PACIFIC JOURNAL OF MATHEMATICS
Vol. 26, No. 3, 1968

UNIFORM APPROXIMATION OF DOUBLY STOCHASTIC OPERATORS

CHOO-WHAN KIM

Let $L_\infty = L_\infty[0, 1]$ be the real Banach space of essentially bounded Lebesgue measurable functions on the unit interval $I = [0, 1]$ with the essential sup-norm. A positive linear operator $T: L_\infty \to L_\infty$ is called doubly stochastic if (1) $T1 = 1$, (2) $\int_I Tfd\ell = \int_I fd\ell$ where ℓ denotes Lebesgue measure on the unit interval. We denote the set of doubly stochastic operators by $\mathscr{D}$. It follows that $\|T\|_\infty = 1$ for each $T \in \mathscr{D}$. Let Φ be the subset of $\mathscr{D}$ induced by measure preserving maps on the unit interval and Φ_1 the subset of Φ induced by invertible measure preserving maps. For each $T_\varphi \in \Phi$ we have $T_\varphi f(x) = f(\varphi)(x), f \in L_\infty$. A regular probability measure μ on the unit square $I \times I$ is called doubly stochastic if $\mu(A \times I) = \mu(I \times A) = \ell(A)$ for each $A \in \mathscr{B}(I)$, the Borel field of the unit interval I. Then there is a one-to-one correspondence between doubly stochastic operators and doubly stochastic measures. If we denote such a correspondence by $T \leftrightarrow \mu_T$, then

$$\int_I g(x)Tf(x)\ell(dx) = \int_{I\times I} g(x)f(y)\mu_T(d(x, y)), \quad f, g \in L_\infty .$$

Thus we will identify each $T \in \mathscr{D}$ with the corresponding doubly stochastic measure $\mu_T \in \Phi$, the doubly stochastic measure $\mu_\varphi = \mu_{T_\varphi}$ is singular with respect to Lebesgue measure ℓ^2 on the unit square. Let L be the set of all $T \in \mathscr{D}$ such that μ_T is absolutely continuous with respect to ℓ^2, i.e., $\mu_T \ll \ell^2$. The metric

$$\rho(T, R) = \sup\left\{\int_I | Tf - Rf | \, d\ell : \|f\|_\infty \leqq 1\right\}, \quad T, R \in \mathscr{D}$$

defines a topology on $\mathscr{D}$ which will be called the uniform topology. The purpose of this paper is to show that each $T \in L$ can be approximated by a convex combination of operators from Φ in the uniform topology, called the uniform approximation theorem.

It is known [3] that every $T \in \mathscr{D}$ arises from a Markov transition function $P(\cdot,\cdot)$ as $Tf(x) = \int_I P(x, dy)f(y), f \in L_\infty$ and $\int_I P(x, A)\ell(dx) = \ell(A)$ for $A \in \mathscr{B}(I)$. For each $T \in \mathscr{D}$ there is a unique $T^* \in \mathscr{D}$ such that

$$\int_I g(x)Tf(x)\ell(dx) = \int_I f(x)T^*g(x)\ell(dx) , \qquad f, g \in L_\infty .$$

T^* is called the adjoint of T. If $T \leftrightarrow \mu_T$, then $T^* \leftrightarrow \mu_{T^*}$ where $\mu_{T^*}(A \times B) = \mu_T(B \times A)$, $A, B \in \mathscr{B}(I)$. If $T \in L$, then $\mu_T \ll \ell^2$, so by the Radon-Nikodym theorem, there is an obvious one-to-one correspondence between the elements of L and those $k \in L_1(I \times I)$, the real Banach space of Lebesgue integrable functions on $I \times I$, such that $k(x, y) \geqq 0$ and $\int_I k(x, y)\ell(dy) = 1$ a.e. (ℓ) and $\int_I k(x, y)\ell(dx) = 1$ a.e. (ℓ). Henceforth we often identify each $T \in L$ with the corresponding element $k \in L_1(I \times I)$, which is called the kernel for T. Thus for each $T \in L$ we have $Tf(x) = \int_I k(x, y)f(y)\ell(dy)$, $f \in L_\infty$. Hereafter D_i^n will denote a dyadic interval of the form $[(i-1)/2^n, i/2^n)$, $1 \leqq i < 2^n$ and $D_{2^n}^n = [1 - 1/2^n, 1]$. 1_A denotes the characteristic function of the set A. The operators U_n defined by the kernels of the form $u_n(x, y) = 2^n \sum_{i=1}^{2^n} 1_{D_i^n \times D_i^n}(x, y)$, $n = 1, 2, \cdots$, are often called the conditional expectation operators. Then $U_n \in L$ and $U_n f = 2^n \sum_{i=1}^{2^n} \left(\int_{D_i^n} f d\ell\right) 1_{D_i^n}$, $f \in L_\infty$.

By the weak (strong) topology in $\mathscr{D}$ we mean the weak (strong) operator topology in $\mathscr{D}$ [2, 8]. We note that the relative uniform topology on Φ_1 coincides with the uniform topology on Φ_1 introduced by Halmos [5]. We will also define the norm topology on $\mathscr{D}$ by the metric

$$\delta(T, R) = \sup \{ |\mu_T(A) - \mu_R(A)| : A \in \mathscr{B}(I \times I) \}, \quad T, R \in \mathscr{D},$$

where $\mathscr{B}(I \times I)$ denotes the Borel field of $I \times I$. We note that the norm topology on $\mathscr{D}$ is the usual norm topology on signed measures [4] restricted to the doubly stochastic measures.

Let $L_p = L_p[0, 1]$, $1 \leqq p < \infty$, be the real Banach space of p-th power Lebesgue integrable functions on the unit interval with the usual norm. We denote the unit ball of L_p by B_p. For each p, $1 \leqq p < \infty$, L_p contains L_∞ as a dense subset in the L_p-norm and so we can extend each $T \in \mathscr{D}$ from L_∞ to L_p. From Jensen's inequality [3], it follows that $\| T \|_p = 1$, $1 \leqq p < \infty$. This extension will be assumed hereafter. For notational convenience we write $\int_I f d\ell$ as $\int f$. We work with real functions on I only.

Certain preliminary result regarding the sets Φ and L are given in § 1. The main results of this paper are stated in § 2. We show the uniform approximation theorems for L and a sharper form of the strong approximation theorem [2, 8] by means of a concrete approach. A simpler proof is given to the weak approximation theorem [2].

1. Preliminaries. The set Φ plays an important role in the study of $\mathscr{D}$. A characterization of Φ is known [2]. We also add the following:

LEMMA 1.1. *Let $e: I \to I$ be such that $e(x) = x$. For each $T \in \mathscr{D}$,*

$$T \in \Phi \Leftrightarrow \int (Te)^2 = \int e^2 .$$

Proof. The necessity of the condition is obvious. It remains to show the sufficiency of the condition. We recall that for each $T \in \mathscr{D}$ we have

$$Tf(x) = \int_I f(y)P(x, dy) , \quad f \in L_\infty .$$

Suppose that for some $x \in I$, $(Te)^2(x) = Te^2(x)$, i.e.,

$$\left(\int_I e(y)P(x, dy)\right)^2 = \int_I e^2(y)P(x, dy) .$$

It follows that the function e is constant a.e. $P(x, \cdot)$, and so $P(x, \cdot) = \varepsilon_{\varphi(x)}(\cdot)$, a probability measure concentrated at some point $\varphi(x)$ of I.

If $\int (Te)^2 = \int e^2$, then by Jensen's inequality [3]: $(Te)^2 \leqq Te^2$, we have $(Te)^2(x) = Te^2(x)$ a.e. $(\measuredangle)$. Thus from the above discussion we have $P(x, \cdot) = \varepsilon_{\varphi(x)}(\cdot)$ a.e., $(\measuredangle)$. φ is defined on I a.e. $(\measuredangle)$, but it can be defined everywhere on I in the usual manner. Thus we have

$$Tf(x) = f(\varphi(x)) , \quad f \in L_\infty .$$

Since $T \in \mathscr{D}$, it follows that φ is measure preserving, and so $T = T_\varphi \in \Phi$.

It is known that $\mathscr{D}$ is metrizable and compact in the weak topology [2]. The following argument also leads us to the assertion. By identifying each element in $\mathscr{D}$ with the corresponding doubly stochastic measure, we can topologize $\mathscr{D}$ by the subspace topology of the weak * topology on $C(I \times I)^*$. By $C(I \times I)$ we mean the space of real continuous functions on $I \times I$, and $C(I \times I)^*$ is the dual space of $C(I \times I)$. We call the topology on $\mathscr{D}$ so defined the weak * topology. By the usual argument we can show that $\mathscr{D}$ is metrizable and compact in the weak * topology. It is also straightforward to prove the equivalence of the weak and the weak * topologies in $\mathscr{D}$. It is interesting to note that the metric topology defined by

$$w(T, R) = \sup \{| \mu_T(A \times B) - \mu_R(A \times B) |: A, B \in \mathscr{J}(I)\} ,$$

where $\mathscr{J}(I)$ denotes the intervals of I, is equivalent to the weak topology. Incidentally the metrics

$$s(T, R) = \sup \{| \mu_T(A \times B) - \mu_R(A \times B) |: A \in \mathscr{B}(I), B \in \mathscr{J}(I)\}$$

and $u(T, R) = \sup\{|\mu_T(A \times B) - \mu_R(A \times B)|: A, B \in \mathscr{B}(I)\}$ define topologies in $\mathscr{D}$ which are equivalent to the strong and the uniform topologies respectively. We write $\mathscr{D}^w(\mathscr{D}^s, \mathscr{D}^u)$ to denote $\mathscr{D}$ endowed with the weak (strong, uniform) topology. The map $T \to T^*$ defined on $\mathscr{D}^w(\mathscr{D}^u)$ into itself is continuous. But this is not the case in $\mathscr{D}^s$, as we will show. First we state without proof a well known fact.

LEMMA 1.2. *The following are equivalent.*

(i) $T_n \xrightarrow{w} T$ and $\int (T_n g)^2 \to \int (Tg)^2$ *for each* $g \in B_\infty$.

(ii) $T_n \xrightarrow{s} T$.

PROPOSITION 1.1. The map $T \to T^*$ defined on $\mathscr{D}^s$ into itself is not continuous on $\Phi - \Phi_1$ but is continuous on Φ^*.

Proof. Given $T \in \Phi - \Phi_1$, by $\Phi = \Phi_1^{-s}$: the closure of Φ_1 in the strong topology [2], there is a sequence $\{T_n\}$ in Φ_1 such that $T_n \xrightarrow{s} T$. Since $\Phi \cap \Phi^* = \Phi_1$, we have $T^* \notin \Phi$, and so by Lemma 1.1,

$$\int (T_n^* e)^2 = \int e^2 > \int (T^* e)^2, \qquad n = 1, 2, \cdots .$$

It follows from Lemma 1.2 that $T_n^* \nrightarrow T^*$ in the strong topology.

Suppose that $T \in \Phi^*$ and $T_n \xrightarrow{s} T$. Then $T_n^* \xrightarrow{w} T^* \in \Phi$ and for each $g \in B_\infty$ as $n \to \infty$,

$$\int g^2 \geqq \| T^* g \|_2 \| T_n^* g \|_2 \geqq \int T^* g T_n^* g \longrightarrow \int (T^* g)^2 = \int g^2$$

and so $\int (T_n^* g)^2 \to \int g^2 = \int (T^* g)^2$. Hence by Lemma 1.2, $T_n^* \xrightarrow{s} T^*$.

In view of Proposition 1.1, our Strong Approximation Theorem 2.2 is sharper than that of [2, 8]. We now prove

THEOREM 1.1. *Φ is a residual set in $\mathscr{D}^w$.*

The proof follows from the following lemma.

LEMMA 1.3. *The identity map $T \to T$ from $\mathscr{D}^w$ to $\mathscr{D}^s$ is continuous at $T \Leftrightarrow T$ is in Φ.*

Proof. ($\Leftarrow$) Suppose that $T \in \Phi$ and $T_n \xrightarrow{w} T$. We have for each $g \in B_\infty$

$$\int Tg T_n g \leqq \| Tg \|_2 \| T_n g \|_2 = \| g \|_2 \| T_n g \|_2 \leqq \int g^2$$

and

$$\int Tg T_n g \longrightarrow \int (Tg)^2 = \int g^2 \qquad \text{as } n \to \infty$$

and so $\int (T_n g)^2 \to \int g^2$ as $n \to \infty$. Hence from Lemma 1.2, $T_n \xrightarrow{s} T$.

($\Rightarrow$) If $T \notin \Phi$, then from $\mathscr{D} = \Phi_1^{-w}$: the closure of Φ_1 in the weak topology [2], there is a sequence $\{T_n\}$ in Φ_1 such that $T_n \xrightarrow{w} T$. Since $T \notin \Phi$, we have from Lemma 1.1, $\int (T_n e)^2 = \int e^2 > \int (Te)^2$, $n = 1, 2, \cdots$, and thus $T_n \nrightarrow T$ in the strong topology.

Proof of Theorem 1.1. Since $\mathscr{D}^w$ and $\mathscr{D}^s$ are metrizable spaces, it follows from Lemma 1.3 that Φ is a G_δ set in $\mathscr{D}^w$. Furthermore, $\mathscr{D}^w$ is a complete metric space and Φ is dense in $\mathscr{D}^w$, and thus Φ is a residual set in $\mathscr{D}^w$.

Similarly we prove that Φ^* is a residual set in $\mathscr{D}^w$. In the remaining part of this section, we will discuss some properties of the set L. The following lemma is obvious.

LEMMA 1.4. (i) *Each U_n is a projection, i.e.,*

$$U_n^2 = U_n, \ U_n^* = U_n$$

(ii) $U_n \xrightarrow{s} I$ *as* $n \to \infty$, *where I denotes the identity operator.*
(iii) $U_n U_m = U_m U_n = U_n$ *if* $n \leqq m$.

Since multiplication on $\mathscr{D}$ is jointly continuous in the strong topology, we have for each $T \in \mathscr{D}$, $U_n T \xrightarrow{s} T$ and $U_n T U_n \xrightarrow{s} T$ as $n \to \infty$. It is worthwhile to point out that $\rho(U_n, I) \geqq 1$, $n = 1, 2, \cdots$. Thus the uniform topology is strictly stronger than the strong topology. It is easily shown that $T \in L^{-u} \Leftrightarrow U_n T \xrightarrow{u} T \Leftrightarrow U_n T U_n \xrightarrow{u} T$. It follows that $L = L^{-n} \subset L^{-u} \subsetneqq L^{-s} = L^{-w} = \mathscr{D}$. In Corollary to Theorem 2.4, we will show that the norm topology is strictly stronger than the uniform topology, but it is not clear whether the same is true on the set L. What is L^{-u}? The question is not completely answered. But we state the following:

THEOREM 1.2. *L is nowhere dense in $\mathscr{D}^u$.*

Proof. It will be enough to show that $S(T, \varepsilon) \notin L^{-u}$ for each $\varepsilon > 0$ and $T \in L^{-u}$, where $S(T, \varepsilon) = \{R: \rho(R, T) < \varepsilon\}$. Let $T_{n_0} = U_{n_0} T U_{n_0}$ be such that $\rho(T, T_{n_0}) < \varepsilon/2$. Then $S(T_{n_0}, \varepsilon/2) \subset S(T, \varepsilon)$. Define $Q_\delta, 0 < \delta < 1$, by $Q_\delta = (1 - \delta) T_{n_0} + \delta I$. It follows that $Q_\delta \xrightarrow{u} T_{n_0}$ as $\delta \to 0$, and so there is δ_ε, $0 < \delta_\varepsilon < 1$, such that $\rho(T_{n_0}, Q_{\delta_\varepsilon}) < \varepsilon/2$. Thus

$$Q_{\delta_\varepsilon} \in S(T_{n_0}, \varepsilon/2) \subset S(T, \varepsilon) \; .$$

But we claim $Q_{\delta_\varepsilon} \notin L^{-u}$. It follows from Lemma 1.4 that

$$U_n Q_{\delta_\varepsilon} = (1 - \delta_\varepsilon) T_{n_0} + \delta_\varepsilon U_n \; , \qquad n \geqq n_0 \; ,$$

and so $U_n Q_{\delta_\varepsilon} - Q_{\delta_\varepsilon} = \delta_\varepsilon(U_n - I), n \geqq n_0$. Since

$$\rho(U_n Q_{\delta_\varepsilon}, Q_{\delta_\varepsilon}) \geqq \delta_\varepsilon \rho(U_n, I) \geqq \delta_\varepsilon > 0 \; , \qquad n \geqq n_0 \; ,$$

we have $U_n Q_{\delta_\varepsilon} \not\to Q_{\delta_\varepsilon}$ in the uniform topology. Hence $Q_{\delta_\varepsilon} \notin L^{-u}$ and $S(T, \varepsilon) \not\subset L^{-u}$.

We will prove

THEOREM 1.3. *$T \subset L^{-u} \Leftrightarrow T(B_\infty)$ is strongly conditionally compact in L_1.*

Proof. ($\Rightarrow$) Suppose $T \in L^{-u}$. It will be enough to show that $T(B_\infty)$ is sequentially compact in the L_1-norm. Let $\{f_j\}$ be any sequence in B_∞. Let $T_n = U_n T U_n, n = 1, 2, \cdots$. Then $T_n \xrightarrow{u} T$ as $n \to \infty$. Since the range of T_n is contained in a finite dimensional subspace of L_1, T_n is a compact operator in L_1 [9]. Thus we have a family of subsequences $\{f_{nj}\}_j, n = 1,2, \cdots$, such that for each n, $\{f_{n+1j}\}_j \subset \{f_{nj}\}_j$ and $T_n f_{nj} \to g_n$ in the L_1-norm as $j \to \infty$ for some $g_n \in L_1$. Thus for each n, $T_n f_{jj} \to g_n$ in the L_1-norm as $j \to \infty$. Then from the inequality:

$$\| Tf_{jj} - Tf_{kk} \|_1 \leqq \| Tf_{jj} - T_n f_{jj} \|_1 + \| T_n f_{jj} - T_n f_{kk} \|_1 + \| T_n f_{kk} - Tf_{kk} \|_1 \leqq 2\rho(T, T_n) + \| T_n f_{jj} - T_n f_{kk} \|_1 \; ,$$

it follows that $\{T_{jj}\}_j$ is a Cauchy sequence in the L_1-norm and so $Tf_{jj} \to g$ in the L_1-norm for some $g \in L_1$.

($\Leftarrow$) Suppose that $T(B_\infty)$ is conditionally compact in the L_1-norm. If $T \notin L^{-u}$, then $U_n T \not\to T$ in the uniform topology. Thus there are $\varepsilon > 0$ and a sequence $\{f_{n_i}\}$ in B_∞ such that

$$\int | U_{n_i} T f_{n_i} - T f_{n_i} | > \varepsilon, i = 1, 2, \cdots \; .$$

From the assumption, there is a subsequence $\{g_{n_i}\}$ of $\{f_{n_i}\}$ such that $Tg_{n_i} \to q$ in the L_1-norm for some $q \in L_1$. But then

$$\varepsilon < \int | U_{n_i} T g_{n_i} - T g_{n_i} | \leqq \| U_{n_i}(T g_{n_i}) - U_{n_i} q \|_1 + \| U_{n_i} q - q \|_1 + \| q - T g_{n_i} \|_1 \leqq 2 \| T g_{n_i} - q \|_1 + \| U_{n_i} q - q \|_1 \longrightarrow 0$$

as $n_i \to \infty$, a contradiction.

2. Approximation theorems. The problem of approximating doubly stochastic operators is discussed by several authors [2, 7, 8]. In this section we discuss various approximation theorems. First we will give a simple proof for the weak approximation theorem [2] and a sharper form of the strong approximation theorem [2, 8].

THEOREM 2.1. *(Brown [2]) For each* $T \in \mathscr{D}$ *there is a sequence* $\{T_n : n = 1, 2, \cdots\}$ *in* Φ_1 *such that* $T_n \xrightarrow{w} T$ *as* $n \to \infty$.

We prove the following lemma.

LEMMA 2.1. *For each* $T \in \mathscr{D}$ *and for* $n = 1.2, \cdots$, *there is* $T_n \in \Phi_1$ *such that*

$$\int 1_{D_i^n} T 1_{D_j^n} = \int 1_{D_i^n} T_n 1_{D_j^n}, \qquad 1 \leqq i, j \leqq 2^n,$$

where $D_i^n = [(i-1)/2^n, i/2^n), 1 \leqq i < 2^n$, *and* $D_{2^n}^n = [1 - 1/2^n, 1]$.

Proof. Let $f_i = 1_{D_i^n}$ and $a_{ij} = \int f_i T f_j$. For each i and for each j,

$$\sum_{k=1}^{2^n} a_{ik} = 2^{-n} \quad \text{and} \quad \sum_{k=1}^{2^n} a_{kj} = 2^{-n}.$$

We put

$$x_{ij} = (i-1)h + \sum_{k=1}^{j} a_{ik}, \; y_{ij} = (i-1)h + \sum_{k=1}^{j} a_{ki},$$

where $h = 2^{-n}$ and $1 \leqq i, j \leqq 2^n$.

Let $\varphi : I \to I$ be such that

$$\varphi(x) = x - x_{ij} + y_{j+1\, i-1} \quad \text{on} \quad [x_{ij}, x_{i\, j+1}),$$

where $1 \leqq i, j \leqq 2^n$. We note that $x_{10} = y_{10} = 0$, $x_{j0} = x_{j-1\, 2n}$ and $y_{j0} = y_{j-1\, 2n}$. Clearly $T_\varphi \in \Phi_1$. It remains to show

$$a_{ij} = \int f_i T_\varphi f_j, \qquad 1 \leqq i, j \leqq 2^n.$$

Since $D_j^n = [y_{j-1\, 2^n}, y_{j\, 2^n})$ and $D_i^n = [x_{i-1\, 2^n}, x_{i\, 2^n})$, we have $\varphi^{-1}(D_j^n) = \bigcup_{k=1}^{2^n} \varphi^{-1}\{[y_{j\, k-1}, y_{jk})\} = \bigcup_{k=1}^{2^n} [x_{k\, j-1}, x_{kj})$ and

$$\ell\{D_i^n \cap \varphi^{-1}(D_j^n)\} = \ell\{[x_{i\, j-1}, x_{ij})\} = a_{ij}.$$

Thus we have

$$\int f_i T_\varphi f_j = \int 1_{D_i^n \cap \varphi^{-1}(D_j^n)} = a_{ij}.$$

If we set $T_n = T_\varphi$, then T_n is a desired one.

Proof of Theorem 2.1. Given n and $T \in \mathscr{D}$, we have from Lemma 2.1, $T_n \in \Phi_1$ and

$$\int hTk = \int hT_n k\ , \qquad h, k \in U_n(B_\infty)\ ,$$

and hence the equality holds for $h, k \in U_m(B_\infty)$, $m \leqq n$. It follows that for $m(\leqq n)$ and $f, g \in B_\infty$,

$$\int fU_m TU_m g = \int fU_m T_n U_m g\ .$$

If a subsequence $\{T_{n_i}\}$ of $\{T_n\}$ converges weakly to Q, then for each fixed m and for every $f, g \in B_\infty$,

$$\int fU_m TU_m g = \int fU_m T_{n_i} U_m g \longrightarrow \int fU_m QU_m g\ , \qquad n_i \longrightarrow \infty$$

and $\int fU_m TU_m g = \int fU_m QU_m g$, and so $T = Q$.

Since $\mathscr{D}$ is compact in the weak topology, we have $T_n \xrightarrow{w} T$ as $n \rightarrow \infty$.

We can also prove Theorem 2.1 by the usual approximation of $f \in L_\infty$ by $U_n f$ without using the weak compactness of $\mathscr{D}$.

We will prove the following strong approximation theorem. Let $\mathrm{co}(A)$ be the convex hull of the set A.

THEOREM 2.2. *For each $T \in \mathscr{D}$ there is a sequence $\{R_n\}$ in $\mathrm{co}(\Phi_1)$ such that*

$$R_n \xrightarrow{s} T \quad \textit{and} \quad R_n^* \xrightarrow{s} T^* \quad \textit{at}\ n \longrightarrow \infty\ .$$

We need to prove

LEMMA 2.2. *For each $T \in \mathscr{D}$ there is a sequence $\{R_n\}$ in $\mathrm{co}(\Phi_1)$ such that*
$U_n TU_n = R_n U_n$ and $U_n T^* U_n = R_n^* U_n$ for $n = 1, 2, \cdots$.

Proof. If we write

$$U_n f = \sum_{i=1}^{2^n} c_i 1_{D_i^n}, \qquad c_i = 2^n \int_{D_i^n} f\ , \qquad f \in L_\infty, n = 1, 2, \cdots\ ,$$

then

$$(U_n TU_n) f = \sum_{i=1}^{2^n} c_i (U_n TU_n) 1_{D_i^n} = \sum_{i,j=1}^{2^n} c_i a_{ij} 1_{D_j^n}\ ,$$

where $a_{ij} = 2^n \int_{D_j^n} T1_{D_i^n}$.

It is easy to see that the matrix $D_n = (a_{ij}: 1 \leqq i, j \leqq 2^n)$ is doubly stochastic [6]. Let $\mathscr{M}_n$ be the subspace of L_∞ spanned by $\{1_{D_i^n}: 1 \leqq i \leqq 2^n\}$. By the linear operator D_n we mean the linear operator on $\mathscr{M}_n$ into itself corresponding to the doubly stochastic matrix D_n. Then the operators U_nTU_n and D_nU_n satisfy $U_nTU_n = D_nU_n$. Similarly we have $U_nT^*U_n = D_n^*U_n$, where the operator (matrix) D_n^* is the adjoint of the operator (matrix) D_n.

From a theorem of Birkhoff [1, 6] the doubly stochastic matrix D_n is a convex combination of permutation matrices P_k, i.e., $D_n = \sum_{k=1}^r c_kP_k$, $1 \leqq r \leqq (2^n - 1)^2 + 1$, and so for adjoint, $D_n^* = \sum_{k=1}^r c_kP_k^*$. By the operator P_k we mean the linear operator on the subspace $\mathscr{M}_n$ corresponding to the matrix P_k.

For each operator P_k we choose $T_k = T_{\varphi_k} \in \Phi_1$ such that $P_k1_{D_i^n} = T_k1_{D_i^n}$, $1 \leqq i \leqq 2^n$. It follows that $P_k^*1_{D_i^n} = T_k^*1_{D_i^n}$ for each i. Thus, $U_nTU_n = (\sum_{k=1}^r c_kT_k)U_n$ and $U_nT^*U_n = (\sum_{k=1}^r c_kT_k)^*U_n$. By setting $R_n = \sum_{k=1}^r c_kT_k$, we establish the lemma.

Proof of Theorem 2.2. Given $T \in \mathscr{D}$, we have from Lemma 2.2 a sequence $\{R_n\}$ in $\mathrm{co}(\Phi_1)$ such that

$$U_nTU_n = R_nU_n \quad \text{and} \quad U_nT^*U_n = R_n^*U_n\,, \qquad n = 1, 2, \cdots .$$

From Lemma 1.4 and the argument following Lemma 1.4,

$$\begin{aligned} \| Tf - R_nf \|_1 &\leqq \| Tf - U_nTU_nf \|_1 + \| R_nU_nf - R_nf \|_1 \\ &\leqq \| Tf - U_nTU_nf \|_1 + \| U_nf - f \|_1 \longrightarrow 0 \qquad \text{as } n \to \infty\,, \end{aligned}$$

and hence $R_n \xrightarrow{s} T$. Similarly, $R_n^* \xrightarrow{s} T^*$.

We state the uniform approximation theorems:

THEOREM 2.3. *If the kernel for $T \in L$ is an element of $L_2(I \times I)$, then there is a sequence $\{T_m\}$ in $\mathrm{co}(\Phi)$ such that*

$$\| T - T_m \|_2 \longrightarrow 0 \qquad as\ m \to \infty\,.$$

THEOREM 2.4. *For each $T \in L$, there is a sequence $\{T_m\}$ in $\mathrm{co}(\Phi)$ such that $T_m \xrightarrow{u} T$ as $m \to \infty$.*

Theorems 2.3 and 2.4 are easily derived from the following four lemmas.

LEMMA 2.3. *For each $T \in \mathscr{D}$ and each $n = 1, 2, \cdots$,*

$$\rho(T^{2n}, E) \leqq \| T^{2n} - E \|_2 \leqq \| T - E \|_2^{2n}\,,$$

where $Ef = \int f$.

The proof is immediate.

We review the notion of the independence of measurable maps (random variables) [3]. Let φ_1 and φ_2 be measurable maps on the unit interval I into itself. Then φ_1 and φ_2 are said to be independent if $\ell(\varphi_1^{-1}(A) \cap \varphi_2^{-1}(B)) = \ell(\varphi_1^{-1}(A))\ell(\varphi_2^{-1}(B))$, $A, B \in \mathscr{B}(I)$. It follows easily that if φ_1 and φ_2 are independent, measurable maps on I, then

$$\int f(\varphi_1) f(\varphi_2) = \int f(\varphi_1) \int f(\varphi_2) \, , \qquad f \in L_1 \, .$$

LEMMA 2.4. *Let φ_1 and φ_2 be independent, measure perserving maps on I. If $T = cT_{\varphi_1} + (1-c)T_{\varphi_2}$, $0 < c < 1$, then $\rho(T^{2n}, E) \leqq \| T^{2n} - E \|_2 \leqq \| T - E \|_2^{2n} \leqq \{c^2 + (1-c)^2\}^n$, $n = 1, 2, \cdots$, and $\rho(T^{2n}, E) \leqq \| T^{2n} - E \|_2 \to 0$ as $n \to \infty$.*

Proof. We observe that for each $T \in \mathscr{D}$,

$$\begin{aligned} \| T - E \|_2 &= \sup \{ \| Tf - Ef \|_2 : f \in B_2 \} = \sup \{ \| T(f - Ef) \|_2 : f \in B_2 \} \\ &\leqq \sup \{ \| Tg \|_2 : g \in B_2, Eg = 0 \} \, . \end{aligned}$$

If $T = cT_{\varphi_1} + dT_{\varphi_2}$, $0 < c < 1$, $d = 1 - c$, then for each g such that $Eg = 0$ and $\| g \|_2 \leqq 1$,

$$\begin{aligned} \| Tg \|_2^2 &= \int (cT_{\varphi_1} g + dT_{\varphi_2} g)^2 \\ &= (c^2 + d^2) \int g^2 + 2cd \Big(\int g \Big)^2 \leqq c^3 + d^2 \, , \end{aligned}$$

and so $\| T - E \|_2 \leqq (c^2 + d^2)^{1/2}$ and $\| T - E \|_2^{2n} \leqq (c^2 + d^2)^n$. Thus Lemma 2.4 follows.

Since there are no invertible measure preserving maps φ_1 and φ_2 on I that are independent, Lemma 2.4 can not be strengthened any further.

LEMMA 2.5. *For each $n = 1, 2, \cdots$, there are measure preserving maps θ_1 and θ_2 on I such that*

$$\| \{ cT_{\theta_1} + (1-c) T_{\theta_2} \}^{2m} - U_n \|_2 \leqq 2^{m/2} \delta^{2m} \, , \qquad m = 1, 2, \cdots \, ,$$

where $\delta^2 = c^2 + (1-c)^2$, $0 < c < 1$.

Proof. Let n be a positive integer. For each integer i, $1 \leqq i \leqq 2^n$, the operator V_{ni} defined by $V_{ni} f = 2^n \int_{D_i^n} f$, $f \in L_\infty(D_i^n)$, is a doubly stochastic operator on $L_\infty(D_i^n)$. We note that V_{ni} is an analogue of the operator E, i.e., $V_{ni}^2 = V_{ni}$ and $V_{ni} = V_{ni}R = RV_{ni}$ for each doubly stochastic operator R on $L_\infty(D_i^n)$. Hence we have an analogue of Lemma 2.3, i.e.,

$$\| R^{2m} - V_{ni} \|_2 \leqq \| R - V_{ni} \|_2^{2m} , \qquad m = 1, 2, \cdots ,$$

for each doubly stochastic operator R on $L_\infty(D_i^n)$, where $\| \cdot \|_2$ denotes the $L_2(D_i^n)$-operator norm. Let $\varphi_{ij}: D_i^n \to D_i^n, j = 1, 2$, be independent, measure preserving maps. We define

$$T_i = cT_{\varphi_{i1}} + (1 - c)T_{\varphi_{i2}} , \qquad 0 < c < 1 .$$

Clearly T_i is doubly stochastic on $L_\infty(D_i^n)$. By a similar argument given in the proof of Lemma 2.4, we have

$$\| T_i^{2m} - V_{ni} \|_2 \leqq \| T_i - V_{ni} \|_2^{2m} \leqq \delta^{2m} , \qquad m = 1, 2, \cdots .$$

Let $\theta_j: I \to I, j = 1, 2$, be such that

$$\theta_j(x) = \varphi_{ij}(x) , \qquad x \in D_i^n , \qquad 1 \leqq i \leqq 2^n .$$

Then θ_1 and θ_2 are measure preserving but not necessarily independent. Let $T = cT_{\theta_1} + (1 - c)T_{\theta_2}$. For each $m = 1, 2, \cdots$ and each $f \in B_2$,

$$\int | T^{2m}f - U_nf |^2 = \sum_{i=1}^{2^n} \int_{D_i^n} | T_i^{2m}f - V_{ni}f |^2 \leqq \sum_{i=1}^{2^n} \| T_i^{2m} - V_{ni} \|_2^2 \leqq 2^n\delta^{4m} ,$$

and so $\| T^{2m} - U_n \|_2^2 \leqq 2^n\delta^{4m}$. It follows that $\| T^{2m} - U_n \|_2 \to 0$ as $m \to \infty$.

Lemma 2.6. *For each $T \in \mathscr{D}$ and each U_n, there exist $S_n \in \mathrm{co}(\Phi_1)$ and $R_n \in \mathrm{co}(\Phi)$ such that*

$$\| U_nTU_n - S_nR_n^{2m} \|_2 \leqq 2^{n/2}\delta^{2m} , \qquad m = 1, 2, \cdots .$$

Proof. For each $T \in \mathscr{D}$ there is $S_n \in \mathrm{co}(\Phi_1)$, by Lemma 2.2, such that $U_nTU_n = S_nU_n$.

By Lemma 2.5, there is $R_n \in \mathrm{co}(\Phi)$ such that

$$\| U_n - R_n^{2m} \|_2 \leqq 2^{n/2}\delta^{2m} ,$$

and thus

$$\| U_nTU_n - S_nR_n^{2m} \|_2 = \| S_nU_n - S_nR_n^{2m} \|_2 \leqq \| U_n - R_n^{2m} \|_2 \leqq 2^{n/2}\delta^{2m} .$$

Proof of Theorem 2.3. Let $k(\cdot, \cdot)$ be the kernel for $T \in L$ and $k_n(\cdot, \cdot)$ the kernel for U_nTU_n. If $k(\cdot, \cdot) \in L_2(I \times I)$, then

$$\| T - U_n T U_n \|_2^2 \leqq \int_{I\times I} | k - k_n |^2 d\ell^2 \longrightarrow 0 \qquad \text{as } n \to \infty \ .$$

By Lemma 2.6, there are $S_n \in \mathrm{co}(\Phi_1)$ and $R_n \in \mathrm{co}(\Phi)$ such that

$$\begin{aligned} \| T - S_n R_n^{2m} \|_2 &\leqq \| T - U_n T U_n \|_2 + \| U_n T U_n - S_n R_n^{2m} \|_2 \\ &\leqq \left\{ \int_{I\times I} | k - k_n |^2 d\ell^2 \right\}^{1/2} + 2^{n/2} \delta^{2m} \ . \end{aligned}$$

Then $S_n R_n^{2m} \in \mathrm{co}(\Phi)$ and the assertion follows by usual argument.

Proof of Theorem 2.4. Let k and k_n be as in the proof of Theorem 2.3. Then

$$\rho(T, U_n T U_n) \leqq \int_{I\times I} | k - k_n | \, d\ell^2 \longrightarrow 0 \qquad \text{as } n \to \infty \ .$$

By using Lemma 2.6, we have

$$\begin{aligned} \rho(T, S_n R_n^{2m}) &\leqq \rho(T, U_n T U_n) + \rho(U_n T U_n, S_n R_n^{2m}) \\ &\leqq \int_{I\times I} | k - k_n | \, d\ell^2 + \| U_n T U_n - S_n R_n^{2m} \|_2 \ , \end{aligned}$$

from which the assertion follows.

As a corollary to Theorem 2.4 we have the following.

COROLLARY. *The norm topology is strictly stronger than the uniform topology.*

Proof. It follows from Theorem 2.4 that there is a sequence $\{T_n\}$ in $\mathrm{co}(\Phi)$ such that $T_n \xrightarrow{n} E$ as $n \to \infty$. Since $\mu_{T_\varphi} \perp \ell^2$ for each $T_\varphi \in \Phi$, we have $\mu_n = \mu_{t_n} \perp \ell^2$ for each $n = 1, 2, \cdots$. Let (A_n, B_n) be a decomposition of $I \times I$ such that $\ell^2(B_n) = 0$ and $\mu_n(A_n) = 0$. Then we have

$$\| \mu_n - \ell^2 \| \geqq | \mu_n(A_n) - \ell^2(A_n) | = \ell^2(A_n) = 1 \ ,$$

and so $\mu_n \nrightarrow \ell^2$ in the norm topology.

It is not clear whether we can choose a sequence $\{T_m\}$ from $\mathrm{co}(\Phi_1)$ instead of $\mathrm{co}(\Phi)$ in Theorems 2.3 and 2.4. By the Approximation Theorems we have

$$\begin{array}{c} \mathrm{co}^{-u}(\Phi_1) \subset \mathrm{co}^{-u}(\Phi) \subset \mathrm{co}^{-s}(\Phi) = \mathrm{co}^{-s}(\Phi_1) = \Phi_1^{-w} = \mathscr{D} . \\ \quad\;\; \cup \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \\ \quad\;\; L \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \end{array}$$

It remains to be determined whether the inclusion relations in

$$\mathrm{co}^{-u}(\Phi_1) \subset \mathrm{co}^{-u}(\Phi) \subset \mathrm{co}^{-s}(\Phi)$$

are proper.

This paper contains a part of the author's Ph. D. thesis submitted to the University of Washington. The work was directed by Professor R. M. Blumenthal to whom the author is deeply indebted for encouragement and guidance. The author also wishes to thank Professor J. V. Ryff for many helpful conversations.

BIBLIOGRAPHY

1. G. Birkhoff, *Three observations on linear algebra*, Univ. Nac. Tucuman. Rev. (A) **5** (1946), 147-151.
2. J. R. Brown, *Approximation theorems for Markov operators*, Pacific J. Math. **16** (1966), 13-23.
3. J. L. Doob, *Stochastic Processes*, John Wiley, New York, 1953.
4. P. R. Halmos, *Measure Theory*, Van Nostrand, New York, 1954.
5. ———, *Lectures on Ergodic Theory*, Math. Soc. Japan, Tokyo, 1956.
6. L. Mirsky, *Results and problems in the theory of doubly stochastic matrices*, Z. Wahr. **1** (1963), 319-334.
7. J. E. L. Peck, *Doubly stochastic measures*, Michigan Math. J. **6** (1959), 217-220.
8. J. V. Ryff, *Orbits of L_1-functions under doubly stochastic transformations*, Trans. Amer. Math. Soc. **117** (1965), 92-100.
9. A. C. Zaanen, *Linear Analysis*, North-Holland, Amsterdam, 1953.

Received April 21, 1966.

SIMON FRASER UNIVERSITY
B. C. CANADA