GAUSSIAN PROCESSES ON SEVERAL PARAMETERS[1]

BY R. M. DUDLEY

University of California

1. Introduction. In this paper, we study certain Gaussian random processes on multidimensional parameter spaces. The main examples, treated in Sections 6–8, are processes with "independent values" and stationary processes, which are Fourier transforms of each other (Section 6); the processes with nearly independent values which arise in connection with "empirical measures" (Section 7); and various generalizations of Brownian motion (Section 8).

We begin in Section 2 with a general continuity theorem for ordinary Gaussian processes on several parameters. Section 3 is devoted to "random linear functionals", including random Schwartz distributions or "generalized random processes". Sections 4 and 5 give continuity results for certain random linear functionals based on the theorem in Section 2. Sections 5, 6 and 7 are independent of each other.

Some remarks on the continuity theorem in Section 2 seem to be in order here. It asserts that if $\{x(t), t \varepsilon R^k\}$ is a Gaussian process on a k-dimensional Euclidean space such that for some $a > 1$ and $C \geqq 0$,

$$E|x(t+h) - x(t)|^2 \leqq C/|\log |h||^a$$

for all sufficiently small h and all t, then the sample functions $x(t)$ may be taken continuous with probability 1. Yu. K. Belaev [1] proved this for stationary processes with $k = 1$, and showed that for $a \leqq 1$ it need not be true. (It appears that his inequality (53) requires the condition $t_1 + h \leqq t_2$.) His proof is based on Fourier transforms and a theorem of Hunt [9].

My proof is more direct and does not use Fourier transforms, and the result is improved. However, the hypothesis is not invariant under homeomorphisms of R^k, while the conclusion is. This inelegance is avoided in work of V. Strassen, who formulates the hypothesis instead in terms of the "ϵ-entropy" of the set of Gaussian variables $x(t)$ for t in a compact set, which need not lie in a Euclidean space. His result easily implies mine and hence Belaev's; I understand that his proof is quite similar to mine. Section 2 is included here only because Strassen's work has not yet appeared.

2. The continuity theorem. If V is a finite-dimensional linear space over the field F of real or complex numbers, a Borel probability measure μ on V will be called *Gaussian* if it is concentrated either in a point or in a linear space $W \subseteq V$, and in the latter case it is of the form

$$d\mu(x) = ce^{-Q(x-w)}\, d\lambda(x),$$

Received 16 September 1964.

[1] Supported by National Science Foundation Fellowship and Grant GP-2 during part of the work on this paper.

where Q is a positive definite quadratic form on W (i.e., $Q(y) = (y, y)$ for some nondegenerate inner product $(\ , \ )$ on W), $w \varepsilon W$, λ is Lebesgue measure in W, and $c > 0$.

Let R^k be the linear space of k-tuples, $t = (t_1, \cdots, t_k)$, of real numbers, with $|t| = (t_1^2 + \cdots + t_k^2)^{\frac{1}{2}}$. Then by *(Gaussian) process* on R^k I shall mean an assignment to each finite set $F \subset R^k$ of a (Gaussian) probability measure μ_F on a finite-dimensional linear space V_F and real linear functionals $x_F(t)(\ )$, $t \varepsilon F$, on V_F such that if $F \subset G$ finite, the joint distributions of the $x_F(t)(\ )$ and the $x_G(t)(\ )$ for $t \varepsilon F$ are the same.

According to a well-known theorem of Kolmogorov (Loève [16] p. 93), there exists for any process a probability measure ν on the set $\mathfrak{F}$ of all functions from R^k to the real numbers, defined on the σ-field $\mathcal{S}$ generated by all sets of the form

$$S(t^{(1)}, \cdots, t^{(n)}; B) = \{f \varepsilon \mathfrak{F}: \langle f(t^{(1)}), \cdots, f(t^{(n)}) \rangle \varepsilon B\},$$

where $t^{(1)}, \cdots, t^{(n)} \varepsilon R^k$ and B is a Borel set in R^n, with $\nu(S(t^{(1)}, \cdots, t^{(n)}; B)) = \mu_F(\langle x_F(t^{(1)}), \cdots, x_F(t^{(n)}) \rangle \varepsilon B)$, where $F = \{t^{(1)}, \cdots, t^{(n)}\}$. We say the process has *continuous sample functions* if $\nu^*(\mathcal{C}) = 1$, where $\mathcal{C}$ is the set of continuous functions in $\mathfrak{F}$ and ν^* is ν-outer measure. Then ν^* is a countably additive probability measure on the sets $A \cap \mathcal{C}$, $A \varepsilon \mathcal{S}$: this includes the Borel sets for the supremum norm in $\mathcal{C}$ on any fixed bounded set.

We return now to standard notation by letting $\Pr = \nu$, $E =$ integral with respect to Pr, and $x(t) =$ the function $f \to f(t)$ on $\mathfrak{F}$.

THEOREM 2.1. *Suppose given a Gaussian process* $(x(t), \Pr)$ *on a finite-dimensional Euclidean space* R^k *such that for some* $a > 1$, $C > 0$, *and all sufficiently small* h,

$$E[(x(t + h) - x(t))^2] \leqq C/|\log |h||^a$$

for all t. Then the process has continuous sample functions.

PROOF. If y is a Gaussian random variable with mean m, then $E(y^2) = E((m + y - m)^2) = m^2 + E[(y - m)^2]$. Thus if $E(x(t)) = m(t)$,

$$(m(t + h) - m(t))^2 \leqq C/|\log |h||^a$$

for h small enough, so that m is continuous, and

$$E[(x(t + h) - m(t + h) - (x(t) - m(t)))^2] \leqq C/|\log |h||^a.$$

Thus we can assume $m(t) \equiv 0$.

If $\mathcal{C}$ has Pr-outer measure less than 1, it is included in a set G in $\mathcal{S}$ with $\Pr(G) < 1$, and G is in the σ-field generated by countably many sets $\{x(t_n) \varepsilon B_n\}$. Thus it suffices to prove uniform continuity with probability one on an arbitrary countable bounded set. Also, we can restrict ourselves to the cube, $0 \leqq t_j \leqq 1$, $j = 1, \cdots, k$.

For $n = 0, 1, \cdots$, let D_n be the set of dyadic rational numbers, $r/2^n$, $r = 0, 1, \cdots, 2^n$, and let L_n be the "lattice" $D_n \times \cdots \times D_n$ in R^k. Let $M_n = L_{2^n}$.

For each n, we consider the set G_n of Gaussian random variables $x(s) - x(t)$, where $t \varepsilon M_{n+1}$ and s is one of the points of M_n at minimal distance from t. There are at most $2k$ such points, so G_n contains less than $(2^{2^{n+1}} + 1)^k \cdot 2k \leqq 2^{2^{n+3k}}$ random variables.

Since $|s - t| < k/2^{2^{n+1}}$, we have for some $K > 0$,

$$E(z^2) \leqq C/|(2^n + 1) \ln 2 - \ln k|^a \leqq K/2^{na}$$

for any $z \varepsilon G_n$ if n is large enough.

Now if y is any Gaussian random variable with mean 0, $E(y^2) = \sigma^2$ and $b \geqq 1$, then

$$\Pr(|y/\sigma| \geqq b) = [1/(2\pi)^{\frac{1}{2}}] \int_{|x| \geqq b} \exp(-x^2/2)\, dx$$
$$\leqq [2/(2\pi)^{\frac{1}{2}}] \int_b^\infty x \exp(-x^2/2)\, dx \leqq \exp(-b^2/2).$$

Thus for each $z \varepsilon G_n$,

$$\Pr(|z| \geqq 1/n^2) \leqq \exp(-2^{na}/Kn^4) \leqq \exp(-2^{nc})$$

for n large enough, where $c = (1 + a)/2 > 1$. Thus

$$\Pr(\max_{z \varepsilon G_n} |z| \geqq 1/n^2) \leqq 2^{2^{n+3k}} \exp(-2^{nc}),$$

which approaches zero faster than exponentially as $n \to \infty$. Thus

$$\sum_{n=1}^{\infty} \Pr(\max_{z \varepsilon G_n} |z| \geqq 1/n^2)$$

converges, and with probability one, $|z| < 1/n^2$ for all $z \varepsilon G_n$ with $n \geqq r = r(x(\ ))$.

Now this implies that with probability 1, $x(t)$ is uniformly continuous on the union of all the M_n, as follows: if $1 \leqq r \leqq n$ each $p \varepsilon M_n$ is in a cube of side $1/2^{2^r}$ with vertices in M_r, possibly on a face of one or more such cubes.

Given $\epsilon > 0$, let u be an integer greater than $\max(r(x(\ )), 6/\epsilon) + 1$. Now if $p \varepsilon M_m$ and $q \varepsilon M_n$ for some m and n and $|p - q| < 1/2^{2^u}$, then p and q must belong to cubes with vertices in M_u and sides $1/2^{2^u}$ having at least one common vertex v. We can assume $m \geqq u$ and $n \geqq u$.

Now we move from p to a nearest point of M_{m-1}, then to a nearest point of M_{m-2}, etc., until we arrive at $p_u \varepsilon M_u$, and similarly from q to $q_u \varepsilon M_u$. We then have

$$|x(p) - x(p_u)| \leqq 1/u^2 + 1/(u + 1)^2 + \cdots \leqq 1/(u - 1),$$

and $|x(q) - x(q_u)| \leqq 1/(u - 1)$.

Now p_u and q_u are vertices of cubes with a common vertex $v \varepsilon M_u$. There is a point p_{u-1} of M_{u-1} such that both p_u and v are nearest to p_{u-1}, and a $q_{u-1} \varepsilon M_{u-1}$ nearest to both v and q_u. Hence $|x(p_u) - x(v)| \leqq 2/(u - 1)^2$ and $|x(q_u) - x(v)| \leqq 2/(u - 1)^2$. Hence $|x(p) - x(q)| \leqq 4/(u - 1)^2 + 2/(u - 1) \leqq 6/(u - 1) < \epsilon$.

Thus, indeed, we have uniform continuity with probability 1 on the union M of all M_n. For any $\epsilon > 0$ and $\delta > 0$, let $A_{\delta\epsilon}$ be the set of all $x(\ )$ such that $|x(s) - x(t)| \geqq \epsilon$ for some s and t in M with $|s - t| < \delta$. Then for any $\epsilon > 0$, there is a $\delta > 0$ such that $\Pr(A_{\delta\epsilon}) < \epsilon$. Now if Q is any countable subset of the unit cube, and $s, t \varepsilon Q$, $|s - t| < \delta$, then since $x(\ )$ is continuous in probability and M is dense in the unit cube, $|x(s) - x(t)| \leqq 2\epsilon$ for almost all $x(\ )$ not in $A_{\delta\epsilon}$. Thus for almost all $x(\ )$ not in $A_{\delta\epsilon}$, $|x(s) - x(t)| \leqq 2\epsilon$ for *all* $s, t \varepsilon Q$ with $|s - t| < \delta$, since there are only countably many such pairs. Thus $x(\ )$ is uniformly continuous with probability one on Q, and the proof is complete.

Our method applies to processes defined only on certain subsets of a given finite dimensional linear space; for example, certain countable unions of closed cubes (e.g. open subsets). Arbitrary subsets and more general spaces, with a suitable modification of the statement, are covered by the more general theorem of V. Strassen (as yet unpublished).

The extension of Theorem 2.1 to processes with values in any finite-dimensional real or complex linear space is trivial. Possibly it can also be extended to processes with values in locally compact groups, Gaussian in the sense of Urbanik [22].

If we weaken the conclusion of Theorem 1 by changing "$a > 1$" to "$a > 3$", the proof becomes somewhat easier, since one can use the lattices L_n rather than M_n. The theorem for $a > 3$ is easily strong enough to cover all the applications below except in Section 5. The simpler proof for $a > 3$ is essentially that of P. Lévy [15] for the continuity of his "Brownian motion on several parameters", which will be discussed further in Section 8.

3. Random linear functionals and random distributions. In probability theory, one very often has to deal with probability measures on a topological linear space T, possibly infinite-dimensional. If T is finite-dimensional, there is a 1-1 correspondence between Borel probability measures and consistent assignments of joint distributions to finite sets of linear functionals on T, belonging to a set S sufficient to separate points of T. In general, however, such an assignment need not define a countably additive measure on T, and becomes an independent object of study. The space S becomes an arbitrary linear space, and T becomes irrelevant. I will sketch the elementary theory of such assignments, or "random linear functionals".

Let S be a linear space over the field F of real or complex numbers, and let S^* be its (algebraic) dual space. For $x_1, \cdots, x_n \varepsilon S$ and B, a Borel set in F^n, let

$$S^*(s_1, \cdots, x_n; B) = \{y \varepsilon S^*: \langle y(x_1), \cdots, y(x_n) \rangle \varepsilon B\}.$$

Now a *random linear functional* (rlf) on S is either

(A) an equivalence-class of maps L from S into measurable functions on a probability space $(\Omega, \mathcal{S}, P)$, such that if $a, b \varepsilon F$ and $f, g \varepsilon S$, $L(af + bg) = aL(f) + bL(g)$ almost everywhere with respect to P; where L and M are equivalent if for any $x_1, \cdots, x_n \varepsilon S$,

$$\langle L(x_1), \cdots, L(x_n) \rangle \quad \text{and} \quad \langle M(x_1), \cdots, M(x_n) \rangle$$

have the same probability distribution on F^n;

or

(B) an assignment of a probability measure $\mu(x_1, \cdots, x_n)(\ )$ on the Borel sets of F^n to each ordered finite set $\langle x_1, \cdots, x_n \rangle$ of elements of S, such that $S^*(x_1, \cdots, x_m; B) = S^*(y_1, \cdots, y_n; C)$ implies

$$\mu(x_1, \cdots, x_m)(B) = \mu(y_1, \cdots, y_n)(C);$$

or

(C) a probability measure P in S^* on the σ-field generated by all sets $S^*(x_1, \cdots, x_n; C)$.

There are natural 1-1 correspondences between the objects defined as rlf's in (A), (B) and (C), as follows: first given L as in (A), let

$$\mu(x_1, \cdots, x_n)(B) = P(\langle L(x_1), \cdots, L(x_n) \rangle \varepsilon B)$$

for any Borel set $B \subset F^n$. The almost-everywhere linearity of L implies easily that $\mu(\ )(\ )$ is an rlf in sense (B), and clearly μ depends only on the equivalence-class of L.

Given an rlf in sense (B), there is a probability measure P on S^* such that

$$P(S^*(x_1, \cdots, x_n; B)) \equiv \mu(x_1, \cdots, x_n)(B);$$

one can prove this by choosing a Hamel basis in S and applying Kolmogorov's theorem on probability measures in product spaces (Loève [16], p. 93 A), or its generalization, Bochner's theorem on inverse limit measures (see Dudley [3], Theorem 1).

Finally, an rlf in sense (C) yields one in sense (A) where $\Omega = S^*$, $P = P$, $\mathcal{S} = \sigma$-field generated by the sets $S^*(x_1, \cdots, x_n; B)$, and $L(x)$ is the map $y \rightarrow y(x)$ of S^* into F for each $x \varepsilon S$.

Below, sense (A) will generally be used, and a particular L will be chosen, leading to the helpful abuse of language, "an rlf $L \cdots$."

If a sense of sequential convergence is defined in S, e.g. by a topology, an rlf L on S is called *continuous* if $x_n \rightarrow x$ implies $L(x_n) \rightarrow L(x)$ in probability (or, using (B),

$$\textstyle\int_F f(t)\mu(x_n)(dt) \rightarrow \int_F f(t)\mu(x)(dt)$$

for each bounded continuous real-valued function f on F).

If S' is a linear subspace of S^*, sufficiently large to separate points of S, let

$$S'(x_1, \cdots, x_n; B) = S' \cap S^*(x_1, \cdots, x_n; B).$$

An rlf P on S^* (sense (C)) defines a "semimeasure" Q on S' by

$$Q(S'(x_1, \cdots, x_n; B)) = P(S^*(x_1, \cdots, x_n; B)).$$

The separating property implies that Q is well-defined, but it need not be countably additive on S'. Clearly it is countably additive if and only if S' has P-outer measure 1 in S^*. If S has a topology, S' is its topological dual space (the continuous linear functionals), and L is an rlf on S such that Q is countably additive, I shall say L "is a measure on S'" or "is countably additive on S'."

Clearly any rlf on a finite-dimensional space S is a measure on S' for S' separating, since then $S' = S^*$. The standard example of an rlf which is not a measure is the "normal distribution", defined as follows. Let $(H, (\ ,\ ))$ be a separable, infinite-dimensional Hilbert space over F. Then the normal rlf L is Gaussian (i.e. each $\mu(f_1, \cdots, f_n)(\ )$ is Gaussian) and $E(L(f)) = 0$, $E(L(f)\overline{L(g)}) = (f, g)$ for all $f, g \in H$. To prove that such an L exists, we can take an orthonormal basis $\{f_n\}$ in H and let $L(f_n)$ be independent, normalized Gaussian random variables (that is, the density of any $\mu(f_1, \cdots, f_n)(\ )$ with respect to Lebesgue measure in F^n is a constant times $\exp(-\frac{1}{2}\sum_{j=1}^{n} |t_j|^2))$. L is not a measure on the topological dual space H' since $\sum_{n=1}^{\infty} |L(f_n)|^2$ diverges with probability 1.

An rlf L on a Hilbert space H is a measure on H' if and only if it can be written $L(f) = M(A(f))$, where M is continuous rlf and A is a Hilbert-Schmidt operator (Kolmogorov [14], Minlos [17], Gross [8]).

An rlf L on a linear space S will be said to have second moments if $E(|L(f)|^2)$ is finite for all $f \in S$; this implies that $B(f, g) = E(L(f)\overline{L(g)})$ is also finite for all $f, g \in S$. B is clearly linear in f and conjugate-linear in g, i.e. "sesquilinear", and conjugate-symmetric.

Now if S is a metrizable linear space of second category and L is a continuous rlf on S, $f_n \to f$ implies $B(f, f) \leqq \liminf B(f_n, f_n)$ by Fatou's lemma. Thus for any $M > 0$ the set of $f \in S$ such that $B(f, f) \leqq M$ is closed. Thus, by a category argument, $f \to B(f, f)$ is continuous at zero in S. It follows that for each $g \in S$, $f \to B(f, g)$ is continuous on S.

If S is an LF-space, i.e. it is a union of locally convex complete metric linear spaces S_n, and S has the strongest locally convex topology such that each injection $S_n \to S$ is continuous, then it clearly remains true that $f \to B(f, g)$ is continuous if L is continuous (in probability).

An important special case of rlf's is the *random distributions*, which are (sequentially) continuous rlf's on the linear space $\mathfrak{D}(R^k)$ of C^∞ functions with compact support with its usual convergence, i.e. $f_n \to f$ if all f_n vanish outside a fixed compact set and $D^p f \to D^p f$ uniformly for each partial derivation

$$D^p = \partial^{p_1+\cdots+p_k}/\partial x_1^{p_1} \cdots \partial x_k^{p_k}.$$

Incidentally, $\mathfrak{D}$ is an LF-space (Schwartz [19], Chapter III, Section 1, Theorem II, p. 66) for a topology which yields the above sequential convergence.

Another property of the topology of $\mathfrak{D}$, its "nuclearity", yields the pleasant fact that every random distribution is a measure on $\mathfrak{D}'$ (Gelfand and Vilenkin [7], Chapter IV, Section 2, No. 4, p. 407). We call such a probability measure on $\mathfrak{D}'$ the "Minlos measure" of the associated random distribution.

For any random distribution L with second moments, $f \to E(L(f)\overline{L(g)})$ is a continuous linear functional $T(g)$ on $\mathfrak{D}$, i.e. a distribution. The map $g \to T(\bar{g})$ is linear from $\mathfrak{D}$ to $\mathfrak{D}'$. If $g_n \to g$ in $\mathfrak{D}$, $T(\bar{g}_n) \to T(\bar{g})$ in the weak* topology of $\mathfrak{D}'$, hence the strong topology (Schwartz [19] Chapter III, Section 3, Theorem XIII, p. 74). Thus $g \to T(\bar{g})$ is a sequentially continuous linear map from $\mathfrak{D}$ to $\mathfrak{D}'$. Since both spaces are bornologic, it is continuous (Dudley [2], Theorems 6.1 and 6.3). Now by the "theorem of kernels" (Schwartz [20], I Proposition 25,

p. 93 or [21], Section 4), there is a distribution $K \varepsilon \mathfrak{D}'(R^k \times R^k)$ such that $B(f, g) = K(f \otimes \bar{g})$, where $(f \otimes \bar{g})(x, y) = f(x)\bar{g}(y)$, $x, y \varepsilon R^k$.

One can show likewise that if L is a random distribution with values in a finite-dimensional vector space V and $E(|\lambda(L(f))|^2) < \infty$ for each $\lambda \varepsilon V'$, then for any $\rho, \sigma \varepsilon V'$ there is a distribution $K_{\rho,\sigma}$ with

$$E(\rho(L(f))\overline{\sigma(L(g))}) = K_{\rho,\sigma}(f \otimes \bar{g}).$$

4. Continuity on blocks. This section is devoted to continuity results for Gaussian random distributions which are rather special, but sufficient to cover the cases in Sections 6–8.

Let μ be a Borel measure on R^k, finite on compact sets. Suppose L is a random distribution with second moments for which

$$E(|L(f)|^2) \leqq \int |f|^2 \, d\mu$$

for any $f \varepsilon \mathfrak{D}$. Then, clearly, L extends uniquely to a continuous rlf on $H = L^2(R^k, \mu)$. In any case, L has a Minlos measure P on $\mathfrak{D}'(R^k)$. I want to show that if L is Gaussian, P is concentrated in the set of Schwartz derivatives

$$\partial^k f / \partial t_1 \cdots \partial t_k = D^k f$$

where f is a function belonging to the space Q_μ to be defined below.

If $u = (u_1, \cdots, u_k) \varepsilon R^k$, we say $s \rightarrow u$ "in an octant" if for each j, $s_j \rightarrow u_j$ and either $s_j \leqq u_j$ or $s_j > u_j$.

Definition. Given a Borel measure μ on R^k, finite on compact sets, Q_μ is the set of complex-valued functions f on R^k with the following properties:

(1) for each $u \varepsilon R^k$, $f(s)$ approaches a limit as $s \rightarrow u$ in any octant;

(2) the limit as $s_j \uparrow u_j$ for all j is $f(u)$;

(3) given u and j, the limit is independent of whether $s_j \uparrow u_j$ or $s_j \downarrow u_j$ unless the hyperplane $t_j = u_j$ has positive μ-measure;

(4) if μ has finite total mass, f is defined on the product of extended real lines $[-\infty, \infty]$ with the same continuity properties.

Now if $f \varepsilon Q_\mu$, f is continuous except on at most countably many hyperplanes. Also f is bounded on each bounded set, since if it is unbounded in every neighborhood of a point u this will remain true in some octant. f is clearly Lebesgue measurable. If μ gives all hyperplanes parallel to the axes measure zero, Q_μ is simply the set of all continuous complex-valued functions on R^k.

For any $t \varepsilon R^k$, let A_t be the indicator function of the block $B(t)$ of all s such that

$$0 \leqq s_j < t_j \quad \text{if} \quad t_j \geqq 0,$$

$$t_j \leqq s_j < 0 \quad \text{if} \quad t_j \leqq 0$$

for $j = 1, \cdots, k$. If L is a (Gaussian) rlf on a space containing the functions A_t we define a (Gaussian) process $x(\ )$ on R^k by

$$x(t) = (-1)^{\text{neg}(t)} L(A_t),$$

where neg (t) is the number of negative t_j. Of course, $x(t) = 0$ if any t_j is zero since then $B(t)$ is the null set. We call $x(\ )$ the *indefinite integral* of L.

First we show that if $x(\ )$ is not too discontinuous, then $L = \partial^k x/\partial t_1 \cdots \partial t_k$ (Theorem 4.1). Then we show that for $E(|L(f)|^2) \leqq \int |f|^2\, d\mu$, $x(t)$ is sufficiently continuous (Theorem 4.2).

Let $\mathcal{R}$ be the class of functions on R^k which are continuous almost everywhere with respect to Lebesgue measure and bounded on bounded sets. Let $D^k = \partial^k/\partial t_1 \cdots \partial t_k$. Then D^k as a Schwartz derivative maps $\mathcal{R} \subset \mathfrak{D}'$ into $\mathfrak{D}'$. The result of applying a distribution T to a function f will be written $[T](f)$.

THEOREM 4.1. *Suppose L is an* rlf *on a space of complex-valued functions on R^k which includes $\mathfrak{D}$ and all functions A_t, and that if f_n all vanish outside a fixed compact set and $f_n \to f$ uniformly, then $L(f_n) \to L(f)$ in probability, on the probability space $(\Omega, \mathcal{S}, P_1)$ of L. Also suppose that the Kolmogorov measure P_x for the process $x(t) = (-1)^{\text{neg}(t)} L(A_t)$ gives $\mathcal{R}$ outer measure 1. Let Q_x be P_x confined to $\mathcal{R}$. Then there is a map $\xi\colon \omega \to \xi_\omega(\ )$ of Ω into $\mathcal{R}$ such that for each $f \in \mathfrak{D}$,*

$$L(f)(\omega) = [D^k \xi_\omega](f)$$

with P_1-probability one. Thus the Minlos measure P of L restricted to $\mathfrak{D}$ is equal to $Q_x \circ (D^k)^{-1}$ on all sets where P is defined.

PROOF. Let E be the set of k-tuples of dyadic rational numbers $r/2^m$, r, m integers. Let $\eta_\omega(t) = x(t)(\omega)$, $t \in E$. For $t \notin E$, let

$$\eta_\omega(t) = \lim_{s \to t, s \in E} \eta_\omega(s)$$

if the limit exists, and $\eta_\omega(t) = 0$ otherwise. By our assumption on P_x, $\eta_\omega \in \mathcal{R}$ with probability one. Let $\xi_\omega = \eta_\omega$ for $\eta_\omega \in \mathcal{R}$, otherwise $\xi_\omega \equiv 0$.

P is defined on the σ-field generated by the ring of all sets

$$\{T \in \mathfrak{D}' : \langle T(f_1), \cdots, T(f_n) \rangle \in B\}$$

for $f_1, \cdots, f_n \in \mathfrak{D}$ and B a Borel set in complex n-space. The inverse image of such a set under D^k is clearly Q_x-measurable in $\mathcal{R}$.

Given $f \in \mathfrak{D}$, let

$$f_n(t_1, \cdots, t_k) = f((2r_1 + 1)/2^{n+1}, \cdots, (2r_k + 1)/2^{n+1}) \qquad \text{for} \quad r_i/2^n \leqq t_i < (r_i + 1)/2^n,$$

where $r_1, \cdots, r_k$ are any integers, $n = 1, 2, \cdots$. For any $\delta > 0$ and any complex-valued function g on R^k let

$$(D_\delta^k g)(t_1, \cdots, t_n) = \textstyle\sum_{s_i = \pm 1} \left(\prod_{i=1}^k s_i\right) g(t_1 + s_1\delta, \cdots, t_k + s_k\delta).$$

For each n, the distribution $D^k f_n$ is a complex-valued measure concentrated in finitely many points of the form

$$q = (r_1/2^n, \cdots, r_k/2^n),$$

and assigning such a point the measure $(D_\delta^k f)(q)$ where $\delta = 1/2^{n+1}$.

It is easy to verify that

$$(D_\delta^k f)(t_1, \cdots, t_n) = \int_{t_1-\delta}^{t_1+\delta} \cdots \int_{t_n-\delta}^{t_n+\delta} (D^k f)(x_1, \cdots, x_n)\, dx_n \cdots dx_1 .$$

Hence $H_\delta = D_\delta^k f/(2\delta)^k - D^k f \to 0$ as $\delta \to 0$. Since $f \varepsilon \mathfrak{D}$, there is a $K > 0$ such that

$$|(D^k f)(s) - (D^k f)(t)| \leqq K\, |s - t|$$

for all $s, t \varepsilon R^k$. Thus H_δ is bounded uniformly in $\delta > 0$. Hence for some $M > 0$,

$$|(D_\delta^k f)(t)| \leqq M\delta^k$$

for all t and $\delta > 0$. Now if R is the diameter of the support of f, the total variation of $D^k f_n$ is at most $M[(R + 2)2^n]^k/2^{k(n+1)}$, which is bounded uniformly in n.

Since $f_n \to f$ in $\mathfrak{D}'$, $D^k f_n \to D^k f$ in $\mathfrak{D}'$. Since $\mathfrak{D}$ is dense in the continuous functions on a neighborhood of the support of f with the supremum norm,

$$\int g\, d(D^k f_n) \to \int g D^k f\, dt$$

for every continuous function g on R^k, and thus for every g bounded and continuous almost everywhere with respect to $|D^k f|\, dt$ (Prokhorov [18], Theorem 1.8), hence for every $g \varepsilon \mathfrak{R}$.

We call a product of left closed, right open intervals, a *block*. The A_t are indicator functions of blocks with a vertex at the origin. By definition of $x(\ )$,

$$\int x(s)\, d(D^k A_t)(s) = (-1)^{k-\operatorname{neg}(t)} x(t) = (-1)^k L(A_t),$$

since $x(t) = 0$ if $t_j = 0$ for some j.

Any indicator function h of a block is a finite linear combination of functions A_t, so $\int x(s)\, d(D^k h)(s) = (-1)^k L(h)$ with probability 1. This also holds if h is a finite linear combination of indicators of blocks, e.g. $h = f_n$. By definition of ξ_ω, we have with probability one

$$\int \xi_\omega(t)\, d(D^k f_n)(t) = \int x(t)(\omega)\, d(D^k f_n)(t),$$

and as $n \to \infty$, $L(f_n) \to L(f)$ in probability. Thus

$$L(f) = (-1)^k \int \xi_\omega\, d(D^k f) = [D^k \xi_\omega](f)$$

with probability one for any $f \varepsilon \mathfrak{D}$. Clearly $P_1 \circ \xi^{-1} = Q_x$, so $P = Q_x \circ (D^k)^{-1}$, q.e.d.

THEOREM 4.2. *If μ is a Borel measure on R^k, finite on bounded sets, and L is a* Gaussian rlf *on $H = L^2(R^k, \mu)$ such that*

$$E\, |L(f)|^2 \leqq \int |f|^2\, d\mu$$

then Q_μ has outer measure 1 for the Kolmogorov measure of $x(\ )$, the indefinite integral of L.

PROOF. It suffices to show that, given $M > 0$, the desired monotone continuity conditions hold on the cube $C: 0 \leqq t_j < M, j = 1, \cdots, k$, since $x(t) = 0$ if any t_j is zero.

We spread μ on C into a continuous measure, as follows: for $j = 1, \cdots, k$, $0 < a < M$, let $f_j(a)$ be the sum of the μ-measures of all those $(k - 1)$-dimensional cubes, $\{t \varepsilon C: t_j = b\}$, which have positive μ-measure and $b < a$. Clearly each f_j is a bounded function. Let $F_j(s) = s + f_j(s)$, $j = 1, \cdots, k$, and let $F(t) = \langle F_1(t_1), \cdots, F_k(t_k) \rangle$. Then F takes C 1-1 into a larger rectangular solid D, taking hyperplanes $t_j =$ constant into hyperplanes of the same form. We define a continuous map $G = (G_1, \cdots, G_k)$ of D onto C as the inverse of F where it is defined, with

$$G_j(s_1, \cdots, s_k) = t_j \qquad \text{for} \quad F_j(t_j) \leqq s_j \leqq F_j(t_j^{+}).$$

We define a measure ν on D by the following construction. Let J be any subset of $\{1, \cdots, k\}$, say of cardinality n. Let N_J be the set of n-tuples $\{\tau_j, j \varepsilon J\}$ such that f_j is discontinuous at τ_j for each $j \varepsilon J$. Then $N_J = \prod_{j\varepsilon J} N_{\{j\}}$. Now for $\alpha \varepsilon N_J$ let $\mu(J; \alpha)$ be μ restricted to the set of $t \varepsilon C$ where $\{t_j, j \varepsilon J\} = \alpha$ and f_j is continuous at t_j for $j \not\varepsilon J$. Then for $j \not\varepsilon J$, no measure $\mu(J; \alpha)$ gives positive mass to any hyperplane $t_j =$ constant.

Let $P_J(t) = \{t_j, j \not\varepsilon J\}$ and let $\nu(J; \alpha)$ be the product measure, $[\mu(J; \alpha) \circ F^{-1} \circ P_J^{-1}] \times \lambda(J; \alpha)$, where $\lambda(J; \alpha)$ is Lebesgue measure, normalized to total mass 1, in the rectangular solid, $F_j(G_j(\alpha_j)) < s_j < F_j(G_j(\alpha_j)^{+})$, $j \varepsilon J$ in an n-dimensional space. Of course $\mu(J; \alpha) \circ F^{-1} \circ P_J^{-1}$ is a measure on a $(k - n)$-dimensional space, and F and P_J commute. Let

$$\nu = \sum \nu(J; \alpha)$$

where the sum is over all sets $J \subset \{1, \cdots, k\}$ and $\alpha \varepsilon N_J$. Then ν is a measure on D. It gives measure zero to any hyperplane $s_j =$ constant since each $\nu(J; \alpha)$ does.

For each J and $\alpha \varepsilon N_J$,

$$\nu(J; \alpha) \circ G^{-1} = \mu(J; \alpha) \circ P_J^{-1} \times \delta_\alpha = \mu(J; \alpha)$$

where δ_α is the unit mass at α in n-space. μ restricted to C, or μ_C, is the sum of the $\mu(J; \alpha)$ over all J and $\alpha \varepsilon N_J$, since $\sum_{\alpha\varepsilon N_J} \mu(J; \alpha)$ is μ_C restricted to the set where f_j is continuous precisely for $j \not\varepsilon J$, and these sets are disjoint with union C.

Thus $\nu \circ G^{-1} = \mu_C$.

Now let M be the rlf on $L^2(D, \nu)$ defined by

$$M(f) = L(f \circ G).$$

Then $E\,|M(f)|^2 \leqq \int |f|^2\,d\nu$ for all $f \varepsilon L^2(D, \nu)$. Letting y be the indefinite integral of M, we have

$$y(s) = x(G(s)), \quad s \varepsilon D.$$

For $j = 1, \cdots, k$ let $H_j(b) = b + \nu\{s \varepsilon D: s_j < b\}$. Then H_j is continuous and strictly increasing for $0 \leqq b < f_j(M)$. Letting $z(H_1(s_1), \cdots, H_k(s_k)) = y(s_1, \cdots s_k)$, we obtain a Gaussian process z on a rectangular solid with

$$E\,|z(u) - z(v)|^2 \leqq k^2\,|u - v|,$$

so that z is almost surely continuous by Theorem 2.1 and hence so is y. Since F has the desired continuity properties on C, so does x.

If μ has finite total mass, we can let $M = +\infty$ and apply the same argument, essentially. The proof is complete.

5. Processes on spaces of polyhedra. A set which is the convex hull of some $k + 1$ points of R^k will be called a $(k + 1)$-*set*. (If it has non-empty interior, it is a simplex in the usual sense.)

A sequence C_n of $(k + 1)$-sets will be called *convergent* to another such set C_0 if for some choice and ordering of $(k + 1)$ points $C_{n,1}, \cdots, C_{n,k+1}$ whose convex hull is C_n for each n, $C_{n,j} \to C_{0,j}$ for each j. The set of $(k + 1)$-tuples of points of R^k is a $k(k + 1)$-dimensional linear space W. The set S_k of all $(k + 1)$-sets in R^k is the factor space of W by an equivalence relation, but rather than trying to take advantage of this I shall simply treat functions on S_k as functions on W.

Now suppose given a Gaussian rlf L on some linear space of real- or complex-valued functions on R^k which contains all indicator functions of $(k + 1)$-sets. Then Theorem 2.1 on W implies almost sure continuity of L on S_k if the following holds:

(∗) If K is any bounded subset of W, there is a $C > 0$ and an $a > 1$ such that if x and y are any two points of K and f and g are the indicator functions of the corresponding simplices,

$$E\,|L(f) - L(g)|^2 \leqq C/|\log|x - y||^a.$$

To obtain (∗) in some cases we have

THEOREM 5.1. *Suppose L is a random distribution on R^k such that if ν is Lebesgue measure on R^k and M is any bounded subset of R^k, there is a $C_1 > 0$ and $a > 1$ such that*

$$E(|L(f)|^2) \leqq C_1/|\log \nu(U)|^a$$

whenever $f \varepsilon \mathfrak{D}$ vanishes outside the open set $U \subset M$ and $|f| \leqq 1$. Then for each M, L extends by continuity to an rlf *on $L^2(M, \nu)$ for which* (∗) *holds.*

PROOF. The extension to an rlf on $L^2(M, \nu)$, for which we may assume M open, is immediate, and we then have the inequality in the hypothesis for any $f \varepsilon L^2(M, \nu)$. It then suffices to show that there is an $m > 0$ such that whenever $x, y, \varepsilon M$ and E, F are the corresponding simplices,

$$\nu(E\Delta F) \leqq m\,|x - y|,$$

where $E\Delta F$ is the symmetric difference

$$(E \sim F) \cup (F \sim E).$$

Let N be the supremum of the $(k - 1)$-dimensional "surface areas" of $(k + 1)$-sets defined by points of M. Then clearly we can take $m = 2N$, and the proof is complete.

The hypothesis of Theorem 5.1 is satisfied if L is a noise or centered noise process whose spectral measure μ is absolutely continuous with respect to Lebesgue measure ν and satisfies the rather mild additional condition that

$$\mu(B) \leqq C_1/|\log \nu(B)|^a$$

where $C_1 > 0$ and $a > 1$ depend only on the choice of a bounded set including the Borel (or open) set B.

The method of reparametrization is unavailable here since only linear transformations will preserve $(k+1)$-sets. Thus the full force of Theorem 2.1 seems to be used.

Continuity on $(k+1)$-sets implies continuity on polyhedra with any fixed number of vertices. It is not only considerably stronger than continuity on rectangular solids parallel to a given set of axes, but invariant under linear transformations. It would be desirable to conclude from continuity on $(k+1)$-sets that $f \to L(f)$ is almost surely continuous on some function space larger than what we have from Section 4 (say $\partial^k f/\partial t_1 \cdots \partial t_k$ is integrable and f has compact support). However, I have no such result at present.

6. Noise processes and stationary Gaussian processes. Let μ be a nonnegative, σ-finite measure on a space X, and H the separable complex Hilbert space $L^2(X, \mu)$. We assume H is infinite-dimensional, i.e., that μ is not concentrated in finitely many atoms.

Let L be the normal rlf on H. Then L has "independent values," in the sense that if f and g in H have disjoint supports ($fg \equiv 0$), then $L(f)$ and $L(g)$ are independent. If X is a Euclidean space and μ is Lebesgue measure, L is called the "white noise" rlf. In general, we say L is the *noise* rlf with *spectral measure* μ.

Now suppose μ is a Borel measure on R^k, finite on bounded sets. Then L restricted to $\mathfrak{D}$ is a random distribution with independent values, as defined by Gelfand and Vilenkin ([7], Chapter III, Section 4).

I shall now describe how noise rlf's arise as the Fourier transforms of stationary Gaussian rlf's. If f is a function on R^k and $x, y \in R^k$, let $f_y(x) = f(x - y)$. A random distribution L on R^k is called *stationary* if for any $f^{(1)}, \cdots, f^{(n)} \in \mathfrak{D}$, $y \in R^k$, the joint probability laws of $L(f^{(1)}), \cdots, L(f^{(n)})$ and $L(f_y^{(1)}, \cdots f_y^{(n)})$ are the same (see Ito [12]).

Suppose L has values in a finite-dimensional complex linear space V, and has second moments. Then, as mentioned in Section 4, for any two linear functionals ρ and σ on V there is a distribution $B_{\rho\sigma}$ on R^{2k} such that

$$E(\rho(L(f))\overline{\sigma(L(g))}) = B_{\rho\sigma}(f \otimes \bar{g}).$$

Now stationarity implies that for any $z \in R^k$, the transformation

$$\langle x, y\rangle \to \langle x + z, y + z\rangle, \qquad x, y \in R^k,$$

of R^{2k} into itself leaves $B_{\rho\sigma}$ invariant. Hence there is a distribution $C_{\rho\sigma}$ on R^k such that for any $h \in \mathfrak{D}(R^{2k})$,

$$B_{\rho\sigma}(h) = [C_{\rho\sigma}]_{(y)}[\textstyle\int_{R^k} h(x + y, x)\, dx]$$

where dx is a Lebesgue measure in R^k (see Schwartz [19], tome I, Chapter II, Section 5, formula 10, p. 57); apply the formula k times).

Now for $\rho = \sigma$ we have for any $f \in \mathfrak{D}(R^k)$,

$$0 \leqq E(|\sigma(L(f)|^2) = B_{\sigma\sigma}(f \otimes \bar{f})$$
$$= C_{\sigma\sigma(y)}[\int_{R^k} f(x+y)\bar{f}(x)\,dx].$$

Thus $C_{\sigma\sigma}$ is of positive type, and according to the generalized Bochner theorem (Schwartz [19], tome II, Chapter VII, Section 9, Theorem XVIII, p. 132) $C_{\sigma\sigma}$ is the Fourier transform of a non-negative tempered measure $\mu_{\sigma\sigma}$ so that for any $f, g \,\varepsilon\, \mathfrak{D}(R^k)$,

$$E(\sigma(L(f))\overline{\sigma(L(g))}) = \int (\int f(x+y)\bar{g}(x)\,dx)^\sim(\lambda)\,d\mu_{\sigma\sigma}(\lambda)$$

where for any Lebesgue integrable function F on R^k, $\tilde{F}(\lambda) = \int_{R^k} e^{-2\pi i(\lambda, y)}F(y)\,dy$, $\lambda \,\varepsilon\, R^k$, and $(\lambda, y) = \sum_{j=1}^k \lambda_j y_j$. Now letting

$$(\Phi * \Psi)(y) = \int_{R^k} \Phi(y-x)\Psi(x)\,dx \quad \text{and} \quad \check{\Psi}(x) = \Psi(-x)$$

for any $\Phi, \Psi \,\varepsilon\, \mathfrak{D}$, we have

$$B_{\sigma\sigma}(f \otimes \bar{g}) = \int (f*(\bar{g})^\vee)^\sim(\lambda)\,d\mu_{\sigma\sigma}(\lambda)$$
$$= \int \tilde{f}(\lambda)(\bar{g})^-(\lambda)\,d\mu_{\sigma\sigma}(\lambda).$$

It is easy to check that for any ρ and σ,

$$4B_{\rho\sigma} = B_{\rho+\sigma,\rho+\sigma} - B_{\rho-\sigma,\rho-\sigma} + iB_{\rho+i\sigma,\rho+i\sigma} - iB_{\rho-i\sigma,\rho-i\sigma}.$$

Thus for any ρ and σ there is a complex-valued measure $\mu_{\rho\sigma}$ with

$$B_{\rho\sigma}(f \otimes \bar{g}) = \int \tilde{f}(\bar{g})^-\,d\mu_{\rho\sigma}(\lambda).$$

We have for any $f \,\varepsilon\, \mathfrak{D}$ and linear functional σ on V,

$$E(\sigma(L(f))) = \sigma(v)\int_{R^k} f(x)\,dx$$

for some fixed $v \,\varepsilon\, V$. Letting $M = L - v$ we obtain a stationary random distribution with second moments and $E(\sigma(M(f))) = 0$ for any $f \,\varepsilon\, \mathfrak{D}$ and linear functional σ on V.

Now if V is one-dimensional, let $\sigma(t) = t$; then for some measure μ on R^k, finite on bounded sets,

$$E(M(f)\overline{M(g)}) = \int \tilde{f}(\bar{g})^-\,d\mu$$

and $E(M(f)) = 0$ for all $f, g \,\varepsilon\, \mathfrak{D}$. Thus we can say that the Fourier transform of M is the noise process with spectral measure μ.

For $k = 1$ and V one-dimensional the representation

$$E(L(f)\overline{L(g)}) = \int \tilde{f}(\bar{g})^-\,d\mu$$

for the covariance of a random distribution when it is translation-invariant was first proved by K. Ito [12]; for $k > 1$ and V multidimensional the representation of $B_{\rho\sigma}$ in the stationary case was found by Ito [13] and Yaglom [24], and has been used by Urbanik [23] and probably others.

Now, for any tempered nonnegative measure μ on R^k there is a stationary Gaussian random distribuion L with

$$E(L(f)) = 0,$$
$$E(L(f)\overline{L(g)}) = \int \tilde{f}(\tilde{g})^{-} \, d\mu$$

for all $f, g \in \mathfrak{D}$. Let N be the noise rlf with spectral measure μ. Then my result [4] implies that the Minlos measure for N is concentrated in the set of distributions $D^k g$ where g is measurable and $\int_B |g|^p \, dx$ is finite for any bounded set B and $0 \leqq p < \infty$. Theorems 4.1 and 4.2 yield a better result in this case, namely that g can be taken in Q_μ, so that it is bounded on bounded sets.

7. Centered noise processes. Another class of examples, closely related to noise rlf's, is as follows:

DEFINITION. If $(X, \mathcal{S}, \mu)$ is a probability space, the *centered noise* rlf L on $H = L^2(X, \mathcal{S}, \mu)$ is a real Gaussian rlf L such that for any $f, g \in H$, $E(L(f)) = 0$ and

$$E(L(f)L(g)) = \int (f - m(f))(g - m(g)) \, d\mu,$$

where $m(h) = \int h \, d\mu$ (defined for any $h \in H$ since μ is a probability measure).

To show that such an L exists, we take an orthonormal basis $\{f_n\}$ of H where f_1 is the constant function 1, and let $L(f_1) = 0$, while for $n > 1$, $L(f_n)$ are independent, normalized Gaussian random variables; then we extend by linearity and continuity. It is clear that, roughly speaking, a centered noise rlf is obtained from the (complex) noise rlf with the same μ by taking real parts, multiplying by $2^{\frac{1}{2}}$, and imposing the restriction $L(1) = 0$.

If X is R^k for some k and $\mathcal{S}$ is the Borel field, a centered noise rlf can be treated, like noise rlf's in the preceding section, as a random distribution, i.e. restricted to $\mathfrak{D} \subset H$. Then by Theorems 4.1 and 4.2 the Minlos measure for L is concentrated in the set of distributions of the form $\partial^k f / \partial t_1 \cdots \partial t_k$, where $f \in Q_\mu$. I don't know whether this representation can be improved using Theorem 5.1 under suitable assumptions on μ.

Centered noise rlf's arise in connection with "empirical measures" as follows: let $x_1, x_2, \cdots$, be independent X-valued random variables with distribution μ, i.e. we construct a countably infinite product Ω of spaces isomorphic to $(X, \mathcal{S}, \mu)$ with x_i as coordinate functions; let the product measure be Pr. Now if f is an integrable function on X, the $f(x_i)$ are independent, integrable random variables with the same distribution, so by Kolmogorov's strong law of large numbers (Loève [16], p. 239) $(f(x_1) + \cdots + f(x_n))/n - \int f \, d\mu$ converges to 0 as $n \to \infty$ with product probability one.

For any $x \in X$, let δ_x be the probability measure concentrated at x, and let

$$\mu_n = (\delta_{x_1} + \cdots + \delta_{x_n})/n,$$

where the δ_{x_i} are regarded as measures on the same space X, i.e. we have a mapping of Ω into probability measures on X for each n. The μ_n will be called "empirical measures" associated with μ. Clearly $\int f \, d(\mu_n - \mu)$ converges to zero with probability one for f integrable on X. If X has a topology for which continuous real-valued functions are $\mathcal{S}$-measurable, this implies that $\mu_n \to \mu$ in the weak* topology of the dual of the Banach space of bounded continuous real functions on X, with supremum norm.

Now, if $f_1, \cdots, f_m \varepsilon H$ the multidimensional central limit theorem implies that if

$$\nu_n = n^{\frac{1}{2}}(\mu_n - \mu),$$

the joint distribution of $\int f_1\, d\nu_n, \cdots, \int f_m\, d\nu_n$ will approach that of $L(f_1), \cdots, L(f_m)$ where L is the centered noise rlf for μ.

If X is the real line, the manner of convergence of the distribution of ν_n to that of L has been intensively studied by several authors, particularly when μ is a continuous measure; all continuous measures are equivalent for this purpose. I have extended some results to $X = R^k$, $k > 1$, and will publish the extensions separately [5].

8. Special examples. In this section we consider various generalizations of the one-dimensional Brownian motion process.

One is Lévy's Brownian motion, a Gaussian process G on R^k which satisfies, for any $s, t \varepsilon R^k$,

$$E[G(s) - G(t)]^2 = |s - t|, E(G(s)) = 0,$$

where (say) $G(0)$ has a given Gaussian distribution. For definiteness we let $G(0)$ be identically zero.

Theorem 2.1 easily implies almost sure continuity of G. As noted at the end of Section 2, Lévy's original proof is somewhat simpler in this case (Lévy [15]).

Here it is unnatural to consider $\partial^k G/\partial t_1 \cdots \partial t_k$. Instead, we use the gradient

$$\langle \partial G/\partial t_1, \cdots, \partial G/\partial t_k \rangle,$$

i.e. we consider the rlf from $\mathfrak{D}(R^k)$ to R^k defined by

$$L(f) = -\langle \int (\partial f/\partial t_1)G(t)\, dt, \cdots, \int (\partial f/\partial t_k)G(t)\, dt \rangle.$$

It is easy to see that L is stationary. The Minlos measure for L is concentrated in the set of vector-valued distributions $\langle \partial x/\partial t_1, \cdots, \partial x/t_k \rangle$ where $x(\ )$ are continuous functions.

As in Section 4, the distribution defined by a function or partial derivative A will be written $[A]$. Ito [13] proved that there is a $c_k > 0$ such that for any i and j and $f \varepsilon \mathfrak{D}$,

$$E([\partial G/\partial t_i](f)[\partial G/\partial t_j](g)^-) = c_k \int_{R^k} (\lambda_i \lambda_j/|\lambda|^{k+1})\tilde{f}(\lambda)(\tilde{g})^-(\lambda)\, d\lambda$$

(see Section 6 for the notation).

It follows easily (see Yaglom [24]) that for any $h \varepsilon \mathfrak{D}$ whose integral over R^k is zero,

$$E|[G](h)|^2 = 4\pi^2 c_k \int |\tilde{h}(\lambda)|^2\, d\lambda/|\lambda|^{k+1}.$$

Let $4\pi^2 c_k = b_k$ and

$$d\mu_k(\lambda) = b_k\, d\lambda/|\lambda|^{k+1}.$$

To evaluate c_k, let $h_n \varepsilon \mathfrak{D}$ have integral zero and be such that as $n \to \infty$,

$$\int g h_n\, dt \to g(1, 0, \cdots, 0) - g(-1, 0, \cdots, 0)$$

for any bounded continuous function g on R^k, with all $h_n(t)$ vanishing for $|t| \geqq 2$. Then

$$\begin{aligned} 2 &= \lim_{n\to\infty} E|[G](h_n)|^2 \\ &= \lim_{n\to\infty} \int |\tilde{h}_n|^2\, d\mu_k \\ &= b_k \int \sin^2 (2\pi\lambda_1)\, d\lambda/|\lambda|^{k+1}, \end{aligned}$$

by the dominated convergence theorem, since for some $K > 0$, $|\tilde{h}_n(\lambda)|^2 \subseteq K\, |\lambda|^2$ for all n and λ.

Now for $k = 1$, this yields $2\pi^2 b_k = 2$, $b_k = 1/\pi^2$. For $k \geqq 2$, let S_k be the total surface measure of a sphere of radius 1 in R^k, and let $r = (\lambda_2^2 + \cdots + \lambda_k^2)^{\frac{1}{2}}$, $w = 1/r^2$. Then

$$\begin{aligned} 2 &= 4b_k S_{k-1} \int_{-\infty}^{\infty} \sin^2 (2\pi\lambda_1)\, d\lambda_1 \int_0^\infty r^{k-2}\, |\lambda|^{k+1}\, dr \\ &= 4b_k S_{k-1} \int_{-\infty}^{\infty} \sin^2 (2\pi\lambda_1)\, d\lambda_1 \int_0^\infty (1 + w\lambda_1^2)^{-(k+1)/2}\, dw/2 \\ &= [4b_k S_{k-1}/(k-1)] \int_{-\infty}^{\infty} [\sin^2 (2\pi\lambda_1)/\lambda_1^2]\, d\lambda_1 \\ &= [4b_k S_{k-1}/(k-1)]\cdot 2\pi^2. \end{aligned}$$

Thus $b_k = [4\pi^2 S_{k-1}/(k-1)]^{-1} = (4\pi^2 V_{k-1})^{-1}$ where V_j is the volume of a ball of radius 1 in R^j.

Now to complete the description of the covariance distribution of G, let $\phi \varepsilon \mathfrak{D}$ have

$$\int \phi(t)\, dt = a \neq 0.$$

Take $g_n \varepsilon \mathfrak{D}$ with $\int g_n h\, dt \to h(0)$ for all bounded continuous h on R^k, with $g_n = 0$ for $|t| \geqq 1$ for all n. Then

$$\lim_{n\to\infty} E|[G](g_n)|^2 = 0,$$

and for any $f \varepsilon \mathfrak{D}$ with integral zero,

$$E([G](ag_n - \phi)\overline{[G](f)}) = \int (a\tilde{g}_n - \tilde{\phi})(\tilde{f})^-\, d\mu_k\,.$$

Letting $n \to \infty$ we get

$$E([G](\phi)\overline{[G](f)}) = \int (\tilde{\phi} - \tilde{\phi}(0))(\tilde{f})^-\, d\mu_k\,.$$

Now

$$E([G](\phi)\overline{[G](ag_n - \phi)}) = \int (\tilde{\phi} - \tilde{\phi}(0))(a\tilde{g}_n - \tilde{\phi})^-\, d\mu_k\,.$$

Again letting $n \to \infty$, we have $E|[G](\phi)|^2 = \int |\tilde{\phi} - \tilde{\phi}(0)|^2\, d\mu_k$. Thus for any $f, g \varepsilon \mathfrak{D}$,

$$E([G](f)\overline{[G](g)}) = \int (\tilde{f} - \tilde{f}(0))(\tilde{g} - \tilde{g}(0))^{-} \, d\mu_k ,$$

so that the covariance distribution of $[G]$ is completely determined.

Another generalization of Brownian motion is the indefinite integral $x(t)$ of the "white noise" rlf L with $E(|L(f)|^2) = \int |f(t)|^2 \, dt$. $x(t)$ is almost surely continuous by Theorem 4.2. For $k = 1$ this is classical, and for $k = 2$ it was proved by J. Yeh [25].

Still another generalization is the "multiple Wiener integral" introduced by Ito ([10], [11]), where we let

$$L_I(f) = \int f(t_1, \cdots, t_k) \, dx_1(t_1) \cdots dx_k(t_k)$$

and $x_1, \cdots, x_k$ are independent real or complex Brownian motion processes on one real parameter. The rlf L_I is clearly defined at least on $\mathfrak{D}$, and satisfies

$$E(L_I(f)) = 0,$$
$$E(L_I(f)\overline{L_I(g)}) = \int f\bar{g} \, dt$$

for any $f, g \in \mathfrak{D}$. Thus L_I can be extended to a continuous rlf on $L^2(R^k, dt)$, and its means and covariances are precisely those of white noise.

However, for $k > 1$ L_I is not Gaussian. For example, if f is the indicator function of a rectangular solid parallel to the axes, $L_I(f)$ is a product of independent Gaussian random variables and hence is not Gaussian. Also, L_I does not have independent values.

Acknowledgments. Parts of this paper have been adapted from my doctoral thesis at Princeton, written with the advice of G. A. Hunt and Edward Nelson. Recent conversations with V. Strassen have also been very helpful.

REFERENCES

[1] Belaev, Yu. K. (1961). Continuity and Hölder's conditions for sample functions of stationary Gaussian processes. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **2** 23–24. Univ. California Press.

[2] Dudley, R. M. (1964). On sequential convergence. *Trans. Amer. Math. Soc.* **112** 483–507.

[3] Dudley, R. M. (1964). Singular translates of measures on linear spaces. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **3** 128–137.

[4] Dudley, R. M. (1965). Fourier analysis of sub-stationary processes with a finite moment. To appear in *Trans. Amer. Math. Soc.*

[5] Dudley, R. M. (1965). Weak convergence of probabilities on nonseparable metric spaces and empirical measures on Euclidean spaces. To appear in *Illinois J. Math.*

[6] Fortet, R. (1958). Recent advances in probability theory. *Some Aspects of Analysis and Probability*. Wiley, New York, 171–240.

[7] Gelfand, I. M. and Vilenkin, N. Ya. (1961). *Some Applications of Harmonic Analysis (Generalized Functions* 4).

[8] Gross, L. (1963). Harmonic analysis in Hilbert space. *Mem. Amer. Math. Soc.* No. 46.

[9] Hunt, G. A. (1951). Random Fourier transforms. *Trans. Amer. Math. Soc.* **71** 38–69.

[10] Ito, Kiyosi (1951). Multiple Wiener integral. *J. Math. Soc. Japan* **3** 156–169.

[11] Ito, Kiyosi (1952). Complex multiple Wiener integral. *Japan. J. Math.* **22** 63–86.

[12] Ito, Kiyosi (1953). Stationary random distributions. *Mem. Coll. Sci. Univ. Kyoto Ser. A Math.* **28** 212–223.

[13] ITO, KIYOSI (1956). Isotropic random current. *Proc. Third Berkeley Symp. Math. Statist. Prob.* **2** 125–132.

[14] KOLMOGOROV, A. N. (1959). A note to the papers of R. A. Minlos and V. Sazonov. *Teor. Veroyatnost. i Primenen* **4** 237–239. (in Russian) (English translation in *Theor. Prob. Appl.* **4** 221–223.

[15] LÉVY, P. (1948). *Processus Stochastiques et Mouvement Brownien.* (Monographies des Probabilités, Fascicule 6) Gauthier-Villars, Paris.

[16] LOÈVE, M. (1960). *Probability Théory.* (2nd ed.). Van Nostrand, Princeton.

[17] MINLOS, R. A. (1959). Generalized random processes and their extensions to measures. *Trudy Moskov. Mat. Obšč.* **8** 497–518 (in Russian) (English translation in *Selected Transl. Math. Statist. Prob.* **3** 291–314).

[18] PROKHOROV, YU. V. (1956). Convergence of random processes and limit theorems in probability. *Teor. Veroyatnost. i Primenen.* **1** 177–238 (in Russian) (English translation in *Theor. Prob. Appl.* **1** 157–214).

[19] SCHWARTZ, L. (1957 and 1959). *Théorie des Distributions.* Hermann, Paris.

[20] SCHWARTZ, L. (1954). Espaces de fonctions différentiables à valeurs vectorielles. *J. Analyse Math.* **4** 88–158.

[21] SCHWARTZ, L. (1957) and (1958). Théorie des distributions à valeurs vectorielles. *Ann. Inst. Fourier Grenoble* **7** 1–142 and **8** 1–210.

[22] URBANIK, K. (1960). Gaussian measures in locally compact abelian topological groups. *Studia Math.* **19** 77–88.

[23] URBANIK, K. (1960). A contribution to the theory of generalized stationary random fields. *Trans. Second Prague Conference Information Theor. Statist. Decision Functions, Random Processes.*

[24] YAGLOM, A. M. (1957). Certain types of random fields in n-dimensional space, similar to stationary stochastic processes. *Teor. Veroyatnost. i Primenen* **2** 292–338 (in Russian) (English translation in *Theor. Prob. Appl.* **2** 273–320).

[25] YEH, J. (1960). Wiener measure in a space of functions of two variables. *Trans. Amer. Math. Soc.* **95** 433–450.