PAPERS COMMUNICATED

12. *The Markoff Process with a Stable Distribution.*

By Kôsaku YOSIDA.

Mathematical Institute, Osaka Imperial University.

(Comm. by T. TAKAGI, M.I.A., March 12, 1940.)

§1. *Introduction and the theorems.* Let R be a space and let $B(R)$ be a completely additive family of "measurable" subsets of R. We assume that R itself belongs to $B(R)$. Let $P(x, E)$ denote the transition probability that the point $x \in R$ is transferred, by a simple Markoff process, into the set $E \in B(R)$ after the elapse of a unit-time. We naturally assume that $P(x, E)$ is completely additive for $E \in B(R)$ if x is fixed, and that $P(x, E)$ is "measurable" (with respect to $B(R)$) in x if E is fixed. Then the transition probability $P^{(n)}(x, E)$ that the point x is transferred into E after the elapse of n unit-times is given by $P^{(n)}(x, E) = \int P^{(n-1)}(x, dy) P(y, E)$[1] $\big(n=1, 2, \ldots; P^{(1)}(x, E) = P(x, E)\big)$. We have surely,

$$(1) \qquad P^{(n)}(x, E) \geqq 0, \quad P^{(n)}(x, R) \equiv 1 \qquad (n=1, 2, \ldots).$$

We now assume that there exists a non-negative set function $\varphi(E)$ which satisfies the conditions:

(2) $\varphi(E)$ is completely additive for $E \in B(R)$ and $\varphi(R)=1$.

(3) The space $B(R)$, if metrised by the distance $d(E_1, E_2) = \varphi(E_1 + E_2 - E_1 \cdot E_2)$, is (complete and) separable.[2]

(4) $\int \varphi(dx) P(x, E) = \varphi(E)$ for any $E \in B(R)$.

We have, from (4), $\int \varphi(dx) P^{(n)}(x, E) = \varphi(E)$ for any n. Hence the mass distribution $\varphi(E)$ is *stable* with respect to the time. Such Markoff process (*with a stable distribution* $\varphi(E)$) is fairly general; it includes the *deterministic transition process* in the ergodic theory of the incomplessible stationary flow, originated by G. D. Birkhoff and J. von Neumann. In fact, let T be a one-to-one point transformation of R on R which maps any set $E \in B(R)$ on the set $T \cdot E \in B(R)$ in measure-preserving way: $\varphi(E) = \varphi(T \cdot E)$. Let $C_E(x)$ be the characteristic function of E and put $P(x, E) = C_E(T \cdot x)$, then it is easy to see that this $P(x, E)$ defines a Markoff process with stable distribution $\varphi(E)$. Another example is given by the Markoff process with *symmetric* φ-density: $P(x, E) = \int_E p(x, y) \varphi(dy)$, $p(x, y) \equiv p(y, x)$. Thus our general

---

1) The definite integral over R will be denoted by $\int \varphi(dx)$.

2) The completeness of the metrical space $B(R)$ follows from the complete additivity of $B(R)$. The separability hypothesis may be taken away, by suitably modifying the proof below. However, for the sake of brevity, I here assume it.

$P(x, E)$ applies to the deterministic transition process $x \rightarrow T\cdot x$ as well as to the *indeterministic* (*probabilistic*) *transition process.*

Let (L) denote the set of all the real-valued measurable functions $f(x)$ which are φ-integrable. (L) is surely a Banach space with the norm $\|f\|=\int|f(x)|\,\varphi(dx)$. We will prove the following theorems.

Theorem 1. For any $f\in(L)$,

$$f^{(m)}(x)=\int P^{(m)}(x, dy) f(y) \qquad (m=1, 2, \ldots)$$

exists φ-*almost everywhere and belongs to* (L) *in such a way that* $\|f^{(m)}\|\leqq\|f\|$. *For any* $f\in(L)$ *and for any* m $(=1, 2, \ldots)$, *there corresponds* $f^{(-m)}\in(L)$, $\|f^{(-m)}\|\leqq\|f\|$, *such that*

$$\int\varphi(dx) f(x) P^{(m)}(x, E)=\int_E f^{(-m)}(x)\,\varphi(dx) \quad \textit{for all} \quad E\in B(R).$$

Theorem 2. For any $f\in(L)$, *there corresponds* $f^{(*)}$, $f^{(-*)}\in(L)$ *such that*

$$(5) \qquad \begin{cases} \lim\limits_{n\to\infty}\displaystyle\int\left|f^{(*)}(x)-\frac{1}{n}\sum_{m=1}^{n}f^{(m)}(x)\right|\varphi(dx)=0\,, \\ f^{(*)}(x)=\displaystyle\int P(x, dy) f^{(*)}(y) \quad \varphi\textit{-almost everywhere,} \end{cases}$$

$$(6) \qquad \begin{cases} \lim\limits_{n\to\infty}\displaystyle\int\left|f^{(-*)}(x)-\frac{1}{n}\sum_{m=1}^{n}f^{(-m)}(x)\right|\varphi(dx)=0\,, \\ \displaystyle\int\varphi(dx) f^{(-*)}(x) P(x, E)=\int_E f^{(-*)}(x)\,\varphi(dx) \quad \textit{for any} \quad E\in B(R)\,. \end{cases}$$

In the deterministic case $P(x, E)=C_E(T\cdot x)$, $f^{(m)}(x)$ and $f^{(-m)}(x)$ corresponds to $f(T^m\cdot x)$ and $f(T^{-m}\cdot x)$ respectively. Theorem 2 is an extension of J. von Neumann's mean ergodic theorem for (L) to the indeterministic case. *Thus the problem of the Markoff process is reduced, in a certain sense, to the determination of all the possible stable distributions.* Under some topological assumptions, such determination is possible. This will be indicated in § 3. As a physical application of the Markoff process with stable distribution, an interpretation may be given to the H-theorem of the statistical mechanics.

Remark. On reading the manuscript, S. Kakutani obtained that,[3] *if* $f(x)$ *is bounded and measurable,* the mean convergence in (5) may be replaced by the φ-almost-everywhere convergence, by modifying J. L. Doob's arguments.[4] This is more precise than (5), since from this result and the theorem 1 (5) may easily be deduced.[5] It is noted, however, that Doob did not obtain the mean convergence (5), nor its "conjugate" (6). The Doob-Kakutani's proof appeals to Birkhoff-Khintchine's ergodic theorem and to the measure theory in infinite product space. Our proof is operator-theoretical; it is a direct adaptation of the mean ergodic theorem in Banach spaces.

---

3) See the following paper of S. Kakutani.

4) Trans. Amer. Math. Soc., **44** (1938), 87-150.

5) See S. Kakutani: loc. cit.

§ 2. *Proof of the theorems.* The set (M') of all the bounded measurable functions $f(x)$ constitutes a Banach space with the norm $\|f\|_{M'}=\sup|f(x)|$. In the same way the set (M) of all the measurable functions $f(x)$ which are φ-essentially bounded also constitutes a Banach space with the norm $\|f\|_M=\text{ess. sup}\,|f(x)|$. Any element $f\in(M')$ $\big((M)\big)$ may be considered as an element $\in(L)$, and (M') $\big((M)\big)$ is dense in (L) by the topology defined by the norm $\|\ \ \|$.

Proof of theorem 1. Let $f\in(M')$, then $f^{(m)}(x)$ is well defined and $\|f^{(m)}\|_{M'}\leqq\|f\|_{M'}$ by (1). We have $\|f^{(m)}\|=\int|f^{(m)}(x)|\,\varphi(dx)\leqq\int\varphi(dx)$ $\left\{\int P^{(m)}(x,dy)\,|f(y)|\right\}=\|f\|_{M'}$ by (4). Next let $f\in(L)$ be non-negative and put $f_n(x)=\min.\big(f(x),n\big)$. Then, since $f_n(x)$ is a bounded function, $f_n^{(m)}(x)$ is well defined and $f_n^{(m)}(x)\leqq f_{n+1}^{(m)}(x)\leqq\cdots$, $\|f_n^{(m)}\|\leqq\|f_n\|\leqq\|f\|$. Hence by Levi's theorem, the finite

$$\lim_{n\to\infty}f_n^{(m)}(x)=\lim_{n\to\infty}\int P^{(m)}(x,dy)\,f_n(y)=\int P^{(m)}(x,dy)\{\lim_{n\to\infty}f_n(y)\}=\int P^{(m)}(x,dy)\,f(y)$$

exists φ-almost everywhere. Thus $f^{(m)}(x)=\lim_{n\to\infty}f_n^{(m)}(x)\in(L)$ and $\|f^{(m)}\|\leqq\|f\|$. This proves the first part of the theorem.

Since $P^{(m)}(x,E)$ is a bounded function, $\int\varphi(dx)f(x)P^{(m)}(x,E)$ exists for any $f\in(L)$ and we have from (4)

$$\text{(7)}\qquad \sup_E\int\varphi(dx)f(x)P^{(m)}(x,E)\leqq\sup_E\int\varphi(dx)\,|f(x)|\,P^{(m)}(x,E)\leqq\|f\|\,.$$

Let $f\in(M)$, then by (4), $\int\varphi(dx)f(x)P^{(m)}(x,E)$ is φ-absolutely continuous: $\left|\int\varphi(dx)f(x)P^{(m)}(x,E)\right|\leqq\|f\|_M\cdot\varphi(E)$, and hence, in this case, $f^{(-m)}$ is defined and $\|f^{(-m)}\|_M\leqq\|f\|_M$. Next let $f\in(L)$ be non-negative and put $f_n(x)=\min.\big(f(x),n\big)$ as above. Since $\int\varphi(dx)f(x)P^{(m)}(x,E)=\lim_{n\to\infty}\int\varphi(dx)f_n(x)P^{(m)}(x,E)$ for any $E\in B(R)$, the left hand member is φ-absolutely continuous, by Vitali-Hahn-Saks' theorem. Thus, in this case also $f^{(-m)}\in(L)$ is defined. By (7), we have $\|f^{(-m)}\|\leqq\|f\|$. The last part of theorem is hereby proved.

Remark. The operation $P^m:f\to f^{(m)}(\bar{P}^m:f\to f^{(-m)})$ is thus a linear operation on (L) to (L) with the norm one. $P^m(\bar{P^m})$ also defines an operation on (M') to (M') $\big($on (M) to $(M)\big)$ with the norm one. This remark is the key to the proof of theorem 2.

Proof of theorem 2. Let $f\in(M')$, then by (1)

$$\text{(8)}\qquad \left\|\frac{1}{n}\sum_{m=1}^{n}f^{(m)}\right\|_{M'}\leqq\|f\|_{M'}\qquad(n=1,2,\ldots).$$

Thus, by (3) and the diagonal method, we obtain a sequence of integers $\{n'\}$, $n' < (n+1)'$, such that

$$\lim_{n'\to\infty}\int_E\left\{\frac{1}{n'}\sum_{m=1}^{n'}f^{(m)}(x)\right\}\varphi(dx)=F^{(*)}(E)\quad\text{exists for all } E\in R(R).$$

By (8) and Vitali-Hahn-Saks' theorem, $F^{(*)}(E)$ is φ-absolutely continuous. Let $F^{(*)}(E)=\int_E f^{(*)}(x)\,\varphi(dx)$, $f^{(*)}\in(L)$. Then it is easy to see that $f^{(*)}$ is the weak limit of the sequence $\left\{\frac{1}{n'}\sum_{m=1}^{n'}f^{(m)}\right\}$: for any $g\in(M)$ we have $\lim_{n'\to\infty}\int\left\{\frac{1}{n'}\sum_{m=1}^{n'}f^{(m)}(x)\right\}g(x)\,\varphi(dx)=\int f^{(*)}(x)\,g(x)\,\varphi(dx)$. Hence by the mean ergodic in Banach spaces,[6] we see that the sequence $\left\{\frac{1}{n}\sum_{m=1}^{n}f^{(m)}\right\}$ itself converges strongly (in the sense of the norm $\|\;\|$) to $f^{(*)}$. Since the sequence of operators $\left\{\frac{1}{n}\sum_{m=1}^{n}P^m\right\}$ are all of (L)—norm $\leqq 1$ and since (M') is strongly dense in (L), $\left\{\frac{1}{n}\sum_{m=1}^{n}P^m\cdot f\right\}$ converges strongly for any $f\in(L)$. That $P\cdot f^{(*)}=f^{(*)}$ is proved at the same time. Thus (5) is proved. The proof of (6) may be obtained analogously.

§ 3. *The determination of the stable distributions.* Let R be a compactum and let $B(R)$ denote the family of all the Borel sets of R. We assume the Markoff process to satisfy the continuity hypothesis:

(9) $\left\{\begin{array}{l}\int P(x,dy)f(y) \text{ is a continuous function of } x \text{ for continuous} \\ \text{function } f(y).\end{array}\right.$

For such Markoff process $P(x,E)$ we may prove the following results.

Let (C) denote the Banach space of all the real-valued continuous functions $f(t)$ on R with the norm $\|f\|_C=\sup|f(x)|$. Then, i) the Markoff process $P(x,E)$ satisfying (9) may be characterised as the linear operation P on (C) to (C) which maps positive element f ($f(x)\geqq 0$ for $x\in R$) onto positive element $P\cdot f$ of (C) in such a way that $P\cdot f=f$ if $f(x)\equiv 1$. ii) Since (C) is separable and since we have (1), the sequence of distributions of the form

$$\left\{\begin{array}{l}P_{n_i,m_i}(x_0,E)=\dfrac{1}{m_i-n_i}\displaystyle\sum_{k=n_i+1}^{m_i}P^{(k)}(x_0,E)\quad(i=1,2,\dots;\ \lim_{i\to\infty}(m_i-n_i)=\infty)\\ \text{for fixed } x_0,\end{array}\right.$$

is compact as a set of continuous linear functionals on (C). That is, there exists a partial sequence $\{P_{n_i',m_i'}(x_0,E)\}$ and a set function $\tilde{P}(x_0,E)$

6) K. Yosida: Proc., **14** (1938), 292–294, S. Kakutani: Proc., **14** (1938), 295–300 and F. Riesz: Journal of London Math. Soc., **13** (1938), 274–278. That mean ergodic theorem is valid for individual point of the Banach space was stressed on by the writer in the preceding note: Proc., **15** (1939), 255–259.

completely additive for $E \in B(R)$ such that $\int \tilde{P}(x_0, dy) f(y) = \lim_{i\to\infty} \int P_{n_i', m_i'}(x_0, dy) f(y)$ for any $f \in (C)$. Such limit distribution $\tilde{P}(x_0, E)$ is stable, since by (9) we have

$$\int \left\{ \int \tilde{P}(x_0, dy) P(y, dz) \right\} f(z) = \int \tilde{P}(x_0, dy) \left\{ \int P(y, dz) f(z) \right\}$$
$$= \lim_{i\to\infty} \int P_{n_i', m_i'}(x_0, dy) \left\{ \int P(y, dz) f(z) \right\}$$
$$= \lim_{i\to\infty} \int P_{n_i', m_i'}(x_0, dz) f(z) = \int \tilde{P}(x_0, dz) f(z)$$

for any $f \in (C)$. iii) Let $\varphi(E)$ be any stable distribution and let a sequence of elements $f_1, f_2, \dots \in (C)$ be dense in (C). By (5) and the diagonal method, we may find a sequence of integers $\{n'\}$ and a set $E_0 \in B(R)$ of φ-measure zero such that $\lim_{n'\to\infty} \frac{1}{n'} \sum_{m=1}^{n'} f_i^{(m)}(x)$ exists for $x \in R - E_0$, $i = 1, 2, \dots$ Hence $\lim_{n'\to\infty} \int P_{1, n'}(x, dy) f(y)$ exists for any $f \in (C)$ and for any $x \in R - E_0$. Put $\lim_{n'\to\infty} \int P_{1, n'}(x, dy) f(y) = \int \tilde{P}(x, dy) f(y)$ for $x \in R - E_0$, $f \in (C)$. Then, since φ is stable,

$$\int \varphi(dy) f(y) = \lim_{n'\to\infty} \int \left\{ \int \varphi(dx) P_{1, n'}(x, dy) \right\}$$
$$f(y) = \int \varphi(dx) \left\{ \lim_{n'\to\infty} \int P_{1, n'}(x, dy) f(y) \right\} = \int \varphi(dx) \left\{ \int \tilde{P}(x, dy) f(y) \right\}$$
$$= \int \left\{ \int \varphi(dx) \tilde{P}(x, dy) \right\} f(y) \quad \text{for any} \quad f \in (C).$$

Thus it is proved that any stable distribution $\varphi(E)$ of $P(x, E)$ may be obtained as a "convex combination" of the limit distributions as $\tilde{P}(x, E)$. In this way we may extend Kryloff-Bogoliouboff's results[7] to the indeterministic transition process. It is to be noted that similar results may be obtained for Markoff process with enumerably infinite number of possible states. However this idea was partly and unconsciously used in a joint work with S. Kakutani.[8]

§ 4. *An interpretation of the H-theorem.* Let R be the interval $(0, 1)$ and let $B(R)$ be the set of all the Borel sets of R. We assume that $P(x, E)$ is given by Borel-measurable density in the following manner:

$$P(x, E) = \int_E p(x, y)\, dy, \quad \int p(x, y)\, dx \equiv 1. \tag{10}$$

The uniform distribution dx is thus stable for $P(x, E)$.

Let a non-negative $f(x)$ be such that $\int f(x)\, dx = 1$, $\int f(x) \overset{+}{\log} f(x)\, dx \neq$ infinite. Then we have

7) Ann. Math., **38** (1937), 65–113.
8) Jap. J. Math., **16** (1939), 47–55.

$$(11) \qquad \int H\big(f(y)\big)dy \geqq \int H\big(f^{(-1)}(y)\big)dy \geqq \int H\big(f^{(-2)}(y)\big)dy \geqq \cdots,$$

$$H(z) = z \log z.$$

The proof follows from (10) and the convexity of the function $H(z)$. If, moreover, we assume the measurability of $p^{(s)}(y) = \inf_x p^{(s)}(x, y)$ $\left(P^{(s)}(x, E) = \int p^{(s)}(x, y)\, dy\right)$ and $\int p^{(s)}(y)\, dy > 0$ for a certain s, we obtain

$$(12) \qquad \sup_{f(x)\geqq 0,\ |f|=1} \int |f^{(-n)}(y) - 1 \,|\, dy \leqq \frac{\delta}{(1+\varepsilon)^n} \qquad (n = 1, 2, \ldots),$$

for certain constants $\delta, \varepsilon > 0$. The proof may be obtained easily by modifying the *ergodic principle* of A. Kolmogoroff.[9]

The above result may serve as an interpretation of the H-theorem in the statistical mechanics. For, by $f(x) \geqq 0$ and $\int f(x)\, dx = 1$, $f(x)$ may be considered as the *weight* of the "state" x at the origin of the time. Then, *irrespective of the initial weight* $f(x)$, the weight $f^{(-n)}(x)$ of the state x after the elapse of n unit-times tends to the uniform weight $\left(f^{(-*)}(x) \equiv 1\right)$ as $n \to \infty$, in such a way that (6), (11) and (12) hold. The hypothesis (10) may be considered as a *generalised symmetric condition*, and the hypothesis concerning $p^{(s)}(y)$ amounts to a kind of *irreducibility* of the transition process $P(x, E)$. Thus the results may serve as a precision to the arguments of R. von Mises in his book.[10]

---

9) Math. Ann., **104** (1931), 415–458.

10) Wahrscheinlichkeitsrechnung (1931). For the discussion of this § I owe to K. Husimi and S. Kakutani.