

MEASURABLE FUNCTIONS ON HILBERT SPACE

BY

LEONARD GROSS⁽¹⁾

1. **Introduction.** Consider a finite dimensional real vector space L , its dual space L^* and a linear mapping F from L^* to the linear space \mathcal{M} of all measurable functions modulo null functions on a probability measure space (S, \mathcal{S}, μ) . If y_1, \dots, y_n is a basis of L^* the set function ν_0 defined on semi-infinite rectangles of L by means of

$$\nu_0(\{x \in L : y_j(x) \leq \alpha_j, j = 1, \dots, n\}) = \mu(\{s \in S : F(y_j)(s) \leq \alpha_j, j = 1, \dots, n\})$$

extends in a well known way to a measure ν on the Borel sets of L , which is independent of the choice of basis y_1, \dots, y_n used in constructing it and depends only on the mapping F . Every probability Borel measure on L arises in this way from such a linear mapping F for if ω is a probability Borel measure then the elements of L^* are measurable functions on the probability space (L, ω) and the identity map on L^* fulfills the role of F , inducing as above a measure ν which may easily be seen to coincide with ω . Clearly two such linear mappings F and F' induce the same measure ν on L if and only if for each finite set y_1, \dots, y_k of elements of L^* (in particular for a basis of L^*) the random variables $F(y_1), \dots, F(y_k)$ on the one hand and $F'(y_1), \dots, F'(y_k)$ on the other hand have the same joint distribution in k -space.

This correspondence between linear mappings F and probability measures ν on L breaks down in case L is not finite dimensional. If L is, say, an infinite dimensional locally convex topological linear space then it is still true that a reasonable probability measure on L gives rise to a linear mapping F from the dual space L^* , of continuous linear functionals on L , to random variables, namely the identity mapping on L^* . However a given linear map from L^* to random variables will not in general induce a measure on L due to the absence of local compactness of L .

By a random variable on a probability space (S, \mathcal{S}, μ) we shall mean an element of the linear space of measurable functions modulo null functions on S .

DEFINITION. A weak distribution on a topological linear space L is an equivalence class of linear maps F from L^* to random variables on a probability space (S, \mathcal{S}, μ) (depending on F) where two such maps F and F' are equivalent

Presented to the Society, October 31, 1959; received by the editors October 18, 1961.

(1) Most of the work on this paper was done while the author was an N.S.F. Fellow.

if for each finite set y_1, \dots, y_k of vectors in L^* , $F(y_1), \dots, F(y_k)$ have the same joint distribution in k -space as $F'(y_1), \dots, F'(y_k)$.

This definition is due to I. E. Segal [7].

Thus a weak distribution on a finite dimensional space is equivalent to the notion of a measure on L while on an infinite dimensional space it is a more general object than a measure on L . Nevertheless much of the usual integration theory in Euclidean spaces can be carried out in a modified form with respect to a weak distribution on an infinite dimensional space. See for example [4; 5; 7; 8; 9]. A *tame* function on L is a function of the form $f(x) = \phi(y_1(x), \dots, y_k(x))$ where y_1, \dots, y_k are in L^* and ϕ is a Baire function on E_k . If F is a representative of a weak distribution on L and

$$f(x) = \phi(y_1(x), \dots, y_k(x))$$

is a same function on L then $f^\sim = \phi(F(y_1), \dots, F(y_k))$ is a random variable and it is not hard to see that f^\sim depends only on the function f and on F and not otherwise on the representation of f in terms of ϕ and y_1, \dots, y_k . The mapping $f \rightarrow f^\sim$ from the algebra of tame functions on L to random variables is a homomorphism. It is not hard to verify that notions such as the integrability of f^\sim , the distribution of f^\sim and the convergence in L_p or in probability of sequences of such functions f_n^\sim depend only on the weak distribution and not on the particular representative F . Thus for a given weak distribution we may consider questions about tame functions involving the above notions of integration theory. We remark that if F is the identity map on L^* then $f^\sim = f$. It is natural and for some purposes necessary to determine what other functions g on L besides tame functions can be made to correspond to random variables g^\sim in such a way that the map $g \rightarrow g^\sim$ is an extension of that for tame functions and is linear and multiplicative while the above intrinsic integration notions such as the distribution, etc., of g^\sim depends only on g and the weak distribution and is independent of other arbitrary choices (such as representative of the weak distribution or of a basis of L).

It is the purpose of this paper to describe a class of continuous functions on L for which such a correspondence with random variables can be established. We shall be concerned only with the case where L is a real Hilbert space and the weak distribution is the canonical normal distribution with variance parameter c , i.e., that weak distribution which is uniquely determined by the condition that $F(y)$ is normally distributed with mean zero and variance $c \|y\|^2$. This weak distribution appears to be the best infinite dimensional analogue of Lebesgue measure in finite dimensional spaces. It is intimately related to the Wiener integral and has been found useful in studying certain operators and concepts that occur in quantum field theory. In two forthcoming papers we shall apply some of the results herein obtained to strengthen a theorem on non-

linear transformations proved in [5] and to the study of harmonic analysis on a Hilbert space.

2. Measurable pseudo-norms. Let H be a real Hilbert space. If f is a tame function on H then as noted in the Introduction the expectation and other intrinsic integration notions applied to the corresponding random variable $f\sim$ does not depend on which representative of the given weak distribution is used to define $f\sim$. Consequently in discussing the convergence in probability or distribution of random variables corresponding to tame functions it is not necessary to specify the representative of the given weak distribution. In fact we shall often leave out the tilde over tame functions f in equations where it is clear that the corresponding random variable $f\sim$ is actually meant. It is clear that a function

on H is a tame function if and only if there is a finite dimensional projection P on H such that $f(x) = f(Px)$ for all x and such that f restricted to the finite dimensional space PH is a Baire function. f is then said to be *based on PH* . If f is based on PH then it is clearly also based on QH whenever $Q \geq P$ and Q is a projection. If f is based on the finite dimensional subspace PH then we note that its expectation with respect to the canonical normal distribution with variance parameter c , which we restrict ourselves to exclusively, is given by

$$(1) \quad E(\tilde{f}) = (2\pi c)^{-n/2} \int_{PH} f(x) \exp[-\|x\|^2/2c] dx$$

when the integral exists where n is the dimension of PH and dx is Lebesgue measure on PH . Finally we remark that a general existence and uniqueness theorem for weak distributions is given in [7] which ensures the existence and uniqueness of a weak distribution satisfying the defining relation for the normal distribution given above in the introduction.

DEFINITION. A pseudo-norm $\|x\|_1$ on a real Hilbert space H is called a measurable pseudo-norm if for every $\varepsilon > 0$ there exists a finite dimensional projection P_0 such that for the tame function $\|Px\|_1$ we have

$$\text{Prob}(\|Px\|_1 \sim > \varepsilon) < \varepsilon$$

whenever P is a finite dimensional projection orthogonal to P_0 where $\text{Prob}(f\sim > \varepsilon)$ refers to the distribution of the random variable $f\sim$.

We shall generally denote pseudo-norms with a subscript.

REMARKS. (1) A pseudo-norm is always continuous on a finite dimensional space so that the function $\|Px\|_1$ is a tame function based on PH and the condition in the definition makes sense.

(2) By applying equation (1) to the function $f(x) = \chi(\|Px\|)$ where χ is the characteristic function of (ε, ∞) it may be seen by changing variables that $\text{Prob}^{(c)}(\alpha\|Px\| > \varepsilon) = \text{Prob}^{(\alpha^2 c)}(\|Px\| > \varepsilon)$ where the superscript on Prob denotes the variance parameter. It follows easily that a pseudo-norm which is measurable

with respect to the normal distribution with variance parameter c is also measurable for any other variance parameter.

(3) It is easy to see that a positive multiple of a measurable pseudo-norm is also a measurable pseudo-norm as is also a sum of two measurable pseudo-norms.

THEOREM 1. *If $\|x\|_1$ is a measurable pseudo-norm on H then the net $\|Px\|_1$ of tame functions converges in probability as P approaches the identity operator through the directed set \mathcal{F} of finite dimensional projections on H . (\mathcal{F} is directed by inclusion of the ranges.) If $\|x\|_1^{\sim}$ denotes the limit random variable then for every $\varepsilon > 0$ $\text{Prob}(\|x\|_1^{\sim} < \varepsilon) > 0$.*

Proof. We may choose an increasing sequence of projections P_n in \mathcal{F} such that $\text{Prob}(\|Px\|_1 > 1/n) < 1/n$ whenever $P \perp P_n$ and P is in \mathcal{F} . If $m > n > k$ then $\text{Prob}(\| \|P_mx\|_1 - \|P_nx\|_1 \| > 1/k) \leq \text{Prob}(\|(P_m - P_n)x\|_1 > 1/k) \leq \text{Prob}(\|(P_m - P_n)x\|_1 > 1/n) \leq 1/n$. Hence $\|P_nx\|_1$ converges in probability. Denote the limit function by g . Then if P is in \mathcal{F} and $P \geq P_n$ we have for $k < n$

$$\begin{aligned} \text{Prob}(|g - \|Px\|_1| > 1/k) &\leq \text{Prob}(|g - \|P_nx\|_1| > 1/k) + \text{Prob}(\| \|P_nx\|_1 - \|Px\|_1 \| > 1/k) \\ &\leq \text{Prob}(|g - \|P_nx\|_1| > 1/k) + \text{Prob}(\|(P - P_n)x\|_1 > 1/n) \\ &\leq \text{Prob}(|g - \|P_nx\|_1| > 1/k) + 1/n. \end{aligned}$$

Hence not only the sequence $\|P_nx\|_1$ but also the whole net $\|Px\|_1$ converges in probability to g .

It remains to show that the essential lower bound of g is zero. Now for any finite dimensional projection $P \{x \in PH : \|Px\|_1 < \varepsilon/2\}$ is a nonempty open convex set in the finite dimensional subspace PH and therefore $\text{Prob}(\|Px\|_1 < \varepsilon/2) > 0$. Let $\varepsilon > 0$ be a continuity point of the distribution function of g and choose n so large that $1/n < \varepsilon/2$ and we may also assume $n > 2$. Let $k > n$. Since $\|(P_k - P_n)x\|_1$ and $\|P_nx\|_1$ are based on orthogonal subspaces they determine independent random variables. Hence

$$\begin{aligned} \text{Prob}(\|P_kx\|_1 < \varepsilon) &\geq \text{Prob}(\|(P_k - P_n)x\|_1 + \|P_nx\|_1 < \varepsilon) \\ &\geq \text{Prob}(\|(P_k - P_n)x\|_1 < \varepsilon/2, \|P_nx\|_1 < \varepsilon/2) \\ &\geq \text{Prob}(\|(P_k - P_n)x\|_1 < \varepsilon/2) \text{Prob}(\|P_nx\|_1 < \varepsilon/2) \\ &\geq 1/2 \text{Prob}(\|P_nx\|_1 < \varepsilon/2). \end{aligned}$$

Since $\text{Prob}(g < \varepsilon) = \lim_{k \rightarrow \infty} \text{Prob}(\|P_kx\|_1 < \varepsilon)$ the theorem follows.

NOTATION. We denote by H_m the topology on H determined by the measurable pseudo-norms. By virtue of Remark(3) the basic open neighborhoods of zero are the unit spheres of measurable pseudo-norms.

DEFINITION. A function f on H is called uniformly continuous near zero in H_m (u.c.n. 0 in H_m) if there exists a sequence $\|x\|_n$ of measurable pseudo-norms such that $\|x\|_n \rightsquigarrow$ converges to zero in probability while f is uniformly continuous in the topology H_m on the unit sphere of $\|x\|_n$ for each n .

REMARK. If H is finite dimensional then the preceding definition is equivalent to ordinary continuity since the Euclidean norm is measurable (so that H_m is the ordinary topology) and a continuous function is uniformly continuous on compact sets.

THEOREM 2. *If f is a complex valued function which is u.c.n. 0 in H_m then the net $f(Px)$ of tame functions converges in probability as P approaches the identity through the directed set \mathcal{F} of finite dimensional projections.*

Proof. The proof is similar to the proof of Theorem 1 of [5]. Given $\varepsilon > 0$ there exists a measurable pseudo-norm $\|x\|_1$ such that $\text{Prob}(\|x\|_1 \rightsquigarrow > 1/2) < \varepsilon/3$ and such that f is uniformly continuous in the topology H_m on the set $S = \{x : \|x\|_1 \leq 1\}$. There exists a measurable pseudo-norm $\|x\|_2$ such that $|f(x) - f(y)| < \varepsilon$ whenever x and y are in S and $\|x - y\|_2 < 1$. Let P_0 be such a large finite dimensional projection that if P is in \mathcal{F} and $P \geq P_0$ then $\text{Prob}(\|Px\|_1 > 1) < \varepsilon/3$ while at the same time $Q \perp P_0$ implies $\text{Prob}(\|Qx\|_2 > 1/2) < \varepsilon/6$ for Q in \mathcal{F} . Now suppose that P and Q are in \mathcal{F} and $P \geq P_0$ and $Q \geq P_0$. Since $\|(P - Q)x\|_2 \leq \|(P - P_0)x\|_2 + \|(Q - P_0)x\|_2$ we have

$$\begin{aligned} \{x : \|(P - Q)x\|_2 \leq 1\} &\supset \{x : \|(P - P_0)x\|_2 \leq 1/2\} \\ &\cap \{x : \|(Q - P_0)x\|_2 \leq 1/2\} \end{aligned}$$

and $\text{Prob}(\|(P - Q)x\|_2 \leq 1) \geq 1 - \varepsilon/3$. But $|f(Px) - f(Qx)| < \varepsilon$ whenever $\|Px\|_1 \leq 1$, $\|Qx\|_1 \leq 1$ and $\|(P - Q)x\|_2 \leq 1$. Since each of these three conditions holds with probability at least $1 - \varepsilon/3$ they hold simultaneously with probability at least $1 - \varepsilon$. Hence $\text{Prob}(|f(Px) - f(Qx)| \geq \varepsilon) \leq \varepsilon$ whenever P and Q bound P_0 . Now let $\varepsilon = 1/n$ and denote by P_n the projection P_0 obtained above for a general ε . The P_n may with no loss of generality be assumed increasing. Then $f(P_n x)$ converges in probability to a function $f \rightsquigarrow$ and as in the proof of Theorem 1 it may be seen that the whole net $\{f \circ P\}$ converges in probability to $f \rightsquigarrow$.

In the following section we shall give some examples of measurable pseudo-norms which may be simply defined in an abstract Hilbert space. In order to investigate more deeply the relation between the normal distribution and the Wiener integral it is useful to establish the measurability of certain pseudo-norms on $L^2(0, 1)$ (see §5). To this end it is convenient to prove first some convergence theorems for measurable pseudo-norms. This we shall do in §4.

3. **A class of measurable pseudo-norms.** Under suitable restrictions on a homogeneous positive polynomial $p(t_1, \dots, t_k)$ of degree $2n$ in k real variables its

$2n$ th root defines a pseudo-norm on E_k . In this section we describe these polynomials and their analogues on a real Hilbert space and obtain from among them some measurable pseudo-norms.

It is convenient to use the terminology of tensors instead of polynomials. Let H be a real Hilbert space. We shall denote by $H^{(n)}$ the Hilbert space of symmetric n -tensors over H and by P_S the projection operator on the space of all n -tensors over H whose range is $H^{(n)}$. If x_1, \dots, x_n are in H then we put $P_S(x_1 \otimes x_2 \otimes \dots \otimes x_n) = x_1 x_2 \dots x_n$ and $x^n = x \otimes \dots \otimes x$ (n factors). An operator A on a Hilbert space is of trace class if trace $[(A^*A)^{1/2}]$ converges on some orthonormal basis. As is known [6] the operators of trace class on a Hilbert space form a Banach space under the norm $\|A\|_1 = \text{trace}[(A^*A)^{1/2}]$.

DEFINITION. A bounded linear operator A from $H^{(n)}$ into $H^{(n)}$ is called *completely symmetric* if for any finite sets of vectors x_1, \dots, x_n and y_1, \dots, y_n in H we have

$$(1) \quad (A x_1 x_2 \dots x_n, y_1 \dots y_n) = (A y_1 x_2 \dots x_n, x_1 y_2 \dots y_n).$$

REMARKS. (1) If A is completely symmetric then it is symmetric. This follows from the fact that vectors of the form $x_1 \dots x_n$ span $H^{(n)}$.

(2) If A is a Hilbert-Schmidt operator on $H^{(n)}$ i.e. trace $A^*A < \infty$ and if w is the tensor in $H^{(n)} \otimes H^{(n)}$ which corresponds to A via the relation $(Au, v) = (w, u \otimes v)$ then w may also be regarded as an element of $H \otimes H \otimes \dots \otimes H$ ($2n$ factors). A is completely symmetric if and only if w is in $H^{(2n)}$. For if A is completely symmetric then $(w, (x_1 \dots x_n) \otimes (x_{n+1} \dots x_{2n}))$ is invariant under all permutations of the x_i , hence $(w, u \otimes v) = (w, P_S(u \otimes v))$ for all $u, v \in H^{(n)}$ where P_S is here the symmetry projection on the $2n$ -tensors over H . Thus $(Au, v) = (P_S w, u \otimes v)$. Since $P_S w$ is in $H^{(2n)} \subset H^{(n)} \otimes H^{(n)}$ and since there is only one w in $H^{(n)} \otimes H^{(n)}$ for which $(Au, v) = (w, u \otimes v)$ it follows that $w = P_S w$. The converse is clear. Much of the following could be phrased in terms of the symmetric $2n$ -tensor w . However operator theoretic properties of A play a dominant role.

THEOREM 3. If A is a completely symmetric operator on $H^{(n)}$ of trace class such that for all x, y in H $(Ax^{n-1}y, x^{n-1}y) \geq 0$ then the function $\|x\|_A = (Ax^n, x^n)^{1/2n}$ is a measurable pseudo-norm on H . Furthermore $E[(\|\cdot\|_A)^{2n}] = c^n 1 \cdot 3 \dots (2n - 1) \text{trace}(A)$ where E denotes expectation with respect to the canonical normal distribution on H with variance parameter c .

LEMMA 3.1. If H is finite dimensional then every completely symmetric operator on $H^{(n)}$ is a finite linear combination of projections of the form $Pv = x^n(v, x^n)$ where x is a unit vector in H .

Proof. The symmetric $2n$ -tensor corresponding as in Remark (2) to the operator $Av = \sum a_j x_j^n(v, x_j^n)$ is $\sum a_j x_j^{2n}$. Thus by Remark (2) it suffices to show that the vectors x^{2n} span $H^{(2n)}$. One may prove easily by induction that the

vectors x^m span $H^{(m)}$. Assuming that w is in $H^{(m)}$ and that $(w, x^m) = 0$ for all x the induction step consists in setting equal to zero the coefficient of s in the expansion $0 = (w, (sx + y)^m) = (w, y^m) + sm(w, xy^{m-1}) + \dots$ thereby obtaining with the help of the induction assumption $(w, P_S(x \otimes u)) = 0$ for all u in $H^{(m-1)}$ and in particular $(w, x_1 x_2 \dots x_m) = 0$ for all x_j in H so that $w = 0$.

The formula in the following corollary is due to Seidman and Friedrichs [3].

COROLLARY 3.1. *If H is a finite dimensional Hilbert space and A is a completely symmetric operator on $H^{(n)}$ and if E is as in the theorem then $E[(Ax^n, x^n)] = c^n 1 \cdot 3 \dots (2n - 1)$ trace A .*

Proof. Writing $Av = \sum a_j x_j^n (v, x_j^n)$ where $x_j, j = 1, \dots, q$ are unit vectors in H we have $E[(Ax^n, x^n)] = \sum a_j E[(x, x_j)^{2n}] = \sum a_j (2\pi c)^{-1/2} \int_{-\infty}^{\infty} t^{2n} \exp[-t^2/2c] dt = c^n 1 \cdot 3 \dots (2n - 1) \sum a_j = c^n 1 \cdot 3 \dots (2n - 1)$ trace A .

LEMMA 3.2. *Let \mathcal{R} be a bounded (in operator norm) set of bounded linear operators on a (real or complex) Hilbert space K . If A is an operator of trace class on K then the map $B \rightarrow BA$ from \mathcal{R} to the Banach space of trace class operators on K is continuous in the strong operator topology on \mathcal{R} .*

Proof. If $S = VC$ is the polar composition of A with $C = (A^*A)^{1/2}$ then the orthogonal complement of the null space of C is separable and has an orthonormal basis x_1, x_2, \dots such that $Cx_j = \lambda_j x_j, j = 1, 2, \dots$. Then $\lambda_j > 0$ and $\sum \lambda_j = \text{tr } C < \infty$. Let $\{y_\alpha\}$ be an orthonormal basis of the null space of C . If $\|B\| \leq r$ for all B in \mathcal{R} and if $D = B - B_0$ then

$$\begin{aligned} \|BA - B_0A\|_1 &= \text{tr} [(DA)^*DA]^{1/2} \\ &= \sum_j ((A^*D^*DA)^{1/2} x_j, x_j) + \sum_\alpha ((A^*D^*DA)^{1/2} y_\alpha, y_\alpha) \\ &\leq \sum_j \|(A^*D^*DA)^{1/2} x_j\| + \sum_\alpha \|(A^*D^*DA)^{1/2} y_\alpha\| \\ &\leq \sum_j (A^*D^*DA x_j, x_j)^{1/2} + \sum_\alpha (A^*D^*DA y_\alpha, y_\alpha)^{1/2} \\ &\leq \sum_j \lambda_j \|DV x_j\|. \end{aligned}$$

If N is so large that $2r \sum_{j>N} \lambda_j < \epsilon/2$ and if B is in the strong neighborhood of B_0 defined by $\|A\| \|(B - B_0)Vx_j\| < \epsilon/2N, j = 1, \dots, N$, then $\|BA - B_0A\|_1 < \epsilon$ which establishes the lemma.

COROLLARY 3.2. *If A is a trace class operator on a (real or complex) Hilbert space K and if P_α is a net of projections which converge strongly to the identity then $P_\alpha AP_\alpha$ converges in trace class norm to A .*

Proof. $\|P_\alpha AP_\alpha - A\|_1 \leq \|P_\alpha AP_\alpha - AP_\alpha\|_1 + \|AP_\alpha - A\|_1 \leq \|P_\alpha A - A\|_1 + \|P_\alpha A^* - A^*\|_1$. Apply Lemma 3.2.

LEMMA 3.3. *If A is a completely symmetric bounded operator on $H^{(n)}$ then*

(Ax^n, x^n) is the $2n$ th power of a pseudo-norm on H if and only if $(Ax^{n-1}y, x^{n-1}y) \geq 0$ for all x, y in H .

Proof. If $\|x\|_A = (Ax^n, x^n)^{1/2n}$ is a pseudo-norm then $\|x + ty\|_A$ is a convex function of t for all x and y hence so is $\|x + ty\|_A^{2n}$. The second derivative of $(A(x + ty)^n, (x + ty)^n)$ with respect to t at $t = 0$ is therefore non-negative i.e. $2n(2n - 1)(Ax^{n-1}y, x^{n-1}y) \geq 0$. Conversely if $(Ax^{n-1}y, x^{n-1}y) \geq 0$ for all x, y then it follows easily that the second derivative of $(A(x + ty)^n, (x + ty)^n)$ is non-negative for all t , that this is therefore a convex function of t , that the set of x for which $(Ax^n, x^n) \leq 1$ is a convex set and that the positive homogeneous function $(Ax^n, x^n)^{1/2n}$ (which is defined because $(Ax^n, x^n) \geq 0$ for all x) has a convex unit sphere and is therefore a pseudo-norm.

Proof of theorem. For each finite dimensional projection Q on H we denote by $Q^{(n)}$ the projection $Q \otimes Q \otimes \dots \otimes Q$ on $H^{(n)}$. As Q approaches the identity through the directed set \mathcal{F} of finite dimensional projections on H , $Q^{(n)}$ converges strongly to the identity on $H^{(n)}$. If A is a trace class operator on $H^{(n)}$ then $Q^{(n)} A Q^{(n)}$ converges to A in trace class norm by Corollary 3.2. Thus if A satisfies the hypotheses of the theorem and if $\varepsilon > 0$ then there exists a finite dimensional projection P_0 on H such that $\|A - P_0^{(n)} A P_0^{(n)}\|_1 < \varepsilon^{2n+1}$. Let P be a finite dimensional projection orthogonal to P_0 then $P^{(n)} P_0^{(n)} = 0$ and $\|P^{(n)} A P^{(n)}\|_1 = \|P^{(n)}(A - P_0^{(n)} A P_0^{(n)})P^{(n)}\|_1 < \varepsilon^{2n+1}$. By Lemma 3.3 $\|x\|_A = (Ax^n, x^n)^{1/2n}$ is a pseudo-norm. Now $P^{(n)} A P^{(n)}$ is a completely symmetric operator on $H^{(n)}$ and for the purpose of evaluating $E[(P^{(n)} A P^{(n)} x^n, x^n)]$ we may regard $P^{(n)} A P^{(n)}$ as an operator on $(PH)^{(n)}$ and apply Corollary 3.1. Thus we have

$$E[\|Px\|_A^{2n}] = c^n 1 \cdot 3 \dots (2n - 1) \operatorname{tr}(P^{(n)} A P^{(n)}) < c^n 1 \cdot 3 \dots (2n - 1) \varepsilon^{2n+1}.$$

Hence $\operatorname{Prob}(\|Px\|_A > \varepsilon) < c^n 1 \cdot 3 \dots (2n - 1) \varepsilon$. Hence $\|x\|_A$ is a measurable pseudo-norm. To derive the final formula of the theorem observe that if P and Q are in \mathcal{F} then $|\|Px\|_A - \|Qx\|_A| \leq \|(P - Q)x\|_A$ and that $E[\|(P - Q)x\|_A^{2n}] = c^n 1 \cdot 3 \cdot 5 \dots (2n - 1) \operatorname{trace} [((P - Q)^{(n)})^2 A]$ which goes to zero as P and Q approach the identity since $(P - Q)^{(n)} \rightarrow 0$ strongly and Lemma 3.2 applies. Thus $\|Px\|_A$ converges to $\|\cdot\|_{\tilde{A}}$ not only in probability but also in L^{2n} . Hence

$$\begin{aligned} E[(\|\cdot\|_{\tilde{A}})^{2n}] &= \lim_{P \rightarrow I} E[\|Px\|_A^{2n}] \\ &= \lim_{P \rightarrow I} c^n 1 \cdot 3 \dots (2n - 1) \cdot \operatorname{trace} P^{(n)} A P^{(n)} = c^n 1 \cdot 3 \dots (2n - 1) \operatorname{trace} A. \end{aligned}$$

EXAMPLES. (1) Let x_1, x_2, \dots be an orthonormal basis of a real Hilbert space H . Let $a_j \geq 0$ be a sequence of real numbers such that $\sum_j a_j < \infty$. Let $A : H^{(n)} \rightarrow H^{(n)}$ be defined by $A = \sum a_j P_j$ where P_j is the one dimensional projection $P_j v = x_j^n(v, x_j^n)$. If $x = \sum_j t_j x_j$ then $\|x\|_A = (\sum_j a_j t_j^{2n})^{1/2n}$ is a measurable pseudo-norm.

(2) If A is a non-negative trace class operator on H then $\|x\|_A = (Ax, x)^{1/2}$ is the typical measurable pseudo-norm used in [5].

4. **General properties of measurable pseudo-norms.** In order to establish that a given pseudo-norm is measurable it is convenient to have some convergence theorems for measurable pseudo-norms. The following theorem and its corollaries will suffice us for the examples of the next section. A limit indicated by $P \rightarrow I$ means the limit as $P \rightarrow I$ through the directed set \mathcal{F} .

THEOREM 4. *Let $\|x\|_n$ be a nondecreasing sequence of pseudo-norms on H which converge for each x in H . Let $\|x\|_0 = \lim_{n \rightarrow \infty} \|x\|_n$. Assume further the following*

- (1) \lim in $\text{Prob}_{P \rightarrow I} \|Px\|_n = h_n$ exists for each n .
- (2) $h = \lim$ in $\text{Prob}_{n \rightarrow \infty} h_n$ exists.
- (3) $\text{Prob}(h \leq \varepsilon) > 0$ for every $\varepsilon > 0$.

Then $\|x\|_0$ is a measurable pseudo-norm and $\|x\|_0^\sim = h$.

LEMMA 4.1. *Let S be a compact convex set in E_n . Let p_1, p_2, p_3 be three parallel hyperplanes in E_n such that p_2 lies between p_1 , and p_3 (i.e. p_2 is in the convex closure of p_1 and p_3). Let $S_i = S \cap p_i$, $i = 1, 2, 3$ and let m denote $n - 1$ dimensional Lebesgue measure. Then $m(S_2) \geq \min \{m(S_1), m(S_3)\}$.*

Proof. This is a simple consequence of the Brunn-Minkowski theorem. (See [1, § 41].)

LEMMA 4.2. *Let F be a subspace of a finite dimensional real Hilbert space K . Denote by μ and μ' Gauss measure⁽²⁾ in K and F respectively. Let B be a convex centrally symmetric set in K and let C be a convex cylinder in K of the form $C = D \times F$ where D is a convex set in F^\perp . If B is contained in C then $\mu(B) \leq \mu(C)\mu'(F \cap B)$. In particular if $C = K$ then $\mu(B) \leq \mu'(F \cap B)$.*

Proof. Let $x \neq 0$ be in F^\perp and denote by P the orthogonal projection on K with range F . Let $r > 0$ and let C_r denote the circular cylinder $\{x : \|Px\| \leq r\}$. If k is the dimension of F then we denote by m, k -dimensional Lebesgue measure. In the $k + 1$ dimensional subspace G spanned by x and F the previous lemma may be applied to the three hyperplanes $-x + F, F$ and $x + F$ and to the convex set $S_r = G \cap B \cap C_r$ to obtain

$$m(F \cap S_r) \geq \min \{m((-x + F) \cap S_r), m((x + F) \cap S_r)\}.$$

(The possible lack of compactness of S_r clearly does not affect the application of Lemma 4.1). Since S_r is centrally symmetric we have $m(F \cap S_r) \geq m((x + F) \cap S_r)$. Hence

(2) Gauss measure on a k -dimensional real space refers to the probability measure $(2\pi)^{-k/2} \exp(-\|x\|^2/2) dx$ where dx is Lebesgue measure.

$$\begin{aligned}
\mu'((x + F) \cap B) &= \mu'((x + F) \cap (G \cap B)) \\
&= (2\pi c)^{-k/2} \int_0^\infty e^{-r^2/2c} d_r m((x + F) \cap S_r) \\
&= (2\pi c)^{-k/2} \int_0^\infty m((x + F) \cap S_r) e^{-r^2/2c} (r/c) dr \\
&\leq (2\pi c)^{-k/2} \int_0^\infty m(F \cap S_r) e^{-r^2/2c} (r/c) dr \\
&\leq \mu'(F \cap B).
\end{aligned}$$

Now let ν be Gauss measure in F^\perp . Then μ is the product of μ' and ν . Now the projection of B on F^\perp is contained in the projection of C on F^\perp which is exactly $C \cap F^\perp = D$. Hence

$$\begin{aligned}
\mu(B) &= \int_{F^\perp} \mu'((x + F) \cap B) d\nu(x) \\
&= \int_D \mu'((x + F) \cap B) d\nu(x) \\
&\leq \int_D \mu'(F \cap B) d\nu(x) \\
&\leq \mu'(F \cap B) \nu(D) = \mu'(F \cap B) \mu(C).
\end{aligned}$$

COROLLARY 4.3. *If $\|x\|_1$ is a pseudo-norm on H and P and Q are in \mathcal{F} and $P \leq Q$ then for all $a > 0$ $\text{Prob}(\|Px\|_1 > a) \leq \text{Prob}(\|Qx\|_1 > a)$.*

Proof. Take $K = QH$, $F = PH$, $B = \{x \in QH : \|Qx\|_1 \leq a\}$. Then the preceding lemma yields $\text{Prob}(\|Qx\|_1 \leq a) = \mu(B) \leq \mu'(F \cap B) = \text{Prob}(\|Px\|_1 \leq a)$.

Proof of theorem. It is clear that $\|x\|_0$ is a pseudo-norm. It suffices to prove that if $\varepsilon > 0$ is a continuity point of the distribution function of h and also of the distribution functions of all the h_n and if δ is given with $0 < \delta < 1$ then there exists a finite dimensional projection P_0 such that $\text{Prob}(\|Px\|_0 > \varepsilon) \leq \delta$ whenever P is in \mathcal{F} and $P \perp P_0$.

Let $\eta = \delta/(3 - \delta)$ and let $a = \text{Prob}(h \leq \varepsilon)$. Now $\text{Prob}(h_m \leq \varepsilon)$ converges to $\text{Prob}(h \leq \varepsilon) = a$ as $m \rightarrow \infty$ and also $\text{Prob}(\|Qx\|_m \leq \varepsilon)$ converges to $\text{Prob}(h_m \leq \varepsilon)$ as Q approaches I through \mathcal{F} for each m . Hence there exists an N and for each $m \geq N$ a finite dimensional projection Q_m such that $|\text{Prob}(\|Qx\|_m \leq \varepsilon) - a| \leq \eta a$ for $m \geq N$ and $Q \geq Q_m$. Let $X(t)$ be the characteristic function of $(-\infty, \varepsilon]$. It is not hard to show that $X(\|Qx\|_N)$ converges to $X(h_N)$ in L^2 as $Q \rightarrow I$ through \mathcal{F} . Hence there exists P_0 in \mathcal{F} which we may and shall take $\geq Q_N$ such that $\|X(\|Qx\|_N) - X(h_N)\|_{L^2}^2 \leq \eta a/2$ for all $Q \geq P_0$ ($Q \in \mathcal{F}$).

Now let P be in \mathcal{F} and let $P \perp P_0$. Let Q be in \mathcal{F} and suppose $Q \geq P_0 + P$

while $m \geq N$. In the space $K = QH$ we consider the following centrally symmetric convex sets

$$\begin{aligned} B_1 &= \{x \in QH : \|P_0x\|_N \leq \varepsilon\}, \\ B_2 &= \{x \in QH : \|x\|_m \leq \varepsilon\}, \\ B_3 &= B_1 \cap B_2, \\ B_4 &= \{x \in QH : \|x\|_N \leq \varepsilon\}. \end{aligned}$$

If μ denotes Gauss measure in K then $\mu(B_1) = \text{Prob}(\|P_0x\|_N \leq \varepsilon)$ and similarly for other probabilities. Let F be the subspace $(Q - P_0)H$ and let μ' be Gauss measure in F . B_1 is a cylinder parallel to F . Since $B_3 \subset B_1$ and B_3 is centrally symmetric we have from Lemma 4.2 $\mu(B_3) \leq \mu(B_1) \mu'(B_3 \cap F)$. As $F \subset B_1$ it follows that $B_3 \cap F = B_2 \cap F$ so that $\mu'(B_3 \cap F) = \mu'(B_2 \cap F) = \text{Prob}(\|(Q - P_0)x\|_m \leq \varepsilon)$. Hence

$$\mu(B_3) \leq \text{Prob}(\|P_0x\|_N \leq \varepsilon) \text{Prob}(\|(Q - P_0)x\|_m \leq \varepsilon).$$

Furthermore $\|x\|_m \geq \|x\|_N$ so that $B_2 \subset B_4$. Thus $\mu(B_2 - B_1) \leq \mu(B_4 - B_1) \leq \mu((B_4 - B_1) \cup (B_1 - B_4)) = E[(X(\|P_0x\|) - X(\|Qx\|_N))^2] \leq \eta a$. Consequently as $\text{Prob}(\|Qx\|_m \leq \varepsilon) = \mu(B_2) = \mu(B_3) + \mu(B_2 - B_1)$ we have

$$(1) \quad \text{Prob}(\|Qx\|_m \leq \varepsilon) \leq \text{Prob}(\|P_0x\|_N \leq \varepsilon) \text{Prob}(\|(Q - P_0)x\|_m \leq \varepsilon) + \eta a$$

for all $Q \geq P_0 + P$ and all $m \geq N$.

Now PH is a subspace of $(Q - P_0)H$ so that from Corollary 4.3 it follows that $\text{Prob}(\|(Q - P_0)x\|_m \leq \varepsilon) \leq \text{Prob}(\|Px\|_m \leq \varepsilon)$. Furthermore if $Q \geq Q_m$ then from the definition of Q_m we have $\text{Pr}(\|Qx\|_m \leq \varepsilon) \geq (1 - \eta)a$ and as $P_0 \geq Q_N$ we have $\text{Pr}(\|P_0x\|_N \leq \varepsilon) \leq (1 + \eta)a$. Making these substitutions in (1) we obtain when $Q \geq Q_m \vee (P_0 + P)$

$$(1 - \eta)a \leq (1 + \eta)a \text{Prob}(\|Px\|_m \leq \varepsilon) + \eta a.$$

Upon cancelling a and collecting terms we obtain

$$\text{Prob}(\|Px\|_m \leq \varepsilon) \geq (1 - 2\eta)/(1 + \eta) = 1 - \delta$$

for $P \perp P_0$ and $m \geq N$. In the finite dimensional space PH the pseudo-norm $\|x\|_0$ has a continuous distribution function with respect to Gauss measure in PH . Since $\|Px\|_m$ converges to $\|Px\|_0$ everywhere in PH it converges in probability and consequently $\text{Prob}(\|Px\|_0 \leq \varepsilon) = \lim_{m \rightarrow \infty} \text{Pr}(\|Px\|_m \leq \varepsilon) \geq 1 - \delta$ whenever $P \perp P_0$. This proves that $\|x\|_0$ is a measurable pseudo-norm.

To conclude the proof of the theorem let $g = \|x\|_0 \tilde{\cdot}$. Observe that if $R \geq P_0$ then $\text{Prob}(\|(R - P_0)x\|_m > \varepsilon) \leq \delta$ for $m \geq N$ and also for $m = 0$. Hence

$$\begin{aligned} \text{Prob}(\|h_m - \|P_0x\|_m > 2\varepsilon) &\leq \lim_{R \rightarrow I} \text{Prob}(\|\|Rx\|_m - \|P_0x\|_m > \varepsilon) \\ &\leq \lim_{R \rightarrow I} \text{Prob}(\|(R - P_0)x\|_m > \varepsilon) \leq \delta. \end{aligned}$$

Similarly for $\|x\|_0$. Hence in virtue of the inequality

$$|g - h| \leq |g - \|P_0x\|_0| + |\|P_0x\|_0 - \|P_0x\|_m| + |\|P_0x\|_m - h_m| + |h_m - h|$$

it follows that $\text{Prob}(|g - h| > 6\varepsilon) \leq 2\delta + \text{Prob}(|\|P_0x\|_0 - \|P_0x\|_m| > \varepsilon) + \text{Prob}(|h_m - h| > \varepsilon)$. The last two terms go to zero as $m \rightarrow \infty$ which establishes that $g = h$.

COROLLARY 4.4. *Let $\|x\|_n$ be a nondecreasing sequence of measurable pseudo-norms. Suppose that $h = \lim$ in $\text{prob}_{n \rightarrow \infty} \|x\|_n \tilde{\sim}$ exists and has the property that $\text{Prob}(h \leq \varepsilon) > 0$ for all $\varepsilon > 0$. Then $\|x\|_0 = \lim_{n \rightarrow \infty} \|x\|_n$ is a measurable pseudo-norm.*

Proof. The corollary will follow immediately from the theorem as soon as we show that $\lim_{n \rightarrow \infty} \|x\|_n$ exists for each x in H . If x_0 is such that $\lim_{n \rightarrow \infty} \|x_0\|_n = \infty$ then also $\|x\|_n$ diverges to infinity on the one dimensional subspace spanned by x_0 except at $x = 0$. Hence if P is the one dimensional projection onto the span of x_0 then for any real number a we have $\text{Prob}(\|Px\|_n \leq a) \rightarrow 0$ as $n \rightarrow \infty$. If Q is a finite dimensional projection dominating P then by Corollary 4.3 we have $\text{Prob}(\|Qx\|_n \leq a) \leq \text{Prob}(\|Px\|_n \leq a)$ and letting $Q \rightarrow I$ we have

$$\text{Prob}(\|x\|_n \tilde{\sim} \leq a) \leq \text{Prob}(\|Px\|_n \leq a)$$

whenever a is a continuity point of the distribution function of $\|x\|_n \tilde{\sim}$. Thus if a is a continuity point of the distribution function of all the random variables $\|x\|_n \tilde{\sim}$ then $\text{Prob}(\|x\|_n \tilde{\sim} \leq a) \rightarrow 0$ as $n \rightarrow \infty$ which contradicts the assumption that the sequence $\|x\|_n \tilde{\sim}$ converges in probability.

COROLLARY 4.5. *Let $\|x\|_0$ be a pseudo-norm on H . Then $\|x\|_0$ is a measurable pseudo-norm if and only if $h = \lim$ in $\text{prob}_{P \rightarrow I} \|Px\|_0 \tilde{\sim}$ exists and has the property that $\text{Prob}(h \leq \varepsilon) > 0$ for every $\varepsilon > 0$.*

Proof. Half of the corollary is the assertion of Theorem 1. The other half follows from Theorem 4 upon considering the sequence $\|x\|_n = \|x\|_0$ which clearly satisfies the conditions of the theorem.

THEOREM 5. *If A is a bounded linear operator on a real Hilbert space H and $\|x\|_1$ is a measurable pseudo-norm on H then the pseudo-norm $\|x\|_2 = \|Ax\|_1$ is measurable and for every $\varepsilon > 0$ $\text{Prob}(\|x\|_2 \tilde{\sim} > \|A\| \varepsilon) \leq \text{Prob}(\|x\|_1 \tilde{\sim} > \varepsilon)$. Furthermore $\|x\|_2 \tilde{\sim} = \lim$ in $\text{prob}_{P \rightarrow I} \|PAx\|_1 \tilde{\sim}$.*

LEMMA 5.1. *Let T be a convex centrally symmetric set in E_n . If B is a symmetric linear transformation on E_n such that $\|Bx\| \geq \|x\|$ for all x then $m(BT \cap S_r) \geq m(T \cap S_r)$ for all $r > 0$ where S_r is the sphere of radius r centered at the origin and m denotes Lebesgue measure.*

Proof. If the lemma holds for B_1 and B_2 and all centrally symmetric convex

sets T then it also holds for $B = B_1 B_2$ for $m((BT) \cap S_r) = m((B_1 B_2 T) \cap S_r) \geq m(B_2 T \cap S_r) \geq m(T \cap S_r)$. It suffices therefore to prove the lemma for a transformation B whose matrix on an orthonormal basis x_1, \dots, x_n is diagonal with diagonal entries $b, 1, 1, \dots, 1$ where $b \geq 1$.

Let l_h be the hyperplane whose equation is $(x, x_1) = h$, and let m' denote $n - 1$ dimensional Lebesgue measure. We note that $m(T \cap S_r) = \int_{-r}^r m'(T \cap S_r \cap l_h) dh$. Let $T_h = T \cap S_r \cap l_h$. Since T and S_r are centrally symmetric it follows that T_{-h} is the reflection of T_h in the origin. Hence $m'(T_h) = m'(T_{-h})$. Let D_h be the convex set generated by T_h and T_{-h} . Then $D_h \subset T \cap S_r$ and $D_h \cap l_h = T_h$ while $D_h \cap l_{-h} = T_{-h}$. Hence by Lemma 4.1 $m'(D_h \cap l_{h/b}) \geq m'(T_h)$. Since $B l_{h/b} = l_h$ we have $S_r \cap BT \cap l_h = S_r \cap B(T \cap l_{h/b}) \supset S_r \cap B(D_h \cap l_{h/b})$. Furthermore it is easy to see that $B(D_h \cap l_{h/b}) \subset S_r$ since D_h is contained in the convex set (a cylinder) generated by $S_r \cap l_h$ and $S_r \cap l_{-h}$. Hence $S_r \cap BT \cap l_h \supset B(D_h \cap l_{h/b})$ and $m'(S_r \cap BT \cap l_h) \geq m'(B(D_h \cap l_{h/b})) = m'(D_h \cap l_{h/b}) \geq m'(T_h)$. Thus $m(BT \cap S_r) = \int_{-r}^r m'(S_r \cap BT \cap l_h) dh \geq \int_{-r}^r m'(T_h) dh = m(T \cap S_r)$.

LEMMA 5.2. *Under the hypotheses of the preceding lemma we have $\mu(T) \leq \mu(BT)$ where μ denotes Gauss measure in E_n .*

Proof.

$$\begin{aligned} (2\pi c)^{n/2} \mu(T) &= \int_0^\infty e^{-r^2/2c} d_r m(T \cap S_r) = \int_0^\infty m(T \cap S_r) e^{-r^2/2c} r/c dr \\ &\leq \int_0^\infty m(BT \cap S_r) e^{-r^2/2c} r/c dr = (2\pi c)^{n/2} \mu(BT). \end{aligned} \quad \text{q.e.d.}$$

Proof of theorem. We shall first establish the following inequality for any finite dimensional projection P and $\varepsilon > 0$.

$$(2) \quad \text{Prob}(\|PAx\|_1 > \|A\| \varepsilon) \leq \text{Prob}(\|Px\|_1 > \varepsilon).$$

By dividing A by $\|A\|$ we may assume $\|A\| = 1$. Let K be the orthogonal complement of the null space of PA . $\|PAx\|_1$ is a tame function based on K . Let $R = \{x \in K : \|PAx\|_1 \leq \varepsilon\}$ and let $T = \{x \in PH : \|x\|_1 \leq \varepsilon\}$. If μ and μ' denote Gauss measure in PH and K respectively then (2) is equivalent to $\mu'(R) \geq \mu(T)$. Now let F be the range of PA . Then PA regarded as a transformation from K to F is nonsingular and we may write its polar composition $PA = CV$ where V is an orthogonal transformation from K onto F and C is a strictly positive transformation on F with $\|C\| \leq 1$. Now $PA(R)$ is the intersection of T with F . Hence if μ'' denotes Gauss measure in F then by Lemmas 4.2 and 5.2 $\mu(T) \leq \mu''(T \cap F) = \mu''(CV(R)) \leq \mu''(C^{-1}CV(R)) = \mu''(V(R)) = \mu'(R)$ since C^{-1} satisfies the hypothesis of Lemma 5.2. This establishes (2).

To show that $\|x\|_2$ is measurable let P_0 be a finite dimensional projection such that $\text{Prob}(\|Px\|_1 > \varepsilon) < \varepsilon$ whenever $P \perp P_0$ and P is in \mathcal{F} . Let Q_0 be the (finite

dimensional) projection onto the orthocomplement of the null space of P_0A . If Q is in \mathcal{F} and Q is orthogonal to Q_0 then the projection P onto the range of AQ is orthogonal to P_0 . Furthermore $PAQ = AQ$. Thus in virtue of Corollary 4.3 and inequality (2) we have $\text{Prob}(\|Qx\|_2 > \|A\|\varepsilon) = \text{Prob}(\|PAQx\|_1 > \|A\|\varepsilon) \leq \text{Prob}(\|PAx\|_1 > \|A\|\varepsilon) \leq \text{Prob}(\|Px\|_1 > \varepsilon) < \varepsilon$. Hence $\|x\|_2$ is measurable.

Now let P'_n be an increasing sequence of projections in \mathcal{F} chosen as in the proof of Theorem 1 so that $\|P'_n x\|_1$ converges to $\|x\|_1 \tilde{\sim}$ in probability and let Q_n be chosen similarly for $\|x\|_2$. Let P''_n be the projection onto the range of AQ_n and let $P_n = P'_n \vee P''_n$. Then $\|P_n x\|_1$ also converges to $\|x\|_1 \tilde{\sim}$ in probability. Furthermore as $P_n AQ_n = AQ_n$ we have

$$\begin{aligned} \text{Prob}(\|Q_n x\|_2 > \|A\|\varepsilon) &= \text{Prob}(\|P_n AQ_n x\|_1 > \|A\|\varepsilon) \leq \text{Prob}(\|P_n Ax\|_1 > \|A\|\varepsilon) \\ &\leq \text{Prob}(\|P_n x\|_1 > \varepsilon) \end{aligned}$$

by a repeated application of Corollary 4.3 and inequality (2). Passing to the limit on n we get the second assertion of the theorem for those ε which are continuity points of the distribution functions of $\|x\|_1 \tilde{\sim}$ and $\|A\|^{-1}\|x\|_2 \tilde{\sim}$ and hence for all ε .

Finally to show that the limit in probability of the net $\|PAx\|_1$ exists and equals the limit h of the net $\|AQx\|_1$, P and Q in \mathcal{F} we observe that

$$\begin{aligned} \text{Prob}(|\|PAx\|_1 - h| > 3\varepsilon) &\leq \text{Prob}(|\|PAx\|_1 - \|AQx\|_1| > 2\varepsilon) + \text{Prob}(|\|AQx\|_1 - h| > \varepsilon). \end{aligned}$$

Furthermore

$$|\|PAx\|_1 - \|AQx\|_1| \leq \|(PA - AQ)x\|_1 \leq \|PA(I - Q)x\|_1 + \|(I - P)AQx\|_1$$

so that

$$\begin{aligned} \text{Prob}(|\|PAx\|_1 - \|AQx\|_1| > 2\varepsilon) &\leq \text{Prob}(\|PA(I - Q)x\|_1 > \varepsilon) + \text{Prob}(\|(I - P)AQx\|_1 > \varepsilon) \\ &\leq \text{Prob}(\|PA(I - Q)x\|_1 > \varepsilon) + \text{Prob}(\|(I - P)x\|_1 \tilde{\sim} > \varepsilon/\|A\|) \end{aligned}$$

where in the last step we have made use of the fact already proven that $\|(I - P)x\|_1$ is a measurable pseudo-norm and $\|AQ\| \leq \|A\|$. Now any measurable pseudo-norm $\|x\|_3$ has the following property: if $\varepsilon > 0$ and $\delta > 0$ and P_0 is in \mathcal{F} and is such that $\text{Prob}(\|Rx\|_3 > \delta) \leq \varepsilon$ whenever R is in \mathcal{F} and $R \perp P_0$ then for any P in \mathcal{F} with $P \geq P_0$ we have

$$\text{Prob}(\|(I - P)x\|_3 \tilde{\sim} > 2\delta) \leq \varepsilon.$$

This follows easily from the relations

$$\|(I - P)x\|_3 \tilde{\sim} = \lim \text{in prob}_{Q \in \mathcal{F}} \|(I - P)Qx\|_3 = \lim \text{in prob}_{Q \in \mathcal{F}} \|(Q - P)x\|_3$$

since $Q \geq P$ implies $Q - P \perp P_0$. Therefore there exists P_0 in \mathcal{F} such that $\text{Prob}(\|(I - P)x\|_1 > \varepsilon/\|A\|) < \varepsilon$ whenever $P \geq P_0$. Thus for $P \geq P_0$ and P in \mathcal{F} we have

$$\begin{aligned} \text{Prob}(\|PAx\|_1 - h > 3\varepsilon) &\leq \varepsilon + \text{Prob}(\|PA(I - Q)x\|_1 > \varepsilon) \\ &\quad + \text{Prob}(\|AQx\|_1 - h > \varepsilon). \end{aligned}$$

Again since $\|PAx\|_1$ is a measurable pseudo-norm the second term on the right of the last inequality goes to zero as $Q \rightarrow I$ through \mathcal{F} while the third does also by the definition of h . This concludes the proof of the theorem.

COROLLARY 5.1. *If A_α is a uniformly bounded net of bounded linear operators on H for which A_α^* converges strongly to A^* and if $\|x\|_1$ is a measurable pseudo-norm on H then $\|A_\alpha x\|_1 \tilde{}$ converges in probability to $\|Ax\|_1 \tilde{}$.*

Proof. Let $M = \sup_\alpha \|A_\alpha\|$. Let P_0 in \mathcal{F} be such that $P \perp P_0$ and P in \mathcal{F} imply $\text{Prob}(\|Px\|_1 > \varepsilon/2M) \leq \delta$. Let $Q \geq P_0$. If B is an operator such that $0 < \|B\| \leq M$ then

$$\begin{aligned} \text{Prob}(\|QBx\|_1 - \|P_0Bx\|_1 > \varepsilon/2) &\leq \text{Prob}(\|(Q - P_0)Bx\|_1 > \varepsilon/2) \\ &\leq \text{Prob}(\|(Q - P_0)x\|_1 > \varepsilon/2 \|B\|) \\ &\leq \text{Prob}(\|(Q - P_0)x\|_1 > \varepsilon/2M) \\ &\leq \delta. \end{aligned}$$

Letting $Q \rightarrow I$ we get $\text{Pr}(\|Bx\|_1 \tilde{} - \|P_0Bx\|_1 > \varepsilon) \leq \delta$. Hence

$$\begin{aligned} \text{Prob}(\|Ax\|_1 \tilde{} - \|A_\alpha x\|_1 \tilde{} > 3\varepsilon) &\leq \text{Prob}(\|Ax\|_1 \tilde{} - \|P_0Ax\|_1 > \varepsilon) + \text{Prob}(\|P_0(A - A_\alpha)x\|_1 > \varepsilon) \\ &\quad + \text{Prob}(\|P_0A_\alpha x\|_1 - \|A_\alpha x\|_1 \tilde{} > \varepsilon) \\ &\leq 2\delta + \text{Prob}(\|x\|_1 \tilde{} > \varepsilon/\|P_0(A - A_\alpha)\|). \end{aligned}$$

Since P_0 is finite dimensional and $\|P_0(A - A_\alpha)\| = \|(A^* - A_\alpha^*)P_0\|$ it follows easily that $\|P_0(A - A_\alpha)\| \rightarrow 0$ as $\alpha \rightarrow \infty$ and therefore the last term of the last inequality goes to zero as α increases. Thus

$$\text{Prob}(\|Ax\|_1 - \|A_\alpha x\|_1 > 3\varepsilon) \leq 3\delta \text{ for sufficiently large } \alpha.$$

REMARK. The appearance of adjoints in the statement of Corollary 5.1 is necessary. Thus if x_1, x_2, \dots is an O.N. basis of H , if $A_n x_j = x_1$ for $j = n$ and equals 0 for $j \neq n$ and if $\|x\|_n = |(x, x_n)|$ then $A_n \rightarrow 0$ strongly and $\|A_n\| \leq 1$ but $\|A_n x\|_1 = \|x\|_n$ does not converge to zero in probability since all the $\|x\|_n$ have the same distribution function.

COROLLARY 5.2. *If P_α and P_β are two nets of finite dimensional projections*

each converging to the identity and $\|x\|_1$ is a measurable pseudo-norm then $\|x\|_1^{\sim} = \lim \text{in prob}_{\alpha \rightarrow \infty} \|P_\alpha x\|_1^{\sim}$. Furthermore $\|(P_\alpha - P_\beta)x\|_1^{\sim}$ converges to zero in probability as $\alpha, \beta \rightarrow \infty$.

Proof. The corollary follows from Corollary 5.1.

COROLLARY 5.3. *If $\{P_\alpha\}_{\alpha \in D}$ is a net of finite dimensional projections converging strongly to the identity and if f is u.c.n. 0 in H_m then $\lim \text{in prob}_{\alpha \rightarrow \infty} f \circ P_\alpha$ exists and equals f^{\sim} .*

Proof. It suffices to prove that $\lim \text{in prob}_{P \in \mathcal{F}, \alpha \in D} |f \circ P_\alpha - f \circ P|$ is zero. This may be proven by the same method used in Theorem 2 once we know that $\lim \text{in prob}_{P \in \mathcal{F}, \alpha \in D} \|(P - P_\alpha)x\|_1$ is zero for any measurable pseudo-norm $\|x\|_1$. The latter is asserted in Corollary 5.2. q.e.d.

COROLLARY 5.4. *If $\|x\|_1$ is a measurable pseudo-norm on H then it is bounded on the unit sphere of H . Furthermore if $\|x\|_n, n = 1, 2, \dots$ is a sequence of measurable pseudo-norms such that $\|x\|_n^{\sim}$ approaches zero in probability then $\|x\|_n$ converges to zero uniformly on the unit sphere of H .*

Proof. Let $a_n = \text{Prob}(\|x\|_n^{\sim} > 1)$. If $\|y\| = 1$ then $Px = (x, y)y$ is a projection. Hence keeping y fixed we have

$$\begin{aligned} \text{Prob}(|(x, y)^{\sim}| \|y\|_n > 1) &= \text{Prob}(\|Px\|_n^{\sim} > 1) \\ &\leq \text{Prob}(\|x\|_n^{\sim} > 1) \\ &\leq a_n. \end{aligned}$$

Now the function $x \rightarrow (x, y)$ has a Gaussian distribution. Let b_n be such that $2(2\pi c)^{-1/2} \int_{b_n}^{\infty} \exp[-t^2/2c] dt = a_n$. Then $\|y\|_n^{-1} \geq b_n$ i.e. $\|y\|_n \leq b_n^{-1}$. Furthermore as $a_n \rightarrow 0, b_n \rightarrow \infty$ and $b_n^{-1} \rightarrow 0$ concluding the proof of the corollary.

COROLLARY 5.5. *If f is u.c.n. 0 in H_m then f^{\sim} is the null function if and only if f is identically zero on H .*

Proof. We note first that if $\|x\|_1$ is any measurable pseudo-norm then the function $g(x) = \|x - x_0\|_1$ is uniformly continuous in H_m and that the associated measurable function g^{\sim} has essential lower bound zero. The latter fact may be proved just as in Theorem 1 by enlarging the projections P_j that occur there so as to include x_0 in their ranges and writing $\|(P_k x) - x_0\|_1 \leq \|(P_n x) - x_0\|_1 + \|(P_k - P_n)x\|_1$ and proceeding as therein described.

Now suppose that $f(x_0) \neq 0$. We may assume $f(x_0) = 3a > 0$. Since f is u.c.n. 0 in H_m there exists a sequence of measurable pseudo-norms converging to zero in probability on the unit ball of each of which f is uniformly continuous in H_m . By Corollary 5.4 one of these pseudo-norms which we shall denote by $\|x\|_2$ has the property that $\|x_0\|_2 \leq 1/2$. Since f is uniformly continuous on $S = \{x : \|x\|_2$

$\leq 1\}$ in the topology H_m there exists another measurable pseudo-norm $\|x\|_3$ such that $|f(x) - f(y)| < a$ whenever $\|x - y\|_3 < 1$ and x and y are in S . Let $\|x\|_1 = \|x\|_2 + \|x\|_3$. Then $\|x\|_1$ is a measurable pseudo-norm. If $\|x - x_0\|_1 < 1/2$ then $\|x\|_3 < 1$ and $\|x - x_0\|_2 < 1$ and hence $|f(x) - f(x_0)| < a$. Put $g(x) = \|x - x_0\|_1$. Thus $f(x) > 2a$ whenever $g(x) < 1/2$. Hence for any P in \mathcal{F} $\text{Prob}(f \circ P > 2a) \geq \text{Prob}(g \circ P < 1/2)$. Since $g \sim$ has essential lower bound zero and $g \circ P \rightarrow g \sim$ in probability the right side of the last inequality is bounded away from zero as $P \rightarrow I$ hence so is the left side. It follows that $\text{Prob}(f \sim > a) > 0$.

5. Examples. (1) The function $\|x\|$ is not a measurable pseudo-norm on H for the distribution of the tame function $\|Px\|$ depends only on the rank of the finite dimensional projection P .

(2) Mere continuity of a function does not ensure that it will define a measurable function by the prescription given in this paper. In fact the function $f(x) = \exp(i\|x\|^2)$ is uniformly continuous on bounded sets whereas the net $f(Px)$, P in \mathcal{F} , does not converge in probability for if $P \geq Q$ then $|f(Px) - f(Qx)| = |\exp(i\|(P - Q)x\|^2) - 1|$. Thus in order for $f(Px)$ to converge in probability it is necessary that for any given $\varepsilon > 0$ there exist a projection P_0 in \mathcal{F} such that $\text{Prob}(|\exp(i\|(P - Q)x\|^2) - 1| > \varepsilon) < \varepsilon$ whenever $P \geq Q \geq P_0$. However since the distribution of $\|(P - Q)x\|^2$ depends only on the rank of the projection $P - Q$ and not otherwise on P or Q we need only find an ε for which the last inequality fails when $\text{rank } P - Q = 1$ in order to conclude that $f(Px)$ does not converge in probability. The existence of such an ε is clear.

(3) Functions of $\|x\|$ behave badly with respect to the normal distribution since $\|x\|$ itself in no reasonable manner defines a measurable function on H (with respect to the normal distribution). This has been pointed out by Segal [9] and follows from the fact that $\|x\|^2 = \sum_j (x_j, x)^2$ where $\{x_j\}$ is an O.N. basis. Any measurable function corresponding to $\|x\|^2$ would be the sum of an infinite series of independent identically distributed random variables. By Kolmogoroff's three-series theorem such a series diverges with probability one. Now consider the function $f(x) = \exp(-\|x\|^2)$. In this case the net $f(Px)$ converges in probability but the limit is zero even though f is not zero. In fact it is easy to show that it converges to zero in L_1 for a simple computation yields $E(f(Px)) = 1/(1 + 2c)^{n/2}$ where $n = \text{rank } P$. Thus as $P \rightarrow I$, $E(f(Px)) \rightarrow 0$. By Corollary 5.5 this cannot happen for a nonzero function which is u.c.n. 0 in H_m .

(4) We shall elaborate here on the relation between certain measurable functions on Wiener space and functions on the space $H = \text{real } L^2(0, 1)$. We denote by \mathcal{C} the Wiener measure space consisting of the continuous functions on $[0, 1]$ which vanish at zero. A representative of the normal distribution on H with variance parameter $c = 1$ is given by the stochastic integral $n(f)(x) = \int_0^1 f(t) dx(t)$. Here f is in H , x is in \mathcal{C} and $n(f)$ is a random variable on \mathcal{C} defined for almost all x in \mathcal{C} . If G is a function on H which is u.c.n. 0 in H_m then by Theorem 2 it

defines a random variable $G \sim F$ on \mathcal{C} . What functions F are obtainable in this way? We shall show that (among others) any function F which is continuous in the sup norm on \mathcal{C} or in the Lip α norm, $0 < \alpha < 1/2$, is equal to $G \sim$ where G is the function defined by $G(f) = F(\int_0^1 f(s) ds)$. G is u.c.n. 0 in H^m . In proving this we shall also construct some more examples of measurable pseudo-norms.

Let \mathcal{C}' be the subset of \mathcal{C} consisting of those absolutely continuous functions x whose derivative x' is in $L^2(0, 1)$. Denote by $t_j, j = 0, 1, \dots, 2^m$ the binary rationals, $t_j = j2^{-m}$. Consider a function F on \mathcal{C} of the form

$$F(x) = \Phi(x(t_1), \dots, x(t_{2^m}))$$

where Φ is a Borel measurable function of 2^m real variables. Let G be the function on H defined by $G(f) = F(\int_0^1 f(s) ds)$. Thus if $\chi_t(s)$ is the characteristic function of $[0, t]$ then $G(f) = \Phi((f, \chi_{t_1}), \dots, (f, \chi_{t_{2^m}}))$ and G is a tame function on H . The corresponding measurable function $G \sim$ on \mathcal{C} is by definition $\tilde{G}(x) = \Phi(n(\chi_{t_1})(x), \dots, n(\chi_{t_{2^m}})(x))$. Now $n(\chi_t)(x) = \int_0^1 \chi_t(s) dx(s) = x(t)$. Hence $\tilde{G}(x) = \Phi(x(t_1), \dots, x(t_{2^m})) = F(x)$. In other words if we identify H with \mathcal{C}' by means of the mapping $f \rightarrow x$ where $x(t) = \int_0^t f(s) ds$ then the above function F is a tame function on the Hilbert space \mathcal{C}' (with inner product $(x, y) = \int_0^1 x'(t)y'(t) dt$) and the corresponding measurable function is just an extension of F from \mathcal{C}' to all of \mathcal{C} . In order to extend this relationship between F and G to other functions F we shall first construct some measurable pseudo-norms on H .

Consider the function $F_m(x) = \sup \{ |x(t_j)| : j = 1, \dots, 2^m \}$. The corresponding function $\|f\|_m = F_m(\int_0^1 f(s) ds)$ is clearly a tame pseudo-norm on H . As $m \rightarrow \infty$ $\|f\|_m$ increases to the limit $\|f\|_0 = \sup \{ |\int_0^t f(s) ds| : 0 \leq t \leq 1 \}$ which is a pseudo-norm on H while the corresponding measurable functions $F_m = \|f\|_m \sim$ converge everywhere on \mathcal{C} to the function

$$F(x) = \sup \{ |x(t)| : 0 \leq t \leq 1 \}.$$

which is well known to have the property that $\text{Prob}(F(x) \leq \epsilon) > 0$ for all $\epsilon > 0$. Hence $\|f\|_0$ is a measurable norm on H by Corollary 4.4 and the corresponding measurable function is $\|f\|_0 \sim = F(x)$.

Other measurable norms which we shall need are as follows. Let $0 < \alpha < 1/2$. Let $F_m^{(\alpha)}(x) = \sup \{ |t_i - t_j|^{-\alpha} |x(t_i) - x(t_j)|, i \neq j, i, j = 0, 1, \dots, 2^m \}$. The same procedure as in the preceding paragraph may be applied to conclude that the norm $\|f\|^{(\alpha)} = \text{Lip } \alpha$ norm of $x(t) = \int_0^t f(s) ds$ is a measurable norm on H .

Let F be a function on \mathcal{C} which is continuous with respect to the uniform topology on \mathcal{C} . Let $G(f) = F(\int_0^1 f(s) ds)$. We shall show that G is u.c.n. 0 in H_m and that $G \sim = F$. We put $\|x\|_0 = \sup_{0 \leq t \leq 1} |x(t)|$ and $\|x\|_\alpha = \sup_{t \neq s} |t-s|^{-\alpha} |x(t) - x(s)|$ for x in \mathcal{C} and x in $\mathcal{C} \cap \text{Lip } \alpha$ respectively. We recall that for any fixed α with $0 < \alpha < 1/2$ almost all x in \mathcal{C} are in $\text{Lip } \alpha$ and at the same time $S_k = \{x \in \mathcal{C} : \|x\|_\alpha \leq k\}$ is compact with respect to $\|x\|_0$. Hence F

is uniformly continuous on S_k with respect to the uniform topology. In particular F is uniformly continuous on $S_k \cap \mathcal{E}'$ with respect to the uniform topology. In other words we have shown that on the unit ball in H of the measurable norm $k^{-1} \|f\|^{(\alpha)}$ G is uniformly continuous with respect to the norm $\|f\|_0$. Since $k^{-1} \|x\|_\alpha$ converges to zero a.e. on \mathcal{E} (hence in probability) as $k \rightarrow \infty$ it follows that G is u.c.n. 0 in the topology H_m . In order to prove that $G \sim F$ we may apply Corollary 5.3 as follows. Let P_m be the projection onto the span of the functions $\chi_{t_1}, \dots, \chi_{t_{2^m}}$. Then $\{P_m\}$ is an increasing sequence of projections on H which converges strongly to the identity. Hence $G \sim \lim \text{in prob. } (G \circ P_m) \sim$. For any x in \mathcal{E} let $x^{(m)}$ be the piecewise linear function with vertices at $t_j, j = 0, 1, \dots, 2^m$ which agrees with x at these points. Let $F_m(x) = F(x^{(m)})$. It is not hard to verify that $F_m(\int_0^{(\cdot)} f(s)ds) = G(P_m f)$. Hence $(G \circ P_m) \sim F_m$. As $x^{(m)} \rightarrow x$ in sup norm (and also in $\text{Lip } \alpha$ for $x \in \text{Lip } \alpha$) it follows that $F_m \rightarrow F$ on \mathcal{E} hence in probability. Thus $G \sim F$.

A similar proof shows that if F is defined on $\mathcal{E} \cap \text{Lip } \beta$ and is continuous in $\text{Lip } \beta$ norm where $0 < \beta < 1/2$ then the function $G(f) = F(\int_0^{(\cdot)} f(s)ds)$ is u.c.n. 0 in H_m and that $G \sim F$. One uses the fact that for $\beta < \alpha < 1/2$ the unit balls of $k^{-1} \|x\|_\alpha$ are compact with respect to $\|x\|_\beta$.

BIBLIOGRAPHY

1. T. Bonnesen and W. Fenchel, *Theorie der konvexen Körper*. Springer, Berlin, 1934.
2. K. O. Friedrichs and H. N. Shapiro, *Integration over Hilbert space and outer extensions*, Proc. Nat. Acad. Sci. U.S.A. **43** (1957), 336–338.
3. K. O. Friedrichs et al., *Integration of functionals*, New York University, Mimeographed notes, 1957.
4. R. K. Gettoor, *On characteristic functions of Banach space valued random variables*, Pacific J. Math. **7** (1957), 885–896.
5. L. Gross, *Integration and nonlinear transformations in Hilbert space*, Trans. Amer. Math. Soc. **94** (1960), 404–440.
6. R. Schatten, *A theory of cross spaces*, Princeton Univ. Press, Princeton, N.J., 1950.
7. I. E. Segal, *Tensor algebras over Hilbert spaces*, Trans. Amer. Math. Soc. **81** (1956), 106–134.
8. ———, *Distributions in Hilbert space and canonical systems of operators*, Trans. Amer. Math. Soc. **88** (1958), 12–41.
9. ———, *Ergodic subgroups of the orthogonal group on a real Hilbert space*, Ann. of Math. **66** (1957), 297–303.

THE INSTITUTE FOR ADVANCED STUDY,
 PRINCETON, NEW JERSEY
 YALE UNIVERSITY,
 NEW HAVEN, CONNECTICUT
 CORNELL UNIVERSITY,
 ITHACA, NEW YORK