*The General Temporally Discrete Markoff Process**

EBERHARD HOPF

Graduate Institute for Applied Mathematics, Indiana University, Bloomington, Indiana

CONTENTS

1. Introduction	14
2. Operator characterization of general Markoff processes.	16
3. Iteration of a Markoff operator.	20
4. The dual L of a Markoff operator.	22
5. Density operators L^* of a Markoff process	24
6. Extreme case of a point transformation.	25
7. The maximal ergodic theorem for Markoff operators Λ and L^*	27
8. Conservative and dissipative part of a general Markoff process	32
9. The conservative part CX . Invariant functions and sets	35
10. On the ergodic theorem. Markoff processes with an invariant distribution.	39
11. On the influence of the dissipative part	42
References.	45

* This paper was supported in part by an ONR research contract.

1. Introduction. A temporally discrete Markoff process in a space X with points $x, y, \dots$ and subsets, $A, B, \dots$ (belonging to a certain countably additive field of subsets of X including X itself) is a sequence of functions $P^n(x, A) \geq 0$, $n = 1, 2, \dots$, $P^n(x, X) = 1$, which are countably additive in A and which are interrelated by the recursion formula

$$P^{n+1}(x, A) = \int_x P^1(y, A) P^n(x, dy).$$

They are called transition probabilities. The process is completely characterized by the generating transition probability

$$P(x, A) = P^1(x, A).$$

A Markoff process determines two functional operators,

$$\Lambda\varphi(A) = \int_x P(x, A)\varphi(dx)$$

which is an operator from countably additive set functions φ to functions $\Lambda\varphi$ of this same sort, and the dual operator

$$Lf(x) = \int_x f(y)P(x, dy)$$

which is an operator from point functions f to functions Lf of the same sort. Both operators are linear and positive. The simplification of the formalism of a Markoff process is rendered evident by the formulae for their iterates,

$$\Lambda^n\varphi(A) = \int_x P^n(x, A)\varphi(dx)$$

and

$$L^n f(x) = \int_x f(y)P^n(x, dy).$$

Besides being positive and linear these operators have the following basic properties. The value $\varphi(X)$ for $A = X$ of a countably additive set function φ is an invariant under Λ . The constant functions are L -invariant.

Taken in full generality a Markoff process comprises the extreme singular case of a point transformation of X into X just as well as the regular case in which the operator L is completely continuous or sufficiently nearly so. The regular case in turn comprises the finite case where X consists only of finitely many points and also the continuous case in which L is a smoothing operator.

The advent of the ergodic theory of point transformations and linear operators has provided powerful tools for the exploration of the asymptotic behavior as $n \rightarrow \infty$ and of the spectral structure of a Markoff process. Thanks to the efforts

of Hostinski, Frechet, Doeblin, Doob, P. Lévy, Kryloff & Bogoliouboff, and Yosida & Kakutani the regular case of a Markoff process can be considered well explored. The work of the last four authors is predominantly operator-theoretical. Kryloff & Bogoliouboff introduced the operator L . Λ was introduced and used along with L by Yosida & Kakutani, whose work [13] has brought the theory of the regular case to high perfection. At the same time the desire was felt for a theory of the general Markoff process which would include the ergodic theory of point transformations. The first steps in this direction were taken by Doob [2], Yosida [11] and Kakutani [7]. The Markoff processes considered in these papers are general in the sense that the case of a point transformation is no longer excluded but they are restricted by the requirement that the process has an invariant distribution π , $\Lambda\pi = \pi$, $\pi \geq 0$, $\pi(X) = 1$. The results obtained are thus generalizations of the ergodic theory of point transformations with an invariant finite measure. On using the basic idea of Doob's beautiful theory of stochastic processes* Kakutani proved the individual ergodic theorem for the means

$$\frac{1}{n} \sum_0^{n-1} L^n f(x)$$

but only for bounded functions f [7]. Ten years later Doob [3] proved it for functions $f \in \mathcal{L}_\pi^p$, $p > 1$, and a still wider class but short of $\mathcal{L}_\pi^1$ which is the widest function space in which one can expect the individual ergodic theorem to hold†. The difficulty in trying to bring the ergodic theorem to the same level of generality as in the extreme singular case of a point transformation is surprising. One would expect the theorem to be true *a fortiori* as soon as smoothing is involved. However, smoothing is effective only as long as the corresponding integral converges. Hence the problem is to prove the finiteness of certain (integral) operations. In the present paper the ergodic theorem is proved within all of $\mathcal{L}_\pi^1$. Part of the proof (which makes no claim to be the simplest one) can presumably be used to prove a general ergodic theorem for ratios of sums in the most general case where the existence of an invariant measure is not assumed, but the author was unable to carry this proof through to the end. It appears that such a general ergodic theorem is a theorem about Λ rather than L . In the case where an invariant distribution exists this fact is not discernible because in this case L happens to be the density operator of another operator of type Λ .

Much as we should have preferred to use the natural approach of Doob, had

* Doob's idea is to associate with a temporally discrete stochastic process the space Ω of infinite sequences $\omega = (\dots, x_{-1}, x_0, x_1, \dots)$ of points of X and to apply the classical ergodic theory to the shift transformation $T\omega = (\dots, x_0, x_1, x_2, \dots)$. In the simplest case of completely independent chance variables this connection was first observed independently by the author and by Doob. The author was the first to point out the strong mixing property of this transformation.

† In the case of a metrically transitive π -invariant point transformation the limit of the means equals $\int_X f(x)\pi(dx)$ almost everywhere. For functions $f \geq 0$ outside of $\mathcal{L}_\pi^1$ this limit would be $+\infty$ almost everywhere.

we been able to carry it through, in this paper we employ instead the approach based upon the operator-concept of a Markoff process. This the author was able to consummate by generalizing the key inequality concealed in G. D. Birkhoff's proof of his ergodic theorem—this inequality was called “the maximal ergodic theorem” by Yosida & Kakutani [12]—to the most general Markoff process. This maximal ergodic theorem is almost as good a tool as the unproven general ergodic theorem itself if one wants to determine the ergodic structure of the process, dissipative and conservative part, L -invariant functions and invariant sets. A new theorem on invariant functions is proved. The theory is however incomplete in one point. The dissipative part has, in general, an influence on the conservative part, which it is important to know. A conjecture about this influence which would fill this gap is stated at the end of the paper.

2. Operator characterization of general Markoff processes. We denote by X a fixed abstract space with points $x, y, \dots$ and by $\mathfrak{F}$ a fixed countably additive field of subsets, $A, B, \dots$ of X . It is supposed throughout this paper that $X \in \mathfrak{F}$. Real-valued point functions in X are denoted by $f(x), g(x), \dots$ and real-valued set functions in $\mathfrak{F}$ by $\varphi(A), \psi(A), \dots$.

An f.c.a. set function $\varphi(A)$ is a $\varphi(A)$ which is finite and countably additive in $\mathfrak{F}$. An f.c.a. $\varphi(A)$ which is ≥ 0 in $\mathfrak{F}$ is a finite measure in $\mathfrak{F}$. For a finite measure π in $\mathfrak{F}$ we always have $\pi(X) < \infty$. A point function $f(x)$ in X is $\mathfrak{F}$ -measurable if and only if every set $\{x \mid f(x) > c\} \in \mathfrak{F}$.

Φ denotes the linear space of all f.c.a. set functions φ in $\mathfrak{F}$.

Φ_π denotes the linear space of all those f.c.a. φ in $\mathfrak{F}$ which are continuous with respect to a given finite measure π in $\mathfrak{F}$ (π -continuous). π -continuity of φ means that $\varphi(A) = 0$ holds for every set A such that $\pi(A) = 0$. The Radon-Nikodym theorem establishes a one-to-one correspondence between the set functions $\varphi \in \Phi_\pi$ and the π -equivalence classes $\mathbf{f}$ of $\mathfrak{F}$ -measurable point functions $f(x)$ which are π -integrable over X ,

$$\varphi(A) = \int_A f(x) \pi(dx), \quad A \in \mathfrak{F}.$$

$f(x)$ is a derivative of φ with respect to π ,

$$f(x) = \frac{\varphi(dx)}{\pi(dx)} = \frac{d\varphi}{d\pi}.$$

Note that in this context that

$$(2.1) \quad \varphi \leq \psi \text{ in } \mathfrak{F} \longleftrightarrow \frac{d\varphi}{d\pi} \leq \frac{d\psi}{d\pi} \text{ in } X$$

π -almost everywhere.

We are interested in certain single-valued operators Λ from a subspace of Φ

to a subspace of Φ , *i.e.* for which both argument and value are f.c.a. set functions in $\mathfrak{F}$.

Definition 2.1. A Markoff operator is an operator Λ with the following properties. α) $\Lambda\varphi$ is defined in a linear subspace Φ' of Φ , and $\Lambda\Phi' \subset \Phi$. β) Λ is linear, $\Lambda(a\varphi + b\psi) = a\Lambda\varphi + b\Lambda\psi$. γ) Λ is positive, $\varphi \geq 0$ in $\mathfrak{F} \rightarrow \Lambda\varphi \geq 0$ in $\mathfrak{F}$. δ) $\Lambda\varphi(X) = \varphi(X)$, $\varphi \in \Phi'$. By $\Lambda\varphi(A)$ we always mean the value at A taken by the set function $\Lambda\varphi$.

If a norm $\|\varphi\|$ is introduced in Φ by the total variation of φ over X (which is always finite for any $\varphi \in \Phi$) then δ) can be equivalently expressed by $\|\Lambda\varphi\| = \|\varphi\|$, $\varphi \in \Phi$.

Linearity and positivity imply the monotonicity of Λ ,

$$(2.2) \quad \varphi \leq \psi \text{ in } \mathfrak{F} \rightarrow \Lambda\varphi \leq \Lambda\psi \text{ in } \mathfrak{F}.$$

We wish to examine the connection between Markoff operators and Markoff processes. A Markoff process over X , $\mathfrak{F}$ is usually (Kakutani [6], Yosida [9]) understood to be a function $P(x, A)$, $x \in X$, $A \in \mathfrak{F}$, with the following properties. $\alpha')$ For each fixed $A \in \mathfrak{F}$, P is $\mathfrak{F}$ -measurable in x ; for each fixed $x \in X$, P is an f.c.a. set function of $A \in \mathfrak{F}$. $\gamma')$ $P \geq 0$. $\delta')$ $P(x, X) = 1$, $x \in X$.

These properties imply that $P \leq 1$. P is the conditional probability that the process carries a point with original position at x into the set A . The process defines an operator Λ by means of the Lebesgue-Stieltjes integral

$$(2.3) \quad \Lambda\varphi(A) = \int_X P(x, A)\varphi(dx).$$

Λ is easily seen to be a Markoff operator with domain of definition $\Phi' = \Phi$. If φ is an initial probability distribution ($\varphi \geq 0$ in $\mathfrak{F}$, $\varphi(X) = 1$) then $\Lambda\varphi$ is the probability distribution into which it is transformed by the process. The fact that $\Phi' = \Phi$, *i.e.* that (2.3) is well defined for *every* f.c.a. set function $\varphi(B)$, $B \in \mathfrak{F}$, is due to the assumption that $P(x, A)$ is defined and countably additive for *every* fixed x . This is a rather stringent assumption and it seems to us more natural to replace it by the following one.

Definition 2.2. Let $\pi(A)$ denote a finite measure in $\mathfrak{F}$. A π -conditioned Markoff process over X , $\mathfrak{F}$ is a function $P(x, A)$ with the following properties. $\alpha'')$ For each fixed $A \in \mathfrak{F}$, P is defined π -almost everywhere in X and $\mathfrak{F}$ -measurable; for any given pairwise disjoint sets $A_i \in \mathfrak{F}$ the relation of countable additivity $P(x, \sum A_i) = \sum P(x, A_i)$ holds for π -almost all $x \in X$. $\gamma'')$ $P \geq 0$. $\delta'')$ $P(x, X) = 1$.

It is easily seen that for a π -conditioned Markoff process over X , $\mathfrak{F}$ the Lebesgue-Stieltjes integral (2.3) is uniquely defined, $A \in \mathfrak{F}$, whenever the f.c.a. set function φ is π -continuous, *i.e.* whenever $\varphi \in \Phi_\pi$. Furthermore, $\Lambda\varphi(A)$ is an f.c.a. set function in $\mathfrak{F}$, $\Lambda\varphi \in \Phi$. In other words, Λ has the property α) of definition 2.1, $\Phi' = \Phi_\pi$. γ), δ) follow directly from the properties $\gamma'')$, $\delta'')$ of P , respectively.

The linearity β of Λ is a property of the formula (2.3) (there is no corresponding property β'' of P).

Consequently, a π -conditioned Markoff process over $X, \mathfrak{F}$ determines by means of (2.3) a unique Markoff operator Λ with domain of definition $\Phi' = \Phi_\pi$. We wish to complete this connection by proving the exactly converse

Theorem 2.1. To every Markoff operator Λ with domain of definition $\Phi' = \Phi_\pi$ where π is a given finite measure in $\mathfrak{F}$ there belongs a π -conditioned Markoff process $P(x, A)$ over $X, \mathfrak{F}$ which satisfies (2.3). For fixed $A \in \mathfrak{F}$, P is unique in the sense of π -equivalence.

The uniqueness follows easily from the fact that the contractions π_c of π ,

$$\pi_c(A) = \pi(CA),$$

into arbitrary sets $C \in \mathfrak{F}$ are in Φ_π . In fact, (2.3) or

$$\Lambda \pi_c(A) = \int_C P(x, A) \pi(dx)$$

must hold for every $C \in \mathfrak{F}$. This fixes P as stated.

To prove the existence of P as stated we use the following two lemmas the first of which reflects a familiar fact.

Lemma 2.1. If $\varphi_1 \leq \varphi_2 \leq \dots$ in $\mathfrak{F}$, $\varphi_n \in \Phi$, and if $\varphi = \lim \varphi_n$, taken setwise in $\mathfrak{F}$, is finite in $\mathfrak{F}$, then $\varphi \in \Phi$.

Lemma 2.2. A Markoff operator $\Lambda\varphi$ with domain of definition $\Phi' (\subset \Phi)$ has the following property of continuity with Φ' . If $\varphi_1 \leq \varphi_2 \leq \dots$, $\varphi_n \in \Phi'$, $\varphi = \lim \varphi_n \in \Phi'$, then

$$\Lambda\varphi = \lim \Lambda\varphi_n \text{ in } \mathfrak{F}.$$

To prove the second lemma note that, by virtue of the hypothesis and of α), definition 2.1, $\varphi'_n = \Lambda\varphi_n \in \Phi$ and $\Lambda\varphi \in \Phi$. Furthermore, by hypothesis and by (2.2),

$$\varphi'_1 \leq \varphi'_2 \leq \dots \leq \Lambda\varphi \text{ in } \mathfrak{F}.$$

$\Lambda\varphi(A)$ is finite in $\mathfrak{F}$ and so is $\varphi'_1(A)$. Consequently,

$$\varphi' = \lim \varphi'_n$$

is finite in $\mathfrak{F}$. On applying lemma 2.1 to the set functions φ'_n, φ' we infer that $\varphi' \in \Phi$. Hence $\Lambda\varphi - \varphi' \geq 0$ is a non-negative f.c.a. set function in $\mathfrak{F}$. The lemma asserts that it vanishes in $\mathfrak{F}$. To prove this it suffices to show that it vanishes for the set $A = X$. But in this case we infer from δ), definition 2.1, that

$$\begin{aligned} \Lambda\varphi(X) - \varphi'(X) &= \varphi(X) - \lim \varphi'_n(X) = \varphi(X) - \lim \Lambda\varphi_n(X) \\ &= \varphi(X) - \lim \varphi_n(X) = 0. \end{aligned}$$

Lemma 2.2 is thereby proved.

Proof of theorem 2.1. Consider the given operator Λ with its properties α)– δ),

$\Phi' = \Phi_\pi$. For a fixed A the contraction $\pi_c(A)$ is countably additive in C , $C \in \mathfrak{F}$. For given pairwise disjoint sets $C_i \in \mathfrak{F}$, $C = \sum C_i \in \mathfrak{F}$ we therefore have

$$\pi_c(A) = \lim \varphi_n(A), \quad \varphi_n(A) = \sum_1^n \pi_{C_i}(A)$$

for each $A \in \mathfrak{F}$. On applying lemma 2.2 to these set functions of A and on using β) we infer that

$$\Lambda \pi_c(A) = \sum_i \Lambda \pi_{C_i}(A).$$

In other words, if A is fixed $\Lambda \pi_c(A)$ is countably additive in C .

On the other hand, by virtue of α), $\Lambda \pi_c$ is an f.c.a. set function of $A \in \mathfrak{F}$ and, by virtue of γ), it is ≥ 0 . Hence and from δ),

$$\Lambda \pi_c(A) \leq \Lambda \pi_c(X) = \pi_c(X) = \pi(CX) = \pi(C).$$

Again keeping A fixed we infer from the last inequality that the f.c.a. set function of C on the left is continuous with respect to the measure $\pi(C)$. Hence it follows by the Radon-Nikodym theorem that

$$(2.4) \quad \Lambda \pi_c(A) = \int_C P(x, A) \pi(dx) = \int_X P(x, A) \pi_c(dx)$$

holds with a function $P(x) = P(x, A)$ defined in X except perhaps in a π -null set and which is $\mathfrak{F}$ -measurable and π -integrable in X . The property γ'') of P , $P \geq 0$, obviously follows from the property γ) of Λ and δ'') follows from δ) as

$$\int_C 1 \pi(dx) = \pi(C) = \pi_c(X) = \Lambda \pi_c(X) = \int_C P(x, X) \pi(dx)$$

must hold for every $C \in \mathfrak{F}$. The countable additivity of $P(x, A)$ as stated in α'') follows similarly from (2.4) as, for $A = \sum A_i$,

$$\begin{aligned} \int_C P(x, A) \pi(dx) &= \Lambda \pi_c(A) = \sum \Lambda \pi_c(A_i), \\ &= \sum \int_C P(x, A_i) \pi(dx) = \int_C \sum P(x, A_i) \pi(dx) \end{aligned}$$

must hold for every $C \in \mathfrak{F}$. It should be noted that γ''), A being fixed, and δ'') follow only for π -almost all $x \in X$. But no change in Λ nor in α'') is effected if, for any given $A \in \mathfrak{F}$, $P(x, A)$ is modified in some π -null set of points x . Hence, P can be chosen such that γ'') and δ'') hold without exception.

So far (2.3) has been shown to hold for every $\varphi = \pi_c$. Hence it holds for any finite linear combination of such φ . Every $\varphi \geq 0$ in Φ_π is however known to be the limit function in $\mathfrak{F}$ of a non-decreasing sequence φ_i of such special set functions,

$$\varphi(B) = \lim \varphi_i(B), \quad B \in \mathfrak{F}.$$

(2.3) holds for each φ_i . As $i \rightarrow \infty$ the left hand side of (2.3) tends to $\Lambda\varphi(A)$ by virtue of lemma 2.2. The right hand side tends to

$$\int_X P(x, A)\varphi(dx)$$

for exactly the same reason since this right hand side had been shown before to represent a Markoff operator $\Lambda'\varphi(A)$ with domain of definition $\Phi' = \Phi_\pi$. Theorem 2.1 is herewith completely proved.

3. Iteration of a Markoff operator. An important feature which is usually attributed to a Markoff process and which we did not have to consider in the previous section is the possibility of unlimited iteration of Λ within its domain of definition Φ' or, in other words, the property that

$$(3.1) \quad \Lambda\Phi' \subset \Phi'.$$

We confine ourselves to the case where Φ' is of the type Φ_π , π being a given finite measure in $\mathfrak{F}$ and where $\Lambda\Phi_\pi \subset \Phi_\pi$.

The question arises, of course, whether this restriction is natural. This question is answered a little further below. The result proved there has the following implication. If Φ' is closed under contractions of f.c.a. set functions and also under monotone convergence and if $\Lambda\Phi' \subset \Phi'$, then Φ' is the union of domains of type Φ_π each of which satisfies $\Lambda\Phi_\pi \subset \Phi_\pi$. In the extreme case $\Phi' = \Phi$ considered by Yosida & Kakutani (3.1) is an automatic consequence of α), definition 2.1. But in general (3.1) represents a new additional condition on Λ .

We begin with the following simple observation.

Lemma 3.1. A Markoff operator Λ with domain of definition Φ_π , π being a fixed finite measure in $\mathfrak{F}$, maps Φ_π upon a subspace of $\Phi_\Lambda\pi$ or, in other words, π -continuity of an f.a.c. φ in $\mathfrak{F}$ implies $\Lambda\pi$ -continuity of $\Lambda\varphi$.

Proof. Suppose that $\Lambda\pi(A) = 0$ for some $A \in \mathfrak{F}$. On applying the representation formula (2.3) to π we infer that the relation $\Lambda\pi(A) = 0$ is equivalent to the relation

$$\pi(B) = 0, \quad B = \{x|P(x, A) \neq 0\}$$

and that $\pi(B') = 0$, $B' \subset B$, $B' \in \mathfrak{F}$. If φ is π -continuous $\varphi(B') = 0$ holds for every such set B' . Hence, by (2.3), $\Lambda\varphi(A) = 0$. The lemma is hereby proved.

Definition 3.1. A Markoff process $P(x, A)$ over X , Λ is called measurable with respect to a given finite measure π in $\mathfrak{F}$ (π -measurable) if it is π -conditioned and if for every π -null set A the set $B = \{x|P(x, A) \neq 0\}$ is also a π -null set.

Lemma 3.2. A necessary and sufficient condition that a Markoff operator Λ with domain of definition Φ_π satisfies $\Lambda\Phi_\pi \subset \Phi_\pi$ is that $\Lambda\pi$ is π -continuous. An equivalent condition is that the associated Markoff process $P(x, A)$ is π -measurable.

Proof. The necessity is trivial and the sufficiency follows from lemma 3.1.

In fact, $\varphi \in \Phi_\pi$ implies that $\Lambda\varphi$ is $\Lambda\pi$ -continuous, but since $\Lambda\pi$ is π -continuous $\Lambda\varphi$ is π -continuous, $\Lambda\varphi \in \Phi_\pi$, $\Lambda\Phi_\pi \subset \Phi_\pi$. That π -continuity of $\Lambda\pi$ and π -measurability of P mean the same thing is obvious from the proof of lemma 3.1. Lemma 3.2 is thereby proved.

The question raised above will now be examined. It seems natural to demand of the domain of definition Φ' of a Markoff operator Λ that it have the following properties. $D_1)$ Φ' is a linear subspace of Φ . $D_2)$ $\varphi \in \Phi' \rightarrow \varphi_C \in \Phi'$, $C \in \mathcal{F}$. $D_3)$ Closure, $\varphi_i \in \Phi'$, $\varphi_i \uparrow \varphi$, $\varphi \in \Phi \rightarrow \varphi \in \Phi'$.

Note that every space Φ_π , π being a finite measure in $\mathcal{F}$, has these three properties. Conversely, a space Φ' which satisfies $D_1)$ - $D_3)$ has the property that

$$(3.2) \quad \pi \in \Phi', \quad \pi \geq 0 \rightarrow \Phi_\pi \subset \Phi'.$$

In fact, $\varphi \in \Phi_\pi$ implies that φ is monotone limit function of linear combinations of measures π_C , $C \in \mathcal{F}$, and it follows immediately from $D_1)$ - $D_3)$ that $\varphi \in \Phi'$. Another property of a space Φ' , $D_1)$ - $D_3)$, is that

$$(3.3) \quad \varphi \in \Phi' \rightarrow |\varphi| \in \Phi'$$

where $|\varphi| (A)$ is the total variation of φ over A ; $|\varphi|$ is well known to be again an f.c.a. set function of A , $A \in \mathcal{F}$. In fact, a set $C \in \mathcal{F}$ is known to exist such that $\varphi_C \geq 0$ in $\mathcal{F}$ and $\varphi_{X-C} \leq 0$ in $\mathcal{F}$ and that $|\varphi| = \varphi_C - \varphi_{X-C}$. This implies (3.3) by virtue of $D_1)$ and $D_2)$.

The question raised above is answered by the following

Lemma 3.3. Suppose that Λ is a Markoff operator with a domain of definition Φ' satisfying $D_1)$ - $D_3)$ and such that $\Lambda\Phi' \subset \Phi'$. Then to any countable subset S of Φ' there belongs a finite measure π in $\mathcal{F}$, $\pi \in \Phi'$, such that $S \subset \Phi_\pi \subset \Phi'$ and $\Lambda\Phi_\pi \subset \Phi_\pi$.

Proof. Let φ_i , $i = 1, 2, \dots$, denote the elements of S and $\pi_i = |\varphi_i|$ their total variations, $\pi_i \in \Phi'$, $\pi_i \geq 0$. By virtue of the assumption $\Lambda\Phi' \subset \Phi'$, $\Lambda^k \pi_i \in \Phi'$, $k \geq 0$. Also $\Lambda^k \pi_i \geq 0$ since Λ is positive. Using a remark of Halmos [5] we consider the measure

$$(3.4) \quad \pi = \sum_{i \geq 1, k \geq 0} \frac{2^{-i-k}}{\pi_i(X)} \Lambda^k \pi_i$$

every partial sum of which is an element of Φ' by virtue of $D_1)$. In consequence of $\delta)$, definition 2.1, we have

$$\Lambda^k \pi_i(X) = \pi_i(X), \quad \pi(X) = 2.$$

Therefore π is finite in $\mathcal{F}$, $\pi \in \Phi$. From $D_3)$ it now follows that $\pi \in \Phi'$, and from (3.2) that $\Phi_\pi \subset \Phi'$. Next we show that $S \subset \Phi_\pi$ i.e. that every φ_i is π -continuous. In fact, this is trivial since

$$\pi \geq \frac{2^{-i}}{\pi_i(X)} \pi_i \geq 0, \quad |\varphi_i(A)| \leq |\varphi_i| (A) = \pi_i(A).$$

It remains to show that $\Lambda\phi_\pi \subset \Phi_\pi$ i.e. that $\Lambda\pi$ is π -continuous. But this follows from (3.4) if lemma 2.2 is applied to the partial sums and to the sum π ,

$$\Lambda\pi = \sum_{i \geq 1, k \geq 0} \frac{2^{-i-k}}{\pi_i(X)} \Lambda^{k+1} \pi_i \leq 2\pi.$$

Lemma 3.3 is hereby completely proved.

(3.4) makes the following supplementary statement obvious.

Lemma 3.4. There exists a best π in lemma 3.3: Every set A which is a simultaneous null set for all the measures $\Lambda^k | \varphi |$, $\varphi \in S$, $k \geq 0$, is a π -null set.

4. The dual L of a Markoff operator. We shall have to deal with operators L in which argument as well as value are real-valued point functions $f(x)$, $f'(x) = Lf(x)$ in X . $Lf(x)$ denotes the value of the function Lf at the point x . Strictly speaking, argument and value of L will be π -equivalence classes $\mathbf{f}$, $\mathbf{f}' = L\mathbf{f}$, of such functions π being a finite measure in $\mathfrak{F}$ which is given in advance and fixed. However, such an operator can always be regarded as a function-to-function operator with the following two properties. It is multivalued, the different values for the same $f = f(x)$ being all the functions $f' = f'(x)$ within the same π -equivalence class $\mathbf{f}'$, and it has the same set of values for every f within the same π -equivalence class $\mathbf{f}$. A relation in which this function-to-function concept for L is used is, of course, to be understood as a π -equivalence. We shall freely use both concepts for L .

$\mathfrak{B}_\pi$ denotes the space whose elements are the π -equivalence classes $\mathbf{f}$ formed by the bounded and $\mathfrak{F}$ -measurable functions $f(x)$ in X . Order relations $\mathbf{f} \leq \mathbf{g}$, addition of such classes and their multiplication by real numbers are defined as usual. $\mathfrak{B}_\pi$ is a linear space.

Theorem 4.1. To any Markoff operator Λ with domain of definition Φ_π and with the property that $\Lambda\Phi_\pi \subset \Phi_\pi$ there belongs a unique operator Lf —the dual of Λ —with the following properties. a) L is defined in $\mathfrak{B}_\pi$ and has values again in $\mathfrak{B}_\pi$; if $f(x) \in \mathbf{f} \in \mathfrak{B}_\pi$ and if $Lf(x)$ denotes any function in the π -equivalence class $L\mathbf{f}$ then the relation

$$(4.1) \quad \int_X Lf(x) \varphi(dx) = \int_X f(x) \Lambda\varphi(dx)$$

holds for all $\varphi \in \Phi_\pi$. b) L is linear. c) L is positive, $\mathbf{f} \geq 0 \rightarrow L\mathbf{f} \geq 0$. d) $L1 = 1$.

Proof. Consider the expression

$$(4.2) \quad \int_X f(x) \Lambda\pi_C(dx), \quad C \in \mathfrak{F}, |f| \leq M,$$

where $f(x)$ is $\mathfrak{F}$ -measurable. As π_C is π -continuous $\Lambda\pi_C$ is, by virtue of lemma 3.1, $\Lambda\pi$ -continuous. In consequence of the hypothesis $\Lambda\Phi_\pi \subset \Phi_\pi$ and of lemma 3.2 $\Lambda\pi_C$ is π -continuous. The Lebesgue integral (4.2) is therefore not affected

by a change of $f(x)$ within a π -null set of points in X , in other words, it depends actually only on the π -equivalence class $\mathbf{f}$ determined by f . In the beginning of the proof of theorem 2.1

$$\Lambda\pi_c(A)$$

was shown to be an f.a.c. set function of C , $C \varepsilon \mathfrak{F}$, A being kept fixed. The same is true for (4.2). This is inferred from the following familiar theorem. If φ_i , φ are finite measures in $\mathfrak{F}(X \varepsilon \mathfrak{F})$ and if $\varphi_i(A) \rightarrow \varphi(A)$ holds for every $A \varepsilon \mathfrak{F}$ then

$$\int_X f(x)\varphi_i(dx) \rightarrow \int_X f(x)\varphi(dx)$$

holds for every bounded and $\mathfrak{F}$ -measurable $f(x)$ in X . We assert that (4.2) as an f.a.c. set function of C is continuous with respect to the measure $\pi(C)$. In fact, from δ), definition 2.1, it follows that

$$\left| \int_X f(x)\Lambda\pi_c(dx) \right| \leq M \Lambda\pi_c(X) = M\pi_c(X) = M\pi(C).$$

The Radon-Nikodym theorem therefore assures the existence of a π -measurable function $f'(x)$, $|f'| \leq M$, such that for all $C \varepsilon \mathfrak{F}$

$$(4.3) \quad \int_X f(x)\Lambda\pi_c(dx) = \int_C f'(x)\pi(dx) = \int_X f'(x)\pi_c(dx).$$

This determines $f'(x)$ uniquely up to π -equivalence so that the π -equivalence class $\mathbf{f}$ containing f actually determines in this way a unique π -equivalence class $\mathbf{f}' = L\mathbf{f}$. Now (4.3) is the relation (4.1) for the case $\varphi = \pi_c$. (4.1), f and Lf being fixed, holds therefore for linear combinations of such φ . But every $\varphi \varepsilon \phi_\pi$, $\varphi \geq 0$ —it suffices to consider the case $\varphi \geq 0$ —is limit function of a sequence $0 \leq \varphi_1 \leq \varphi_2 \leq \dots$ of such special functions. By virtue of lemma 2.2 we have also $\Lambda\varphi_i \rightarrow \Lambda\varphi$ in $\mathfrak{F}$. (3.1) now follows for φ on letting $i \rightarrow \infty$ and on using the theorem quoted above. *a*) is hereby completely proved. The properties *b*), *c*) and *d*) of L reflect the properties β), γ) and δ) of Λ , respectively. They follow readily from them by means of the duality relation (4.1). The uniqueness proof for L is already implicitly contained in our argument (application of (4.1) to $\varphi = \pi_c$). Theorem 4.1 is thereby completely proved.

The operator L depends on the underlying measure π insofar as the duality relation has been proved only for $\varphi \varepsilon \phi_\pi$ and $f \varepsilon \mathfrak{B}_\pi$ but the form of this relation suggests that it cannot have anything else to do with π . In fact, the following statement is easily proved. If two finite measures π , π' in $\mathfrak{F}$ are continuous with respect to each other then $\Phi_\pi = \Phi_{\pi'}$, $\mathfrak{B}_\pi = \mathfrak{B}_{\pi'}$, and $L \equiv L'$ holds for the corresponding L -operators within $\mathfrak{B}_\pi$.

In the case where Λ is defined in all of Φ Kakutani could define L simply by the formula

$$(4.4) \quad Lf(x) = \int_x P(x, dy)f(y).$$

In our case $\Phi' = \Phi_\pi$ this can be shown to hold as a π -equivalence. We shall need it, however, only for the case where f is a characteristic function, and in this case (4.4) is most easily proved by means of (2.3) and (4.1).

5. Density operators L^* of a Markoff process. From a Markoff operator $\Lambda\varphi$, $\varphi \in \Phi_\pi$, based upon some given finite measure π in $\mathfrak{F}$ and with the property that

$$(5.1) \quad \Lambda\Phi_\pi \subset \Phi_\pi$$

there arises another operator L^* from point function to point function or rather from π -equivalence class to π -equivalence class. L^* is merely another form of Λ which one obtains on characterizing all set functions involved by their derivatives (densities) with respect to the measure π . L^* essentially depends on the underlying measure π , in contrast to the dual L of Λ .

Let $\mathfrak{L}_\pi$ denote the linear space whose elements are the different π -equivalence classes $\mathbf{f}$ formed by the $\mathfrak{F}$ -measurable and π -integrable functions $f(x)$ in X . As mentioned before the relations

$$(5.2) \quad \varphi(A) = \int_A f(x)\pi(dx), \quad A \in \mathfrak{F}, \quad f(x) = \frac{\varphi(dx)}{\pi(dx)}$$

furnish a one-to-one mapping between Φ_π and $\mathfrak{L}_\pi$. If $\Lambda\varphi$, $\varphi \in \Phi_\pi$, is a Markoff operator with the property (5.1), the relations

$$(5.3) \quad \Lambda\varphi(A) = \int_A L^*f(x)\pi(dx), \quad A \in \mathfrak{F}, \quad L^*f(x) = \frac{\Lambda\varphi(dx)}{\pi(dx)}$$

define together with (5.2) a unique operator L^* with domain of definition $\mathfrak{L}_\pi$ and such that

$$(5.4) \quad L^*\mathfrak{L}_\pi \subset \mathfrak{L}_\pi.$$

This follows from lemmas 3.1 and 3.2. It is easily seen that L^* has the following four properties which are the exact respective equivalents of the four properties of Λ , definition 2.1, designated by the corresponding Greek letters, 5.1 being added to α). $a^*)$ (5.4) holds. $b^*)$ L^* is linear. $c^*)$ L^* is positive. Finally, $d^*)$

$$\int_X L^*f(x)\pi(dx) = \int_X f(x)\pi(dx), \quad f \in \mathfrak{L}_\pi.$$

Conversely, it is easily seen that any given operator L^* with the properties $a^*)-d^*)$ defines, by means of (5.2) and (5.3), a unique Markoff operator $\Lambda\varphi$, $\varphi \in \Phi_\pi$, with the property (5.1). In other words, for a given finite measure π in $\mathfrak{F}$ the relations (5.2) and (5.3) define a one-to-one mapping between Markoff operators Λ in Φ_π , $\Lambda\Phi_\pi \subset \Phi_\pi$, and operators L^* in $\mathfrak{L}_\pi$ with the properties $a^*)-d^*)$.

In frequent cases a Markoff process is defined by a single-valued operator L^* (plus a given measure π) from point function to point function. In order that such an operator may be regarded as an operator L^* in the previous sense it is evidently sufficient that in addition L^* transforms π -equivalent functions again into π -equivalent functions or merely that the L^* -transform of a function $f(x)$ which vanishes π -almost everywhere in X have again this property.

On substituting from (5.2) and (5.3) into the duality relation (4.1), with f replaced by g , we infer that

$$(5.5) \quad \int_X f(x) Lg(x) \pi(dx) = \int_X g(x) L^* f(x) \pi(dx)$$

holds for every bounded and $\mathfrak{F}$ -measurable $g(x)$ and for every $\mathfrak{F}$ -measurable and π -integrable $f(x)$ in X .

6. Extreme case of a point transformation. Let Tx denote a single-valued point transformation of X into precisely X , $TX = X$. T need not be one-to-one. $T^{-1}x$ is defined in X and in general multivalued. TA is well defined for all sets A and so is $T^{-1}A = \{x \mid Tx \in A\}$. By hypothesis, $T^{-1}X = X$.

First we wish to write T^{-1} as a Markoff process, and we impose for this purpose the condition that $A \in \mathfrak{F} \rightarrow T^{-1}A \in \mathfrak{F}$ or, briefly, that T^{-1} is $\mathfrak{F}$ -measurable,

$$(6.1) \quad T^{-1}\mathfrak{F} \subset \mathfrak{F}.$$

The answer is given by

$$(6.2) \quad \Lambda'\varphi(A) = \varphi(T^{-1}A).$$

In fact, by virtue of (6.1) the right hand side is defined for $A \in \mathfrak{F}$ whenever the set function $\varphi(A)$ is defined in $\mathfrak{F}$. Furthermore, T^{-1} is an isomorphism between $\mathfrak{F}$ and $T^{-1}\mathfrak{F}$ which preserves not only sum but also intersection and disjointness of sets (the latter property fails to hold in general for T itself). Hence $\varphi(T^{-1}A)$ is additive in A whenever $\varphi(A)$ is. (6.2) is easily seen to be a Markoff operator with all of Φ for its domain of definition, $\Lambda'\Phi \subset \Phi$. By the way, if φ is a probability distribution over X then $\Lambda'\varphi$ is the distribution resulting from it by the point transformation T . The probability kernel $P'(x, A)$ of Λ' can be simply written in terms of the characteristic function

$$e_B(x) \equiv \epsilon_x(B)$$

of a set B —two notations are used according as we wish to regard the characteristic function as a point function or f.c.a. set function. We have

$$(6.3) \quad \Lambda' \varphi(A) = \int_A P'(x, A) \varphi(dx), \quad P' = e_A(Tx) = e_{T^{-1}A}(x).$$

Consider now a fixed finite measure π in $\mathfrak{F}$. What does it mean that $\Lambda' \Phi_\pi \subset \Phi_\pi$? By virtue of lemma 3.2 it means that $\pi(T^{-1}A)$ is π -continuous or, in other words, that T^{-1} transforms every π -null set again into such a set.

Secondly we wish to write T itself as a Markoff process. We suppose that both T and T^{-1} are $\mathfrak{F}$ -measurable,

$$(6.4) \quad T\mathfrak{F} \subset \mathfrak{F}, \quad T^{-1}\mathfrak{F} \subset \mathfrak{F}.$$

If T is one-to-one the corresponding Markoff operator Λ is simply the inverse of (6.2), $\Lambda\varphi(A) = \varphi(TA)$. But in general this cannot be the answer since $\varphi(TA)$ will not be additive in A .

We consider a fixed finite measure π in $\mathfrak{F}$ and we suppose that $\pi(A)$ be continuous with respect to $\pi(T^{-1}A)$ in $\mathfrak{F}$ or, in other words, that T transforms all π -null sets of the form $T^{-1}A$ again into π -null sets. This assumption can also be phrased as follows. The set function $\pi(TA')$ which is easily seen to be a measure as long as $A' \in T^{-1}\mathfrak{F}$ is $\pi(A')$ -continuous in $T^{-1}\mathfrak{F}$. This implies the existence of a $T^{-1}\mathfrak{F}$ -measurable and π -integrable function $r(x) \geq 0$ in X such that

$$(6.5) \quad \pi(TA') = \int_{A'} r(x) \pi(dx), \quad A' \in T^{-1}\mathfrak{F}.$$

The condition of $T^{-1}\mathfrak{F}$ -measurability determines r uniquely in the sense of π -equivalence. If, however, this condition is replaced by the weaker one of $\mathfrak{F}$ -measurability—this is weaker by virtue of (6.4)—then (6.5) can in general be satisfied by many essentially different functions $r(x)$. An example is given below. By the way, what does $T^{-1}\mathfrak{F}$ -measurability mean for r ? A moment's reflection shows that this is tantamount to $r(x) = s(y)$, $y = Tx$, where $s(y)$ is single-valued and $\mathfrak{F}$ -measurable in X . (6.5) can be written

$$\pi(A) = \int_{T^{-1}A} r(x) \pi(dx) = \int_X r(x) e_A(Tx) \pi(dx), \quad A \in \mathfrak{F}.$$

In other words, the relation

$$(6.6) \quad \int_X f(x) \pi(dx) = \int_X r(x) f(Tx) \pi(dx)$$

holds for every $\mathfrak{F}$ -measurable characteristic function in X . By means of familiar arguments it is then proved to hold for every $\mathfrak{F}$ -measurable and π -integrable

function $f(x)^*$. At the same time the integrability of the right hand integrand follows. In other words, the positive linear operator

$$(6.7) \quad L^* f(x) = r(x)f(Tx)$$

maps $\mathfrak{L}_\pi$ upon itself or part of itself (properties a^* – c^*) in Section 5). (6.6) is the property d^*). Hence,

$$(6.8) \quad f(x) = \frac{d\varphi}{d\pi}, \quad \Lambda\varphi(A) = \int_A r(x)f(Tx)\pi(dx)$$

defines a Markoff operator, $\varphi \in \Phi_\pi$, $\Lambda\Phi_\pi \subset \Phi_\pi$. (6.7) or (6.8) is what we looked for. Λ is a right inverse of Λ' ,

$$\Lambda'\Lambda\varphi \equiv \varphi \text{ in } \mathfrak{F}.$$

In order to show this it suffices to consider contractions $\varphi = \pi_c$ or $f = e_c$. By virtue of (6.2), the left hand side evaluated for the set A equals

$$\Lambda\varphi(T^{-1}A) = \int_{T^{-1}A} r(x)e_c(Tx)\pi(dx) = \int_{T^{-1}(AC)} r(x)\pi(dx) = \pi(AC) = \pi_c(A).$$

The assertion is thereby proved.

For an example consider the case where X is the circular line of all angles $x \bmod 2\pi$ and where $\mathfrak{F}$ is the field of all Borel sets on X . T is the 2-to-1 transformation $Tx = 2x \bmod 2\pi$. $T^{-1}\mathfrak{F}$ is the collection of all those Borel sets which are invariant under the covering transformation $x' = x + \pi \bmod 2\pi$. (6.4) is satisfied. The ordinary Lebesgue measure π , $\pi(dx) = dx$, is seen to be invariant under T^{-1} , $\pi(T^{-1}A) = \pi(A)$. (6.5) says that

$$r(x + \pi) + r(x) = 2$$

holds almost everywhere. The solution $r = 1$ is essentially the only one which is $T^{-1}\mathfrak{F}$ -measurable, *i.e.*, which is measurable and invariant under the covering transformation, $r(x + \pi) = r(x)$. However, there are infinitely many other solutions r , for instance, $r(x) = 2 \sin^2(x/2)$.

7. The maximal ergodic theorem for Markoff operators Λ and L^* . Consider a Markoff operator Λ which is defined in a domain

$$(7.1) \quad \Phi' = \Phi_\mu$$

where μ is a given finite measure in $\mathfrak{F}$. We suppose that the associated Markoff process is μ -measurable, *i.e.*, (lemma 3.2) that

$$(7.2) \quad \Lambda\Phi' \subset \Phi'.$$

* The construction of a $T^{-1}\mathfrak{F}$ -measurable function r for which (6.6) holds is due to Y. N. DOWKER [4].

The operators

$$(7.3) \quad \Lambda_k = \sum_0^{k-1} \Lambda^v, \quad k = 1, 2, \dots, \quad \Lambda_1 \varphi \equiv \varphi,$$

are well defined in Φ' , linear and positive. (7.1), (7.2) and the linearity of the space Φ' imply that

$$(7.4) \quad \Lambda_k \Phi' \subset \Phi'.$$

Hence, every $\Lambda_k \mu$ is μ -continuous. The converse statement that μ is continuous with respect to any $\Lambda_k \mu$ is trivially true.

Now consider two f.c.a. set functions φ and $\pi \geq 0$. We suppose that

$$(7.5) \quad \varphi \varepsilon \Phi'$$

and that

$$(7.6) \quad \pi \geq 0, \quad \Phi_\pi = \Phi' = \Phi_\mu$$

which means precisely that the two measures μ and π are continuous with respect to each other. Consider the point functions

$$(7.7) \quad q_k(x) = \frac{\Lambda_k \varphi(dx)}{\Lambda_k \pi(dx)}, \quad k = 1, 2, \dots$$

in X . The statement about the mutual continuity of μ , $\Lambda_k \mu$ holds just as well for π in place of μ . Hence, the two measures μ , $\Lambda_k \pi$ are continuous with respect to each other. Therefore, the functions $q_k(x)$ are well defined Radon-Nikodym derivatives, and they are all defined in the sense of μ -equivalence.

Suppose now that we start with an operator $L^* f$ instead, where L^* is defined, linear and positive in $\mathfrak{L}_\mu$,

$$L^* \mathfrak{L}_\mu \subset \mathfrak{L}_\mu$$

and where

$$\int_X L^* f(x) \mu(dx) \equiv \int_X f(x) \mu(dx)$$

in $\mathfrak{L}_\mu$ (properties $a^*)$ - d^*), §3). With two functions $f(x)$, $p(x)$

$$(7.8) \quad f \varepsilon \mathfrak{L}_\mu, \quad p \varepsilon \mathfrak{L}_\mu, \quad p > 0,$$

we consider the quotients

$$(7.9) \quad q_k(x) = \frac{L_k^* f(x)}{L_k^* p(x)}$$

where

$$(7.10) \quad L_k^* = \sum_0^{k-1} L^{*\nu}, \quad k \geq 1.$$

We may then write (7.9) in the form (7.7) if we let

$$\varphi(A) = \int_A f(x) \mu(dx), \quad \Lambda \varphi(A) = \int_A L^* f(x) \mu(dx),$$

$A \in \mathfrak{F}$, or

$$\frac{\varphi(dx)}{\mu(dx)} = f(x), \quad \frac{\Lambda \varphi(dx)}{\mu(dx)} = L^* f(x)$$

and analogously with φ, f replaced by π, p . Obviously Λ^ν corresponds to $L^{*\nu}$. Note that (7.6) is a consequence of (7.8).

Our task is to examine the limit properties as $k \rightarrow \infty$ of the sequence of functions $q_k(x)$ in X .

Theorem 7.1. Under the assumptions made above and for any given integer $n \geq 1$ the set

$$A = \{x \mid \sup_{k \leq n} q_k(x) \geq 0\}$$

satisfies the inequality $\varphi(A) \geq 0$.

Proof. Consider the pairwise disjoint sets

$$(7.11) \quad A_i = \{x \mid q_i(x) \geq 0, q_j(x) < 0, j < i\},$$

$i = 1, \dots, n, A_1 = \{x \mid q_1 \geq 0\}$. Obviously $A_i \in \mathfrak{F}$. For $i > n$ we define

$$A_i = 0, \quad i > n.$$

We write $e_i(x)$ for the characteristic function of A_i and

$$(7.12) \quad e^i(x) = e_i(x) + e_{i+1}(x) + \dots, \quad i \geq 1,$$

for the characteristic function of $A_i + A_{i+1} + \dots$. Obviously e^1 is the characteristic function of

$$A = A_1 + A_2 + \dots,$$

and

$$(7.13) \quad e_i = e^i = 0, \quad i > n.$$

By virtue of (7.3), (7.7) and (7.11) the set functions

$$\sum_0^{j-1} \Lambda^v \varphi, \quad 0 < j \leq i \leq n,$$

are, within A_i , ≥ 0 for $j = i$ and ≤ 0 for $j < i$. Hence the set functions

$$\sum_j^{i-1} \Lambda^v \varphi, \quad 0 \leq j < i \leq n,$$

are all ≥ 0 within A_i . In other words,

$$(7.14) \quad \int_B e_i(x) \sum_j^{i-1} \Lambda^v \varphi(dx) \geq 0, \quad B \in \mathfrak{F}, \quad 0 \leq j < i.$$

The condition $i \leq n$ may be dropped by virtue of (7.13). On summing up with respect to i , $i > j$, j being fixed, and on using (7.12) we find that

$$(7.15) \quad \int_B \sum_{v \geq j} e^{v+1}(x) \Lambda^v \varphi(dx) \geq 0, \quad B \in \mathfrak{F}, \quad j \geq 0.$$

Next we define successively for decreasing i the $\mathfrak{F}$ -measurable functions $h_i(x)$, $i \geq 0$,

$$(7.16) \quad \begin{aligned} h_i &= 0, & i &\geq n, \\ h_i &= e^{i+1} + (1 - e^{i+1})Lh_{i+1} = (1 - Lh_{i+1})e^{i+1} + Lh_{i+1}, \end{aligned}$$

where L is the dual of Λ (§4). By virtue of properties *b*), *c*) and *d*) of this operator $0 \leq h \leq 1$ implies that $0 \leq Lh \leq 1$. Hence it follows successively that

$$(7.17) \quad 0 \leq h_i \leq 1, \quad i \geq 0.$$

On introducing the functions

$$(7.18) \quad g_i = (1 - Lh_{i+1})e^{i+1} = h_i - Lh_{i+1}$$

we find that

$$(7.19) \quad h_0 = g_0 + Lg_1 + L^2g_2 + \dots.$$

Obviously

$$(7.20) \quad g_i \geq 0, \quad i \geq 0.$$

From (7.12) and the recursion formula (7.16) we infer that

$$h_i - h_{i+1} = e_{i+1}(1 - Lh_{i+1}) + (1 - e^{i+2})L(h_{i+1} - h_{i+2}), \quad i \geq 0.$$

Hence and from (7.17) it follows by induction that

$$(7.21) \quad h_{i+1} \leq h_i, \quad i \geq 0.$$

We assert that

$$(7.22) \quad e^{j+1}(g_j - g_{j-1}) \geq 0, \quad j \geq 0,$$

if we let $g_{-1} = 0$. In fact, on using the relations

$$(7.23) \quad e^{i+1} e^{k+1} = e^{k+1}, \quad 0 \leq i \leq k,$$

we infer from the first relation (7.18) that for $j > 0$

$$e^{j+1}(g_j - g_{j-1}) = e^{j+1} L(h_j - h_{j+1}).$$

(7.22) follows from this and from (7.21) for $j > 0$. For $j = 0$ it follows from (7.20) and from $g_{-1} = 0$. On using (7.23) for $i = j$, $k = \nu$ we can now conclude from (7.15) and (7.22) that

$$\int_X \sum_{j \geq 0} (g_j(x) - g_{j-1}(x)) \sum_{\nu \geq j} e^{\nu+1}(x) \Lambda^\nu \varphi(dx) \geq 0.$$

The integrand may be written

$$\sum_{\nu \geq 0} g_\nu(x) e^{\nu+1}(x) \Lambda^\nu \varphi(dx)$$

or, by virtue of (7.18), $g_\nu e^{\nu+1} = g_\nu$,

$$\sum_{\nu \geq 0} g_\nu(x) \Lambda^\nu \varphi(dx).$$

Hence we infer on using the duality relation and (7.19) that

$$\int_X h_0(x) \varphi(dx) = \sum_{\nu \geq 0} \int_X L^\nu g_\nu(x) \varphi(dx) = \sum_{\nu \geq 0} \int_X g_\nu(x) \Lambda^\nu \varphi(dx) \geq 0.$$

By virtue of (7.11), $q_1 = d\varphi/d\pi < 0$ holds within $X - A \subset X - A_1$. Hence, $\varphi \leq 0$ holds within $X - A$. Also, $h_0 \geq 0$ by (7.17). Within $A = A_1 + A_2 + \dots$, however, we find that $e^1 = 1$ and, consequently, that

$$h_0 = e^1 + (1 - e^1)Lh_1 = 1.$$

Therefore, *a fortiori*

$$0 \leq \int_A h_0 d\varphi = \varphi(A).$$

Theorem 7.1 is herewith proved.

Maximal ergodic theorem. Under the assumptions made about Λ , φ , π and with (7.7), or those about L^* , f , p and with (7.9), the inequality $\varphi(A) \geq a\pi(A)$ is satisfied by any real number a and by the corresponding set

$$A = \{x \mid \sup_{i \geq 1} q_i(x) \geq a\}.$$

The inequality $\varphi(B) \leq b\pi(B)$ is satisfied by any real number b and by the corresponding set

$$B = \{x \mid \inf_{i \geq 1} q_i(x) \leq b\}.$$

Proof. This follows from theorem 7.1 by simple and familiar arguments. It suffices to consider the case $a = 0$ because everything is reduced to this case if φ is replaced by $\varphi - a\pi$. $\Lambda_k\varphi$ goes then over into $\Lambda_k\varphi - a\Lambda_k\pi$ and $q_k(x)$ into $q_k(x) - a$. But for $a = 0$ the present theorem follows easily from the preceding one if we let $k \rightarrow \infty$. The second statement follows from the first if $-\varphi$ is considered instead of φ .

8. Conservative and dissipative part of a general Markoff process. In this and the following paragraphs the structure of a π -measurable Markoff process L^* is investigated. The maximal ergodic theorem serves as the main tool in this investigation. We begin with

Lemma 8.1. If $f \geq 0$, $p > 0$ in X , both functions being π -integrable in X , then $f(x)$ vanishes π -almost everywhere in the set $A = \{x \mid \inf_n q_n(x) = 0\}$ where $q_n(x)$ is given by (7.9) and (7.10).

Proof. Positivity of L^* and the hypothesis, $f \geq 0$, imply that $q_n \geq 0$. Therefore, $A = \{x \mid \inf_n q_n(x) \leq 0\}$. Hence and from the maximal ergodic theorem it follows that $\int_A f d\pi \leq 0$, which implies the lemma since $f \geq 0$.

Theorem 8.1. With respect to a given π -measurable Markoff process L^* , X splits into two disjoint sets $\varepsilon \mathfrak{F}$, $X = CX + DX$, the conservative and the dissipative part. Every π -integrable function $p(x)$ in X , $p \geq 0$, satisfies

$$L_\infty^* p(x) = \sum_0^\infty L^{*\nu} p(x) < \infty \text{ in } DX$$

π -almost everywhere. Every π -integrable function $p(x)$ in X , $p > 0$, satisfies

$$L_\infty^* p(x) = \sum_0^\infty L^{*\nu} p(x) = \infty \text{ in } CX$$

π -almost everywhere.

Proof. Consider a fixed π -integrable $p > 0$ in X . The series

$$\sum_0^\infty L^{*\nu} p(x)$$

has non-negative terms. Let CX denote the set $(\varepsilon \mathfrak{F})$ in which its sum is infinite, and let $DX = X - CX$. Consider now an arbitrary π -integrable $p' > 0$ in X . In consequence of lemma 8.1, $f = p' > 0$, and on setting

$$L_n^* = \sum_0^{n-1} L^{*\nu}$$

we obtain:

$$\inf_n \frac{L_n^* p'(x)}{L_n^* p(x)} > 0$$

holds in X minus some π -null set. On interchanging p, p' we find that

$$\sup_n \frac{L_n^* p'(x)}{L_n^* p(x)} < \infty$$

holds in X minus some π -null set. Theorem 8.1 readily follows from these statements, however, the first statement follows only if $p > 0$. But if $p \geq 0$ this statement is *a fortiori* true. In fact, this follows from the positivity of L^* if $p + 1 > 0$ is considered instead of p . Theorem 8.1 is thereby proved.

An immediate consequence of the theorem and of the inequality

$$|L^{*\nu} f| \leq L^{*\nu} |f|$$

is the following statement. If f is π -integrable in X then

$$\sum_0^\infty L^{*\nu} f(x)$$

converges absolutely π -almost everywhere in DX .

Lemma 8.2. Let $g \geq 0$ be π -integrable in X and let $A = \{x \mid L_\infty^* g(x) = \infty\}$. Then, for every π -integrable $f(x)$ such that $f = 0$ holds π -almost everywhere in $X - A$, $L^* f$ too vanishes π -almost everywhere in $X - A$. $P(x, X - A) = 0$ holds for π -almost all $x \in A$.

Proof. By virtue of $g \geq 0$,

$$0 \leq L_1 g(x) \leq L_2 g(x) \leq \dots$$

and

$$(8.1) \quad L^* L_n^* g(x) = L_{n+1}^* g(x) - g(x) \leq L_{n+1}^* g(x).$$

To prove the theorem it obviously suffices to deal with the case in which

$$f \geq 0 \text{ in } X.$$

On letting

$$(8.2) \quad h_n(x) = \inf (f(x), L_n^* g(x))$$

we can now infer that

$$0 \leq h_n(x) \uparrow f(x)$$

holds in A as well as in $X - A$. Hence and from lemma 2.2, translated into the L^* -language,

$$L^*h_n(x) \uparrow L^*f(x)$$

for π -almost all $x \in X$. From (8.1) and (8.2),

$$L^*h_n \leq L^*L_n^*g \leq L_{n+1}^*g \leq L_\infty^*g,$$

and hence

$$L^*f \leq L_\infty^*g$$

in X minus some π -null set. This inequality holds for any $f \geq 0$ satisfying the hypothesis and therefore for cf , $c > 0$,

$$L^*f \leq \frac{1}{c} L_\infty^*g.$$

Since c is arbitrary it follows from the definition of the set A that $L^*f = 0$ must hold for π -almost all $x \in X - A$. The first part of the lemma is thereby proved. As to the second part, consider the set functions

$$\varphi(B) = \int_B f \, d\pi, \quad \Lambda\varphi(B) = \int_B L^*f \, d\pi$$

where $f = 0$ in $X - A$. On using the first part we infer that

$$0 = \Lambda\varphi(X - A) = \int_X P(x, X - A)\varphi(dx) = \int_A P(x, X - A)f(x)\pi(dx).$$

The second part follows if we let $f(x) = 1$ in A . The lemma is herewith completely proved.

Lemma 8.3. If $f(x)$ is ≥ 0 and π -integrable in X , then either $f(x) = 0$ or $L_\infty^*f(x) = \infty$ holds within CX minus some π -null set.

Proof. Let $B' = CX - B \cdot CX$ where $B = \{x \mid L_\infty^*f(x) = \infty\}$. $L_\infty^*f < \infty$ holds in B' . If p is > 0 and π -integrable in X then, by theorem 8.1, $L_\infty^*p = \infty$ holds π -almost everywhere in CX and hence in B' . This implies that $\inf_n q_n(x) = 0$ holds π -almost everywhere in B' if $q_n = L_n^*f/L_n^*p$. Hence and from lemma 8.1 it follows that $f = 0$ holds in B' minus some π -null set. The lemma is thereby proved.

Lemma 8.4. If $f(x)$ is ≥ 0 and π -integrable in X the function $L_\infty^*f(x)$ can, within CX , take values different from 0 and ∞ only in a π -null set.

Proof. Consider the sets B, B' in the preceding proof. If lemma 8.3 is applied to L^*f instead of f these sets might change. But in virtue of the relations

$$L_n^*L^*f = L_{n+1}^*f - f, \quad L_\infty^*L^*f = L_\infty^*f - f$$

they can change only by a π -null set. On repeating this argument we infer from lemma 8.3 that $f = L^*f = L^{**}f = \dots = 0$ holds π -almost everywhere in B' . This implies the validity of the present lemma.

Theorem 8.2. If f is π -integrable in X and if $f = 0$ holds in DX minus some π -null set then $L^*f = 0$ too holds in DX minus some π -null set. The probability function $P(x, DX)$ vanishes for π -almost all $x \in CX$. If $f = 0$ in DX , $f > 0$ in CX then $L_\infty^*f = \infty$ holds in CX minus some π -null set.

The meaning of the first and second statement is that the Markoff process represents, within the conservative part CX , again a full-fledged Markoff process. The third statement means that, within CX , this Markoff process has its full x -space CX for its conservative part. In the simple case where the process is a one-to-one point transformation of X into X the process splits into two completely independent processes, one in CX and one in DX . In general, however, the dissipative part has some influence on the conservative part. This influence is studied further down.

Proof of theorem 8.2. We prove the last statement first. $DX = \{x \mid f = 0\}$ holds by hypothesis. On letting $B = \{x \mid L_\infty^*f = \infty\}$ we infer from lemma 8.3 that

$$CX \subset B + DX$$

disregarding π -null sets. DX may be dropped on the right. On the other hand by virtue of theorem 8.1, $B \subset CX$. Therefore, the two sets B and CX can differ only by a π -null set. The third statement is hereby proved. The first and second statements are immediate consequences of the third and of lemma 8.2. Theorem 8.2 is herewith completely proved.

9. The conservative part CX . Invariant functions and sets. CX is the vital part as far as ergodic theory is concerned. Notions of invariance appear to have significance in this respect only as far as they are tied up with CX .

Definition 9.1. A bounded and $\mathfrak{F}$ -measurable function $h(x)$ in X is called L -invariant in CX if $Lh = h$ holds in CX except in some π -null set. A subset $A \in \mathfrak{F}$ of CX is called invariant (under the Markoff process) if its characteristic function $e_A(x)$ is L -invariant in CX or, equivalently, if $P(x, A) = e_A(x)$ holds for π -almost all $x \in CX$ (see 4.4)).

Any constant is L -invariant, even in X (property d of L).

Lemma 9.1. If g vanishes in CX minus some π -null set the same is true for Lg , i.e., g is L -invariant in CX . In other words, the values of Lf in CX do not depend on the values of f in DX .

Proof. From (5.5) and theorem 8.2 it follows that

$$(9.1) \quad \int_{CX} gL^*f \, d\pi = \int_{CX} fLg \, d\pi, \quad f = 0 \text{ in } DX,$$

holds for any bounded g . Lemma 9.1 is now obvious since f is arbitrary in CX . It is the dual form of the first statement in theorem 8.2.

For invariant functions this has the following immediate consequences.

Lemma 9.2. A function h in X which is L -invariant in CX retains this property if its values are altered arbitrarily in DX .

CX is an invariant set. This follows from the definition and from lemma 9.2 if the L -invariant function 1 is changed to zero within DX . We have, moreover,

$$(9.2) \quad Le_{CX} \begin{cases} = e_{CX} \text{ in } CX, \\ \geq e_{CX} \text{ in } DX. \end{cases}$$

The following lemma is not needed later:

Lemma 9.3. If f is ≥ 0 and π -integrable in X then the set

$$A = \{x \mid L_{\infty}^* f = \infty\}$$

is invariant.

Proof. In virtue of theorem 8.1, $A \subset CX$. By means of the duality relation the following dual equivalent of lemma 8.2 is readily proved. If $g = 0$ in A then $Lg = 0$ in A . Hence $L(1 - e_A) = 1 - Le_A = 0$ in A , or $Le_A = e_A$ in A . But $Le_A \geq 0 = e_A$ in $X - A$. Hence $Le_A \geq e_A$, and by lemma 9.2, $Le_A = e_A$ in CX .

The following lemma is basic for the theory of invariant functions and sets:

Lemma 9.4. If a bounded and $\mathfrak{F}$ -measurable function h in X satisfies the inequality $Lh \geq h$ in CX minus some π -null set then it is L -invariant in CX .

Proof. Formula (9.1) holds as well for the powers of L^* , L and their sums, L_n^* , L_n ,

$$(9.3) \quad \int_{CX} g L_n^* f \, d\pi = \int_{CX} f L_n g \, d\pi, \quad f = 0 \text{ in } DX.$$

Let

$$g = Lh - h, \quad f = 1 \text{ in } CX.$$

If $|h| \leq M$ then

$$|L_n g| = |L^{n+1} h - h| \leq L^{n+1} |h| + |h| \leq 2M.$$

The left hand side of (9.3) therefore stays bounded as $n \rightarrow \infty$. But, by hypothesis, $g \geq 0$ in CX and, by the last part of theorem 8.2, $L_n^* f \rightarrow \infty$ holds in CX . Consequently, within CX , g can be $\neq 0$ only in a π -null set. q.e.d.

If a function h is L -invariant in CX so is $|h|$. In fact, $h = Lh$ in CX and the positiveness and linearity of L imply that

$$|h| = |Lh| \leq L|h|$$

holds in CX . Familiar arguments lead from here and from the linearity of the space of L -invariant functions to the fact that any function (within a wide class) of an L -invariant function in CX is again L -invariant in CX . The proof also makes use of the following theorem which is most easily derived from (5.5). If $h_n \uparrow h$ where all these functions are bounded and f -measurable then $Lh_n \uparrow Lh$. If h is L -invariant in CX so is $F(h)$, where

$$F(u) = \begin{cases} 1, & u > c, \\ 0, & u \leq c. \end{cases}$$

This function $F(h(x))$ is the characteristic function of the set $\{x \mid h(x) > c\}$. Within CX it is the characteristic function of the intersection $CX \cdot \{x \mid h > c\}$. This latter characteristic function regarded as a function in X differs from $F(h(x))$ only within DX . By virtue of lemma 9.2 it is therefore L -invariant in CX . Hence, if $h(x)$ is L -invariant within CX then every set $CX \cdot \{x \mid h > c\}$ is an invariant set.

Theorem 9.1. A bounded and $\mathfrak{F}$ -measurable function $h(x)$ in X is L -invariant in CX if and only if every set

$$CX \cdot \{x \mid h > c\}$$

is an invariant set.

Proof. It remains to prove the “if”. We suppose that everyone of the sets mentioned is invariant. This means that their characteristic functions are L -invariant in CX . Therefore, by lemma 9.2, the characteristic functions of the sets $\{x \mid h > c\}$ are L -invariant in CX . But h is monotone limit in X of linear combinations of such characteristic functions, *i.e.*, of L -invariant functions in CX . Therefore, h itself is L -invariant in CX . Theorem 9.1 is hereby proved.

The functions h which are L -invariant in CX form an algebraic field: sum, product and quotient (with a denominator bounded away from zero) of two such functions are again such functions. All this follows by familiar reasoning. With h , $F(h) = h^2$, $F(h) = 1/h$ are L -invariant in CX , and $2hg = (h + g)^2 - (h - g)^2$.

The invariant sets form a countably additive field of subsets of CX including CX itself. This is a consequence of the immediately preceding statement concerning invariant functions and of the continuity property of L mentioned above.

We conclude this section with the following new

Theorem 9.2. If h is L -invariant in CX then, for any bounded and f -measurable function $w(x)$ in X ,

$$L(hw) = hLw$$

holds in CX minus some π -null set, and, for any π -integrable $w'(x)$ in X which vanishes in DX ,

$$L^*(hw') = hL^*w'$$

holds in X minus some π -null set.

Proof. To prove the first relation we need only consider characteristic functions

$$w = e_A, \quad h = e_I$$

h again being L -invariant in CX . The relation becomes then

$$L(e_I e_A) = Le_{IA} = e_I Le_A$$

or, in virtue of (4.4),

$$(9.4) \quad P(x, IA) = e_I(x)P(x, A).$$

Before we prove it we give it another equivalent form. Let $I' = X - I$. On substituting $I = X - I'$ and on using the additivity of P we find that (9.4) is equivalent to

$$(9.5) \quad P(x, I'A) = e_{I'}(x)P(x, A).$$

Every x which satisfies one relation satisfies the other one. It suffices therefore to prove (9.4) for almost all $x \in I'$ and εCX . Now we make use of the hypothesis that e_I is L -invariant in CX . This means (see (4.4)) that

$$P(x, I) = e_I(x)$$

for almost all $x \in CX$. If as assumed $x \in I'$ and εCX we must have $P(x, I) = 0$ for almost all these x . Hence both sides of (9.4) are zero for these x . The first relation is thereby proved. The second relation is obtained from the first by means of the duality relation (5.5). In fact,

$$\int_X w' L(hw) d\pi = \int_X w' hLw d\pi$$

holds for arbitrary w, w' as stated, $w' = 0$ in DX . The left hand side equals

$$\int_X L^*(w')hw d\pi$$

whereas the right hand side is equal to

$$\int_X L^*(w'h)w d\pi.$$

Their equality for an arbitrary w implies the second relation. Theorem 9.2 is hereby proved.

10. On the ergodic theorem. Markoff processes with an invariant distribution.

Suppose that π is a finite measure in $\mathfrak{F}$ and that L^*f is an operator with the properties $a^*)-d^*)$, §5. Let the two functions f, p be π -integrable in X , $p > 0$. Is it then true that

$$(10.1) \quad \lim_{n \rightarrow \infty} \frac{\sum_0^{n-1} L^{*\nu} f(x)}{\sum_0^{n-1} L^{*\nu} p(x)}$$

exists for π -almost every $x \in X$? Or is it only true under an additional restrictive assumption? The existence of the limit within the dissipative part DX is trivial and uninteresting and it is in CX where the real problem lies.

We confine ourselves in this paragraph to a Markoff process with a given invariant distribution π ,

$$(10.2) \quad \Lambda\pi = \pi.$$

The process is supposed to be π -conditioned (Λ defined in ϕ_π). Together with (10.2) this implies π -measurability ($\Lambda\phi_\pi \subset \phi_\pi$). The density operator L^*f with respect to π has, in consequence of (5.2) and (5.3), the property that

$$(10.3) \quad L^*1 = 1$$

holds for π -almost all $x \in X$. We may just as well start with an operator L^* which has, in addition to the properties $a^*)-d^*)$, §5, the property (10.3). The Markoff operator $\Lambda\phi$ determined by L^* and π has then automatically the property (10.2).

In the case of an invariant distribution π the two operators L^*, L are fully symmetric in their basic properties. Not only does L^* share the property $d)$ with L but L shares also the property $d^*)$ with L^* ,

$$(10.4) \quad \int_X Lg \, d\pi = \int_X g \, d\pi.$$

This follows from (5.5) and (10.2). The only asymmetry left is in the original regions of definition $\mathfrak{B}_\pi, \mathfrak{L}_\pi$ for L, L^* respectively. However, it is possible to extend L with all its basic properties $a)-d)$, (10.4) from $\mathfrak{B}_\pi$ to $\mathfrak{L}_\pi$,

$$L\mathfrak{L}_\pi \subset \mathfrak{L}_\pi.$$

In order to keep the duality relation (5.5) valid the region of definition $\mathfrak{L}_\pi$ of L^* must, of course, be narrowed down to $\mathfrak{B}_\pi$.

L may be regarded as the density operator of another Markoff operator Λ^* with respect to π ,

$$\Lambda^* \gamma(A) = \int_A Lg(x) \pi(dx), \quad g = \frac{d\gamma}{d\pi},$$

which leaves π again invariant. We may write

$$\Lambda^* \gamma(A) = \int_X P^*(x, A) \gamma(dx).$$

P^* is what Kakutani calls "the inverse Markoff process." On letting $g = e_B$ and on using (4.4), $f = e_B$, we obtain from the last two formulae Kakutani's relation

$$\int_A P(x, B) \pi(dx) = \int_B P^*(x, A) \pi(dx),$$

$A \in \mathfrak{F}$, $B \in \mathfrak{F}$. Of course, "Inverse" does not mean that the operators Λ , Λ^* or L^* , L are inverses of each other.

We turn our attention to the ergodic theorem in the present case and we let $p = 1$ in (10.1).

Ergodic theorem. Let π be a finite measure in $\mathfrak{F}$ (remember that $X \in \mathfrak{F}$ and hence $\pi(X) < \infty$) and let L^* be an operator from $\mathfrak{L}_\pi$ to $\mathfrak{L}_\pi$ which has, in addition to a^* - d^*), §5, the property (10.3). Then, if $f(x)$ is π -integrable in X ,

$$(10.5) \quad \lim_{n \rightarrow \infty} \frac{1}{n} L_n^* f(x), \quad L_n^* = \sum_0^{n-1} L^{*n},$$

exists π -almost everywhere in X . The limit function $\bar{f}(x)$ is again π -integrable and simultaneously L^* - and L -invariant in X . The relation

$$(10.6) \quad \int_X f h \, d\pi = \int_X \bar{f} h \, d\pi$$

holds for every bounded and L -invariant $h(x)$ in X .

Proof. Note first that $DX = 0$. In fact, this follows from (10.3), $L_\infty^* 1 = \infty$ almost everywhere in X , and from theorem 8.1.

First we show that (10.5) exists almost everywhere in X for each function $f(x)$ out of a set of functions which is dense in $\mathfrak{L}_\pi$ in the sense of the metric

$$(10.7) \quad \|f\| = \int_X |f| \, d\pi.$$

We assert that the functions f of the special form

$$(10.8) \quad f(x) = L^* F(x) - F(x) + G(x),$$

where F and G are bounded in X and where

$$(10.9) \quad LG(x) = G(x)$$

in X , are dense in $\mathfrak{L}_\pi$. This can be shown in various ways.

$$\int_X (L^*F - F)G \, d\pi \equiv \int_X F(LG - G) \, d\pi$$

implies that a function G is orthogonal to all functions $L^*F - F$ if and only if $LG = G$. Hence the closures in the Hilbert space $\mathfrak{L}_\pi^2$ of the two linear spaces $\{G\}$, $LG = G$, and $\{L^*F - F\}$ are orthogonal complements of each other in $\mathfrak{L}_\pi^2$. The functions (10.8) are therefore dense in $\mathfrak{L}_\pi^2$ in the sense of the Hilbert metric and, *a fortiori*, in the sense of (10.7). But $\mathfrak{L}_\pi^2$ is dense in $\mathfrak{L}_\pi$ in the sense of (10.7). This proves the assertion. Now we show that (10.5) exists for every f of the form (10.8). In fact, (10.9) implies by theorem 9.2, $w' = 1$, that $L^*G = G$ holds almost everywhere in X . Hence,

$$\frac{1}{n} L_n^* f = \frac{1}{n} (L^{*n} F - F) + G$$

holds in X minus some π -null set which can be chosen independent of n . Since F is bounded, $|F| < M$, it follows that

$$|L^{*n} F| \leq L^{*n} |F| \leq M L^{*n} 1 = M.$$

Hence (10.5) equals G almost everywhere.

To prove the existence of (10.5) almost everywhere for any π -integrable $f(x)$ we have to make use of the following theorem of Banach [1]. Let $L_n f$ be a sequence of continuous linear operators from $\mathfrak{L}_\pi$ to the linear space of $\mathfrak{F}$ -measurable functions $f'(x)$ in X with the metric

$$d(f', g') = \int_X \frac{|f' - g'|}{1 + |f' - g'|} \, d\pi.$$

If

$$(10.10) \quad \lim_{n \rightarrow \infty} L_n f(x)$$

exists almost everywhere for each $f(x)$ out of a set of functions dense in $\mathfrak{L}_\pi$ and if

$$(10.11) \quad \sup_n L_n f(x) < \infty$$

holds almost everywhere for every $f \in \mathfrak{L}_\pi$, then (10.10) exists almost everywhere for every $f \in \mathfrak{L}_\pi$.

In the present case, $L_n = L_n^*/n$, the continuity of L_n follows easily from the fact that L^* preserves the norm (10.7). The existence of (10.10) in a dense subset of $\mathfrak{L}_\pi$ has just been proved. But (10.11) we can only infer from the maximal ergodic theorem (see the proof of theorem 8.1). The existence of (10.5) almost everywhere in X is thereby proved for every π -integrable $f(x)$.

The proof of the remainder of the theorem is outlined only quite briefly. It is here that we have to use the extension of the dual operator L from $\mathfrak{B}_\pi$ to $\mathfrak{L}_\pi$. The extension of the notion of L -invariance in X to π -integrable functions $h(x)$

in X is straightforward. Lemma 9.4 and its important consequences, theorems 9.1 and 9.2, stay valid with obvious modifications. However, the proof of the lemma uses the boundedness of h , and it must therefore be modified. In the present case, π being Λ -invariant, $L^* 1 = 1$, it becomes almost trivial. Let $g = Lh - h \geq 0$ and $f = 1$ in X ($CX = X$, $DX = 0$). We have

$$\int_X g \, d\pi = \int_X 1(Lh - h) \, d\pi = \int_X h(L^* 1 - 1) \, d\pi = 0$$

and hence $g = 0$ in X . Lemma 9.4 is hereby proved. Every π -integrable and L -invariant $h(x)$ in X is automatically L^* -invariant. This follows from theorem 9.2 on letting $w' = 1$. The converse is also true, as follows from the symmetry of L and L^* which is valid here.

The assertions about the limit function $\tilde{f}$ of (10.5) can be proved in various known ways. They are easiest derived from the mean ergodic theorem in $\mathfrak{L}_\pi$ which Kakutani [6] deduced from the individual ergodic theorem for bounded functions: for every $f \in \mathfrak{L}_\pi$ there exists an $\tilde{f} \in \mathfrak{L}_\pi$ such that

$$\left\| \frac{1}{n} \sum_0^{n-1} L^{*v} f - \tilde{f} \right\| \rightarrow 0.$$

$\tilde{f}(x)$ is necessarily the limit function of our individual ergodic theorem because a sequence which converges in the mean contains a subsequence which converges almost everywhere. $L^* \tilde{f} = \tilde{f}$ follows easiest from the mean ergodic theorem. Strong convergence implies weak convergence and, therefore,

$$\int_X f \, d\pi = \int_X \tilde{f} \, d\pi.$$

If h is L^* -invariant and bounded then we consider hf instead of f . By theorem 9.2, $L^{*v}(hf) = hL^{*v}f$. Hence $\tilde{f}$ goes over into $h\tilde{f}$. (10.6) and the ergodic theorem are thereby proved.

11. On the influence of the dissipative part. We suppose that the Markoff process is π -measurable. Invariance of π is no longer assumed. Consider the contractions of the operator Λ into CX and DX ,

$$\Lambda_C \varphi = (\Lambda \varphi)_{CX}, \quad \Lambda_D \varphi = (\Lambda \varphi)_{DX}$$

or the corresponding density operators

$$L_C^* f = e_{CX} L^* f, \quad L_D^* f = e_{DX} L^* f.$$

We have

$$L^* = L_C^* + L_D^*$$

and, by virtue of theorem 8.2,

$$L_D^* L_C^* = 0.$$

In general, however, $L_C^* L_D^*$ does not vanish. The operator

$$(11.1) \quad g_n f = L_C^* L_{D,n}^* f, \quad L_{D,n}^* = \sum_0^{n-1} L_D^{*\nu},$$

measures, physically speaking, the contribution of φ or f to CX accumulated from DX in n successive trials. We wish to examine its limit properties as $n \rightarrow \infty$.

From

$$\begin{aligned} L_C^* L_{D,n}^* f &= L^* L_{D,n}^* f - L_D^* L_{D,n}^* f \\ &= L^* L_{D,n}^* f - L_{D,n+1}^* f + f \end{aligned}$$

we find, on using d^*), §5, that

$$\begin{aligned} \int_X g_n f d\pi &= \int_X L_{D,n}^* f d\pi - \int_X L_{D,n+1}^* f d\pi + \int_X f d\pi \\ &= \int_X f d\pi - \int_X L_D^{*n} f d\pi. \end{aligned}$$

We transform the second integral on the right by means of successive application of the duality relation and on using the formula $e_{DX} L^\nu e_{DX} = L^\nu e_{DX}$, $\nu > 0$, which follows easily from (9.2). It follows that

$$\begin{aligned} \int_X g_n f d\pi &= \int_X f d\pi - \int_X f L^n e_{DX} d\pi \\ &= \int_X f(1 - L^n e_{DX}) d\pi = \int_X f(L^n e_{CX}) d\pi. \end{aligned}$$

The quantity in parentheses in the integral on the right always lies between 0 and 1. Hence,

$$(11.2) \quad f \geq 0 \text{ in } X \rightarrow \int_X g_n f d\pi \leq \int_X f d\pi.$$

Therefore

$$g_\infty f = \lim_{n \rightarrow \infty} L_C^* \sum_0^{n-1} L_D^{*\nu} f$$

exists π -almost everywhere and is π -integrable in X . This is consequently true for any π -integrable f , irrespective of sign. The operator g_∞ from $\mathfrak{L}_\pi$ to $\mathfrak{L}_\pi$ is thus

positive, linear and of $\mathfrak{L}_\pi$ -norm ≤ 1 . It indicates the accumulated contribution of the dissipative to the conservative part.

The following problem arises here. Does

$$\lim_{n \rightarrow \infty} \frac{\sum_0^{n-1} L^{*\nu}(f - f')}{\sum_0^{n-1} L^{*\nu} p} = 0$$

hold π -almost everywhere in CX whenever p and f are π -integrable in X , $p > 0$ and

$$f' = \mathfrak{g}_\infty f ?$$

If the answer were yes this would show that the limit behavior of a quotient

$$q_n = \frac{\sum_0^{n-1} L^{*\nu} f}{\sum_0^{n-1} L^{*\nu} p}$$

in CX would not be affected if f and p would be replaced by $f' = \mathfrak{g}_\infty f$ and $p' = \mathfrak{g}_\infty p$ — p' is > 0 in CX —respectively. As $f' = p' = 0$ holds in DX , a satisfactory separation of the conservative part from the dissipative one would be achieved in this way, satisfactory at least as far as the limit behavior of the quotients q_n is concerned.

REFERENCES

1. ST. BANACH, *Sur la convergence presque partout des fonctionnelles linéaires*, Bull. Sci. Math. vol. **50** (1926), pp. 36–43. S. MAZUR & W. ORLICZ, *Über Folgen linearer Operationen*, Studia Math. vol. **4** (1933), pp. 152–157. The proof of these two authors is also found in the following paper: N. DUNFORD & D. S. MILLER, *On the ergodic theorem*, Trans. Amer. Math. Soc. vol. **60** (1946), pp. 538–549.
2. J. L. DOOB, *Stochastic processes with an integral-valued parameter*, Trans. Amer. Math. Soc. vol. **44** (1938), pp. 87–150.
3. ———, *Asymptotic properties of Markoff transition probabilities*, Trans. Amer. Math. Soc. vol. **63** (1948), pp. 393–421.
4. Y. N. DOWKER, *A new proof of the general ergodic theorem*, Acta. Sci. Math. Szeged vol. **12** (1950), pp. 162–166.
5. P. R. HALMOS, *An ergodic theorem*, Proc. Nat. Acad. Sci. U. S. A. vol. **32** (1946), pp. 156–161.
6. S. KAKUTANI, *Some results in the operator-theoretical treatment of the Markoff process*, Proc. Imp. Acad. Tokyo vol. **15** (1939), pp. 260–264.
7. ———, *Ergodic theorems and the Markoff process with a stable distribution*, Proc. Imp. Acad. Tokyo vol. **16** (1940), pp. 49–54.
8. ———, *Ergodic theory*, Proc. Internat. Congress of Math. 1950 vol. II, pp. 128–142.
9. K. YOSIDA, *Operator-theoretical treatment of the Markoff process*, Proc. Imp. Acad. Tokyo vol. **14** (1938), pp. 363–367.
10. ———, *Operator-theoretical treatment of Markoff's process II*, Proc. Imp. Acad. Tokyo vol. **15** (1939), pp. 127–130.
11. ———, *The Markoff process with a stable distribution*, Proc. Imp. Acad. Tokyo vol. **16** (1940), pp. 43–48.
12. K. YOSHIDA & S. KAKUTANI, *Birkhoff's ergodic theorem and the maximal ergodic theorem*, Proc. Imp. Acad. Tokyo vol. **15** (1939), pp. 165–168.
13. ———, *Operator-theoretical treatment of Markoff process and mean ergodic theorem*, Ann. of Math. vol. **42** (1941), pp. 188–228.