CONVERGENCE IN MEASURE AND RELATED RESULTS IN FINITE RINGS OF OPERATORS

BY

A. R. PADMANABHAN

Dedicated to Professor Ganapathy Iyer on his sixtieth birthday

Introduction. The foundations of a noncommutative integration theory were laid by Segal in [6], in connection with investigations in quantum mechanics, operator algebras and harmonic analysis on groups. In the aforesaid disciplines, there arise systems which are (noncommutative) analogues of measurable functions on a conventional measure space, the simplest instances of such systems being the so-called factors of type I_n, e.g., the ring of all complex $n \times n$ matrices together with the trace. If the trace is normalized so as to assume the value unity, on the identity operator, then this system becomes the noncommutative analogue of complex random variables, on a probability space generated by n atoms.

Let G be a locally compact unimodular group (such as, for some fixed n, the group of all real $n \times n$ nonsingular matrices), μ the Haar measure on G, and A the algebra of all bounded, integrable functions on G, with multiplication defined as convolution. The pair (A, μ) provides an example of the sort described in the beginning.

Let $H = L_2(G, \mu)$ where G is the group of all transformations $cx+d$, c, d rational, and μ is the counting measure on G. For f, g in H, and a, b in G, let U_a be defined thus: $U_a f = g$, where $g(b) = f(ba)$. Let M be the ring of all bounded operators which commute with U_a for all a. Every bounded operator A in H is representable in the form of a bounded numerical matrix $A \sim \|\eta_{a,b}\|$, a, b in G. For any A in M, set $\tau(A) = \eta_{e,e}$ (e being the identity of G). Let I denote the identity operator. For arbitrary C, D in M, $\tau(CD) = \tau(DC)$, and $\tau(I) = 1$. M, a factor of type II_1, is the noncommutative analogue of bounded, complex random variables on a nonatomic probability space.

A special case of a type II_1 factor, bearing the appellation "approximately finite factor," arises in a natural way, in the theory of Fermi-Dirac quantization as described in the papers [7], [8], and [9].

The integration theory, developed by Segal in the most general setting, may be epitomized thus. A ring of operators, with a trace defined on some elements thereof is given. The trace is then extended to a wider class via suitable convergence concepts. And, for this enlarged ensemble (whose elements are called integrable) analogues of standard measure-theoretic results are obtained.

In an arbitrary gage space, Segal introduced, and made a fairly extensive study of, the concepts of measurable operators, convergence nearly everywhere, and

Received by the editors September 27, 1966.

integrable and square-integrable operators, besides obtaining extensions for such basic results as the Riesz-Fischer, Radon-Nikodym and Lebesgue monotone convergence theorems and for a reformulation of the Fubini Theorem. Thus a development of the theory of rings, collateral to measure theory was rendered possible by Segal's work.

The kind of work initiated by Segal was further pursued by Stinespring [10]. He defined convergence in measure in gage spaces, studied the interrelation between convergence in measure and convergence nearly everywhere, and proved, inter alia, noncommutative versions of Fatou's Lemma and Fubini's Theorem. Deserving of special mention are his results which state that under some mild restrictions, a continuous function preserves convergence in the L_2-mean and that a sequence $\{T_n\}$ of selfadjoint measurable operators converges in measure to a selfadjoint measurable operator T, if and only if, for any real continuous function Φ with compact support on the real line $\{\Phi(T_n)\}$ converges in measure to $\Phi(T)$.

Drawing freely on the works of Segal and Stinespring, we have, in the present paper, extended certain standard results in probability theory to finite rings.

1. **Summary and preliminaries.** Throughout this paper, the notation and terminology will be the same as those of [6] and [10]. Let (H, α, m) be a gage space in the sense of [6], I the identity operator and m, a gage on α with $m(I)=1$, and which is regular, i.e., for any projection P, $m(P)=0$ implies $P=0$. By an operator we shall always mean an operator measurable with respect to (w.r.t.) α is the sense of [6].

This paper is divided into several sections.

In §2, we prove that convergence in measure is preserved by a continuous function which is expressible as the sum of a finite number of monotonic continuous functions. We indicate a few applications of this result including one to operator-entropy.

In §3, we prove a noncommutative version of Egoroff's Theorem and its converse. We then obtain a necessary and sufficient condition for convergence in measure and convergence nearly everywhere to coincide.

In §4, we introduce the notion of the distribution function of a selfadjoint measurable operator with respect to a faithful state (positive, normal, linear functional) ϕ of α. Let F_n be the distribution function of T_n, $n=1, 2, \ldots$, and F that of T, with respect to ϕ. Let x be any continuity-point of F (i.e., a point at which F is continuous). We prove that if $\{T_n\}$ converges in measure to T, then $F_n(x) \to F(x)$. Let $\phi=m$, and let F_n, F, and x be as above. Then we also show that $\{T_n\}$ converges to T in measure if and only if, for each such x, $\{P_n^x\}$ converges in measure to P^x, where P_n^x and P^x are spectral projections of T_n and T respectively corresponding to the infinite closed interval $(-\infty, x]$. We then deduce a few simple corollaries.

In §5, we prove two dominated convergence theorems, which, in the case of a finite gage space, are stronger than the corresponding results of Stinespring. As applications, we show, in §6, that if $\{T_n\}$ is a sequence of nonnegative square-integrable operators converging in the L_2-mean to an operator T, then the operator-entropy of T_n tends in the L_1-mean to that of T, and for any bounded selfadjoint operator R, the information about T_n contained in R (as defined by Umegaki and Nakamura) tends to the information about T contained in R.

Although the notation and terminology of this paper will be the same as those of the papers [6] and [10], we shall define some concepts and explain some symbols which we shall repeatedly deal with.

Let (H, α, m) be the underlying gage space. For any operator A in α, $\|A\|$ will denote the operator norm of A. For any measurable operator T (not necessarily bounded), $|T|$ will denote $(T^*T)^{1/2}$. The extension of the gage m to the class of all integrable operators will also be denoted by m. The L_1-norm of T, denoted by $\|T\|_1$, is the number $m(|T|)$. The L_2-norm of T, denoted by $\|T\|_2$, is the number $[m(T^*T)]^{1/2}$. For any two projections P and Q in α, $P \vee Q$ and $P \wedge Q$ will denote respectively the lattice-sum and lattice-product of P and Q. A projection and its range will be denoted by the same symbol. For any two measurable operators A and B, $A+B$ and $A \cdot B$ will denote respectively the strong-sum and strong-product of A and B.

A sequence $\{T_n\}$ will be said to *converge in measure* to T, if given any $\varepsilon > 0$, there exists a sequence $\{Q_n\}$ of projections in α, such that $\|(T_n - T)Q_n\| < \varepsilon$ for all n and $m(Q_n) \to 1$.

It will be said to *converge nearly everywhere* to T if, given any $\varepsilon > 0$, there exists a sequence $\{Q_n\}$ of projections in α, such that $\|(T_n - T)Q_n\| < \varepsilon$ for all n and $Q_n \uparrow I$.

It will be said to *converge to T in the L_p-mean* if $\|T_n - T\|_p \to 0$, as $n \to \infty$, $p = 1, 2$, and *converge to T almost uniformly* if, given any $\delta > 0$, one can find a projection P in α with $m(P) \geqq 1 - \delta$ and such that

$$\|(T_n - T)P\| \to 0 \quad \text{as } n \to \infty.$$

The following results of Stinespring, which we shall refer to as $[D_1]$, $[D_2]$, and $[D_3]$ respectively, will be frequently used in this paper.

$[D_1]$ [10, p. 32, Corollary 5.2]. "A sequence $\{A_n\}$ of measurable operators converges in measure to a measurable operator A if and only if, for any $\varepsilon > 0$, $m(R_n^\varepsilon) \to 0$ as $n \to \infty$, where R_n^ε is the spectral projection of $|A_n - A|$ corresponding to (ε, ∞)."

$[D_2]$ [10, p. 33, Theorem 5.5]. "A sequence $\{T_n\}$ of selfadjoint measurable operators converges in measure to a selfadjoint measurable operator T if and only if, for any continuous function σ with compact support on the real line, $\sigma(T_n) \to \sigma(T)$ in the L_2-mean."

$[D_3]$ [10, p. 24]. "Convergence in the L_p-mean implies convergence in measure," $p = 1, 2$.

2. **Preservation of convergence in measure by continuous functions.**

THEOREM 2.1. *Let $\{T_n\}$ be a sequence of operators converging in measure to an operator T. Let R_n^ε denote the spectral projection of $|T_n - T|$ corresponding to the interval (ε, ∞). Then there exists a subsequence $\{n_k\}$ such that $\sum_{k=1}^\infty m(R_{n_k}^\varepsilon) < \infty$.*

Proof. In the case of an arbitrary gage space, it has been proved by Stinespring [10, p. 23] that $\{T_n\}$ has a subsequence which converges nearly everywhere. His methods together with the finiteness of the gage yield the desired result.

THEOREM 2.2. *Let $\{T_n\}$ be a sequence of selfadjoint operators converging in measure to a selfadjoint operator T. Let Φ be a real-valued continuous function, expressible as the sum of a finite number of real and monotonic continuous functions. Then $\{\Phi(T_n)\}$ converges in measure to $\Phi(T)$.*

Proof. *Case* 1. Φ is strictly increasing and continuous, and its range is the whole real line.

Proof. Let $\Phi(T_n)=S_n$ and $\Phi(T)=S$. Then S and S_n are selfadjoint. Let "a" be any continuous function with compact support. Let $a\cdot\Phi$ denote the composite map defined thus: For any real number λ, $(a\cdot\Phi)(\lambda)=a(\Phi(\lambda))$. It is easy to verify that $(a\cdot\Phi)(T_n)=a(\Phi(T_n))=a(S_n)$. Since Φ is strictly increasing and its range is the whole real line, $a\cdot\Phi$ is once again a continuous function with compact support. Since $\{T_n\}$ converges in measure to T, it follows by $[D_2]$ that $\{(a\cdot\Phi)(T_n)\}$ converges in the L_2-mean to $(a\cdot\Phi)(T)$, i.e., $\{a(S_n)\}$ converges in the L_2-mean to $a(S)$. As "a" is arbitrary, by applying $[D_2]$ again, it follows that $\{S_n\}$ converges in measure to S.

Case 2. Φ is continuous and strictly increasing, and its range is a bounded interval.

Proof. Let $a(\lambda)=\lambda$. Let $b(\lambda)=a(\lambda)+\Phi(\lambda)$. Then, by case 1, $\{b(T_n)\}$ converges in measure to $b(T)$, so that $\{b(T_n)-T_n\}$ converges in measure to $b(T)-T$, i.e., $\{\Phi(T_n)\}$ converges in measure to $\Phi(T)$.

Case 3. Φ is strictly increasing and continuous, and its range is an unbounded interval with a finite left-hand endpoint.

Proof. Let $\Phi(0)=k$. Without loss of generality, we may assume $k=0$, as otherwise we may consider $f(\lambda)=\Phi(\lambda)-k$. Define a new function $g(\lambda)$ thus: For $\lambda>0$, $g(\lambda)=\Phi(\lambda)$. For $\lambda=0$, and $\lambda<0$, $g(\lambda)=\lambda$. Let $b(\lambda)=g(\lambda)+\Phi(\lambda)$. By case 1, $\{g(T_n)\}$ converges in measure to $g(T)$ and $\{b(T_n)\}$ to $b(T)$. Hence $\{\Phi(T_n)\}$ converges in measure to $\Phi(T)$. The case where the range of Φ is an interval bounded on the right, but not on the left, can be disposed of similarly. Thus the Theorem has been proved for any strictly increasing continuous function Φ. Now let Φ be continuous and increasing (but not necessarily strictly). Let $g(\lambda)=\Phi(\lambda)+\lambda$. Then, since g is strictly increasing $\{g(T_n)\}$ converges in measure to $g(T)$, so that $\{\Phi(T_n)\}=\{g(T_n)-T_n\}$ converges in measure to $g(T)-T=\Phi(T)$. Similarly the theorem can be proved when Φ is decreasing and continuous. Hence the theorem is true when Φ is a finite linear combination of monotonic continuous functions.

APPLICATIONS. 1. Let S be a selfadjoint operator, P the spectral projection of S corresponding to $[0,\infty)$, and $Q=I-P$. The operator SP will be denoted by S^+ and the operator $-SQ$ by S^-. Let $\{T_n\}$ converge in measure to T. Then one can show that $\{T_n^+\}$ converges in measure to T^+ and $\{T_n^-\}$ in measure to T^-.

Proof. Let $g(\lambda)=\lambda$ $(\lambda\geqq 0)$ and $=0$ $(\lambda<0)$.

Then, by the above theorem, $\{g(T_n)\}$ converges in measure to $g(T)$. But $g(T_n)=T_n^+$ and $g(T)=T^+$. The other part can be proved similarly.

2. Let $\{T_n\}$ be an arbitrary sequence of operators (not necessarily selfadjoint) converging in measure to T. Then $\{|T_n|\}$ converges in measure to $|T|$.

Proof. It is known [10, pp. 28 and 32] that if $\{T_n\}$ converges in measure to T, then $\{T_n^*T_n\}$ converges in measure to T^*T. Let $g(\cdot)$ be defined thus: $g(\lambda)=0$, $\lambda\leqq 0$, and $g(\lambda)=+(\lambda)^{1/2}$, for $\lambda\geqq 0$. Let $S_n=T_n^*T$ and $S=T^*T$. Then $\{g(S_n)\}$ converges in measure to $g(S)$.

3. Let $\{T_n\}$ be a sequence of nonnegative operators. For each n, $-T_n\log T_n$ is called the operator entropy of T_n, and $m(-T_n\log T_n)$ is called the numerical entropy of T_n. Let $\{T_n\}$ converge in measure to T. Then T can be shown to be nonnegative so that $T\log T$ can be defined. Now the function $\lambda\log\lambda$ is expressible as the sum of a finite number of monotonic, continuous functions so that, by the above theorem, the operator-entropy of T_n converges in measure to the operator-entropy of T. Let now $\{T_n\}$ be uniformly bounded, i.e., there exists a positive integer k such that $\|T_n\|\leqq k$ for all n. In this case, there is convergence even in the L_1-mean. Hence $m(-T_n\log T_n)\to m(-T\log T)$. In other words, the entropy function is continuous in bounded sets of α in the topology of convergence in the L_1-mean.

4. Let g be a strictly monotonic and real function, continuous everywhere on the real line. Let $g(0)=0$. Also let g be bounded, i.e., for some positive integer k, $|g(\lambda)|<k$ for all real λ. For any two measurable operators T and S, define $p(T,S)=m(g(|T-S|))$. Then $\{T_n\}$ converges in measure to T if and only if $p(T_n,T)\to 0$ as $n\to\infty$.

Proof. First let $\{T_n\}$ converge in measure to T. This implies that $\{|T_n-T|\}$ converges in measure to zero. Let $S_n=|T_n-T|$. As $\{S_n\}$ converges in measure to zero, it follows by Theorem 2 that $\{g(S_n)\}$ converges in measure to $g(0)=0$. Since $\|g(S_n)\|\leqq k$ for all n, it even follows that $m(g(S_n))\to 0$, i.e., $p(T_n,T)\to 0$.

Conversely, let $p(T_n,T)\to 0$. This implies that $m(g(S_n))\to 0$. As $g(S_n)$ is nonnegative for each n, it follows that $\{g(S_n)\}$ converges to zero in measure. Now, if possible, let $\{S_n\}$ not converge in measure to zero. By [D_1] there exists therefore some $\varepsilon>0$ such that the sequence $\{m(R_n^\varepsilon)\}$ does not converge to zero, where R_n^ε is the spectral projection of S_n corresponding to (ε,∞). Hence there exists at least one subsequence, denoted by $\{m(R_{n_k}^\varepsilon)\}$, which converges to a strictly positive number L. As g is strictly monotonic, $g(\varepsilon)>g(0)=0$. Now

$$\begin{aligned} 0 < g(\varepsilon)L &= g(\varepsilon)\operatorname*{Lt}_{n_k\to\infty} m(R_{n_k}^\varepsilon) \\ &= \operatorname*{Lt}_{n_k\to\infty} m(g(\varepsilon)\cdot R_{n_k}^\varepsilon) \\ &< \operatorname*{Lt}_{n_k\to\infty} m(g(S_n)) \\ &= 0. \end{aligned}$$

Thus $L=0$. And this contradiction shows that $\{S_n\}$ converges in measure to zero.

3. Egoroff's Theorem and the various modes of convergence.

THEOREM 3.1. *A sequence $\{T_n\}$ converges nearly everywhere to T if and only if it converges to T almost uniformly.*

REMARK. This theorem generalizes Egoroff's Theorem and its converse.

Proof. Let $\{T_n\}$ converge to T nearly everywhere. We shall show that it converges to T almost uniformly. Without loss of generality we may take $T=0$. Let $\{\varepsilon_k\}$ be a sequence of positive numbers converging to zero. By the definition of nearly everywhere convergence, corresponding to ε_j $(j=1, 2, \ldots)$ there exists a sequence $\{Q_{jn}\}$ such that $\|T_n Q_{jn}\| < \varepsilon_j$ for all n and for a fixed j, $Q_{jn} \uparrow I$, as $n \to \infty$. Given any positive number $\delta > 0$, one can, in view of the latter property, find a projection Q_{1n_1} such that $m(Q_{1n_1}) > 1-\delta/2$ and similarly a projection Q_2n_2 such that $m(Q_{2n_2}) > 1-\delta/2^2, \ldots$ and in general a projection Q_{jn_j} (from the sequence $\{Q_{jn}\}$), such that $m(Q_{jn_j}) > 1-\delta/2^j$. For convenience, write $Q_j = Q_{j,n_j}$. Let $S = \bigwedge_{j=1}^{\infty} Q_j$. Let $S^{\perp} = I - S$. Then $m(S^{\perp}) = m(\bigvee_{j=1}^{\infty} Q_j^{\perp}) \leqq \sum_j m(Q_j^{\perp}) = \sum_j \delta/2^j = \delta$. Hence $m(S) \geqq 1-\delta$. Let an arbitrary positive number ε be given. Since $\varepsilon_k \to 0$ as $k \to \infty$, $\varepsilon_k < \varepsilon$ for all $k \geqq$ some positive integer L. Hence for any $k \geqq n_L$, $\|T_k S\| \leqq \|T_k Q_{L,n_L}\| \leqq \varepsilon_L < \varepsilon$. And $m(S) > 1-\delta$.

This being true of any arbitrary $\delta > 0$, it follows that $\{T_n\}$ converges to 0 nearly everywhere.

CONVERSE. Let $\{T_n\}$ converge to 0 almost uniformly. We shall show that $\{T_n\}$ converges to 0 nearly everywhere.

Proof. Let δ be a positive number less than 1. In view of our assumption, there exists a sequence $\{S_k\}$ of projections such that $m(S_k) > 1-\delta/2^k$ and for each fixed k, $\|T_n S_k\| \to 0$ as $n \to \infty$. Let $R_1 = \bigwedge_{k=1}^{\infty} S_k$, $R_2 = \bigwedge_{k=2}^{\infty} S_k, \ldots, R_n = \bigwedge_{k=n}^{\infty} S_k, \ldots$. Then $m(R_n^{\perp}) \leqq \sum_{j=n}^{\infty} \delta/2^j$. So $m(R_n) \to 1$. As m is regular, and $R_k < R_{k+1}$ for each k, it follows that $R_n \uparrow I$. Now for each k, $R_k < S_k$. Since for each fixed k, $\|T_n S_k\| \to 0$ as $n \to \infty$, given k, and $\varepsilon > 0$, one can find a positive integer N_k such that $\|T_n S_k\| < \varepsilon$ for all $n \geqq N_k$ $(k=1, 2, 3, \ldots)$. Define a sequence $\{E_k\}$ of projections thus:

$$E_1 = E_2 = \cdots = E_{N_1-1} = 0 \text{ (the zero projection)},$$

$$E_{N_1} = E_{N_1+1} = \cdots = E_{N_1+N_2-1} = R_1,$$

$$E_{N_1+N_2} = E_{N_1+N_2+1} = \cdots = E_{N_1+N_2+N_3-1} = R_2.$$

And, in general,

$$E_{N_1+N_2+\cdots+N_k} = \cdots = E_{N_1+N_2+\cdots+N_{k+1}-1} = R_k, \ldots, \qquad k = 1, 2, \ldots,$$

The sequence $E_n \uparrow I$. For $n < N_1$, $\|T_n E_n\| = 0$ and so is less than ε. For $n \geqq N_1$, n lies between two integers $N_1 + \cdots + N_k$ and $N_1 + N_2 + \cdots + N_{k+1} - 1$ so that $\|T_n E_n\| = \|T_n R_k\| < \|T_n S_k\| < \varepsilon$ (since $R_k \leqq S_k$). Thus for all n, $\|T_n E_n\| < \varepsilon$. For any given $\varepsilon > 0$, the choice of one such sequence $\{E_n\}$ being possible, it follows that $\{T_n\}$ converges nearly everywhere to 0. Hence the theorem is proved.

In what follows a projection P in α will be said to be minimal if, for any projection Q in α, $Q \leqq P$ implies $Q = P$ or $Q = 0$.

THEOREM 3.2. *Convergence in measure and convergence nearly everywhere are equivalent if and only if each projection in* α *contains a minimal projection.*

Proof. Let each projection contain a minimal projection. We shall show that these two notions of convergence coincide. Let $I=P_1+P_2+\cdots+P_n+\cdots$ be some resolution of the identity into a sequence of pairwise orthogonal minimal projections. Let $m(P_i)=\delta_i$. As $\delta_i\geqq 0$, and $\sum_{i=1}^{\infty}\delta_i=1$, we can assume without loss of generality that $\delta_1\geqq\delta_2\geqq\cdots\geqq\delta_n\geqq\cdots$. Let $\{T_n\}$ be a sequence of operators converging in measure to an operator T. We may take $T=0$. By the definition of convergence in measure, given ε there exists a sequence $\{Q_n\}$ of projections such that $m(Q_n)\to 1$ and $\|T_nQ_n\|<\varepsilon$ for all n. As $m(Q_n)\to 1$, there exists a positive integer N_1 such that $m(Q_n)>1-\delta_1/2$ for all $n\geqq N_1$, and similarly a positive integer N_2 such that $m(Q_n)>1-\delta_1/2-\delta_2/2$ for all $n\geqq N_2,\ldots$, and in general a positive integer N_k such that $m(Q_n)>1-\delta_1/2-\delta_2/2\cdots-\delta_k/2$ for all $n\geqq N_k$, k ranging over all positive integral values. Obviously for $n\geqq N_1$, $Q_n\wedge P_1$ is nonnull. But as P_1 is minimal, this implies $P_1\leqq Q_n$ for $n\geqq N_1$. Similarly $P_1\vee P_2\leqq Q_n$ for $n\geqq N_1+N_2$ etc.

$$R_1 = 0 = \cdots = R_{N_1-1},$$
$$R_{N_1} = P_1 = \cdots = R_{N_1+N_2-1},$$
$$R_{N_1+N_2} = P_1P_2 = \cdots = R_{N_1+N_2+N_3-1}, \text{ etc.}$$

As $\sum_{i=1}^{\infty} m(P_i)=1$, it follows that $R_n\uparrow I$ and $\|T_nR_n\|<\varepsilon$ for all n. Hence convergence in measure implies convergence nearly everywhere. Since, in the case of a finite gage space, convergence in measure is always implied by convergence nearly everywhere, it follows that they are equivalent in this case.

CONVERSE. Let convergence in measure and convergence nearly everywhere coincide. We shall show that each projection has to contain a minimal projection. If not, let there exist a nonnull projection P, with the following property: $m(P)=\delta$ and for any β with $0<\beta<\delta$, there exists a projection $R<P$, with $m(R)=\beta$. We can show that this implies that for any n, P can be expressed as the sum of n pairwise orthogonal projections $P_{1n},\ldots,P_{nn}$, such that $m(P_{in})=m(P)/n$, $i=1,2,\ldots,n$. Now clearly the sequence $P_{11}, 2P_{21}, 2P_{22}, 3P_{31}, 3P_{32}, 3P_{33},\ldots, nP_{n1}, nP_{n2},\ldots, nP_{nn},\ldots$ converges in measure to 0. If possible, let this converge nearly everywhere to 0. By the noncommutative version of Egoroff's Theorem, one can find a projection S with $m(S)>1-m(P)$ and such that $\|nP_{ni}S\|\to 0$. Let $\varepsilon<1$. Hence there exists a positive integer N such that $\|nP_{ni}S\|<\varepsilon\cdots(1)$; for all $n\geqq N$. $P\wedge S$ is nonnull, as $m(S)>1-m(P)$. Let x be a unit vector in $P\wedge S$. Then for any $n\geqq N$, $1=\|x\|^2=\|Px\|^2=\|P_{n1}x\|^2+\cdots+\|P_{nn}x\|^2$.

But by (1), $\|P_{ni}x\|^2<\varepsilon^2/n^2$, $i=1,2,\ldots,n$. Hence

$$1 < \varepsilon^2/n^2+\cdots+\varepsilon^2/n^2\ (n \text{ times}) = \varepsilon^2/n < \varepsilon^2 < 1.$$

This contradiction proves the result.

4. **Weak convergence.** Let (X,Γ,P) be a probability space. By a random variable is meant a real, almost everywhere finite-valued Γ-measurable function.

With each random variable ξ, one can associate a probability measure μ defined on the Borel sets of line as $\mu(E)=P[\xi^{-1}(E)]$. This μ is known as the probability measure corresponding to ξ. Let $(-\infty, x]=\{y : -\infty<y\leqq x\}$. The function F defined by $F(x)=\mu\{(-\infty, x]\}$ is called the distribution function of ξ. F is always right-continuous. A point x at which F is also left-continuous is known as a continuity-point of F. It can easily be seen that x is a continuity-point of F if and only if $\mu(x)=0$. The set of continuity-points of a distribution function is known to be dense on the real line.

Now let T be any selfadjoint operator and P_E its spectral projection corresponding to the Borel set E. We shall call the measure μ, defined by $\mu(E)=m(P_E)$, the probability measure associated with T. And the point function F, defined at any point x by $F(x)=\mu\{(-\infty, x]\}$, we shall call the distribution function of T. Following the measure-theoretic case, we define that a sequence $\{T_n\}$ of selfadjoint operators with distribution functions $\{F_n\}$ converges weakly to a selfadjoint operator T with distribution function F if, at every continuity point x of F, $F_n(x)\rightarrow F(x)$. (As before, if μ_n is the probability measure associated with T_n, $n=1, 2, \ldots$, and μ is the measure associated with T, then it is known that a necessary and sufficient condition for weak convergence [1, p. 33] is that $\int_R g\, d\mu_n \rightarrow \int_R g\, d\mu$, where g is any bounded continuous function, and R denotes the real line) or, what is the same, $m(g(T_n))\rightarrow m(g(T))$.

More generally, let σ be any faithful state of α, with $\sigma(I)=1$. Let T be any selfadjoint operator. Since σ is completely additive, using σ one can as before associate a distribution function G with T. This G we shall call the distribution function of T with respect to σ. As σ is faithful, it can easily be seen that a point x is a continuity point of G if and only if the spectral projection of T corresponding to the singleton x is the zero projection. When we say simply the distribution function of an operator, we mean its distribution function with respect to the gage m.

In what follows $\{T_n\}$ will denote a sequence of selfadjoint operators converging in measure to a selfadjoint operator T, F the distribution function of T, and F_n that of T_n $(n=1, 2, \ldots)$. We shall now state and prove four theorems.

THEOREM 4.1. *Let x be an arbitrary continuity point of F. Then $F_n(x)\rightarrow F(x)$ as $n\rightarrow\infty$.*

THEOREM 4.2. *Let σ be any faithful state. Let G be the distribution function of T with respect to σ, and G_n that of T_n with respect to σ; $n=1, 2, \ldots$. Let x be an arbitrary continuity point of G. Then $G_n(x)\rightarrow G(x)$.*

THEOREM 4.3. *Let $\{A_n\}$ and $\{B_n\}$ be two sequences of selfadjoint operators such that $\{A_n-B_n\}$ converges in measure to zero. Let $\{A_n\}$ converge weakly to a selfadjoint operator A (i.e., if H_n is the distribution of A_n, and H that of A, then $H_n(x)\rightarrow H(x)$ at every point x which is a continuity point of H). Then $\{B_n\}$ also converges weakly to A.*

THEOREM 4.4. *Let x be an arbitrary continuity point of F. Let P_n^x, for each n, denote the spectral projection of T_n corresponding to the interval $(-\infty, x]$ and P^x that of T corresponding to the same interval. Then a necessary and sufficient condition for $\{T_n\}$ to converge in measure to T is that, corresponding to each continuity point x of F, $\{P_n^x\}$ converges in measure to P^x.*

REMARKS. Theorems 4.1 and 4.2 are consequences of Theorem 4.4; they are stated separately, however, because the proof of Theorem 4.4 depends on that of Theorem 4.1. Theorem 4.2 is more general than Theorem 4.1 but cannot be proved directly and has only to be deduced from Theorem 4.4.

Proof of Theorem 4.1.

Case 1. Let T_n's be uniformly bounded, i.e., there exists a positive integer k such that $\|T_n\| < k$ for all n. In this case, the proof is exceedingly simple. Clearly for each n, the spectrum of T_n is contained in the closed interval $[-k, k]$. Let μ_n be the probability measure associated with T_n. It is easy to verify that all the μ_n's vanish outside the compact set $[k, k]$. Hence, in this case, to establish weak convergence it suffices to consider continuous functions g with compact support. Now, $\int_R g\, d\mu_n = m(g(T_n))$ and $\int_R g\, d\mu = m(g(T))$. But by the result [$D_2$], $\{g(T_n)\}$ converges in the L_2-mean to $g(T)$; and hence, in particular,

$$m(g(T_n)) \to m(g(T)), \quad \text{i.e.,} \int_R g\, d\mu_n \to \int_R g\, d\mu.$$

Hence we have weak convergence in this case.

Case 2 (*General case*). Let $\{T_n\}$ be an arbitrary sequence of selfadjoint operators converging in measure to a selfadjoint operator T. Let F_n be the distribution function of T_n, and F that of T. Let x be an arbitrary continuity point of F. To prove $F_n(x) \to F(x)$. Corresponding to $(-\infty, x]$, let P_n^x be the spectral projection of T_n, and P^x that of T. Then $F_n(x) = m(P_n^x)$ and $F(x) = m(P^x)$.

REMARK. The proof in the measure-theoretic case, as given in standard textbooks on probability theory, such as [2] and [3, p. 168], does not directly extend to the noncommutative case. The argument given in those books is of the following type: Let (X, β, P) be a probability space. Let f_n and f be two random variables, c any fixed point on the real line, and $\varepsilon > 0$. Let $A = f_n^{-1}\{(-\infty, c]\}$ and $B = f^{-1}\{(c+\varepsilon, \infty)\}$. Let $C = X - B$. Then,

$$A = A(\overline{\quad})B + A(\overline{\quad})C. \tag{1}$$

This is the crucial decomposition on which the proof in the measure-theoretic case hinges. In the general case, let P_n be the spectral projection of T_n corresponding to $(-\infty, x]$, P the spectral projection of T corresponding to $(c+\varepsilon, \infty)$, and $Q = I - P$. The equality $P_n = P_n \wedge P + P_n \wedge Q$ is not in general valid, since the associative law does not in general hold for the lattice of projections. In other words, the crucial decomposition (1) does not generalize to the case of rings of operators. Hence, we shall furnish a proof which is different from the one in the measure-theoretic case.

Let x be any fixed continuity point of F. We shall show that $m(P^x)$ is the only limit point of the bounded sequence $\{m(P_n^x)\}$. (Note that for all n, $0 \leqq m(P_n^x) \leqq 1$.) Let L be a limit point of $\{m(P_n^x)\}$. Hence there exists a subsequence, also denoted by $\{m(P_n^x)\}$, which converges to L. If possible, let $L < m(P^x)$. So, for some $\delta > 0$, one can write $L = m(P^x) - 2\delta$. As x is a continuity point of F, one can find for some $\varepsilon > 0$, a point $x - \varepsilon$ (to the left of x) such that $F(x) - F(x-\varepsilon) < \delta/2$. (Note that if $P^{x-\varepsilon}$ denotes the spectral projection of T corresponding to $(-\infty, x-\varepsilon]$, then $F(x-\varepsilon) = m(P^{x-\varepsilon})$.) Hence $L < m(P^{x-\varepsilon})$. Since $m(P_n^x) \to L$, it follows that there exists a positive integer N_1 such that for all $n \geqq N_1$, $m(P_n^x) < L + \delta/4 < m(P^x) - \delta$. Let $A_n^x = I - P_n^x$. It follows that for

$$(2) \qquad n \geqq N_1, \qquad m(A_n^x \wedge P^{x-\varepsilon}) \geqq \delta.$$

Also, as $\{T_n\}$ converges in measure to T it follows, by a result mentioned above, that if S_n denotes the spectral projection of $|T_n - T|$ corresponding to the interval $[0, \varepsilon/2)$, then $m(S_n) \to 1$ as $n \to \infty$. So there exists a positive integer N_2 such that $m(S_n) > 1 - \delta/2$ for all $n \geqq N_2$. Let $N = \max(N_1, N_2)$. Then, for any $n \geqq N$, $K_n = S_n \wedge A_n^x \wedge P^{x-\varepsilon}$ is nonnull. Let c_n be any unit vector in K_n. Since c_n belongs to K_n and so to $P^{x-\varepsilon}$, $(Tc_n, c_n) \leqq x - \varepsilon$. And since c_n is in A_n^x, one has $(T_n c_n, c_n) \geqq x$. As c_n belongs to S_n, one has $\| \, |T_n - T| \, c_n \| \leqq \varepsilon/2$. Thus

$$\begin{aligned} \varepsilon/2 \geqq \| \, |T_n - T| \, c_n \| &= \|(T_n - T)c_n\| \geqq |((T_n - T)c_n, c_n)| \\ &= |(T_n c_n, c_n) - (Tc_n, c_n)| \geqq x - (x - \varepsilon) = \varepsilon. \end{aligned}$$

This contradiction shows that the limit point L cannot be less than $m(P^x)$. Similarly by using the fact that F is right-continuous, one can show that the assumption that $L > m(P^x)$ will also lead to a contradiction. Thus $m(P^x)$ is the only limit point of the bounded sequence $\{m(P_n^x)\}$. Hence $m(P_n^x) \to m(P^x)$, i.e., $F_n(x) \to F(x)$. Hence the theorem is proved.

COROLLARY 4.1. *Let x_1 and x_2 be two continuity points of F ($x_1 < x_2$.) Let R_n be the spectral projection of T_n corresponding to the interval $(x_1, x_2]$ (open at x_1 and closed at x_2), and R the spectral projection of T corresponding to the same interval. Then $m(R_n) \to m(R)$.*

The proof is easy and is omitted.

Theorem 4.2 cannot be proved in the same way as above since, unlike a gage, a state may not be subadditive, i.e., for any two arbitrary projections P and Q, the inequality $\sigma(P \vee Q) \leqq \sigma(P) + \sigma(Q)$ is not in general valid. We shall deduce Theorem 4.2 from Theorem 4.4 after proving the latter.

Proof of Theorem 4.3. Theorem 4.3 is not a consequence of Theorem 4.1 since the notion of weak convergence is not additive even in the commutative case. However, a proof can be given along the following lines: Let F_n be the distribution function of A_n, G_n the distribution function of B_n ($n = 1, 2, \ldots$) and F that of A. Let x be any continuity point of F. Corresponding to the interval $(-\infty, x]$, let P_n be the spectral projection of A_n, Q_n that of B_n, and P that of A. By assumption

$m(P_n) \to m(P)$. Let L be a limit point of the sequence $\{m(Q_n)\}$. As such there exists a subsequence, also denoted by $\{m(Q_n)\}$, which converges to L. If possible, let $L < m(P)$. So, for some $\delta > 0$, $L = m(P) - 2\delta$. As in the argument of Theorem 4.1, one can find a number $\varepsilon > 0$ such that $x - \varepsilon$ is a continuity point of F, and such that if R is the spectral projection of A corresponding to $(-\infty, x-\varepsilon]$, then $m(R) > m(P) - \delta/2$, so that $L < m(R) - \delta$. As $m(Q_n) \to L$, one can find a positive integer M such that for all $n \geqq M$, $m(Q_n) < L + \delta < m(R) - \delta/2$. For each n, let R_n be the spectral projection of A_n corresponding to $(-\infty, x-\varepsilon]$. Since $x-\varepsilon$ is a continuity point of F, and $\{A_n\}$ converges weakly to A, it follows that $m(R_n) \to m(R)$. Hence one can find a positive integer N_1 such that $m(Q_n) < m(R_n) - \delta/4$, for all $n \geqq N_1$. Let S_n denote the spectral projection of $|A_n - B_n|$ corresponding to the interval $[0, \varepsilon/2)$. Since $\{A_n - B_n\}$ converges in measure to zero, $m(S_n) \to 1$ and so is $\geqq 1 - \delta/8$ for all $n \geqq$ some positive integer N_2. Let $N = \max(N_1, N_2)$. Then for any $n \geqq N$, $S_n \wedge Q_n^{\perp} \wedge R_n$ is nonnull. Let c_n be any unit vector in $S_n \wedge Q_n^{\perp} \wedge R_n$. Then as c_n is in S_n, $\| \, |A_n - B_n| \, c_n \| \leqq \varepsilon/2$. But as c_n is in $Q_n^{\perp}$, $(B_n c_n, c_n) \geqq x$. And as c_n is R_n, $(A_n c_n, c_n) < x - \varepsilon$. Thus $\varepsilon/2 > \|(A_n - B_n)c_n\| \geqq |(A_n c_n, c_n) - (B_n c_n, c_n)| = |(x-\varepsilon) - x| = \varepsilon$. This contradiction shows that L cannot be less than $m(P)$. Similarly one can show that L cannot be greater than $m(P)$. Thus $m(P)$ is the only limit point of the bounded sequence $\{(m(Q_n)\}$, which proves the theorem.

As a consequence of the theorem, we obtain the following corollary, which, in the commutative case, has been proved by Slutsky.

COROLLARY 4.2. *Let F_n, for each n, be the distribution function of a selfadjoint operator S_n, and let F be that of a selfadjoint operator S. Let $\{S_n\}$ converge to S weakly, and let $\{D_n\}$ be another sequence of selfadjoint operators converging in measure to cI (c some real number, and I the identity operator). Let G_n be the distribution function of $S_n + D_n$ and G that of $S + cI$. Let x be any continuity point of F. Then $G_n(x+c) \to G(X+c)$.*

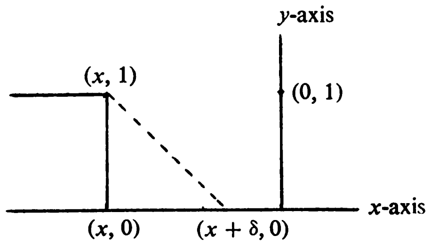
The proof of this corollary rests on the following proposition:

"Let N be any selfadjoint operator with distribution function H. Let c and I be as in the above theorem. Let J be the distribution function of $N + cI$. Then, for any point y on the real line, $H(y) = J(y+c)$."

The proof is as follows: For any Borel set E on the line, let $E + c$ denote the set of points $x + c$, where x is any point of E. Then one can verify that the spectral projection of N corresponding to E, is the same as the spectral projection of $N + cI$ corresponding to $E + c$.

Proof of the corollary. Set $A_n = S_n + D_n - cI$, $B_n = S_n$. Then $\{A_n - B_n\}$ converges in measure to zero. Also $\{B_n\}$ converges weakly to S. Hence by Theorem 4.3 $\{A_n\}$ converges weakly to S. Let H_n be the distribution function of A_n and G_n that of $A_n + cI$. Then $H_n(x) \to F(x)$ where x is any continuity point of F. By the above proposition $H_n(x) = G_n(x+c)$ and $F(x) = G(x+c)$. Hence $G_n(x+c) \to G(x+c)$ which proves the corollary.

Proof of Theorem 4.4. Let $\{T_n\}$ converge in measure to T. Let x be any continuity point of the distribution function F of T. We shall now show that $\{P_n^x\}$ converges in measure to P^x. Let $\psi(\cdot)$ be the *characteristic function* of the set $(-\infty, x]$. Then $P_n^x = \psi(T_n)$ and $P^x = \psi(T)$. We shall prove that $m(|\psi(T_n) - \psi(T)|) \to 0$ which will imply that $\{\psi(T_n)\}$ converges to $\psi(T)$ in the L_1-mean and hence in measure. Let an arbitrary positive number ε be given. The continuity points of F being dense on the line, one can find a sufficiently small δ $(\delta > 0)$ such that $x+\delta$ is a continuity point of F and $F(x+\delta) - F(x) < \varepsilon \cdots (1)$. Let R^x denote the spectral projection of T, corresponding to $(x, x+\delta]$. As x and $x+\delta$ are both continuity points of F, it follows by Corollary 4.1 that $m(R_n^x) \to m(R^x)$. Hence one can find a positive integer N_1 such that $|m(R_n^x) - m(R^x)| < \varepsilon$ for $n \geqq N_1$, i.e., $m(R_n^x) < 2\varepsilon$ for $n \geqq N_1$. Now let σ be the continuous function defined thus:



Graph of σ

For all $\lambda < x$, $\sigma(\lambda) = 1$. For any $\lambda \geqq x+\delta$, $\sigma(\lambda) = 0$. In the open interval $(x, x+\delta)$, the graph of σ is the straight line joining the two points $(x, 1)$ and $(x+\delta, 0)$. Clearly, for any λ, $|\sigma(\lambda) - \psi(\lambda)| < 1$, so that $\|\sigma(T) - \psi(T)\| \leqq 1$. Note that $\sigma(T) - \psi(T)$ is $\geqq 0$, so that

$$\begin{aligned} m(|\sigma(T) - \psi(T)|) &= m(\sigma(T) - \psi(T)) \\ &= m((\sigma(T) - \psi(T))R^x) \leqq \|\sigma(T) - \psi(T)\| \cdot m(R^x) < \varepsilon. \end{aligned}$$

Similarly for $n > N_1$,

$$\begin{aligned} m(|\sigma(T_n) - \psi(T_n)|) &= m(\sigma(T_n) - \psi(T_n)) \\ &= m((\sigma(T_n) - \psi(T_n)) \cdot R_n^x) \leqq \|\sigma(T_n) - \psi(T_n)\| \cdot m(R_n^x) \leqq m(R_n^x) < 2\varepsilon. \end{aligned}$$

As σ is continuous and decreasing, it follows that $\{\sigma(T_n)\}$ converges in measure to $\sigma(T)$. Since this sequence is uniformly bounded, (bounded in norm by 1), convergence in measure implies convergence in the L_1-mean. So, given ε, there exists a positive integer N_2 such that $m(|\sigma(T_n) - \sigma(T)|) < \varepsilon$ for $n \geqq N_2$. Let $N = N_1 + N_2$. Now for any $n \geqq N$ one has

$$\begin{aligned} m(|\psi(T) - \psi(T_n)|) &\leqq m(|\psi(T) - \sigma(T)|) + m(|\sigma(T) - \sigma(T_n)|) + m(|\sigma(T_n) - \psi(T_n)|) \\ &\leqq \varepsilon + \varepsilon + 2\varepsilon = 4\varepsilon, \end{aligned}$$

i.e., $m(|\psi(T_n) - \psi(T)|) \to 0$, i.e., $\{P_n^x\}$ converges to P^x in the L_1-mean and hence in measure. And x being any arbitrary continuity-point of F, the desired result follows.

CONVERSE. For any continuity-point x of F, let P_n^x denote the spectral projection of T_n and P^x that of T, corresponding to $(-\infty, x]$. Let $\{P_n^x\}$ converge in measure to P^x. We shall now show that for any continuous function σ with compact support on the real line, $\{\sigma(T_n)\}$ converges in the L_2-mean to $\sigma(T)$, and this will imply by the result [D_2] that $\{T_n\}$ converges in measure to T.

Let σ have compact support $[-k, k]$. Hence $\sigma(-k)=0=\sigma(k)$. As σ is uniformly continuous in $[-k, k]$, given any $\varepsilon>0$, one can find a positive number δ, such that for any two points x and y in $[-k, k]$, one has $|\sigma(x)-\sigma(y)|<\varepsilon$, whenever $|x-y|<2\delta$. Choose a point x_1 such that (1) $x_1>-k$ (2) the distance between x_1 and $-k$ exceeds $\delta/2$ but is less than δ and (3) x_1 is a continuity point of F. Choose now successively points $x_2, x_3, \ldots, x_N$ such that (4) $x_2<x_3<\cdots<x_{N-1}<k$ and $x_N>k$ (5) $\delta \geqq |x_{i+1}-x_1| \geqq \delta/2$ and (6) each x_i is a continuity point of F, $i=1, 2, \ldots, N-1$. The choice of such a finite sequence $x_1, x_2, \ldots, x_N$ is possible as the continuity points of F are dense on the real line.

Define a new function $\psi(\lambda)$ thus:

$$\begin{aligned}
\psi(\lambda) &= 0 \quad \text{for} \quad \lambda \leqq x_1,\\
\psi(\lambda) &= \sigma(x_1), \qquad x_1 < A \leqq x_2,\\
\psi(\lambda) &= \sigma(x_2), \qquad x_2 < \lambda \leqq x_3, \quad \text{and}\\
\psi(\lambda) &= \sigma(x_{N-1}), \qquad x_{N-1} < \lambda \leqq x_N,\\
\psi(\lambda) &= 0, \qquad \lambda > x_N.
\end{aligned}$$

Clearly for any λ,

$$|\psi(\lambda)-\sigma(\lambda)| \leqq \varepsilon.$$

Corresponding to $(x_{i-1}, x_i]$, let R_n^i denote the spectral projection of T_n and R^i that of T $(i=2, \ldots, N)$. By assumption, $\{R_n^i\}$ converges in measure to R^i $(i=1, 2, \ldots, N)$, so that $\psi(T_k)=\sum_{i=2}^N \sigma(\lambda_{i-1})R_k^i$ converges in measure to $\sum_{i=2}^N \sigma(\lambda_{i-1})R^i=\psi(T)$. It is easy to verify that $|\psi(\lambda)-\sigma(\lambda)|<\varepsilon$ for all λ, so that $\|\psi(T_n)-\sigma(T_n)\|<\varepsilon$ for all n. Since $\{\psi(T_n)\}$ converges in measure to $\psi(T)$ and is uniformly bounded in norm, it follows that $\psi(T_n)\to\psi(T)$ in the L_2-mean, i.e., $\|\psi(T_n)-\psi(T)\|_2\to 0$ as $n\to\infty$. As such, there exists a positive integer N_1 such that $\|\psi(T_n)-\psi(T)\|_2<\varepsilon$ for all $n\geqq N_1$. Again, for each n, $\|\psi(T_n)-\sigma(T_n)\|_2<\|\psi(T_n)-\sigma(T_n)\|\cdot m(I)=\varepsilon$. Hence, for any $n\geqq N_1$, $\|\sigma(T_n)-\sigma(T)\|_2\leqq\|\sigma(T)-\psi(T)\|_2+\|\psi(T)-\psi(T_n)\|_2+\|\psi(T_n)-\sigma(T_n)\|_2<\varepsilon+\varepsilon+\varepsilon=3\varepsilon$.

Hence the theorem is proved.

We shall now deduce Theorem 4.2 from Theorem 4.4. Let x be any continuity point of the distribution function of T. Then by Theorem 4.4, $\{P_n^x\}$ converges in measure and because of uniform boundedness, converges in the L_2-mean, to P^x, i.e., $m((P_n^x-P^x)^*(P_n^x-P^x))\to 0$. Since σ is absolutely continuous with respect to m, it follows that $\sigma((P_n^x-P^x)^*(P_n^x-P^x))\to 0$. And this implies $\sigma(P_n^x-P^x)\to 0$, or $\sigma(P_n^x)\to\sigma(P^x)$, which completes the proof of theorem.

We conclude this section with a necessary and sufficient condition for weak convergence and convergence in measure to coincide.

THEOREM 4.5. *Let $\{T_n\}$ be an arbitrary sequence of selfadjoint operators converging weakly to a selfadjoint operator T. Then $\{T_n\}$ converges in measure to T if and only if $T=cI$ for some real constant c.*

Proof. First we shall show that the given condition is sufficient. Let $T=cI$ for some real c. We shall prove that $\{T_n\}$ converges in measure to T. Let F_n be the distribution function of T_n and F that of T. It is easy to see that $F(x)=0$, if $x<c$, and $F(x)=1$ if $x \geqq c$. Also any point $(y \neq c)$ is a continuity point of F. Corresponding to $(-\infty, y]$, let P_n^y denote the spectral projection of T_n and P^y that of T $(n=1,2,\ldots)$. When $y<c$, $P^y=0$, and when $y \geqq c$, $P^y=I$. In view of weak convergence, $m(P_n^y) \to m(P^y)=0=P^y$ whenever $y<c$, i.e., $\{P_n^y\}$ converges in measure to $0=P^y$ whenever $y<c$. Whenever $y>c$, $m(P^y)=1$, so that $m(P_n^y) \to m(P^y)=1$, which implies that $\{P_n^y\}$ converges in measure to I. Thus, for any arbitrary continuity point y of F, $\{P_n^y\}$ converges in measure to P^y. Hence by Theorem 4.4, $\{T_n\}$ converges in measure to T.

CONVERSE. We shall give an example of a sequence of projections converging weakly to a projection, but still not converging in measure.

Let α be a continuous finite factor and m the faithful normal trace on α with $m(I)=1$. Let P, Q, R, and S be four mutually orthogonal projections in α, with $m(P)=\frac{1}{4}=m(Q)=m(R)=m(S)$. Let $\{A_n\}$ be a decreasing sequence of projections contained in Q, such that $m(A_n) \to 0$. Let $R_n=P+A_n$. Then $m(R_n) \to \frac{1}{4}=m(S)$ and $m(R_n^{\perp}) \to \frac{3}{4}=m(S^{\perp})$. Let F be the distribution function of S and F_n that of R_n. All the involved operators being projections, the spectrum of each one of them is concentrated at the two points 0 and 1. Thus, for any x with $x<0$, $F_n(x)=0=F(x)$. For $0 \leqq x<1$, $F_n(x)=m(R_n^{\perp}) \to \frac{3}{4}=m(S^{\perp})$ and $F(x)=m(S^{\perp})$. For $x \geqq 1$, $F_n(x)=1=F(x)$. And any point x other than 0 and 1 is a continuity point of F. Hence for any such x (in fact for all x), $F_n(x) \to F(x)$. But $\{R_n\}$ cannot converge in measure to S. For, if it does, then $\{R_n \cdot S\}$ will converge in measure to $S \cdot S=S$. But $R_n \cdot S=0$ for all n, while $S \neq 0$. Since $S \neq c \cdot I$ for any c, this counterexample completes the proof.

COROLLARY 4.3. *Let $\{T_n\}$ converge in measure to cI. Let Φ be a real, continuous function (not necessarily expressible as the sum of a finite number of monotonic continuous functions). Then $\{\Phi(T_n)\}$ converges in measure to dI, where $d=\Phi(c)$.*

REMARK. This corollary is nontrivial and does not immediately follow from the definition of convergence in measure.

Proof. First we shall establish the following proposition:

"Let $\{T_n\}$ converge weakly to T. Then $\{\Phi(T_n)\}$ converges weakly to $\Phi(T)$."

To prove this, let g denote any real, bounded and continuous function. Let h denote the composite map $g \cdot \Phi$. Clearly h is bounded and continuous. Now $m(g(\Phi)(T_n))=m(h(T_n)) \to m(h(T))$ since $\{T_n\}$ converges to T weakly. And the arbitrariness of g implies that $\{\Phi(T_n)\}$ converges weakly to $\Phi(T)$. Hence the proposition.

Now, the corollary is an immediate consequence of Theorem 4.5 and the above proposition.

5. **Some dominated convergence theorems.** In this section we shall state and prove some dominated convergence theorems which, in the case of a finite gage space, are stronger than the corresponding results of Stinespring [10].

DEFINITIONS. A sequence $\{A_n\}$ of operators is said to be *U-continuous* if, given any $\varepsilon > 0$, there exists a $\delta > 0$, such that for any projection P with $m(P) < \delta$, one has $|m(A_nP)| < \varepsilon$ for all n, and *V-continuous* if $\|A_nP\|_2 < \varepsilon$ for all n. Our definition of gross convergence is the same as that of Stinespring [10, pp. 23 and 26]. It is known [10, p. 32] that convergence in measure always implies gross convergence and that these two concepts are equivalent when the gage of the identity is finite. The orthogonal complement of any projection P will be denoted either by $P^\perp$ or by $I-P$. For any operator T, Re T will denote $(T+T^*)/2$ and Im T will denote $(T-T^*)/2i$.

For a sequence $\{A_n\}$ of operators converging grossly to an operator A, the following two theorems have been proved by Stinespring (without assuming the finiteness of the gage).

THEOREM 5.1 [10, P. 29]. *If there exists a nonnegative integrable operator T such that $-T \leqq \operatorname{Re} A_n \leqq T$, and $-T \leqq I_m A_n \leqq T$ for all n, then $A_n \to A$ in the L_1-mean.*

THEOREM 5.2 [10, P. 30, AND THE REMARK IN P. 31]. *If there exists a nonnegative integrable T with $(\operatorname{Re} A_n)^2 \leqq T$ and $(\operatorname{Im} A_n)^2 \leqq T$ for all n, then $A_n \to A$ in the L_2-mean.*

However, even in finite gage spaces there are cases which are not covered by these theorems, i.e., one can construct a sequence $\{A_n\}$ of nonnegative integrable operators converging in measure to an integrable operator A, and $m(A_n) \to m(A)$; but there does not exist any integrable operator T with $A_n \leqq T$ for all n.

E.g., let α be a continuous finite factor, and let m be the faithful, normal trace on α with $m(I)=1$. Let P_1 be a projection with $m(P_1)=1/2$. Let P_2 be a projection contained in $I-P_1$ with $m(P_2)=1/2^2, \ldots$, and in general let P_n be a projection contained in $I-P_1-\cdots-P_{n-1}$ with $m(P_n)=1/2^n$. Let $A_n=2^n/n.\ P_n$. Each A_n is integrable and $\{A_n\}$ converges in measure and hence grossly to zero and $m(A_n)=1/n \to 0=m(0)$. If possible let there exist a nonnegative operator T such that $A_n \leqq T$ for all n. Then it is easy to see that $m(T) \geqq m(TP_1)+\cdots+m(TP_n) \geqq \sum_{k=1}^n 1/k$ for all n, which shows that T cannot be integrable. In order to cover such exceptional cases also, we shall state and prove the following theorems and show how they apply to the above example.

In what follows, $\{T_n\}$ will denote an arbitrary sequence of operators converging in measure to an operator T, and $\{S_n\}$ will denote an arbitrary sequence of nonnegative integrable operators converging in measure to an integrable operator S, and such that $m(S_n) \to m(S)$.

THEOREM 5.3. *If, corresponding to each n, $-S_n \leqq \operatorname{Re} T_n \leqq S_n$ and $-S_n \leqq \operatorname{Im} T_n \leqq S_n$, then $T_n \to T$ in the L_1-mean.*

THEOREM 5.4. *If, corresponding to each* n, $(\mathrm{Re}\, T_n)^2 \leqq S_n$, *and* $(\mathrm{Im}\, T_n)^2 \leqq S_n$, *then* $T_n \to T$ *in the* L_2*-mean.*

REMARK. In the case of a finite gage space, these results are stronger than the corresponding results of Stinespring (Theorems 5.1 and 5.2 mentioned at the beginning of this section). Some dominated convergence theorems in arbitrary gage spaces have also been proved in [5].

LEMMA 5.1. *Let* $\{A_n\}$ *be a sequence of nonnegative integrable operators converging in measure to an operator* A, *and let* $m(A_n) \to m(A)$. *Then* $A_n \to A$ *in the* L_1*-mean.*

Proof. This lemma is well known in the commutative case. However, the method by which it is proved in standard textbooks on probability theory, [3, p. 140, Problem 17] does not directly extend to this general case, for which the proof is as follows. Let an arbitrary positive number ε be given. For any n, let $C_n = A - A_n$, and G_n^ε, F_n^ε, and H_n^ε denote the spectral projections of C_n corresponding to the intervals $[\varepsilon, \infty)$, $(-\infty, -\varepsilon]$, and $(-\varepsilon, +\varepsilon)$ respectively. It is easily seen that H_n^ε is the spectral projection of $|C_n|$ corresponding to the interval $[0, \varepsilon)$. Now

$$C_n = C_nG_n^\varepsilon + C_nF_n^\varepsilon + C_nH_n^\varepsilon \cdots \tag{1}$$

and

$$|C_n| = C_nG_n^\varepsilon - C_nF_n^\varepsilon + |C_n|H_n^\varepsilon \cdots. \tag{2}$$

Hence $0 \leqq m(C_nG_n^\varepsilon) \leqq m(G_n^\varepsilon AG_n^\varepsilon)$. Since $\{C_n\}$ converges in measure to zero, $m(G_n^\varepsilon) \to 0$ as $n \to \infty$, and since A is integrable, this implies that $m(AG_n^\varepsilon) \to 0$. Hence $m(C_nG_n^\varepsilon) \to 0$. Again $m(|C_n|H_n^\varepsilon) \leqq \| \, |C_n| \cdot H_n^\varepsilon \| = \varepsilon$, and similarly $|m(C_nH_n^\varepsilon)| < \varepsilon$. As $m(C_n) \to 0$, it follows from (1) that

$$\limsup_{n\to\infty} |m(C_nF_n^\varepsilon)| \leqq \varepsilon.$$

Hence,

$$\limsup_{n\to\infty} m(|C_n|) \leqq \varepsilon + \varepsilon = 2\varepsilon.$$

ε being arbitrary, the desired result follows.

LEMMA 5.2. *Let* $\{A_n\}$ *be any sequence of nonnegative, integrable operators. Then* $A_n \to A$ *in the* L_1*-mean if and only if* $\{A_n\}$ *is U-continuous and converges in measure to* A.

Proof. If $A_n \to A$ in the L_1-mean, then it is known that it converges to A in measure, and it is easily verified that $\{A_n\}$ is U-continuous. For the converse, let an arbitrary positive number ε be given. By assumption there exists a "δ" > 0 such that, for any projection P with $m(P) < \delta$, one has $m(A_nP) < \varepsilon$ for all n. The sequence $\{PA_nP\}$ converges in measure to PAP. The involved operators being nonnegative, it follows by the noncommutative version of Fatou's Lemma [9, p. 31] that $m(PAP) \leqq \liminf_{n\to\infty} m(PA_nP) \leqq \varepsilon$, i.e., $m(AP) \leqq \varepsilon$. Let P_n be the spectral projection of $|A_n - A|$ corresponding to the interval $[0, \varepsilon)$ and $Q_n = I - P_n$. Then $m(P_n) \to 1$.

So, for all $n \geqq$ some positive integer N, $m(P_n) \geqq 1-\delta$. Thus for any $n \geqq N$,

$$\begin{aligned}\|A_n - A\|_1 &= m(|A_n - A|) \\ &= m(|A_n - A|P_n) + m(|A_n - A|Q_n) \\ &\leqq \varepsilon + m(|A_n - A|Q_n).\end{aligned}$$

Let U_n and W_n denote the spectral projection of $A_n - A$ corresponding to $[\varepsilon, \infty)$ and $(-\infty, -\varepsilon]$ respectively. Then $Q_n = U_n + W_n$, so that $m(U_n) \leqq \delta$ and $m(W_n) \leqq \delta$. And

$$\begin{aligned}m(|A_n - A|Q_n) &= m((A_n - A)U_n) - m((A_n - A)V_n) \\ &\leqq |m(A_n U_n)| + |m(AU_n)| + |m(A_n V_n)| + |m(AV_n)| \\ &\leqq \varepsilon + \varepsilon + \varepsilon + \varepsilon = 4\varepsilon.\end{aligned}$$

Hence $\|A_n - A\|_1 \leqq \varepsilon + 4\varepsilon = 5\varepsilon$. ε being arbitrary, the desired result follows.

LEMMA 5.3. *For any square-integrable operator A, and $\varepsilon > 0$, one can find a "δ" > 0, such that for any projection P with $m(P) < \delta$, one has $\|AP\|_2 < \varepsilon$.*

Proof. "A" being square-integrable, there exists a bounded operator B with $\|A - B\|_2 < \varepsilon/2$. Let $\|B\| = k$, let $\delta = \varepsilon/2k$, and P any projection with $m(P) < \delta$. Then $\|BP\|_2 \leqq \|B\| \cdot m(P) \leqq k \cdot \varepsilon/2k = \varepsilon/2$. As $\|AP - BP\|_2 < \varepsilon/2$, it follows that $\|AP\|_2 \leqq \varepsilon/2 + \varepsilon/2 = \varepsilon$. Hence the lemma.

LEMMA 5.4. *For any square-integrable operator A and projection P,*

$$m(|A|^2 P) = (\|AP\|_2)^2.$$

Proof.

$$\begin{aligned}(\|AP\|_2)^2 &= m(AP)^*(AP) \\ &= m((PA^*)(AP)) \\ &= m(P(A^*A)P) \\ &= m(P|A|^2 P) \\ &= m(|A|^2 P).\end{aligned}$$

LEMMA 5.5. *Let $\{A_n\}$ be a sequence of square-integrable operators. Then $A_n \to A$ in the L_2-mean, if and only if $\{A_n\}$ converges in measure to A, and is also V-continuous.*

Proof. Let $A_n \to A$ in the L_2-mean. Then clearly $\{A_n\}$ converges in measure to A. Also, since $A_n - A$ and A_n are square-integrable, it follows that A is square-integrable. By Lemma 5.3, there exists a "δ" such that $\|AP\|_2 < \varepsilon/2$, for all projections P with $m(P) < \delta_1$. As $A_n \to A$ in the L_2-mean, there exists a positive integer N, with $\|A_n - A\|_2 < \varepsilon/2$ for all $n \geqq N$. Hence $\|A_n P\|_2 < \varepsilon$ for all $n > N$. Also $A_1, \ldots, A_N$ being each square-integrable, and N being finite, there exists a δ_2 such that for any projection P with $m(P) < \delta_2$, we have $\|A_i P\|_2 < \varepsilon/2$, $i = 1, 2, \ldots, N$. Let $\delta = \min(\delta_1, \delta_2)$. Then for any projection P with $m(P) < \delta$, we have $\|A_i P\|_2 < \varepsilon$, $i = 1, 2, \ldots$, which shows that $\{A_n\}$ is V-continuous. For the converse note that, since $\{A_n\}$ converges in measure to A, $\{|A_n|^2\}$ converges in measure to $|A|^2$. Let an arbitrary positive number ε be given. Since $\{A_n\}$ is V-continuous, it follows that there exists a $\delta>0$, such that for any projection P with $m(P)<\delta$, $\|A_nP\|_2<\varepsilon$ for all n. Hence $(\|A_nP\|_2)^2=m(|A_n|^2P)<\varepsilon^2$. By Fatou's Lemma, it follows that $m(|A|^2P)<\varepsilon^2$, i.e., $(\|AP\|_2)^2<\varepsilon^2$. As $\{A_n\}$ converges in measure to A, there exists a sequence $\{Q_n\}$ of projections such that $\|(A_n-A)Q_n\|<\varepsilon$ for all n, and $m(Q_n)\to 1$. Hence for all $n\geqq$ some positive integer N, $m(Q_n^\perp)<\delta$, so that for any $n\geqq N$,

$$\begin{aligned}\|A_n-A\|_2 &\leqq \|(A_n-A)Q_n\|_2+\|(A_n-A)Q_n^\perp\|_2\\ &\leqq \varepsilon+\|A_nQ^\perp\|_2+\|AQ_n^\perp\|_2\\ &\leqq \varepsilon+\varepsilon+\varepsilon = 3\varepsilon.\end{aligned}$$

Hence the lemma is proved.

Proof of Theorem 5.3. It suffices to prove the theorem assuming each T_n to be selfadjoint, as the general case is reducible to this. Since $-S_n\leqq T_n\leqq S_n$, it follows that $-S\leqq T\leqq S$, which shows that T is integrable. Again, it follows by Lemma 5.1 that $S_n\to S$ in the L_1-mean. Therefore it also follows that $\{S_n\}$ and hence $\{T_n\}$ are U-continuous. Thus, given an arbitrary positive number ε, there exists a $\delta>0$, such that for any projection P with $m(P)<\delta$, one has $|m(T_nP)|<\varepsilon$ for all n. Let G_n and F_n denote the spectral projections of T_n-T corresponding to the intervals $[\varepsilon/2,\infty)$ and $(-\infty,-\varepsilon/2]$. Let $K_n=I-(G_n+F_n)$. Then K_n is the spectral projection of $|T_n-T|$ corresponding to $[0,\varepsilon/2)$. In view of convergence in measure, $m(K_n)\to 1$ as $n\to\infty$; hence there exists a positive integer N such that for $n\geqq N$, $m(K_n)>1-\varepsilon$, so that $m(G_n)$ and $m(F_n)$ are both $<\delta$. Now, for any $n\geqq N$,

$$\begin{aligned}\|T_n-T\|_1 &= m(|T_n-T|)\\ &= m(|T_n-T|K_n)+m(|T_n-T|G_n)+m(|T_n-T|F_n)\\ &\leqq \varepsilon+m(|T_n-T|G_n)+m(|T_n-T|F_n)\\ &= \varepsilon+m((T_n-T)G_n)-m((T_n-T)F_n)\\ &\leqq \varepsilon+|m(T_nG_n)|+|m(TG_n)|+|m(T_nF_n)|+|m(TF_n)|\\ &\leqq \varepsilon+\varepsilon+\varepsilon+\varepsilon+\varepsilon = 5\varepsilon.\end{aligned}$$

Hence $T_n\to T$ in the L_1-mean. The general case follows by applying the previous result separately to the sequences $\{\text{Re } T_n\}$ and $\{\text{Im } T_n\}$.

Proof of Theorem 5.4. To begin with, let us assume that each T_n is selfadjoint. As $S_n\to S$ in the L_1-mean, it follows that $\{S_n\}$ is U-continuous. Since $T_n^2\leqq S_n$, $\{T_n^2\}$ is also U-continuous. Further, for any projection P, $m(T_n^2P)=(\|T_nP\|_2)^2$ so that the sequence $\{T_n\}$ is V-continuous. Also $\{T_n\}$ converges in measure to T. Hence by Lemma 5.5, $T_n\to T$ in the L_2-mean.

The general case can be proved by applying the previous result, separately to $(\text{Re } T_n)^2$ and $(\text{Im } T_n)^2$.

We shall now show how our theorems apply to the example given at the beginning of this section. Let $A_n, P_n, \ldots$, etc., be as in that example. Define $B_n=A_n+P_n^\perp$. Then $\{B_n\}$ converges in measure to I and $m(B_n)\to 1$ and it follows by Theorem 5.3 that $A_n\to 0$ in the L_1-mean.

6. **Applications to operator-entropy.** I. Let $\{T_n\}$ be a sequence of nonnegative, square-integrable operators, converging in the L_2-mean to an operator T. Then the operator-entropy of T_n $(= -T_n \log T_n)$, tends in the L_1-mean, to the operator-entropy of T $(= -T \log T)$, and in particular, the numerical entropy of T_n $(= m(-T_n \log T_n))$ tends to the numerical entropy of T.

Proof. Let $\sigma(\lambda) = \lambda \log \lambda$, where $\sigma(0)$ is defined to be zero. Hence $\sigma(\lambda)$ is continuous in the closed unit interval so that there exists a positive constant c, such that $-c < \sigma(\lambda) < c$, whenever $0 \leqq \lambda \leqq 1$. Hence it follows that for any nonnegative operator A

$$-A^2 - cI \leqq A \log A \leqq A^2 + cI.$$

Since $\{T_n\}$ converges in the L_2-mean to T, it converges in measure to T. Hence $\{T_n^2\}$ converges in measure to T^2. Also as $\|T_n - T\|_2 \to 0$, it follows that $m(T_n^2) \to m(T^2)$. Thus if we set $S_n = T_n^2 + cI$, and $S = T^2 + cI$, then $\{S_n\}$ is a sequence of nonnegative integrable operators converging in measure to a nonnegative integrable operator S and also $m(S_n) \to m(S)$. Moreover, $-S_n \leqq T_n \log T_n \leqq S_n$ for each n. As $\{T_n\}$ converges in measure to T, it follows by Theorem 2 of §2 that $\{T_n \log T_n\}$ converges in measure to $T \log T$. Now as a consequence of our dominated convergence theorem, it follows that $\{T_n \log T_n\}$ converges in the L_1-mean to $T \log T$.

Before proceeding further, we shall prove a lemma.

LEMMA 6.1. *Let β be any ring contained in α. Let A be any square-integrable operator and let $K = E[A|\beta]$ denote the conditional expectation of A, given β in the sense of Umegaki* [11]. *Then one has $\|K\|_p \leqq \|A\|_p$, $p = 1, 2$.*

Proof. We shall first show this for $p = 1$. Note that for any U in β, $m(AU) = m(KU)$. Let U_1 denote the unit sphere of α and U_2 that of β. Then,

$$\begin{aligned}\|K\|_1 &= \sup_{U \text{ in } U_2} (|m(KU)|) \\ &= \sup_{U \text{ in } U_2} (|m(AU)|) \\ &\leqq \sup_{U \text{ in } U_1} (|m(AU)|) \\ &= \|A\|_1.\end{aligned}$$

For $p = 2$: $G = L_2(H, \alpha, m)$ is a Hilbert space, of which $N = L_2(H, \beta, m)$ is a closed subspace, "A" is an element of G, and K is its projection on the subspace N. Hence the norm of K regarded as an element of N $(= \|K\|_2)$ is less than or equal to the norm of A, regarded as an element of G $(= \|A\|_2)$. Thus $\|K\|_2 \leqq \|2\|_2$.

Using these two results, we can rewrite as follows the two dominated convergence theorems which we have proved:

THEOREM 6.1. *Let all the assumptions of Theorem* 5.3 (*Theorem* 5.4) *be satisfied. Then, in the notation of Theorem* 5.3 (*Theorem* 5.4), $E(T_n|\beta) \to E(T|\beta)$ *in the L_1-mean* (*in the L_2-mean*).

II. Let T be any nonnegative square-integrable operator and R any bounded selfadjoint operator. Let β be the ring generated by R, and let $S=E(T|\beta)$. Then the number $m(T\log T)-m(S\log S)$ is defined by Nakamura and Umegaki [4] to be the information about T contained in R. This we shall denote by $I(T; R)$. $I(T; R)$ is finite, T being square-integrable.

As another application of our theorems, we state the following result:

Let $\{T_n\}$ be a sequence of nonnegative, and square-integrable operators, converging in the L_2-mean to T. Then, in the above notation, $I(T_n; R) \to I(T; R)$.

Proof. Let $E(T_n|\beta)=S_n$ and $E(T|\beta)=S$. Then, by Lemma 6.1, $S_n \to S$ in the L_2-mean. Hence, by an earlier result, $m(T_n \log T_n) \to m(T\log T)$ and $m(S_n \log S_n) \to m(S\log S)$. Thus

$$\begin{aligned} I(T_n; R) &= m(T_n \log T_n) - m(S_n \log S_n) \\ &\to m(T\log T) - m(S\log S) \\ &= I(T; R). \end{aligned}$$

ACKNOWLEDGEMENTS. The author is grateful to Professor I. E. Segal for having gone through the material in this paper and given his comments; to Professor Umegaki for having gone through parts of this paper and suggested Example 3 in §2; and to Dr. Sethuraman and Mr. U. S. R. Murthy for valuable suggestions for improving the presentation of this paper. He also profited from conversations with Mr. Viswanath and suggestions of the referee.

REFERENCES

1. B. V. Gnedenko and A. N. Kolmogorov, *Limit distributions for sums of independent random variables*, Addison-Wesley, Reading, Mass., 1954.

2. A. N. Kolmogorov, *Foundations of the theory of probability*, Chelsea, New York, 1956.

3. M. Loève, *Probability theory*, 3rd ed., Van Nostrand, New York, 1963.

4. M. Nakamura and H. Umegaki, *A note on the entropy for operator algebra*, Proc. Japan Acad. **37** (1961), 149–151.

5. A. R. Padmanabhan, *Some dominated convergence theorems in a von Neumann algebra*, Proc. Japan Acad. **42** (1966), 347–350.

6. I. E. Segal, *A non-commutative extension of abstract integration*, Ann. of Math. (2) **57** (1953), 401–457.

7. ———, *Tensor algebras over Hilbert spaces*. II, Ann. of Math. (2) **63** (1956), 160–175.

8. ———, *Distributions in Hilbert space and canonical systems of operators*, Trans. Amer. Math. Soc. **88** (1958), 12–41.

9. D. Shale and W. F. Stinespring, *States of the Clifford algebra*, Ann. of Math. (2) **80** (1964), 365–381.

10. W. F. Stinespring, *Integration theorems for gages and duality theorems for unimodular groups*, Trans. Amer. Math. Soc. **90** (1959), 15–56.

11. H. Umegaki, *Conditional expectation in an operator algebra*, Tôhoku Math. J. **6** (1954), 177–181.

INDIAN STATISTICAL INSTITUTE
CALCUTTA-35, INDIA
INDIANA UNIVERSITY,
BLOOMINGTON, INDIANA