

ON THE INTEGRO-DIFFERENTIAL EQUATIONS OF PURELY DISCONTINUOUS MARKOFF PROCESSES

BY
WILLY FELLER

1. Introduction. In the following we are concerned with stochastic processes depending on a continuous time parameter t , that is to say, with some entity (chance variable) whose state is specified by a point $X(t)$ varying in some space E according to some probability law. The process is called a *Markoff process*⁽¹⁾ if the probability distribution of $X(t)$ is completely determined for all $t > \tau$ by the knowledge of the state $X(\tau)$, and in particular is independent of the development of the process for $t < \tau$ ⁽²⁾. Analytically a Markoff process is completely determined by its transition probabilities $P(\tau, x; t, \Lambda)$, giving the conditional probability of $X(t)$'s being contained, at the moment t , in the set $\Lambda \subset E$ under the hypothesis that at a fixed moment $\tau < t$ the state $X(\tau)$ coincided with the point x of E .

In strict terms, we shall suppose that there is specified, in the space E , a Borel field \mathfrak{B} of sets (on which probabilities are defined) such that $E \in \mathfrak{B}$ and also any set consisting of a single point belongs to \mathfrak{B} . It is then required that $P(\tau, x; t, \Lambda)$ is, for fixed $\tau, t > \tau$ and $x \in E$, a non-negative and completely additive function of sets on \mathfrak{B} , with

$$(1) \quad P(\tau, x; t, E) \equiv 1.$$

Moreover, we shall always assume that for fixed τ, t, Λ the function $P(\tau, x; t, \Lambda)$ is measurable with respect to \mathfrak{B} , that is to say, that for any $a > 0$ the set where $P(\tau, x; t, \Lambda) < a$ belongs to \mathfrak{B} . Finally we shall, for the sake of simplicity, restrict ourselves to $P(\tau, x; t, \Lambda)$ depending, for fixed other arguments, continuously on both τ and t ⁽³⁾. This implies in particular that as either $t \rightarrow \tau + 0$ or $\tau \rightarrow t - 0$

Presented to the Society, February 24, 1940; received by the editors March 5, 1940.

⁽¹⁾ This name was chosen in accordance with the now common terminology in the case of processes with an integral-valued parameter t . Kolmogoroff [6] calls such processes stochastically definite, and this terminology I had also adopted previously. Markoff processes are, sometimes, also described as being "without after effect," or as being submitted to an "influence globale" (Pólya).

⁽²⁾ This is, of course, not meant to be a strict definition; as a matter of fact, we shall be concerned only with the function $P(\tau, x; t, \Lambda)$, which will be defined purely analytically.

⁽³⁾ It may be pointed out that this does not imply the continuity of the movement of $X(t)$. We shall, on the contrary, be concerned only with states $X(t)$ changing abruptly by jumps. The continuity of $P(\tau, x; t, \Lambda)$ means that the probability of a jump during a small time-interval is small.

$$(2) \quad P(\tau, x; t, \Lambda) \rightarrow \delta(x, \Lambda) = \begin{cases} 1 & \text{if } x \in \Lambda, \\ 0 & \text{otherwise.} \end{cases}$$

Subdividing now the interval (τ, t) by a point s and considering all possible states $X(s)$, we readily get the identity

$$(3) \quad P(\tau, x; t, \Lambda) \equiv \int_E P(\tau, x; s, dE_y) P(s, y; t, \Lambda)$$

known as the equation of Chapman and Kolmogoroff⁽⁴⁾. We shall take these relations as the analytic definition of a Markoff process and consider any $P(\tau, x; t, \Lambda)$ of the described sort as defining the transition probabilities of such a process⁽⁵⁾.

In the special case of the space E containing at most an enumerable number of points, we shall denote these points by x_k and write

$$(4) \quad P(\tau, x_i; t, x_k) = P_{ik}(\tau, t).$$

By (1) we have $\sum_k P_{ik}(\tau, t) \equiv 1$, while (2) and (3) are respectively equivalent to

$$(5) \quad P_{ik}(\tau, t) \rightarrow \delta_{ik} = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{if } i \neq k \end{cases} \quad (t - \tau \rightarrow +0),$$

and

$$(6) \quad P_{ik}(\tau, t) \equiv \sum_j P_{ij}(\tau, s) P_{jk}(s, t) \quad (\tau < s < t).$$

Now a *purely discontinuous process* may be described by the following property: if, at the moment t , the actual state is given by the point x , then there is a probability $1 - p(t, x)\Delta t + o(\Delta t)$ that no change of state will occur during $(t, t + \Delta t > t)$; and if a change occurs, the probability of $X(t + \Delta t)$'s being contained in the set Λ is given, except for terms of $o(1)$, by a probability distribution $\Pi(t, x, \Lambda)$ ⁽⁶⁾. In strict terms we shall say that *the Markoff process*

⁽⁴⁾ Cf. Kolmogoroff [6] where the foundations of the general theory of Markoff processes have been laid.

⁽⁵⁾ This is the natural point of view for the purposes of the present paper. From an axiomatic point of view, however, any stochastic process corresponds to a measure in the space of real functions defined on E . Even in the case of Markoff processes there are problems (especially the problem of ruin, playing an important rôle in the theory of risk) which require a deeper penetration in the theory of the functional space. For the treatment of stochastic processes in terms of measure, the reader is referred to J. C. Doob's fundamental paper [1].

⁽⁶⁾ Essentially this definition was given by Feller [3]; cf. also Dubrovski [2]. This kind of processes was mentioned also by Kolmogoroff [6].

Examples of such processes are furnished by the theory of radioactive processes and the theory of automatic telephone offices; by the transport of stones by rivers (treated by quite different methods of Pólya [8]); by the mathematical theory of struggle for life (Feller [4]), etc. Perhaps the most important application is furnished by the theory of risk.

There is no general definition for "purely continuous" processes in abstract E . In the Euclidean space such processes were defined by Kolmogoroff [6] and somewhat more generally by

defined by $P(\tau, x; t, \Lambda)$ is purely discontinuous if for small $t - \tau > 0$

$$(7) \quad P(\tau, x; t, \Lambda) = \{1 - p(t, x)(t - \tau)\} \delta(x, \Lambda) + p(t, x)(t - \tau) \Pi(t, x, \Lambda) + o(t - \tau),$$

where $\delta(x, \Lambda)$ was defined by (2) and the exact assumptions as to $p(t, x)$ and $\Pi(t, x, \Lambda)$ will be specified in §2, (i)–(ii); in general, $o(t - \tau)$ will depend on x and Λ .

The main problem with which we are confronted is to determine whether or not to any two functions $p(t, x)$ and $\Pi(t, x, \Lambda)$, subjected only to the conditions §2, (i)–(ii), there corresponds a Markoff process, whose transition probabilities $P(\tau, x; t, \Lambda)$ satisfy (7); and if so, whether this process is uniquely determined.

It will be shown (§2) that there is a subclass \mathfrak{B}_1 , of sets $\Lambda \in \mathfrak{B}$ such that for all $\Lambda \in \mathfrak{B}_1$, all τ , and almost all t the partial derivatives $\partial P(\tau, x; t, \Lambda) / \partial t$ and $\partial P(\tau, x; t, \Lambda) / \partial \tau$ exist; for all those Λ , all τ and almost all t , the integro-differential equations

$$(8) \quad \frac{\partial P(\tau, x; t, \Lambda)}{\partial t} = - \int_{\Lambda} p(t, y) P(\tau, x; t, dE_y) + \int_E p(t, y) \Pi(t, y, \Lambda) P(\tau, x; t, dE_y)$$

and

$$(9) \quad \frac{\partial P(\tau, x; t, \Lambda)}{\partial \tau} = p(\tau, x) \left\{ P(\tau, x; t, \Lambda) - \int_E P(\tau, y; t, \Lambda) \Pi(\tau, x, dE_y) \right\}$$

hold, implying the existence of the integrals for all $\Lambda \in \mathfrak{B}_1$, and almost all t . The class \mathfrak{B}_1 contains, among others, sequences of sets $\Lambda_n \uparrow E$.

Thus the problem is reduced to the integration of (8)–(9). It will be shown (§§3–5) that there is a function $P(\tau, x; t, \Lambda)$ which satisfies (8)–(9) for all $\Lambda \in \mathfrak{B}_1$ and almost all t and τ ; this $P(\tau, x; t, \Lambda)$ is uniquely determined by each of the equations (8)–(9)⁽⁷⁾ and has all properties described above, except perhaps (1); one has always

$$(10) \quad 0 \leq P(\tau, x; t, \Lambda) \leq 1,$$

but there are cases with

Feller [3]. This type is illustrated by the diffusion processes: there is a probability equal to 1 that some change of $X(t)$ will occur during any time-interval, but the chance is near to 1 that the variation will be, in a specified sense, small for small intervals. This type is described by partial differential equations of parabolic type. There is also a “mixed type” leading to the equation (15) and its adjoint.

⁽⁷⁾ It should be understood that, in general, a solution of (8) is not uniquely determined by the initial values (2), not even in the case of enumerable spaces. The uniqueness mentioned is a consequence of the additional hypothesis that $0 \leq P(\tau, x; t, \Lambda) \leq 1$.

$$(11) \quad P(\tau, x; t, E) < 1.$$

This exceptional case arises only if $p(t, x)$ is unbounded (cf. Theorem 6), but can occur also in the case of enumerable spaces E .

The existence of positive bounded solutions that conform with all other requirements of the theory, including (3), but still fail to be distribution functions, is most striking, and an analysis of this phenomenon was the primary object, and constitutes the most delicate part, of the present investigation. In the case of temporally homogeneous processes, that is, in the case of $p(t, x)$ and $\Pi(t, x, \Lambda)$ not depending on t , we give in §6 a *necessary and sufficient condition that the solution $P(\tau, x; t, \Lambda)$ be a proper probability distribution*, that is to say, that (1) holds. This condition is rather complicated, but can be interpreted in terms of the ergodic properties of the system; and it shows in particular that the exceptional case (11) can arise only in highly dissipative systems. The simplest example for the phenomenon will be given in §7.

In the case of an enumerable space E we write corresponding to (4)

$$(12) \quad p(t, x_i) = p_i(t), \quad \Pi(t, x_i, x_k) = \Pi_{ik}(t).$$

Equations (8) and (9) are then equivalent with

$$(13) \quad \frac{\partial P_{ik}(\tau, t)}{\partial t} = -p_k(t)P_{ik}(\tau, t) + \sum_j p_j(t)\Pi_{jk}(t)P_{ij}(\tau, t)$$

and

$$(14) \quad \frac{\partial P_{ik}(\tau, t)}{\partial \tau} = p_i(\tau) \left\{ P_{ik}(\tau, t) - \sum_j \Pi_{ij}(\tau)P_{jk}(\tau, t) \right\}.$$

In this case the condition (7) is obviously only a regularity restriction, and there exists only the type of purely discontinuous processes. It follows from the results of the present paper that (7) implies the existence of $\partial P_{ik}(\tau, t)/\partial t$ for almost all t , and hence also the convergence of the sum in (13) for almost all t . However, this sum may diverge for special values of t . It is easy to impose on $p_i(t)$ and $\Pi_{ik}(t)$ further restrictions ensuring the convergence of the sum in (13) for all t (cf. §2, (23)).

Equations (13)–(14) were derived by Kolmogoroff [6] under some slight additional hypothesis on the passage to the limit in (7). The case of finitely many x_i was dealt with by several different methods: a full account of them is to be found in Fréchet's treatise [5]. In the case of infinitely many x_i , the first attempt was made by Kolmogoroff, who found a sufficient condition for the existence of a solution of (13) with the initial condition (5)⁽⁸⁾. From the results of the present paper it readily follows, however, that Kolmogoroff's solution is not necessarily a probability distribution, since it is possible that

⁽⁸⁾ Kolmogoroff [6], §10. The usual notation is: $-p_i(t) = A_{ii}(t)$, $p_i(t)\Pi_{ik}(t) = A_{ik}(t)$. With this notation Kolmogoroff's condition requires that, putting $B_k^{(0)} = 1$ and $B_k^{(n+1)} = \sum_j B_j^{(n)} |A_{jk}|$, all $B_k^{(n)}$ exist and that $\sum_n B_k^{(n)} x^n/n!$ converges for some $x > 0$ and all k .

$\sum_k P_{ik} < 1$. On the other hand, Kolmogoroff's assumptions are rather restrictive.

The case of E 's being the real axis or any Borel set on it was dealt with by Feller [3]. Equations (8)–(9) were, however, derived from (7) under additional hypothesis, and integrated only in the case of a bounded $p(t, x)$ (in which case (11) cannot occur). It may be pointed out that this covers also the special case of enumerable E in the case of bounded coefficient in (13)–(14)⁽⁹⁾. As Dubrowski [2] has shown, Feller's results and proofs can be transferred almost literally to the case of an arbitrary abstract space E ⁽¹⁰⁾.

The present method of dealing with equations (8)–(9) is more general than that used loc. cit. [2, 3], but affords at the same time a considerable simplification. The same simplification can be made in the treatment of the more general integro-differential equation of parabolic type:

$$(15) \quad \frac{\partial P(\tau, x; t, \Lambda)}{\partial \tau} + a(\tau, x) \frac{\partial^2 P(\tau, x; t, \Lambda)}{\partial x^2} + b(\tau, x) \frac{\partial P(\tau, x; t, \Lambda)}{\partial x} \\ = p(\tau, x) \left\{ P(\tau, x; t, \Lambda) - \int_E P(\tau, y; t, \Lambda) \Pi(\tau, x, dE_y) \right\},$$

where E is the real axis, $a(\tau, x) > 0$. This equation and its adjoint describe the mixed type of a Markoff process⁽¹¹⁾.

2. Preliminaries. *The following assumptions on $p(t, x)$ and $\Pi(t, x, \Lambda)$ will be made throughout the paper:*

(i) $p(t, x)$ is finite and non-negative for all points x of E and all t of some finite or infinite interval $T_0 < t < T_1$. For x fixed, $p(t, x)$ is a continuous function of t , and for t fixed it is measurable with respect to \mathfrak{B} .

(ii) $\Pi(t, x, \Lambda)$ is defined for $T_0 < t < T_1$, for all $x \in E$ and all sets $\Lambda \in \mathfrak{B}$. For fixed x, Λ it is a continuous function of t ; for fixed t, Λ it is measurable with respect to \mathfrak{B} , and for fixed t, x it is a non-negative completely additive func-

⁽⁹⁾ It suffices namely to interpret the points x_k as integers. It was with a view of this case that $p(t, x)$ was, in [3], supposed only to be measurable with respect to x . The point was not, however, mentioned explicitly and seems to have been generally overlooked.

⁽¹⁰⁾ Added in proof: In a recent paper [9] (which became accessible to the author only after the present paper was submitted for publication), W. Doeblin investigated essentially the same class of stochastic processes with which we are concerned here. It may be remarked that Doeblin's methods as well as his results are different from ours. He proceeds by a direct and careful analysis of the stochastic movement itself, and arrives at a characterization of the process by means of two functions $U(\tau, t, x)$ and $V(\tau, x; t, \Lambda)$ which may, roughly, be described, respectively, as the probability that the moving point $X(t)$ will remain in its initial position x during (τ, t) , and the compound probability that it will undergo a change such that the first jump takes it into the set Λ . These functions must satisfy the functional equations $U(\tau, t, x) = U(\tau, s, x)U(s, t, x)$ and $V(\tau, x; t, \Lambda) = V(\tau, x; s, \Lambda) + U(\tau, s, x)V(s, x; t, \Lambda)$ for $\tau < s < t$. It is shown that except for these equations and some trivial additional restrictions the functions U and V can be prescribed arbitrarily. The occurrence of the exceptional case (11) is ruled out by a uniformity condition.

⁽¹¹⁾ For the definition see §2 and for the integration of (15), §5 of Feller [3].

tion of sets $\Lambda \in \mathfrak{B}$ with

$$(16) \quad \Pi(t, x, E) = 1.$$

Finally, for the set $\Lambda = x$ we suppose that

$$(17) \quad \Pi(t, x, x) = 0.$$

Throughout this paper the parameters t and τ are restricted so that

$$T_0 < \tau < t < T_1$$

where (T_0, T_1) is the interval specified above. x, y, z will denote points of E . Any function of points will be supposed, or is easily seen to be, measurable with respect to \mathfrak{B} . A set $\Lambda \in \mathfrak{B}$ will be called bounded if $p(t, x)$ is uniformly bounded for all t and $x \in \Lambda$. In particular, we shall write

$$(18) \quad \Lambda_a = \underset{x}{E} \{ p(t, x) < a \},$$

where $a > 0$. By (i) obviously $\Lambda_a \in \mathfrak{B}$ and $\Lambda_a \uparrow E$ as $a \uparrow \infty$. Any finite set is bounded, and in the case of an enumerable E it is more convenient to consider finite sets instead of bounded. A similar remark applies if E is equipped with a metric.

By (i) and (ii) integrals of the type $J(t, x, \Lambda) = \int_{\Lambda} p(t, y) \Pi(t, x, dE_y)$ have a meaning; if, in particular, Λ is a bounded set, $J(t, x, \Lambda)$ is for fixed x a continuous function of t , and for fixed t a function of x which is measurable with respect to \mathfrak{B} . Now any set $\Lambda \in \mathfrak{B}$ is the limit of an increasing sequence of bounded sets, and hence any function of the type $J(t, x, \Lambda)$ is the limit of a monotonic sequence of functions which are, for fixed other arguments, continuous with respect to t and measurable with respect to \mathfrak{B} . This remark applies to all integrals which will be used in the sequel, and enables us in particular to use repeated integrals. We shall also frequently have to interchange the order of integration. To legitimate this procedure once for all the following may be remarked.

Only two different types of inversions will be used. Sometimes both integrations will be with respect to time-parameters: in such cases the elementary theory of repeated integrals will suffice to justify the change in the order of integration. In all other cases the inversion will be based on the following

LEMMA⁽¹²⁾. Let $E^{(1)}$ and $E^{(2)}$ be two spaces, and let $\mathfrak{B}^{(i)}$ be a Borel field of subsets of $E^{(i)}$, $i = 1, 2$. Denote by $x^{(i)}$ a point varying in $E^{(i)}$, and by $\Lambda^{(i)}$ a set belonging to $\mathfrak{B}^{(i)}$. Let $f(x^{(1)})$ and $g(x^{(2)})$ be two non-negative and bounded functions, measurable with respect to $\mathfrak{B}^{(1)}$ and to $\mathfrak{B}^{(2)}$, respectively. Let $F(\Lambda^{(1)})$ be a completely additive function of sets $\Lambda^{(1)} \in \mathfrak{B}^{(1)}$, with $0 \leq F(\Lambda^{(1)}) \leq 1$. Finally, let $G(x^{(1)}, \Lambda^{(2)})$ be defined for all $x^{(1)} \in E^{(1)}$ and $\Lambda^{(2)} \in \mathfrak{B}^{(2)}$ so that it is, for fixed $x^{(1)}$,

⁽¹²⁾ Added in proof: A similar theorem for the case of the real axis was announced by R. H. Cameron and W. T. Martin, but is not yet published; see the abstract presented to the American Mathematical Society, 46-3-162.

a completely additive function of sets $\Lambda^{(2)}$, and for fixed $\Lambda^{(2)}$ measurable with respect to $\mathfrak{B}^{(1)}$, and so that for all values of the arguments $0 \leq G(x^{(1)}, \Lambda^{(2)}) \leq 1$. Then for any two fixed sets $\Gamma^{(1)} \in \mathfrak{B}^{(1)}$ and $\Gamma^{(2)} \in \mathfrak{B}^{(2)}$

$$(19) \quad \int_{\Gamma^{(1)}} f(x^{(1)})F(dE_x^{(1)}) \int_{\Gamma^{(2)}} g(x^{(2)})G(x^{(1)}, dE_x^{(2)}) \\ = \int_{\Gamma^{(2)}} g(x^{(2)}) \int_{\Gamma^{(1)}} f(x^{(1)})G(x^{(1)}, dE_x^{(2)})F(dE_x^{(1)}).$$

Before proving this theorem let us remark that it is much simpler than Fubini's theorem, but is not contained in it. In our applications either both spaces $E^{(i)}$ will coincide with E , or else $E^{(1)}$ will be the real time-axis and $E^{(2)}$ the space E . It is clear that

$$\int_{\Gamma^{(2)}} g(x^{(2)})G(x^{(1)}, dE_x^{(2)})$$

is, as a function of $x^{(1)}$, measurable with respect to $\mathfrak{B}^{(1)}$, since it is the limit of measurable functions. Similarly

$$\int_{\Gamma^{(1)}} f(x^{(1)})G(x^{(1)}, \Lambda^{(2)})F(dE_x^{(1)})$$

is a completely additive function of sets $\Lambda^{(2)} \in \mathfrak{B}^{(2)}$. Thus both sides in (19) have a meaning.

The lemma is easily proved by a decomposition $\Gamma^{(2)} = \sum_n \Gamma_n^{(2)}$ where $\Gamma_n^{(2)}$ is the set of all points $x^{(2)}$ where $(n-1)\epsilon \leq g(x^{(2)}) < n\epsilon$. Since $g(x^{(2)})$ is bounded, only finitely many $\Gamma_n^{(2)}$ are not empty. Hence

$$\int_{\Gamma^{(2)}} g(x^{(2)}) \int_{\Gamma^{(1)}} f(x^{(1)})G(x^{(1)}, dE_x^{(2)})F(dE_x^{(1)}) \\ \cong \sum_n n\epsilon \int_{\Gamma^{(1)}} f(x^{(1)})G(x^{(1)}, \Gamma_n^{(2)})F(dE_x^{(1)}) \\ \cong \int_{\Gamma^{(1)}} f(x^{(1)})F(dE_x^{(1)}) \sum_n \left\{ \epsilon G(x^{(1)}, \Gamma_n^{(2)}) + \int_{\Gamma_n^{(2)}} g(x^{(2)})G(x^{(1)}, dE_x^{(2)}) \right\} \\ \cong \int_{\Gamma^{(1)}} f(x^{(1)})F(dE_x^{(1)}) \int_{\Gamma^{(2)}} g(x^{(2)})G(x^{(1)}, dE_x^{(2)}) + \epsilon \int_{\Gamma^{(1)}} f(x^{(1)})F(dE_x^{(1)}),$$

and the last integral is bounded. This proves (19) with the sign \cong instead of the equality. In the same way, however, we get also the opposite limitation, and this accomplishes the proof.

A word has still to be said about the derivation of the equations (8) and (9) and the relations between them, though this is by no means necessary for

the understanding of the following existence theorems. Accordingly, the reader can pass over directly to §3.

Equation (8) is more natural than (9) since, roughly speaking, (9) describes the process in its dependence on the initial values. (8) leads also to a representation of $P(\tau, x; t, \Lambda)$ which is in most cases more useful than the representation deduced from (9). The later equation is, nevertheless, simpler than (8) since the integrals in (9) converge for all sets $\Lambda \in \mathfrak{B}$ and a derivative $\partial P/\partial \tau$ exists for all τ whereas the integrals in (8) will, in general, converge only for bounded sets and almost all t , so that also $\partial P/\partial t$ exists only for bounded sets and almost all t .

In previous papers⁽¹³⁾ the equations (8) and (9) were derived under the assumption that the passage to the limit in (7) takes place uniformly with respect to x . For a general theory, however, such an assumption is not only an unnecessary restriction, but is also dangerous since it can be shown by examples that it is not realized for the actual solutions⁽¹⁴⁾.

To deduce (9) we observe that by (3) we have for $\Delta\tau > 0$

$$P(\tau - \Delta\tau, x; t, \Lambda) = \int_E P(\tau - \Delta\tau, x; \tau, dE_y)P(\tau, y; t, \Lambda)$$

or, splitting the space of integration into x and $E-x$,

$$\begin{aligned} & - \frac{1}{\Delta\tau} \{ P(\tau - \Delta\tau, x; t, \Lambda) - P(\tau, x; t, \Lambda) \} \\ (20) \quad & = P(\tau, x; t, \Lambda) \frac{P(\tau - \Delta\tau, x; \tau, x) - 1}{-\Delta\tau} \\ & \quad - \frac{1}{\Delta\tau} \int_{E-x} P(\tau, y; t, \Lambda) P(\tau - \Delta\tau, x; \tau, dE_y). \end{aligned}$$

Now by (7) and (17)

$$- \frac{1}{\Delta\tau} \{ P(\tau - \Delta\tau, x; \tau, x) - 1 \} \rightarrow p(\tau, x),$$

and using (16) and (17) it is seen that also

$$\frac{1}{\Delta\tau} P(\tau - \Delta\tau, x; \tau, E - x) \rightarrow p(\tau, x).$$

Hence, for fixed τ, x , the ratio $P(\tau - \Delta\tau, x; \tau, \Lambda)/\Delta\tau$ is uniformly bounded for all sets Λ not containing x , and by (7) this quantity tends to $p(\tau, x)\Pi(\tau, x, \Lambda)$.

⁽¹³⁾ Feller [3], Dubrovski [2].

⁽¹⁴⁾ It may be remarked that the occurrence of solutions satisfying (11) has nothing whatsoever to do with the nonuniformity of the passage to the limit in (7) (or with the circumstance that the derivatives of $P(\tau, x; t, \Lambda)$ are not bounded).

The right-hand member of (20) is thus seen to tend to the limit given by (9), and it follows from (20) that a left-hand derivative $\partial P(\tau, x; t, \Lambda)/\partial \tau$ exists for all τ, x, t, Λ and that with this derivative (9) holds. The actual (and unique) solution of (9) will show that this left-hand derivative actually is the derivative in the usual sense.

It seems impossible to give a strict proof also for (8) in an equally simple way. One can easily render (8) plausible by writing, according to (3),

$$(21) \quad \frac{1}{\Delta t} \{P(\tau, x; t + \Delta t, \Lambda) - P(\tau, x; t, \Lambda)\} \\ = \int_E P(\tau, x; t, dE_y) \{P(t, y; t + \Delta t, \Lambda) - \delta(y, \Lambda)\} / \Delta t,$$

and going formally to the limit applying (7). In a strict sense, however, one gets by this procedure only a partial result. Denote namely by \mathfrak{B}' the class of sets such that $\Lambda \in \mathfrak{B}'$ if, and only if, Λ is bounded and there is a constant $a > 0$ such that

$$(22) \quad \frac{1 - P(t, x; t + \Delta t, x)}{\Delta t} < a, \quad \frac{P(t, x; t + \Delta t, E - x)}{\Delta t} < a$$

for all t , all $x \in \Lambda$, and all $\Delta t > 0$. Denoting then by D_t the upper right-hand derivative with respect to t , it follows easily from (21) and (22) that for all sets $\Lambda \in \mathfrak{B}'$

$$D_t P(\tau, x; t, \Lambda) \geq - \int_{\Lambda} p(t, y) P(\tau, x; t, dE_y) \\ + \int_E p(t, y) \Pi(t, y, \Lambda) P(\tau, x; t, dE_y).$$

Now here the first integral converges, since Λ is a bounded set. Thus for fixed $\Lambda \in \mathfrak{B}'$ and for fixed x , $D_t P(\tau, x; t, \Lambda)$ is uniformly bounded from below; and $D_t P(\tau, x; t, \Lambda) = \infty$ for all values of t for which the second integral diverges. Since $0 \leq P(\tau, x; t, \Lambda) \leq 1$ it follows that, for $\Lambda \in \mathfrak{B}'$, the second integral must converge for almost all t , that it is to say, that a finite right-hand derivative $\partial P(\tau, x; t, \Lambda)/\partial t$ exists for all $\tau, \Lambda \in \mathfrak{B}'$, and almost all t , and furthermore that with this derivative

$$\frac{\partial P(\tau, x; t, \Lambda)}{\partial t} \geq - \int_{\Lambda} p(t, y) P(\tau, x; t, dE_y) + \int_E p(t, y) \Pi(t, y, \Lambda) P(\tau, x; t, dE_y).$$

Actually the sign of equality in (8) holds not only for all $\Lambda \in \mathfrak{B}'$ but even for *all bounded* sets and almost all t . For the sake of simplicity we prefer, however, to prove this assertion in an indirect way:

We shall namely prove that *there is* (under the assumptions (i)–(ii) on $p(t, x)$ and $\Pi(t, x, \Lambda)$) *one and only one function* $P(\tau, x; t, \Lambda)$ *satisfying* (9) *with the initial condition* (2) *and which is, for fixed* τ, x, t *a completely additive function of sets* $\Lambda \in \mathfrak{B}$ *with* $0 \leq P(\tau, x; t, \Lambda) \leq 1$. *For bounded sets* Λ *this function will be shown to be an absolutely continuous function of* t , *satisfying* (8) *for almost all* t .

Moreover, it will be shown that with *this solution* (3) *also holds*. This gives a uniqueness theorem for our general problem, but an existence theorem will be given essentially only for uniformly bounded $p(t, x)$ (cf. Theorem 6), since sometimes instead of (1) only (11) holds.

This result shows in particular that we may use, instead of the class \mathfrak{B}' considered above, the class \mathfrak{B}_1 of all bounded sets. This by itself does not imply that (22) holds for any bounded set and some suitable a . It may, however, be remarked that this is actually the case, as is readily seen from the representation of the solution given below.

It may still be pointed out that it can be shown by examples that not even for bounded sets Λ does the derivative $\partial P(\tau, x; t, \Lambda)/\partial t$ need to exist for all t . It is, however, easy to make additional assumptions on $p(t, x)$ which assure the existence of $\partial P(\tau, x; t, \Lambda)/\partial t$ for all bounded sets and all t . Such a hypothesis is, for instance, that

$$(23) \quad \frac{p(t_1, x)}{1 + p(t_2, x)} < M$$

uniformly for all values x, t_1, t_2 . This hypothesis is in particular fulfilled in the case of temporally homogeneous processes.

3. **Solution of (8).** We shall define a new completely additive function of sets $\Lambda \in \mathfrak{B}$ by

$$(24) \quad \Pi^*(\tau, x; t, \Lambda) = \int_{\Lambda} \exp \left\{ - \int_{\tau}^t p(s, y) ds \right\} \Pi(\tau, x, dE_y).$$

Obviously $0 \leq \Pi^*(\tau, x; t, \Lambda) \leq 1$; furthermore for any bounded set Λ

$$(25) \quad \begin{aligned} \frac{\partial \Pi^*(\tau, x; t, \Lambda)}{\partial t} &= - \int_{\Lambda} p(t, y) \exp \left\{ - \int_{\tau}^t p(s, y) ds \right\} \Pi(\tau, x, dE_y) \\ &= - \int_{\Lambda} p(t, y) \Pi^*(\tau, x; t, dE_y). \end{aligned}$$

THEOREM 1. $P_{ut}^{(1b)}$

$$(26) \quad P^{(0)}(\tau, x; t, \Lambda) = \delta(x, \Lambda) \exp \left\{ - \int_{\tau}^t p(s, x) ds \right\},$$

^(1b) $\delta(x, \Lambda)$ was defined by (2).

and for $n \geq 1$

$$(27) \quad P^{(n)}(\tau, x; t, \Lambda) = \int_{\tau}^t d\sigma \int_E p(\sigma, y) \Pi^*(\sigma, y; t, \Lambda) P^{(n-1)}(\tau, x; \sigma, dE_y).$$

Let

$$(28) \quad P(\tau, x; t, \Lambda) = \sum_{n=0}^{\infty} P^{(n)}(\tau, x; t, \Lambda);$$

(i) the function $P(\tau, x; t, \Lambda)$ is for fixed $\tau, t, x \in E$ a completely additive function of sets $\Lambda \in \mathfrak{B}$ with $0 \leq P(\tau, x; t, \Lambda) \leq 1$; (ii) $P(\tau, x; t, \Lambda)$ is for fixed τ, x, Λ an absolutely continuous function of t ; for any bounded Λ and almost all t the derivative $\partial P / \partial t$ is finite and satisfies (8) with the initial condition (2).

Remark. It will be seen that for any bounded Λ and almost all t

$$(29) \quad \begin{aligned} \frac{\partial P^{(0)}(\tau, x; t, \Lambda)}{\partial t} &= - p(t, x) P^{(0)}(\tau, x; t, \Lambda), \\ \frac{\partial P^{(n)}(\tau, x; t, \Lambda)}{\partial t} &= - \int_{\Lambda} p(t, y) P^{(n)}(\tau, x; t, dE_y) \\ &\quad + \int_E p(t, y) \Pi(t, y, \Lambda) P^{(n-1)}(\tau, x; t, dE_y); \end{aligned}$$

the integrals on the right side converging for almost all t . These equations are in close analogy with (8), and afford the interpretation of $P^{(n)}(\tau, x; t, \Lambda)$ as the compound probability that during (τ, t) the state X will change by exactly n jumps and that $X(t) \in \Lambda$, if it is known that $X(\tau) = x$.

In the special case of an *enumerable* E it is, of course, sufficient to determine the quantities $P_{ik}^{(n)}(\tau, t)$. For these, (29) reduces to the ordinary differential equations

$$(30) \quad \begin{aligned} \frac{\partial P_{ik}^{(0)}(\tau, t)}{\partial t} &= - p_k(t) P_{ik}^{(0)}(\tau, t), \\ \frac{\partial P_{ik}^{(n)}(\tau, t)}{\partial t} &= - p_k(t) P_{ik}^{(n)}(\tau, t) + \sum_j p_j(t) \Pi_{jk}(t) P_{ij}^{(n-1)}(\tau, t), \end{aligned}$$

and (26)–(27) to

$$P_{ij}^{(n)}(\tau, t) = \sum_i \int_{\tau}^t \exp \left\{ - \int_{\sigma}^t p_k(s) ds \right\} p_j(\sigma) \Pi_{jk}(\sigma) P_{ij}^{(n-1)}(\tau, \sigma) d\sigma.$$

If the $p_j(\sigma)$ are not subjected to a further restriction analogous to (23), the derivative $\partial P_{ik}^{(n)}(\tau, t) / \partial t$ will exist only for almost all t .

Proof. Suppose, by induction, that $P^{(n)}(\tau, x; t, \Lambda)$ exists for some fixed $n \geq 0$, and all values of the arguments, and that it is a completely additive function of sets $\Lambda \in \mathfrak{B}$ with $0 \leq P^{(n)}(\tau, x; t, \Lambda) \leq 1$; furthermore that

$$(31) \quad L^{(n)}(\tau, x, t) = \int_{\tau}^t d\sigma \int_E \dot{p}(\sigma, y) P^{(n)}(\tau, x; \sigma, dE_y)$$

is finite. This is certainly true for $n=0$ and

$$(32) \quad P^{(0)}(\tau, x; t, E) + L^{(0)}(\tau, x, t) = 1.$$

It follows then from (27) that also $P^{(n+1)}(\tau, x; t, \Lambda)$ exists and $0 \leq P^{(n+1)}(\tau, x; t, \Lambda) \leq L^{(n)}(\tau, x, t)$. For any *bounded* set Λ , therefore, we get from (25) and (27)

$$\begin{aligned} & \int_{\Lambda} \dot{p}(t, y) P^{(n+1)}(\tau, x; t, dE_y) \\ &= - \int_{\tau}^t d\sigma \int_E \dot{p}(\sigma, y) \frac{\partial \Pi^*(\sigma, y; t, \Lambda)}{\partial t} P^{(n)}(\tau, x; \sigma, dE_y); \end{aligned}$$

the left-hand member is obviously a continuous function of t , and we get

$$\begin{aligned} & \int_{\tau}^t d\sigma_1 \int_{\Lambda} \dot{p}(\sigma_1, y) P^{(n+1)}(\tau, x; \sigma_1, dE_y) \\ &= - \int_{\tau}^t d\sigma_1 \int_{\tau}^{\sigma_1} d\sigma \int_E \dot{p}(\sigma, y) \frac{\partial \Pi^*(\sigma, y; \sigma_1, \Lambda)}{\partial \sigma_1} P^{(n)}(\tau, x; \sigma, dE_y); \end{aligned}$$

inverting the order of integration and observing that $\Pi^*(\sigma, y; \sigma, \Lambda) = \Pi(\sigma, y, \Lambda)$, we get

$$\begin{aligned} & \int_{\tau}^t d\sigma_1 \int_{\Lambda} \dot{p}(\sigma_1, y) P^{(n+1)}(\tau, x; \sigma_1, dE_y) \\ &= \int_{\tau}^t d\sigma \int_E \dot{p}(\sigma, y) \Pi(\sigma, y, \Lambda) P^{(n)}(\tau, x; \sigma, dE_y) \\ &\quad - \int_{\tau}^t d\sigma \int_E \dot{p}(\sigma, y) \Pi^*(\sigma, y; t, \Lambda) P^{(n)}(\tau, x; \sigma, dE_y), \end{aligned}$$

or by (27) finally

$$(33) \quad \begin{aligned} & P^{(n+1)}(\tau, x; t, \Lambda) + \int_{\tau}^t d\sigma \int_{\Lambda} \dot{p}(\sigma, y) P^{(n+1)}(\tau, x; \sigma, dE_y) \\ &= \int_{\tau}^t d\sigma \int_E \dot{p}(\sigma, y) \Pi(\sigma, y, \Lambda) P^{(n)}(\tau, x; \sigma, dE_y). \end{aligned}$$

This is, essentially, the relation (29) of the remark following Theorem 1. Equation (33) holds for any bounded set Λ . We apply (33) in particular to

$\Lambda = \Lambda_a$ (see (18)) and let $a \uparrow \infty$. Since $\Pi(\sigma, y, \Lambda_a) \leq 1$ the right-hand member is bounded by $L^{(n)}(\tau, x, t)$ (cf. (31)). Hence we get

$$(34) \quad P^{(n+1)}(\tau, x; t, E) + L^{(n+1)}(\tau, x, t) = L^{(n)}(\tau, x, t).$$

It is thus seen that both $P^{(n+1)}(\tau, x; t, \Lambda)$ and $L^{(n+1)}(\tau, x, t)$ exist. Moreover, since $P^{(n+1)}(\tau, x; t, \Lambda) \geq 0$, we have $L^{(n+1)}(\tau, x, t) \leq L^{(n)}(\tau, x, t)$. Thus, for all $n \geq 0$,

$$(35) \quad 1 \geq L^{(0)}(\tau, x, t) \geq L^{(1)}(\tau, x, t) \geq \dots \geq L^{(n)}(\tau, x, t) \rightarrow L(\tau, x, t).$$

By (34) and (35) also $P^{(n+1)}(\tau, x; t, \Lambda) \leq 1$, and thus the assumptions of the inductive argument hold for all n .

It may be remarked that it can be shown by examples that the integrand of (31), $\int_E p(\sigma, y) P^{(n)}(\tau, x; \sigma, dE_y)$, sometimes diverges for some values of σ . The proof shows, however, that it converges for almost all σ , and it is readily seen that it converges for all σ if (23) holds.

Now we get from (34) and (32)

$$(36) \quad \sum_{n=0}^N P^{(n)}(\tau, x; t, E) = 1 - L^{(N)}(\tau, x, t),$$

and thus $0 \leq P(\tau, x; t, \Lambda) \leq 1$.

Hence we readily deduce from (33) for any bounded set $\Lambda \in \mathfrak{B}$

$$(37) \quad \begin{aligned} P(\tau, x; t, \Lambda) + \int_{\tau}^t d\sigma \int_{\Lambda} p(\sigma, y) P(\tau, x; \sigma, dE_y) \\ = \int_{\tau}^t d\sigma \int_E p(\sigma, y) \Pi(\sigma, y, \Lambda) P(\tau, x; \sigma, dE_y); \end{aligned}$$

this proves (8) for almost all t . If (23) holds, (37) can obviously be differentiated for all t and (8) holds for any bounded set and all t .

From (35) and (37) we get also the following

COROLLARY. *The necessary and sufficient condition that $P(\tau, x; t, E) = 1$ for all t is that*

$$(38) \quad L(\tau, x, t) = \lim_{N \rightarrow \infty} L^{(N)}(\tau, x, t) = 0$$

for all t .

Incidentally, it is quite obvious that (7) holds at least for all bounded sets Λ , since by (26) and (27)

$$\begin{aligned} \lim_{h \rightarrow +0} \frac{1}{h} \{ \delta(x, \Lambda) - P^{(0)}(\tau, x; \tau + h, \Lambda) \} &= p(\tau, x) \delta(x, \Lambda), \\ \lim_{h \rightarrow +0} \frac{1}{h} P^{(1)}(\tau, x; \tau + h, \Lambda) &= p(\tau, x) \Pi(\tau, x, \Lambda) \end{aligned}$$

and by (10) and (16)

$$\begin{aligned} \limsup_{h \rightarrow +0} \frac{1}{h} \sum_{n=2}^{\infty} P^{(n)}(\tau, x; \tau + h, E) \\ \leq \lim_{h \rightarrow +0} \frac{1}{h} \{1 - P^{(0)}(\tau, x; \tau + h, E) - P^{(1)}(\tau, x; \tau + h, E)\} \\ = 0. \end{aligned}$$

That (7) holds for any $\Lambda \in \mathfrak{B}$ will be proved in §5.

4. **Solution of (9).** We now prove the following theorem.

THEOREM 2. Put

$$(39) \quad Q^{(0)}(\tau, x; t, \Lambda) = \delta(x, \Lambda) \exp \left\{ - \int_{\tau}^t p(s, x) ds \right\},$$

and for $n \geq 1$

$$(40) \quad \begin{aligned} Q^{(n)}(\tau, x; t, \Lambda) = & \int_{\tau}^t p(\sigma, x) \\ & \exp \left\{ - \int_{\tau}^{\sigma} p(s, x) ds \right\} d\sigma \int_E Q^{(n-1)}(\sigma, y; t, \Lambda) \Pi(\sigma, x, dE_y). \end{aligned}$$

Then⁽¹⁶⁾, for any fixed $\tau, x, t,$

$$(41) \quad P(\tau, x; t, \Lambda) = \sum_{n=0}^{\infty} Q^{(n)}(\tau, x; t, \Lambda)$$

is a completely additive function of sets $\Lambda \in \mathfrak{B}$ and $0 \leq P(\tau, x; t, \Lambda) \leq 1$. Furthermore $P(\tau, x; t, \Lambda)$ is a solution of (9) with the initial values given by (2).

Remark. Obviously the $Q^{(n)}(\tau, x; t, \Lambda)$ are solutions of the equations

$$(42) \quad \begin{aligned} \frac{\partial Q^{(0)}(\tau, x; t, \Lambda)}{\partial \tau} &= p(\tau, x) Q^{(0)}(\tau, x; t, \Lambda), \\ \frac{\partial Q^{(n)}(\tau, x; t, \Lambda)}{\partial \tau} &= p(\tau, x) \left\{ Q^{(n)}(\tau, x; t, \Lambda) \right. \\ &\quad \left. - \int_E Q^{(n-1)}(\tau, y; t, \Lambda) \Pi(\tau, x, dE_y) \right\}, \end{aligned}$$

which can be treated as ordinary differential equations.

Proof. Put

$$S^{(n)}(\tau, x; t, \Lambda) = \sum_{k=0}^n Q^{(k)}(\tau, x; t, \Lambda).$$

⁽¹⁶⁾ It will be seen (Theorem 4) that the functions defined by (41) and (28) are actually identical.

Then $0 \leq S^{(0)}(\tau, x; t, \Lambda) \leq 1$. Let us suppose that $0 \leq S^{(0)} \leq S^{(1)} \leq \dots \leq S^{(n-1)} \leq 1$. Then by (39) and (40)

$$\begin{aligned}
 S^{(n)}(\tau, x; t, \Lambda) &= \exp \left\{ - \int_{\tau}^t p(s, x) ds \right\} \left\{ \delta(x, \Lambda) + \int_{\tau}^t p(\sigma, x) \right. \\
 (43) \quad &\left. \exp \left\{ \int_{\sigma}^t p(s, x) ds \right\} d\sigma \int_{\mathcal{E}} S^{(n-1)}(\sigma, y; t, \Lambda) \Pi(\sigma, x, dE_y) \right\} \\
 &\leq \exp \left\{ - \int_{\tau}^t p(s, x) ds \right\} \left\{ 1 \right. \\
 &\quad \left. + \int_{\tau}^t p(\sigma, x) \exp \left\{ \int_{\sigma}^t p(s, x) ds \right\} d\sigma \right\} = 1.
 \end{aligned}$$

On the other hand obviously $S^{(n)}(\tau, x; t, \Lambda) \geq S^{(n-1)}(\tau, x; t, \Lambda)$. Hence $S^{(n)}(\tau, x; t, \Lambda) \uparrow P(\tau, x; t, \Lambda) \leq 1$. That $P(\tau, x; t, \Lambda)$ is a solution of (9) follows immediately from (42), and also the initial condition (2) is obviously satisfied.

THEOREM 3 (Uniqueness theorem⁽¹⁷⁾). *Consider some fixed t and a function $P^*(\tau, x; t, \Lambda)$ which (i) for fixed τ, x, t is a completely additive function of sets $\Lambda \in \mathfrak{B}$ with $0 \leq P^*(\tau, x; t, \Lambda) \leq 1$; (ii) for fixed x, t, Λ is an absolutely continuous function of τ satisfying for almost all τ the equation (9) with the initial value (2) as $\tau \rightarrow t - 0$. Then $P^*(\tau, x; t, \Lambda) = P(\tau, x; t, \Lambda)$, where $P(\tau, x; t, \Lambda)$ is the function defined by Theorem 2.*

Proof. (i) We first show that

$$(44) \quad P^*(\tau, x; t, \Lambda) \geq P(\tau, x; t, \Lambda);$$

this remains true also if the assumption $P^*(\tau, x; t, \Lambda) \leq 1$ be replaced by the weaker one that $P^*(\tau, x; t, \Lambda)$ is uniformly bounded. In fact, treating (9) as an ordinary differential equation, we get by (2)

$$\begin{aligned}
 P^*(\tau, x; t, \Lambda) &= \exp \left\{ - \int_{\tau}^t p(s, x) ds \right\} \left\{ \delta(x, \Lambda) + \int_{\tau}^t p(\sigma, x) \right. \\
 (45) \quad &\left. \exp \left\{ - \int_{\sigma}^t p(s, x) ds \right\} d\sigma \int_{\mathcal{E}} P^*(\sigma, y; t, \Lambda) \Pi(\sigma, x, dE_y) \right\}.
 \end{aligned}$$

Since the last term is non-negative, we see by comparison of (45) with (39) that $P^*(\tau, x; t, \Lambda) \geq Q^{(0)}(\tau, x; t, \Lambda) = S^{(0)}(\tau, x; t, \Lambda)$. Comparing, then, (45) with (43) it is readily seen that $P^*(\tau, x; t, \Lambda) \geq S^{(n)}(\tau, x; t, \Lambda)$ for any n , which proves (44).

(ii) Put

⁽¹⁷⁾ Cf. footnote 7.

$$(46) \quad D(\tau, x; t, \Lambda) = P^*(\tau, x; t, \Lambda) - P(\tau, x; t, \Lambda).$$

By (44), $D(\tau, x; t, \Lambda)$ is a completely additive function of sets, and $0 \leq D(\tau, x; t, \Lambda) \leq 1$. Now, assuming that $D(\tau, x; t, E)$ takes on the value $\alpha > 0$ somewhere, denote by τ_0 the least upper bound of all τ for which $D(\tau, x; t, E) \geq \alpha$, so that

$$(47) \quad D(\tau, x; t, E) < D(\tau_0, x; t, E) = \alpha \quad \text{for } \tau_0 < \tau < t.$$

Now $(1/\alpha)D(\tau, x; t, \Lambda)$ is a solution of (9), which vanishes as $\tau \rightarrow t - 0$. Hence

$$(48) \quad \frac{1}{\alpha} D(\tau, x; t, E) = \int_{\tau}^t p(\sigma, x) \exp \left\{ - \int_{\tau}^{\sigma} p(s, x) ds \right\} d\sigma \int_E \frac{1}{\alpha} D(\sigma, y; t, \Lambda) \Pi(\sigma, x, dE_y).$$

Combining (47) and (48), we get

$$\begin{aligned} 1 &= \frac{1}{\alpha} D(\tau_0, x; t, E) \leq \int_{\tau_0}^t p(\sigma, x) \exp \left\{ - \int_{\tau_0}^{\sigma} p(s, x) ds \right\} d\sigma \\ &= 1 - \exp \left\{ - \int_{\tau_0}^t p(s, x) ds \right\} < 1. \end{aligned}$$

Thus the assumption $D(\tau, x; t, E) = \alpha > 0$ leads to a contradiction. Hence, by (44) and (46), $D(\tau, x; t, \Lambda) = 0$ for all sets Λ , and this accomplishes the proof.

5. Properties of the solutions. We now prove

THEOREM 4. (i) *With the functions defined by Theorems 1 and 2 one has*

$$(49) \quad P^{(n)}(\tau, x; t, \Lambda) = Q^{(n)}(\tau, x; t, \Lambda)$$

identically; thus equations (28) and (41) define the same function $P(\tau, x; t, \Lambda)$.

(ii) *This function satisfies the fundamental assumption (7).*

Proof. (i) Put $P^{(n+1)}(\tau, x; t, \Lambda) = AP^{(n)}(\tau, x; t, \Lambda)$, where A is a linear operator on (t, Λ) ; and similarly $Q^{(n+1)}(\tau, x; t, \Lambda) = BQ^{(n)}(\tau, x; t, \Lambda)$ where the operator B works on (τ, x) . Using the lemma of §2, we readily see that the operators A and B are permutable.

Now obviously $P^{(0)}(\tau, x; t, \Lambda) = Q^{(0)}(\tau, x; t, \Lambda)$ and $P^{(1)}(\tau, x; t, \Lambda) = Q^{(1)}(\tau, x; t, \Lambda)$. Assuming, then, (49) to be true for some $n \geq 1$, we get $P^{(n+1)} = AP^{(n)} = AQ^{(n)} = ABQ^{(n-1)} = BAP^{(n-1)} = BP^{(n)} = BQ^{(n)} = Q^{(n+1)}$.

(ii) To prove the second part we use the representation (39)–(41). Then

$$\frac{\delta(x, \Lambda) - Q^{(0)}(\tau, x; t, \Lambda)}{t - \tau} \rightarrow p(t, x)\delta(x, \Lambda),$$

obviously, as $\tau \rightarrow t - 0$. Moreover

$$\sum_{n=1}^{\infty} Q^{(n)}(\tau, x; t, E) \leq 1 - Q^{(0)}(\tau, x; t, E)$$

and thus

$$(50) \quad \limsup_{\tau \rightarrow t} \frac{1}{t - \tau} \sum_{n=1}^{\infty} Q^{(n)}(\tau, x; t, E) \leq p(t, x).$$

From (40) we get however

$$(51) \quad \liminf_{\tau \rightarrow t} \frac{1}{t - \tau} \sum_{n=1}^{\infty} Q^{(n)}(\tau, x; t, \Lambda) \geq \liminf_{\tau \rightarrow t} \frac{1}{t - \tau} Q^{(1)}(\tau, x; t, \Lambda) \\ = p(t, x)\Pi(t, x, \Lambda).$$

Applying now (51) both for Λ and $E - \Lambda$, we get by (50)

$$\lim_{\tau \rightarrow t} \frac{1}{t - \tau} \sum_{n=1}^{\infty} Q^{(n)}(\tau, x; t, \Lambda) = p(t, x)\Pi(t, x, \Lambda)$$

which proves the theorem.

THEOREM 5. For $\tau < \lambda < t$ one has identically

$$(52) \quad Q^{(n)}(\tau, x; t, \Lambda) = \sum_{k=0}^n \int_E Q^{(k)}(\tau, x; \lambda, dE_y) Q^{(n-k)}(\lambda, y; t, \Lambda)$$

where $Q^{(n)}(\tau, x; t, \Lambda)$ was defined by (39)–(40).

The solution $P(\tau, x; t, \Lambda)$ of Theorems 1 and 2 satisfies the equation (3) of Chapman-Kolmogoroff⁽¹⁸⁾.

Proof. The second part of the theorem is an immediate consequence of the first part.

Equation (52) is trivial for $n=0$. Assuming it to be true for some $n \geq 0$, we get by (39)–(40)

$$\sum_{k=0}^{n+1} \int_E Q^{(k)}(\tau, x; \lambda, dE_y) Q^{(n+1-k)}(\lambda, y; t, \Lambda) \\ = \exp \left\{ - \int_{\tau}^{\lambda} p(s, x) ds \right\} Q^{(n+1)}(\lambda, x; t, \Lambda) \\ + \sum_{k=1}^{n+1} \int_{\tau}^{\lambda} p(\sigma, x) \exp \left\{ - \int_{\tau}^{\sigma} p(s, x) ds \right\} d\sigma \\ \cdot \int_{E_{\sigma}} \Pi(\sigma, x, dE_z) \int_{E_y} Q^{(k-1)}(\sigma, z; \lambda, dE_y) Q^{(n+1-k)}(\lambda, y; t, \Lambda)$$

⁽¹⁸⁾ It should be observed that Theorem 5 is valid even in cases where $P(\tau, x; t, \Lambda)$ is not a proper probability distribution, i.e., where $P(\tau, x; t, E) < 1$.

$$\begin{aligned}
 &= \exp \left\{ - \int_{\tau}^{\lambda} p(s, x) ds \right\} Q^{(n+1)}(\lambda, x; t, \Lambda) \\
 &\quad + \int_{\tau}^{\lambda} p(\sigma, x) \exp \left\{ - \int_{\tau}^{\sigma} p(s, x) ds \right\} d\sigma \int_{E_x} \Pi(\sigma, x, dE_x) Q^{(n)}(\sigma, z; t, \Lambda) \\
 &= \exp \left\{ - \int_{\tau}^{\lambda} p(s, x) ds \right\} Q^{(n+1)}(\lambda, x; t, \Lambda) + Q^{(n+1)}(\tau, x; t, \Lambda) \\
 &\quad - \int_{\lambda}^t p(\sigma, x) \exp \left\{ - \int_{\tau}^{\sigma} p(s, x) ds \right\} d\sigma \int_{E} \Pi(\sigma, x, dE_x) Q^{(n)}(\sigma, z; t, \Lambda) \\
 &= Q^{(n+1)}(\tau, x; t, \Lambda).
 \end{aligned}$$

THEOREM 6. *If there is some $\alpha > 1$ and a function $\pi(t) \in L^{\alpha}$ such that uniformly*

$$p(t, x) < \pi(t),$$

then the solution $P(\tau, x; t, \Lambda)$ of Theorems 1 and 2 is a probability distribution, i.e. (1) holds.

Proof. With the notation of the proof of Theorem 1 we have by (35) and (31) for any n

$$\begin{aligned}
 L(\tau, x, t) &\leq \int_{\tau}^t d\sigma \int_{E} p(\sigma, y) P^{(n)}(\tau, x; \sigma, dE_y) \leq \int_{\tau}^t \pi(\sigma) P^{(n)}(\tau, x; \sigma, E) d\sigma \\
 &\leq \left\{ \int_{\tau}^t [\pi(\sigma)]^{\alpha} d\sigma \right\}^{1/\alpha} \left\{ \int_{\tau}^t P^{(n)}(\tau, x; \sigma, E)^{\alpha/(\alpha-1)} d\sigma \right\}^{(\alpha-1)/\alpha},
 \end{aligned}$$

and since $0 \leq P^{(n)}(\tau, x; \sigma, E) \leq 1$ it follows that

$$(53) \quad \int_{\tau}^t P^{(n)}(\tau, x; \sigma, E) d\sigma \geq h(\tau, t) L(\tau, x, t)^{\alpha/(\alpha-1)},$$

where $h(\tau, t) > 0$ is independent of n . But the left-hand member in (53) is the general term of a convergent series, and therefore $L(\tau, x, t) = 0$. This proves the proposition in view of the corollary to Theorem 1.

6. The temporally homogeneous process. So far it has been shown that there is always a function $P(\tau, x; t, \Lambda)$ satisfying all requirements of the theory except, perhaps, (1). That (1) does not necessarily hold will be shown by means of a simple example in §7. This is a surprising result and requires a better understanding of the mechanism of the process. We shall confine ourselves to the temporally homogeneous processes; but, at least as far as sufficiency is concerned, the condition of the following theorem can easily be extended to some more general cases.

We begin with some preliminary remarks and notations. In the case of

$p(t, x)$ and $\Pi(t, x, \Lambda)$ not depending on t the solution $P(\tau, x; t, \Lambda)$ of Theorems 1 and 2 obviously depends only on $t - \tau$ and we write

$$P(\tau, x; t, \Lambda) = P(t - \tau, x, \Lambda).$$

Similarly we write

$$(54) \quad p(t, x) = p(x), \quad \Pi(t, x, \Lambda) = \Pi(x, \Lambda).$$

$\Pi(x, \Lambda)$ defines in the usual way an ordinary *Markoff chain*, that is to say, a sequence of probability distributions defined by

$$(55) \quad \begin{aligned} \Pi^{(0)}(x, \Lambda) &= \delta(x, \Lambda), \\ \Pi^{(n)}(x, \Lambda) &= \int_E \Pi^{(n-1)}(y, \Lambda) \Pi(x, dE_y), \quad (n \geq 1). \end{aligned}$$

Obviously this chain is closely related to our stochastic process, and in particular the ergodic properties of the original process will be regulated by the ergodic properties of the chain (55). Roughly speaking, $\Pi^{(n)}(x, \Lambda)$ gives the conditional probability distribution of the state $X(t)$ under the assumption that $X(0) = x$ and that a change of state occurred during $(0, t)$ exactly n times—the time of occurrence of these jumps being left out of account.

If Λ and Ω are any two sets of \mathfrak{B} , we put

$$(56) \quad \begin{aligned} \Pi_{\Omega}^{(0)}(x, \Lambda) &= \delta(x, \Lambda \cap \Omega), \\ \Pi_{\Omega}^{(n)}(x, \Lambda) &= \int_{\Omega} \Pi_{\Omega}^{(n-1)}(x, dE_y) \Pi(y, \Lambda), \quad (n \geq 1). \end{aligned}$$

In terms of the Markoff chain (55) $\Pi_{\Omega}^{(n)}(x, \Lambda)$ is the probability that the moving point, starting from $x \in \Omega$ would remain in Ω for the $n - 1$ first steps and would be taken into some point of Λ by the n th step. Obviously $\Pi_{\Omega}^{(n)}(x, \Lambda) = 0$ for all sets Λ if x is not contained in Ω . For fixed x and Ω the sequence $\Pi_{\Omega}^{(n)}(x, \Omega)$ is never increasing: $\Pi_{\Omega}^{(n)}(x, \Omega) \downarrow \alpha$. If $\alpha > 0$, there is a positive probability of never leaving the set Ω , if we have started from the point $x \in \Omega$. For further application we note that for $x \in \Omega$ we have

$$(57) \quad \sum_{k=0}^n \Pi_{\Omega}^{(k)}(x, E - \Omega) + \Pi_{\Omega}^{(n)}(x, \Omega) = 1.$$

Finally we introduce the notation

$$(58) \quad \Omega_+ = E_{x \in \Omega} \{p(x) > 0\},$$

that is to say, Ω_+ consists of those points $x \in \Omega$ for which $p(x) \neq 0$.

THEOREM 7. *Suppose that $p(t, x)$ and $\Pi(t, x, \Lambda)$ are of the form (54).*

(i) In order that the solution $P(\tau, x; t, \Lambda)$ of Theorems 1 and 2 satisfy (1) it is necessary and sufficient that whenever for some point x and some set $\Omega \in \mathfrak{B}$ with $\Omega = \Omega_+$ the inequality

$$(59) \quad \Pi_{\Omega}^{(n)}(x, \Omega) > \alpha > 0$$

holds for all n , then the series

$$(60) \quad \sum_{n=0}^{\infty} \int_{\Omega} \frac{1}{p(y)} \Pi_{\Omega}^{(n)}(x, dE_y)$$

diverges⁽¹⁹⁾.

(ii) In this statement the series (60) can be replaced by

$$(61) \quad \sum_{n=0}^{\infty} \int_{\Omega} \frac{1}{p(y)} \Pi^{(n)}(x, dE_y).$$

COROLLARY. In order that $P(t, x, E) = 1$ it is necessary that for any point $x \in E_+$ the series

$$\sum_{n=1}^{\infty} \int_{E_+} \frac{1}{p(y)} \Pi^{(n)}(x, dE_y)$$

diverges. (This condition is, however, not sufficient, as will be shown by an example in §7.)

Proof. Condition (ii) is stronger than (i). We have, therefore, to prove that the divergence of (61) is a sufficient, the divergence of (60) a necessary, condition.

(i) *Sufficiency.* This part of the proof will rest mainly on the representation of $P(t, x, \Lambda)$ given by Theorem 2.

In the case of a temporally homogeneous process the function $L^{(n)}(\tau, x, t)$ defined by (31) depends only on x and $t - \tau$, and we write

$$L^{(n)}(\tau, x, t) = L^{(n)}(t - \tau, x),$$

so that for $t > 0$

$$(62) \quad L^{(n)}(t, x) = \int_0^t d\sigma \int_E p(y) P^{(n)}(\sigma, x, dE_y).$$

Now, using the notation (54), we get from (42)

$$(63) \quad P^{(0)}(t, x, \Lambda) + p(x) \int_0^t P^{(0)}(\sigma, x, \Lambda) d\sigma = \delta(x, \Lambda),$$

⁽¹⁹⁾ In the sense that the series is to be considered as divergent if some of the integrals in (60) are divergent.

and for $n \geq 1$

$$(64) \quad \begin{aligned} P^{(n)}(t, x, \Lambda) + p(x) \int_0^t P^{(n)}(\sigma, x, \Lambda) d\sigma \\ = p(x) \int_0^t d\sigma \int_E \Pi(x, dE_y) P^{(n-1)}(\sigma, y, \Lambda). \end{aligned}$$

Combining (64) with (62) we get, for $n \geq 1$,

$$(65) \quad \int_E p(y) P^{(n)}(t, x, dE_y) + p(x) L^{(n)}(t, x) = p(x) \int_E L^{(n-1)}(t, y) \Pi(x, dE_y).$$

Now, for all points $x \in E - E_+$ (that is to say, if $p(x) = 0$) we have $L^{(n)}(t, x) = 0$ for all n . Hence, using an inductive argument, it readily follows from (65) and (32) that for $x \in E_+$ and any $n \geq 0$

$$(66) \quad L^{(n)}(t, x) = 1 - \sum_{k=0}^n \int_{E_+} \frac{1}{p(y)} \Pi^{(k)}(x, dE_y) \int_{E_+} p(z) P^{(n-k)}(t, y, dE_z).$$

Integrating (66) we get

$$\sum_{k=0}^n \int_{E_+} \frac{1}{p(y)} L^{(n-k)}(t, y) \Pi^{(k)}(x, dE_y) = \int_0^t \{1 - L^{(n)}(\sigma, x)\} d\sigma \leq t,$$

and since by (35) $L^{(n)}(t, y) \downarrow L(t, y)$, it follows that for $x \in E_+$

$$(67) \quad \sum_{k=0}^n \int_{E_+} \frac{1}{p(y)} L(t, y) \Pi^{(k)}(x, dE_y) \leq t.$$

Suppose now that there is some point x_0 and some t such that $P(t, x_0, E) < 1$. By the corollary to Theorem 1 this implies that

$$(68) \quad 1 > L(t, x_0) = \alpha > 0.$$

Denote, then, by Ω the set of all points x with

$$(69) \quad L(t, x) \geq \alpha.$$

Obviously $\Omega = \Omega_+$, since $p(x) = 0$ implies $L(t, x) = 0$. Now by (65) we have

$$L(t, x) \leq \int_{E_+} L(t, y) \Pi(x, dE_y),$$

and consequently

$$\begin{aligned} \alpha = L(t, x_0) &\leq \int_{E_+} L(t, y) \Pi(x_0, dE_y) \leq \alpha \Pi(x_0, E - \Omega) + \int_{\Omega} L(t, y) \Pi(x_0, dE_y), \\ &= \alpha \Pi_{\Omega}^{(1)}(x_0, E - \Omega) + \int_{\Omega} L(t, y) \Pi_{\Omega}^{(1)}(x_0, dE_y) \end{aligned}$$

so that by an inductive argument, using (56) we have

$$(70) \quad \alpha \leq \alpha \sum_{k=0}^n \Pi_{\Omega}^{(k)}(x_0, E - \Omega) + \int_{\Omega} L(t, y) \Pi_{\Omega}^{(n)}(x_0, dE_y).$$

Here the sign of equality can hold only if $\Pi_{\Omega}^{(k)}(x_0, E - \Omega) = 0$ for $k = 0, \dots, n$. Since $L(t, y) \leq 1$ we get from (70) using (57)

$$(71) \quad \alpha \leq \alpha(1 - \eta) + \eta,$$

where η is defined by

$$\Pi_{\Omega}^{(n)}(x_0, \Omega) \downarrow \eta.$$

Again, in (71) the sign of equality can hold only if $\Pi_{\Omega}^{(k)}(x_0, E - \Omega) = 0$ for all k ; but by (57) we have in this case $\eta = 1$. Otherwise $\alpha < \alpha(1 - \eta) + \eta$ so that certainly $\eta > 0$. Thus $\Omega = \Omega_+$ is a set with the property stated in the theorem. However, by (67), (68) and (69) we get

$$\begin{aligned} \alpha \sum_{k=0}^{\infty} \int_{\Omega} \frac{1}{p(y)} \Pi^{(k)}(x_0, dE_y) &\leq \sum_{k=0}^{\infty} \int_{\Omega} \frac{1}{p(y)} L(t, y) \Pi^{(k)}(x_0, dE_y) \\ &\leq \sum_{k=0}^{\infty} \int_{E_+} \frac{1}{p(y)} L(t, y) \Pi^{(k)}(x_0, dE_y) \leq t, \end{aligned}$$

which means that the series (61) is convergent. Thus the divergence of the series implies $\alpha = 0$ or $P(t, x_0, E) = 1$.

(ii) *Necessity.* This part of the proof will mainly use the representation of $P(t, x, \Lambda)$ given in Theorem 1.

Suppose that condition (i) of the theorem does not hold, that is to say, that there is a set $\Omega = \Omega_+$ for which (59) holds for some fixed $x_0 \in \Omega$ and for which

$$(72) \quad \sum_{k=0}^{\infty} \int_{\Omega} \frac{1}{p(y)} \Pi_{\Omega}^{(k)}(x_0, dE_y) = a < \infty.$$

(72) implies in particular that all the integrals occurring converge. For this fixed set Ω and all points $x \in \Omega$ we define, in analogy to (27), an additive function $P_{\Omega}(t, x, \Lambda)$ of sets $\Lambda \in \mathfrak{E}$ by the recurrence formula

$$(73) \quad \begin{aligned} P_{\Omega}^{(0)}(t, x, \Lambda) &= P^{(0)}(t, x, \Lambda), \\ P_{\Omega}^{(n)}(t, x, \Lambda) &= \int_0^t d\sigma \int_E p(y) \Pi^*(t - \sigma, y, \Lambda \Omega) P_{\Omega}^{(n-1)}(\sigma, x, dE_y), \end{aligned}$$

where

$$\Pi^*(t, x, \Lambda) = \int_{\Lambda} \exp \{ -tp(y) \} \Pi(x, dE_y).$$

Obviously

$$(74) \quad 0 \leq P_{\Omega}^{(n)}(t, x, \Lambda) \leq P^{(n)}(t, x, \Lambda),$$

and putting

$$(75) \quad W_{\Omega}^{(n)}(t, x, \Lambda) = P^{(n)}(t, x, \Lambda) - P_{\Omega}^{(n)}(t, x, \Lambda)$$

it is readily seen that both $W_{\Omega}^{(n)}(t, x, \Lambda)$ and $P_{\Omega}^{(n)}(t, x, \Lambda)$ are non-negative completely additive functions of sets $\Lambda \in \mathfrak{B}$. Furthermore

$$(76) \quad P(t, x, \Lambda) = \sum_{n=0}^{\infty} P_{\Omega}^{(n)}(t, x, \Lambda) + \sum_{n=0}^{\infty} W_{\Omega}^{(n)}(t, x, \Lambda).$$

$P^{(n)}(t, x, \Lambda)$ can be interpreted as the compound probability that the state $X(t)$, starting the point $x \in \Omega$ at $t=0$, will during the time t change by exactly n jumps and in such a way that it remains contained in Ω during the whole time and is contained in Λ at the moment t . Of course $P_{\Omega}^{(n)}(t, x, E - \Omega) = 0$ for $x \in \Omega$.

Now (33) reads in our present notation

$$(77) \quad \begin{aligned} P^{(n+1)}(t, x, \Lambda) + \int_0^t d\sigma \int_{\Lambda} p(y) P^{(n+1)}(\sigma, x, dE_y) \\ = \int_0^t d\sigma \int_E p(y) \Pi(y, \Lambda) P^{(n)}(\sigma, x, dE_y). \end{aligned}$$

It is easily seen that the same calculations lead, for $P_{\Omega}^{(n+1)}(t, x, \Lambda)$, to the analogous formula (supposing, of course, $x \in \Omega$)

$$(78) \quad \begin{aligned} P_{\Omega}^{(n+1)}(t, x, \Lambda) + \int_0^t d\sigma \int_{\Lambda} p(y) P_{\Omega}^{(n+1)}(\sigma, x, dE_y) \\ = \int_0^t d\sigma \int_E p(y) \Pi(y, \Lambda \cap \Omega) P_{\Omega}^{(n)}(\sigma, x, dE_y). \end{aligned}$$

Subtracting (78) from (77), we get by (75)

$$(79) \quad \begin{aligned} W_{\Omega}^{(n+1)}(t, x, \Lambda) + \int_0^t d\sigma \int_{\Lambda} p(y) W_{\Omega}^{(n+1)}(\sigma, x, dE_y) \\ = \int_0^t d\sigma \int_E p(y) \Pi(y, \Lambda - \Lambda \cap \Omega) P_{\Omega}^{(n)}(\sigma, x, dE_y) \\ + \int_0^t d\sigma \int_E p(y) \Pi(y, \Lambda) W_{\Omega}^{(n)}(\sigma, x, dE_y). \end{aligned}$$

Now for any $x \in \Omega$ and any Λ

$$\int_0^t P_{\Omega}^{(0)}(\sigma, x, \Lambda) d\sigma \leq \frac{1}{p(x)} \delta(x, \Lambda\Omega) = \int_{\Lambda+} \frac{1}{p(y)} \Pi_{\Omega}^{(0)}(x, dE_y),$$

and from (78) we get

$$\int_{\Lambda} p(y) \int_0^t P_{\Omega}^{(n+1)}(\sigma, x, dE_y) d\sigma \leq \int_E p(y) \Pi(y, \Lambda\Omega) \int_0^t P_{\Omega}^{(n)}(\sigma, x, dE_y) d\sigma,$$

and thus by induction

$$(80) \quad \int_{\Lambda} p(y) \int_0^t P_{\Omega}^{(n)}(\sigma, x, dE_y) d\sigma \leq \Pi_{\Omega}^{(n)}(x, \Lambda\Omega)$$

or

$$(81) \quad \int_0^t P_{\Omega}^{(n)}(\sigma, x, \Lambda) d\sigma \leq \int_{\Lambda+} \frac{1}{p(y)} \Pi_{\Omega}^{(n)}(x, dE_y)$$

(the convergence of the right-hand member being guaranteed by (72)).

It follows from (81) and (72) that

$$\sum_{n=0}^{\infty} \int_0^t P_{\Omega}^{(n)}(\sigma, x_0, \Omega) d\sigma \leq a,$$

and since $P^{(n)}(\sigma, x_0, E - \Omega) = 0$ we can also write

$$(82) \quad \sum_{n=0}^{\infty} \int_0^t P_{\Omega}^{(n)}(\sigma, x_0, E) d\sigma \leq a.$$

Next, we deduce a limitation for $W_{\Omega}^{(n)}(t, x_0, E)$. Putting for $n \geq 1$

$$\alpha_n = \int_0^t d\sigma \int_E p(y) \Pi(y, E - \Omega) P_{\Omega}^{(n-1)}(\sigma, x_0, dE_y)$$

we readily get by (80)

$$0 \leq \alpha_n \leq \int_{\Omega} \Pi(y, E - \Omega) \Pi_{\Omega}^{(n-1)}(x_0, dE_y) = \Pi_{\Omega}^{(n)}(x_0, E - \Omega).$$

Hence, by (57) and (59),

$$\sum_{n=1}^{\infty} \alpha_n \leq 1 - \alpha < 1.$$

By definition $W_{\Omega}^{(0)}(t, x_0, E) = 0$. Hence we obtain from (79)

$$\sum_{n=0}^N W_{\Omega}^{(n)}(t, x_0, E) = \sum_{n=1}^N \alpha_n - \int_0^t d\sigma \int_{E-\Omega} p(y) W_{\Omega}^{(N)}(\sigma, x_0, dE_y) \leq 1 - \alpha$$

or

$$(83) \quad \sum_{n=0}^{\infty} W_{\Omega}^{(n)}(t, x_0, E) \leq 1 - \alpha.$$

Combining (83) with (82) we get finally, using (76),

$$\int_0^t P(\sigma, x_0, E) d\sigma \leq a + (1 - \alpha)t$$

or, for t sufficiently large,

$$\int_0^t P(\sigma, x_0, E) d\sigma < t.$$

It follows that $P(t, x_0, E) \neq 1$ which proves the necessity of our condition.

A few words may be added about the meaning of the conditions of Theorem 7. Suppose that there is a set $\Omega = \Omega_+$ such that $\Pi_{\Omega}^{(n)}(x_0, \Omega) > \eta > 0$ for some $x_0 \in \Omega$, and such that (61) converges. Consider, then, a random point moving at given moments in E by jumps according to the probability laws expressed by the ordinary Markoff chain (55). It is obvious that, if the point remained in Ω during the first n steps, the probability of never leaving Ω tends to 1. Thus, for any $\epsilon > 0$, there are points $x \in \Omega$ for which $\Pi_{\Omega}^{(n)}(x, \Omega) > 1 - \epsilon$. The proof given for the sufficiency of our condition shows that also for all these points the series (61) will converge. Thus, in the statement of the Theorem 7 α can be replaced by $1 - \epsilon$.

Denote now, as before, by Λ_a the set of points x , with $p(x) \leq a$. The convergence of (61) implies the convergence of

$$\sum_{n=0}^{\infty} \Pi^{(n)}(x, \Lambda_a \Omega)$$

for any fixed a , and accordingly there is some sequence $a_n \uparrow \infty$ such that even

$$\sum_{n=0}^{\infty} \Pi^{(n)}(x, \Lambda_{a_n} \Omega)$$

converges. This means, however, that there is a probability $\eta > 0$ for our moving point to be contained for all n after n steps in $\Omega - \Omega \Lambda_{a_n}$, that is to say, in a part of Ω , where $p(x) > a_n \uparrow \infty$. In other words, there is a positive probability that our moving point will move, in the mean, towards points with increasing $p(x)$; and if it did so for the first n steps, the probability that it will continue tends to unity as $n \rightarrow \infty$. Thus, in terms of the ergodic properties of the Markoff chain (55), the convergence of (61) is only possible if the point x is contained in a dissipative part of E .

Now the same reasoning applies also to the change of the state $X(t)$ under

the influence of our stochastic process (cf. the interpretation of (55) on page 506). Roughly speaking, if $P(t, x, E) < 1$, the difference $1 - P(t, x, E)$ can be interpreted as the probability that the state X will, starting from the point x , change during the time t by infinitely many jumps. The n th jump takes X in the set $\Omega_n \subset \Omega$ and $\Omega_n \rightarrow 0$. It follows from Theorem 7 in particular that we have $P(t, x, E) = 1$ for any point x belonging to an ergodic part of E —that is to say, if there is some bounded set Λ such that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \Pi^{(n)}(x, \Lambda) > 0.$$

7. Examples.

(i) Consider the case of an enumerable E , with the points x_0, x_1, \dots , and of a temporally homogeneous process. Let the p_i be any given positive constants and

$$\Pi_{ik} = \begin{cases} 1 & \text{for } k = i + 1, \\ 0 & \text{for } k \neq i + 1. \end{cases}$$

That is to say, from x_i only a transition to x_{i+1} is possible, and the probability of such a transition during an interval of length Δt is $p_i \Delta t + o(\Delta t)$.

The differential equations (13) of the process take on the form

$$(84) \quad P'_{ik}(t) = -p_k P_{ik}(t) + p_{k-1} P_{i,k-1}(t)$$

so that

$$(85) \quad P_{ik}(t) = 0 \text{ for } k < i, \quad P_{ii}(t) = e^{-p_i t}$$

and

$$P_{ik}(t) = p_{k-1} \int_0^t \exp\{-p_k^{(t-\sigma)}\} P_{i,k-1}(\sigma) d\sigma \text{ for } k > i.$$

In the case that $p_i \neq p_k$ for $i \neq k$, the explicit solution is for $k > i$

$$P_{ik}(t) = (-1)^{k-i} p_i p_{i+1} \cdots p_{k-1} \sum_{\nu=i}^n \frac{e^{-p_\nu t}}{(p_\nu - p_i)(p_\nu - p_{i+1}) \cdots (p_\nu - p_{\nu-1})(p_\nu - p_{\nu+1}) \cdots (p_\nu - p_n)};$$

it can be verified by means of the Lagrange interpolation formula but is of little use. The solution in the case in which $p_i \neq p_k$ for $i \neq k$ is not necessarily true, follows by the usual passage to the limit. We have

$$\Pi_{ik}^{(n)} = \begin{cases} 0 & \text{for } k \neq i + n \\ 1 & \text{for } k = i + n \end{cases}$$

and thus, by Theorem 7, *the necessary and sufficient condition for $\sum_k P_{ik}(t) = 1$ is that $\sum_{n=0}^{\infty} 1/p_{i+n}$ diverges⁽²⁰⁾*. This can also be easily verified directly. Putting

$$(86) \quad L_{i,k}(t) = \int_0^t p_k P_{ik}(\sigma) d\sigma,$$

it follows from (84) for $k > i$ that

$$(87) \quad P_{ik}(t) + L_{ik}(t) = L_{i,k-1}(t)$$

while $P_{ii}(t) + L_{ii}(t) = 1$ and thus $L_{i,i+n}(t) \downarrow L_i(t)$ as $n \rightarrow \infty$. On the other hand we have by (87)

$$\sum_{k=i}^{i+n} P_{ik}(t) = 1 - L_{i,i+n}(t)$$

or

$$(88) \quad \sum_{k=0}^{\infty} P_{ik}(t) = 1 - L_i(t) \leq 1.$$

But by (86) and (88), $L_i(t) > 0$ would imply

$$(89) \quad t \geq \sum_{k=0}^{\infty} \int_0^t P_{ik}(\sigma) d\sigma \geq L_i(t) \sum_{k=i}^{\infty} \frac{1}{p_k},$$

that is to say, the convergence of $\sum_k 1/p_{k+i}$. Conversely, by (89) the divergence of $\sum_k 1/p_{k+i}$ implies that $L_i(t) = 0$, or by (88) that $\sum_k P_{ik}(t) = 1$. A similar argument can be applied even in the case that the p_i depend on t .

The stochastic process just described plays an important role for different applications. In the case that all p_i are equal, $p_i = p$, it reduces to the classical Poisson process

$$P_{ik}(t) = e^{-pt} \frac{(pt)^{k-i}}{(k-i)!} \quad (k \geq i).$$

The general case was used by Lundberg [7] in the theory of invalidity insurance, and by Feller [4] to describe the growth of some biological populations. In both cases it is natural to assume that $p_i \rightarrow \infty$ as $i \rightarrow \infty$. The same stochastic process was also applied to describe radioactive processes, x_i standing for the "elementary probability" of its disintegration; but here, of course, the space E contains only a finite number of points (or, what amounts to the same, some $p_k = 0$).

(ii) Finally we give an example which proves that the condition of the corollary to Theorem 7 is not sufficient.

Let E consist of the points $x_i, i = 0, \pm 1, \pm 2, \pm 3, \dots$. The process is

⁽²⁰⁾ The vanishing of any particular p_n obviously implies that $P_{ik}(t) = 0$ for any couple (i, k) with $i \leq n < k$; and it is readily seen that $\sum_k P_{ik}(t) = 1$ for any $i \leq n$.

again temporally homogeneous. For $i < 0$ only the transition $x_i \rightarrow x_{i-1}$ is possible; for $i \geq 0$ both $x_i \rightarrow x_{i+1}$ and $x_i \rightarrow x_{i-1}$ are possible, the corresponding probabilities being $1 - \pi_i > 0$ and $\pi_i > 0$. In other words we suppose that

$$\Pi_{ik} = \begin{cases} 1 & \text{if } i < 0, k = i - 1, \\ 1 - \pi_i & \text{if } i \geq 0, k = i + 1, \\ \pi_i & \text{if } i \geq 0, k = -i - 1, \\ 0 & \text{otherwise.} \end{cases}$$

Let us now suppose that (i) the product $\prod(1 - \pi_i) = \alpha > 0$, (ii) $p_i = 1$ for $i \leq 0$, and (iii) $\sum_{i=1}^{\infty} 1/p_i = a$ converges. Then the condition of the corollary to Theorem 7 is satisfied. For obviously we have if $n > 0, i \geq 0$,

$$\Pi_{i, i-n}^{(n)} = \pi_i + (1 - \pi_i)\pi_{i+1} + \cdots + (1 - \pi_i)(1 - \pi_{i+1}) \cdots (1 - \pi_{i+n-1})\pi_{i+n}$$

and if $i < 0$

$$\Pi_{i, i-n}^{(n)} = 1.$$

Hence

$$\sum_n \sum_k \frac{1}{p_k} \Pi_{ik}^{(n)} \geq \sum_n \Pi_{i, -|i|-n}^{(n)} \geq \sum_n \pi_i$$

diverges. But, taking for Ω the set of all points x_i with $i \geq 0$, it is readily seen that (59) holds for any $x = x_i, i \geq 0$, and nevertheless the series converges.

REFERENCES

1. J. L. Doob, *Stochastic processes depending on a continuous parameter*, these Transactions, vol. 42 (1937), p. 107.
2. W. Dubrovski, *Eine Verallgemeinerung der Theorie der rein unstetigen stochastischen Prozesse von W. Feller*, Comptes Rendus (Doklady) de l'Académie des Sciences de l'URSS, vol. 19 (1938), p. 439.
3. W. Feller, *Zur Theorie der stochastischen Prozesse (Existenz- und Eindeigkeitssätze)*, Mathematische Annalen, vol. 113 (1936), p. 113.
4. ———, *Die Grundlagen der Volterraschen Theorie des Kampfes ums Dasein in wahrscheinlichkeitstheoretischer Behandlung*, Acta Biotheoretica, vol. 5 (1939), p. 11.
5. M. Fréchet, *Recherches Théoriques Modernes sur le Calcul des Probabilités*, Part II (Traité du Calcul des Probabilités, vol. 1, no. 3), 1938.
6. A. Kolmogoroff, *Ueber die analytischen Methoden in der Wahrscheinlichkeitsrechnung*, Mathematische Annalen, vol. 104 (1931), p. 415.
7. O. Lundberg, forthcoming dissertation, Stockholm.
8. G. Pólya, *Sur la promenade au hasard dans un réseau des rues*, Lecture at the "Colloque Consacré à la Théorie des Probabilités," Geneva, 1937, Actualités Scientifiques et Industrielles, no. 734, 1938, p. 25.
9. Added in proof; cf. the footnote on page 492: W. Doeblin, *Sur certains mouvements aléatoires discontinus*, Skandinavisk Aktuarietidskrift, 1939, p. 211.

BROWN UNIVERSITY,
PROVIDENCE, R. I.