PROCEEDINGS *of the* FIFTH BERKELEY SYMPOSIUM ON MATHEMATICAL STATISTICS AND PROBABILITY

Held at the Statistical Laboratory
University of California
June 21–July 18, 1965
and
December 27, 1965–January 7, 1966

with the support of
University of California
National Science Foundation
National Institutes of Health
Air Force Office of Scientific Research
Army Research Office
Office of Naval Research

VOLUME II
PART 1

CONTRIBUTIONS TO PROBABILITY THEORY

EDITED BY LUCIEN M. LE CAM
AND JERZY NEYMAN

UNIVERSITY OF CALIFORNIA PRESS
BERKELEY AND LOS ANGELES
1967

ABSTRACT WIENER SPACES

LEONARD GROSS
CORNELL UNIVERSITY

1. Introduction

Advanced integral calculus in infinite dimensions was initiated and developed by R. H. Cameron, W. T. Martin, and their associates in a sequence of papers beginning in 1944.

The underlying space for the integral calculus was the Banach space C consisting of the continuous functions on [0, 1] which vanish at zero. The space C carries the probability measure induced by a one-dimensional Brownian motion. The resulting measure space, generally known as Wiener space, has topological, linear, and measure theoretic structures which are well related to one another for the purposes of analysis over C.

The subset C' consisting of the absolutely continuous functions in C with square integrable derivative forms a Hilbert space with respect to the inner product $(x, y) = \int_0^1 x'(t)y'(t)\,dt$. Here a prime denotes derivative. Although C' is a set of Wiener measure zero, the Euclidean structure of this Hilbert space determines the form of the formulas developed by the above authors, and, to a large extent, also the nature of the hypotheses of their theorems. However, it only became apparent with the work of I. E. Segal [11], [12], dealing with the normal distribution on a real Hilbert space, that the role of the Hilbert space C' was indeed central, and that in so far as analysis on C is concerned, the role of C itself was auxiliary for many of Cameron and Martin's theorems, and in some instances even unnecessary. Thus Segal's theorem ([12], theorem 3) on the transformation of the normal distribution under affine transformations, which is formulated for an arbitrary real Hilbert space H, extends and clarifies the corresponding theorem of Cameron and Martin [1], [2] when H is specialized to C'. This is an extreme case in which consideration of the Banach space structure of C, as opposed to merely the Hilbert space structure of C', contributes little or nothing to a proper understanding of the above theorem. For some other theorems, however, the role of C is not negligible, but nevertheless, it is the relation between C and C' which remains important. Specifically, C is the completion of C' with respect to a norm (the sup norm) on C' which is much weaker than the Hilbert norm on C': $\|x\|^2 = \int_0^1 x'(t)^2\,dt$ and enjoys the property of being a measurable norm on C'.

In this paper we shall abstract this relationship by replacing C' by an arbitrary real separable Hilbert space H and the sup norm by its generalization—

This work was supported by N.S.F. contract G. P. 3754.

a measurable norm on H. Our principal result asserts that on the completion of H with respect to a measurable norm the normal distribution becomes countably additive. Our interest in this type of study arose from consideration of regularity theorems for potential theory on a Hilbert space. Such regularity theorems will be studied elsewhere. It turns out that the measurable norms are the right norms with respect to which one should define Hölder conditions. Abstract Wiener spaces thereby enter naturally into this context.

2. Preliminaries

In this section we shall survey some of the basic notions concerning integration over locally convex vector spaces and Hilbert space in particular. Much of the theory of integration over infinite dimensional linear spaces has been surveyed by Prohorov [9]. Consequently, we shall restrict ourselves to material needed in this paper and largely disjoint from that surveyed by Prohorov, and we shall rely on his paper when necessary.

Let $\mathcal{L}$ be a locally convex real linear space and $\mathcal{L}^*$ its topological dual space. For each finite dimensional subspace K of $\mathcal{L}^*$, we denote by π_K the linear map of $\mathcal{L}$ onto the dual space K^* of K given by $\pi_K(x)y = \langle y, x\rangle$ for x in $\mathcal{L}$ and y in K. Let $\mathcal{R}$ be the collection of subsets of $\mathcal{L}$ which have the form $C = \pi_K^{-1}(E)$ where E is a Borel set in K^*. Such a set C will be called a *tame set* (also known as a cylinder set) and will be said to be based on K. The class $\mathcal{R}$ is a ring and the family $\mathcal{S}_K$ of sets in $\mathcal{R}$ which are based on K is a σ-ring.

DEFINITION 1. *A real-valued nonnegative finitely additive function μ on $\mathcal{R}$ is called a cylinder set measure on $\mathcal{L}$ if μ is countably additive on each of the σ-rings $\mathcal{S}_K$ and $\mu(\mathcal{L}) = 1$.*

DEFINITION 2. *A tame function on $\mathcal{L}$ is an $\mathcal{R}$ measurable function f such that $f = g \circ \pi_K$ for some finite dimensional subspace $K \subset \mathcal{L}^*$ and function g on K^*. Such a function f is said to be based on K.*

It is not hard to see that if f is a tame function based on K, then f has the form $f(x) = \varphi(y_1(x), \cdots, y_n(x))$ where $y_1, \cdots, y_n$ is any basis of K and φ is a Baire function on R_n.

A cylinder set measure is referred to by Prohorov [9] as a weak distribution which terminology we shall use for an equivalent but somewhat differently formulated concept. If μ is a cylinder set measure on $\mathcal{L}$ and $y_1, \cdots, y_n$ is a finite set of elements in the finite dimensional space $K \subset \mathcal{L}^*$, then $y_1, \cdots, y_n$ may be regarded as random variables on the probability space $(\mathcal{L}, \mathcal{S}_K, \mu)$. Thus when dealing with tame functions on $\mathcal{L}$ which are based on a fixed finite dimensional subspace of $\mathcal{L}^*$, one is in a countably additive situation. It is sometimes necessary to deal with other functions, however, and the first step toward achieving a suitable degree of countable additivity is to realize all the elements of $\mathcal{L}^*$ simultaneously, as random variables on a (countably additive) probability space (Ω, m). More precisely, this means constructing a linear map F from $\mathcal{L}^*$ to the linear space of random variables (that is, measurable functions modulo null functions) over a probability space (Ω, m) with the property that for any finite set of elements $y_1, \cdots, y_n$ in $\mathfrak{L}^*$, the random variables $F(y_1), \cdots, F(y_n)$ on Ω have the same joint distribution as do $y_1, \cdots, y_n$ as random variables over $(\mathfrak{L}, \mathfrak{R}, \mu)$. The existence of such a map F is easily established in a variety of ways (for example, take a Hamel basis e_α of $\mathfrak{L}^*$ and apply Kolmogorov's theorem to the random variables e_α). The space (Ω, m), and the map F, is of course not unique, but any one with the above property will do. Moreover, given the map F, the cylinder set measure μ can clearly be recovered from it. Thus, a cylinder set measure is equivalent to a weak distribution in the following sense of this phrase.

DEFINITION 3. *A weak distribution over $\mathfrak{L}$ is an equivalence class of linear maps F from $\mathfrak{L}^*$ to the space of random variables over a probability space (Ω, m) (depending on F). Two such maps F_1, F_2 are equivalent if for any finite set $y_1, \cdots, y_n$ in $\mathfrak{L}^*$ the joint distribution of $F_j(y_1), \cdots, F_j(y_n)$ in R_n is the same for $j = 1$ or 2.*

This definition is due to I. E. Segal [11].

Of the two equivalent concepts—weak distribution and cylinder set measure—it is sometimes more convenient to use one and sometimes the other, and sometimes it is convenient to use both, as in this paper.

The weak distribution which will be of interest to us in the remainder of this paper is the normal distribution on a real Hilbert space H defined as follows. If F is a representative of the normal distribution, then for each element y in H^*, $F(y)$ is normally distributed with mean zero and variance $\|y\|^2$. As is well known, this implies that if $y_1, \cdots, y_n$ are orthogonal, then $F(y_1), \cdots, F(y_n)$ are stochastically independent. A tame set in H can be described as a set of the form $C = P^{-1}(E)$ where P is a finite dimensional orthogonal projection on H with range L, say, and E is a Borel set in L. The cylinder set measure ν associated with the normal distribution is called Gauss measure on H, and for the above tame set C we have

$$\nu(C) = (2\pi)^{-n/2} \int_E e^{-\|x\|^2/2}\, dx \tag{2.1}$$

where n is the dimension of L. The set function ν is not countably additive on $\mathfrak{R}$ when H is infinite dimensional.

We consider a fixed representative F of the normal distribution. If f is a tame function on H, then $f(x) = \varphi(y_1(x), \cdots, y_n(x))$ as noted above. Then $\tilde{f} = \varphi(F(y_1), \cdots, F(y_n))$ is a random variable on Ω such that $\int_H f\, d\nu = \int_\Omega \tilde{f}\, dm$, as follows readily from the definitions. The map $f \to \tilde{f}$ is an isomorphism from the algebra of continuous (real- or complex-valued) tame functions on H into the algebra of random variables on Ω. The second step in achieving a useful degree of countable additivity is to extend this isomorphism to functions other than tame functions. This is our main objective in the remainder of this section. This isomorphism does not extend to all continuous functions on H in a reasonable manner. However, we shall describe a class of continuous functions to which this isomorphism does extend. Tame functions play a basic role similar to that of simple functions in general measure theory. If for a given function f

on H we wish to give meaning to $\tilde{f}$ as a random variable on Ω, we must approximate f by a sequence f_n of tame functions such that $\tilde{f}_n$ converges in some sense on Ω. Now if f is continuous, then one can manufacture a sequence f_n of continuous tame functions which converge to f on H by taking any sequence P_n of finite dimensional projections which converge strongly to the identity operator, and define $f_n(x) = f(P_n x)$. This appears to be the easiest systematic way of constructing such a sequence. Unfortunately, the sequence $\tilde{f}_n$ of random variables need not converge in probability on Ω. For example, if $f(x) = \exp[i\|x\|^2]$ and the sequence P_n is taken as an increasing sequence, then for $n > m$ we have

$$|f_n(x) - f_m(x)| = |\exp[i\|(P_n - P_m)x\|^2] - 1|, \tag{2.2}$$

and it is readily seen that the probability that this is greater than a given ϵ depends only on the rank of $P_n - P_m$, and in fact, it approaches one as the rank of $P_n - P_m$ goes to infinity. Even when $\tilde{f}_n$ does converge, it may converge to zero though f may be nowhere zero. This is the case with the function $f(x) = \exp[-\|x\|^2]$ for which a simple computation yields $E(f(Px)^\sim) = 3^{-m/2}$ where m is the rank of P and E denotes expectation.

We proceed to describe a class of functions for which the associated sequence $(f \circ P_n)^\sim$ always converges in probability to a random variable $\tilde{f}$ such that the map $f \to \tilde{f}$ has suitable isomorphism properties. All of the following is taken from [4]. Some of the above material has also been surveyed in [6].

DEFINITION 4. *A seminorm $\|x\|_1$ on H is called a measurable seminorm if for every real number $\epsilon > 0$ there exists a finite dimensional projection P_0 such that for every finite dimensional projection P orthogonal to P_0 we have*

$$\text{Prob}\,(\|Px\|_1^\sim > \epsilon) < \epsilon \tag{2.3}$$

where $\|Px\|_1^\sim$ denotes the random variable on Ω corresponding to the tame function $\|Px\|_1$, and Prob *refers to the probability of the indicated event with respect to the probability measure m associated with the normal distribution.*

We note that the condition (2.3) can also be written $\nu(\{x: \|Px\|_1 > \epsilon\}) < \epsilon$ where ν is Gauss measure on H.

A measurable norm is a measurable seminorm which is a norm.

EXAMPLE 1. If A is a trace class operator, that is, nuclear operator, on H and is nonnegative, then $\|x\|_1 = (Ax, x)^{1/2}$ is a measurable seminorm. It is a measurable norm if $Ax = 0$ implies $x = 0$. This type of norm has proven significant in harmonic analysis on a Hilbert space [5], [9], [10].

EXAMPLE 2. Let H be the Hilbert space C' described in the introduction. If $\|x\|_1 = \sup\{|x(t)|: 0 \leq t \leq 1\}$, then $\|x\|_1$ is a measurable norm on C'. The completion of C' in this norm is identifiable with Wiener space C.

Denote by $\mathfrak{F}$ the directed set of finite dimensional projections on H directed by inclusion of the ranges. The significance of measurable seminorms for the proposed extension of the above described injection $f \to \tilde{f}$ is contained in the following two theorems and definition.

THEOREM. *If $\|x\|_1$ is a measurable seminorm on H, then the net $\|Px\|_1^\sim$ converges in probability on Ω as P converges to the identity through $\mathfrak{F}$.*

We denote the limit in the theorem by $\|x\|_1^\sim$.

The topology $\mathfrak{I}_m$ on H determined by all measurable seminorms is called the measurability topology.

DEFINITION 5. *A function f on H is called uniformly continuous near zero in $\mathfrak{I}_m$(u.c.n.0 in $\mathfrak{I}_m$) if there exists a sequence $\|x\|_n$ of measurable seminorms such that $\|x\|_n^\sim$ converges to zero in probability, while f is uniformly continuous in the topology, $\mathfrak{I}_m$ on $\{x: \|x\|_n \leq 1\}$ for each n.*

In a finite dimensional space this definition reduces to ordinary continuity since $\|x\|$ is measurable in that case, and any continuous function is uniformly continuous on each of the sets $\{x: \|x\|/n \leq 1\}$.

THEOREM. *If f is a complex-valued function on H which is u.c.n.0 in $\mathfrak{I}_m$, then the net $(f \circ P)^\sim$ of random variables on Ω converges in probability as $P \to I$ through $\mathfrak{F}$. If $\tilde{f}$ denotes the limit, then $\tilde{f} = 0$ a.e. if and only if $f(x)$ is identically zero on H.*

The map $f \to \tilde{f}$ from the algebra of bounded real functions on H which are u.c.n.0 in $\mathfrak{I}_m$ to random variables is an algebraic isomorphism. Thus the expectation of f with respect to the normal distribution may be defined as the integral of $\tilde{f}$ over Ω:

$$E(f) = \int_\Omega \tilde{f}(w)m(dw). \tag{2.4}$$

The expectation of f thus defined is clearly independent of which representative of the normal distribution is used.

It is a fairly immediate consequence of the definition of measurable seminorm that if there exists a measurable norm on H, then H must be separable.

3. Abstract Wiener spaces

Let H be a real separable Hilbert space and denote by $\|x\|_1$ a measurable norm on H. Let B be the completion of H with respect to $\|\cdot\|_1$. Then B is a Banach space, and H is dense in B. If y is in B^*, the topological dual space of B, then the restriction of y to H is continuous on H, since a measurable norm on H is always weaker than the H norm by corollary 5.4 of [4]. Moreover, if $y = 0$ on H, then $y = 0$ on B. Hence, restriction to H is a one-to-one linear map of B^* into H^*. We shall thus identify B^* with a subset of H^*, but to avoid confusion, we will not identify H^* with H. The space B^* is dense in H^* since B^* separates points of H.

The normal distribution on H induces a weak distribution on B simply by restricting the defining map F to B^*. The weak distribution on B so obtained defines a cylinder set measure μ on the ring $\mathfrak{R}$ of tame sets of B.

THEOREM 1. *Let H be a real separable Hilbert space. Let $\|x\|_1$ be a measurable norm on H and denote by B the completion of H in this norm. Let μ be the cylinder set measure on the ring $\mathfrak{R}$ of tame sets of B induced by the normal distribution on H. Then μ is countably additive on $\mathfrak{R}$.*

LEMMA 1. *Let $\|x\|_1$ be a measurable norm on H. Let $\{a_j\}_{j=0,1,\cdots}$ be an arbitrary* *sequence of strictly positive real numbers. Then there exists a sequence* $\{Q_j\}_{j=0,1,\cdots}$ *of mutually orthogonal finite dimensional projections on* H *with sum equal to the identity operator such that the sum* $\sum_{j=0}^{\infty} a_j \|Q_j x\|_1$ *converges pointwise on* H *to a measurable norm* $\|x\|_2$.

PROOF. We remark that the interest in this lemma lies in the case where the a_j approach $+\infty$.

Let Ω denote the measure space of some representative of the normal distribution on H. From the definition of measurable norm there exists for each integer $n \geq 1$ a finite dimensional projection P_n on H such that

$$\text{Prob}\,(\|Px\|_1^\sim > 1/(a_n 2^n)) < \tfrac{1}{2}^n \tag{3.1}$$

whenever P is a finite dimensional projection orthogonal to P_n. Moreover, the projections P_n may be taken to be increasing and to converge strongly to the identity operator.

Let $Q_0 = P_1$ and $Q_n = P_{n+1} - P_n$ for $n = 1, 2, \cdots$. Then the projections Q_n are mutually orthogonal finite dimensional projections and $\sum_{j=0}^{\infty} Q_j = I$. Moreover, for $n \geq 1$, Q_n is orthogonal to P_n; so

$$\text{Prob}\,(a_n \|Q_n x\|_1^\sim > 2^{-n}) < 2^{-n}, \qquad n = 1, 2, \cdots. \tag{3.2}$$

If $\epsilon > 0$ and $2^{-k} \leq \epsilon$, then

$$\begin{aligned} \text{Prob}\left(\sum_{n=k+1}^{m} a_n \|Q_n x\|_1^\sim > \epsilon\right) &\leq \text{Prob}\left(\sum_{n=k+1}^{m} a_n \|Q_n x\|_1^\sim > \sum_{n=k+1}^{m} 2^{-n}\right) \\ &\leq \sum_{n=k+1}^{m} \text{Prob}\,(a_n \|Q_n x\|_1^\sim > 2^{-n}) \\ &< 2^{-k} \\ &< \epsilon. \end{aligned} \tag{3.3}$$

Hence the series $\sum_{n=0}^{\infty} a_n \|Q_n x\|_1^\sim$ converges in probability. Let h denote the sum.

In view of corollary 4.4 of [4], it suffices to show that the essential lower bound of h is zero in order to show that the series $\sum_{n=0}^{\infty} a_n \|Q_n x\|_1$ converges on H to a measurable seminorm $\|x\|_2$. Let ϵ be a strictly positive real number. For a sufficiently large integer N, we have $\text{Prob}\,(\sum_{n>N} a_n \|Q_n x\|_1^\sim \leq \epsilon/2) > 0$. Let $f = \sum_{n>N} a_n \|Q_n x\|_1^\sim$ and let $a = \text{Prob}\,(f \leq \epsilon/2)$. Let $g = \sum_{n=0}^{N} a_n \|Q_n x\|_1^\sim$ and let $b = \text{Prob}\,(g \leq \epsilon/2)$. Then $b > 0$, since g is a seminorm based on a finite dimensional subspace of H. Since the projections Q_n are mutually orthogonal, the random variables f and g are mutually independent. Moreover, $h = f + g$. Thus,

$$\begin{aligned} \text{Prob}\,(h \leq \epsilon) &= \text{Prob}\,(f + g \leq \epsilon) \\ &\geq \text{Prob}\,(f \leq \epsilon/2 \quad \text{and} \quad g \leq \epsilon/2) \\ &= \text{Prob}\,(f \leq \epsilon/2)\,\text{Prob}\,(g \leq \epsilon/2) \\ &= ab \\ &> 0. \end{aligned} \tag{3.4}$$

Finally we note that $\|x\|_2$ is a norm, for if $x \neq 0$, then for some j, $Q_j x \neq 0$ and consequently, $\|x\|_2 \neq 0$.

LEMMA 2. *Let $\|x\|_1$ be a measurable norm on H and let B be the completion of H in this norm. There exists a measurable norm $\|x\|_2$ on H such that for each real number $r > 0$, the closure in B of the set $S_r = \{x \in H: \|x\|_2 \leq r\}$ is compact in B.*

PROOF. Let $\{a_j\}_{j=0,1,\cdots}$ be a sequence of strictly positive real numbers such that $\sum_{j=0}^{\infty} a_j^{-1}$ is finite. Let $\|x\|_2 = \sum_{j=0}^{\infty} a_j \|Q_j x\|_1$ where the Q_j are those projections given in lemma 1.

It suffices to show that if x_n is a sequence in H with $\|x_n\|_2 \leq r$ for $n = 1, 2, \cdots$, then some subsequence is Cauchy with respect to $\|\cdot\|_1$. Now the restriction of $\|Q_j x\|_1$ to the range K_j of the finite dimensional projection Q_j is a norm on K_j equivalent to the Euclidean norm, and since $\|Q_j x_n\|_1 \leq r a_j^{-1}$ for all n, there is a subsequence of the sequence x_n such that $Q_j x_n$ is Cauchy with respect to $\|\cdot\|_1$. By diagonalization and dropping to a subsequence, we may assume that $Q_j x_n$ is Cauchy for all j. A measurable norm is strongly continuous by corollary 5.4 of [4]. Consequently, $\|x_n - x_m\|_1 \leq \sum_{j=0}^{\infty} \|Q_j(x_n - x_m)\|_1$. Each term of the sum goes to zero as n and m go to infinity and is dominated by $2ra_j^{-1}$. Hence, $\|x_n - x_m\|_1 \to 0$.

PROOF OF THEOREM. The proof follows a by now well-known pattern. The set function μ is countably additive on $\mathfrak{R}$ if and only if it is continuous from below at B. For a cylinder set measure, the measure of a tame set can be approximated from above by open tame sets. Consequently, μ is countably additive if and only if for every covering of B by a sequence of open tame sets T_n there holds $\sum_{n=1}^{\infty} \mu(T_n) \geq 1$. In order for this condition to hold, it is sufficient that for every real number $\epsilon > 0$ there exists a weakly compact set C_ϵ in B such that $\mu(T) < \epsilon$ for any tame set T disjoint from C_ϵ. Indeed, if T_n is a covering of B by a sequence of open tame sets, then since the T_n are also weakly open, a finite number, say $T_1, \cdots, T_N$, cover C_ϵ, and consequently,

$$\begin{aligned} \sum_{n=1}^{\infty} \mu(T_n) &\geq \sum_{n=1}^{N} \mu(T_n) \\ &\geq \mu\left(\bigcup_{n=1}^{N} T_n\right) \\ &= 1 - \mu\left(B - \bigcup_{n=1}^{N} T_n\right) \\ &\geq 1 - \epsilon, \end{aligned} \tag{3.5}$$

which in view of the arbitrariness of ϵ, implies that $\sum_{n=1}^{\infty} \mu(T_n) \geq 1$.

Before proceeding further, we remark that the preceding argument is the basic one used in much of the literature [3], [7], [8], [9], [13] to prove countable additivity of cylinder set measures. When the underlying space B is itself a dual space of a Banach space or nuclear space, the sets C_ϵ are then taken as closed balls in B. However, in the present case, B is not necessarily a dual space, since example 2 of the preceding section shows that Wiener space is a special case of the space B. Our proof from here parallels Wiener's proof [13] of the countable additivity of Wiener measure. The sets C_ϵ will be strongly compact.

Let $\|x\|_2$ be the measurable norm on H whose existence is asserted in lemma 2. Given $\epsilon > 0$, choose r such that Prob $(\|x\|_2^\sim > r) < \epsilon$. Let C_ϵ be the closure in B of $\{x \in H: \|x\|_2 \leq r\}$. By lemma 2, C_ϵ is (strongly) compact in B. Let T be a tame set of B disjoint from C_ϵ and suppose that T is based on the finite dimensional subspace K of B^*. The set K is also a subspace of H^*. Let L be the finite dimensional subspace of H which corresponds to K under the usual isomorphism between H and H^* induced by the inner product on H. Then L is naturally isomorphic to K^*, and in fact, the isomorphism is an orthogonal transformation between these Euclidean spaces and is given by the restriction to L of the map π_K defined in the preceding section.

In particular, Gauss measure in L is carried by $\pi_K|L$ into Gauss measure in K^*. Thus, if $T = \pi_K^{-1}(E')$ where E' is a Borel set in K^* and if E is the unique Borel set in L with $\pi_K(E) = E'$, then we have $T \cap L = E$. Hence, $\mu(T) = \nu(E) = \nu(T \cap L)$ where ν is Gauss measure in L. But $T \cap L$ is disjoint from $C_\epsilon \cap L$. Therefore, $\nu(T \cap L) \leq 1 - \nu(C_\epsilon \cap L)$. Furthermore, $C_\epsilon \cap L \supset \{x \in L: \|x\|_2 \leq r\}$. Denoting by P the projection of H onto L we have $\nu(C_\epsilon \cap L) \geq$ Prob $(\|Px\|_2^\sim \leq r)$. Hence, $\mu(T) = \nu(T \cap L) \leq$ Prob $(\|Px\|_2^\sim > r) \leq$ Prob $(\|x\|_2^\sim > r)$ by theorem 5 of [4]. Thus $\mu(T) < \epsilon$. This concludes the proof of the theorem.

COROLLARY 1. *In the notation of theorem* 1 *let m denote the countably additive extension of μ to the Borel field $\mathcal{S}$ of B. The identity map on B^* regarded as a densely defined map of H^* into random variables over the probability space $(B, \mathcal{S}, m)$ extends to a representative of the normal distribution over H in a unique manner.*

PROOF. If y is in B^*, then by the definition of μ and m, y is a normally distributed random variable over $(B, \mathcal{S}, m)$ with mean zero and variance $\|y\|^2$ where the norm of y used is the H^* norm. Thus the linear map F_0: $B^* \subset H^* \to L^2(B, \mathcal{S}, m)$ is continuous on a dense set in H^*. Its unique continuous extension F to H^* again assigns to each y in H^* a random variable $F(y)$ over $(B, \mathcal{S}, m)$ which is normally distributed with mean zero and variance $\|y\|^2$. This characterizes F as a representative of the normal distribution. Since any representative of the normal distribution on H over $(B, \mathcal{S}, m)$ is continuous from H^* to $L^2(B, \mathcal{S}, m)$, the asserted uniqueness of the extension of the identity map on B^* follows.

REMARK 1. Corollary 1 may be regarded as an abstract extension of the stochastic integral. For if H is specialized to C' as in example 2 of section 2, then a function y in $L^2(0, 1)$ defines an element of $H^* = (C')^*$ by means of $\langle y, x\rangle = \int_0^1 y(t)x'(t)\,dt$. Although this expression defines a continuous linear functional on C if and only if y is of bounded variation, nevertheless, it exists as a stochastic integral for all y in $L^2(0, 1)$.

COROLLARY 2. *Continuing the notation of corollary* 1 *let $\|x\|_2$ be an arbitrary measurable seminorm on H. Let C be the closure in B of the set $\{x \in H: \|x\|_2 \leq r\}$ where r is a positive real number. Then $m(C) \geq$* Prob $(\|x\|_2^\sim \leq r)$.

PROOF. The set C is the closure of a convex set, hence is itself a closed convex set in B. Since B is separable, the complement of C is a countable union of closed balls B_n, which may be obtained, for example, by taking a dense sequence x_n in the complement of C and taking B_n to be the ball centered at x_n, and with radius equal to half the distance from x_n to C. Each ball B_n can be separated from C by a continuous linear functional y_n on B. It follows that C is the intersection of a sequence T_n of tame sets (in fact half spaces), and $B - \bigcap_{n=1}^{N} T_n$ is a tame set disjoint from C. Thus, if $a = \text{Prob}\,(\|x\|_2^{\sim} \leq r)$, then, as shown in the proof of theorem 1, we have $m(B - \bigcap_{n=1}^{N} T_n) \leq 1 - a$. In view of the countable additivity of m, we thus obtain $m(B - C) \leq 1 - a$ upon letting $N \to \infty$, and consequently, $m(C) \geq a$.

REMARK 2. For any separable real Banach space B, there is a real Hilbert space H and measurable norm $\|\cdot\|_1$ on H such that B is (isometrically isomorphic to) the completion of H in this norm. For since B is separable, there exists an increasing sequence of finite dimensional subspaces F_n of B such that F_n is n-dimensional and such that $K = \bigcup_{n=1}^{\infty} F_n$ is dense in B. Let $z_1, z_2, \cdots$ be a basis for K such that $z_1, \cdots, z_n$ is a basis for F_n. Let S be the open unit ball of B. We construct by induction a sequence of positive real numbers α_n and a new basis y_n of the form $y_n = \alpha_n z_n$, $n = 1, 2, \cdots$ such that $\sum_{j=1}^{n} \beta_j y_j$ is in S whenever $\sum_{j=1}^{n} \beta_j^2 \leq 1$. Choose α_1 such that y_1 is in S. Having chosen $\alpha_1, \cdots, \alpha_{n-1}$ such that $\sum_{j=1}^{n-1} \beta_j^2 \leq 1$ implies $\sum_{j=1}^{n-1} \beta_j y_j$ is in S, we observe that the map f: $(\beta_1, \cdots, \beta_n) \to \sum_{j=1}^{n-1} \beta_j y_j + \beta_n z_n$ is continuous from E_n into B and that $f^{-1}(S)$ contains the closed disk D: $\{(\beta_1, \cdots, \beta_n)\colon \sum_1^{n-1} \beta_j^2 \leq 1, \beta_n = 0\}$ and therefore, also a neighborhood of D. In particular, for some positive number $\alpha_n f^{-1}(S)$ contains the closed set $\sum_{j=1}^{n-1} \beta_j^2 + (\beta_n/\alpha_n)^2 \leq 1$. Thus α_n has been satisfactorily chosen. The space K is a pre-Hilbert space in the inner product for which $y_1, y_2, \cdots$ is an orthonormal set. If $\|x\|'$ denotes the norm on K associated with this inner product and $\|x\|_1$ denotes the given B norm, then clearly $\|x\|_1 < \|x\|'$ for x in K. If $\{\beta_j\}$ is a sequence of real numbers such that $\sum_{j=1}^{\infty} \beta_j^2 < \infty$, then the sequence of partial sums of $\sum_{j=1}^{\infty} \beta_j y_j$ is Cauchy in $\|\cdot\|'$ norm, hence also in $\|\cdot\|_1$ norm. Thus the series converges in $\|\cdot\|_1$ norm to an element of B. Hence the completion of K may be identified with a subset H' of B; H' is a separable Hilbert space. Let A be a one-to-one Hilbert-Schmidt operator on H'. Its range H is a Hilbert space in the norm $\|x\| = \|A^{-1}x\|'$, and moreover, H is dense in B. Since $\|x\|_1 \leq \|x\|' = \|Ax\|$ and since $\|x\|'$ is a measurable norm on H, so is $\|x\|_1$, which concludes the proof of our assertion.

In the following we shall use the representative of the normal distribution constructed in corollary 1. Thus for suitable functions g on H, $\tilde{g}$ denotes a random variable on the probability space $(B, \mathcal{S}, m)$. Suppose that f is a tame function on B. Then f has the form $f(x) = \varphi(y_1(x), \cdots, y_n(x))$ where y_j is in B^*, $j = 1, \cdots, n$. Since the y_j are also in H^*, the restriction g of f to H is also a tame function on H. Moreover, $\tilde{g} = f$ since $y_j^{\sim} = y_j$ by definition. This being said it is natural to ask whether $\tilde{g} = f$ for functions f other than tame functions on B where g is again the restriction of f to H. We showed that this is the case for Wiener space ([4], p. 390) where the proof was facilitated by the existence of a suitable sequence of finite dimensional projections converging strongly to the identity operator on Wiener space. It is not known whether such a sequence exists in a general Banach space such as we are dealing with here (see remark 2), and the proof that $\tilde{g} = f$ (as random variables) must be modified somewhat.

COROLLARY 3. *Let f be a continuous real- or complex-valued function on B. Let g be the restriction of f to H. Then g is u.c.n.0 in $\mathfrak{F}_m$ and $\tilde{g} = f$ almost everywhere with respect to m.*

PROOF. Let $\|x\|_2$ be a measurable norm such that for all $r > 0$, $S_r \equiv \{x \in H: \|x\|_2 \leq r\}$ is precompact in B. Then f, and hence g, is uniformly continuous on S_r for each r with respect to $\|x\|_1$. Since $(\|x\|_2/n)^\sim$ goes to zero in probability as $n \to \infty$, g is u.c.n.0 in $\mathfrak{F}_m$. Thus $\tilde{g}$ makes sense as a random variable on $(B, \mathcal{S}, m)$. Let P_k be a sequence of finite dimensional projections on H converging strongly to the identity operator and such that $P_k H \subset B^*$, $k = 1, 2, \cdots$ when H and H^* are identified in the usual way. In this case each P_k extends to a continuous projection Q_k of B into $P_k H$, as may be seen by taking an orthonormal basis $e_1, \cdots, e_n$ of $P_k H$ and noting that the map $x \to \sum_{j=1}^n (x, e_j)e_j$ on H is continuous in $\|\cdot\|_1$. Now $f \circ Q_k$ is a tame function on B whose restriction to H is exactly $f \circ P_k$, which is by definition $g \circ P_k$. Thus by the discussion preceding the corollary, we have $(g \circ P_k)^\sim = f \circ Q_k$.

We cannot assert that $f \circ Q_k(x) \to f(x)$ for each x in B because the Q_k cannot be arranged to converge strongly to the identity operator. It suffices however to show that $f \circ Q_k$ converges to f in probability. Choose n so large that Prob $(\|x\|_2^\sim > n) < \epsilon/3$. Let C be the closure in B of $\{x \in H: \|x\|_2 \leq n\}$. Then C is compact in B, and by corollary 2, $m(C) \geq 1 - (\epsilon/3)$. Since f is uniformly continuous on C, there is a number $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ when x and y are in C and $\|x - y\|_1 \leq \delta$. By corollary 5.1 of [4], $\|(I - P_k)x\|_1^\sim$ converges to zero in probability as $k \to \infty$. Hence, for some integer k_0, the relation $k > k_0$ implies Prob $(\|(I - P_k)x\|_1^\sim > \delta) < \epsilon/3$. Let D_k be the closure in B of $\{x \in H: \|(I - P_k)x\|_1 \leq \delta\}$. It is easily seen that $D_k = \{x \in B: \|(I - Q_k)x\|_1 \leq \delta\}$. Again by corollary 2, $m(D_k) \geq 1 - (\epsilon/3)$ for $k \geq k_0$. Finally, let $C_k = \{x \in B: Q_k x \in C\}$. Since C_k is closed and contains $\{x \in H: \|P_k x\|_2 \leq n\}$ and since Prob $(\|P_k x\|_2^\sim > n) < \epsilon/3$, it follows that $m(C_k) \geq 1 - (\epsilon/3)$. Thus $|f(x) - f(Q_k x)| < \epsilon$ when x is in $C \cap C_k \cap D_k$. Hence, $m(\{x \in B: |f(x) - f(Q_k x)| < \epsilon\}) \geq 1 - \epsilon$ for $k \geq k_0$, which concludes the proof of the corollary.

COROLLARY 4. *The measure m assigns positive measure to open sets in B.*

PROOF. If U is an open set in B, then there exists a nonnegative somewhere positive bounded continuous function f on B with support in U. Since H is dense in B, the restriction g of f to H is somewhere positive on H. By the preceding corollary, $\tilde{g} = f$ a.e. By corollary 5.5 of [4], $\tilde{g}$ is not the zero random variable. Hence, $f > 0$ on a set of positive measure and $m(U) > 0$.

The next corollary is significant for regularity theorems in potential theory over H. However, technically it belongs here.

COROLLARY 5. *Let A be a bounded operator from B into B^*. Denote by i and j*

the injections of H into B and B^* into H^* respectively. If $E = jAi$, then the symmetric part of E (identifying H^* with H) is a trace class operator and the skew symmetric part is of Hilbert-Schmidt type.

PROOF. Let $C = (E + E^*)/2$ and $D = (E - E^*)/2$. The injection i is a compact operator since for some real number r, the unit ball of H is contained in the ball of radius r of the measurable norm constructed in lemma 2, and the latter set is precompact in B. Hence, E is a compact operator and so are C and D. Thus there is an orthonormal basis of H, $e_1, e_2, \cdots$ such that $Ce_n = \lambda_n e_n$ for all n. Now the function $f(x) = \langle Ax, x\rangle$ is a continuous function on B. Hence, by corollary 3, if g is its restriction to H and P_k is the projection of H onto span $(e_1, \cdots, e_n)$, then $(g \circ P_k)^\sim$ converges in probability as $k \to \infty$. But for x in H, $g(x) = (Cx, x)$. Hence, $g \circ P_k(x) = \sum_{n=1}^{k} \lambda_n x_n^2$ where $x_n = (x, e_n)$. Thus if $\xi_n = (x, e_n)^\sim$, then the ξ_n form a sequence of independent normally distributed random variables with mean zero and variance one, and moreover, $(g \circ P_k)^\sim = \sum_{n=1}^{k} \lambda_n \xi_n^2$. Thus $\sum_1^\infty \lambda_n \xi_n^2$ converges in probability, and therefore, with probability one. Moreover, since the basis $e_1, e_2, \cdots$ can be rearranged arbitrarily without affecting the convergence of $(g \circ P_k)^\sim$, it follows that $\sum_{n=1}^\infty \lambda_n \xi_n^2$ also remains convergent after any rearrangement. Hence from the three-series theorem, it follows that $\sum_1^\infty |\lambda_n| < \infty$. Thus C is trace class.

Now in view of the identification of H^* with H, the operator A^2 is meaningful as a bounded operator from B to B^*. Its restriction to H is $(C + D)^2$ whose symmetric part is $C^2 + D^2$. Hence, D^2 is a trace class operator, and consequently, D is of Hilbert-Schmidt type.

REFERENCES

[1] R. H. CAMERON and W. T. MARTIN, "Transformations of Wiener integrals under translations," *Ann. of Math.*, Vol. 45 (1944), pp. 386–396.

[2] ———, "Transformation of Wiener integrals under a general class of linear transformations," *Trans. Amer. Math. Soc.*, Vol. 58 (1945), pp. 184–219.

[3] I. M. GELFAND and N. YA. VILENKIN, "Some applications of harmonic analysis—augmented Hilbert spaces," *Generalized Functions*, Vol. 4, New York, Academic Press, 1964.

[4] L. GROSS, "Measurable functions on Hilbert space," *Trans. Amer. Math. Soc.*, Vol. 105 (1962), pp. 372–390.

[5] ———, "Harmonic analysis on Hilbert space," *Mem. Amer. Math. Soc.*, No. 46 (1963), 62 pp.

[6] ———, "Classical analysis on a Hilbert space," *Analysis in Function Space*, Chapter 4, pp. 51–68, M.I.T. Press, 1964.

[7] L. LE CAM, "Convergence in distribution of stochastic processes," *Univ. California Publ. Statist.*, Vol. 2 (1957), pp. 207–236.

[8] E. MOURIER, "Éléments aléatoires dans un espace de Banach," *Ann. Inst. H. Poincaré*, Vol. 13 (1953), pp. 161–244.

[9] YU. V. PROHOROV, "The method of characteristic functionals," *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1961, Vol. 2, pp. 403–419.

[10] V. SAZONOV, "Remarks on characteristic functionals," *Teor. Verojatnost. i Primenen.*, Vol. 3 (1958), pp. 201–205. (In Russian.)

[11] I. E. SEGAL, "Tensor algebras over Hilbert spaces," *Trans. Amer. Math. Soc.*, Vol. 81 (1956), pp. 106–134.
[12] ———, "Distributions in Hilbert space and canonical systems of operators," *Trans. Amer. Math. Soc.*, Vol. 88 (1958), pp. 12–41.
[13] N. WIENER, "The average value of a functional," *Proc. London Math. Soc.*, Vol. 22 (1922), pp. 454–467.