SAMPLE BEHAVIOR OF
GAUSSIAN PROCESSES

M. B. MARCUS\textsuperscript{1}
Northwestern University
and
L. A. SHEPP
Bell Telephone Laboratories

1. Introduction

Doob remarked in his 1953 book that, "very few facts specifically true of Gaussian processes are known." This statement is no longer true; the field is very active, and Gaussian processes now form a very special class. The various zero-one laws discovered for Gaussian processes show that their sample functions behave almost deterministically. Indeed, the deterministic-like properties of Gaussian models even appear nonphysical. (A well-known example is the property of the Wiener and other Gaussian processes that the sample quadratic variation on an interval is constant.) However, we will not pursue this point.

This work surveys some recent results on sample behavior of Gaussian processes and continues the study of the supremum $\|X\|$ of a bounded Gaussian process $X(t)$. It is proved in particular that

\begin{equation}
\lim_{t \to \infty} \frac{1}{t^2} \log P(\|X\| > t) = - (2\sigma^2)^{-1},
\end{equation}

where $\sigma^2$ is the supremum of the variances of the individual $X(t)$. This extends work of Fernique [10] and Landau and Shepp [17].

2. A survey of sample behavior

The question due to Kolmogorov (see [7], which stimulated much of the interest in this area, asks which stationary Gaussian processes have continuous sample paths. A stationary process $X$ has covariance $R(s, t) = R(s - t) = EX(s)X(t)$. The question is then for which nonnegative definite functions $R$ is $X$ a.s. continuous. It is, of course, necessary that $R$ be continuous, and at first it is surprising that this is not also sufficient. Kolmogorov must have had examples of stationary processes with discontinuous paths, probably based on random Fourier series.

\textsuperscript{1}The work of this author was supported in part by National Science Foundation Grant GP20043.
When \( R \) is continuous, we can choose a measurable version of the process and then

\[
E \int_a^b X^2(t) \, dt = \int_a^b R(s, s) \, ds < \infty,
\]

and so \( X \in L^2[a, b] \) a.s. A more interesting question is whether \( X \in C[a, b] \) a.s. It may be made precise by requiring, equivalently, any one of the following:

(a) \( P^0 \) (continuous functions) = 1, \( P^0 = \) outer measure;
(b) there exists a continuous version of the process;
(c) \( P \) is \( \sigma \)-additive on the continuous functions.

However, the simplest way of making the problem precise is the way one actually uses the notion of continuity:

(d) \( X(t) \) is uniformly continuous for \( t \) restricted to the rational points in \([a, b] \), or any countable dense subset of \([a, b] \).

Belyayev [1] proved that if \( X \) is stationary and not a.s. continuous then \( X \) is unbounded on every interval; this was an earlier conjecture of Kolmogorov, and typical of the zero-one laws for Gaussian processes [15]. Eaves [8] has given a simple proof.

Fernique [9] made much progress by determining the form of sufficient conditions on \( R \) for \( X \) to be continuous. Apparently, he was stimulated by work of Delporte [5].

Fernique’s sufficient condition is given in terms of the incremental variance

\[
\sigma^2(s, t) = E((X(s) - X(t))^2) \text{ of } X, \text{ which is not necessarily stationary.}
\]

**Theorem 2.1 (Fernique).** If for \( 0 \leq s \leq t \leq \varepsilon \) there is a function \( \Psi \) for which

\[
E(X(s) - X(t))^2 \leq \Psi^2(t - s), \text{ where } \Psi \text{ is nondecreasing on } [0, \varepsilon]\]

then \( X \) is continuous a.s.

This result was announced by Fernique [9] without a proof. The first published proof belongs to Dudley [6] and is based on carefully chosen partitions of the time set, suggested in [9]. There are now several proofs [21] of Fernique’s theorem; perhaps the most elegant one is due to Garsia, Rodemich, and Rumsey [11].

Fernique also made a valuable contribution toward a converse of Theorem 2.1 by giving a method of constructing examples of discontinuous stationary processes using random lacunary Fourier series,

\[
X(t) = \sum_{n=1}^{\infty} a_n (\Re_n \cos \theta_n t + \Im_n \sin \theta_n t),
\]

where \( \Re \) and \( \Im \) are independent sequences of independent standard normal variables; it is a theorem of Sidon (see [21]) that bounded lacunary Fourier series converge absolutely, and so if \( \theta_n = 2^n \) and \( \Sigma a_n = \infty \), then \( X \) is a.s. discontinuous.
The following result from [21] is dual to Theorem 2.1 and a partial converse. 
Again $X$ need not be stationary.

**Theorem 2.2.** If for $0 \leq s \leq t \leq \epsilon$ there is a function $\Psi$ for which $E(X(s) - X(t))^2 \geq \Psi^2(t - s)$, where $\Psi$ is nondecreasing on $[0, \epsilon]$ and Fernique's integral, $I(\Psi) = \infty$ then $X$ is discontinuous a.s.

In particular, in the stationary case, if the incremental variance $\sigma^2(h) = E(X(t + h) - X(t))^2$ is nondecreasing in some interval $0 \leq h \leq \epsilon$, then $I(\sigma) < \infty$ is necessary and sufficient for sample continuity.

Theorems (2.1) and (2.2) essentially settle Kolmogorov's original question. Simple examples show that $I(\sigma)$ may be infinite and $X$ continuous [21], but no example is known to us of a discontinuous stationary process with $I(\sigma) < \infty$. What is clear is that no integral condition of the form of Theorem 2.1 with $\Psi$ replaced by $\sigma$ can be exactly equivalent to sample continuity because such a condition would have to reduce to $I(\sigma) < \infty$ which, as we have stated above, is not necessary for sample continuity. It seems that the situation is somewhat similar to that in determining the upper class for the general law of the iterated logarithm; one must impose monotonicity conditions in order to give simple general statements.

Actually, Theorem 2.2 is proved in [21] only under the additional assumption that $X$ is stationary. However, in Section 3 of [21], it is shown that if $\Psi$ satisfies the hypothesis of Theorem 2.2 there is a process $Y$ (a random lacunary Fourier series) which is discontinuous and has incremental variance

$$E(Y(s) - Y(t))^2 \leq \Psi^2(t - s)$$

for $0 \leq s \leq t \leq \epsilon$. Theorem 2.2, as stated, then follows from the following lemma, which also answers a question raised in [11].

**Lemma 2.1.** Let $X$ and $Y$ be Gaussian processes for which $E(Y(s) - Y(t))^2 \leq E(X(s) - X(t))^2$ for $0 \leq s \leq t \leq 1$. Then if $X$ is continuous, so is $Y$.

Lemma 2.1 is proved in Section 5 as a corollary of the next lemma, due to Slepian [33] (see also [21]) which is playing an increasing role in sample behavior [2], [25], [26].

**Lemma 2.2 (Slepian).** Let $X_1$ and $X_2$ be separable zero mean Gaussian processes such that for all $s$ and $t$,

$$EX_1(s)X_1(t) \geq EX_2(s)X_2(t),$$

$$EX_1^2(t) = EX_2^2(t).$$

Then for any $M$, $-\infty < M < \infty$,

$$P(X_1(t) > M \text{ for some } t) \leq P(X_2(t) > M \text{ for some } t).$$

Next, let us reconsider Kolmogorov's problem from the spectral point of view, perhaps the more natural for stationary processes. First for the random Fourier series (2.3) with $\theta_n \equiv n$, Kahane [14] (see also [6]) in 1960 found a condition equivalent to sample continuity under monotonicity restrictions. Earlier work had been done by Hunt [12].
Theorem 2.3 (Kahane). In Equation (2.3) set \( \theta_n = n \), and set

\[
2n+1 \sum_{k=2n+1}^{2n+1} (a_k).
\]

(i) If \( X \) is continuous a.s., then \( \sum s_n < \infty \).

(ii) If \( s_n \leq \Psi_n \) where \( \Psi_n \) is nonincreasing and \( \sum \Psi_n < \infty \), then \( X \) is continuous a.s. In particular, if \( s_n \) is eventually nonincreasing, then a necessary and sufficient condition for continuity is that \( \sum s_n < \infty \).

For a general stationary Gaussian process \( X \), we ask for conditions for sample continuity in terms of the spectral distribution function \( F \) determined by

\[
EX(t + s)X(s) = \int_0^\infty \cos tx \, dF(x).
\]

Nisio [24] (see also [17]) proved that Theorem 2.3 remains valid with \( s_n = F(2n+1) - F(2n) \), which agrees with the former definition of \( s_n \) when \( X \) is a Fourier series. Actually, Nisio proved only that Theorem 2.3 holds with the phrase “\( X \) is continuous a.s.” replaced by

\[
E \sup_{0 \leq t \leq 1} |X(t)| < \infty
\]

in both (i) and (ii). It is clear that (2.9) implies that \( X \) is bounded, and hence continuous by the Belyayev theorem. In the opposite direction, if \( X \) is continuous then it is bounded on finite intervals and (2.9) is then a corollary of the following theorem of Fernique [10] and Landau and Shepp [17].

Theorem 2.4. Let \( X_1, X_2, \cdots \) be a sequence of Gaussian variables with arbitrary covariance and means. If \( P(\sup |X_n| < \infty) > 0 \), then \( P(\sup |X_n| < \infty) = 1 \), and there is an \( \varepsilon > 0 \) for which for all sufficiently large \( t \),

\[
P(\sup |X_n| > t) < \exp (-\varepsilon t^2).
\]

Roughly, the theorem asserts that the supremum of a bounded Gaussian process has Gaussian like tails. In Section 5, we give a very neat proof of the theorem due to Fernique [10].

The conclusion of Theorem 2.4 can be strengthened still further. Let \( \sigma_n^2 \) be the variance of \( X_n \) and set \( \sigma^2 = \sup_n \sigma_n^2 \).

Theorem 2.5. Under the hypothesis of Theorem 2.4, \( \sigma^2 < \infty \) and equation (2.10) holds for any \( \varepsilon < (2\sigma^2)^{-1} \).

This result is sharp in the sense that for any \( \varepsilon > (2\sigma^2)^{-1} \), equation (2.10) becomes false for large \( t \) simply because for every \( n \),

\[
P(\sup |X_n| > t) \geq \sup P(|X_n| > t) \geq \exp (-\varepsilon t^2)
\]

for every \( \varepsilon > (2\sigma_n^2)^{-1} \) and \( t \) sufficiently large. On the other hand, with further knowledge of the covariance of \( X \), one may obtain more precise bounds on the tail of the supremum (see Lemma 3.1 and Proposition 1 in [9]).
Let us emphasize that even though the hypothesis that the paths are bounded is often difficult to verify (for example, in the stationary case the verification is precisely Kolmogorov's problem), nevertheless, Theorems 2.4 and 2.5 can be useful. We have already seen Theorem 2.4 applied to Nisio's theorem. Similarly, Theorem 2.4 applies immediately to remove the conditions of some interesting work of M. Pincus on the asymptotics of function space integrals (see Theorem 1, page 202 of [28]). As an example of an application of Theorem 2.5, we obtain the following very special result from the work of Marlow (see (8), page 7 of [22]).

**Theorem 2.6 (Marlow).** Let \( Y = \int_0^1 |W(t)| \, dt \), where \( W \) is the Wiener process. Then

\[
\lim_{t \to \infty} \frac{1}{t^2} \log P\{Y > t\} = -\frac{3}{2}.
\]

Noting that

\[
Y = \sup_{\varphi} \int_0^1 W(t) \varphi(t) \, dt,
\]

where the supremum is taken over step functions \( \varphi \) with values \( \pm 1 \) based on intervals with rational endpoints, Theorem 2.6 follows easily from Theorem 2.5.

Sato [30] has extended Theorem 2.4 to Banach valued \( X \) by writing, similar to (2.13), an arbitrary Banach norm in terms of the supremum of a family of linear functionals.

**Theorem 2.7 (Sato).** Let \( \mu \) be a Gaussian measure on a Banach space \( B \) and assume that \( B^* \) is separable. Then there exists \( \varepsilon > 0 \) such that

\[
\int_B \exp \{\varepsilon \|x\|^2\} \mu(dx) < \infty.
\]

The proof of Theorem 2.5, given in Section 5, uses the following simple lemma, attributed in [11] to Rodemich.

**Lemma 2.3 (Rodemich).** If \( X_1, \ldots, X_i, \ldots \) are any random variables for which

(i) \( Y_i \) is independent of \( (X_1, \ldots, X_i) \) for \( i = 1, \ldots \),

(ii) \( Y_i \) is symmetric for \( i = 1, 2, \ldots \),

then for any \( \theta > 0 \)

\[
P\{\sup_i |X_i| \geq \theta\} \leq 2P\{\sup_i |X_i + Y_i| \geq \theta\}.
\]

Lemma 2.3 is used in [11] to prove that the Karhunen–Loève expansion of a sample continuous process converges uniformly. This result follows also from theorems of Itô and Nisio [13], who generalize, in an elegant way, the three series theorem to abstract valued random variables. Symmetry plays a key role in [13] as well as in Lemma 2.3. Itô and Nisio were more concerned with proving
that

\[ \sum_{j=1}^{\infty} \eta_j \int_0^t \varphi_j \to W(t) \]

uniformly on \(0 \leq t \leq 1\), where \(\{\varphi_1, \varphi_2, \cdots\}\) is a complete orthonormal system, \(\{\eta_1, \eta_2, \cdots\}\) is a standard normal sequence, and \(W\) a standard Wiener process.

The first proof of (2.16) using abstract valued processes is due to J. B. Walsh [35]. The assertion (2.16), itself, apparently first appeared in [31].

Dudley [6], and also Posner, Rodemich, and Rumsey, Jr. [29] have used notions of \(\varepsilon\)-entropy, or efficient coverings by small sets of the part of function space on which the process is carried, in order to prove sample continuity. Neither [6] nor [29] is principally concerned with sample behavior, although the ideas involved are useful for gaining insight and essentially best known results can be obtained by this method. However, Sudakov [34] has shown that \(\varepsilon\)-entropy alone is not sufficient to decide continuity.

Dudley [6] poses an elegant problem. Its solution would probably shed light on the structure of compact subsets of Hilbert space as well as on Gaussian processes. Given any compact subset \(T\) of Hilbert space, define \(X(t)\), \(t \in T\), to be the real valued Gaussian process with mean zero and covariance \(EX(s)X(t) = \langle s, t \rangle\) for \(s \in T\), \(t \in T\). Dudley asks to determine the classes

\[ GB = \{T : \sup_T |X(t)| < \infty \text{ a.s.}\} \]

\[ GC = \{T : X(t) \text{ is continuous on } T \text{ a.s.}\}. \]

Slepian's inequality and Lemma 2.1 can clearly be used to prove either continuity or discontinuity once one constructs an appropriate comparison process. The first person to use Slepian's inequality after Slepian himself was apparently Berman [2]. He showed that if

\[ r_n = EX_{n+m}X_m = o \left( \frac{1}{\log n} \right) \]

for a stationary Gaussian sequence \(X\), or if \(\Sigma r_n^2 < \infty\), then

\[ \{(2 \log n)^{1/2} [\max (X, \cdots, X_n) - (2 \log n)^{1/2}] - \frac{1}{2} \log \log n - \text{constant}\} \]

converges to \(\exp \{- \exp (-x)\}\) in law. Since then much work has been done on such problems for Gaussian processes [25], [36]. They are called growth rate problems and are closely related to sample behavior problems [23]. In particular, Pickands extended many of Berman's results to continuous time [27], and also showed [26] that \(r_n \to 0\) is not enough in (2.18). It is because we do not want to go too far afield that we limit our discussion of growth rates to this paragraph. We also completely omit a discussion of the extensive work connected with level crossings of Gaussian processes, much of which is closely related to sample behavior.
3. Modulus of continuity of sample functions

Let \( X(t), t \in [0, 1] \) be a sample continuous Gaussian process with mean zero and set

\[
\rho^2(h) = \rho_X^2(h) = \sup_{|s-t| \leq h} E(X(s) - X(t))^2.
\]

A function \( f(1/h) \uparrow \infty \) as \( h \downarrow 0 \) is called a uniform modulus of continuity for \( X \) if there exist constants \( 0 < C_0 \) and \( C_1 < \infty \) so that, with probability one,

\[
C_0 \leq \limsup_{|s-t|=h \to 0} \frac{|X(s) - X(t)|}{(\rho^2(h)f(1/h))^{1/2}} \leq C_1.
\]

The function \( f(1/h) \) is called a local modulus of continuity for \( X \) at \( t \) if (3.2) holds with \( t \) fixed. In each case, the problem is to find an \( f \) that works. Given an \( f \) that works, the \( \limsup \) in (3.2) is, by a zero-one law, a constant and so we may take \( C_0 = C_1 = C \). However, in general, \( C \) may be difficult to determine.

### Table I

**Moduli of Continuity**

<table>
<thead>
<tr>
<th>Case</th>
<th>( \sigma^2(h) )</th>
<th>Uniform ( f )</th>
<th>Local ( f )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( h^a \log h \beta )</td>
<td>( 2 \log 1/h )</td>
<td>( 2 \log \log 1/h )</td>
</tr>
<tr>
<td></td>
<td>( 0 &lt; \alpha &lt; 1, -\infty &lt; \beta &lt; \infty )</td>
<td>( C = 1 )</td>
<td>( C = 1 )</td>
</tr>
<tr>
<td>2</td>
<td>( \exp[-</td>
<td>\log h</td>
<td>\beta (\log \log 1/h)^\alpha] )</td>
</tr>
<tr>
<td></td>
<td>( 0 &lt; \alpha &lt; 1, -\infty &lt; \beta &lt; \infty )</td>
<td>( C = 1 )</td>
<td>( (\log \log 1/h)^\beta )</td>
</tr>
<tr>
<td>3</td>
<td>(</td>
<td>\log h</td>
<td>^{-\alpha} (\log \log 1/h)^\beta )</td>
</tr>
<tr>
<td></td>
<td>( 1 &lt; \alpha &lt; \infty, -\infty &lt; \beta &lt; \infty )</td>
<td>( \log 1/h )</td>
<td>( \log 1/h )</td>
</tr>
</tbody>
</table>

Table I gives a sample of the results known [18], [19]. Here \( X \) is stationary and \( E(X(t+h) - X(t))^2 = \sigma^2(h) \). Where \( C \) is not indicated, it is unknown. Finally, in case \( \sigma^2(h) = |\log h|^{-\alpha} (\log \log 1/h)^{-b} \) (where \( b > 2 \) for sample continuity), even \( f \) is unknown. It is remarkable that in Case 1, \( f \) does not depend on \( \alpha \) or \( \beta \), and that in Case 3 the uniform \( f \) is the same as the local \( f \).

To obtain the results in Table I, the required upper bounds in (3.2) were gotten by estimating the supremum of the increments of the process and using the Cantelli lemma [18]. Here, the basic idea is to study the increments based on a sequence of partitions refining at a certain rate depending on \( \rho \), and then to use Fernique's lemma [9]. The following lemma due to Fernique [9] is proved in [20].
LEMMA 3.1 (Fernique). If $I(p) < \infty$ in Theorem 2.1, set $\Gamma(\delta) = \sup [EX^2(t) : t \in 0, \delta]$. Let $c(p) = n^2p$, $n = a$ fixed integer > 1, and $a > (4 \log n)^{1/2}$. Then for any $\delta > 0$,

$$P \left( \sup_{t \in [0, \delta]} |X(t)| \geq a \left[ \Gamma(\delta) + \sqrt{2} \sum_{p=1}^{\infty} \rho(\delta/c(p)) 2^{p/2} \right] \right) \leq Cn^2 \int_{a}^{\infty} \exp \left\{-\frac{1}{4}u^2\right\} du,$$

where $C$ is a constant independent of $\delta$.

Lemma 3.1 implies Theorem 2.5 for the special case when $I(p) < \infty$. Applying Lemma 3.1 to intervals $[0, \delta_k]$ for a sequence $\delta_k \to 0$ fast enough and for values of $n$ dependent upon $\rho(\delta_k)$ enables one to obtain the best known upper bounds for the local modulus of continuity.

It seems to be more difficult to get good lower bounds. When $\sigma^2(h)$ is concave for small $h$, the increments are negatively correlated and even though they are not independent, the Chung–Erdős lemma [4] allows Borel–Cantelli arguments to be used. The following lemma based on Slepian’s inequality (Lemma 2.2) and proved in Section 5, relaxes the concavity assumption on $\sigma^2$, for the local modulus of continuity.

**Lemma 3.2.** Let $X$ and $Y$ be Gaussian and suppose that for $|t - s| \leq \delta$,

$$E(Y(s) - Y(t))^2 \leq E(X(s) - X(t))^2.$$

Suppose further that $\phi$ is a function such that

$$\lim \sup_{|h| \to 0} \frac{|X(h) - X(0)|}{\phi(h)} < 1,$$

a.s.,

where $\rho(h) = \rho_X(h) = o(\phi(k))$ as $h \to 0$ with $\rho$ defined as in (3.1). Then (3.5) holds with $Y$ replacing $X$.

An immediate corollary of Lemma 3.2 is that if $X$ and $Y$ have stationary increments and their incremental variances are asymptotic at zero.

$$\sigma^2_X(h) \sim \sigma^2_Y(h) \quad \text{as} \quad h \to 0,$$

then $X$ and $Y$ have the same local modulus of continuity. In particular, the results in Table I remain valid if $\sigma^2(h)$ is merely asymptotic to any of the $\sigma^2(h)$ given there. Thus, Lemma 3.2 enables us to remove the restriction to concave $\sigma^2$.

Oddly enough, no analogue to Lemma 3.2 for the uniform modulus of continuity is known. It is also unknown whether or not (3.6) implies that $X$ and $Y$ have the same uniform modulus of continuity for processes with stationary increments. Besides the methods already discussed for obtaining lower bounds in the uniform case, there is an approach due to Berman [3], especially noteworthy because concavity of $\sigma^2$ is not required. Even though Berman’s results are so far not as sharp as those obtained by the other methods in the case of
concave $\sigma^2$, his method is interesting and perhaps can be sharpened. The basic idea is that a smooth local time means rough sample paths, as seen below.

For any function $X$, not necessarily random, and any interval $I$ define, with $\lambda$ Lebesgue measure,

$$F_I(x) = \lambda\{t \in I | X(t) \leq x\},$$

(3.7)

the occupation time distribution function of $X$ on $I$. When $F_I$ is absolutely continuous, its density $\phi_I(x)$ is called the local time of $X$ at $x$, since $\phi_I(x) \, dx$ is the amount of time $t \in I$ that $X(t) \in dx$. Finally, define the Fourier transform $f_I$ of local time of $X$,

$$f_I(u) = \int_I \exp\{iuX(t)\} \, dt = \int_{-\infty}^{\infty} \exp\{iux\} \phi_I(x) \, dx.$$

(3.8)

If $X(t) = X(t, \omega)$, then $F_I(x) = F_I(x, \omega)$ and $\phi_I(x) = \phi_I(x, \omega)$. If $X$ is Gaussian and

$$E(X(s) - X(t))^2 \geq C(t - s)^\beta, \quad 0 \leq s \leq t \leq 1, \quad 0 < \beta,$$

(3.9)

then for $p = 2/\beta - 1$, it follows that

$$E \int_{-\infty}^{\infty} |u|^p |f_{[0,1]}(u)|^2 \, du < \infty,$$

(3.10)

which in turn implies that $\phi_I(\cdot, \omega) \in L^2(-\infty, \infty)$ a.s. (if $p \geq 0$) and has $p/2$ derivatives. For $p$ an even integer, we have

$$|I| = \text{total time in } I = \int_{-\infty}^{\infty} \phi_I(x, \omega) \, dx = \int_{b}^{c} \phi_I(x) \, dx$$

$$= \frac{(-1)^{p/2}}{(p/2)!} \int_{b}^{c} (x - b)^{p/2} \phi_I(x, \omega) \, dx,$$

where $(b, c) = (b(\omega), c(\omega))$ is the range of $X(t, \omega)$, $t \in I$, because $\phi_I(x, \omega) = 0$ outside the range of $X$. Using Schwarz's inequality, the fact that $\phi_I(x) \leq \phi_{[0,1]}(x)$ if $I \subset [0, 1]$, and the Parseval identity, we obtain

$$|I|^2 \leq (c - b)^{p+1} \int_{-\infty}^{\infty} |u|^p |f_{[0,1]}(u)|^2 \, du.$$

(3.12)

The range of $X(t), t \in I$ has length $c - b$ and so if the length of $I$ is $h$, taking $(p + 1)$th roots in (3.12), we have

$$\sup_{|t - s| \leq h} |X(t) - X(s)| \geq h^{2(1+p)} \left[ \int_{-\infty}^{\infty} |u|^p |f_{[0,1]}(u)|^2 \, du \right]^{-1}.$$

(3.13)

Thus, for $p = 2/\beta - 1$ an even integer, (3.13) provides a lower bound on the modulus of continuity of $X$, since the denominator of the right side of (3.13) is a.s. finite by (3.10).
Upper bounds on the uniform modulus of continuity have recently been obtained by Garsia, Rodemich, and Rumsey, Jr. [11] by a neat method which gives the best known results. We omit the details of this method which are given elsewhere in this volume. It is shown in [11] that every sample continuous Gaussian process on a compact interval does indeed have a uniform modulus of continuity satisfying (3.2). Under the further restriction that Fernique's condition (in Theorem 2.1) \( I(\rho) < \infty \) holds, they show [11] that for \( 0 \leq s \leq t \leq 1 \),

\[
|X(s) - X(t)| \leq 16D(\omega)\rho(t - s) + 16\sqrt{2} \int_0^{t-s} \left( \frac{1}{u} \right)^{1/2} d\rho(u),
\]

where

\[
E \exp D^2 \leq 4\sqrt{2}.
\]

Dividing (3.14) by the integral on the right and letting \( t - s \to 0 \), for \( s \) and \( t \in [0, 1] \), we obtain

\[
\limsup_{|t-s|=h \downarrow 0} \frac{|X(t) - X(s)|}{\int_0^{t-s} \left( \frac{1}{u} \right)^{1/2} d\rho(u)} \leq 16\sqrt{2}.
\]

It is easy to check that (3.16) gives the upper bounds needed in Table I except for the value of the constant.

At first glance (3.14) appears stronger than (3.16) because (3.14) holds uniformly in \( s \) and \( t \). However, it is easy to see that (3.16) implies (3.14) for some finite \( D(\omega) \). On the other hand, (3.15) does not seem to follow from (3.16) even using Theorem 2.4. Thus, it seems that (3.14) is a new and interesting global formulation of the uniform modulus of continuity deserving further study.

The problem of the uniform and local modulus of continuity can be posed in a sharper way. Call a monotone nondecreasing function \( \varphi \) a member of the upper class with respect to uniform continuity of \( X(t, \omega) \) if there is a \( \delta(\omega) > 0 \) so that for almost all \( \omega \), \( 0 \leq |t - s| \leq \delta(\omega) \) implies

\[
|X(t, \omega) - X(s, \omega)| \leq \sigma(|t - s|)\varphi\left(\frac{1}{|t - s|}\right),
\]

that is, almost all sample functions have \( \sigma(h)\varphi(1/h) \) as a modulus of continuity. The upper class with respect to local continuity is defined similarly.

Following the work of Chung, Erdős, and Sirao [4] on Brownian motion, Sirao and Watanabe [32] have given an integral test for the upper class. They consider stationary processes for which \( \sigma^2(h) \) is concave and

\[
C_0h^\alpha|\log h|^\beta \leq \sigma^2(h) \leq C_1h^\alpha|\log h|^\beta
\]

for \( h \in [0, \delta] \), where \( 0 < \delta \leq 1 \), \( 0 < \alpha \leq 1 \), \( -\infty < \beta < \infty \), and \( 0 < C_0 < C_1 < \infty \). They obtain the following theorem, among others.
Theorem 3.1 (Sirao and Watanabe). A nondecreasing continuous function $\varphi$ belongs to the upper class with respect to uniform continuity if and only if

\[(3.19) \quad \int_{-\infty}^{\infty} \varphi(t)^{4/\alpha - 1} \exp \left\{ -\frac{1}{2} \varphi^2(t) \right\} dt < \infty,\]

where $\alpha$ is given by (3.18).

A similar test is given for local continuity. It seems likely that if $\sigma_1^2(h) \sim \sigma^2(h)$, where $\sigma^2$ is concave and satisfies (3.18), then (3.19) also holds for $\sigma_1^2$. However, Lemma 3.2 is not strong enough to prove this even in the local case. Finally, it would be very interesting to find the upper class for some $\sigma^2$ with $\alpha = 0$, which would be closer to the borderline with discontinuous processes.

Kôno [16] has recently improved and extended parts of [18] and [32] to multidimensional time. Nisio [23] has obtained lower bounds for the modulus of uniform continuity under conditions somewhat weaker than concavity of $\sigma^2$, but only for $\alpha > 0$ (Case 1 of Table I).

4. Inequalities for Gaussian measure of convex sets

Let $Q$ be a Gaussian measure on $\mathbb{R}^n$ with mean zero and let $C$ be a convex set in $\mathbb{R}^n$. Suppose that

\[(4.1) \quad \Phi(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{s} \exp \left\{ -\frac{1}{2} u^2 \right\} du, \quad -\infty < s < \infty\]

and note that $\Phi(s) > \frac{1}{2}$ for $s > 0$.

**Lemma 4.1** (Landau and Shepp [17]). If $Q(C) = \Phi(s)$ for $s > 0$, then $Q(aC) \geq \Phi(as)$ for $a > 1$, where $aC = \{ax : x \in C\}$.

**Remark.** Note that $aC$ is increasing in $a$ because the origin 0 belongs to $C$ (if $0 \notin C$, $C$ would be contained in a half space not containing 0 and $Q(C) < \frac{1}{2}$).

Lemma 4.1 is the basis of the proof of Theorem 2.4 given in [17] and is sharp in the sense that equality holds whenever $C$ is a half space.

When $C$ is known to be symmetric ($x \in C$ implies $-x \in C$) in addition, one would expect that the conclusion of Lemma 4.1 could be strengthened. The following unproved statement is likely to be true.

**Conjecture 4.1.** If $C$ is convex, symmetric, and if $Q(C) = 2\Phi(s) - 1$, then $Q(aC) \geq 2\Phi(as) - 1$ for $a > 1$.

Conjecture 4.1 is sharp in the sense that equality holds when $C$ is a symmetric slab. A weaker version of Conjecture 4.1 has been proved by Fernique [10].

**Lemma 4.2** (Fernique). If $C$ is convex, symmetric, and $q = Q(C) > \frac{1}{2}$, then for $a > 1$,

\[(4.2) \quad Q(aC) \geq 1 - q \exp \left\{ -\frac{a^2}{24} \log \frac{q}{1 - q} \right\}.\]
Fernique [10] states his lemma in terms of seminorms rather than in the form of Lemma 4.2.

Lemma 4.3 (Fernique). Let \( Q \) be a Gaussian measure on \( \mathbb{R}^n \) with mean zero, and let \( N \) be a seminorm on \( \mathbb{R}^n \). Suppose that \( q = Q(x : N(x) \leq s) > \frac{1}{2} \). Then for \( a > 1 \), the inequality (4.2) holds with the left side replaced by \( Q(x : N(x) \leq as) \).

Of course, Lemmas 4.2 and 4.3 are seen to be equivalent under the one to one correspondence,

\[
C = \{ x : N(x) \leq s \},
\]

\[
N(x) = \inf \{ \lambda : \lambda x \in AC \},
\]

between a closed, convex, symmetric set \( C \) and a seminorm \( N \). If Conjecture 4.1 is true, the restriction \( q > \frac{1}{2} \) could be removed in Lemmas 4.2 and 4.3 and the conclusion strengthened. On the other hand, as we shall see, Lemma 4.3 is good enough to prove Theorem 2.4 and so from the point of view of Theorem 2.4 there is no need to get a sharp form of the inequality (4.2)—except for the sake of elegance. This discussion is included only to clarify the relationship between [10] and [17]. We give Fernique’s short proof of Lemma 4.3 below. It is completely elementary compared to the proof in [17] of Lemma 4.1 which depends on the difficult Brunn–Minkowski–Schmidt inequality in spherical geometry.

It is of course simpler to work with Gaussian vectors rather than Gaussian measures, so let \( X_1 \) and \( X_2 \) be independent \( n \) vectors with distribution \( Q \). Then \( (X_1 + X_2)/\sqrt{2} \) and \( (X_1 - X_2)/\sqrt{2} \) are also independent with distribution \( Q \). For any \( s \) and \( t \),

\[
P(N(x) \leq s)P(N(x) > t) = P\left( N\left(\frac{X_1 + X_2}{\sqrt{2}}\right) \leq s, N\left(\frac{X_1 - X_2}{\sqrt{2}}\right) > t \right)
\]

\[
\leq P\left( |N(X_1) - N(X_2)| \leq s, N(X_1) + N(X_2) > t\sqrt{2} \right)
\]

\[
\leq P\left( N(X_1) > \frac{t - s}{\sqrt{2}}, N(X_2) > \frac{t - s}{\sqrt{2}} \right)
\]

\[
= P\left( N(X) > \frac{t - s}{\sqrt{2}} \right)^2.
\]

Define \( t_0 = s > 0 \), \( t_{n+1} = s + \sqrt{2}t_n \), \( n = 1, 2, \ldots \), \( q = P(N(X) \geq s) \), and for \( n \geq 0 \), \( x_n = P(N(X) > t_n)/q \). It follows from (4.4) that \( x_{n+1} \leq x_n^2 \) and so \( x_n \leq (x_0)^{2^n} \). Thus, expressing \( t_n \) in terms of \( s \),

\[
P(N(X) > (2^{(n+1)/2} - 1)(\sqrt{2} + 1)s) \leq q \exp\left\{-2^n \log \frac{q}{1 - q} \right\}.
\]

Now if \( a > 1 \), then for some \( n = 0, 1, 2, \ldots \),

\[
(2^{(n+1)/2} - 1)(\sqrt{2} + 1) \leq a < (2^{(n+2)/2} - 1)(\sqrt{2} + 1)
\]

and (4.2) follows from (4.5) after a short calculation.
In particular, taking \( N(x) = \max (|x_1|, \cdots , |x_n|) \), we obtain from (4.2) that if \( q_n = P(\max (|X_1|, \cdots , |X_n|) \leq s) > \frac{1}{2} \), where \( X_1, \cdots , X_n \) are Gaussian random variables, then

\[
(4.7) \quad P(\max_i |X_i| > t) \leq \exp \left\{ - \frac{t^2}{24s^2} \log \frac{q_n}{1 - q_n} \right\}.
\]

If \( X \) is infinite dimensional, \( X = (X_1, X_2, \cdots) \), and \( s \) is such that \( q = P(\sup_i |X_i| \leq s) > \frac{1}{2} \), then \( q_n > \frac{1}{2} \) and passing to the limit \( n \to \infty \) in (4.7), we obtain

\[
(4.8) \quad P(\sup_i |X_i| > t) \leq \exp \left\{ - \frac{t^2}{24s^2} \log \frac{q}{1 - q} \right\}.
\]

Relation (4.8) gives information about the distribution \( F \) of the supremum of an arbitrary bounded Gaussian process; it states that the tail of \( F \) is at most Gaussian. However, the central part of the distribution \( F \) need not be Gaussian-like at all, as the following example shows. Let \( X_1, X_2, \cdots \) be independent with means zero and variances \( \lambda_1^2, \lambda_2^2, \cdots \). Then for each \( u \),

\[
(4.9) \quad P(\sup_i |X_i| < u) = \prod_{i=1}^{\infty} \left( 1 - 2\Phi\left( - \frac{u}{\lambda_i} \right) \right).
\]

Choosing \( 1/\lambda_i^2 = 2 \log i + a \log \log i \) with \( a > 1 \), we find easily that

\[
(4.10) \quad P(\sup_i |X_i| < 1) = 0, \quad P(\sup_i |X_i| \leq 1) > 0,
\]

so that \( F \) has an atom at unity. On the other hand, it seems unlikely that \( F \) can have more than one atom.

5. Proofs

**Proof of Lemma 2.1.** We prove first that \( Y \) is continuous at each point \( t \), say \( t = 0 \). It is no loss of generality to assume \( X(0) = Y(0) = 0 \) because \( X(t) - X(0) \) and \( Y(t) - Y(0) \) also satisfy the hypothesis of Lemma 2.1. Then \( X(t) \to 0 \) in law, and so

\[
(5.1) \quad \rho^2(\delta) = \sup_{|t| \leq \delta} EX^2(t) \to 0 \quad \text{as } \delta \to 0.
\]

Let \( \delta \) be fixed and set

\[
(5.2) \quad f^2(t) = \rho^2(\delta) - EX^2(t) + EY^2(t), \quad |t| \leq \delta,
\]

\[
\hat{Y}(t) = Y(t) + \eta \rho(\delta), \quad |t| \leq \delta,
\]

\[
\hat{X}(t) = X(t) + \eta' f(t), \quad |t| \leq \delta,
\]

where \( \eta \) and \( \eta' \) are standard normal and independent of each other as well as of \( X \) and \( Y \). Note that \( f^2(t) \geq 0 \). We then have
(5.3) \[ E(\bar{Y}(s) - \bar{Y}(t))^2 = E(Y(s) - Y(t))^2 \leq E(X(s) - X(t))^2 \leq E(\bar{X}(s) - \bar{X}(t))^2 \]

and

(5.4) \[ E\bar{Y}(t)^2 = E\bar{X}(t)^2. \]

Thus, \( E\bar{Y}(s)\bar{Y}(t) \geq E\bar{X}(s)\bar{X}(t) \) for all \( s \) and \( t \) and the hypothesis of Lemma 2.2 is satisfied. Therefore, for any \( \xi > 0 \), by Lemma 2.2,

(5.5) \[ P(\sup |Y(t)| > \xi) \leq 2P(\sup \bar{Y}(t) > \xi) \leq 2P(\sup \bar{X}(t) > \xi), \]

where in each case the sup is taken over \( |t| < \delta \). Since \( EY^2(t) \leq EX^2(t) \leq \rho^2(\delta) \to 0 \), we have

(5.6) \[ \sup_{|t| \leq \delta} |f(t)| \to 0, \quad \sup_{|t| \leq \delta} |\bar{X}(t)| \to 0, \]

a.s. as \( \delta \to 0 \).

By (5.5) and (5.6), \( P(\sup [|Y(t)| : |t| \leq \delta] > \xi) \to 0 \) for each \( \xi > 0 \) as \( \delta \to 0 \) and so

(5.7) \[ \sup_{|t| \leq \delta} |\bar{Y}(t)| \to 0 \]

a.s. as \( \delta \to 0 \).

Hence, \( Y(t) \to 0 \) as \( t \to 0 \) a.s. and \( Y \) is continuous at each point.

The following argument supplied by Dudley, shows that if a Gaussian process \( Y \) is a.s. continuous at each point it is sample continuous. For each \( \varepsilon > 0 \) and each subinterval \( I \) of \([0, 1]\), let \( A_\varepsilon(I) \) be the event that at some \( t_0 \in I \),

(5.8) \[ \limsup_{t \to t_0} Y(t) > \liminf_{t \to t_0} Y(t) + \varepsilon, \]

where \( t \) runs through the rational points of \([0, 1]\). It is a simple zero-one law ([32], Theorem 1) based on the Karhunen–Loève expansion that \( P(A_\varepsilon(I)) \) equals zero or one for each fixed \( \varepsilon \) and \( I \). If \( Y \) is not sample continuous, then there is an \( \varepsilon > 0 \) for which \( P(A_\varepsilon([0, 1])) = 1 \). Since

(5.9) \[ A_\varepsilon[0, \frac{1}{2}] \cup A_\varepsilon[\frac{1}{2}, 1] = A_\varepsilon[0, 1], \]

one of the events on the left has probability one also. Continuing in this way, we find a sequence of closed intervals \( I_n \) nesting to some point \( t_\infty \in [0, 1] \) with \( P(A_\varepsilon(I_n)) = 1 \). Since (5.8) holds for some point \( t_0 = t_0(\omega, n) \in I_n \), it follows that (5.8) must also hold with probability one at \( t_0 = t_\infty \). But \( Y \) is continuous at \( t_\infty \) a.s. The contradiction proves Lemma 2.1.

**Proof of Theorem 2.4.** First observe that if \( P(\sup |X_n| < \infty) > 0 \), then the sequence \( EX_n^2 \) of variances must be bounded because otherwise for each \( M \),

(5.10) \[ P(\sup |X_n| \leq M) \leq \inf_n P(|X_n| \leq M) = 0. \]
Using the Gram–Schmidt procedure, for example, we can find numbers $c_{i,j}$ and a sequence $\eta_1, \eta_2, \cdots$ of independent standard normal variables so that

\begin{equation}
X_i = \sum_{j=1}^{\infty} c_{i,j} \eta_j.
\end{equation}

If now $P(\sup |X_n| < \infty) > 0$, then, as seen above, there is a number $M < \infty$ for which $EX_i^2 \leq M^2$ for all $i$ and so $|c_{i,j}| \leq M$ for all $i$ and $j$. Thus, for any $n$,

\begin{equation}
\left| \sum_{j=1}^{n} c_{i,j} \eta_j \right| \leq M \sum_{j=1}^{n} |\eta_j|.
\end{equation}

Therefore, the event that $\{|X_n|\}$ is bounded is measurable on $\{\eta_{n+1}, \eta_{n+2}, \cdots\}$ for any $n$ and so has measure zero or one. Since zero is excluded the first assertion of Theorem 2.4 is proved.

To prove the second assertion of Theorem 2.4, note that we have just proved the existence of an $s$ for which

\begin{equation}
q = P(\sup |X_i| \leq s) > \frac{1}{2}.
\end{equation}

An application of (4.8) now proves (2.10), since $q > \frac{1}{2}$ and Theorem 2.4 is proved. For an alternate proof of Theorem 2.4 based on Lemma 4.1 see [17].

**Proof of Theorem 2.5.** There is clearly no loss of generality in assuming $\sigma = 1$, so that in (5.11) we have

\begin{equation}
\sum_{j=1}^{\infty} c_{i,j}^2 \leq 1, \quad i = 1, 2, \cdots.
\end{equation}

Define for all $i$ and $n$,

\begin{equation}
x_{i,n} = \sum_{j=1}^{n} c_{i,j} \eta_j,
y_{i,n} = \sum_{j=n+1}^{\infty} c_{i,j} \eta_j,
\end{equation}

so that $X_i = x_{i,n} + y_{i,n}$. Applying Lemma 2.3 to $x_{i,n}$ and $y_{i,n}$ using the ordering $x_{1,1}, x_{2,1}, \cdots, x_{k,1}; \cdots; x_{1,k}, x_{2,k}, \cdots, x_{k,k}$ and letting $k \to \infty$, we obtain that for each $s > 0$,

\begin{equation}
P(\sup_{i,n} |x_{i,n}| > s) \leq 2P(\sup_{i} |X_i| > s).
\end{equation}

By Theorem 2.4, $P(\sup |X_i| < \infty) = 1$, so we can choose $s$ so large that $P(\sup |X_i| > \frac{1}{2}s) < \frac{1}{3}$. Since

\begin{equation}
\sup_{i,n} |y_{i,n}| \leq \sup_{i} |X_i| + \sup_{i,n} |x_{i,n}|,
\end{equation}

we have from (5.16) that

\begin{equation}
P(\sup_{i,n} |y_{i,n}| > s) \leq P(\sup_{i} |X_i| > \frac{1}{2}s) + P(\sup_{i,n} |x_{i,n}| > \frac{1}{2}s)
\leq 3P(\sup_{i} |X_i| > \frac{1}{2}s) < 1.
\end{equation}
It follows that
\[(5.19) \quad P(\limsup_{n \to \infty} \sup_i |y_{i,n}| > s) = 0\]
because the event on the left of (5.19) is a tail event contained in the event on
the left of (5.18). Thus, for any \(q, \frac{1}{2} < q < 1\), there is some value of \(n\) for which
\[(5.20) \quad P(\sup_i |y_{i,n}| \leq s) \geq q.\]
Again appealing to (4.8), using \(y_{i,n}\) in place of \(X_i\), we obtain, for this same value
of \(n\),
\[(5.21) \quad P(\sup_i |y_{i,n}| > t) \leq \exp \left\{ -\frac{t^2}{24s^2} \log \frac{q}{1 - q} \right\}.\]
Now fix a number \(\delta > 0\), define \(s\) as above (below (5.16)), and choose \(q\) so close
to one that \(\theta > \frac{1}{2}\), where
\[(5.22) \quad \theta = \frac{\delta^2}{24s^2} \log \frac{q}{1 - q}.\]
Choosing the same \(n\) which makes (5.21) valid, we have
\[(5.23) \quad P(\sup_i |X_i| > (1 + \delta)t) \leq P(\sup_i |y_{i,n}| > \delta t) + P(\sup_i |x_{i,n}| > t)
\leq \exp \{-t^2\theta\} + P(\sup_i |x_{i,n}| > t).\]
Since \(n\) is fixed in the last term, and from (5.14)
\[(5.24) \quad |x_{i,n}| \leq \left( \sum_{j=1}^{n} \eta_j^2 \right)^{1/2}, \quad i = 1, 2, \ldots,\]
we have
\[(5.25) \quad P(\sup_i |x_{i,n}| > t) \leq P\left( \left( \sum_{j=1}^{n} \eta_j^2 \right)^{1/2} > t \right).\]
The right side of (5.25) is the tail of the \(\chi^2\) distribution and as is easily seen, satisfies
\[(5.26) \quad P\left( \left( \sum_{j=1}^{n} \eta_j^2 \right)^{1/2} > t \right) \leq t^{n-1} \exp \{-\frac{1}{2}t^2\}\]
for sufficiently large \(t\). Using (5.25) and (5.26) to bound the last term in (5.23),
we see that for any \(\delta > 0\) and any \(\epsilon < \frac{1}{2}\) for \(t\) sufficiently large,
\[(5.27) \quad P(\sup_i |X_i| > (1 + \delta)t) \leq \exp \{-t^2\epsilon\}.\]
Finally, replacing \(t\) by \(t/(1 + \delta)\) in (5.27), yields Theorem 2.5.

**Proof of Lemma 3.2.** As in the proof of Lemma 2.1, we may assume
\(X(0) = Y(0) = 0\). Fix \(\delta > 0\) and define \(\rho^2(\delta), f^2(t), \bar{Y}(t), \) and \(\bar{X}(t)\) as in (5.1)
and (5.2). Noting that
BEHAVIOR OF GAUSSIAN PROCESSES 439

\( E\overline{Y}(s)\overline{Y}(t) \geq E\overline{X}(s)\overline{X}(t), \)
\( E\overline{Y}(t)^2 = E\overline{X}(t)^2, \)

for \(-\delta \leq s \leq t \leq \delta\), we may apply Slepian's lemma (Lemma 2.2) to \( \overline{Y}(t)/\varphi(t) \) and \( \overline{X}(t)/\varphi(t) \) with \( M = 1 \). We obtain

\[ P(\overline{Y}(t) > \varphi(t), \text{ for some } |t| \leq \delta) \leq P(\overline{X}(t) > \varphi(t), \text{ for some } |t| \leq \delta). \]

Because \( \eta \) is symmetric,

\[ P(|Y(t)| > \varphi(t), \text{ for some } |t| \leq \delta) \leq 2P(|\overline{Y}(t)| > \varphi(t), \text{ for some } |t| \leq \delta) \leq 4P(\overline{Y}(t) > \varphi(t), \text{ for some } |t| \leq \delta). \]

where the second inequality follows from the symmetry of \( \overline{Y} \). By hypothesis (3.4) and the fact that \( f(t) \leq \sqrt{2p_X(t) = o(\varphi(t))}. \)

\[ \limsup \frac{\overline{X}(t)}{\varphi(t)} \leq \limsup \frac{X(t)}{\varphi(t)}. \]

By (5.31) and (3.5), we see that the right and hence the left side of (5.29) tends to zero with \( \delta \). Thus,

\[ \limsup \frac{|Y(t)|}{\varphi(t)} \leq 1 \quad \text{a.s.} \]

Finally, by the zero-one law, (3.5) also holds with the right side replaced by a number less than unity and so equality in (5.32) is impossible. Thus Lemma 3.2 is proved.

REFERENCES


SIXTH BERKELEY SYMPOSIUM: MARCUS AND SHEPP


