CONVEX COMBINATIONS OF MARKOV TRANSITION FUNCTIONS[1]

S. R. FOGUEL

1. **Introduction.** Let $(\Omega, \Sigma, \lambda)$ be a measure space with $\lambda(\Omega)=1$. Consider the set $\mathfrak{A}$ of all the operators, P, on $L_2(\Omega, \Sigma, \lambda)$ such that:

(1.1) $\|P\|=1$

(1.2) $P1 \ =1$

(1.3) if $f\in L_2$ and $f\geqq 0$ then $Pf\geqq 0$ a.e.

These operators are given by a Markov transition function for which λ is an invariant measure.

The set $\mathfrak{A}$ is a weakly closed convex set and is selfadjoint.

We shall study several notions that are related to "mixing" properties and will show:

a convex combination mixes better than its generators.

Many of the results can be phrased, and their proofs are identical, for general contractions on a Hilbert space. This will be mentioned without details to avoid repetitions.

Let us repeat the following definitions from [**1**]. For each $P\in\mathfrak{A}$:

(1.4) $K(P)=\{f: f\in L_2(\lambda), \|P^nf\|=\|P^{*n}f\|=\|f\|,\ n=1, 2, \cdots\}$.

(1.5) $H_0(P)=\{f: f\in L_2(\lambda),\quad \text{weak}\lim P^nf=0\}$.

(1.6) $H_1(P)=H_0(P)^\perp$.

It is proved in [**1**, Theorems 2 and 4] that these are subspaces and $H_1(P)\subset K(P)$ and

(1.7) $K(P)=L_2(\Omega_1, \Sigma_1, \lambda)\ \Sigma_1\subset\Sigma$ and $\cup\Sigma_1=\Omega_1$.

Except for (1.7) all these definitions and results are valid for general contractions on a Hilbert space.

Our aim will be to find how "small" is $H_1(P)$. Thus if Σ_1 is atomic the "Limit Theorem" [**1**, Theorem 8] holds and if $H_1=$constant functions, then P is strongly mixing.

The notions of ergodic and strongly mixing operators are defined in [**3**].

2. Convex combination of finitely many Markov transition functions.

THEOREM 1. *Let $P_1, \cdots, P_m$ belong to $\mathfrak{A}$ and $P=\Sigma\alpha_iP_i$ where $\alpha_i>0$ and $\Sigma\alpha_i=1$. Then*

Received by the editors November 18, 1966.

[1] The research reported in this document has been sponsored by the Office of Scientific Research, OAR under Grant AF EOAR 66-18 through the European Office of Aerospace Research, United States Air Force and by NSF-GP 3906 and NSF-GP 4033 in the University of Minnesota, Minneapolis.

$$K(P) = \bigcap_{i=1}^{m} K(P_i) \cap \bigcap_{n=1}^{\infty} \{f\colon P_i^n f = P_j^n f,\ P_i^{*n} f = P_j^{*n} f \quad \text{for all } 1 \leqq i, j \leqq m\}. \tag{2.1}$$

PROOF. It is enough to consider the case $m=2$ since a convex combination of $P_1, \cdots, P_m$ can be represented as a convex combination of P_1 and Q where Q is a convex combination of $P_2, \cdots, P_m$.

Let $f \in K(P)$ then

$$\|f\| = \|[\alpha P_1 + (1-\alpha)P_2]^n f\| \leqq \Sigma \alpha^j (1-\alpha)^{n-j} \|P_1^{h_1} P_2^{k_1} \cdots P_1^{h_n} P_2^{k_n} f\|$$
$$\leqq \Sigma \alpha^j (1-\alpha)^{n-j} \|f\| = \|f\|$$

where $h_i k_i$ assume the values 0, 1 and $\Sigma h_i = j$, $\Sigma k_i = n-j$ and the sum is taken over all such choices when j ranges between 0 and n. Thus each of the terms in the sum must have the norm $\|f\|$ (since the operators P_1 and P_2 are contractions). In particular $\|P_1^n f\| = \|P_2^n f\| = \|f\|$. Also a Hilbert space is strictly convex and equality can occur in the triangular inequality only when the terms are proportional. Thus $P_1^n f = \gamma P_2^n f$ for some $\gamma \geqq 0$ and since they have the same norms $P_1^n f = P_2^n f$. Finally if $f \in K(P)$ then $\|P^{*n} f\| = \|f\|$ and the same argument would apply to P_1^* and P_2^*.

Conversely let $P_1^n f = P_2^n f$, $P_1^{*n} f = P_2^{*n} f$ and $f \in K(P_1)$. Then

$$P_1^{h_1} P_2^{k_1} \cdots P_1^{h_n} P_2^{k_n} f = P_1^{h_1} P_2^{k_1} \cdots P_1^{h_n} P_1^{k_n} f = P_1^{h_1} P_2^{k_1} \cdots P_2^{h_n + k_n} f$$

continuing in this way we shall get $P_1^n f$ and thus $P^n f = P_1^n f$. Now since $f \in K(P_1)$, $\|P^n f\| = \|P_1^n f\| = \|f\|$. The result for P^* is analogous.

THEOREM 2. *Let $P_1 \cdots P_m \in \mathfrak{A}$ and $P = \Sigma \alpha_i P_i$ where $\alpha_i > 0$ and $\Sigma \alpha_i = 1$. Then*

$$H_1(P) = \bigcap_{i=1}^{m} H_1(P_i) \cap \bigcap_{n=1}^{\infty} \{f\colon P_i^n f = P_j^n,\ P_i^{*n} f = P_j^{*n} f \quad \text{for all } 1 \leqq i, j \leqq m\}. \tag{2.2}$$

PROOF. If $f \in H_1(P)$ then $f \in K(P)$ hence $P^n f = P_i^n f$ for every $1 \leqq i \leqq m$. By [2, Theorem 3.1] $f \in H_1(P_i)$. The converse is proved in the same way.

Note that Theorems 1 and 2 hold for any contraction in a Hilbert space.

From (2.1) follows that, under the assumptions of Theorem 1, $\Sigma_1(P) \subset \bigcap_{i=1}^m \Sigma_1(P_i)$. Hence if at least one of the fields $\Sigma_1(P_i)$ is atomic, then so is $\Sigma_1(P)$.

From (2.2) follows, under the assumptions of Theorem 2, that if at least one of the operators P_i is strongly mixing so is P. Thus if P_1 is strongly mixing and P_2 any operator in $\mathfrak{A}$ then $\alpha P_1+(1-\alpha)P_2$ is strongly mixing for any choice of $0<\alpha<1$. Thus P_2 can be approximated in norm by strongly mixing operators. Let Q be an invertible ergodic transformation in $\mathfrak{A}$. By the Second Category Theorem [3, p. 78] such transformations exist. Put $P=\frac{1}{2}(Q+Q^2)$. Now if $f\in K(P)$ then $Qf=Q^2f$ by Theorem 1 hence $f=Qf$ and f is a constant. This shows that there is at least one strongly mixing operator in $\mathfrak{A}$ and hence a dense set of $\mathfrak{A}$. Notice that we did not show the existence of strongly mixing transformations but only operators in $\mathfrak{A}$.

Let us conclude this chapter with the following remark: *Let P and P_1 belong to $\mathfrak{A}$. There exists an operator P_2 in $\mathfrak{A}$ such that P is a convex combination of P_1 and P_2 iff for some $0<\alpha<1$ $Pf\geqq\alpha P_1f$ for every $f\geqq 0$.*

Clearly the condition is necessary. Now if $Pf\geqq\alpha P_1 f$ for every $f\geqq 0$ define $P_2f=(1-\alpha)^{-1}(Pf-P_1f)$. Then P_2 satisfies (1.2) and (1.3). In order to prove (1.1) it is enough to observe that P_2 is a contraction on L_∞ and if $f=\Sigma c_i 1_{A_i}$, where A_i are disjoint sets and 1_{A_i} denote their characteristic functions, then

$$\int |P_2f|\,d\lambda \leqq \Sigma |c_i| (1-\alpha)^{-1}\int (P-\alpha P_1)1_{A_i}d\lambda = \Sigma|c_i|\lambda(A_i)$$

since

$$\begin{aligned}\int (P-\alpha P_1)1_{A_i}d\lambda &= \langle (P-\alpha P_1)1_{A_i}, 1\rangle = \langle 1_{A_i}, (P^*-\alpha P_1^*)1\rangle \\ &= (1-\alpha)\langle 1_{A_i}, 1\rangle = (1-\alpha)\lambda(A_i),\end{aligned}$$

where $\langle f, g\rangle$ is the inner product of f and g.

3. **Integral averages.** Following Choquet's theory let us consider, throughout this chapter, the following setup:

(3.1) *Let μ be a regular positive measure, of total mass* 1, *defined on the Borel subsets of $\mathfrak{A}$ with its weak operator topology. Put*

$$Q=\int_{\mathfrak{A}} P\mu(dP).$$

The operator Q is defined by $\langle Qf, g\rangle=\int_{\mathfrak{A}}\langle Pf, g\rangle\mu(dP)$. The integral exists since for every pair of vectors f, g the function $\phi(P)=\langle Pf, g\rangle$ is continuous in the weak operator topology. Thus $\int_{\mathfrak{A}}\langle Pf, g\rangle\mu(dP)$ defines a bilinear form and hence is equal to $\langle Qf, g\rangle$ for some operator Q. It is easy to check that Q belongs to $\mathfrak{A}$.

Let us consider all the open subsets of $\mathfrak{A}$ on which μ vanishes. Since μ is regular μ vanishes also on the union of all these sets. Denote by $\mathfrak{B}$ (support of μ) the complement of this set. Thus

(3.2) $P_0 \in \mathfrak{B}$ *iff* μ *does not vanish on any neighborhood of* P_0.

THEOREM 3. *Given* Q *by* (3.1) *then*

$$K(Q) = \bigcap_{P \in \mathfrak{B}} K(P) \cap \bigcap_{P \in \mathfrak{B}} \{f: P^n f = Q^n f,\ P^{*n} f = Q^{*n} f \text{ for all } n\}. \tag{3.3}$$

PROOF. Let $f \in K(Q)$ and $P_0 \in \mathfrak{B}$ and let $\mathfrak{D}$ be any weak neighborhood of P_0. Put

$$Q = \mu(\mathfrak{D})\left(\mu(\mathfrak{D})^{-1} \int_{\mathfrak{D}} P\mu(dP)\right) + \mu(\mathfrak{B} - \mathfrak{D})\left(\mu(\mathfrak{B} - \mathfrak{D})^{-1} \int_{\mathfrak{B}-\mathfrak{D}} P\mu(dP)\right).$$

Then, by Theorem 1, $(\mu(\mathfrak{D})^{-1}\int_{\mathfrak{D}} P\mu(dP))(f) = Qf$. Or, for any $g \in L_2$, $\mu(\mathfrak{D})^{-1}\int_{\mathfrak{D}}\langle Pf, g\rangle\mu(dP) = \langle Qf, g\rangle$. If for some g $\langle P_0 f, g\rangle \neq \langle Qf, g\rangle$ say $\langle P_0 f, g\rangle < \langle Qf, g\rangle$ then taking for $\mathfrak{D}$ the open set $\{P: \langle Pf, g\rangle < \langle Qf, g\rangle\}$ we shall get a contradiction. Thus for every $P \in \mathfrak{B}$, $Pf = Qf$. Now $P^n f = P^{n-1} Qf$ and $Qf \in K(Q)$ too and by an induction argument $P^n f = Q^n f$. The argument for P^* is analogous.

Conversely, if f belongs to the right side of (3.3) then take any P in $\mathfrak{B}$ and $\|Q^n f\| = \|P^n f\| = \|f\|$, since $f \in K(P)$.

REMARK. Theorem 3 can be viewed as a necessary condition for an element P of $\mathfrak{A}$ to belong to the support of any representation of the type (3.1).

THEOREM 4. *Given* Q *by* 3.1 *then*

$$H_1(Q) = K(Q) \cap \bigcap_{P \in \mathfrak{B}} H_1(P). \tag{3.4}$$

The proof is identical to the proof of Theorem 2.

4. **Semigroup of contractions.** Let us conclude this note with a study of convergence of iterates of the resolvent of a semigroup of contractions. The situation is somewhat similar to the one studied in Chapter 3 but much stronger results are valid.

Let P_t be a strongly continuous semigroup of contractions in the Hilbert space H. Let $R_\lambda = \int_0^\infty e^{-\lambda t} P_t dt$, $\lambda > 0$. Thus R_λ, $\lambda > 0$, is the resolvent of the infinitesimal generator, A, of P_t at the point λ. Let U_t be the strong dilation of the semigroup see [**4**, Theorem IV]. Then U_t is a strongly continuous semigroup of unitary operators. Let the infinitesimal generator of U_t be iB. Then B is selfadjoint [**5**, p. 385]. Thus

$$\int_0^\infty e^{-\lambda t} U_t dt = (\lambda - iB)^{-1} \qquad \lambda > 0.$$

The spectrum of $\lambda(\lambda - iB)^{-1}$ is included in $\{\lambda(\lambda - it)^{-1}: t \text{ is real}\}$. This set is inside the unit circle and touches the circumference of the unit circle at the point 1 only. Now $\lambda(\lambda - iB)^{-1}$ is a normal operator and, from the Spectral Theorem and the above description of the spectrum, follows that $(\lambda(\lambda - iB)^{-1})^n f$ is strongly convergent for every f in the larger space where the U_t are defined.

THEOREM 5. *Let* $R_\lambda = \int_0^\infty e^{-\lambda t} P_t dt$, $\lambda > 0$. *Then* $\lim(\lambda R_\lambda)^n f =$ *projection of* f *on the set* $\{g: P_t g = g \quad \text{for all } t > 0\}$.

PROOF. Let $L = \{g: P_t g = g \quad \text{for all } t\}$ and $f = f_1 + f_2$ where $f_1 \in L$ and $f_2 \perp L$. Clearly $\lambda R_\lambda f_1 = f_1$ and we shall consider f_2 only. Now $\|P_t\| \leqq 1$ and thus $P_t g = g$ if and only if $P_t^* g = g$ or $L^\perp$ is invariant under P_t. Now

$$(\lambda R_\lambda)^n = \lambda^n \int_0^\infty \cdots \int_0^\infty e^{-\lambda(t_1 + \cdots + t_n)} P_{t_1 + \cdots + t_n} dt_1 \cdots dt_n$$

and is equal to the projection of $[\lambda(\lambda - iB)^{-1}]^n$ on H. Thus $(\lambda R_\lambda)^n f_2$ is strongly convergent too. Let its limit be h. Then $\lambda R_\lambda h = h$. Hence h belongs to the domain of definition of A and $(A - \lambda)h = \lambda(A - \lambda)R_\lambda h = \lambda h$ or $Ah = 0$. Thus $h \in L$ but f_2 and $(\lambda R_\lambda)^n f_2$ belong to L. Therefore $h = 0$.

REFERENCES

1. S. R. Foguel, *The ergodic theorem for Markov processes*, Israel J. Math. **4** (1966), 11–22.

2. ———, *Invariant subspaces of a measure preserving transformation*, Israel J. Math. **2** (1964), 198–200.

3. P. R. Halmos, *Lectures on ergodic theory*, The Mathematical Society of Japan, 1956.

4. B. Sz.-Nagy, *Prolongements des transformations de l'espace de Hilbert qui sortent de cet espace* (*Appendice au livre Leçons d'analyse fonctionelle*), Akadémiai Kiadó, Budapest, 1955.

5. F. Riesz and B. Sz.-Nagy, *Functional analysis*, Ungar, New York, 1955.

THE HEBREW UNIVERSITY OF JERUSALEM AND
UNIVERSITY OF MINNESOTA, MINNEAPOLIS