Operators and Matrices
Volume 3, Number 2 (2009), 227–233

STOCHASTIC OPERATORS AND EXTREME POINTS

DON HADWIN, ERIC NORDGREN AND HEYDAR RADJAVI

(*Communicated by J. W. Helton*)

Abstract. The operators on $L^p, 1 \leqslant p < \infty$ that preserve positivity and the constants are shown to have the composition operators as extreme points. In the case of the unit interval with Lebesgue measure they constitute the closed convex hull of these extreme points, but this is not true of all measure spaces.

1. Introduction

Suppose $(X, \mathscr{F})$ is a standard Borel space and m is a probability measure on the Borel sets $\mathscr{F}$ of X. An operator on $L^p(m), 1 \leqslant p < \infty$ is called *stochastic* in case the image of every nonnegative function is a nonnegative function and the constant one function is its own image. In the special case where X is a set with n points of equal measure $L^p(m)$ may be identified with $\mathbb{C}^n$ and the definition coincides with the usual one for row stochastic matrices. It is easy to see that for a row stochastic matrix A whose entries consist of zeros and ones, where the one in row i lies in column $\phi(i)$, the image Af of a vector f, i.e. a function on the set $X = \{1, 2, \ldots, n\}$, is $f \circ \phi$. In other words A is a composition operator. It is a familiar fact that these special 0,1 stochastic matrices, i.e. the composition operators, are the extreme points of the set of row stochastic matrices and every row stochastic matrix is a convex combination of these extreme points. Our first objective is to prove that the composition operators are the extreme points of the stochastic operators. Also, we show in the special case when m is Lebesgue measure on $[0,1]$ that, relative to the weak operator topology, the convex hull of the composition operators is dense in the set of stochastic operators on L^2, but this is not true for all probability spaces.

Operators related to what we are calling stochastic operators have a long history. Of particular importance is Iwanik's paper [2] of 1976 in which he shows that the composition operators P on the L^p space of a standard Borel space are characterized by the properties (i) $P|f| = |Pf|$ and (ii) $P1 = 1$. He also points out that these two properties imply that characteristic functions are carried to characteristic functions. As mentioned by Iwanik, it had been proved earlier [3] by Phelps in 1963 that the class of operators from one L^∞ space to another that are contractive and satisfy $0 \leqslant P1 \leqslant 1$ have the operators that send characteristic functions to characteristic functions as its extreme points, and these are precisely the operators that preserve products.

Mathematics subject classification (2000): Primary 47L07; Secondary 47B33.
Keywords and phrases: positivity preserving operators, extreme points.

Paper OaM-03-13

2. Composition operators are extreme

An operator T on $L^p = L^p(m)$, where $1 \leqslant p < \infty$ and m is a probability measure on a standard Borel space $(X, \mathscr{F})$, is called *stochastic* in case $Tf \geqslant 0$ whenever $f \geqslant 0$ and $T1 = 1$. (Iwanik [2, Lemma 1] points out that the positivity condition implies boundedness.) Let $\mathfrak{S}$ be the set of all stochastic operators. Sourour [6, Theorem 1 and following Remark] shows that a bounded operator on $L^p, 1 \leqslant p < \infty$ that sends every function with nonnegative values to another such function is necessarily a pseudo-integral operator (see also [1]). Thus there is a family of measures $\{\mu_x\}_{x \in X}$ on X such that for every $f \in L^p(m)$ and almost every $x \in X$

$$Tf(x) = T_\mu f(x) = \int_X f \,\mathrm{d}\mu_x,$$

and the function $x \mapsto \mu_x(A)$ is measurable for every $A \in \mathscr{F}$. A family of measures μ_x will be called *measurable* if it satisfies the second of these conditions. In particular, every stochastic operator T is a pseudo-integral operator, and, in addition, since $T1 = 1$, almost every μ_x is a probability measure. The paper of Sourour, [6], should be consulted for interesting examples and properties of pseudointegral operators. We may summarize our observations as follows.

THEOREM 1. (Sourour) *A bounded operator on $L^p, 1 \leqslant p < \infty$ is stochastic if and only if it is a pseudointegral operator that is induced by a measurable family of probability measures.*

If T is induced by a family of measures μ_x, then we can define a measure μ on $X \times X$ by

$$\mu(E) = \int_X \int_X \chi_E(x,y)\,\mathrm{d}\mu_x(y)\,\mathrm{d}m(x)$$

for each measurable set E in the product space, and then

$$\langle Tf, g\rangle = \int_{X \times X} f(y)\overline{g(x)}\,\mathrm{d}\mu(x,y) \tag{1}$$

whenever $f \in L^p$ and $g \in L^q$ where p and q are conjugate indices. Conversely, if μ is a measure on $X \times X$ such that it induces a bounded operator satisfying equation (1), then one obtains a measurable family of measures from μ by taking its disintegration with respect to m.

It is clear that if a family $\{\nu_x\}_{x \in X}$ of measures is measurable and is dominated by a measurable family $\{\mu_x\}_{x \in X}$, in the sense that $\nu_x \leqslant M\mu_x$ a.e.-$[m]$ for some constant M, then $\{\nu_x\}_{x \in X}$ induces a bounded operator on $L^p(m)$ whenever $\{\mu_x\}_{x \in X}$ does.

Among the stochastic operators on $L^p(m)$ are the composition operators. These are operators of the form $C_\phi f = f \circ \phi$ for some measurable transformation ϕ of X into itself. Here $C_\phi = T_\mu$ with $\mu_x = \delta_{\phi(x)}$, where for any y, δ_y is the unit point mass at y. The composition operators are extreme points of the set of stochastic operators $\mathfrak{S}$. This follows from the familiar fact that point masses are extreme points of the set of probability measures. For if C_ϕ is an average of two stochastic operators T_λ and

T_ν, then $2\delta_{\phi(x)} = \lambda_x + \mu_x$ a.e.-$[m]$. Since all the measures are probability measures, $\lambda_x = \nu_x = \delta_{\phi(x)}$ a.e.-$[m]$, and hence $T_\lambda = T_\nu = C_\phi$. Thus C_ϕ is extreme. The converse is contained in the following theorem.

THEOREM 2. *A stochastic operator on* $L^p, 1 \leqslant p < \infty$ *is an extreme point of the family of all stochastic operators if and only if it is a composition operator.*

Proof. Sufficiency has already been established. For necessity suppose $T \in \mathfrak{S}$, but T is not a composition operator. Then T is induced by a measurable family $\{\mu_x\}_{x\in X}$ such that μ_x is not a point mass for all x in a set F_0 with $m(F_0) > 0$.

The σ-algebra $\mathscr{F}$ is generated by a countable sequence of sets (A_n), and consequently to every $x \in F_0$ corresponds at least one n such that $0 < \mu_x(A_n) < 1$. The function t_n on F_0 defined by $t_n(x) = \mu_x(A_n)$ is measurable for each n, and for every $x \in F_0$ there is an n for which $t_n(x) \in (0,1)$. Thus there is a single n such that $t_n(x) \in (0,1)$ on a subset F_1 of F_0 of positive measure. In fact, by choosing a sufficiently small $\varepsilon \in (0,1/2)$, we may suppose $t_n(x) \in (\varepsilon, 1-\varepsilon)$ for all x in a set $F \subset F_1$ of positive measure. We will from here on write simply A for this A_n and put $B = X \setminus A$. Thus for all $x \in F$ we have both $\varepsilon < \mu_x(A) < 1-\varepsilon$ and $\varepsilon < \mu_x(B) < 1-\varepsilon$, and also $\mu_x(A) + \mu_x(B) = 1$.

Our objective is to define measurable families of probability measures $\{\lambda_x\}$ and $\{\nu_x\}$, both different from $\{\mu_x\}$, such that for some fixed $\tau \in (0,1)$ we have $\mu_x = \tau\lambda_x + (1-\tau)\nu_x$ for all $x \in X$. First define probability measures α_x and β_x as follows: if $D \in \mathscr{F}$, then

$$\alpha_x(D) = \begin{cases} \dfrac{\mu_x(D\cap A)}{\mu_x(A)} & \text{if } x \in F, \\ \mu_x(D) & \text{if } x \neq F \end{cases}$$

and

$$\beta_x(D) = \begin{cases} \dfrac{\mu_x(D\cap B)}{\mu_x(B)} & \text{if } x \in F, \\ \mu_x(D) & \text{if } x \neq F. \end{cases}$$

Then clearly $\mu_x = \mu_x(A)\alpha_x + \mu_x(B)\beta_x$ for $x \in F$.

Let $\lambda_x = (1-\varepsilon)\alpha_x + \varepsilon\beta_x$. Then for $x \in F$, $\mu_x(A) < 1-\varepsilon = \lambda_x(A)$ and $\lambda_x(B) = \varepsilon < \mu_x(B)$, so λ_x is not a multiple of μ_x. Also, define ν_x by

$$\nu_x = \begin{cases} \dfrac{\mu_x(A) - \varepsilon + \varepsilon^2}{1-\varepsilon}\alpha_x + \dfrac{\mu_x(B) - \varepsilon^2}{1-\varepsilon}\beta_x & \text{if } x \in F, \\ \mu_x & \text{if } x \notin F. \end{cases}$$

(This definition is obtained by putting $\tau = \varepsilon$ and solving the equation $\mu_x = \tau\lambda_x + (1-\tau)\nu_x$ for ν_x.) It is easy to check that the coefficients of α_x and β_x in the above equation are positive and add up to one, so $\{\nu_x\}$ is a measurable family of probability measures. Also, as just mentioned, the family $\{\mu_x\}$ is a convex combination of the families $\{\lambda_x\}$ and $\{\nu_x\}$, i.e. for all x,

$$\mu_x = \tau\lambda_x + (1-\tau)\nu_x.$$

Finally, for $x \in F$, we have $\lambda_x(A) = 1 - \varepsilon > \mu_x(A)$, so $\lambda_x \neq \mu_x$.

Since $\{\lambda_x\}$ is dominated by a multiple of $\{\mu_x\}$ ($\varepsilon\lambda_x \leqslant \mu_x$), it induces a stochastic operator T_λ, and ν_x induces a stochastic operator T_ν. These operators satisfy $T = \tau T_\lambda + (1-\tau)T_\nu$, and hence T is not extreme. Thus only composition operators are extreme.

A proof of the preceding theorem could also have been constructed by employing a measurable cross section theorem. Combining the preceding theorem with Iwanik's Theorem 1 in [2], we obtain the following very slight improvement of his characterization.

THEOREM 3. *For a stochastic operator T on $L^p, 1 \leqslant p < \infty$ the following are equivalent:*

(a) T is an extreme point of the set of all stochastic operators;

(b) T is a composition operator;

(c) T is multiplicative in the sense that if f, g and fg belong to L^p, then $T(fg) = (Tf)(Tg)$;

(d) T preserves squares in the sense that if f and f^2 are in L^p, then $T(f^2) = (Tf)^2$;

(e) T preserves the class of characteristic functions.

(f) $T1 = 1$ and $|Tf| = T|f|$ for all $f \in L^p$.

Proof. The preceding theorem shows the equivalence of (a) and (b). Iwanik proves the equivalence of (b), (e) and (f). Finally, it is clear that each of (b), (c) and (d) implies its successor.

3. Convex hull of extreme points

We wish to show that the set $\mathfrak{S}$ of stochastic operators is the strong operator topology closure of the convex hull of its set of extreme points. The set $\mathfrak{S}$ is not bounded, and therefore is not compact in either the strong or weak operator topologies. Thus the Krein Milman Theorem is not directly applicable.

THEOREM 4. *The strong operator topology closed convex hull of the set of composition operators on L^2 of the unit interval with Lebesgue measure contains all stochastic operators.*

Proof. Given a stochastic operator $T = T_\mu$, we will approximate μ by measures μ_n supported on lines of slope one. The corresponding stochastic operators T_n will be shown to converge to T in the weak operator topology, and it will be shown that each T_n is a convex combination of composition operators, and $\|T_n\| \leqslant \|T\|$. To accomplish this we will first introduce a sequence of partitions of the unit square J and with each partition a family of measures m_{ij}.

Fix a natural number n, and subdivide J into 2^{2n} squares J_{ij}, each of side length 2^{-n}. Thus, if $I_i = [(i-1)2^{-n}, i2^{-n})$ for $1 \leqslant i \leqslant 2^n$, then $J_{ij} = I_i \times I_j$. If $D_{ij} = \{((i-1)2^{-n}+x, (j-1)2^{-n}+x) : 0 \leqslant x < 2^{-n}\}$ is the diagonal of J_{ij}, then define m_{ij} by $m_{ij}(E) = m(\pi(E \cap D_{ij}))$, where π is the projection of J onto the x-axis, so $\pi(x,y) = x$. Then m_{ij} is a measure of total mass 2^{-n} that is supported on D_{ij} for $1 \leqslant i, j \leqslant 2^n$. It is easy to see that the marginal measures of m_{ij} are the restrictions of m to the intervals I_i and I_j respectively. We remark that if E_{ij} is the pseudo-integral operator induced by m_{ij}, then E_{ij} is a partial isometry with initial space $L^2(I_j)$ and final space $L^2(I_i)$. If $f \in L^2(I_j)$, so f is an L^2 function supported on I_j, then $E_{ij}f$ is just the translate of f by the appropriate amount and is supported on I_i. The collection of all E_{ij} is a set of matrix units corresponding to the decomposition $L^2 = \bigoplus_{k=1}^{2^n} L^2(I_k)$

Define numbers a_{ij} corresponding to the J_{ij} by $a_{ij} = \mu(J_{ij})$, and let μ_n be the measure on J defined by

$$\mu_n = \sum_{i,j=1}^{2^n} a_{ij} 2^n m_{ij}.$$

The factor 2^n in each term is required to make each $2^n m_{ij}$ into a probability measure and thus μ_n is an approximation to μ that agrees with it on the σ-algebra $\mathscr{F}_n$ generated by the intervals I_i. We will show that if T_n is the pseudointegral operator generated by μ_n, then

(a) $T_n = \sum_{i,j=1}^{2^n} a_{ij} 2^n E_{ij}$,

(b) T_n is a convex combination of composition operators,

(c) $\|T_n\| \leqslant \|T\|$, and

(d) the sequence (T_n) converges to T in the weak operator topology.

The equation (a) is immediate from the definition of μ_n. To see that (b) holds observe that for each i,

$$\sum_{j=1}^{2^n} a_{ij} 2^n = \sum_{j=1}^{2^n} \mu(J_{ij}) 2^n = \mu(I_i \times I) 2^n = 1.$$

The last equality follows because μ_x is a probability measure for each x. Therefore the matrix $(a_{ij}2^n)$ is a stochastic matrix and is consequently a convex combination of extreme points (b_{ij}). The extreme points of the stochastic matrices are 0,1 matrices, so if (b_{ij}) is such a matrix, then for each row i there is a unique column $j(i)$ containing a one. Put

$$\nu = \sum_{i,j=1}^{2^n} b_{ij} 2^n m_{ij} = \sum_{i=1}^{2^n} m_{i,j(i)}.$$

The corresponding operator T_ν is a composition operator since ν_x is a point mass for each x. The convex combination of matrices (b_{ij}) that yields $(a_{ij}2^n)$ corresponds to a convex combination of operators of the form T_v that yields T_n. Thus T_n is a convex combination of composition operators.

To see that (c) holds note first that since the E_{ij} form a system of matrix units, (a) above implies $\|T_n\| = \|(a_{ij}2^n)\|$, where $(a_{ij}2^n)$ acts on $\mathbb{C}^{2^n}$. Let $e_1, e_2, \ldots, e_{2^n}$ be the standard orthonormal basis for $\mathbb{C}^{2^n}$, where the entries of e_i are all 0 except for a 1 in position i. Define $V : \mathbb{C}^{2^n} \to L^2$ by putting $Ve_i = 2^{n/2}\chi_{I_i}$ for $1 \leqslant i \leqslant 2^n$ and extending linearly. Then V is an isometry from C^{2^n} onto the subspace of L^2 spanned by the χ_{I_i}, i.e. the $\mathscr{F}_n$-measurable functions in L^2. Also, for each $f \in L^2$, $V^*f = \sum_{i=1}^{2^n} c_i e_i$, where $c_i = \langle f, 2^n\chi_{I_i}\rangle$. Observe that by equation (1)

$$\langle V^*TVe_j, e_i\rangle = 2^n\langle T\chi_{I_j}, \chi_{I_i}\rangle = 2^n a_{ij},$$

and thus

$$\|T_n\| = \|(2^n a_{ij})\| = \|V^*TV\| \leqslant \|T\|.$$

Finally, take arbitrary unit vectors f and g in L^2, and consider $\langle (T - T_n)f, g\rangle$. For arbitrary n define V as in the preceding paragraph and put $P_n = VV^*$. The projections P_n converge strongly to the identity operator, so for given $\varepsilon > 0$ and sufficiently large n the lengths of $f - P_nf$ and $g - P_ng$ will be less than ε. Equation (a) above implies that the range of P_n is invariant under T_n, and thus equation(3) implies

$$\langle (T - T_n)P_nf, P_ng\rangle = 0.$$

Thus

$$\langle (T_n - T)f, g\rangle = \langle (T_n - T)(f - P_nf), P_kg\rangle + \langle (T_n - T)f, g - P_ng\rangle,$$

and each of the two terms on the right hand side has an absolute value less than $2\|T\|\varepsilon$. Hence T is the weak operator topology limit of the T_n.

Consider the case where $X = [0,1]$ and m is Lebesgue measure on $[0,1/2]$ plus a point mass of $1/2$ at 1. If C_ϕ is a composition operator on $L^2(m)$, then $\phi(1) = 1$, and consequently $\langle C_\phi\chi_{[0,1/2]}, \chi_{\{1\}}\rangle = 0$. Let T be the stochastic operator induced by the family of measures where μ_x is twice Lebesgue measure on $[0,1/2]$. Then $\langle T\chi_{[0,1/2]}, \chi_{\{1\}}\rangle = 1/2$, but for any convex combination S of composition operators $\langle S\chi_{[0,1/2]}, \chi_{\{1\}}\rangle = 0$. Thus T can not be a limit of such convex combinations.

Acknowledgement. We are indebted to the referee for the reference to Iwanik's paper.

REFERENCES

[1] W. ARVESON, *Operator algebras and invariant subspaces*, Ann. Math., (2) **100** (1974), 433–532. **MR**0365167 (51 #1420)

[2] A. IWANIK, *Pointwise induced operators on L_p-spaces*, Proc. Amer Math Soc., **58** (1976), 173–178. **MR**0412883 (54 #1004)

[3] R.R. PHELPS, *Extreme positive operators and homomorphisms*, Trans. Amer. Math. Soc., **108** (1963), 265–274. **MR**0156224 f27 #6153.

[4] R. SIKORSKI, *On the inducing of homomorphisms by mappings*, Fund. Math., **36**, (1949), 7–22. **MR**0031535 (11,166a)

[5] R. SIKORSKI, *Boolean algebras*, Second edition, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 25, Springer-Verlag New York Inc., New York, 1969. **MR**0242724 (39 #4053)

[6] A. SOUROUR, *Pseudo-Integral Operators*, Trans. Amer. Math. Soc., **253** (1979), 339–363. **MR**0536952 (80k:47056)

(Received January 18, 2008)

Don Hadwin
Department of Mathematics and Statistics
University of New Hampshire
Durham, NH 03824
e-mail: don@math.unh.edu

Eric Nordgren
Department of Mathematics and Statistics
University of New Hampshire
Durham, NH 03824
e-mail: ean@math.unh.edu

Heydar Radjavi
Department of Pure Mathematics
University of Waterloo
Waterloo, Ontario
Canada N2L 3G1
e-mail: hradjavi@uwaterloo.ca