SOME CONTINUITY PROPERTIES
OF BROWNIAN MOTION WITH THE
TIME PARAMETER IN HILBERT SPACE([1])

BY

SIMEON M. BERMAN

1. **Introduction and summary.** Let H be a real separable Hilbert space with inner product function (x, y) and norm $\|x\| = (x, x)^{1/2}$, $x, y \in H$. Let $X(x)$, $x \in H$, be the Brownian motion process of P. Lévy: for any finite set of elements $x_1, \ldots, x_k$ in H, the real random variables $X(x_1), \ldots, X(x_k)$ have a joint Gaussian distribution with expectations 0 and covariances

$$(1.1) \qquad EX(x_i)X(x_j) = \tfrac{1}{2}[\|x_i\| + \|x_j\| - \|x_i - x_j\|].$$

Lévy introduced this process in [5, Chapter 8]. A bibliography of his later work is given in the second edition. The sample functions of this process are continuous when H is finite dimensional; however, when H is infinite dimensional, the sample functions are not only discontinuous but also unbounded on every sphere. The local irregularity of the sample functions is caused by the property of infinite-dimensional Hilbert space that a point can be approached from any one of infinitely many mutually orthogonal directions: a sample function has a negligible chance of being well behaved on all of the paths leading to a point. We ask: what are some sets upon which the sample functions are continuous? A simple example is a finite-dimensional subspace, as we have already indicated; it is therefore natural to propose sets which are topologically "almost finite-dimensional" in the sense that the mutual distances of their points are determined mostly by finite-dimensional projections of these points. Such a set is a cube. P. T. Strait proved the continuity of the sample functions on cubes whose edges have lengths $1, 2^{-1}, 2^{-2}, \ldots$ [10]. Here we shall study the continuity of the sample functions on ellipsoids. Our results are:

(i) A necessary condition for the continuity on an ellipsoid is that the lengths of the semiaxes converge to 0.

(ii) The above condition is not sufficient: we can construct such an ellipsoid upon which the sample functions are unbounded.

(iii) A sufficient condition for continuity is that the lengths of the semiaxes form a pth power summable sequence for some $p < 1$. This implies Strait's result

because every cube of the given type can be enclosed in an ellipsoid of the given type. The proof of our theorem is based on

(a) a general criterion of V. Strassen for the continuity of the sample functions of a Gaussian process whose parameter set is a totally bounded pseudo-metric space; and

(b) an estimate of the ε-entropy of an ellipsoid.

We construct a series expansion for the process which converges in quadratic mean and with probability 1 at each point of H. The expansion is constructed from independent identically distributed Brownian motions of a real time parameter: each term is a "Fourier coefficient" of a "rescheduled" Brownian motion with respect to an orthonormal system of functions in H, where orthonormality is defined by means of Segal's infinite-dimensional product Gaussian distribution on H. The series representation is used to prove a dichotomous characteristic of the process: on each compact subset of H the process has continuous sample functions, or, if not, then there is a point in the set at which the sample functions are discontinuous. In the former case the series converges uniformly with probability 1. The latter case is not vacuous because in (ii) above we have constructed a compact ellipsoid upon which the sample functions are not continuous. Finally we prove a generalization of Strassen's lemma giving a criterion for the weak compactness of a family of measures corresponding to Gaussian processes on a compact metric space with continuous sample functions.

I thank R. M. Dudley for sending me the theorem of Strassen on the continuity of sample functions of Gaussian processes before its publication [2]. I also thank S. R. S. Varadhan for some very helpful conversations on measures in linear spaces and their connections with stochastic processes. The referee was most generous with his suggestions: he proposed the sharper inequality (2.4), and a considerable improvement of Theorem 4.2. H. P. McKean read the draft, and I thank him for the improvements he suggested.

2. **Continuity on an ellipsoid.** Let S be a topological space and $X(x)$, $x \in S$ a real-valued Gaussian process on S. According to the fundamental theorem of Kolmogorov, the finite-dimensional distributions of the process determine a unique probability measure P in the product space R^S of all real-valued functions on S. The probability measure is defined on the σ-field $\mathscr{F}$ of subsets of R^S generated by the sets of functions f of the form $\{f : f(x) \leqq r\}$, where $x \in S$ and r is real. Let C be the subset of R^S consisting of all continuous functions on S to the real line. The process $X(\cdot)$ is said to have continuous sample functions if the P-outer measure of C is 1.

Let $\{e_n\}$ be a fixed orthonormal basis of H, and $\{\lambda_n\}$ a sequence of positive numbers. A set in H of the form $\{x : x \in H, \sum_{n=0}^{\infty} (x, e_n)^2/\lambda_n^2 \leqq 1\}$ is called an ellipsoid of semiaxes lengths $\lambda_1, \lambda_2, \ldots$. It is well known that the ellipsoid is compact if $\lambda_n \to 0$ [3, p. 30]. Our first result is

LEMMA 2.1. *A necessary condition for the continuity of the sample functions of the Brownian motion process on an ellipsoid is that* $\lambda_n \to 0$. *This is not a sufficient condition; indeed, there exists an ellipsoid satisfying it but where the sample functions of the process are all unbounded along a fixed sequence of points converging to the origin in* **H**.

Proof. Consider an ellipsoid for which the condition of the lemma is not satisfied; then for some $\varepsilon > 0$ and some subsequence $\{m\}$ we have $\lambda_m > \varepsilon$ for all m. Let B be the unit ball in the subspace spanned by the subset $\{e_m\}$ of the basis of **H**. The ball εB is a subset of the ellipsoid; hence, by Lévy's classical result [5, p. 345 (2nd ed.)] which we mentioned in §1, the process is unbounded almost surely.

We now construct a sequence $\{\lambda_n\}$ which converges to 0 and is such that $\{X(\lambda_n e_n)\}$ is unbounded almost surely. We recall a certain representation of a finite or denumerably infinite family of equally (positively) correlated Gaussian random variables: if $X_1, X_2, \ldots$ have mean 0 and common variance σ^2, and if all pairs (X_i, X_j), $i \neq j$ have common correlation coefficient ρ, $\rho < \sigma^2$, then the sequence $\{X_n\}$ is stochastically equivalent to the sequence $\{U_n + Y\}$, where $Y, U_1, U_2, \ldots$ are mutually independent Gaussian random variables with common mean 0 and where $EY^2 = \rho\sigma^2$, $EU_1^2 = EU_2^2 = \cdots = \sigma^2(1-\rho)$. (This can be verified by checking the first and second order moments.) This representation implies that the sequence $\{X_n\}$ is unbounded almost surely. (Lévy has a different proof [5, p. 345 (2nd ed.)].) The sequence $\{X(\varepsilon_1 e_n)\}$ is equally correlated for any $\varepsilon_1 > 0$; hence, for any pair of positive numbers B, δ, there exists an integer N_1 (depending on ε_1, δ, B) such that

$$P\{\max (|X(\varepsilon_1 e_1)|, \ldots, |X(\varepsilon_1 e_n)|) \leq B\} < \delta, \qquad n \geq N_1.$$

Now repeat this construction with a number $\varepsilon_2 < \varepsilon_1$ in place of ε_1, with δ^2 in the place of δ, and with the "decapitated" sequence $\{e_n, n > N_1\}$ in the place of the original sequence: there exists an integer $N_2 > N_1$ (depending on N_1, ε_2, δ, B) such that

$$P\{\max (|X(\varepsilon_2 e_{N_1 + 1})|, \ldots, |X(\varepsilon_2 e_n)|) \leq B\} < \delta^2, \qquad n \geq N_2.$$

Proceeding in this way with a sequence of numbers $\{\varepsilon_n\}$ converging to 0, we get a subsequence $\{N_k\}$ such that

$$P\{\max [|X(\varepsilon_k e_j)| : N_k < j \leq N_{k+1}] \leq B\} < \delta^k,$$

for $k = 1, 2, \ldots$. Summing the terms on the left-hand side of this inequality over all k and picking $\delta < 1$, we find that the resulting series converges; hence, by the Borel-Cantelli lemma, there is probability 1 that at most finitely many of the inequalities

$$\max [|X(\varepsilon_k e_j)| : N_k < j \leq N_{k+1}] \leq B, \qquad k = 1, 2, \ldots$$

occur. Define the sequence $\{\lambda_n\}$ as

$$\lambda_n = \varepsilon_1, \qquad 1 \leq n \leq N_1, \qquad \lambda_n = \varepsilon_2, \qquad N_1 < n \leq N_2, \ldots \text{ etc.};$$

then $|X(\lambda_n e_n)| > B$ infinitely often almost surely but $\lambda_n \to 0$. The proof of Lemma 2.1 is complete.

Suppose that S is a compact metric space. For every $\varepsilon > 0$, there exists a finite set of balls in S of radius at most ε such that S is covered by these balls because S is compact. Let $N(\varepsilon)$ be the least number of such balls; then $\log N(\varepsilon)$ is called the ε-entropy of S.

Lemma 2.2 is due to V. Strassen. It was noted by R. M. Dudley who proved similar results [1], [12]. We shall not give a proof of the lemma now because it is implicit in Theorem 5.1 below:

LEMMA 2.2 (STRASSEN). *Suppose that S has the metric*

$$d(x, y) = \{E[X(x) - X(y)]^2\}^{1/2}, \qquad x, y \in S.$$

If for some constant $K > 0$ and some constant β, $0 < \beta < 1$, the inequality

$$N(\varepsilon) \leqq \exp [K\varepsilon^{-2\beta}]$$

holds for all sufficiently small $\varepsilon > 0$, then $X(\cdot)$ has continuous sample functions on S.

An asymptotic bound for the ε-entropy of certain ellipsoids is due to Gel'fand and Vilenkin [3, p. 89]:

Suppose that the lengths of the semiaxes of an ellipsoid satisfy the condition

(2.1) $$\sum_{n=1}^{\infty} \lambda_n^c < \infty$$

for some c, $0 < c < 1$; then,

(2.2) $$\limsup_{\varepsilon \to 0} \frac{\log \log N(\varepsilon)}{\log (1/\varepsilon)} \leqq 2c.$$

This bound can be sharpened, and can be replaced by

(2.3) $$\lim_{\varepsilon \to 0} \frac{\log \log N(\varepsilon)}{\log (1/\varepsilon)} = c;$$

this means that for any number $c' > c$, the inequality

(2.4) $$N(\varepsilon) \leqq \exp [\varepsilon^{-c'}]$$

holds for all sufficiently small $\varepsilon > 0$. The equation (2.3) is obtained in the following way. Without loss of generality, the sequence $\{\lambda_n\}$ may be assumed to satisfy $\lambda_1 \geqq \lambda_2 \geqq \cdots$. The largest distance from a point of the ellipsoid to the n-dimensional subspace $x_j = 0$, $j \geqq n$ is λ_n (not $(\sum_{j=n}^{\infty} \lambda_j^2)^{1/2}$); thus, in [3, p. 89], the estimate (20) is not the best one to use. By a modification of their method (using $p = n(\varepsilon/2)$ instead of $p = n(\varepsilon^2/4a)$ in their derivation) we find that their equation (22) can be replaced by $\rho_{km} = \lambda_{km}$.

THEOREM 2.1. *The sample functions of $X(\cdot)$ are continuous on an ellipsoid if the lengths of the semiaxes satisfy (2.1) for some c, $0 < c < 1$.*

Proof. Consider the metric

$$d(x, y) = \{E[X(x) - X(y)]^2\}^{1/2} = \|x - y\|^{1/2}$$

on H; although it generates the same topology on H as the inner product norm, it has a different ε-entropy function. A ball of radius ε in the norm metric of H is of radius $\varepsilon^{1/2}$ in the metric d. Since the ε-entropy of the ellipsoid in the *norm metric* satisfies (2.4) for any $c' > c$, where $0 < c < 1$, it follows that there is a number β, $0 < \beta < 1$ such that (2.4) is satisfied for $c' = \beta$ by the ε-entropy of the *metric d*. Lemma 2.2 (with $K = 1$) now implies the conclusion of the theorem.

3. **A series representation for the process.** We shall now construct a series representation of the process. It resembles the expansion of the Brownian motion over a finite-dimensional space, constructed by H. P. McKean [6]. While the latter series is in terms of the orthonormal basis of the space of functions square-integrable with respect to the uniform measure on the unit sphere (spherical functions) our expansion is in terms of an orthonormal system of functions on H square integrable with respect to the infinite-dimensional product Gaussian distribution of Segal [9]. Later we use the representation to prove the dichotomy result for the sample functions.

Let μ be the "weak" product Gaussian distribution on H: there is a linear mapping L from H into the random variables on some probability space such that for any fixed elements $x_1, \ldots, x_k$ in H, the random variables $L(x_1), \ldots, L(x_k)$ have a joint Gaussian distribution with means 0 and covariance matrix $((x_i, x_j))$. We shall use the property of μ that if $f(u_1, \ldots, u_k)$ is a nonnegative measurable function on R^k then

$$\int_H f(L(x_1), \ldots, L(x_k)) \, d\mu$$

is defined as the integral of f with respect to the joint Gaussian distribution of $(L(x_1), \ldots, L(x_k))$ in R^k [9].

Let $\{h_n(\cdot)\}$ be an orthonormal basis of the space of functions on the real line which are square-integrable with respect to the one-dimensional Gaussian distribution with mean 0 and variance 1; for example, $\{h_n\}$ may be taken as the normalized Hermite polynomials. Let $\mathcal{H}_k$ be the family of functions on R^k of the form

$$h_{n_1}(x_1) h_{n_2}(x_2) \cdots h_{n_k}(x_k), \qquad (x_1, \ldots, x_k) \in R^k$$

where each of the indices $n_1, \ldots, n_k$ range over all the nonnegative integers. The family $\mathcal{H}_k$ is countable and is an orthonormal basis for the space of functions on R^k square-integrable with respect to the product standard Gaussian distribution in R^k. The union of the sets $\mathcal{H}_k$, $k = 1, 2, \ldots$, is countable; we denote it as a sequence $\{\phi_n(\cdot)\}$ of functions on H.

For the purpose of stating the theorem that follows, we shall have to define certain integrals with respect to μ over H. In writing such integrals we shall omit

the symbol H from under the sign of integration whenever the integration is over the whole space. Let $f(u)$ be a measurable function on the real line such that

$$\int_{-\infty}^{\infty} |f(u)| \exp(-u^2/2)\, du < \infty.$$

Let us denote by R^k the (finite-dimensional) subspace spanned by $e_1, \ldots, e_k$, and let ϕ be a function on H which depends on a point of H only through its projection on R^k. We shall now define the integral

$$(3.1) \qquad \int f(L(x))\phi\, d\mu, \qquad x \in H.$$

Write $x = x' + x''$, where x' is the projection of x onto R^k, and x'' the orthogonal projection; put

$$\Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^{x} \exp(-y^2/2)\, dy;$$

then, the integral (3.1) is defined as

$$(3.2) \qquad \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f\left(\sum_{i=1}^{k} x_i y_i + \|x''\| y_{k+1}\right) \phi(y_1, \ldots, y_k) \prod_{i=1}^{k+1} d\Phi(y_i),$$

where $x_1, \ldots, x_k$ and $y_1, \ldots, y_k$ are coordinates of the projections on R^k of x and y respectively, and where the integral is a $(k+1)$-fold multiple integral. For a nonnegative, measurable function $f = f(u, v)$ on the plane we define the double integral

$$(3.3) \qquad \iint f(L_1(x_1), L_2(x_2))(\phi \cdot \phi)\, d(\mu \cdot \mu), \qquad x_1, x_2 \in H,$$

as

$$(3.4) \qquad \begin{aligned} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} & f\left(\sum_{i=1}^{k} x_{i1} u_i + \|x_1''\| u_{k+1}, \sum_{i=1}^{k} x_{i2} v_i + \|x_2''\| v_{k+1}\right) \\ & \cdot \phi(u_1, \ldots, u_k)\phi(v_1, \ldots, v_k) \prod_{i=1}^{k+1} d\Phi(u_i)\, d\Phi(v_i), \end{aligned}$$

where $x_{1j}, \ldots, x_{kj}$ are the coordinates of the projection on R^k of x_j, $j = 1, 2$, and where x_j'' are the corresponding orthogonal projections. We note that when ϕ is taken as the function identically equal to 1, then the integral (3.1) has the value

$$(3.5) \qquad \int_{-\infty}^{\infty} f(z\|x\|)\, d\Phi(z).$$

THEOREM 3.1. *Let* $X_1(t), X_2(t), \ldots, t \geq 0$ *be independent standard Brownian motions on the real line. Put*

$$(3.6) \qquad Y(x) = \sum_{n=1}^{\infty} \int X_n(L^+(x))\phi_n\, d\mu, \qquad x \in H,$$

where $L^+(x) = [L(x)]^+$ and $c^+ = \max(c, 0)$. The series converges in quadratic mean (and with probability 1) for each $x \in H$; furthermore, it is stochastically equivalent to $(2/\pi)^{1/4} X(x)$, $x \in H$.

Proof. First we shall prove that

$$(3.7) \qquad \iint \min[L_1^+(x), L_2^+(y)](|\phi \cdot \phi|)\, d(\mu \cdot \mu) < \infty$$

for any function ϕ in the sequence $\{\phi_n\}$. Put

$$\varepsilon(t) = 0, \qquad t < 0, \qquad \varepsilon(t) = 1, \qquad t \geqq 0;$$

then we write

$$\min(s^+, t^+) = \int_0^\infty \varepsilon(t - w)\varepsilon(s - w)\, dw,$$

and write (3.7) as

$$\iint \left\{ \int_0^\infty \varepsilon(L_1(x) - w)\varepsilon(L_2(y) - w)\, dw \right\} (|\phi \cdot \phi|)\, d(\mu \cdot \mu).$$

Although μ is not a countably additive measure, Fubini's theorem is still applicable· to the *latter* integral because of its definition as a multiple integral over Euclidean space. Since the integrand is nonnegative we may invert the order of integration. Integrating first with respect to $d(\mu \cdot \mu)$, we have

$$\int_0^\infty \left\{ \int \varepsilon(L(x) - w)|\phi|\, d\mu \right\} \left\{ \int \varepsilon(L(y) - w)|\phi|\, d\mu \right\} dw.$$

The Cauchy-Schwarz inequality is valid for our integrals also; since $|\phi|^2$ has μ-integral 1, the expression above is dominated by

$$\int_0^\infty \left\{ \int \varepsilon(L(x) - w)\, d\mu \right\}^{1/2} \left\{ \int \varepsilon(L(y) - w)\, d\mu \right\}^{1/2} dw.$$

A second application of that inequality shows that this integral is dominated by the square root of

$$\int_0^\infty \int \varepsilon(L(x) - w)\, d\mu\, dw \cdot \int_0^\infty \int \varepsilon(L(y) - w)\, d\mu\, dw.$$

Interchanging the order of integration, we find this product equal to

$$\int L^+(x)\, d\mu \cdot \int L^+(y)\, d\mu;$$

this, by (3.5), is equal to $\|x\| \cdot \|y\|/2\pi$.

Next we note the inequalities

$$\iint E|X_n(L_1^+(x))X_m(L_2^+(x))|(|\phi_n \cdot \phi_m|)\, d(\mu \cdot \mu)$$

$$\leqq \int (L^+(x))^{1/2}|\phi_n|\, d\mu \cdot \int (L^+(x))^{1/2}|\phi_m|\, d\mu$$

$$\leqq \int L^+(x)\, d\mu, \qquad m, n = 1, 2, \ldots;$$

these follow from two successive applications of the Cauchy-Schwarz inequality, from the form of the variance of the Brownian motion, and from the fact that ϕ_n^2 has the integral 1. By virtue of these inequalities, each term of the series (3.6) has finite variance; furthermore, the covariance function of the process defined by the nth term of the series is given by the integral

$$\iint \min [L_1^+(x), L_2^+(y)](\phi_n \cdot \phi_n) \, d(\mu \cdot \mu).$$

In accordance with the discussion following the inequality (3.7), this double integral may be put in the form

(3.8)
$$\int_0^\infty g_n(x, w) g_n(y, w) \, dw$$

where

$$g_n(x, w) = \int \varepsilon(L(x) - w) \phi_n \, d\mu.$$

We shall show that

(3.9)
$$\int \sum_{n=1}^\infty g_n^2(x, w) \, dw \leq (2/\pi)^{-1/2} \|x\|, \qquad x \in H,$$

which implies that the series (3.6) converges in quadratic mean, and, by the mutual independence of the summands, with probability 1. We have

(3.10)
$$\sum_{n=1}^\infty g_n^2(x, w) \leq \int \varepsilon(L^+(x) - w) \, d\mu.$$

Indeed, the integrand on the right-hand side is a square-integrable function on the sample space of the random variables $L(e_1)$, $L(e_2)$, ...; furthermore, $\{\phi_n\}$ is a complete orthonormal system for the corresponding L_2 space; hence, (3.10) is the result of the classical Bessel inequality.

Now we find the covariance function of the series $Y(x)$; we have

(3.11)
$$E \, Y(x) Y(y) = \sum_{n=1}^\infty \int_0^\infty g_n(x, w) g_n(y, w) \, dw$$

by virtue of the mutual independence of the terms of the series (3.6) and the representation (3.8). A standard argument involving the inequality (3.9) permits the interchange of the order of summation and integration. We shall evaluate the series $\sum g_n(x, w) g_n(y, w)$ and then find the covariance function by integration over w. By Parseval's theorem (see argument following (3.10)) we have

(3.12)
$$\sum_{n=1}^\infty g_n(x, w) g_n(y, w) = \int \varepsilon(L(x) - w) \varepsilon(L(y) - w) \, d\mu.$$

We integrate both sides of (3.12) with respect to w: using Fubini's theorem, we obtain the expression

$$(3.13) \qquad\qquad \int \min(L^+(x), L^+(y)) \, d\mu.$$

This is equal to

$$(3.14) \qquad\qquad \tfrac{1}{2}(2\pi)^{-1/2}[\|x\| + \|y\| - \|x-y\|]$$

(and this confirms the assertion of the theorem about stochastic equivalence). To prove that the expressions (3.13) and (3.14) are equal, we note: if X and Y have a joint Gaussian distribution with 0 means, then the expected value of $\min(X^+, Y^+)$ is equal to $\tfrac{1}{2}(2\pi)^{-1/2}\{(EX^2)^{1/2} +)EY^2)^{1/2} - (E(X-Y)^2)^{1/2}\}$; this can be verified by expressing $\min(X^+, Y^+)$ as $XI_{[0 < X < Y]} + YI_{[0 < Y < X]}$, where $I_{[\cdots]}$ is the indicator function of $[\cdots]$.

The following result will be used in §4.

LEMMA 3.1. *Each term of the series* (3.6) *is a Gaussian process with continuous sample functions on* **H**.

Proof. A typical term of the series is

$$Z(x) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} X_n\left(\left[\sum_{i=1}^{k} x_i z_i + \|x''\| z_{k+1}\right]^+\right) \phi_n(z_1, \ldots, z_k) \prod_{i=1}^{k+1} d\Phi(z_k),$$

where $x_1, \ldots, x_k$ are the first k coordinates of x and where x'' is the projection on the orthogonal complement of R^k; hence, such a term depends on a $(k+1)$-dimensional parameter which itself is a continuous function on **H**. It is sufficient to show that $Z(\cdot)$, as a process on $(k+1)$ parameters, has continuous sample functions. We use a criterion of Dudley to prove this [1]. Put $t = (t_1, \ldots, t_{k+1})$ and $h = (h_1, \ldots, h_{k+1})$, and call the process $Z(t)$; then, by Fubini's theorem, the Cauchy-Schwarz inequality, and the form of the variance of Brownian motion $X_n(\cdot)$, we find

$$E[Z(t+h) - Z(t)]^2 \leqq \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left|\left(\sum_{i=1}^{k+1}(t_i + h_i)z_i\right)^+ - \left(\sum_{i=1}^{k+1} t_i z_i\right)^+\right| \prod_{i=1}^{k+1} d\Phi(z_i).$$

Apply the elementary inequality $|a^+ - b^+| \leqq |a-b|$: the multiple integral is dominated by

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left|\sum_{i=1}^{k+1} h_i z_i\right| \prod_{i=1}^{k+1} d\Phi(z_i) = \|h\|/(2\pi)^{1/2}.$$

By Dudley's criterion, the sample functions are continuous.

4. A dichotomous property of the process. We return to the terminology of §2. Suppose that C has P-outer measure 1. We shall restrict P to the set C: we replace the σ-field $\mathscr{F}$ by the σ-field $\mathscr{F} \cap C$ consisting of all sets of the form $F \cap C$, $F \in \mathscr{F}$,

and we define the probability measure of $F \cap C$ as $P(F)$. Suppose S is compact. Furnished with the supremum norm, C is a separable Banach space. Let C^* be the space of continuous linear functionals L on C. Let us denote by $\mathscr{F}_1$ the σ-field in C generated by the sets of the form $\{f : L(f) \leq r\}$, where $L \in C^*$ and r is real. Since C is separable, $\mathscr{F}_1$ is the σ-field of Borel sets of C [7]; we also have $\mathscr{F}_1 = \mathscr{F} \cap C$ from the following considerations. The inclusion $\mathscr{F} \cap C \subset \mathscr{F}_1$ follows from the fact that the mapping $f \to f(x)$ (for a fixed x in S) is a continuous linear functional on C. To prove $\mathscr{F}_1 \subset \mathscr{F} \cap C$, we use the Riesz representation theorem. Let L be a positive continuous linear functional on C; then there exists, by Riesz's theorem [4, p. 247], a Borel measure λ on S such that

$$L(f) = \int_S f \, d\lambda, \quad f \in C.$$

If λ is concentrated on a finite point-set of S, then L is $\mathscr{F} \cap C$-measurable. The argument for *every* positive continuous L is that the Borel measures concentrated on finite point-sets of S constitute a dense subset of C^* in the weak $*$ topology: for every λ there exists a sequence $\{\lambda_n\}$ of measures concentrated on finite point-sets of S such that

$$\int_S f \, d\lambda_n \to \int_S f \, d\lambda, \quad f \in C.$$

Now every continuous linear functional L may be expressed as the difference of two positive continuous linear functionals [4, p. 249]; hence every L is $\mathscr{F} \cap C$-measurable, and so $\mathscr{F}_1 \subset \mathscr{F} \cap C$.

The result of this discussion is that a stochastic process with continuous sample functions on S determines a unique probability measure on the Borel sets of the space C (this is well known when S is a real interval [8]). We may treat such a process as a random element in C. The correspondence between stochastic processes with continuous sample functions and random elements in the Banach space C of continuous functions on S is linear: if $Y_1(\cdot)$ and $Y_2(\cdot)$ are jointly distributed processes on S, and $\tilde{Y}_1$ and $\tilde{Y}_2$ the corresponding random elements, then $\tilde{Y}_1 + \tilde{Y}_2$ is the random element corresponding to the process $Y_1(\cdot) + Y_2(\cdot)$; indeed, the finite-dimensional distribution of the latter process is identical with those generated by linear functionals of the random elements $\tilde{Y}_1 + \tilde{Y}_2$.

We shall say that a stochastic process has discontinuous sample functions on S if every compact subset of C has P-measure 0. A sequence of stochastic processes with continuous sample functions on S (defined on a common probability space) is said to converge uniformly if the corresponding random elements in C converge with probability 1.

Let S be an arbitrary compact subset of H. To each term of the series (3.6) there corresponds, by Lemma 3.1, a random element in C; furthermore, to different terms there correspond independent random elements. If the series converges

uniformly, then the process represented by the series has continuous sample functions. Under the zero-one law, the only other possibility is that the corresponding series in C converges with probability 0. We shall shortly show that in the latter case, the process represented by the series (as a quadratic mean limit) has discontinuous sample functions at some point of S.

It is well known that the set of probability measures on a complete separable metric space is itself a complete separable metric space in the topology of weak convergence. For the compactness of a family (P_i) of probability measures on C it is necessary and sufficient that for every $\varepsilon > 0$ there exist a compact subset K of C such that $\inf_i P_i(K) \geqq 1 - \varepsilon$ [8]. A compact subset of the Banach space of continuous functions on S is, characteristically, a uniformly bounded, equicontinuous set of functions, or, equivalently, an equicontinuous set of functions bounded at one point of S.

We present a result on series of independent symmetric random elements in C. (A random element X is symmetric if it has the same distribution as $-X$.) It is an adaptation of Tortrat's generalization of a classical result of Lévy [11]:

LEMMA 4.1. *Let S_n, $n = 1, 2, \ldots$ be the sequence of partial sums of a series of independent symmetric random elements in C. There are two alternatives: either*

(a) *the sequence of distributions of $\{S_n\}$ is weakly compact: or*

(b) *the concentration function of the distribution of the "tail" of the series is 0; more precisely, putting*

$$Q_{nm}(A) = \sup_{x \in C} P\{S_m - S_n + x \in A\}, \qquad A \in C \cap \mathscr{F}, \qquad n < m,$$

$$Q_n(A) = \lim_{m \to \infty} Q_{nm}(A),$$

$$Q(A) = \sup_n Q_n(A),$$

we have $\sup \{Q(K) : K \text{ compact}\} = 0$.

Proof. It is known that the only alternative to (b) is that there exists a sequence $\{c_n\}$ of fixed elements in C such that the distributions of $\{S_n - c_n\}$ are weakly compact [11]. In the latter case the "symmetrized" sequence of distributions is weakly compact, that is, if $\{S_n'\}$ is stochastically equivalent to S_n and independent of it, then the distributions of $\{S_n - S_n'\}$ are weakly compact. These are the same as the distributions of $\{S_n + S_n'\}$ because of the independence of $\{S_n\}$ and $\{S_n'\}$ and their symmetry. It follows that the sequence of distributions of $\{S_n + S_n'\}$ has a Cauchy subsequence (in the space of distributions); therefore, the same must be true for $\{S_n\}$ because a convolution of two measures in a Banach space is concentrated near 0 if and only if the same is true of each factor. The completeness of the space of distributions now implies the alternative (a).

THEOREM 4.1. *Let $X(\cdot)$ be a Brownian motion process on a compact set S. Either*

the sample functions are continuous and the series (3.6) *converges uniformly on S, or the sample functions are discontinuous.*

Proof. Let P be the probability measure on $\mathscr{F}$ (see §2) generated by the distribution of the series (3.6); and let P_n and W_n be the measure corresponding to the sum of the first n terms and to the sum of the terms from $n+1$ on, respectively. Under the definition of addition of functions as pointwise addition, the space R^S is a linear space. Let ℓ be a generic element of R^S. The following convolution equation (which is intuitively evident from probabilistic considerations) relates P, P_n and W_n for every n:

$$(4.1) \qquad P(A) = \int_{R^S} W_n(A - \ell)\, dP_n(\ell), \qquad A \in \mathscr{F}.$$

The equation is true for all sets A which are finite disjoint unions of "rectangles" in R^S, for this is just the convolution equation for distributions in a finite-dimensional space; furthermore, the domain of sets A for which the equation is true is closed under the formation of limits of monotone sequences of sets A; hence, the domain is $\mathscr{F}$.

Let B be an arbitrary compact subset of C. By equation (4.1), $P(B)$ is dominated by the essential supremum of the function $W_n(B - \ell)$ as a function of ℓ, where the essential supremum is defined with respect to the measure P_n. Lemma 3.1 implies that C has P_n-outer measure 1 for every n. It can be concluded that the essential supremum with respect to P_n is dominated by the supremum over C; hence,

$$P(B) \leqq \sup_C W_n(B - \cdot), \qquad \text{for every } n.$$

We now apply Lemma 4.1 to the random elements in C corresponding to the terms of the series (3.6). Since

$$\sup_C W_n(B - \cdot) \leqq Q_{nm}(B), \qquad m > n,$$

it follows under alternative (b) that

$$\limsup_{n \to \infty} \sup_C W_n(B - \cdot) \leqq Q(B) = 0;$$

hence $P(B) = 0$.

Now consider alternative (a) of Lemma 4.1. The sequence of probability measures corresponding to the partial sums of the series (3.6) is weakly compact over C. Since the finite-dimensional distributions of the series converge (by Theorem 3.1), the sequence of probability measures converges weakly over C. The sequence of partial sums of the random elements in C even converges with probability 1; indeed, the weak convergence of the distributions of sums of independent random elements in a Banach space implies the probability 1 convergence of the sum [11].

The conclusion of Theorem 4.1 can be strengthened: if the sample functions are discontinuous on a compact set, then they are actually discontinuous at some fixed

point in the set in the sense to be described. For $\varepsilon > 0$ a function f is said to be uniformly ε-continuous on a metric space S having a metric $d = d(x, y)$, x, $y \in S$, if there exists a number $\delta > 0$ such that $d(x, y) < \delta$ implies $|f(x) - f(y)| < \varepsilon$. It follows that if f is not uniformly ε-continuous then there exist two sequences $\{x_n\}$ and $\{y_n\}$ in S such that $d(x_n, y_n) \to 0$ and $|f(x_n) - f(y_n)| \geqq \varepsilon$ for all n. A stochastic process on S is said to have uniformly ε-continuous sample functions on S if the P-outer measure of the set of uniformly ε-continuous functions in R^S is equal to 1. If the P-outer measure of this set is 0, the process is said to have ε-discontinuous sample functions on S. If the sample functions are ε-discontinuous on every neighborhood of a point, the sample functions are said to be ε-discontinuous at that point.

THEOREM 4.2. *If the sample functions of the Brownian motion $X(\cdot)$ are discontinuous on a compact set S, then they are ε-discontinuous at some point of the set, for some constant $\varepsilon > 0$.*

Proof. First we shall show that for any countable subset $A \subseteq S$, the set of functions on S which are uniformly ε-continuous on the subset A (is $\mathscr{F}$-measurable and) has P-measure 0 or 1 for any $\varepsilon > 0$. By virtue of the representation of $X(\cdot)$ as a series of independent processes with (uniformly) continuous sample functions on S (Theorem 3.1 and Lemma 3.1) it follows that the set of functions on S which are uniformly ε-continuous on A is measurable with respect to a tail σ-field of independent σ-fields, and so has measure 0 or 1; indeed, if a function g is uniformly continuous on A, then $f + g$ is uniformly ε-continuous on A if and only if f is.

If F is an arbitrary $\mathscr{F}$-measurable set containing the uniformly ε-continuous functions on S, then it also contains the functions on S which are uniformly ε-continuous on some A; indeed, F necessarily lies in a sub-σ-field of $\mathscr{F}$ generated by a countable collection of random variables $X(\cdot)$. It follows that either (i) every measurable cover of the uniformly ε-continuous functions has measure 1, or (ii) some measurable cover has measure 0.

If $X(\cdot)$ has discontinuous sample functions on S, then, for some *fixed* $\varepsilon > 0$, it has ε-discontinuous sample functions on S. As a matter of fact, if the set of uniformly ε-continuous sample functions on S had *positive* P-outer measure for every $\varepsilon > 0$ it would have P-outer measure 1 for every $\varepsilon > 0$ and so the sample functions would be continuous on S. Since S is compact, it is a finite union of relatively open subsets (of arbitrarily small diameter) and the sample functions are necessarily ε-discontinuous on one of these. It follows that there exists a point in S and a sequence of sets shrinking to that point such that the sample functions are ε-discontinuous on each of the sets. The proof is complete.

5. A criterion for the weak compactness of a family of measures in C corresponding to Gaussian processes. The content of this section is only indirectly related to the previous ones. It consists of a generalization of Lemma 2.2; as an application, it provides a sufficient condition for the uniform convergence of the series (3.6) on a compact set, and, in particular, provides another proof of Theorem 2.1.

We shall now consider a family $\{Y_i(x), x \in S\}$, $i \in I$, of Gaussian processes on S with continuous sample functions; here I represents an arbitrary index set. For each process $Y_i(\cdot)$ there is a probability measure P_i in C defined on the Borel sets. We shall write the P_i-measure of a set in C of the form $\{f : f$ satisfies condition such and such$\}$ as Pr $\{Y_i(\cdot)$ satisfies condition such and such$\}$.

THEOREM 5.1. *Let* $Y_i(\cdot)$, $i \in I$ *be a family of Gaussian processes with means* 0 *on a compact metric space* S. *Suppose that*
  (a) *for some point* z *in* S

$$\lim_{n \to \infty} \sup_i \text{Pr} \{|Y_i(z)| > n\} = 0;$$

  (b) *there is a metric* $d = d(x, y)$ *for* S *whose* ε-*entropy satisfies the exponential bound of Lemma 2.1; and*
  (c) $\sup_i E|Y_i(x) - Y_i(y)|^2 \leq d^2(x, y)$, $x, y \in S$.
Then the family (P_i) *is compact over* C.

  Proof. Each P_i is defined in C by virtue of Condition (c) and Lemma 2.2. Condition (a) implies that there is a set of functions in C bounded at the point z carrying almost all of every measure P_i. In accordance with the weak compactness criterion referred to in §4 it suffices to prove that there is an equicontinuous set in C which carries almost all of every measure P_i. The construction closely follows that in the proof of Strassen's theorem (cf. [1]).

  Let M_n be a finite set of points in S of minimum cardinality such that d-spheres of radius at most $2^{-(n+3)}$ centered at these points cover S; put $L_n = M_1 \cup \cdots \cup M_n$. For each i, let $G_{n,i}$ be the set of random variables

$$\{Y_i(x) - Y_i(y), d(x, y) \leq 2^{-n}, x, y \in L_n\}.$$

Since L_n has at most

$$n \exp(K 2^{2\beta(n+3)})$$

elements, G_{ni} has at most

$$\exp(2K 2^{2\beta(n+3)} + 2n)$$

elements. Every element of G_{ni} has a Gaussian distribution with mean 0 and variance at most equal to

$$\max_{x,y \in G_{ni}} E|Y_i(x) - Y_i(y)|^2 \leq \max_{x,y \in G_{ni}} d^2(x, y) \leq 2^{-2n};$$

thus, for $Z \in G_{ni}$, we have

$$\text{Pr} \{|Z| \geq n^{-2}\} \leq \exp(-2^{2n}/2n^4), \qquad i \in I;$$

therefore,

$$\text{Pr} \left\{ \max_{G_{ni}} |Z| \geq n^{-2} \right\} \leq \exp(2K 2^{2\beta(n+3)} + 2n - 2^{2n}/2n^4).$$

Since $\beta < 1$, the series

$$\sum_{n=1}^{\infty} \Pr\left\{\max_{G_{ni}} |Z| \geqq n^{-2}\right\}$$

converges uniformly in $i \in I$, that is, for every $\delta > 0$ there exists an integer n_0 sufficiently large so that

$$\sum_{n=n_0}^{\infty} \Pr\left\{\max_{G_{ni}} |Z| \geqq n^{-2}\right\} \leqq \delta, \qquad i \in I.$$

Let A be any countable subset of S. We shall prove there exists a sequence $\{d_n\}$ of positive numbers such that

(5.1) $\quad \sup_i \Pr\{\sup\left[|Y_i(x) - Y_i(y)| : d(x, y) \leqq d_n, x, y \in A\right] \geqq 2^{-n}$

$$\text{for some } n > N\} \leqq 2^{-N}$$

for every integer $N > 0$. Let $\delta > 0$ be arbitrary; choose n_0 so large that

$$\sum_{n=n_0}^{\infty} n^{-2} < \delta/3; \qquad \sup_i \sum_{n=n_0}^{\infty} \Pr\left\{\max_{G_{ni}} |Z| \geqq n^{-2}\right\} < \delta.$$

For a point x in A, let $L_n(x)$ be an element of L_n such that $d(x, L_n(x)) \leqq 2^{-(n+3)}$; such a point exists by virtue of the definition of the set L_n. If $d(x, y) \leqq 2^{-(n+3)}$, then

$$d(L_n(x), L_n(y)) \leqq d(L_n(x), x) + d(x, y) + d(y, L_n(y))$$

$$\leqq 2^{-(n+3)} + 2^{-(n+3)} + 2^{-(n+3)} \leqq 2^{-n}, \qquad x, y \in A;$$

hence $Y_i(L_n(x)) - Y_i(L_n(y))$ is an element of G_{ni}. For any $x \in A$, it is also true that

$$d(L_n(x), L_{n+1}(x)) \leqq d(L_n(u), u) + d(L_{n+1}(u), u) \leqq 2^{-(n+1)};$$

therefore, $Y_i(L_n(x)) - Y_i(L_{n+1}(x))$ is an element of $G_{n+i, i}$. If the sample function $Y_i(\cdot)$ is outside the set

(5.2) $$\bigcup_{n=n_0}^{\infty} \left\{\max_{G_{ni}} |Z| \geqq n^{-2}\right\},$$

then

$$\sum_{n=n_0}^{\infty} [Y_i(L_{n+1}(x)) - Y_i(L_n(x))] \leqq \sum_{n=n_0}^{\infty} n^{-2} \leqq \delta/3,$$

and so $Y_i(L_n(x))$ converges for $n \to \infty$. Its limit is necessarily equal to $Y_i(x)$ because $Y_i(L_n(x))$ converges to $Y_i(x)$ in probability; in fact $Y_i(L_n(x)) - Y_i(x)$ has mean 0 and variance at most equal to $2^{-2(n+3)}$. If $d(x, y) \leqq 2^{-n_0-3}$, then for every $n > n_0$ the following inequalities hold for sample functions outside the set (5.2):

$$|Y_i(x) - Y_i(y)| \leqq \sum_{j=n_0}^{n} |Y_i(L_{j+1}(x)) - Y_i(L_j(x))|$$

$$+ |Y_i(L_{n+1}(x)) - Y_i(x)| + |Y_i(L_{n_0}(x)) - Y_i(L_{n_0}(y))|$$

$$+ \sum_{j=n_0}^{n} |Y_i(L_{j+1}(y)) - Y_i(L_j(y))| + |Y_i(L_{n+1}(y)) - Y_i(y)|$$

$$\leqq |Y_i(L_{n+1}(x)) - Y_i(x)| + |Y_i(L_{n+1}(y)) - Y_i(y)| + \delta.$$

These inequalities hold under passage to the limit as $n \to \infty$:

$$|Y_i(x) - Y_i(y)| \leq \delta;$$

thus, for any $\delta > 0$ there exists a number n_0 sufficiently large so that

$$\sup_i \Pr \{\sup [|Y_i(x) - Y_i(y)| : d(x, y) \leq 2^{-n_0-3}, x, y \in A] > \delta\} \leq \delta.$$

In this inequality let us replace δ by the sequence of numbers $\{2^{-n}\}$, and 2^{-n_0-3} by a suitable sequence $\{d_n\}$; then, for any integer $N > 0$, we obtain

$$\sum_{n=N+1}^{\infty} \sup_i \Pr \{\sup [|Y_i(x) - Y_i(y)| : d(x, y) \leq d_n, x, y \in A] > 2^{-n}\} \leq 2^{-N}.$$

This implies the inequality (5.1).

Let us denote by K the set

$$(5.3) \qquad \bigcap_{n=N+1}^{\infty} \{f : f \in C, \sup [|f(x) - f(y)| : d(x, y) \leq d_n, x, y \in S] \leq 2^{-n}\}.$$

It is a Borel set in C and so belongs to $\mathscr{F} \cap C$; therefore it is representable as $K = F \cap C$, for some $F \in \mathscr{F}$. Since K is a subset of C, it is equal to $F \cap K$ and so $K \in \mathscr{F}$; therefore, K is in a sub-σ-field of $\mathscr{F}$ generated by sets of the form $\{f : f(x) \leq r\}$ where x runs through a *countable* subset of S and r runs through the rationals; therefore, K is equal to a set of the same form as (5.3) but where the parameter set S is replaced by some countable subset A. We conclude from the inequality (4.1) that

$$\inf_i P_i(K) \geq 1 - 2^{-N},$$

and complete the proof.

We remark that the reasoning in the last paragraph shows that C has P_i-inner measure 1 for every i; we have thereby proved Lemma 2.2.

Here are some applications.

1. Let S be a compact set. The partial sums of the series (3.6) are processes with continuous sample functions on S. Let $Y_n(\cdot)$, $n = 1, 2, \ldots$ be the processes corresponding to the partial sums. If they satisfy the conditions of Theorem 5.1, then the corresponding sequence of distributions converges weakly over C. In accordance with the alternatives of Lemma 4.1 and the proof of Theorem 4.1, the series converges uniformly. Theorem 2.2 can be proved in this way.

2. The above method can be used to prove that the series of McKean for Brownian motion on a finite-dimensional space converges uniformly on every compact set. That series represents the Brownian motion $X(\cdot)$ as the sum of independent Gaussian processes with means 0 and with continuous sample functions; hence, the variance of $X(x) - X(y)$ dominates the variances corresponding to the partial sums of the series; therefore, they satisfy the conditions of Theorem 5.1. The weak convergence of the partial sums over C then implies uniform convergence.

S. M. BERMAN

REFERENCES

1. R. M. Dudley, *Gaussian processes on several parameters*, Ann. Math. Statist. **36** (1965), 771–788.

2. ———, *The sizes of compact subsets of Hilbert space and continuity of Gaussian processes*, J. Functional Analysis (to appear).

3. I. M. Gel'fand and N. Ya. Vilenkin, *Generalized functions*. Vol. 4. *Applications of harmonic analysis*, Academic Press, New York, 1964.

4. P. R. Halmos, *Measure theory*, Van Nostrand, New York, 1950.

5. P. Lévy, *Processus stochastiques et mouvement brownien*, 2nd ed., Gauthier-Villars, Paris, 1966.

6. H. P. McKean, *Brownian motion with a several-dimensional time*, Theor. Probability Appl. **4** (1963), 357–378.

7. E. Mourier, *Eléments aléatoires dans un espace de Banach*, Ann. Inst. Henri Poincaré **13** (1952) 161–244.

8. Yu. V. Prokhorov, *The convergence of random processes and limit theorems in probability theory*, Theor. Probability Appl. **1** (1956), 157–214.

9. I. E. Segal, *Distributions in Hilbert space and canonical systems of operators*, Trans. Amer. Math. Soc. **88** (1958), 12–42.

10. P. T. Strait, *Sample function regularity for Gaussian processes with the parameter in a Hilbert space*, Pacific J. Math. **19** (1966), 159–173.

11. A. Tortrat, *Lois de probabilité sur en espace topologique complètement regulier et produits infinis à termes indépendants dans un groupe topologique*, Ann. Inst. Henri Poincaré **1** (B) (1965), 217–237.

NEW YORK UNIVERSITY,
    NEW YORK, NEW YORK