Pacific Journal of Mathematics

# WEAK AND STRONG CONVERGENCE FOR MARKOV PROCESSES

SHAUL FOGUEL

Vol. 10, No. 4 December 1960

# WEAK AND STRONG CONVERGENCE FOR MARKOV PROCESSES

S. R. FOGUEL

**1. Introduction.** Let $(\Omega, \Sigma, P)$ be a probability space and $x_t(\omega)$ a Markov process defined on it. For every Borel set on the real line $P_t(\omega, A)$ is the conditional probability that $x_t \in A$ given $x_0$. The purpose of this paper is to study the limiting behavior, of the family of functions, $p_t(\omega, A)$, for $t \to \infty$ and $A$ fixed.

In § 3 we investigate conditions for the weak convergence, in the sense of $L_2(\Omega, \Sigma, P)$, of $p_t(\omega, A)$. The classical result on Markov processes, as presented in [2] p. 353, is generalized to functions $x_t(\omega)$ with nondiscrete ranges. Under the additional assumption of existence of finite stationary measures.

It should be noted that

$$p_{ij}^{(n)} = \frac{(p_n(\omega, \{j\}), \chi_{x_0} = i)}{P(x_0 = i)}$$

where the parenthesis stand for scalar product and $\chi_{x_0} = i$ is the characteristic function of the set $x_0(\omega) = i$. Thus weak convergence of $p_n(\omega, \{j\})$ implies ordinary convergence of $p_{ij}^{(n)}$.

In § 4 the strong convergence in $L_2(\Omega, \Sigma, P)$ is studied. Our results are similiar to Theorem 11 of [4] though the exact relation between the two theories is not clear to us.

The paper deals with real processes and $L_2$ is the real Hilbert space.

Throughout the paper a weak form of the definition of Markov processes is used. We do not assume any of the regularity properties which are usually imposed.

**2. Notation and general background.** Let $x_t(\omega)$ be a set of measurable functions, defined on $\Omega$, where $t$ runs over $[0, \infty)$ or the positive integers. This set of functions, will be called a Markov process if whenever $t_1 < t_2 < t_3$ then *conditional probability that* $x_{t_3} \in A$ *given* $x_{t_1}$ *and* $x_{t_2}$, *is equal to the conditional probability that* $x_{t_3} \in A$ *given* $x_{t_2}$.

In order to simplify this condition let us observe the following:

*If* $\Sigma_1$ *is a sub* $\sigma$ *algebra of* $\Sigma$ *and* $f \in L_2(\Omega, \Sigma, P)$ *then the conditional expectation of* $f$ *with respect to* $\Sigma_1$ *is equal* a.e. *to* $E_1 f$ *where* $E_1$ *is the self adjoint projection on the subspace of* $L_2$ *generated by characteristic functions of sets in* $\Sigma_1$.

With the Markov process, $x_t(\omega)$, associate a collection of subspaces,

Received November, 23, 1959, and in revised from March 2, 1960.

$B_t$ of $L_2(\Omega, \Sigma, P)$, where $B_t$ is the closed subspace spanned by characteristic functions of sets of the form $x_t^{-1}(A)$, $A$ a Borel set on the line. Let $E_t$ be the self adjoint projection on $B_t$.

THEOREM 2.1. *If the set of functions* $x_t(\omega)$ *is a Markov process, then*

$$E_{t_1}E_{t_2}E_{t_3} = E_{t_1}E_{t_3} \quad \textit{for} \quad t_1 < t_2 < t_3 . \tag{2.1}$$

*Proof.* Let $t_1 < t_2 < t_3$. If $g \in B_{t_3}$ then $g - E_{t_2}g$ is orthogonal to $B_{t_1}$. Thus

$$E_{t_1}(E_{t_3} - E_{t_2}E_{t_3}) = 0 .$$

DEFINITION. *A Collection of spaces* $B_t \subset L_2(\Omega)$, *is a Markov class if equation* 2.1 *holds.*

From the above definition follows:

THEOREM 2.2. *Let* $B_t$ *be a Markov class. If* $f \in B_{t_1} \cap B_{t_2}$ *and* $t_1 < t < t_2$ *then* $f \in B_t$.

*Proof.* If $f = E_{t_1}f = E_{t_2}f$ then

$$\begin{aligned} \| E_t f \|^2 &= (E_t f, f) = (E_t E_{t_2} f, E_{t_1} f) = (E_{t_1} E_t E_{t_2} f, f) \\ &= \| (E_{t_1} E_{t_2} f, f) = \| f \|^2 . \end{aligned}$$

Thus $f = E_t f \in B_t$ .

DEFINITION. *A Markov process is called stationary if*

$$P(x_{t_1+\alpha} \in A_1 \cap x_{t_2+\alpha} \in A_2) = P(x_{t_1} \in A_1 \cap x_{t_2} \in A_2) . \tag{2.2}$$

In particular for a stationary Markov process

$$P(x_t \in A) = P(x_0 \in A) . \tag{2.3}$$

Let $T_t$ be the transformation from $B_0$ to $B_t$ defined for characteristic functions in $B_0$ by

$$T_t \chi_{x_0 \in A} = \chi_{x_t \in A} . \tag{2.4}$$

LEMMA 2.4. *Let* $x_t(\omega)$ *be a stationary Markov process. The transformation* $T_t$ *can be extended in a unique way to all of* $B_0$ *such that*

(a) $$\| T_t x \| = \| x \| \quad \textit{if} \quad x \in B_0$$

(b) $$T_t B_0 = B_t$$

(c) $$(T_{t_1+\alpha} x, T_{t_2+\alpha} y) = (T_{t_1} x, T_{t_2} y)$$

*for every* $x \in B_0$, $y \in B_0$ *and* $\alpha > 0$.

*Proof.* In order to consider $T_t$ as a transformation in $B_0$ we have to show that:

If $A_1$ and $A_2$ are two Borel sets and $\chi_{x_0^{-1}(A_1)}$, $\chi_{x_0^{-1}(A_2)}$ differ by a set of measure zero, then

$$\chi_{x_t^{-1}(\omega)_{(A_1)}} = \chi_{x_t^{-1}(A_2)}(\omega) \quad \text{a.e.}$$

Now by assumption

$$\| \chi_{x_0^{-1}(A_1)} \| = \| \chi_{x_0^{-1}(A_2)} \| = \| \chi_{x_0^{-1}(A_1 \cap A_2)} \| .$$

But by 2.3

$$\| \chi_{x_t^{-1}(A_1)} \| = \| \chi_{x_t^{-1}(A_2)} \| = \| \chi x_{x_t^{-1}(A_1 \cap A_2)} \|$$

which means

$$\chi_{x_t^{-1}(A_1)} = \chi_{x_t^{-1}(A_2)} \quad \text{a.e.}$$

Let us extend $T_t$ to linear combinations of characteristic functions by additivity. If conditions $a$ and $c$ are satisfied for this dense set, we will be able to extend $T_t$ by continuity to all of $B_0$ and $T_t$ will satisfy $a, b$ and $c$. It is enough to show that the extension of $T_t$ to linear combinations is unique. For then $c$ follows from 2.2, and a holds because every linear combination of characteristic functions in $B_0$, can be written with disjoint characteristic functions. Let us assume, then, that there exists numbers $a_i$ and Borel sets $A_i$ such that

$$\Sigma a_i \chi_{x_0^{-1}(A_i)} = 0 \quad \text{but} \quad \Sigma a_i \chi_{x_t^{-1}(A_i)} \neq 0 .$$

Thus there are $k$ integers $i_1, \cdots, i_k$ with

$$\chi_{x_t^{-1}(B \cap A_i)} = 0 \quad \text{a.e.,} \ i \neq i_j$$

where

$$B = \bigcap_{j=1}^{k} A_{i_j}, \ P(x_t^{-1}(B)) > 0$$

and

$$\sum_{i=1}^{k} a_{i_i} \neq 0 .$$

But then, by 2.3,

$$\chi_{x_0^{-1}(B \cap A_i)} = 0 \quad \text{a.e.}$$

if $i \neq i_j$ and for $\omega \in x_0^{-1}(B)$

$$\Sigma a_i \chi_{x_0^{-1}(A_i)}(\omega) = \sum_{j=1}^{k} a_{i_j} \neq 0 .$$

This contradicts our assumption for

$$P(x_0^{-1}(B)) = P(x_t^{-1}(B)) \neq 0 .$$

REMARK. From $a$ follows that $T_t$ preserves inner products.

DEFINITION. *A Markov class is called stationary if there exist transformations $T_t$ from $B_0$ to $B_t$ satisfying a, b and c of Lemma* 2.4.

In the rest of the paper we will use the notation

$$\chi_{t,A} = \chi_{x_t^{-1}(A)}$$

## 3. Weak convergence.

The main tool in this section is:

LEMMA 3.1. *Let $B_t$ be a stationary Markov class. If $\bigcap_{n=0}^{\infty} B_n = 0$ then*

$$\text{weak lim } T_n x_0 = 0$$

*for every $x_0 \in B_0$.*

For the proof we need the following.

LEMMA 3.2. *Let $B_t$ be a stationary Markow class, and $\bigcap_{n=0}^{\infty} B_n = 0$. If for some subsequence $n_i$, of the integers,*

$$\text{weak lim } T_{n_i} x_0 = x \neq 0$$

then

$$x = E_0 x + \sum_{n=1}^{\infty} (E_n - E_{n-1}) x$$

*and the terms of the sum are mutually orthogonal.*

*Proof.* Let $n < m$ then

$$(*) \qquad E_n E_m x = \text{weak} \lim_{i \to \infty} E_n E_m T_{n_i} x_0 = \text{weak} \lim_{i \to \infty} E_n T_{n_i} x_0 = E_n x$$

by Equation 2.1 Thus

$$(**) \qquad E_n(E_m x - E_{m-1} x) = E_n x - E_n x = 0 .$$

Now

$$\| E_N x \|^2 = \| E_0 x + \sum_{n=1}^{N} (E_n - E_{n-1}) x \|^2 = \| E_0 x \|^2 + \sum_{n=1}^{N} \| (E_n - E_{n-1}) x \|^2$$

hence the sum converges. Let

$$y = E_0 x + \sum_{n=1}^{\infty} (E_n - E_{n-1}) x \ .$$

If $z = E_n z \in B_n$ then by (**)

$$(y, z) = (E_n y, z) = (E_n x, z) = (x, z) \ .$$

Also if $z$ is orthogonal to all the spaces $B_n$ then

$$(y, z) = (x, z) = 0 \ .$$

Thus $y = x$.

LEMMA 3.3. *Under the same conditions, there exists a subsequence* $n'_i$, *of* $n_i$, *such that if* $z_n \in B_0$ *is defined by*

$$(***) \qquad T_n z_n = E_n x / \| x \|$$

*Then*

$$\text{weak lim } z_{n'_i} = 0 \ .$$

*Proof*. Let $z_{n'_i}$ converges weakly to $z$. Such subsequence exists because a Hilbert space is weakly sequentially compact. Now $z \in B_0$, we shall prove that $z \in B_k$, for all $k$, and thus $z = 0$. Now, by equations (***) and 2.2

$$(T_k z_{n+k}, z_n) = (T_{n+k} z_{n+k}, T_n z_n) = (E_{n+k} x / \| x \|, E_n x / \| x \|) \xrightarrow[n \to \infty]{} 1 \ .$$

Hence

$$\| T_k z_{n+k} - z_n \|^2 \leqq 2 - 2(T_k z_{n+k}, z_n) \to 0 \ .$$

If $u \in L_2(\Omega)$ then

$$(T_k z_{n'_i+k}, u) = ((T_k z_{n'_i+k} - z_{n'_i}), u) + (z_{n'_i}, u) \to (z, u)$$

or

$$\text{weak lim } T_k z_{n'_i+k} = z$$

and by Hahn Banach Theorem $z \in B_k$.

*Proof of Lemma* 3.1. It is enough to show that for any subsequence $n_i$, there exists a subsequence $n'_i$, of $n_i$, such that

$$\text{weak lim } T_{n'_i} x_0 = 0 \ .$$

We may assume that $T_{n_i}x_0$ converges weakly to $x$. Let $n_i'$ be chosen by Lemma 4.3. Then

$$\begin{aligned} 0 &= \lim_{i\to\infty} (z_{n_i'}, x_0) = \lim_{i\to\infty} (T_{n_i'}z_{n_i'}, T_{n_i'}x_0) \\ &= \lim_{i\to\infty} (E_{n_i'}x/\|x\|, T_{n_i'}x_0) = \|x\| \end{aligned}$$

For $E_{n_i}x$ tends strongly to $x$, by Lemma 3.2, and by assumption $T_{n_i}x_0$ converges weakly to $x$.

COROLLARY. *Let $x_t$ be a stationary Markov process. If $\bigcap_{n=0}^{\infty} B_n = \{1\}$ then*

$$\text{weak}\lim \chi_{n,A} = \|\chi_{0,A}\|^2 1 \ .$$

*Proof.* The Markov class $B_t - \{1\}$ satisfies the conditions of Lemma 3.1, hence

$$\text{weak}\lim \chi_{n,A} - \|\chi_{n,A}\|^2 1 = 0 \ .$$

In the rest of this section let $x_t$ be a given stationary Markov process. Let

$$C_0 = \bigcap_{n=0}^{\infty} B_n \ .$$

By Theorem 2.2

$$C_0 = \bigcap_{i=0}^{\infty} B_{t_i}$$

wherever $t_0 = 0$ and $t_i \to \infty$. Let

$$C_m = \bigcap_{n=m}^{\infty} B_n \quad \text{and} \quad D_m = B_m - C_m \ .$$

REMARK. $\{1\}$ stands for the space of constants. Also if $B$ and $C$ are subspaces $B - C$ is the orthogonal complement of $C$ in $B$.

LEMMA 3.4. *For every integer $n$*

$$T_nC_0 = C_n \ , \quad T_nD_0 = D_n$$

*and*

$$C_n \subset C_{n+1} \ .$$

*Proof.* Let $x = T_mx_0$. The vector $x$ belongs to $C_m$, if and only if, for every integer $k$ there exists a vector $x_k \in B_0$ such that

$$x = T_{m+k}x_k \ .$$

But then

$$\| x \|^2 = (T_{m+k}x_k, T_m x_0) = (T_k x_k, x_0)$$

and $\| x_0 \| = \| x \| = \| T_k x_k \|$. Hence $x_0 = T_k x_k$ and $x_0 \in B_k$ for all $k$: $x_0 \in C_0$. Now $y \in D_m$ if and only if $y = T_m y_0$ and

$$(y, x) = 0 \quad \text{if} \quad x \in C_m .$$

This is equivalent to

$$(T_m y_0, T_m x_0) = 0 \quad \text{if} \quad x_0 \in C_0, \quad \text{or} \quad (y_0, x_0) = 0 .$$

Thus $y \in D_m$ if and only if $y_0 \in D_0$.

LEMMA 3.5. *Both $C_m$ and $D_m$ are stationary Markov classes.*

*Proof.* The class $C_m$ is Markov because $C_m \subset C_{m+1}$. Now let $F_m$ be the projection on $C_m$ and $G_m$ the projection on $D_m$. Then

$$G_m = E_m(I - F_m) .$$

If $n \geqq m$ then $E_n F_m = F_m$ hence $E_n$ and $I - F_m$ commute. Let $m_1 < m_2 < m_3$ then

$$\begin{aligned} G_{m_1}G_{m_2}G_{m_3} &= E_{m_1}(I - F_{m_1})E_{m_2}(I - F_{m_2})E_{m_3}(I - F_{m_3}) \\ &= E_{m_1}E_{m_2}E_{m_3}(I - F_{m_1})(I - F_{m_2})(I - F_{m_3}) \\ &= E_{m_1}E_{m_3}(I - F_{m_1})(I - F_{m_3}) = G_{m_1}G_{m_3} . \end{aligned}$$

We used Equation 2.1 and the fact that $I - F_m$ decreases with $m$.

THEOREM 3.6. *If $x \in D_0$ then $T_n x$ tends weakly to zero.*

*Proof.* The Markov class $D_m$ satisfies the conditions of Theorem 3.1 for

$$\bigcap_{n=0}^{\infty} D_m \subset D_0 \cap \bigcap_{n=0}^{\infty} B_n = 0 .$$

It remains to study the monotone stationary Markov class $C_m$. Define

$$C_{-m} = T_m^{-1} C_0, \quad H = \bigcap_{m=1}^{\infty} C_{-m} .$$

REMARK. If $C_0$ is finite dimensional then $C_0 \subset C_m$ and both have same dimension:

$$C_0 = C_m \quad \text{and} \quad H = C_0 .$$

THEOREM 3.7. *If $x \in C_0$ is orthogonal to $H$ then*

$$\text{weak} \lim_{n\to\infty} T_n x = 0$$

*Proof.* If $m > k$ then $C_{-m} \subset C_{-k}$: if $x \in C_{-m}$ then $T_m x \in C_0$. Let $y_0 \in C_0$ and $T_{m-k} y_0 = T_m x$ then

$$\| T_m x \|^2 = (T_m x, T_{m-k} y_0) = (T_k x, y_0)$$

Thus $y_0 = T_k x \in C_0$.

Now if $F_{-m}$ is the projection of $C_0$ on $C_{-m}$ then for each $x \in C_0$ $F_{-m}x$ converges to the projection of $x$ on $H$ (See [3] p. 266). Thus

$$x = \lim(I - F_{-m})x$$

or $x$ is the limit of vectors orthogonal to $C_{-m}$.

Let us prove that

$$\text{weak} \lim_{n\to\infty} T_n x = 0$$

if $x$ is orthogonal to $C_{-m}$, and because this is a dense set the theorem will follow.

The vector $x$ is orthogonal to $C_{-m}$, and hence to $C_{-m-p}$ for all $p$. Now

$$(T_{rm+d}x, T_d x) = (T_{rm}x, x)$$

but $x \in C_0$ and for some $y_0 \in C_0$, $x = T_{rm} y_0$ thus

$$(T_{rm+d}x, T_d x) = (T_{rm}x, T_{rm}y_0) = (x, y_0) = 0$$

for $y_0 \in C_{-rm}$. Thus the $m$ sequences

$$\{T_{rm+d}x\} d = 0, 1, \cdots, m-1$$

consist of mutually orthogonal elements and thus converge weakly to zero.

It remains to study $T$ on $H$.

THEOREM 3.8. *On the space $H$, $T$ is a unitary operator and $T_n = T^n$.*

*Proof.* If $x \in H$ then $T_n x \in C_0$ for all $n$ and it is possible to take $T_m(T_n x)$. But then

$$(T_{n+m}x, T_n(T_m x)) = \| T_m x \|^2$$

thus $T_{n+m}x = T_n(T_m x)$, or $T_n x = T^n x$. Thus if $y = Tx \in C_0$ then $T_n y = T_{n+1}x \in C_0$ and $y \in H$.

In order to show that $T$ is unitary we have to show that it is onto. Let $x \in H$ then for some $x_0 \in C_0$ $Tx_0 = x$. But then $T_n x_0 = T_{n-1}x \in C_0$ and $x_0 \in H$.

In general the powers of a unitary operator do not converge. However the operator $T$ has some special properties.

LEMMA 3.9. *If* $f \in L_2(\Omega)$ *and* $f \in H$ *then* $\chi_{f^{-1}(A)} \in H$ *for every Borel set* $A$.

*Proof.* In order to prove this we have to go back to the definitions of $H$ and $T$. Now, if $f \in B_n$ and $A$ is a Borel set, then $f^{-1}(A) = x_n^{-1}(A_n)$ for some $A_n$ and thus $\chi_{f^{-1}(A)} \in B_n$. Thus $f \in C_0$ implies that $\chi_{f^{-1}(A)} \in C_0$. But $f \in H$ so $T_n f \in H$. The Lemma will be proved if we show that

$$T_n \chi_{f^{-1}(A)} = \chi_{(T_n f)^{-1}(A)} \quad \text{a.e.}$$

If $M \leqq f \leqq N$ then $M \leqq T_n f \leqq N$, thus it is enough to prove the above equation under the assumption that $A$ is a bounded set and $f$ a bounded function. If $f$ is bounded (hence $T_n f$ is bounded also) it defines a self adjoint operator on $L_2(\Omega)$,: the multiplication operator. Thus as an operator

$$f = \int \lambda \chi_{f^{-1}}(d\lambda)$$

$$T_n f = \int \lambda T_n \chi_{f^{-1}(d\lambda)} = \int_{(T_n f)^{-1}(d\lambda)} .$$

Now $T_n$ transforms characteristic functions to characteristic functions and $T_n \chi_{f^{-1}(A)}, \chi_{(T_n f)^{-1}(A)}$ are both the spectral measure of $T_n f$. Thus

$$T_n \chi_{f^{-1}(A)} = \chi_{(T_n f)^{-1}(A)} . \quad \text{a.e.}$$

This lemma shows that $H$ is generated by characteristic functions. Let us study the limits of $T_n x$ when $x$ is a characteristic function.

LEMMA 3.10. *Let $H$ be generated by a countable number of disjoint characteristic functions $\chi_i$. For a given $\chi_i$ there is an integer $m$: $T_m \chi_i = \chi_i$ and then*

$$T_{rm+d} \chi_i = T_d \chi_i .$$

*Proof.* For every $n$ $T_n \chi_i$ is a characteristic function, hence either $T_n \chi_i = \chi_i$ or

$$(T_n \chi_i, \chi_i) = 0 .$$

If $(T_n \chi_i, \chi_i) = 0$ for all $n$ then $(T_m \chi_i, T_n \chi_i) = (T_{m-n} \chi_i, \chi_i) = 0$ thus there exist infinitely many disjoint sets of equal measure which is impossible.

Now if for some $m$, $T_m \chi_i = \chi_i$, let $m$ be the smallest integer that this happens. Then

$$T_{rm+d}\chi_i = T^d T^{rm}\chi_i = T^d\chi_i = T_d\chi_i .$$

THEOREM 3.11. *Let $x_t$ be a stationary Markov process. If $H$ is generated by a countable collection of disjoint characteristic functions $\{\chi_i\}$ then for every $y \in B_0$ such that $(y, \chi_i) \neq 0$ for finitely many $i$'s ($y$ has a "finite" support), there exists an integer $m$ such that the $m$ sequences*

$$\{T_{km+d}y\} \qquad d = 1, 2 \cdots, m$$

*converge weakly.*

*Proof.* From Theorems 3.6 and 3.7 it follows that

$$\text{weak lim } T_n(y - \Sigma(y, \chi_i) \| \chi_i \|^{-2}\chi_i) = 0 .$$

Let $\chi_{i_1}, \chi_{i_2}, \cdots, \chi_{i_n}$ be those functions for which $(y, \chi_i) \neq 0$. Now $T^{m_j}\chi_{i_j} = \chi_{i_j}$. Choose $m$ to be the product of this $m_j$. Thus

$$T_{km+d}\chi_{i_j} = T^d\chi_{i_j} .$$

Hence

$$\begin{aligned}(3.1) \qquad \text{weak} \lim_{k\to\infty} T_{km+d}y &= \text{weak} \lim_{k\to\infty} T_{km+d}\Sigma(y, \chi_i) \| \chi_i \|^{-2}\chi_i \\ &= \Sigma(y, \chi_i) \| \chi_i \|^{-2} T^d\chi_i .\end{aligned}$$

COROLLARY 1. *Equation* 3.1 *holds if the function $x_0$ has countable range.*

This is a classical theorem see [2] p. 353.

COROLLARY 2. *If there exists a finite measure $\varphi$, on the line, such that, for some $\varepsilon > 0$, $\varphi(A) \leqq \varepsilon$ implies that*

$$E_0\chi_{n,A} \neq \chi_{n,A}$$

*for some $n$, then the space $H$ is generated by a finite number of disjoint characteristic functions. Thus an integer $m$ exists, such that Equation* 3.1 *holds for all $y \in B_0$.*

*Proof.* Let $k$ be an integer greater or equal to $\varphi(\Omega)\varepsilon$. If $\chi_{0, A_i} \in H$ $i = 1, \cdots, k$ where the $A_i$ are disjoint then

$$\varphi(\Omega) \geqq \Sigma\varphi(A_i) \geqq \min(\varphi(A_i))k$$

or $\varphi(A_{i_0}) \leqq \varphi(\Omega)/k \leqq \varepsilon$ for some $i_0$. But then, for some $n$, $\chi_{n,A_{i_0}} \notin H$ hence

$$\chi_{0,A_{i_0}} \notin H .$$

Thus there are at most $k-1$ disjoint characteristic functions that generate $H$.

REMARK. This last corollary is similiar to Doeblin's condition as given in [1] page 192.

4. **Strong convergence.** Throughout this section we assume:

4.1. *There exists a real number $t_0 > 0$ such that the space $B_0 \cap B_{t_0}$ is finite dimensional and there is a positive angle between $B_{t_0} - B_0 \cap B_{t_0}$ and $B_0 - B_0 \cap B_{t_0}$.*

Two subspaces, $B^*$ and $B^{**}$, are said to have a positive angle between them if

$$\sup \{(b^*, b^{**}) \mid \| b^* \| = \| b^{**} \| = 1 \text{ and } b^* \in B^*, b^{**} \in B^{**}\} < 1 .$$

CONDITION 4.1. Is equivalent to each of the following.

(a) The point 1 is not in the essential spectrum of $E_0 E_{t_0} E_0$(or $E_{t_0} E_0 E_{t_0}$).

(b) The operator $E_0 E_{t_0} E_0$(or $E_{t_0} E_0 E_{t_0}$) is quasi compact.

(c) The operator $E_0 E_{t_0} E_0$(or $E_{t_0} E_0 E_{t_0}$) is a sum of a compact operator and an operator of norm less than 1.

(d) The norm of $E_0$ restricted to $B_{t_0} - B_0 \cap B_{t_0}$ is less than one.

LEMMA 4.1. *If $t > t_0$ then Condition* 4.1 *is satisfied when $B_{t_0}$ is replaced by $B_t$.*

*Proof.* Let us use the form given in $c$ for 4.1. Now

$$E_t E_0 E_t = E_t (E_{t_0} E_0 E_{t_0}) E_t$$

by Equation 2.1, hence it is a sum of a compact and an operator of norm less than 1.

Now from Theorem 2.2 it follows that $B_0 \cap B_t$ decreases with $t$. Let $t_1$ be such that

$$\dim (B_0 \cap B_{t_1}) \leqq \dim (B_0 \cap B_t) \text{ for all } t.$$

It is easy to see that $B_0 \cap B_{t_1}$ is generated by a finite number of disjoint characteristic functions. Let them be $\chi_1, \cdots, \chi_k$, thus

$$B_0 \cap B_{t_1} = B_0 \cap B_t = \operatorname{span} \{\chi_1, \cdots, \chi_k\} \qquad t > t_1 .$$

because by Theorem 2.2

$$B_0 \cap B_{t_1} \supset B_0 \cap B_t$$

and they have the same dimension.

LEMMA 4.2. *If $t > 0$ then*

$$T_t(B_0 \cap B_{t_1}) = B_0 \cap B_{t_1}$$

*and*

$$T_t(B_0 - B_0 \cap B_{t_1}) = B_t - B_0 \cap B_{t_1} = B_t - B_0 \cap B_t .$$

*Proof.* A vector $x \in B_0 \cap B_{t_1}$, if and only if, $x \in B_0$ and $x = T_{t_1} y$ for some $y \in B_0$. But then

$$(T_t x, T_{t+t_1} y) = (x, T_{t_1} y) = \| x \|^2 = \| T_t x \|^2$$

or

$$T_t x = T_{t+t_1} y: \quad T_t x \in B_t \cap B_{t+t_1} .$$

Thus

$$T_t(B_0 \cap B_{t_1}) = B_t \cap B_{t+t_1} \supset B_0 \cap B_{t+t_1} = B_0 \cap B_{t_1}$$

by Theorem 2.2 and the remark above. This shows that

$$T_t(B_0 \cap B_{t_1}) = B_0 \cap B_{t_1} .$$

Let $x \in B_0$ be orthogonal to $B_0 \cap B_{t_1}$. If $y \in B_0 \cap B_{t_1}$, then $y = T_t z$ where $z \in B_0 \cap B_{t_1}$. Thus

$$(T_t x, y) = (T_t x, T_t z) = (x, z) = 0 .$$

THEOREM 4.3. *Let $x \in B_0$ and let $c =$ norm of $E_0$ restricted to $B_{t_1} - B_0 \cap B_{t_1}$.*
*Then $c < 1$ and*

$$\| E_0 T_t x - \sum_{i=1}^{k} (x, \chi_i) \| \chi_i \|^{-2} T_t \chi_i \| \leqq c^n \| x \| \tag{4.2}$$

*where $n$ is an integer such that $nt_1 < t$.*

*Proof.* The vector $x - \sum_{i=1}^{k} (x, \chi_i) \| \chi_i \|^{-2} \chi_i$ is orthogonal to $B_0 \cap B_{t_1}$ and hence so is

$$y = T_t x - \Sigma (x, \chi_i) \| \chi_i \|^{-2} T_t \chi_i .$$

Thus

$$\| E_0 y \| = \| E_0 E_t y \| = \| E_0 E_{t_1} E_{2t_1} \cdots E_{nt_1} y \| .$$

Now the norm $E_{jt_1}$ restricted to $B_{(j+1)t_1} - B_0 \cap B_{t_1}$ is equal to $c$ hence

$$\| E_0 y \| \leqq c^n \| y \| \leqq c_n \| x \| .$$

It becomes now interesting to study $T_t \chi_i$.

THEOREM 4.4. *For each given $t$ there is a permutation of the integer $1, 2, \cdots, k, \pi_t$, such that*

$$T_t\chi_i = \chi_{\pi_t i} .$$

*Also there exists an integer $m$ such that*

$$T_{mt}\chi_i = \chi_{(\pi_t i)^m} = \chi_i \quad \text{for all } i .$$

*Proof.* Let us use induction on $k$. Let $\chi_{i_1}, \chi_{i_2}, \cdots, \chi_{i_j}$ be a subset of $\chi_i$, $i = 1, \cdots, k$, with minimum norm: $\|\chi_{i_r}\| \leqq \|\chi_i\|$. Then $T_t\chi_{i_r}$ is a characteristic function in $B_0 \cap B_{t_1}$ with norm smaller or equal to the norm of $\chi_1, \chi_2, \cdots, \chi_k$: $T_t\chi_{i_r} \in (\chi_{i_1}, \cdots, \chi_{i_j})$.

This shows that $T_t$ maps the set $(\chi_{i_1}, \cdots, \chi_{i_k})$ into, therefore onto, itself. If $\chi_i$ is not in this set then $T_t\chi_i$ will be also, orthogonal to $\chi_{i_r}$. In the remaining set there are less than $k$ functions and by induction the first part of the theorem is proved. The second part is an easy result on permutations.

The last two theorems include the classical result on Markov processes with a finite number of states. There might be a connection to Theorem 11 of [4].

If $\dim B_0 \cap B_{t_1} = 1$ then

$$\| T_t x - (x, 1)1 \| \leqq c^n \| x \|$$

where $nt_1 < t$ and 1 is $\chi_\Omega$. This is a similiar to the case of independent functions. Let us conclude this section by studying this case. Thus let $B_1$ and $B_2$ be two subspaces of $L_2(\Omega)$ generated by characteristic functions $\chi_A$ and $\chi_{A'}$, where $A$ and $A'$ belong to some $\sigma$ subalgebras of $\Sigma$. The cosine of the angle between $B_1 - \{1\}$ and $B_2 - \{1\}$, $c$, is given by

$$(*) \qquad c = \sup\{(\Sigma a_i\chi_{A_i}, \Sigma a'_i\chi A'_i) \,|\, 1 = \Sigma a_i^2 P(A_i) = \Sigma a_i^2 P(A'_i)$$

and

$$\Sigma a_i P(A_i) = \Sigma a'_i P(A'_i) = 0\} .$$

THEOREM 4.5. *The number $c$ is smaller than*

1. $\sup |(P(A \cap A') - P(A)P(A'))P(A \cap A')^{-1}| = c_1 .$

2. $\sup |(P(A \cap A') - P(A)P(A')(P(A)P(A'))^{-1}| = c_2 .$

*Where $A$ and $A'$ belong to the $\sigma$ subalgebras generating $B_1$ and $B_2$ respectively.*

*Proof.* Let us show that $c \leqq c_1$, the other inequality is proved in a similiar way. Now let $a_i$, $a'_i$, $A_i$ and $A'_i$ satisfy the conditions of equation (*). Then

$$\sum_{i,j} a_i a'_j P(A_i \cap A_j) = \sum_{i,j} a_i a'_j (P(A_i \cap A'_j) - P(A_i)P(A'_j)) + \sum_{i,j} a_i a_j P(A_i)P(A_j) .$$

The second term is equal to zero. Thus

$$\begin{aligned} | \Sigma a_i b'_j P(A_j \cap A'_j) | &\leqq c_1 \sum_{i,j} | a_i a'_j | P(A_i \cap A'_j) \\ &\leqq c_1 \Big( \sum_{i,j} a_i^2 P(A_i \cap A'_j) \Big)^{1/2} \Big( \sum_{i,j} a'^2_j P(A_i \cap A'_j) \Big)^{1/2} \\ &= c_1 \Big( \sum_i a_i^2 P(A_i) \Big)^{1/2} \Big( \sum_j a'^2_j P(A'_j) \Big)^{1/2} = c_1 . \end{aligned}$$

A more convenient form of the conditions of Lemma 3.2 is

1. $c_1$ is the largest number for which

$$(1 + c_1)^{-1} \leqq P(A \cap A')(P(A)P(A'))^{-1} \leqq (1 - c_1)^{-1} .$$

2. $c_2$ is the largest number for which

$$1 - c_2 \leqq P(A \cap A')(P(A)P(A'))^{-1} \leqq 1 + c_2 .$$

## Bibliography

1. J. L. Doob, *Stochastic Processes*, Wiley, New York, 1953.
2. W. Feller, *An Introduction to Probability Theory*, Wiley New York, 1957.
3. F. Riesz, and B. Sz. Nagy, *Lecons D'analyse Fonctionelle*, Budapest, 1953.
4. K. Yosida, and S. Kakutani, *Operator theoretical treatment of Markoff's process and mean ergodic theorem*, Ann. Math. **42**, (1941) 188-228.

UNIVERSITY OF CALIFORNIA
BERKELEY CALIFORNIA

---

Mathematical papers intended for publication in the *Pacific Journal of Mathematics* should be typewritten (double spaced), and the author should keep a complete copy. Manuscripts may be sent to any one of the four editors. All other communications to the editors should be addressed to the managing editor, L. J. Paige at the University of California, Los Angeles 24, California.

50 reprints per author of each article are furnished free of charge; additional copies may be obtained at cost in multiples of 50.

---


The *Pacific Journal of Mathematics* is published quarterly, in March, June, September, and December. The price per volume (4 numbers) is $12.00; single issues, $3.50. Back numbers are available. Special price to individual faculty members of supporting institutions and to individual members of the American Mathematical Society: $4.00 per volume; single issues, $1.25.

Subscriptions, orders for back numbers, and changes of address should be sent to Pacific Journal of Mathematics, 2120 Oxford Street, Berkeley 4, California.

Printed at Kokusai Bunken Insatsusha (International Academic Printing Co., Ltd.), No. 6, 2-chome, Fujimi-cho, Chiyoda-ku, Tokyo, Japan.

PUBLISHED BY PACIFIC JOURNAL OF MATHEMATICS, A NON-PROFIT CORPORATION

# Pacific Journal of Mathematics

## Vol. 10, No. 4 December, 1960