

OPERATOR LIMIT THEOREMS

BY
NORTON STARR⁽¹⁾

0. Introduction. In this paper certain almost everywhere convergence and maximal theorems are proved, extending a type of limit theorem developed by Rota [22] and generalized by Doob [9].

Given a sequence $\{T_n\}_{n=1}^\infty$ of linear operators on $L^p(X, \Sigma, \mu)$ into itself for some $p \in [1, \infty)$ and some totally σ -finite positive measure space (X, Σ, μ) , we say the T_n satisfy a *maximal* theorem if

$$\int_X \left(\sup_n |Tf| \right)^p d\mu \leq A_p \int_X |f|^p d\mu,$$

for some constant A_p and all $f \in L^p(X, \Sigma, \mu)$. We say a sequence $\{f_n\}_{n=1}^\infty$ of functions in $L^p(X, \Sigma, \mu)$ converges *boundedly* in L^p if $\lim_{n \rightarrow \infty} f_n$ exists almost everywhere and $\int (\sup_n |f_n|)^p d\mu < \infty$.

Among the actively developed theories of bounded convergence are the ergodic theory, martingale theory, singular integral theory, the behavior of monotone sequences of operators [13], and the study of powers of self-adjoint positive definite operators [5], [24], [1].

The starting point for our investigations is the following new type of limit theorem due to Rota [22]: given doubly stochastic operators S_n , $n = 1, 2, 3, \dots$, on functions integrable over a probability space, $\lim_{k \rightarrow \infty} S_1^* S_2^* \dots S_k^* S_k S_{k-1} \dots S_1 f$ exists a.e. and boundedly in L^p for any $f \in L^p$ ($1 < p < \infty$). In deriving our results, we apply the techniques used by Doob [9] in proving a generalization, in the form of a ratio theorem, of the above theorem.

We remark that for the special case in which all the operators S_k , S_k^* of Rota's theorem coincide with a single operator T of the form $T = E_1 \cdot E_2$, where each E_i is a conditional expectation, bounded almost everywhere convergence in L^2 was obtained earlier by Burkholder and Chow [5]. As a preliminary to this result, Burkholder and Chow actually obtained bounded almost everywhere convergence in L^2 of $T^{2^n} f$, assuming only that T be a self-adjoint operator on L^2 which maps

Received by the editors January 8, 1965.

(1) The results of this paper were contained in the author's doctoral dissertation, submitted to the Massachusetts Institute of Technology, January, 1964. We express our sincere appreciation to Professor Gian-Carlo Rota for suggesting the problems considered here and for his valuable encouragement and advice. For financial support we are grateful to the National Science Foundation, the General Electric Foundation, and the Air Force Office of Scientific Research.

L^1 into L^1 with operator norm ≤ 1 and $L^1 \cap L^\infty$ into $L^1 \cap L^\infty$ without increasing L^∞ norms (the measure space could be infinite, too). These latter hypotheses are essentially those of Stein [24], who proved bounded almost everywhere convergence in L^p ($1 < p < \infty$) for $T^n f$, under the additional assumption that $(Tg, g) \geq 0$ ($g \in L^2$). The methods of Burkholder and Chow, Rota, and Stein differ markedly from one another. All these results are implied by our work, which is dependent on the theorems of Rota and Doob.

Before presenting a concrete description of our work, we record at this point some conventions. All measure spaces (X, Σ, μ) considered are totally σ -finite positive measure spaces unless the contrary is explicitly stated. We work with spaces of real valued functions only. All operators are assumed linear. Equations, inequalities, and limits are understood to hold almost everywhere (a.e.) unless otherwise noted.

By $\log^+ |f|$ we mean $\log (\max (1, |f|))$. We frequently consider the class of functions $f \in \bigcup_{1 \leq p < \infty} L^p$ which also have the property $\int_X |f| \log^+ |f| d\mu < \infty$. Note that $f \in \bigcup_{1 < p < \infty} L^p$ implies f has this property, and that on a probability space the functions having this property are integrable.

For any function f measurable on (X, Σ, μ) , $\|f\|_p$ denotes $(\int_X |f|^p d\mu)^{1/p}$ for $1 \leq p < \infty$ and $\|f\|_\infty$ denotes $\text{ess sup}_{x \in X} |f(x)|$.

Given subsets A, B, \dots of X and functions ϕ, ψ, \dots on X , $\mathcal{B}\{A, B, \dots\}$ and $\mathcal{B}\{\phi, \psi, \dots\}$ denote the σ -fields on X generated by $\{A, B, \dots\}$ and determined by $\{\phi, \psi, \dots\}$ respectively.

For fixed $p: 1 \leq p \leq \infty$, a linear operator P from $L^p(X_1, \mu_1)$ to $L^p(X_2, \mu_2)$ is called *positive* if $f \geq 0$ a.e. (μ_1) implies $Pf \geq 0$ a.e. (μ_2) and a *contraction* if $\|Pf\|_p \leq \|f\|_p$ for all $f, g \in L^p(X_1, \mu_1)$.

The linear operators T that occur most frequently in this work are of the following sort: T maps $\bigcup_{1 < p < \infty} L^p(X_1, \Sigma_1, \mu_1)$ into $\bigcup_{1 < p < \infty} L^p(X_2, \Sigma_2, \mu_2)$ and is a contraction on $L^p(X_1, \Sigma_1, \mu_1)$ into $L^p(X_2, \Sigma_2, \mu_2)$ for each $p: 1 < p < \infty$. We denote this class of operators by $\mathcal{D}((X_1, \Sigma_1, \mu_1), (X_2, \Sigma_2, \mu_2))$ or by any of $\mathcal{D}, \mathcal{D}(X_1, X_2)$, etc., where the terms omitted are understood from the context. The following remarks serve to extend the domain of definition of T and to establish a well-defined adjoint operation. (i) Given $T \in \mathcal{D}(X_1, X_2)$ there exists $T^* \in \mathcal{D}(X_2, X_1)$ such that regarded as an operator defined on any of the $L^p(X_2)$ spaces ($1 < p < \infty$), T^* is the adjoint of T . (ii) T has a unique extension to a contraction on $L^1(X_1)$ into $L^1(X_2)$, which we also denote T . A similar remark holds for T^* . (iii) As an operator on $L^1(X_2)$ into $L^1(X_1)$, the contraction T^* has for adjoint a contraction on $L^\infty(X_1)$ into $L^\infty(X_2)$ which agrees with T on $L^p(X_1) \cap L^\infty(X_1)$ for any $p: 1 \leq p < \infty$. We therefore finally extend T to a contraction on $L^\infty(X_1)$ into $L^\infty(X_2)$. Similar remarks apply to T^* . Denoting the extensions also by T and T^* , we call T^* the adjoint of T (hence T is the adjoint of T^*). An operator in $\mathcal{D}(X, X)$ is said to be self-adjoint if it is equal to its own adjoint in the above sense of the term.

An operator $S \in \mathcal{D}(X_1, X_2)$ is *doubly stochastic* if S is a positive operator such

that $S1 = 1$ a.e. (μ_2) and $S^*1 = 1$ a.e. (μ_1) , where 1 is the constant function assuming everywhere the value 1. (For an operator T on $L^\infty(X_1)$ into $L^\infty(X_2)$ such that $T1 = 1$, positivity of T is equivalent to contractivity (in L^∞ norm) [19].) Note that if S is doubly stochastic, so is S^* . It follows that if $\mu_1(X_1) = \mu_2(X_2) < \infty$, then any operator $S \in \mathcal{D}(X_1, X_2)$ for which $S1 = 1$ is doubly stochastic.

Given $f \in \bigcup_{1 \leq p \leq \infty} L^p$ and a σ -field $\mathcal{F} \subset \Sigma$ for which (X, \mathcal{F}, μ) is σ -finite, the conditional expectation of f with respect to \mathcal{F} , denoted $E\{f|\mathcal{F}\}$, is defined as that function measurable on (X, \mathcal{F}) for which

$$\int_B E\{f|\mathcal{F}\} d\mu = \int_B f d\mu \quad (B \in \mathcal{F}, \mu(B) < \infty).$$

By the theory of Radon-Nikodym derivatives [11], $E\{f|\mathcal{F}\}$ is well defined—unique to within sets of μ -measure zero (cf. [15], [7]). By the argument of [8, p.22] if $f \in L^p(X, \Sigma, \mu)$, $h \in L^q(X, \mathcal{F}, \mu)$, $1/p + 1/q = 1$, then $E\{fh|\mathcal{F}\} = hE\{f|\mathcal{F}\}$. It is readily verified (cf. [19]) that conditional expectation is a self-adjoint, idempotent, doubly stochastic operator.

A decreasing martingale on (X, Σ, μ) is a sequence $\{f_n\}_{n=1}^\infty$ of functions and a sequence $\{\mathcal{F}_n\}_{n=1}^\infty$ of σ -subfields of Σ satisfying

- (i) (X, \mathcal{F}_n, μ) is σ -finite,
 - (ii) $\mathcal{F}_n \supset \mathcal{F}_{n+1}$,
 - (iii) $f_1 \in \bigcup_{1 \leq p \leq \infty} L^p(X, \Sigma, \mu)$, and
 - (iv) $E\{f_n|\mathcal{F}_{n+1}\} = f_{n+1}$, $(n = 1, 2, \dots)$
- (cf. [8], [15], [7] and [13]).

A decreasing submartingale is defined the same as a decreasing martingale, except for condition (iv), which reads

$$(iv)_s \quad E\{f_n|\mathcal{F}_{n+1}\} \geq f_{n+1}, \quad (n = 1, 2, \dots).$$

The basic operator composition used is $T_1^* T_2^* \dots T_n^* T_n T_{n-1} \dots T_1$, where each $T_k \in \mathcal{D}(X_{k-1}, X_k)$, so following [9] we abbreviate: T_{1n} denotes $T_n T_{n-1} \dots T_1$, and T_{1n}^* denotes $T_1^* T_2^* \dots T_n^*$, $(n = 1, 2, \dots)$. We usually denote doubly stochastic operators, positive operators, and unspecified operators by S , P , and T respectively. T_{10} , P_{10} , and S_{10} denote the identity operator I .

By way of introduction we now state the following results:

THEOREM (ROTA [21], [22]). Let (Ω, P) be a probability space, and

$$S_n \in \mathcal{D}(\Omega, P) \quad (n = 1, 2, \dots)$$

be doubly stochastic. For each $p \in [1, \infty)$ and $f \in L^p(\Omega, P)$ satisfying $\int_\Omega |f| |\log^+ |f|| dP < \infty$, $\lim_{n \rightarrow \infty} S_{1n}^* S_{1n} f$ exists a.e. (dP) and boundedly in $L^p(\Omega, P)$.

THEOREM (DOOB [9]). Let P_n be a positive contraction on $L^1(X_{n-1}, \mu_{n-1})$ into $L^1(X_n, \mu_n)$ $(n = 1, 2, \dots)$. Let $h \in L^1(X_0, \mu_0)$, $h \geq 0$. Then if $d\mu'_n = (P_{1n} h) d\mu_n$, and if f is a function on X_0 satisfying

$$\int_{X_0} |f| (\log^+ |f|) h d\mu_0 < \infty,$$

it follows that

$$\lim_{n \rightarrow \infty} P_1^* P_2^* \cdots P_n^* \left(\frac{P_{1n}(fh)}{P_{1n}(h)} \right)$$

exists a.e. (μ_0) where $h > 0$ and also in the $L^1(X_0, h d\mu_0)$ topology. Here P_k^* is the linear extension of L^1 norm ≤ 1 of P_k^* to an operator from $L^1(X_k, (P_{1k}h)d\mu_k)$ into $L^1(X_{k-1}, (P_{1,k-1}h)d\mu_{k-1})$.

The main results of the present work are summarized in the following

MAIN THEOREM. Let $T_n \in \mathcal{D}(X, X)$ ($n = 1, 2, \dots$). Then if $g \in L^p$ ($1 < p < \infty$),

$$(0.1) \quad \int_X \left(\sup_{n \geq 0} |T_{1n}^* T_{1n} g| \right)^p d\mu \leq \left(\frac{p}{p-1} \right)^p \int_X |g|^p d\mu.$$

If $g \in L^1$ and satisfies $\int_X |g| \log^+ |g| d\mu < \infty$, then

$$(0.2) \quad \int_X \left(\sup_{n \geq 0} |T_{1n}^* T_{1n} g| \right) h d\mu \leq \frac{e}{e-1} \int_X (h + |g| \log^+ |g|) d\mu < \infty,$$

for any $h \in L^1$ satisfying $1 \geq h \geq \min(|g|, 1)$.

Under the additional hypothesis that the operators T_n are positive operators (which we now denote P_n) it follows that

$$(0.3) \quad \lim_{n \rightarrow \infty} P_{1n}^* P_{1n} g \text{ exists a.e. } (\mu)$$

for any $g \in \bigcup_{1 \leq p < \infty} L^p$ satisfying $\int_X |g| \log^+ |g| d\mu < \infty$.

The theorem of Doob is based on that of Rota. Both authors transfer the setting of the problem of convergence to a representation space. In this context they respectively show that $\{S_{1n}f\}_0^\infty$ and $\{P_{1n}(fh)/P_{1n}(h)\}_0^\infty$ are martingales with respect to a decreasing family of subfields, whose suprema are integrable. They then note that $S_{1n}^*, P_1^* P_2^* \cdots P_n^*$ are conditional expectations with respect to a fixed subfield independent of n , and these are known to preserve bounded almost everywhere convergence in L^1 .

Our theorem is based on an argument (Theorem 2 below) which is a submartingale version of the proof of Doob's theorem. Given $g \in L^1$ satisfying $\int_X |g| \log^+ |g| d\mu < \infty$, we factor g as the product $f \cdot h$, where f and h are as in Doob's theorem, and $h \leq 1$. We then show that the denominators $\{P_{1n}h\}_{n=0}^\infty$ of Doob's ratio form a uniformly bounded submartingale, whence

$$P_1^* P_2^* \cdots P_n^* \left((P_{1n}h) \frac{P_{1n}(fh)}{P_{1n}h} \right) = P_{1n}^* P_{1n}(fh)$$

converges as desired. A vector-valued Schwarz inequality of Rota [20] is used to

prove the submartingale property. In obtaining the maximal estimates, we make use of a lemma of Dunford and Schwartz [10] to associate with an operator $T \in \mathcal{D}(X, X)$ a positive operator $P \in \mathcal{D}(X, X)$ for which $|Tf| \leq P(|f|)$ a.e. (μ) ($f \in \bigcup_{1 \leq p \leq \infty} L^p$).

In §1 we prove Rota's theorem directly on a space of infinite measure (Theorem 1), following Doob's proof. That is, we obtain the *pointwise* and *maximal* results ((0.3) and (0.1)) for $p > 1$ of the Main Theorem, under the hypothesis that all operators are *doubly stochastic*. The machinery of the proof of Theorem 1 is used in §2 to prove Theorem 2, which gives the *bounded convergence* ((0.2) and (0.3)) stated in the Main Theorem for $g \in L^1$ satisfying $\int |g| \log^+ |g| d\mu < \infty$, under the hypothesis that all operators are *positive operators* in \mathcal{D} . Under this latter hypothesis, Corollary 1 gives the *pointwise convergence* ((0.3)) for functions in L^p ($1 < p < \infty$). In §3 the assumption of positivity is dropped, and the *maximal inequalities* ((0.1) and (0.2)) of the Main Theorem are obtained (Theorem 3). §4 contains a quick proof of the strong convergence of $T_{1n}^* T_{1n}$, based on the particular form of the operator composition. Also, properties of the limit operator are discussed. In §5 the continuous parameter versions of our results are stated. The Appendix contains the appropriate statements of certain auxiliary results.

1. Convergence on a space of infinite measure, for doubly stochastic operators.

Our argument here is simply the infinite measure space analog of the argument Doob used in his proof of Rota's theorem [9]. The latter argument followed the basic method of Rota but utilized a different representation theory and applied to the more general case of doubly stochastic operators between different spaces.

THEOREM 1. *Let $S_n \in \mathcal{D}((X_{n-1}, \Sigma_{n-1}, \mu_{n-1}), (X_n, \Sigma_n, \mu_n))$ be doubly stochastic ($n = 1, 2, \dots$). Then given $p: 1 < p < \infty$ and $f \in L^p(X_0, \Sigma_0, \mu_0)$,*

- (i) $\lim_{n \rightarrow \infty} S_{1n}^* S_{1n} f$ exists a.e. (μ_0) and
- (ii) $\int_{X_0} (\sup_{n \geq 0} |S_{1n}^* S_{1n} f|)^p d\mu_0 \leq (p/(p-1))^p \int_{X_0} |f|^p d\mu_0$.

Proof. We use the representation theory of the Appendix, in which a one-to-one, linear, order preserving map from $\bigcup_{1 \leq p \leq \infty} L^p(X_n)$ onto $\bigcup_{1 \leq p \leq \infty} L^p(X_n^\#)$ such that $g \leftrightarrow g^\#$ preserves L^p norms ($1 \leq p \leq \infty$) is established ($n = 0, 1, \dots$). Moreover, pointwise a.e. convergence of sequences of functions is preserved under the correspondence $g \leftrightarrow g^\#$. Consequently the limiting and maximal behavior we seek to establish is invariant under the representation. Throughout this proof we work on the representation spaces $X_n^\#$ ($n = 0, 1, \dots$), dropping the notation “ $\#$ ”. As noted in the Appendix, the operators S_k^* can be represented as kernel operators (indeed this is the reason for passing to the representation spaces): for any $f \in \bigcup_{1 \leq p \leq \infty} L^p(X_k)$

$$(1.1) \quad (S_k^* f)(\xi_{k-1}) = \int_{X_k} f(\xi_k) Q_k(\xi_{k-1}, d\xi_k) \quad \text{a.e. } (\mu_{k-1}).$$

$Q_k(\cdot, \cdot)$ here is a real valued function on $X_{k-1} \times \Sigma_k$: for each fixed $\xi_{k-1} \in X_{k-1}$ it is a probability measure on (X_k, μ_k) , and for each fixed $A \in \Sigma_k$ it is measurable on (X_{k-1}, μ_{k-1}) — indeed nonnegative and essentially bounded also.

By an easy extension of the Ionescu Tulcea theorem [14] on product space measures,

$$P(\Lambda) = \int_{X_0} \mu_0(dx_0) \int_{X_1} Q_1(x_0, dx_1) \int \cdots \int_{X_k} \chi_{\Lambda|k}(x_0, \dots, x_k) Q_k(x_{k-1}, dx_k)$$

(where $\Lambda \in \Sigma_0 \times \cdots \times \Sigma_k \times X_{k+1} \times X_{k+2} \times \cdots$, $\Lambda|k$ is the projection of Λ on $X_0 \times \cdots \times X_k$, and $\chi_A(\cdot)$ assumes the value 1 on A and vanishes elsewhere) determines a totally σ -finite positive measure P on $\Omega = X_0 \times X_1 \times \cdots$. Letting x_n denote the n th coordinate function on Ω , we note that

$$(1.2) \quad \int_{\Omega} g(x_0, \dots, x_n) dP = \int_{X_0} d\mu_0 \int_{X_1} Q_1(x_0, dx_1) \cdots \int_{X_{n-1}} Q_{n-1}(x_{n-2}, dx_{n-1}) \int_{X_n} g(x_0, \dots, x_n) Q_n(x_{n-1}, dx_n)$$

for any function g measurable on $(X_0 \times \cdots \times X_n, \Sigma_0 \times \cdots \times \Sigma_n)$ and satisfying $g(x_0, \dots, x_n) \in L^1(\Omega, P)$: (1.2) is true for finite linear combinations of characteristic functions of sets of form $A_1 \times \cdots \times A_n \in \Sigma_1 \times \cdots \times \Sigma_n$, $\mu_i(A_i) < \infty$, hence for all such g by the monotone convergence theorem.

(1.2) is valid in particular for g of form $\chi_A(x_0)h(x_n)$, where $h \in L^\infty(X_n, \mu_n)$ and $A \subset X_0$, $\mu_0(A) < \infty$. Therefore by (1.1) and (1.2),

$$\int_{\Omega} \chi_A(x_0)h(x_n) dP = \int_{\Omega} \chi_A(x_0)S_{1n}^*h(x_0) dP$$

so that $(S_{1n}^*h)(x_0)$ is the conditional expectation $E\{h(x_n)|x_0\}$ of $h(x_n)$ (on Ω) with respect to the σ -field $\mathcal{B}\{x_0\}$ (in Ω). We now show

$$(1.3) \quad (S_{1n}^*h)(x_0) = E\{h(x_n)|x_0\} \text{ a.e. } (dP) \left(h \in \bigcup_{1 \leq p \leq \infty} L^p(X_n, \mu_n) \right).$$

(1.3) is true for $h \in L^\infty(X_n, \mu_n)$ (i.e. $h(x_n) \in L^\infty(\Omega, P)$). The dominated convergence theorem for positive operators (Appendix) can be applied to each side of (1.3) to prove (1.3) true for $h \in \bigcup_{1 \leq p < \infty} L^p(X_n, \mu_n)$, provided such an h satisfies

$$h(x_n) \in \bigcup_{1 \leq p < \infty} L^p(\Omega, P)$$

(\subset the domain of definition of the conditional expectation operator):

The σ -field on Ω determined by x_0 is σ -finite, since x_0 has distribution μ_0 . Similarly, the σ -fields on Ω determined by each x_n , $n = 1, 2, \dots$, are σ -finite, since the distribution of each x_n is μ_n , a σ -finite measure:

$$\begin{aligned}
 (1.4) \quad P(\{\omega: x_n(\omega) \in A\}) &= \int_{\Omega} E\{\chi_A(x_n) | x_0\} dP = \int_{x_0} S_{1n}^* \chi_A d\mu_0 \\
 &= \int_X \chi_A S_{1n}(1) d\mu_n = \mu_n(A) \quad (A \in \Sigma_n).
 \end{aligned}$$

Thus for any $h \in \bigcup_{1 \leq p < \infty} L^p(X_n, \mu_n)$, $\int_{\Omega} |h(x_n)|^p dP = \int_{x_n} |h|^p d\mu_n$. Therefore (1.3) holds for any such function h .

Next we note that for any $f \in \bigcup_{1 \leq p \leq \infty} L^p(X_0, \mu_0)$,

$$(1.5) \quad (S_{1n}f)(x_n) = E\{f(x_0) | x_n\} \text{ a.e. } (dP).$$

Let $A \in \Sigma_n$, $\mu_n(A) < \infty$. Then

$$\begin{aligned}
 \int_{\Omega} f(x_0) \chi_A(x_n) dP &= \int_{\Omega} E\{f(x_0) \chi_A(x_n) | x_0\} dP \\
 &= \int_{x_0} f S_{1n}^* \chi_A d\mu_0 = \int_{x_n} (S_{1n}f) \chi_A d\mu_n \\
 &= \int_{\Omega} S_{1n}f(x_n) \chi_A(x_n) dP, \text{ proving (1.5).}
 \end{aligned}$$

Finally we prove the following lemma which establishes $\{S_{1n}f(x_n)\}_{n=0}^{\infty}$ as a decreasing martingale, the source of all decent limiting behavior here.

LEMMA 1. Given $f \in \bigcup_{1 \leq p \leq \infty} L^p(X_0, \Sigma_0, \mu_0)$,

$$(1.6) \quad (S_{1n}f)(x_n) = E\{f(x_0) | x_n, x_{n+1}, \dots\} \text{ a.e. } (dP), \quad n = 0, 1, \dots$$

Proof⁽²⁾. We must show the reverse Markov identity

$$(1.7) \quad E\{f(x_0) | x_n\} = E\{f(x_0) | x_n, x_{n+1}, \dots\} \text{ a.e. } (dP)$$

for the process consisting of coordinate functions on a product space whose measure is defined in terms of a sequence of regular conditional probabilities and a totally σ -finite initial measure μ_0 . A set $A \subset X_0$ of finite measure determines a subspace $(\Omega_A, \Sigma_A, P_A)$ of (Ω, Σ, P) , where $\Omega_A = A \times X_1 \times X_2 \times \dots$, Σ_A consists of those subsets of Ω_A belonging to Σ , and P_A is the restriction of P to (Ω_A, Σ_A) . Now (1.7) is well known in the case of a finite initial measure [8]. Hence using the definition of conditional expectation there exists a function $\phi_A \in \bigcup_{1 \leq p < \infty} L^p(\Omega, \Sigma, P)$, measurable with respect to $\mathcal{B}\{x_n\}$, and vanishing on $\Omega - \Omega_A$, which satisfies

$$\int_{B \cap \Omega_A} \phi_A(x_n) dP = \int_{B \cap \Omega_A} f(x_0) dP \quad (B \in \mathcal{B}\{x_n, x_{n+1}, \dots\}, P(B) < \infty).$$

Now letting A_n increase to X_0 , $\mu_0(A_n) < \infty$, it is clear that $\phi = \lim_{n \rightarrow \infty} \phi_{A_n}$ is measurable with respect to $\mathcal{B}\{x_n\}$ and satisfies

(2) We thank the referee for suggesting the following argument, which is substantially shorter and clearer than our original line of reasoning.

$$\int_B \phi(x_n) dP = \int_B f(x_0) dP \quad (B \in \mathcal{B}\{x_n, x_{n+1}, \dots\}, P(B) < \infty),$$

proving (1.7).

Thus for $f \in L^p(X_0, \Sigma_0, \mu_0)$, $1 < p < \infty$, $\{S_{1n}f(x_n)\}_{n=0}^\infty$ is a decreasing martingale on (Ω, P) . By [13, p. 113] there exists an idempotent operator $T_\infty \in \mathcal{D}(\Omega, \Omega)$ such that

$$\lim_{n \rightarrow \infty} S_{1n}f(x_n) = T_\infty f \quad \text{a.e. } (dP)$$

for all $f \in \bigcup_{1 \leq p < \infty} L^p(X_0, \mu_0) (\subset \bigcup_{1 \leq p < \infty} L^p(\Omega, P))$. Moreover, by an argument similar to that of [8, pp. 317, 318], the maximal inequality of [13, p.113] implies

$$(1.8) \quad \int_\Omega \left(\sup_{n \geq 0} |(S_{1n}f)(x_n)| \right)^p dP \leq \left(\frac{p}{p-1} \right)^p \int_\Omega |f(x_0)|^p dP$$

for $f \in L^p(X_0, \mu_0)$, $1 < p < \infty$. The dominated convergence theorem for positive operators now implies, using (1.3),

$$\begin{aligned} \lim_{n \rightarrow \infty} S_{1n}^* S_{1n} f(x_0) &= \lim_{n \rightarrow \infty} E \left\{ (S_{1n}f)(x_n) \mid x_0 \right\} = E \left\{ \lim_{n \rightarrow \infty} (S_{1n}f)(x_n) \mid x_0 \right\} \\ &= E \{ T_\infty f \mid x_0 \} \quad \text{a.e. } (dP). \end{aligned}$$

Hence $\lim_{n \rightarrow \infty} S_{1n}^* S_{1n} f$ exists a.e. (μ_0) . This proves statement (i) of Theorem 1 on the representation spaces, hence on the original spaces.

To prove (ii) of Theorem 1 observe first that by (1.3),

$$\int_{X_0} \left(\sup_{n \geq 0} |S_{1n}^* S_{1n} f| \right)^p d\mu_0 = \int_\Omega \left(\sup_{n \geq 0} |E\{S_{1n}f(x_n) \mid x_0\}| \right)^p dP.$$

The latter integral is less than or equal to

$$\int_\Omega \left(E \left\{ \sup_{n \geq 0} |S_{1n}f(x_n)| \mid x_0 \right\} \right)^p dP$$

which in turn is less than or equal to

$$\int_\Omega \left(\sup_{n \geq 0} |S_{1n}f(x_n)| \right)^p dP$$

since conditional expectation is a contraction on L^p into itself. Applying (1.8) to this last integral gives (ii):

$$\int_{X_0} \left(\sup_{n \geq 0} |S_{1n}^* S_{1n} f| \right)^p d\mu_0 \leq \left(\frac{p}{p-1} \right)^p \int_{X_0} |f|^p d\mu_0$$

for each $f \in L^p(X_0, \mu_0)$, $1 < p < \infty$, completing the theorem.

REMARKS. (1) Convergence for $f \in L^1(X_0, \mu_0)$ is in general false, as Burkholder [3] has shown. For integrable f satisfying in addition $\int_{X_0} |f| \log^+ |f| d\mu_0 < \infty$, however, pointwise convergence holds, as we shall prove in §2. On a finite measure space, the arguments of [9], [21] give this result as well as bounded convergence in L^1 .

(2) Theorem 1 is stated for the case of doubly stochastic operators S_k between different function spaces $\bigcup_{1 \leq p \leq \infty} L^p(X_{k-1})$, $\bigcup_{1 \leq p \leq \infty} L^p(X_k)$ rather than on a given function space to itself, because this formulation is used in Theorem 2 below. (Actually, Theorem 2 uses only Doob's finite measure version of Theorem 1.) By taking a direct sum of the measure spaces $\{(X_k, \Sigma_k, \mu_k)\}_{k=0}^\infty$ and appropriately extending the operators S_k , the situation in Theorem 1 can be included in that of doubly stochastic operators on a given function space to itself.

EXAMPLE. Decreasing martingales. Let (f_1, \mathcal{F}) , (f_2, \mathcal{F}) , \dots be a decreasing, martingale on (X, Σ, μ) , and set $S_n = E\{\cdot | \mathcal{F}_n\}$ ($n = 1, 2, \dots$). Then $S_{1n}^* S_{1n} = S_n$, so the behavior of a decreasing martingale, which was used to determine that of the products $S_{1n}^* S_{1n}$, can in turn be described by such expressions. Examples of Jerison [15] of decreasing martingales on an infinite measure space show that

- (i) $S_{1n}^* S_{1n} g$ need not converge for $g \in L^\infty$; and
- (ii) $S_{1n}^* S_{1n} g$ can converge pointwise a.e. yet fail to converge in the L^1 norm, for some $g \in L^1 \cap L^\infty$. In this case, $(\sup_{n \geq 0} |S_{1n}^* S_{1n} g|)$ can not be integrable.

Neither of these situations can arise if $\mu(X) < \infty$ [8].

2. Pointwise convergence for positive contraction operators. The assumption of Theorem 1 that the operators S_n, S_n^* map 1 into 1 is now dropped, and similar limiting behavior is proved for the resulting positive contraction operators, which we denote P_n, P_n^* . As noted above it is no loss of generality to assume all functions are defined on the same measure space (X, Σ, μ) .

In order to prove our theorem for positive contraction operators, we represent $P_{1n}^* P_{1n} g$ as a conditional expectation (S_{1n}^*) acting on the product of a uniformly bounded submartingale $(P_{1n} h)$ and a martingale $(S_{1n} f)$, all on a finite measure space $((\Omega, P))$. This representation uses that of the finite measure version of Theorem 1, in which case bounded pointwise convergence in L^1 holds for a special subset of L^1 as remarked above. The transfer of our problem to the above context is effected by the techniques Doob [9] used in relating his ratio theorem to Rota's theorem.

THEOREM 2. Let $P_n \in \mathcal{D}(X, X)$ be a positive operator ($n = 1, 2, \dots$). Let $g \in L^1(X, \mu)$ satisfy $\int_X |g| \log^+ |g| d\mu < \infty$. Then

- (i) $\lim_{n \rightarrow \infty} P_{1n}^* P_{1n} g$ exists a.e. (μ) and
- (ii) $\int_X (\sup_{n \geq 0} |P_{1n}^* P_{1n} g|) h d\mu \leq (e/(e-1)) \int_X (h + |g| \log^+ |g|) d\mu < \infty$ for any $h \in L^1(X, \mu)$ satisfying $1 \geq h \geq \min(|g|, 1)$.

REMARK. The maximal inequality (ii) assumes a more useful form in special cases: if $\mu(X) < \infty$, h may be chosen $\equiv 1$; if $\{x: g(x) \neq 0\} \subset A$ and $\mu(A) < \infty$, h may be chosen $= \chi_A$.

We first prove the following lemma, the bulk of which is due to Doob [9] and which is of independent interest insofar as it illuminates the relation between positive contraction and doubly stochastic operators. In this connection we remark that the lemma is valid for any $g \in L^1(X, \mu)$ and any nonnegative $h \in L^1(X, \mu)$ which vanishes nowhere on the set $\{x: g(x) \neq 0\}$.

LEMMA 2. *Let g, h as in Theorem 2 be fixed. Then setting $f = g/h$ on $\{x: h(x) > 0\}$ and 0 elsewhere, there exist finite measure spaces (X_k, μ'_k) for which $X \subset X_k$ ($k = 0, 1, \dots$), $X_0 = X$, and doubly stochastic operators*

$$S_k \in \mathcal{D}((X_{k-1}, \mu'_{k-1}), (X_k, \mu'_k))$$

for which

$$(2.1) \quad S_{1k}^*[P_{1k}(h) S_{1k}(f)] = P_{1k}^* P_{1k} g \quad \text{a.e. } (\mu'_0) \quad (k = 1, 2, \dots).$$

Proof of Lemma 2. As a first step in returning to the context of Rota's theorem, namely doubly stochastic operators, consider the operators

$$P'_n(\cdot) = \frac{P_n((\cdot)P_{1,n-1}h)}{P_{1n}h}$$

and the measures

$$d\mu_n = (P_{1n}h) d\mu$$

of Doob's Ratio Theorem [9] (note $P_{10}h = h$). Note $\mu_n(X) < \infty$. For any $\phi \in L^\infty(X, \mu)$,

$$|P_n(\phi P_{1,n-1}h)| \leq \|\phi\|_\infty P_{1n}|h| = \|\phi\|_\infty P_{1n}h \quad \text{a.e. } (\mu),$$

so by the dominated convergence theorem for positive operators, $P_n(\psi P_{1,n-1}h)$ vanishes a.e. (μ) where $P_{1n}h$ does, for any $\psi \in \bigcup_{1 \leq p \leq \infty} L^p(X, \mu)$. P'_n is therefore a positive contraction on $L^\infty(X, \mu)$ into $L^\infty(X, \mu)$ mapping 1 into 1 a.e. (μ_n) . Moreover, P'_n is a (positive) contraction from $L^1(X, \mu_{n-1})$ into $L^1(X, \mu_n)$, since for $\Phi \geq 0$ and $\Phi \cdot P_{1,n-1}h \in L^1(X, \mu)$,

$$\int P'_n(\Phi) d\mu_n = \int P_n(\Phi \cdot P_{1,n-1}h) d\mu = \int \Phi \cdot P_{1,n-1}h \cdot P_n^* 1 d\mu \leq \int \Phi d\mu_{n-1}.$$

Therefore $P'_n \in \mathcal{D}((X, \mu_{n-1}), (X, \mu_n))$.

A similar computation shows that if $\phi \in L^\infty(X, \mu)$ ($\subset L^1(X, \mu_n)$) and $\psi \in L^\infty(X, \mu)$ ($\subset L^1(X, \mu_{n-1})$) then

$$\int \phi(P'_n\psi) d\mu_n = \int (P_n^*\phi)\psi d\mu_{n-1}.$$

Therefore the positive contraction $P_n'^*$ on $L^1(X, \mu_n)$ to $L^1(X, \mu_{n-1})$ agrees with P_n^* on a common set dense in $L^1(X, \mu_n)$, hence is the extension thereto of P_n^* . By this fact there can be no measure zero irregularities.

We now observe that the measure spaces $(X, \mu_k) = (X, P_{1k}h \, d\mu)$ ($k = 0, 1, \dots$) are *finite* measure spaces and

$$\mu_k(X) = \int_X P_{1k}h \, d\mu = \int_X P_{1,k-1}h P_k^* 1 \, d\mu \leq \int_X P_{1,k-1}h \, d\mu = \mu_{k-1}(X).$$

There cannot exist doubly stochastic operators $S_n \in \mathcal{D}((X, \mu_{n-1}), (X, \mu_n))$ unless $\mu_n(X) = \mu_{n-1}(X)$ ($n = 1, 2, \dots$), for such operators and their adjoints map 1 into 1. Consequently atoms ρ_1, ρ_2, \dots are now adjoined to X to obtain a sequence $\{(X_n, \Sigma_n, \mu'_n)\}_{n=0}^\infty$ of finite measure spaces, each of measure $\mu_0(X)$, as follows [9]:

Let

$$X_n = X \cup \rho_n,$$

$$\Sigma_n = \mathcal{B}\{\Sigma, \rho_n\},$$

$$\mu'_n \text{ agree with } \mu_n \text{ on measurable subsets of } X,$$

$$\mu'_n(\rho_n) = \mu_0(X) - \mu_n(X) \quad (n = 1, 2, \dots);$$

and

$$\mu'_0 = \mu_0, X_0 = X, \Sigma_0 = \Sigma.$$

Note $d\mu'_0$ on X_0 is just $h \, d\mu$ on X .

Again following [9] we define a sequence of doubly stochastic operators $S_n \in \mathcal{D}((X_{n-1}, \Sigma_{n-1}, \mu'_{n-1}), (X_n, \Sigma_n, \mu'_n))$, extending first the adjoint operators:

$$S_n^* \phi = \begin{cases} \phi(\rho_n)(1 - P_n'^* 1) + P_n'^* \phi & \text{on } X_{n-1} - \rho_{n-1} \quad (= X), \\ \phi(\rho_n) & \text{at } \rho_{n-1} \end{cases}$$

for $\phi \in L^1(X_n, \mu'_n)$. S_n^* is clearly a positive operator mapping 1 into 1 a.e. (μ'_{n-1}) , and is a contraction on $L^1(X_n, \mu'_n)$ into $L^1(X_{n-1}, \mu'_{n-1})$. Since $\{(X_k, \mu'_k)\}_{k=0}^\infty$ are finite measure spaces of equal measure, S_n^* (hence S_n) is doubly stochastic ($n = 1, 2, \dots$). It can be shown that S_n satisfies:

$$S_n \psi = \begin{cases} P_n' \psi & \text{on } X_n - \rho_n \quad (= X), \\ \left\{ \int_X (1 - P_n'^* 1) \psi \, d\mu_{n-1} + \psi(\rho_{n-1}) [\mu_0(X) - \mu_{n-1}(X)] \right\} / (\mu_0(X) - \mu_n(X)) & \text{at } \rho_n \end{cases}$$

for $\psi \in L^1(X_{n-1}, \mu'_{n-1})$. In the expressions for S_n and S_n^* above, the domain and range of $P_n', P_n'^*$ are understood to be among $\bigcup_{1 \leq p \leq \infty} L^p(X, \mu_{n-1}), \bigcup_{1 \leq p \leq \infty} L^p(X, \mu_n)$. We note in particular the important relation

$$(2.2) \quad (P_{1k}h)(\rho_k) = 0 \quad (k = 1, 2, \dots).$$

Using (2.2), the definitions of P_n', S_n , and S_n^* , and the fact that $P_n'^*$ is the extension of P_n^* , we obtain

$$\begin{aligned}
S_{1n}^*[P_{1n}(h)S_{1n}(f)] &= P_{1n}^*[P_{1n}(h)S_{1n}(f)] \\
&= P_{1n}^*[P_{1n}(h)P'_{1n}(f)] = P_{1n}^* \left[P_{1n}(h) \frac{P_{1n}(fh)}{P_{1n}(h)} \right] \\
&= P_{1n}^*P_{1n}(fh) = P_{1n}^*P_{1n}(g) \text{ a.e. } (\mu'_0) \text{ on } X_0.
\end{aligned}$$

This concludes the proof of the lemma.

Proof of Theorem 2. Let g, h satisfying the hypotheses of the theorem be fixed. Since among the h satisfying these hypotheses there are some which vanish nowhere, to prove (i) it is enough (noting $d\mu'_0 = h d\mu$) to show:

$$(i') \quad \lim_{n \rightarrow \infty} S_{1n}^*[P_{1n}(h)S_{1n}(f)] \text{ exists a.e. } (\mu'_0).$$

Moreover, since $|f|h = |g|$ and $\log^+ |f| = \log^+ |g|$, to prove (ii) it suffices to show:

$$\begin{aligned}
(ii') \quad & \int_{X_0} \left(\sup_{n \geq 0} |S_{1n}^*[P_{1n}(h)S_{1n}(f)]| \right) d\mu'_0 \\
& \leq \frac{e}{e-1} \int_{X_0} (1 + |f|\log^+ |f|) d\mu'_0 < \infty.
\end{aligned}$$

Applying the representation theoretic technique used in proving Theorem 1, we transfer the setting of the problem to the representation space $(\Omega^\#, P^\#)$ used there for the study of the doubly stochastic operators S_n . Again $\Omega^\# = X_0^\# \times X_1^\# \times \cdots$, $P^\#$ is the measure on $\Omega^\#$ determined by the initial (now finite) measure $\mu_0'^\#$ and the regular conditional probabilities $Q_n^\#(x_{n-1}, dx_n)$, the x_n are the coordinate functions on $\Omega^\#$, with distribution $\mu_n'^\#$ (by (1.4)), and

$$S_n^{*\#} \psi^\#(\xi_{n-1}) = \int_{X_n^\#} \psi^\#(\xi_n) Q_n^\#(\xi_{n-1}, d\xi_n) \quad (\psi^\# \in L^1(X_n^\#, \mu_n'^\#)).$$

Moreover, the correspondence $g \leftrightarrow g^\#$ between functions in $\bigcup_{1 \leq p \leq \infty} L^p(X_n, \mu_n')$ and in $\bigcup_{1 \leq p \leq \infty} L^p(X_n^\#, \mu_n'^\#)$ preserves L^p norms and pointwise a.e. sequential convergence. Thus the limiting behavior claimed for $S_{1n}^*[P_{1n}(h)S_{1n}(f)]$ in (i') and (ii') is equivalent to that of $S_{1n}^{*\#}[P_{1n}^\#(h^\#)S_{1n}^\#(f^\#)]$ on $(X_0^\#, \mu_0'^\#)$. We therefore prove Theorem 2 in the context of the representation space, and omit, for the rest of the proof, the “ $\#$ ” notation.

From the proof of Theorem 1, $\{S_{1n}f(x_n)\}_{n=0}^\infty$ is a decreasing martingale on (Ω, P) with respect to the sequence of subfields $\mathcal{B}\{x_n, x_{n+1}, \dots\}$ ((1.6)). Since

$$\int_{\Omega} |f(x_0)| \log^+ |f(x_0)| dP = \int_X (|f| \log^+ |f|) h d\mu = \int_X |g| \log^+ |g| d\mu < \infty$$

and

$$P(\Omega) = \int_{\Omega} dP = \int_{X_0} d\mu'_0 = \int_X h d\mu < \infty,$$

martingale theory on finite measure spaces [8] implies that $\{S_{1n}f(x_n)\}_{n=0}^{\infty}$ converges a.e. (dP) and

$$(2.3) \quad \int_{\Omega} \sup_{n \geq 0} |S_{1n}f(x_n)| dP \leq \frac{e}{e-1} \int_{\Omega} (1 + |f(x_0)| \log^+ |f(x_0)|) dP < \infty.$$

We now show (ii'). From the inequality $|P_{1n}h(x_n)| \leq 1$ a.e. (dP) and the fact ((1.3)) that $S_{1n}^* \phi(x_0) = E\{\phi(x_n) | x_0\}$ for $\phi \in L^1(X_n, \mu_n)$ ($n = 0, 1, \dots$),

$$\begin{aligned} & \int_{\Omega} \left(\sup_{n \geq 0} |S_{1n}^*[P_{1n}(h) S_{1n}(f)](x_0)| \right) dP \\ & \leq \int_{\Omega} \sup_{n \geq 0} (S_{1n}^* |S_{1n}f|(x_0)) dP \\ & = \int_{\Omega} \sup_{n \geq 0} E\{|S_{1n}f(x_n)| \mid x_0\} dP. \end{aligned}$$

Using now the positivity of the conditional expectation operation and (2.3), the last term above is less than or equal to

$$\begin{aligned} & \int_{\Omega} E \left\{ \sup_{n \geq 0} |S_{1n}f(x_n)| \mid x_0 \right\} dP \\ & = \int_{\Omega} \sup_{n \geq 0} |S_{1n}f(x_n)| dP \\ & \leq \frac{e}{e-1} \int_{\Omega} (1 + |f(x_0)| \log^+ |f(x_0)|) dP < \infty. \end{aligned}$$

Since integrals on (Ω, P) of functions of x_0 are integrals over (X_0, μ'_0) , these inequalities prove (ii') on the representation space $(X^{\#})$, hence on the original space (X) .

We finally show that

$$\lim_{n \rightarrow \infty} S_{1n}^*[P_{1n}(h) S_{1n}(f)](x_0) \text{ exists a.e. } (dP),$$

implying (i') above, hence (i) of our theorem (since the distribution of x_0 is $\mu'_0 = h d\mu$). Since $\{S_{1n}f(x_n)\}_{n=0}^{\infty}$ is boundedly convergent in $L^1(\Omega, P)$ and conditional expectation preserves bounded L^1 convergence,

$$P_{1n}^* P_{1n} g(x_0) = S_{1n}^*(P_{1n} h S_{1n} f)(x_0) = E\{(P_{1n} h S_{1n} f)(x_n) | x_0\}$$

will converge as desired if it is shown that $\{P_{1n}h(x_n)\}_{n=0}^{\infty}$ converges a.e. (dP), since $\{P_{1n}h(x_n)\}_{n=0}^{\infty}$ is uniformly bounded (by one). To complete the theorem we now show $\{P_{1n}h(x_n)\}_{n=0}^{\infty}$ is a decreasing submartingale on (Ω, P) . By submartingale theory on a finite measure space ([8], where a submartingale is called a semi-martingale), it will follow that $\lim_{n \rightarrow \infty} P_{1n}h(x_n)$ exists a.e. (dP).

LEMMA 3. $\{P_{1n}h(x_n)\}_{n=0}^\infty$ is a decreasing submartingale on (Ω, P) with respect to the sequence of subfields $\mathcal{B}\{x_n, x_{n+1}, \dots\}$.

Proof. We first remark that by (1.4)

$$\begin{aligned} \int_{\Omega} g(x_n) S_n f(x_n) dP &= \int_{X_n} g S_n f d\mu'_n \\ &= \int_{X_{n-1}} f S_n^* g d\mu'_{n-1} = \int_{\Omega} f(x_{n-1}) S_n^* g(x_{n-1}) dP \\ (f \in L^p(X_{n-1}, \mu'_{n-1}), g \in L^q(X_n, \mu'_n), 1 \leq p \leq \infty, 1/p + 1/q = 1). \end{aligned}$$

Let C belong to $\mathcal{B}\{x_k\}$, $C|_k (\in \Sigma_k)$ be the projection of C on X_k . In the following we denote both these sets simply by C , interpreting $\chi_C(x_k)$ as $\chi_{C|_k}(x_k)$.

Next we note

$$(2.4) \quad (S_{r+1}^* S_{r+2}^* \cdots S_n^* \phi)(x_r) = E\{\phi(x_n) | x_r\} \text{ a.e. } (dP)$$

for all $\phi \in L^\infty(X_n, \mu'_n)$, $n > r$. This is immediate upon considering the stochastic kernel representation (1.1) of the S_k^* and the following (cf. (1.2)):

$$\begin{aligned} \int_{\Omega} \phi(x_n) \chi_A(x_r) dP &= \int_{X_0} \mu'_0(d\xi_0) \int_{X_1} Q_1(\xi_0, d\xi_1) \cdots \\ &\quad \int_{X_r} \chi_A(\xi_r) Q_r(\xi_{r-1}, d\xi_r) \cdots \int_{X_n} \phi(\xi_n) Q_n(\xi_{n-1}, d\xi_n) \end{aligned}$$

where $A \in \mathcal{B}\{x_r\}$.

From (2.4) we infer

$$(2.5) \quad E\{f(x_k) | x_n\} = (S_n S_{n-1} \cdots S_{k+1} f)(x_n) \text{ a.e. } (dP) \quad (n > k)$$

for $f \in L^\infty(X_k, \mu'_k)$, since

$$\begin{aligned} \int_{\Omega} f(x_k) \chi_B(x_n) dP &= \int_{\Omega} E\{f(x_k) \chi_B(x_n) | x_k\} dP \\ &= \int_{\Omega} f(x_k) (S_{k+1}^* S_{k+2}^* \cdots S_n^* \chi_B)(x_k) dP \\ &= \int_{\Omega} S_n S_{n-1} \cdots S_{k+1} f(x_n) \chi_B(x_n) dP \end{aligned}$$

for any $B \in \mathcal{B}\{x_n\}$.

From the Markovian nature of $\{x_n\}_{n=0}^\infty$,

$$E\{P_{1r}h(x_r) | x_n\} = E\{P_{1r}h(x_r) | x_n, x_{n+1}, \dots\},$$

so to show $\{P_{1n}h(x_n)\}_{n=0}^\infty$ is a decreasing submartingale it suffices to prove

$$E\{P_{1r}h(x_r) | x_n\} \geq P_{1n}h(x_n)$$

a.e. (dP) , $(r < n)$. This we do by proving that for each $A \in \mathcal{B}\{x_n\}$,

$$\begin{aligned} \int_{\Omega} \chi_A(x_n) P_{1r}h(x_r) dP &= \int_X \chi_A \cdot P_n P_{n-1} \cdots P_{r+1} ((P_{1r}h)^2) d\mu \\ (2.6) \qquad \qquad \qquad &\geq \int_X \chi_A \cdot (P_n P_{n-1} \cdots P_1 h)^2 d\mu \\ &= \int_{\Omega} \chi_A(x_n) P_{1n}h(x_n) dP. \end{aligned}$$

Since $P_{1r}h(x_r) \geq 0$ a.e. (dP) and $P_{1n}h(\rho_n) = 0$ ((2.2)), we may assume $\rho_n \notin A$, or $A \subset X_n - \rho_n = X$.

The inequality in (2.6) follows from the observation of Rota [20, p.557] that

$$P_k(\phi^2) \geq (P_k\phi)^2 \text{ a.e. } (\mu) \quad (\phi \in L^\infty(X, \mu), k = 1, 2, \dots)$$

(simply apply Schwarz's inequality to the kernel representation of P_k afforded by the Appendix).

To prove the equalities in (2.6), we use first (2.5), obtaining

$$\begin{aligned} \int_{\Omega} \chi_A(x_n) P_{1r}h(x_r) dP &= \int_{\Omega} \chi_A(x_n) E\{P_{1r}h(x_r) | x_n\} dP \\ &= \int_{X_n} \chi_A S_n S_{n-1} \cdots S_{r+1} (P_{1r}h) d\mu'_n. \end{aligned}$$

Taking adjoints and remarking $P_{1r}h(\rho_r) = 0$, this is equal to

$$\int_{X_r - \rho_r} (S_{r+1}^* S_{r+2}^* \cdots S_n^* \chi_A) (P_{1r}h) d\mu_r,$$

for $X_r - \rho_r$ is just X , and the restriction of μ'_r to X is μ_r . Expressing $d\mu_r = (P_{1r}h) d\mu$ and using the definition of S_k^* , this becomes

$$\int_X (P_{r+1}'^* P_{r+2}'^* \cdots P_n'^* \chi_A) (P_{1r}h)^2 d\mu,$$

since $\chi_A(\rho_n) = 0$. But by definition, $P_k'^*$ agrees with P_k^* on $L^\infty(X_k, \mu_k)$, so the above integral is just

$$\int_X (P_{r+1}^* P_{r+2}^* \cdots P_n^* \chi_A) (P_{1r}h)^2 d\mu,$$

which in turn equals

$$\int_X \chi_A P_n P_{n-1} \cdots P_{r+1} ((P_{1r}h)^2) d\mu,$$

giving the first equality in (2.6). Using the definition of μ_n , μ'_n , and the fact that $\rho_n \notin A$,

$$\begin{aligned} \int_X \chi_A(P_n P_{n-1} \cdots P_1 h)^2 d\mu &= \int_{X - \rho_n} \chi_A(P_n P_{n-1} \cdots P_1 h) d\mu_n \\ &= \int_{\Omega} \chi_A(x_n) P_{1n} h(x_n) dP, \end{aligned}$$

giving the second equality. This completes the proof of the lemma, hence of the theorem.

COROLLARY 1. *Let $P_n \in \mathcal{D}(X, X)$ be a positive operator ($n = 1, 2, \dots$). Then given $p: 1 < p < \infty$ and $g \in L^p(X, \mu)$, $\lim_{n \rightarrow \infty} P_{1n}^* P_{1n} g$ exists a.e. (μ).*

Proof. If, in addition, $g \in L^1$, the convergence follows by Theorem 2 (for g then satisfies $\int |g| \log^+ |g| d\mu < \infty$). Since $L^1 \cap L^p$ is dense in L^p , by a theorem of Banach [10, pp. 332, 333] it suffices to show $\mu\{x: \sup_{n \geq 0} |P_{1n}^* P_{1n} f(x)| = \infty\} = 0$ for each $f \in L^p$.

Given $f \in L^p$, let $A = \{x: |f(x)| > 1\}$, $B = \{x: |f(x)| \leq 1\}$. Then $f = f\chi_A + f\chi_B$ and $\|f\chi_B\|_{\infty} \leq 1$. Therefore $\sup_{n \geq 0} |P_{1n}^* P_{1n} f| \leq \sup_{n \geq 0} (P_{1n}^* P_{1n} |f\chi_A|) + 1$ a.e. (μ). $|f\chi_A|$ satisfies the hypotheses of Theorem 2, since it belongs to $L^1 \cap L^p$. From (ii) of Theorem 2 we infer $\sup_{n \geq 0} (P_{1n}^* P_{1n} |f\chi_A|) < \infty$ a.e. (μ), since h may be chosen to be > 0 everywhere. This completes the proof.

REMARK. The ratio theorem of Doob [9, Theorem 6.1], giving a pointwise limit, assumed only that the operators P_k were positive contractions in L^1 . We use the additional hypothesis of contractivity in L^{∞} both in proving pointwise convergence and in obtaining maximal inequalities: The uniform boundedness of $\{P_{1n} h\}_{n=0}^{\infty}$ is used just below (2.3), above Lemma 3, and below (3.5). The inequality $P_{1k} \leq 1$ is used to show $P_k(\phi^2) \geq (P_k \phi)^2$ just below (2.6). If the hypothesis $\|P_k\|_{\infty} \leq 1$ is dropped, a counterexample is provided by the following set of positive contractions on functions integrable over the unit interval with Lebesgue measure:

$$P_k f = \phi_k \int_0^1 f, \text{ where } \phi_k \geq 0, \int_0^1 \phi_k = 1, \text{ and } \int_0^1 \phi_k^2 > k.$$

Since $P_k^* f = \int_0^1 \phi_k f$, $P_k^* P_k f = \int_0^1 \phi_k^2 \int_0^1 f$. Therefore $|P_{1n}^* P_{1n} f| \geq n |\int_0^1 f|$.

3. Maximal inequalities for contraction operators.

THEOREM 3. *Let $T_n \in \mathcal{D}(X, X)$ ($n = 1, 2, \dots$). Then*

$$(3.1) \quad \int_X \left(\sup_{n \geq 0} |T_{1n}^* T_{1n} g| \right)^p d\mu \leq \left(\frac{p}{p-1} \right)^p \int_X |g|^p d\mu$$

if $g \in \bigcup_{1 < p < \infty} L^p(X, \mu)$ and

$$(3.2) \quad \int_X \left(\sup_{n \geq 0} |T_{1n}^* T_{1n} g| \right) h \, d\mu \leq \frac{e}{e-1} \int_X (h + |g| \log^+ |g|) \, d\mu < \infty$$

if $g \in L^1(X, \mu)$, $\int |g| \log^+ |g| \, d\mu < \infty$, and h is any function in $L^1(X, \mu)$ satisfying $1 \geq h \geq \min(|g|, 1)$.

Cf. Remark following Theorem 2.

Proof. We show first that it is no loss of generality to assume the T_k are positive operators. By a lemma of Dunford and Schwartz [10, p. 672], cf. [6], given $T \in \mathcal{D}$, there exists an operator $P \in \mathcal{D}$ satisfying

$$(3.3) \quad |T^n f| \leq P^n |f| \quad \text{a.e. for } n = 1, 2, \dots \text{ and } f \in L^1.$$

Now for any set $E \subset X$ of finite measure and any $g \in \bigcup_{1 \leq p \leq \infty} L^p$,

$$- \int |g| P^n(\chi_E) \, d\mu \leq \int g T^n(\chi_E) \, d\mu \leq \int |g| P^n(\chi_E) \, d\mu,$$

hence

$$- \int \chi_E P^{*n} |g| \, d\mu \leq \int \chi_E T^{*n}(g) \, d\mu \leq \int \chi_E P^{*n} |g| \, d\mu.$$

We conclude that $|T^{*n} g| \leq P^{*n} |g|$ a.e. for such g and that P^* is a positive operator in \mathcal{D} . From the symmetry of the argument it follows that (3.3) holds for $f \in \bigcup_{1 \leq p \leq \infty} L^p$ and that P is a positive operator in \mathcal{D} .

REMARK. If T is self-adjoint, so is P : by [10], $Pf = \sup_{|g| \leq f} |Tg|$ for $0 \leq f \in L^1 \cap L^\infty$, $g \in L^\infty$. Now for $|g| \leq f$, $P^* f \geq P^* |g| \geq |T^* g| = |Tg|$. Thus $(P^* - P)f \geq 0$ for $0 \leq f \in L^1 \cap L^\infty$, which, considering adjoints, implies

$$(P - P^*)f \geq 0$$

for such f . Therefore $P = P^*$ (cf. [1]).

Applying the above to each of $T_1^*, T_2^*, \dots, T_n^*, T_n, T_{n-1}, \dots$, and T_1 , there exist positive operators $P_k \in \mathcal{D}(X, X)$ for which

$$|T_{1n}^* T_{1n} g| \leq P_1^* |T_2^* T_3^* \dots T_n^* T_n T_{n-1} \dots T_1 g| \leq \dots \leq P_{1n}^* P_{1n} |g|$$

a.e., ($g \in \bigcup_{1 \leq p \leq \infty} L^p$). Consequently (3.2) holds by Theorem 2. We prove (3.1) by showing that for any fixed p : $1 < p < \infty$ and any $g \in L^p$,

$$(3.4) \quad \int_X \left(\sup_{n \geq 0} |P_{1n}^* P_{1n} g| \right)^p \, d\mu \leq \left(\frac{p}{p-1} \right)^p \int_X |g|^p \, d\mu.$$

Since $|P_{1n}^* P_{1n} g| \leq P_{1n}^* P_{1n} |g|$ we may and shall assume $g \geq 0$. Letting $A_k = \{x: g(x) \geq 1/k\}$ and $g_k = g \chi_{A_k}$ ($k = 1, 2, \dots$), g_k increases to g a.e. as $k \rightarrow \infty$. For each positive integer N ,

$$\begin{aligned} \sup_{0 \leq n < N} (P_{1n}^* P_{1n} g) &= \lim_{k \rightarrow \infty} \sup_{0 \leq n < N} (P_{1n}^* P_{1n} g_k) \leq \lim_{k \rightarrow \infty} \sup_{n \geq 0} (P_{1n}^* P_{1n} g_k) \\ &\leq \sup_{n \geq 0} (P_{1n}^* P_{1n} g), \end{aligned}$$

using the dominated convergence theorem for positive operators. As $N \rightarrow \infty$ the first term tends to the last, giving equality among all terms. Consequently if (3.4) holds for each g_k , by the Lebesgue monotone convergence theorem it holds for $g = \lim_{k \rightarrow \infty} g_k$.

For fixed k , we now prove (3.4) for g_k by arguing as in the proof of Theorem 2. First let the sequence $\{B(m)\}_{m=1}^\infty$ of measurable sets in X satisfy

$$A_k = B(1) \subset B(2) \subset \dots, \bigcup_{m=1}^\infty B(m) = X,$$

and $\mu(B(m)) < \infty$ for all m . Next set $h_m = \chi_{B(m)}$ ($m = 1, 2, \dots$) and $f = g_k$. Note $g_k = h_m f$, $f \in L^1 \cap L^p$, and $h_m \in L^1 \cap L^\infty$, for all m .

Fixing m , h_m and f are now used as were h and f in the proof of Theorem 2. We again suppress the “ $\#$ ” notation for the representations associated with the Stone spaces.

Since $0 \leq f(x_0) \in L^p(\Omega, P)$ and $\{S_{1n}f(x_n)\}_{n=0}^\infty$ is a decreasing martingale on the finite measure space (Ω, P) , it follows [8] that

$$(3.5) \quad \int_{\Omega} \left(\sup_{n \geq 0} S_{1n}f(x_n) \right)^p dP \leq \left(\frac{p}{p-1} \right)^p \int_{\Omega} f(x_0)^p dP.$$

Noting that $0 \leq (P_{1n}h_m)(x_n) \leq 1$ a.e. (dP), we have by Lemma 2:

$$\begin{aligned} \int_X \left(\sup_{n \geq 0} P_{1n}^* P_{1n} g_k \right)^p h_m d\mu &= \int_{\Omega} \left(\sup_{n \geq 0} (P_{1n}^* P_{1n} g_k)(x_0) \right)^p dP \\ &= \int_{\Omega} \left(\sup_{n \geq 0} (S_{1n}^* [P_{1n} h_m S_{1n} f])(x_0) \right)^p dP \\ &\leq \int_{\Omega} \left(\sup_{n \geq 0} (S_{1n}^* S_{1n} f)(x_0) \right)^p dP. \end{aligned}$$

By (1.3) the last integral equals

$$\int_{\Omega} \left(\sup_{n \geq 0} E\{S_{1n}f(x_n) | x_0\} \right)^p dP.$$

Using the fact that conditional expectation is a positive contraction on $L^p(\Omega, P)$ and (3.5), the above expression is less than or equal to

$$\begin{aligned} \int_{\Omega} \left(E \left\{ \sup_{n \geq 0} S_{1n}f(x_n) | x_0 \right\} \right)^p dP &\leq \int_{\Omega} \left(\sup_{n \geq 0} S_{1n}f(x_n) \right)^p dP \\ &\leq \left(\frac{p}{p-1} \right)^p \int_{\Omega} f(x_0)^p dP. \end{aligned}$$

From the definition of f , h_m , and P , this last term equals

$$\left(\frac{p}{p-1}\right)^p \int_X (g_k)^p h_m d\mu = \left(\frac{p}{p-1}\right)^p \int_X (g_k)^p d\mu.$$

Therefore

$$\int_X \left(\sup_{n \geq 0} P_{1n}^* P_{1n} g_k \right)^p h_m d\mu \leq \left(\frac{p}{p-1}\right)^p \int_X (g_k)^p d\mu$$

for each m . Letting $m \rightarrow \infty$, we obtain (3.4) for g_k , completing the proof.

REMARKS. (1) The above theorems raise the natural question as to the pointwise convergence of $\{T_{1n}^* T_{1n} g\}_{n=0}^\infty$, for $g \in \bigcup_{1 < p < \infty} L^p$. Since (3.1) implies

$$\mu \left\{ x: \sup_{n \geq 0} |T_{1n}^* T_{1n} g(x)| = \infty \right\} = 0$$

for any such g , it would suffice by the theorem of Banach used in Corollary 1 to show convergence on a dense subset of any of the L^p spaces ($1 < p < \infty$). This question remains unanswered.

(2) Stein [24] has proved the following maximal theorem: Let T be a self-adjoint operator in \mathscr{D} . Then

$$\int \left(\sup_{n \geq 0} |T^n g| \right)^p d\mu \leq B_p \int |g|^p d\mu \quad \left(g \in \bigcup_{1 < p < \infty} L^p \right),$$

where the constant B_p depends on p . This result is an obvious consequence of Theorem 3 above: noting $\sup_{n \geq 0} |T^n g| \leq \sup_{n \geq 0} |T^{2n} g| + \sup_{n \geq 0} |T^{2n+1} g|$, we obtain Stein's result with $B_p = 2(p/(p-1))^p$. Yet another proof of this maximal result, based on techniques of Rota [22] and Burkholder [4, p. 82], is contained in the thesis of Al-Hussaini [1].

(3) The principal open question relating to the problems considered in this paper is that of the pointwise convergence of $T^n g$ for $g \in L^1$, where T is a self-adjoint operator in \mathscr{D} satisfying $(Tf, f) \geq 0$ for all $f \in L^2$. (Positive definiteness is necessary: consider rotation through π radians on the unit circle.) Equivalent formulations of this problem are given by Burkholder [4], using a characterization he proves there of certain conditions under which one of the two major hypotheses of Banach's theorem will hold (in this case the hypothesis that $\mu\{x: (\sup_{n \geq 0} |T^n g(x)|) = \infty\} = 0$). (Al-Hussaini [1] has given an example of two distinct such operators T_1, T_2 for which $T_2^n T_1^n f$ diverges a.e. for some $f \in L^1$.)

4. Strong convergence and identification of the limit operator. The preceding results imply the strong convergence of the positive operators $P_{1n}^* P_{1n} \in \mathscr{D}(X, X)$

considered as bounded operators on $L^p(X)$ into itself, for any fixed $p: 1 < p < \infty$. The same is true for $p = 1$ in case $\mu(X) < \infty$. A more direct proof of a more general result is given here.

Denoting by 0 and I the zero and the identity operators, we use the following known

THEOREM (VIGIER [18]). *If $\{A_n\}_{n=1}^\infty$ is a sequence of positive definite operators on a Hilbert space such that $I \geq A_1 \geq A_2 \geq \cdots \geq 0$, then $\{A_n\}$ converges strongly to a self-adjoint positive definite contraction operator A .*

On the basis of this, we state and prove

PROPOSITION 1. *Let $T_n \in \mathcal{D}(X, X)$, $(n = 1, 2, \dots)$. There exists a self-adjoint operator $\Pi_\infty \in \mathcal{D}(X, X)$ to which $T_{1n}^* T_{1n}$ converges in each of the strong topologies on bounded operators on $L^p(X)$, $1 < p < \infty$.*

Given any $g \in L^1(X)$ and any $h \in L^p(X) \cap L^\infty(X)$ for some $p \in [1, \infty)$, $T_{1n}^ T_{1n} g \rightarrow \Pi_\infty g$ in the $L^1(h d\mu)$ norm: $\|(T_{1n}^* T_{1n} g - \Pi_\infty g)h\|_1 \rightarrow 0$ as $n \rightarrow \infty$.*

REMARKS. If $\mu(X) < \infty$, $T_{1n}^* T_{1n} \rightarrow \Pi_\infty$ in the strong topology on bounded operators on $L^1(X)$: set $h = 1$. Π_∞ is clearly positive definite on L^2 .

Proof. Throughout this proof we denote $T_{1n}^* T_{1n}$ by Π_n . Using only the fact of contractivity of each T_n in the case $p = 2$, the sequence $\{\Pi_n\}_{n=1}^\infty$ clearly satisfies the hypotheses of Vigier's theorem. Consequently $\{\Pi_n\}$ converges in the strong topology on bounded operators on $L^2(X)$ to a self-adjoint positive definite contraction Π_∞ on $L^2(X)$. We now show Π_∞ has a unique extension to a self-adjoint operator in $\mathcal{D}(X, X)$, to which $\{\Pi_n\}$ converges strongly in L^p ($1 < p < \infty$).

(i) For each $p \in [2, \infty)$ and each $g \in L^p$, $\{\Pi_n g\}$ converges in the L^p norm: For each $p \in [2, \infty)$ we show Cauchy convergence in L^p of $\{\Pi_n f\}$ for $f \in L^2 \cap L^\infty$. Since $L^2 \cap L^\infty$ is dense in L^p , Cauchy convergence for $g \in L^p$ follows by the triangle inequality and contractivity of each Π_k in L^p . Given $f \in L^2 \cap L^\infty$, we may assume $\|f\|_\infty < \frac{1}{2}$. Then

$$\begin{aligned} \|(\Pi_n - \Pi_m)f\|_p^p &= \int_X |(\Pi_n - \Pi_m)f|^p d\mu \leq \int_X |(\Pi_n - \Pi_m)f|^2 d\mu \\ &= \|(\Pi_n - \Pi_m)f\|_2^2 \rightarrow 0 \end{aligned}$$

as $m, n \rightarrow \infty$, where we have used the L^∞ contractivity of each Π_k .

(ii) For each $p \in (1, 2)$ and each $g \in L^p$, $\{\Pi_n g\}$ converges in the L^p norm: As in (i) it suffices to prove L^p norm convergence on a dense set, in this case $L^1 \cap L^\infty$. For $f \in L^1 \cap L^\infty$, $(\Pi_n - \Pi_m)f \in L^1 \cap L^\infty$ for each $m, n = 1, 2, \dots$. Let $p = 1 + \varepsilon$ ($0 < \varepsilon < 1$). Setting $a = \varepsilon/2$, $b = 1 + \varepsilon/2$, $r = 4/\varepsilon$, and $s = 4/(4 - \varepsilon)$, note $a + b = p$, $ar = 2$, and $bs \geq 1$. In particular, $|(\Pi_n - \Pi_m)f|^a \in L^r$, $|(\Pi_n - \Pi_m)f|^b \in L^s$, and $1/r + 1/s = 1$. Therefore

$$\begin{aligned}
\int |(\Pi_n - \Pi_m)f|^p d\mu &= \int |(\Pi_n - \Pi_m)f|^{a+b} d\mu \\
&\leq \| |\Pi_n f - \Pi_m f|^a \|_r \cdot \| |\Pi_n f - \Pi_m f|^b \|_s \\
&\leq \| |\Pi_n f - \Pi_m f|^a \|_r \cdot \| 2f \|_{bs}^b < \infty.
\end{aligned}$$

Since $ar = 2$, $\| |\Pi_n f - \Pi_m f|^a \|_r = \| (\Pi_n - \Pi_m)f \|_{ar}^a \rightarrow 0$ as $m, n \rightarrow \infty$, so $\{\Pi_n f\}$ is Cauchy convergent in L^p .

(iii) For any $g \in L^1 \cap L^\infty$, $\{\Pi_n g\}$ converges to $\Pi_\infty g$ in L^p , for each $p: 1 < p < \infty$: For fixed $p \in (1, \infty)$, denote the L^p norm limit of $\{\Pi_n g\}$ by $S\Pi_\infty g$. We show $S\Pi_\infty g = \Pi_\infty g$. For each set $E \subset X$ of finite measure,

$$\begin{aligned}
&\| (\Pi_\infty g - S\Pi_\infty g)\chi_E \|_1 \\
&\leq \| ((\Pi_\infty - \Pi_n)g)\chi_E \|_1 + \| (\Pi_n g - S\Pi_\infty g)\chi_E \|_1 \\
&\leq \| (\Pi_\infty - \Pi_n)g \|_2 \cdot (\mu(E))^{1/2} + \| (\Pi_n g - S\Pi_\infty g) \|_p \cdot (\mu(E))^{1/q} \rightarrow 0
\end{aligned}$$

as $n \rightarrow \infty$ by Vigier's result and norm convergence in L^p , respectively ($1/p + 1/q = 1$).

(iv) $\| \Pi_\infty g \|_p \leq \| g \|_p$ for any $g \in L^1 \cap L^\infty$, any $p \in (1, \infty)$: Letting $g \in L^1 \cap L^\infty$, note

$$\| \Pi_\infty g \|_p - \| \Pi_n g \|_p \leq \| (\Pi_\infty - \Pi_n)g \|_p \rightarrow 0 \quad \text{as } n \rightarrow \infty, \quad \text{by (iii).}$$

Since $\| \Pi_n g \|_p \leq \| g \|_p$ for any n , (iv) is true as claimed.

Π_∞ is thus a densely defined contraction on each L^p ($1 < p < \infty$) and so has a unique extension to an operator in $\mathcal{D}(X, X)$, which we also denote Π_∞ . From (iii) it is evident that $\Pi_n \rightarrow \Pi_\infty$ strongly with respect to each of the topologies induced by L^p ($1 < p < \infty$).

To show Π_∞ is self-adjoint as an operator in $\mathcal{D}(X, X)$, let $f \in L^p$ ($1 < p < \infty$) and $g \in L^q = (L^p)^*$. Then

$$\begin{aligned}
&|(\Pi_\infty f, g) - (f, \Pi_\infty g)| \\
&\leq |((\Pi_\infty - \Pi_n)f, g)| + |(\Pi_n f, g) - (f, \Pi_n g)| + |(f, (\Pi_n - \Pi_\infty)g)| \\
&\leq \| (\Pi_\infty - \Pi_n)f \|_p \| g \|_q + 0 + \| f \|_p \| (\Pi_n - \Pi_\infty)g \|_q \rightarrow 0
\end{aligned}$$

as $n \rightarrow \infty$. (We have denoted the action of the functional $\phi \in L^q$ on $\Psi \in L^p$ by (ϕ, Ψ) .)

Finally, given h as in the statement of Proposition 1, we note $h \in L^{2p}$. Hence $\| (\Pi_n g - \Pi_\infty g)h \|_1 \rightarrow 0$ for any $g \in L^1 \cap L^q$, where $1/2p + 1/q = 1$, by the L^q norm convergence of $\Pi_n g$ to $\Pi_\infty g$. Since Π_n and Π_∞ are L^1 contractions and $L^1 \cap L^q$ is dense in L^1 , $\| (\Pi_n g - \Pi_\infty g)h \|_1 \rightarrow 0$ for any $g \in L^1$, by a standard approximation argument.

COROLLARY 2. Let $P_n \in \mathcal{D}(X, X)$ be a positive operator, ($n = 1, 2, \dots$). Let $g \in \bigcup_{1 \leq p < \infty} L^p$ satisfy $\int |g| \log^+ |g| d\mu < \infty$. Then $\lim_{n \rightarrow \infty} P_{1n}^* P_{1n} g = \Pi_\infty g$ a.e. (μ),

where Π_∞ is a self-adjoint positive operator in $\mathcal{D}(X, X)$, and positive definite as an operator on L^2 .

Sketch of proof. Π_∞ is clearly the strong limit operator provided by Proposition 1, namely the extension to all L^p of the limit operator on L^2 given by Vigier's theorem. The corollary follows readily, using Theorems 2,3, Corollary 1, Proposition 1, and Lebesgue's Dominated Convergence Theorem.

REMARK. Since Π_∞ is the operator provided in Proposition 1, $P_{1n}^* P_{1n} g \rightarrow \Pi_\infty g$ in the appropriate topologies as noted there.

EXAMPLES. (1) Iterates of conditional expectation operators. Burkholder and Chow [5] have given the following explicit identification of the limit operator in the case of powers of products of conditional expectation operators. Let E_i denote $E\{\cdot | \mathcal{F}_i\}$, $\mathcal{F}_i \subset \Sigma$, and $\overline{\mathcal{F}_i}$ denote the completion [16, p. 90] of \mathcal{F}_i ($1 \leq i \leq k$). Then for $g \in L^2(X, \Sigma, \mu)$,

$$\lim_{n \rightarrow \infty} (E_1 E_2 \cdots E_k E_{k-1} \cdots E_1)^n g = E\{g | \overline{\mathcal{F}_1} \cap \overline{\mathcal{F}_2} \cap \cdots \cap \overline{\mathcal{F}_k}\}$$

a.e. (μ) and boundedly in $L^2(X, \Sigma, \mu)$. (Actually, convergence for a somewhat more general sequence, containing that given here, is shown (see [5, §3]).)

(2) If each P_k is doubly stochastic, Π_∞ need not be. Consider a decreasing martingale $\{f_n, \mathcal{F}_n\}_{n=0}^\infty$ on a totally σ -finite positive measure space (X, Σ, μ) , such that $\bigcap_{n=0}^\infty \mathcal{F}_n = \{\phi, X\}$. Letting $P_n = E\{\cdot | \mathcal{F}_n\}$, it follows that Π_∞ is the zero operator. Thus strong limits in the L^p topology ($1 < p < \infty$) of doubly stochastic operators need not be doubly stochastic. It is easy to see that this situation can occur only on spaces of infinite measure.

(3) Π_∞ need not be idempotent, even if each P_k is a doubly stochastic operator on functions integrable over the unit interval with Lebesgue measure. To see this let

$$S_1 f(x) = f(2x) \pmod{1} \quad \text{and} \\ S_2 f(x) = \begin{cases} 2 \left(\int_0^{1/2} f \right) & \text{on } \left[0, \frac{1}{2}\right), \\ f(x) & \text{on } \left[\frac{1}{2}, 1\right]. \end{cases}$$

S_2 is just a conditional expectation operator. S_1 has for adjoint (and left inverse) the operator defined by

$$S_1^* f(x) = \frac{1}{2} f(x/2) + \frac{1}{2} f((x+1)/2) \pmod{1}.$$

S_1 and S_2 are clearly doubly stochastic. Letting $S_n = I$ ($n \geq 3$),

$$\Pi_\infty = S_1^* S_2^* S_2 S_1 = S_1^* S_2 S_1.$$

If $f(x) = x$,

$$\Pi_{\infty} f(x) = 1/4 + x/2,$$

$$\Pi_{\infty}^2 f(x) = 1/4 + 1/2(1/4 + x/2) = 3/8 + x/4.$$

Hence $\Pi_{\infty}^2 \neq \Pi_{\infty}$. The operators S_1, S_1^* were brought to our attention by J. Ryff in connection with the study of extremal doubly stochastic operators.

REMARK. The results of §§2, 3, and the present section are valid in an arbitrary positive measure space, since only countably many operations on functions vanishing off a set of σ -finite measure are involved.

5. Continuous parameter theorems. For each real $t \geq s \geq 0$ let $T(t, s)$ be an operator in $\mathcal{D}(X, X)$ which satisfies

$$(5.1) \quad T(t, s) = T(t, r) T(r, s) \quad (t \geq r \geq s), \quad T(t, t) = I.$$

Note that $T^*(t, s) = T^*(r, s)T^*(t, r)$.

These operators, which form an inhomogeneous semigroup, arise in the theory of nonstationary Markov processes as positive operators [8, Chapter VI]. For this reason (5.1) is often called the Chapman-Kolmogorov equation. If $T(t, s) = T(t - s, 0)$ for all $t \geq s \geq 0$, the operators form a semigroup, the analog of a stationary Markov process in the case of positive operators. Such semigroups arise in a variety of other situations, notably in differential equation theory (cf. [24]).

We consider here the limiting and maximal behavior of $T^*(t, 0)T(t, 0)$. By the theory of separable processes [8, Chapter II], such behavior is essentially determined by that of the discrete case considered above. First, however, we state a result on operator convergence, for which separability theory is superfluous. The proof of this result, which depends on Vigier's theorem and is similar to that of Proposition 1, is omitted. (Vigier's theorem applies to bounded increasing, as well as decreasing, sequences.)

PROPOSITION 2. *Let $r > 0$ be fixed. Then there exists a self-adjoint operator $\Pi_{\infty}(r^-) \in \mathcal{D}(X, X)$ to which $T^*(t, 0)T(t, 0)$ converges in the strong topology on bounded operators on L^p , for each $p: 1 < p < \infty$, as $t \rightarrow r^-$.*

For any $g \in L^1$ and any $h \in L^p \cap L^{\infty}$ for some $p: 1 \leq p < \infty$, $T^(t, 0)T(t, 0)g$ converges in the $L^1(h d\mu)$ norm to $\Pi_{\infty}(r^-)g$ as $t \rightarrow r^-$.*

Similarly, there is an operator $\Pi_{\infty}(r^+)$ to which $T^(t, 0)T(t, 0)$ converges in each of the above topologies as $t \rightarrow r^+$. In this case r may equal 0.*

REMARKS. (1) If $\mu(X) < \infty$, strong convergence in L^1 holds: set $h = 1$. (2) Letting $T(t, 0) = E\{\cdot | \mathcal{F}_t\}$, with $\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2}$ if $t_1 \geq t_2$, one sees that continuous parameter closed martingales [16] can be described in terms of $T^*(t, 0)T(t, 0)$. (Hence $\Pi_{\infty}(r^+)$ need not equal $\Pi_{\infty}(r^-)$.)

The following proposition is the continuous parameter version of the main results of the preceding sections. Since the proof is a relatively direct, though tedious, consequence of the separability theory of [8, Chapter II], it is omitted.

PROPOSITION 3. *Let $g \in \bigcup_{1 \leq p < \infty} L^p$ satisfy $\int |g| \log^+ |g| d\mu < \infty$. Then there exists a separable process $\{(T^*(t, 0)T(t, 0)g)^\sim\}_{t \in [0, \infty)}$ which differs from $\{T^*(t, 0)T(t, 0)g\}_{t \in [0, \infty)}$ at most on a set of μ -measure zero in each t , such that*

$$(i) \quad \int_X \left(\sup_{t \in [0, \infty)} |(T^*(t, 0)T(t, 0)g)^\sim| \right)^p d\mu \leq \left(\frac{p}{p-1} \right)^p \int_X |g|^p d\mu$$

if $g \in L^p$ ($1 < p < \infty$),

$$(ii) \quad \int_X \left(\sup_{t \in [0, \infty)} |(T^*(t, 0)T(t, 0)g)^\sim| \right) h d\mu \leq \frac{e}{e-1} \int_X (h + |g| \log^+ |g|) d\mu$$

if $g \in L^1$, $\int |g| \log^+ |g| d\mu < \infty$, and $h \in L^1$ satisfies $1 \geq h \geq \min(|g|, 1)$. (In case $\mu(X) < \infty$, note h may be chosen $\equiv 1$.)

(iii) If $r \in [0, \infty)$ and if $T(t, s)$ are in addition positive operators, then

$$\lim_{t \rightarrow r^-} (T^*(t, 0)T(t, 0)g)^\sim = \Pi_\infty(r^-)g \quad \text{a.e.,}$$

where $\Pi_\infty(r^-)$ is the strong limit operator of Proposition 2. Moreover, if any other process $\{(T^(t, 0)T(t, 0)g)^\sim\}_{t \in [0, \infty)}$, differing from $\{T^*(t, 0)T(t, 0)g\}_{t \in [0, \infty)}$ at most on a set of μ -measure zero in each t , converges a.e. on a set of positive measure, the limit must agree with $\Pi_\infty(r^-)g$.*

REMARKS. (1) For a one parameter semigroup $\{T^t, t \geq 0\}$ of self-adjoint operators in \mathscr{D} , possessing strong continuity at $t=0$ in the L^2 norm, Stein [24] has proved maximal and pointwise theorems for $\{T^t f\}_{0 < t < \infty}$, ($f \in \bigcup_{1 < p < \infty} L^p$). ($T^t f$ was redefined on a set of μ -measure zero for each t in such a manner that for almost every x , $T^t f(x)$ is a continuous function of t .) By (i) of Proposition 3, the maximal theorem is a corollary to Theorem 3 above.

(2) The results of Proposition 3 are possible because of the "inhomogeneous" character of our discrete limit theorems: we have at our disposal theorems about products $T_{1n}^* T_{1n}$ of distinct operators T_1, T_2, \dots , and it is just such products that arise as the result of applying separability theory.

APPENDIX.

1. Representation theory. The following is a sketch of the σ -finite version of a representation theory found in Doob [9]; details come from [9], Maharam [17], and Halmos [11], [12], and are omitted.

Let (X, Σ, μ) be a totally σ -finite positive measure space. The Boolean σ -algebra of measurable sets mod null sets in (X, Σ, μ) is isomorphic to the algebra of clopen (simultaneously closed and open) sets in a totally disconnected compact Hausdorff

space $X^\#$. The σ -field $\Sigma^\#$ generated by the clopen sets is the σ -field of Baire sets and the clopen sets are a base for the topology on $X^\#$. Regarding sets of the first category in $X^\#$ as null sets, each bounded Baire function differs from a continuous function at most on a null set, i.e. $L^\infty(X^\#, \Sigma^\#)$ is naturally identified with $C(X^\#)$. Moreover, equivalence classes mod null functions of $L^\infty(X, \Sigma)$ can be put in a one-to-one correspondence with $C(X^\#)$ in a natural fashion. This correspondence is a Banach space isometry, is multiplicative, order preserving, and preserves pointwise a.e. sequential convergence.

μ on Σ induces naturally a totally σ -finite positive measure $\mu^\#$ on $\Sigma^\#$, which vanishes on sets of the first category. The correspondence between $L^\infty(X, \Sigma, \mu)$ and $L^\infty(X^\#, \Sigma^\#, \mu^\#)$ can now be extended to an order preserving Banach space isometry of each of the corresponding L^p spaces ($1 \leq p < \infty$) (actually of the equivalence classes mod null functions). Again a.e. convergence of sequences is preserved.

For $T \in \mathcal{D}(X_1, X_2)$, there is a corresponding $T^\# \in \mathcal{D}(X_1^\#, X_2^\#)$, and $(T^*)^\# = (T^\#)^*$. If $W \in \mathcal{D}(X_0, X_1)$, $(TW)^\# = T^\# W^\#$.

As shown in Doob [9], the image of a doubly stochastic operator $S \in \mathcal{D}(X_1, X_2)$ can be regarded as a kernel operator in the sense that

$$S^\#(f^\#)(x_2) = \int_{X_1^\#} f^\#(x_1) Q^\#(x_2, dx_1) \quad \text{a.e. } (\mu_2),$$

where $Q^\#$ is a real valued function on $X_2^\# \times \Sigma_1^\#$ which for each fixed $x_2 \in X_2^\#$ is a probability measure on $(X_1^\#, \Sigma_1^\#)$, and for each fixed $A^\# \in \Sigma_1^\#$ is a (continuous) nonnegative bounded measurable function on $(X_2^\#, \Sigma_2^\#)$. (Given a clopen set $A^\# \subset X_1^\#$, let $A \in \Sigma_1$ have image $A^\#$. $Q^\#(\cdot, A^\#)$ is defined to be the continuous image of $S\chi_A$ under the isomorphism of $L^\infty(X_2)$ and $C(X_2^\#)$ noted above.)

Note that the kernel representation for doubly stochastic operators holds for positive operators $P \in \mathcal{D}(X_1, X_2)$, the only change being that $Q(x_2, \cdot)$ becomes a positive measure bounded by 1.

2. Dominated convergence theorem for positive operators. Let $f_n \in L^p(X_1)$, $\lim_{n \rightarrow \infty} f_n = f$ a.e. (μ_1) , $\sup_{n \geq 1} |f_n| \in L^p(X_1)$, for some p : $1 \leq p \leq \infty$. Then if $P \in \mathcal{D}(X_1, X_2)$ is a positive operator, $\lim_{n \rightarrow \infty} P f_n = P f$ a.e. (μ_2) .

For a direct proof of a more general result, see [23].

In the case of probability measure spaces, Blackwell and Dubins [2] and Al-Hussaini [1] have given converses to the above proposition.

BIBLIOGRAPHY

1. A. N. Al-Hussaini, *Some almost everywhere convergence theorems for positive definite operators*, Ph. D. thesis, University of Illinois, Urbana, Ill., 1964.
2. D. Blackwell and L. E. Dubins, *A converse to the dominated convergence theorem*, Illinois J. Math. 7 (1963), 508-515.

3. D. L. Burkholder, *Successive conditional expectations of an integrable function*, Ann. Math. Statist. **33** (1962), 887–893.
4. ———, *Maximal inequalities as necessary conditions for almost everywhere convergence*, Z. Wahrscheinlichkeitstheorie und Veru. Gebiete **3** (1964), 75–88.
5. D. L. Burkholder and Y. S. Chow, *Iterates of conditional expectation operators*, Proc. Amer. Math. Soc. **12** (1961), 490–495.
6. R. V. Chacón and U. Krengel, *Linear modulus of a linear operator*, Proc. Amer. Math. Soc. **15** (1964), 553–560.
7. Y. S. Chow, *Martingales in a σ -finite measure space indexed by directed sets*, Trans. Amer. Math. Soc. **97** (1960), 254–285.
8. J. L. Doob, *Stochastic processes*, Wiley, New York, 1952.
9. ———, *A ratio operator limit theorem*, Z. Wahrscheinlichkeitstheorie und Veru. Gebiete **1** (1963), 288–294.
10. N. Dunford and J. T. Schwartz, *Linear operators*, Part I, Interscience, New York, 1958.
11. P. R. Halmos, *Measure theory*, Van Nostrand, New York, 1950.
12. ———, *Lectures on Boolean algebras*, Van Nostrand, New York, 1963.
13. A. Ionescu Tulcea and C. Ionescu Tulcea, *Abstract ergodic theorems*, Trans. Amer. Math. Soc. **107** (1963), 107–124.
14. C. Ionescu Tulcea, *Mesures dans les espaces produits*, Atti Acad. Naz. Lincei. Rend. Cl. Sci. Fis. Mat. Nat. **7** (1950), 208–211.
15. M. Jerison, *Martingale formulation of ergodic theorems*, Proc. Amer. Math. Soc. **10** (1959), 531–539.
16. M. Loève, *Probability theory*, 2nd ed., Van Nostrand, New York, 1960.
17. D. Maharam, *The representation of abstract measure functions*, Trans. Amer. Math. Soc. **65** (1949), 279–330.
18. F. Riesz and B. Sz.-Nagy, *Functional analysis*, Ungar, New York, 1955.
19. G.-C. Rota, *On the representation of averaging operators*, Rend. Sem. Mat. Univ. Padova **30** (1960), 52–64.
20. ———, *On the eigenvalues of positive operators*, Bull. Amer. Math. Soc. **67** (1961), 556–558.
21. ———, *Symposium on ergodic theory*, Tulane Univ., New Orleans, La., 1961.
22. ———, *An “alternierende Verfahren” for general positive operators*, Bull. Amer. Math. Soc. **68** (1962), 95–102.
23. N. Starr, *Operators preserving bounded convergence* (to be submitted).
24. E. M. Stein, *On the maximal ergodic theorem*, Proc. Nat. Acad. Sci. U.S.A. **47** (1961), 1894–1897.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY,
CAMBRIDGE, MASSACHUSETTS