu \varphi_\nu(s) \varphi_\nu(t).$$

However, although the random variables $\theta_n(\omega)$ will have joint distributions which are completely determined by the joint distributions of the sample functions of the given process, all we can say in the general case is that the $\theta_n(\omega)$'s form an orthonormal set. Nevertheless, that is all that is required for our application.

We can state the relevant theorem in this case in the following form:

Theorem 2.2. Suppose that for the Mercer kernel

$$R(s, t) = \sum_{\nu=1}^{\infty} \lambda_\nu \varphi_\nu(s) \varphi_\nu(t)$$

we have

$$\int_0^1 \int_0^1 \frac{\Delta R(s, t)}{p^2(s - t)} \, ds \, dt < \infty,$$

with a function $p(u)$ which, in addition to the conditions set forth in Section 1, *is such that*

$$\int_0^1 \frac{dp(u)}{u} < \infty.$$

Then, for any orthonormal set of random variables $\{\theta_n(\omega)\}$ there is a sequence of partial sums

$$X_t^{(n_k)}(\omega) = \sum_{\nu=1}^{n_k} \lambda_\nu \varphi_\nu(t) \theta_\nu(\omega)$$

which converges uniformly in [0, 1] *with probability one. Furthermore, there is a random variable $B(\omega)$ with finite expectation such that, for all $s, t \in$* [0, 1] *and for all k*

2.12 $$|X_s^{(n_k)}(\omega) - X_t^{(n_k)}(\omega)| \leqq 16(B(\omega))^{1/2} \int_0^{|s-t|} \frac{dp(u)}{u}.$$

Proof. By a theorem of Marcinkiewicz [13], there is an increasing sequence of integers $\{n_k\}$ and a constant c, depending on $\{\theta_n(\omega)\}$, such that for all k and for all $a_1, a_2, \cdots, a_n, \cdots$

2.13 $$E\left(\max_{1 \leq h \leq k} \left| \sum_{\nu=1}^{n_h} a_\nu \theta_\nu(\omega) \right|^2 \right) \leqq cE\left(\left| \sum_{\nu=1}^{n_k} a_\nu \theta_\nu(\omega) \right|^2 \right).$$

Using this sequence $\{n_k\}$ we define

$$B(\omega) = \sup_{n_k} \int_0^1 \int_0^1 \left(\frac{X_s^{(n_k)}(\omega) - X_t^{(n_k)}(\omega)}{p(s - t)} \right)^2 ds \, dt$$

and Lemma 1.1 immediately gives 2.12.

Thereafter the proof can follow very much the lines of our previous argument, except that we set now

$$P_k(s, t; \omega) = \frac{X_s^{(n_k)}(\omega) - X_t^{(n_k)}(\omega)}{p(s - t)}.$$

From 2.13 we then obtain

2.14 $$E(\max_{h \leq k} P_h^2(s, t; \omega)) \leqq cE(P_k^2(s, t; \omega)).$$

Using the orthonormality of the $\theta_n(\omega)$'s,

$$p^2(s - t)E(P_k^2(s, t; \omega)) = \sum_{\nu=1}^{n_k} \lambda_\nu(\varphi_\nu(s) - \varphi_\nu(t))^2 \leqq \Delta R(s, t).$$

Thus, if we integrate 2.14 with respect to s and t from zero to one, by Fubini's theorem, we get

$$E\left(\int_0^1 \int_0^1 \max_{h \leq k} P_h^2(s, t; \omega)\, ds\, dt\right) \leqq c \int_0^1 \int_0^1 \frac{\Delta R(s, t)}{p^2(s - t)}\, ds\, dt.$$

Finally, from this inequality, by monotone convergence, we deduce

$$E(B(\omega)) \leqq c \int_0^1 \int_0^1 \frac{\Delta R(s, t)}{p^2(s - t)}\, ds\, dt < \infty.$$

This implies $B(\omega)$ is almost surely finite and our arguments can be easily completed.

Remark 2.1. A somewhat neater result than that stated in theorem 2.2 can be obtained when $\{\theta_n(\omega)\}$ is a system of "martingale differences". That is when we have

2.15 $$E(\theta_n \mid \theta_1, \theta_2, \cdots, \theta_m) = 0 \quad \text{for all} \quad m < n.$$

Indeed, under these hypotheses, from the martingale inequalities we obtain

2.16 $$E(\max_{m \leq n} |a_1\theta_1 + \cdots + a_m\theta_m|^2) \leqq 4E(|a_1\theta_1 + \cdots + a_n\theta_n|^2)$$

for all sequences of reals $a_1, a_2, \cdots, a_n, \cdots$.

If we modify the proof of theorem 2.2 so as to make use of the inequalities in 2.16 rather than those in 2.13 we can easily obtain the following result.

Theorem 2.3. *Under the hypotheses of Theorem 2.2, if the system $\{\theta_n\}$ satisfies also the condition 2.15 then the partial sums*

$$X_t^{(n)}(\omega) = \sum_{\nu=1}^{n} (\lambda_\nu)^{1/2} \varphi_\nu(t)\theta_\nu(\omega)$$

converge uniformly in [0, 1] *with probability one. Furthermore there is a random variable $B(\omega)$ with finite expectation such that for all $s, t \in [0, 1]$ and for all $n \geqq 1$*

$$|X_s^{(n)}(\omega) - X_t^{(n)}(\omega)| \leqq 16\sqrt{B(\omega)} \int_0^{|s-t|} \frac{dp(u)}{u}.$$

3. Some further remarks. It is quite natural at this point to wonder whether or not the continuity of paths, at least in the Gaussian case, can always be obtained by proving the uniform convergence of the Karhunen–Loeve expansion.

This is the case, and in fact we can show directly that the partial sums of this expansion are equicontinuous with probability one if and only if the process has continuous paths with probability one.

More precisely we have

Theorem 3.1. *Let $\{X_t(\omega)\}$ be a measurable and separable model of a mean continuous real Gaussian process on* [0, 1]. *Let*

$$R(s, t) = \sum_{\nu=1}^{\infty} \lambda_\nu \varphi_\nu(s)\varphi_\nu(t)$$

be the Mercer's expansion of the covariance function

$$R(s, t) = E(X_s X_t).$$

Set

3.1
$$\theta_n(\omega) = \frac{1}{(\lambda_n)^{1/2}} \int_0^1 X_t(\omega)\varphi_n(t)\, dt.$$

Then

a) *The paths $X_t(\omega)$ are continuous with probability one if and only if*

3.2
$$X_t^{(n)}(\omega) = \sum_{\nu=1}^{n} (\lambda_\nu)^{1/2}\varphi_\nu(t)\theta_\nu(\omega) \rightarrow X_t(\omega)$$

uniformly for $t \in [0, 1]$ with probability one

b) *Furthermore, if the paths are* a.s. *continuous, there are two functions $p(\delta)$ and $M(\delta)$, both tending to zero as $\delta \rightarrow 0$, such that for each $0 < \delta < 1$, outside an event of probability $p(\delta)$, we have:*

$$|X_t^{(n)}(\omega) - X_s^{(n)}(\omega)| \leqq M(|t - s|),$$

for all n *and* for all $|t - s| \leqq \delta$.

Note. A partial result of this nature has been shown by J. Walsh [15] and the full result, as we understand has been independently discovered by other authors, in particular, Ito. (We owe this reference to R. Getoor.) We shall indicate here its proof for sake of completeness and since it doesn't seem to have appeared anywhere in the literature.

Proof. It is clear that we need only show b). To this end, let us set

$$M^*(\omega, \delta) = \sup_n \sup_{|s-t| \leqq \delta} |X_s^{(n)}(\omega) - X_t^{(n)}(\omega)|.$$

This function is clearly a random variable, since the "sup" need only be taken over an appropriate countable set of couples (s_k, t_k). We shall show that for all $\epsilon, \delta > 0$

3.3 $$P[M^*(\omega, \delta) > \epsilon] \leqq 2P[M(\omega, \delta) > \epsilon],$$

where

$$M(\omega, \delta) = \sup_{|s-t| \leq \delta} |X_s(\omega) - X_t(\omega)|.$$

We set for each n

$$Y_t^{(n)}(\omega) = X_t(\omega) - X_t^{(n)}(\omega),$$

The basic idea here is to deduce 3.3 from the independence of $Y_t^{(n)}(\omega)$ and $X_t^{(n)}(\omega)$ together with the symmetry of the "tails" $Y_t^{(n)}(\omega)$.

For clarity we shall illustrate the method by applying it to a simpler situation.

Lemma 3.1. *Let $X_1, X_2, \cdots, X_n; Y_1, Y_2, \cdots, Y_n$ be given random variables with Y_i (for $1 \leqq i \leqq n$) independent of $X_1, X_2, \cdots, X_i$. Suppose that each Y_i is symmetric, so that*

3.4 $$P[Y_i \geqq 0] = P[Y_i \leqq 0] \geqq \tfrac{1}{2}.$$

Then, setting

$$Z_i = X_i + Y_i$$

we have, for all $\epsilon > 0$

3.5 $$P[|Z_1| \vee |Z_2| \vee \cdots \vee |Z_n| > \epsilon] \geqq \tfrac{1}{2}P[|X_1| \vee |X_2| \vee \cdots \vee |X_n| > \epsilon].$$

Proof. Let

$$E_\nu = \{|X_1| \vee |X_2| \vee \cdots \vee |X_{\nu-1}| \leqq \epsilon, |X_\nu| > \epsilon\}.$$

Then, in E_ν $|X_\nu| > \epsilon$ and $|Z_\nu| = |X_\nu + Y_\nu| > \epsilon$ if $X_\nu > 0$ and $Y_\nu \geqq 0$ or if $X_\nu < 0$ and $Y_\nu \leqq 0$. Thus, if we set

$$E_\nu^+ = E_\nu \cap \{X_\nu > 0\}, \qquad E_\nu^- = E_\nu \cap \{X_\nu < 0\}$$

we have

$$\sum_{\nu=1}^n E_\nu^+ \cap \{Y_\nu \geqq 0\} + \sum_{\nu=1}^n E_\nu^- \cap \{Y_\nu \leqq 0\} \subset \{\max_\nu |Z_\nu| > \epsilon\}.$$

Therefore, by the independence assumption and 3.4, we deduce

$$\frac{1}{2} \sum_{\nu=1}^n P(E_\nu^+) + \frac{1}{2} \sum_{\nu=1} P(E_\nu^-) \leqq P\{\max_\nu |Z_\nu| > \epsilon\}.$$

But this is 3.5.

The more complex inequality 3.3 is obtained in a similar way, but some additional organization is needed. First of all we fix a suitable countable set of couples (s_k, t_k) with $|s_k - t_k| \leqq \delta$, and take all suprema only over these couples.

We then set

$$X_k^{(n)}(\omega) = X_{s_k}^{(n)}(\omega) - X_{t_k}^{(n)}(\omega),$$

and

$$E_{\nu,k} = \{\max_{\mu<\nu} \sup_h |X_h^{(\mu)}| \leqq \epsilon; \max_{h<k} |X_h^{(\nu)}| \leqq \epsilon, |X_k^{(\nu)}| > \epsilon\}.$$

This given, the proof of 3.3 can be obtained by operating upon these events $E_{\nu,k}$ as we did upon the E_ν's in the proof of Lemma 3.1.

To complete the proof of Theorem 3.1 we need to exhibit the functions $M(\delta)$ and $p(\delta)$ having the stated properties. To this end we pick a decreasing sequence $\{\epsilon_k\}$ of positive numbers such that $\sum_{n=1}^{\infty} \epsilon_k < \infty$, and determine another decreasing sequence $\{\delta_k\}$ such that for each k

$$P[M(\delta_k, \omega) > \epsilon_k] \leqq \epsilon_k,$$

we can do this since, by our assumptions, $M(\delta, \omega) \to 0$ a.s. as $\delta \to 0$.

We then define

$$M(\delta) = \epsilon_k \quad \text{if} \quad \delta_{k+1} < \delta \leqq \delta_k,$$

and obtain (from 3.3)

$$P\left[\sup_{\delta'<\delta} \frac{M^*(\delta', \omega)}{M(\delta')} > 1\right] \leqq P\left[\sup_{\delta_h \leqq \delta_k} \frac{M^*(\delta_h, \omega)}{\epsilon_h} > 1\right] \leqq 2 \sum_{h=k}^{\infty} \epsilon_k$$

The proof of the theorem is thus completed if we set

$$p(\delta) = 2 \sum_{h=k}^{\infty} \epsilon_h \quad \text{for} \quad \delta_{k+1} < \delta \leqq \delta_k .$$

We have made repetitive use of the fact that an equicontinuous sequence of functions on [0, 1] that converges to zero in the L_2 sense must necessarily converge uniformly. It might be worthwhile at this point to point out a simple inequality that establishes this fact.

Lemma 3.2. *Let $\{X_n(s)\}$ be a sequence of functions on* [0, 1] *such that, for some $M(\delta) \downarrow 0$ as $\delta \downarrow 0$*

$$|X_n(s) - X_n(t)| \leqq M(|t - s|) \quad \text{for all} \quad |t - s| \leqq \delta_0 \quad \text{and for all} \quad n.$$

Then, for all $\delta \leqq \delta_0$ we have

3.6
$$\max_s |X_n(s)| \leqq M(\delta) + \left(\frac{1}{\delta}\int_0^1 X_n^2(t)\,dt\right)^{1/2}.$$

Proof. Let I be an interval of length δ contained in [0, 1] and containing a given point $s \varepsilon [0, 1]$. Trivially

$$\delta M^2(\delta) \geqq \int_I (X_n(s) - X_n(t))^2\,dt$$

$$\geqq X_n^2(s)\,\delta - 2\,|X_n(s)| \int_I |X_n(t)|\,dt + \int_I X_n^2(t)\,dt$$

So, by Schwarz's Inequality, we get

$$\delta M^2(\delta) \geqq \left(X_n^2(s)\,\delta - 2\,|X_n(s)|\left(\delta \int_I |X_n(t)|^2\,dt\right)^{1/2} + \int_I X_n^2(t)\,dt\right)$$

$$\geqq \left(|X_n(s)|(\delta)^{1/2} - \left(\int_I |X_n(t)|^2\,dt\right)^{1/2}\right)^2.$$

from which 3.6 is easily derived.

4. Historical comments. We should hasten to say that, with the possible exception of some of the consequences of Theorem 2.2, none of our results in Section 2 yields any sharpening of the best results of this nature that can be found in the literature.

So, the "raison d'être" for the present paper rests mainly in the novelty of the methods of proof. Indeed, condition 2.5 can be shown to be equivalent to that stated by Fernique in [7]. Our original motivation, however, was not to obtain conditions on $R(s, t)$ which assure the continuity of paths but to prove the uniform convergence of the Karhunen–Loève expansion under such conditions. Nevertheless, even this further fact can be found hidden in a paragraph of a recent paper by Delporte [4].

Delporte's paper is perhaps the most all-inclusive, in these matters, that can be found in the literature. Fernique's paper [7] in this respect is a more elegant and digestible recasting of some of Delporte's work. Fernique, however, among other things, brings forth in [8] a convincing argument that his condition (which is equivalent to ours in 2.5) is the best possible of its kind.

Fernique's work essentially indicates that if our function $p(u)$ is such that the integral in 2.5 diverges, then there may be a covariance function $R(s, t)$ such that

$$\sup_{|s-t|\leqq\delta} \Delta R(s, t) \leqq p(\delta)$$

whose corresponding Gaussian processes $\{X_t\}$ have discontinuous paths with probability one. Delporte's paper [4] contains some rather ingenious estimates and facts, presented however in a functional analysis setting. This, turns out to be a fortunate circumstance, as far as the present paper is concerned, for, had we been acquainted with Delporte's work when we set out to do our investigations we might have never been led to the present proof of these results.

Some further references worth mentioning in this connection are the paper of Ciesielski [2] and that of Belayev [1]. In both of these papers conditions on $\Delta R(s, t)$ are given which assure the a.s. Hölder continuity of paths. These particular results of theirs can be deduced from our Theorem 2.2.

Conditions of a nature different from 2.5 that assure the continuity of paths, have been given by Kahane [10] and extended by M. Nisio [14]. These conditions have the pleasant feature that they are necessary and sufficient but are quite definitely restricted to the stationary case.

Roughly speaking, both Kahane and Nisio show that under certain rather general circumstances, the condition

4.1 $$E(\sup_{t\in[0,1]} |X_t(\omega)|) < \infty$$

is necessary and sufficient for a.s. continuity of paths of a stationary process.

It is to be noted that the sufficiency of 4.1 is an immediate consequence of a result of Kolmogorov (proved by Belayev in [1]), to the effect that for every mean continuous stationary process one of the following alternatives hold: either with probability one the sample functions are continuous or with probability one the paths are essentially unbounded in every interval.

In the non-stationary case matters can be quite different. Indeed, there are some such processes for which much more than 4.1 holds true and yet their paths are discontinuous with probability one.

An example in point can be put together as follows. We pick $\{a_n\}$ so that

$$1 \geqq a_0 > a_1 > \cdots > a_n \downarrow 0.$$

Further we let $\{\varphi_n(t)\}$ be a sequence of continuous functions on [0, 1] with

a) $0 \leqq \varphi_n(t) \leqq 1$,
b) $\max_t \varphi_n(t) = 1$,
c) support $\varphi_n(t) \subset [a_{n+1}, a_n]$.

This given, if $\{\theta_n(\omega)\}$ is a sequence of independent standard normal variables, then the process,

$$X_t(\omega) = \sum_{n=2}^{\infty} \frac{\varphi_n(t)\theta_n(\omega)}{(\log n)^{1/2}},$$

can be shown to satisfy the condition

$$E(\exp [a \max_t |X_t(\omega)|^2]) < \infty,$$

for all sufficiently small $a > 0$. Nevertheless it is easy to see that $X_t(\omega)$ as a.s. discontinuous at 0.

A more elaborate version of the same basic idea can be used to construct processes whose set of discontinuities, with probability one, fill any preassigned countable union of closed nowhere dense sets [16].

Before closing, we would like to point out an interesting question that we have not been able to settle. It would seem from the above considerations and especially from condition 2.5 that the "size" of $\Delta R(s, t)$ in a neighborhood of the "diagonal" $s = t$ of the square $[0, 1] \times [0, 1]$ is a determining factor for the continuity of paths of the corresponding Gaussian processes. It would seem then that if for two covariances $R_1(s, t)$, $R_2(s, t)$ we have for all (s, t) in $[0, 1] \times [0, 1]$

4.2 $$\Delta R_1(s, t) \leqq \Delta R_2(s, t),$$

and if the R_2 Gaussian process has continuous paths with probability one so will the R_1 process.

It is interesting to note that, if instead of 4.2 we require the stronger condition that the difference

$$R(s, t) = R_2(s, t) - R_1(s, t)$$

be itself a covariance function, then this implication can be drawn.

The basic reason is that under this assumption we can produce a separable R_2-process of the form

4.3 $$X_t^{(2)}(\omega) = X_t^{(1)}(\omega) + X_t(\omega)$$

with $\{X_t^{(1)}\}$ and $\{X_t\}$ independent separable processes having covariances R_1 and R respectively.

This given, it is not difficult to show, by using another modified version of lemma 3.1, that if $\{X_t^{(2)}\}$ has continuous paths then both $\{X_t^{(1)}\}$ and $\{X_t\}$ must have continuous paths as well.

The general question however, seems beyond the reach of these methods and as far as we know it is still open.

REFERENCES

[1] Y. K. BELAYEV, *Continuity and Hölder conditions for sample functions of stationary Gaussian processes, Proc. 4th Berkeley Symp. on Math. Stat. and Prob.*, **2** (1961) 23–33.

[2] Z. CIESIELSKI, *Hölder conditions for realizations of Gaussian processes, Am. Math. Soc. Trans.*, **99** (1961) 403–413.

[3] J. DELPORTE, *Extension des conditions suffisantes pour la construction de fonctions aléatoires normales presque sûrement continues, C. R. Acad. Sci. Paris Sér. A-B.*, **256** (1963) 3816–3819.

[4] J. DELPORTE, *Fonctions aléatoires presque sûrement continues sur un intervalle fermé, Ann. Inst. Henri Poincaré*, **B 1** (1964/65) 111–215.

[5] R. L. DOBRUSHIN, *Continuity properties of sample functions of stationary Gaussian processes, Teor. Veroyatnost i Primenen*, **5** (1960) 132–134.

[6] R. M. DUDLEY, *The sizes of compact subsets of Hilbert space and continuity of Gaussian processes, J. Functional Anal.*, **1** (1967) 290–330.

[7] X. FERNIQUE, *Continuité des processus Gaussiens, C. R. Acad. Sci. Paris Sér. A-B.*, **258** (1964) 6058–6060.

[8] X. FERNIQUE, *Series de variables aléatoires independantes, Publ. Inst. Stat. Univ. Paris*, **16** (1967) no. 1 35–46.

[9] K. ITO & M. NISIO, *On the convergence of sums of independent Banach space valued random variables, Osaka J. Math.*, **5** (1968) 35–48.

[10] J. P. KAHANE, *Series de Fourier aléatoires, Sém. Math. Supér. Univ. de Montréal* (1963).

[11] M. KAČ & A. J. F. SIEGERT, *An explicit representation of a statistical Gaussian process, Am. Math. Stat.*, **18** (1947) 438–442.

[12] K. KARHUNEN, *Sur Spektral theorie stochastischer Prozesse, Ann. Acad. Sci. Fennicae, Ser. A I Math. Phys.*, **34** (1946).

[13] J. MARCINKIEWICZ, *Sur la convergence des séries orthogonales, Studia Math.*, **6** (1936) 39–45.

[14] M. NISIO, *On the continuity of stationary Gaussian processes*, (to appear).

[15] J. B. WALSH, *A note on uniform convergence of stochastic processes, Proc. Amer. Math. Soc.* **18** (1967) 129–132.

[16] E. RODEMICH & H. RUMSEY JR., *Gaussian Processes with Bounded Paths* (to appear).

University of California
Jet Propulsion Lab.

Date communicated: JANUARY 6, 1970