INFINITESIMAL OPERATORS OF MARKOV PROCESSES

E. B. DYNKIN

In 1931 A. N. Kolmogorov [9] showed that a wide class of one-dimensional Markov processes can be described by differential equations of the form

\[ \frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + b \frac{\partial u}{\partial x}. \]

Kolmogorov's work left unanswered the question of whether there exist one-dimensional Markov processes which can be described by differential equations of the form

\[ \frac{\partial u}{\partial t} = \mathcal{A}u, \]

where \( \mathcal{A} \) is a differential operator of order greater than 2. Various investigators attempted to derive similar equations of higher order, using various assumptions with respect to the transition probabilities. These attempts were of a highly doubtful character, however, since no one was able to construct an example to show that such processes can actually exist.

The question was solved completely in 1954--1955 by W. Feller [11, 12], who showed that a one-dimensional process with a continuous path function can always be described by (II) where \( \mathcal{A} \) is a generalized second derivative.

Feller's purely analytic method is related essentially to the one-dimensionality of the problem, and it is not easily extended to processes in 2 and higher dimensional spaces.

In the present work we develop another method, based on a consideration of the probability properties of path functions. This method is applicable both to one-dimensional and many-dimensional problems, and can be used to prove that a continuous Markov process in a space of any dimensionality is given by (II) where \( \mathcal{A} \) is a generalized elliptic differential operator of second order. The method is valid not only for continuous processes, but for any Markov process whose path function is continuous on the right.

In Feller's work it is assumed that the operator semigroup of the Markov process transforms continuous functions into continuous ones. The class of processes which is so specified is indeed very important. We shall call it the class of Feller processes. Our method is, however, valid for a wider class of processes, which we call strong Markov processes. The usual definition of a Markov process requires that its development after a time \( t \) is independent of its history if its state at time \( t \) is known. A strong Markov process shall be one...
for which this condition is satisfied not only for a constant \( t \) but also for a random \( t \) which is "independent of the future". All Feller processes are strong Markov processes \([6]\), but the converse is not in general true.

Our method is applicable also to strong Markov processes which are not Feller processes.

The present work starts with an introduction, in which we review briefly the basic facts concerning Markov and strong Markov processes and their operator semigroups. This is based on the results of two of the author's previous works \([5, 6]\). The basic method for calculating the infinitesimal operators of Markov processes is given in Section 2. Feller processes are studied in Section 3. The operator \( \mathcal{A} \), constructed in this paragraph plays in general the same role as the differential operator on the right side of (1) plays in diffusion processes, or the transition probability density matrix plays in processes with a denumerable set of states. The operator \( \mathcal{A} \) alone is not in general sufficient to define a Markov process. It may also be necessary to establish certain subsidiary conditions, such as the boundary conditions in the case of a diffusion process, the probability distribution at the commencement of transitions in the case of a process with a denumerable set of states, etc. In the language of semigroup theory this means that in general the infinitesimal operator \( A \) of a Markov process is a contraction of the operator \( \mathcal{A} \) (see Theorem 4). It is proved in Theorem 5, however, that if the set of states is compact, then \( \mathcal{A} = A \). In Section 4 we study Feller processes with continuous path functions. It is shown that for this case the operator \( \mathcal{A} \) constructed in Section 3 is the only generalization possible of an elliptic differential operator of order 2.

We go into more detail with respect to continuous Feller processes on a straight line, and it is shown that for such processes \( \mathcal{A} \) reduces to the Feller generalized second derivative. This clarifies the probability theoretical meaning of the two functions used to construct the generalized derivative. Section 5 is devoted to the consideration of some discontinuous processes \(^1\).

The results of the present work can be used to classify all one-dimensional strong Markov processes (including those which are not Feller processes) \(^1\). These results have been formulated elsewhere \([8]\). They shall be proved in a separate article.

### 1. Introduction

1. Let \( x(t, \omega) \) be a function of the real variable \( t (0 \leq t < \infty) \) and the point \( \omega \) of the subsidiary "space of elementary events" \( \Omega \). This function lies in the metric space \( \mathcal{E} \). We shall denote by \( \mathcal{B} \) the \( \sigma \)-algebra generated by the open sets of \( \mathcal{E} \), by \( \mathcal{B} \) the \( \sigma \)-algebra generated by the intervals on the line \( 0 \leq t < \infty \), and by \( \mathcal{W} \) the \( \sigma \)-algebra of submanifolds of \( \Omega \) which is generated by the sets \( \{ x(t, \omega) \in \Gamma \} \ (t \geq 0, \Gamma \in \mathcal{W}) \). We shall assume that on \( \mathcal{W} \) there is given a family of probability measures \( P_\omega(x \in \mathcal{E}) \), such that \( P_\omega(x(0, \omega) = x) = 1 \) and that for any \( \Gamma \in \mathcal{W} \), the function \( P_\omega(x(t, \omega) \in \Gamma) \) is, in \( \mathcal{B} \times \mathcal{W} \), a measurable function of \( t \) and \( x \). Finally we shall assume that for all \( \omega \in \Omega \) the function

\(^1\) Some of the results of the present article have already been published \([7]\) without detailed proofs.
$x(t, \omega)$ is continuous on the right. If all these assumptions are fulfilled, we shall say that we have specified in $\mathcal{F}$ a random process $x(t, \omega)$ continuous on the right.

Let $\zeta(\omega)$ be an $\mathcal{M}$-measurable function on $\Omega$. We shall denote by $M_{x(\cdot)}$ the integral

$$\int_{\Omega} \zeta(\omega) P_x(d\omega).$$

We note that $x(t, \omega)$ is measurable in $\mathbb{B} \times \mathbb{B}$. Indeed, let us write $x_n(t, \omega) = x((k+1)/n, \omega)$ for $k/n \leq t < (k+1)/n$, $(k = 0, 1, 2, \cdots)$. The $x_n(t, \omega)$ are measurable functions, and

$$x(t, \omega) = \lim_{n \to \infty} x_n(t, \omega).$$

We note further that for any neighborhood $U$ of an arbitrary point $x$ we have

$$\lim_{t \to 0} P\{x(t) \in U\} = 1.$$ (1)

We shall denote by $\mathcal{M}_t$ the $\sigma$-algebra generated by the sets $\{x(u, \omega) \in \Gamma\}$ $(0 \leq u \leq t, \Gamma \in \mathcal{B})$. We shall say that the nonnegative random quantity $\tau = \tau(\omega)$ is independent of the future if for any $t$ the set $\{t(\omega) \leq t\}$ belongs to the $\sigma$-algebra $\mathcal{M}_t$. The sets $A \in \mathcal{M}$ which for arbitrary $t$ satisfy the condition $t\{\tau(\omega) \leq t\} \cap A \in \mathcal{M}$ form a $\sigma$-algebra which we shall call $\mathcal{M}_\tau$. This notation agrees with the above; that is to say if the random variable $\tau$ is some constant $t$, then $\mathcal{M}_\tau$ coincides with $\mathcal{M}_t$. The conditional probabilities with respect to the $\sigma$-algebra $\mathcal{M}_\tau$ will be denoted by the symbol $P(\cdot | x(u)_{u \leq \tau})$.

A random process is called a Markov process if for any positive constant $\tau$ and for any $\Gamma_1, \cdots, \Gamma_n \in \mathcal{B}$, $0 \leq t_1 < t_2 < \cdots < t_n$ we have with given probability

$$P(x(\tau+t_1) \in \Gamma_1, \cdots, x(\tau+t_n) \in \Gamma_n | x(u)_{u \leq \tau})$$ (2)

$$= P_{x(\tau)}(x(t_1) \in \Gamma_1, \cdots, x(t_n) \in \Gamma_n).$$

If this condition is fulfilled not only for every constant, but also for any random variable $\tau = \tau(\omega)$ which is independent of the future, then we shall call the random process a strong Markov process.

The relations between Markov and strong Markov processes have been studied by the author and Iushkevich [6]. In particular, we have shown that a Markov process continuous on the right is a strong Markov process if $g(x) = M_{x(t)}$ is a continuous function of $x$ for all continuous and bounded functions $f(x)$.

We shall rewrite condition (2) that a process be a strong Markov process in a form which we will find more convenient to use in what follows.

We denote by $\Omega_\tau$ the set of all $\omega$, such that $\tau(\omega) < \infty$ (the function $\tau = \tau(\omega)$ may take on the value $\infty$ in addition to finite values). We define the homomorphism $\theta_\tau$ of the $\sigma$-algebra $\mathcal{M}$ into itself by the following condition: if $A = \{x(t, \omega) \in \Gamma\}$ $(t \geq 0, \Gamma \in \mathcal{B})$, then $\theta_\tau A = \Omega_\tau \cap \{x(\tau(\omega)+t, \omega) \in \Gamma\}$. Let

$^2$ The definitions and properties of conditional probabilities and conditional mathematical expectations with respect to a $\sigma$-algebra can be found in Doob's work [2]. We drop the index $x$ in the expression $P_x(A | B)$ if this probability does not depend on $x$.

$^3$ We should, perhaps, add "homogeneous in time." Since, however, the present work treats only processes homogeneous in time, we shall make no special mention of this fact.
\( \xi(\omega) \) and \( \eta(\omega) \) be two \( \mathcal{M} \)-measurable functions. We shall write \( \eta = \theta_\tau \xi \) if \( \theta_\tau \{ \xi(\omega) = c \} = \Omega_\tau \cap \{ \eta(\omega) = c \} \) for any \( c \). (It is clear that \( \eta \) is defined by \( \xi \) on the set \( \Omega_\tau \), but that it remains undefined outside of \( \Omega_\tau \).)

We shall now show that if \( x(t, \omega) \) is a strong Markov process and if \( \tau \) is a random variable independent of the future, then for all \( A \in \mathcal{M}_\tau \) \( (A \subset \Omega_\tau) \) and \( \mathfrak{B} \in \mathcal{M} \) we have

\[
P_\tau(A \cap \theta_\tau B) = \int_A P_\tau(\tau | B) P_\tau(d\omega).
\]

For all \( \mathcal{M}_\tau \)-measurable functions \( \xi \) and \( \mathcal{M} \)-measurable functions \( \zeta(\omega) \), if \( P_\tau(\xi > 0, \tau = \infty) = 0 \), then

\[
M_\tau[\xi \theta_\tau \zeta] = M_\tau[\xi M_\tau(\tau) \zeta]_*.
\]

To prove (3) we need only note that the left and right sides of this equation are countably additive functions of \( B \), and if \( B = \{ x(t_1) \in \Gamma_1, \cdots, x(t_n) \in \Gamma_n \} \) \( (0 \leq t_1 < \cdots < t_n, \Gamma_1, \cdots, \Gamma_n \in \mathfrak{B}) \), then (3) is equivalent to (2) according to the definition of conditional probability. In order now to derive (4), we note the following: (a) if \( \xi \) and \( \zeta \) are characteristic functions of any sets, then (4) becomes (3); (b) both sides of (4) depend linearly on \( \xi \) and \( \zeta \); (c) if (4) is fulfilled for any functions \( \xi_n, \zeta_n \) such that \( \xi_1 \leq \xi_2 \leq \cdots, \zeta_1 \leq \zeta_2 \leq \cdots \), then it is also fulfilled for the limit functions \( \lim \xi_n, \lim \zeta_n \). If we are dealing with a Markov process, but not a strong Markov process, then (3) and (4) are fulfilled for \( \tau = \text{const.} \)

2. Let us now denote by \( B \) the Banach space of \( \mathfrak{B} \)-measurable bounded functions on \( \mathcal{E} \) whose norm is \( ||f|| = \sup_{x \in \mathcal{E}} |f(x)| \) and let us set

\[
T_t f(x) = \int_\Omega f(x(t, \omega)) P_\tau(d\omega) = M_\tau f(x(t)).
\]

This formula defines uniquely a one-parameter family of linear operators in \( \mathcal{E} \) and for these operators \( ||T_t f|| \leq ||f|| \). If \( x(t, \omega) \) is a Markov process, then for all \( s \geq 0, t \geq 0 \) we have

\[ T_{s+t} = T_s T_t, \]

so that the \( T_t \) operators form a contraction semigroup.

Let \( f_n, f \in B \). We write \( f = s \lim f_n \) if \( ||f_n - f|| \to 0 \), and \( f = w \lim f_n \) if \( f_n(x) \to f(x) \) for all \( x \in \mathcal{E} \) and if \( ||f_n|| \) are bounded. The set of all \( f \in B \) for which \( s \lim_{t \to 0} T_t f = f \) shall be denoted by \( B_0 \). The set of all \( f \in B \) for which \( w \lim_{t \to 0} T_t f = f \) shall be denoted \( \bar{B}_0 \). If

\[
g(x) = s \lim_{t \to 0} \frac{T_{-t} f(x) - f(x)}{t} \quad (f, g \in B_0),
\]

we shall write \( f \in D_A, A f = g \). If

\[
g(x) = w \lim_{t \to 0} \frac{T_{-t} f(x) - f(x)}{t} \quad (f, g \in \bar{B}_0),
\]

we shall write \( f \in D_{\bar{A}}, \bar{A} f = g \). The operators \( A \) and \( \bar{A} \) are called the infinitesimal operators of the Markov process. Of these, \( A \) is called strong and \( \bar{A} \) is called weak.

The set of bounded linear operators \( R_\lambda (\lambda > 0) \) defined by

\[
R_\lambda h(x) = \int_0^\infty e^{-\lambda t} T_t h(x) dt \quad (h \in B),
\]

is called the resolvent of the semigroup \( T_t \).
The following facts are fundamental [5].

A. A Markov process continuous on the right is defined uniquely by its infinitesimal operator, either strong or weak (which means that either $A$ or $\bar{A}$ can be used to establish uniquely the family of probability measures $P_z$ in $\Omega$).

B. For the sets and infinitesimal operators we may write

$$B_0 \subseteq \bar{B}_0, \quad A \subseteq \bar{A}$$

(the second of these means that $D_A \subseteq D_{\bar{A}}$, and if $f \in D_A$, then $Af = \bar{A}f$).

C. For any $\lambda > 0$ the operator $\lambda E - A$ is a one-to-one mapping of $D_A$ onto $B_0$. The operator $\lambda E - \bar{A}$ is a one-to-one mapping of $D_{\bar{A}}$ onto $\bar{B}_0$. The operator $R_\lambda$ is equal on $B_0$ to $(\lambda E - A)^{-1}$ and on $\bar{B}_0$ it is equal to $(\lambda E - \bar{A})^{-1}$.

2. The Resolvent and Infinitesimal Operators of a Strong Markov Process

3. Comparing equations (5) and (8), we obtain

$$R_\lambda h(x) = \int_0^\infty \int_0 e^{-\lambda t} h[x(t, \omega)] P_z(\omega) dt$$

$$= \int A \int_0^\infty e^{-\lambda t} h(x(t, \omega)] P_z(\omega) dt = M_x \int_0^\infty e^{-\lambda t} h[x(t)] dt.$$

(Changing the order of integration is valid, since the integrand is measurable and absolutely integrable.)

**Theorem 1.** Let $x(t, \omega)$ be a strong Markov process continuous on the right. Let $\tau = \tau(\omega)$ be a random variable independent of the future, and let $f(x) = R_\lambda h(x)$ $(h \in B, \lambda > 0)$. Then

$$M_x \{e^{-\lambda \tau} f[x(\tau)]\} - f(x) = -M_x \int_0^\tau e^{-\lambda t} h[x(t)] dt.$$

**Proof.** From (9) we have

$$f(x) = M_x \int_0^\infty e^{-\lambda t} h[x(t)] dt + M_x \int_0^\infty e^{-\lambda t} h[x(t)] dt.$$

Substituting for the variable of integration, we may write

$$\int_\tau^\infty e^{-\lambda t} h[x(t)] dt = e^{-\lambda \tau} \int_0^\infty e^{-\lambda s} h[x(s + \tau)] ds = e^{-\lambda \tau} \theta_{\tau} \left[ \int_0^\infty e^{-\lambda s} h[x(s)] ds \right].$$

Since $e^{-\lambda \tau}$ is measurable with respect to $\mathbb{M}_\tau$ and since $P_x(e^{-\lambda t} > 0, \tau = \infty) = 0$, according to equation (4),

$$M_x \left[ e^{-\lambda \tau} \theta_{\tau} \left[ \int_0^\infty e^{-\lambda s} h[x(s)] ds \right] \right] = M_x \left[ e^{-\lambda \tau} f[x(\tau)] \right].$$

Equations (11)—(13) give (10).

**Theorem 2.** Let $x(t)$ be a strong Markov process continuous on the right, and be a random variable which is independent of the future and for which $M_x \tau < \infty$. Let $\bar{A}$ be the weak infinitesimal operator of the process $x(t)$, and write $\bar{A}f = g$. Then

Since $A \subseteq \bar{A}$, the theorem is valid also for the strong infinitesimal operator. Furthermore, our proof is valid both for the strong and weak case.
Infinitesimal operators of Markov processes

39

\( M_x[f(x)] - f(x) = M_x \int_0^\tau g[x(t)] dt \).

**Proof.** Let us write \( h = \lambda f - g \). In view of remark C of paragraph 2, \( R_x h = f \), and according to Theorem 1,

\[ M_x[e^{-\lambda \tau} f(x)] - f(x) = -M_x \int_0^\tau e^{-\lambda \tau} h[x(t)] dt. \]

Going to the limit \( \tau \to 0 \), we obtain (14).

4. Let \( U \) be an arbitrary open set in \( \mathcal{F} \). We shall denote by \( \tau_U \) the exact lower bound of all instants of time \( t \) such that the distance of the path function segment \( x(s) \) (0 \( \leq s \leq t \)) from the set \( \mathcal{F} - U \) vanishes. From the fact that \( x(s) \) is continuous on the right, it follows that the lowest bound is attained, so that \( \tau_U \) is the first instant of time at which the distance between \( x(s) \) (0 \( \leq s \leq t \)) and \( \mathcal{F} - U \) vanishes. (If \( x(s) \) is continuous, \( \tau_U \) is the first instant at which it reaches the set \( \mathcal{F} - U \).) Let us write \( V_n \) for the set of all points \( y \in \mathcal{F} \) such that \( \rho(y, \mathcal{F} - U) < 1/n \). Then

\[ \{ \tau_U(\omega) \leq t \} = \bigcap_{n=1}^{\infty} \bigcup_{r \leq t} \{ x(r, \omega) \in V_n \} \cup \{ x(t, \omega) \in \mathcal{F} - U \}, \]

where \( r \) takes on rational values. It is seen from that that \( \tau_U \) is a random variable independent of the future.

**Theorem 3.** Let \( \tilde{A} \) be the weak infinitesimal operator of a strong Markov process \( x(t) \) continuous on the right. If \( \tilde{A} f(x) \) is discontinuous at a point \( x \), and in a certain neighborhood \( U_0 \) of the point \( x \) we have \( M_x \tau_{U_0} < \infty \), then

\[ \tilde{A} f(x) = \lim_{\delta(U) \to 0} \frac{M_x f(x(\tau_U)) - f(x)}{M_x \tau_U}. \]

(\( d(U) \) denotes the diameter of the neighborhood \( U \)).

**Proof.** The random variable \( \tau_U \) is independent of the future. If \( U \subseteq U_0 \), then \( \tau_U \leq \tau_{U_0} \) so that \( M_x \tau_{U_0} < \infty \). Therefore Theorem 2 can be applied to \( \tau_U \). If \( g(x) \) is continuous at \( x \), then for all \( \varepsilon > 0 \), there is a \( \delta > 0 \) such that for arbitrary \( y \in U \) we have \( |g(y) - g(x)| < \varepsilon \), and therefore \( |g(x(t)) - g(x)| < \varepsilon \) for \( t < \tau_U \). Therefore the right side of (14) differs from \( g(x) M_x \tau_U \) by less than \( \varepsilon M_x \tau_U \) which proves (15).

5. **Lemma 1.** Let \( x(t, \omega) \) be a strong Markov process continuous on the right, and let \( U_0 \) be some region of \( \mathcal{F} \). We write

\[ m(y) = M_x \tau_{U_0} \quad (y \in \mathcal{F}). \]

If \( m(x) < \infty \) for some point \( x \) of \( U_0 \), then for any neighborhood \( U \) of a point \( x \) contained in \( U_0 \) we may write

\[ M_x \tau_U = -M_x m(x(\tau_U)) + m(x). \]

**Proof.** We note that \( \theta_{\tau_U}(\tau_{U_0}) = \tau_{U_0} - \tau_U \). Therefore in view of (4),

\[ M_x(\tau_{U_0} - \tau_U) = M_x M_x(\tau_{U_0} - \tau_U) = M_x m(x(\tau_U)), \]

which is equivalent to (16).

\[ ^5 \text{By } \rho(x, y) \text{ we mean the distance in } \mathcal{F}. \]

\[ ^* \text{See the footnote 4.} \]
If (16) for $M_z \tau_D$ is inserted into (15), we obtain the following version of Theorem 3.

**Theorem 3’.** If the conditions of Theorem 3 are fulfilled, then

$$
\tilde{A}f(x) = -\lim_{\alpha(t) \to 0} \frac{M_z[f(x(t))] - f(x)}{M_z[m(x(t))] - m(x)} = -\lim_{\alpha(t) \to 0} \frac{\int [f(y) - f(x)] \Pi_U(x, dy)}{\int [m(y) - m(x)] \Pi_U(x, dy)},
$$

where

$$
\Pi_U(x, \Gamma) = P_x(x(\tau_U) \in \Gamma).
$$

6. The proof that $M_z \tau_D$ is finite is often simplified by the following lemma.

**Lemma 2.** Let $x(t)$ be a Markov process continuous on the right, and $U$ an open set of $\mathcal{E}$. If for some $t > 0$, $\alpha > 0$ we have $P_x(\tau_U > t) < 1 - \alpha$ for all $x \in U$, then $M_z \tau_U < K < \infty$ for all $x \in U$.

**Proof.** Let us set $A_s = \{\tau_U > s\}$. For all $s \geq 0$, we have

$$
A_{s+t} = A_s \cap \theta_s A_t.
$$

Therefore in view of (3),

$$
P_x(A_{s+t}) = \int A_s P_\omega(A_t) P_x(d\omega).
$$

For $\omega \in A_s$, $x(s, \omega) \in U$ and by assumption $P_{x(s)}(A_t) = P_{x(s)}(\tau_U > t) < 1 - \alpha$. Therefore $P_x(A_{s+t}) < (1 - \alpha)^n$ and we have

$$
M_z \tau_U = \int_0^\infty P_x(\tau_U > s) ds = \sum_{n=0}^\infty \int_{nt}^{(n+1)t} P_x(\tau_U > s) ds < t \sum_{n=0}^\infty P_x(\tau_U > nt) = t \sum_{n=0}^\infty (1 - \alpha)^n = \frac{t}{\alpha}.
$$

3. Feller Processes

7. We shall call a Markov process $x(t, \omega)$ in a metric space $\mathcal{E}$ a Feller process if the space $C$ of all bounded continuous functions on $\mathcal{E}$ is invariant under the operators $T_x$. This condition means that the motion is in some sense stable. In other words the distribution of probabilities $P_x(x(t) \in \Gamma)$ for the position of a moving particle at time $t$ changes by a small amount (in the sense of weak convergence of measures) if the original point $x$ is changed by a small amount. (For more details see Feller [11].)

For Feller processes it is convenient to change somewhat the definition of the infinitesimal operators $A$ and $\tilde{A}$ of paragraph 2. In particular, we shall say that $f \in D_A$, $A f = g$ (or $f \in D_{\tilde{A}}$, $A f = g$) if in addition to (6) or (7) we have $\int_{B_0} f P_{x(0)}(d\omega)$ valid for this definition of $A$ and $\tilde{A}$ if the spaces $B_0$ and $\mathcal{B}_0$ are replaced by $C_0 = C \cap B_0$ and $\mathcal{C}_0 = C \cap \mathcal{B}_0$. (See the previous article [5], Section 4.)

All Feller processes continuous on the right are strong Markov processes (see Theorem 2 of an earlier work [6]), and the results of Section 2 can be applied to them.

8. **Lemma 3.** Let $x(t)$ be a Feller process and assume that for some $x \in \mathcal{E}$, $t > 0$ and some region $V$ we have $P_x(x(t) \in V) = \alpha > 0$. Then there exists a neighborhood $U$ of the point $x$ such that $P_y(x(t) \in V) > \alpha/2$ for all $y \in U$. 

PROOF. Let us denote by $\Gamma_n$ the set of all points $z$ such that $\rho(z, \mathcal{E} - V) \geq 1/n$. It is obvious that $V = \bigcup_n \Gamma_n$. Therefore

$$\lim_{n \to \infty} P_x\{x(t) \in \Gamma_n\} = P_x\{x(t) \in V\} = \alpha$$

and there exists an $n_0$ such that

$$P_x\{x(t) \in \Gamma_{n_0}\} > \frac{3}{4}\alpha.$$ 

Let us consider the continuous function

$$f(y) = \begin{cases} n_0 \rho(y, \mathcal{E} - V) & (y \in \Gamma_{n_0}) \\ 1 & (y \notin \Gamma_{n_0}). \end{cases}$$

In view of the fact that we are dealing with a Feller process, the function $g(y) = M_yf(x(t))$ is also continuous and there therefore exists a neighborhood $U$ of the point $x$ such that for $y \in U$ this function satisfies the condition $g(y) > g(x) - \alpha/4$. We have

$$g(y) < P_y\{x(t) \in V\}, \quad g(x) > P_x\{x(t) \in \Gamma_{n_0}\} > \frac{3}{4}\alpha.$$ 

Therefore

$$P_y\{x(t) \in V\} > \frac{1}{2}\alpha.$$ 

We shall say that a point $x \in \mathcal{E}$ is an absorption point for the process $x(t)$ if for all $t$ we have $P_x\{x(t) = x\} = 1$. In this case $T_1f(x) = M_xf(x(t)] = f(x)$ for all $f$. Therefore $\hat{A}f(x) = 0$ for all $f \in D_\hat{A}$.

**Lemma 4.** If $x$ is not an absorption point for the Feller process $x(t)$ continuous on the right, there exists a neighborhood $U$ of $x$ such that $m(y) = M_y< \infty$ for all $y \in U$.

**Proof.** There exists a region $V$ such that $\rho(x, V) > 0$ and such that for some $t$ we may write $P_x\{x(t) \in V\} = \alpha > 0$. In view of Lemma 3 we may construct a neighborhood $U$ of the point $x$ such that $P_y\{x(t) \in V\} > \alpha/2$ for all $y \in U$.

Let us choose $U$ so that $U \cap V = 0$. Then $P_y\{\tau_U > t\} < P_y\{x(t) \notin V\} < 1 - \alpha/2$ for all $y \in U$ and according to Lemma 2 we may write $M_y\tau_U < \infty$ for all $y \in U$.

It follows from Lemma 4 that in a Feller process the absorption points form a closed set.

9. Let us set

$$\mathcal{A}f(x) = \lim_{d(y) \to 0} \frac{M_xf(x(\tau_U)) - f(x)}{M_x\tau_U}.$$ 

If this limit exists for all $x$ and if $f$ and $\mathcal{A}f \in C$ we shall say that $f \in D_{\mathcal{A}}$.

**Theorem 4.** For any Feller process continuous on the right $A \subseteq \hat{A} \subseteq \mathcal{A}$.

**Proof.** The relation $A \subseteq \hat{A}$ is fulfilled according to remark B of paragraph 2. We shall show that $\hat{A} \subseteq \mathcal{A}$. Let $f \in D_{\hat{A}}$. If $x$ is an absorption point, then

\[ \text{Thus if } x \text{ is an absorption point, then } \mathcal{A}f(x) = 0. \]
\( \tilde{A}f(x) = 0 \). On the other hand we have by assumption \( \mathcal{U}f(x) = 0 \). If \( x \) is not an absorption point, then according to Lemma 2 there exists a neighborhood \( U_0 \) of \( x \) such that \( M_x \tau_{U_0} < \infty \) and according to Theorem 3, \( \mathcal{U}f(x) = \tilde{A}f(x) \). Thus \( \mathcal{U}f(x) = \tilde{A}f(x) \) for all \( x \in \mathcal{E} \) and \( f \in D_{\mathcal{A}} \).

**Theorem 5.** Let \( x(t, \omega) \) be a Feller process continuous on the right. If \( \mathcal{E} \) is compact, then

\[
A = \tilde{A} = \mathcal{A}.
\]

**Proof.** According to Theorem 5 of the previous article \([5, A] \) and \( \mathcal{C}_0 = C_0 = C \). According to Theorem 4, \( \tilde{A} \subseteq \mathcal{A} \), and therefore it is sufficient to show that \( D_{\mathcal{A}} \subseteq D_{\tilde{A}} \).

Let \( f \in D_{\mathcal{A}} \). We write \( h = f - \mathcal{A}f, \tilde{f} = R_n h = \int_0^\infty e^{-\epsilon t} h \, dt \). Since \( h \in C = \mathcal{C}_0 \), it follows that (see remark C of paragraph 2) \( \tilde{f} \in D_{\tilde{A}} \) and \( \tilde{f} - \tilde{A}\tilde{f} = h \). Since \( \tilde{A} \subseteq \mathcal{A} \), we have \( \tilde{f} - \mathcal{A}\tilde{f} = h \). It is seen from this that \( \varphi = f - \tilde{f} \) satisfies the equation

\[
(\text{is})
\]

In view of the fact that \( \mathcal{E} \) is compact, \( \varphi(x) \) takes on its maximum value at some point \( x_0 \). From (17) it is seen that \( \mathcal{U}\varphi(x_0) \leq 0 \). According to (18), \( \varphi(x_0) \leq 0 \) so that \( \varphi(x) \leq 0 \) for all \( x \in \mathcal{E} \). It can be similarly proved that for all \( x \in \mathcal{E} \) we have \( \varphi(x) \leq 0 \). Therefore \( \varphi = 0 \) and \( f = \tilde{f} \in D_{\tilde{A}} \).


10. In this section we shall consider only continuous 8 Feller processes.

As has already been pointed out in paragraph 4, if \( x(t) \) is continuous, \( \tau_U \) is the first instant at which this function reaches the closed set \( \mathcal{E} - U \). It is clear that \( x(\tau_U) \) belongs to the boundary \( U' \) of the region \( U \), and (17) can be written in the form

\[
\mathcal{U}f(x) = \lim_{a(U) \to 0} \frac{\int_{U'} [f(y) - f(x)] H_U(x, dy)}{c_{U}(x)},
\]

where

\[
H_U(x, \Gamma) = P_x(x(\tau_U) \in \Gamma); \quad c_{U}(x) = M_x \tau_U.
\]

If \( U_0 \) is a neighborhood of \( x \) such that \( m(x) = M_x \tau_{U_0} < \infty \), then according to paragraph 5, (19) can be written in the form

\[
\mathcal{U}f(x) = -\lim_{a(U) \to 0} \frac{\int_{U'} [f(y) - f(x)] H_U(x, dy)}{\int_{U'} [m(y) - m(x)] H_U(x, dy)}.
\]

The operator \( \mathcal{U} \) defined by (19) and (20) has the following properties:

1) If \( f_1(x) = f_2(x) \) in some neighborhood of \( x_0 \), then \( \mathcal{U}f_1(x_0) = \mathcal{U}f_2(x_0) \).

2) If \( f(x) \) takes on its maximum value at \( x_0 \), then \( \mathcal{U}f(x_0) \geq 0 \).

3) Assume that \( \mathcal{U} \) is defined at \( x_0 \) for some continuous functions \( \varphi_i (i = 1, 2, \cdots, n) \) and their products by pairs \( \varphi_i \varphi_j (i, j = 1, 2, \cdots, n) \)

---

8 A process \( x(t, \omega) \) is called continuous if for any \( \omega \in \Omega \) the function \( x(t, \omega) \) is continuous in \( t \).
and let $F(y_1, \cdots, y_n)$ be any function of $n$ real variables with continuous first and second derivatives in the neighborhood of $\varphi_1(x_0), \cdots, \varphi_n(x_0)$. Then $\mathfrak{A}$ is defined at $x_0$ for the function $F(\varphi_1, \cdots, \varphi_n)$ and we have

$$
\mathfrak{A}F(\varphi_1, \cdots, \varphi_n) = \sum_{i=1}^{n} a_i \frac{\partial F}{\partial \varphi_i} + \sum_{i,j=1}^{n} b_{ij} \frac{\partial^2 F}{\partial \varphi_i \partial \varphi_j},
$$

where

$$a_i = \mathfrak{A}(\varphi_i - \varphi_i^0), \quad b_{ij} = \mathfrak{A}(\varphi_i - \varphi_i^0) (\varphi_j - \varphi_j^0) \quad (\varphi_i = \varphi_i(x), \varphi_i^0 = \varphi_i(x_0))$$

and the derivatives are taken at the point $(\varphi_1^0, \cdots, \varphi_n^0)$. The numbers $b_{ij}$ form a nonnegative matrix.

Properties 1 and 2 are obvious. We shall prove only 3. Let us first assume that all the second derivatives $\frac{\partial^2 F}{\partial \varphi_i \partial \varphi_j}$ are different from zero at the point $(\varphi_1^0, \cdots, \varphi_n^0)$. Then, by Taylor's formula we have

$$
F(\varphi_1, \cdots, \varphi_n) - F(\varphi_1^0, \cdots, \varphi_n^0) = \sum_i \frac{\partial F}{\partial \varphi_i} \Delta_i + \sum_{i,j} \frac{\partial^2 F}{\partial \varphi_i \partial \varphi_j} (1 + \varepsilon_{ij}) \Delta_i \Delta_j,
$$

where $\Delta_i = \varphi_i - \varphi_i^0$, the derivatives are taken at $(\varphi_1^0, \cdots, \varphi_n^0)$ and $\varepsilon_{ij} = \varepsilon_{ij}(\varphi_1, \cdots, \varphi_n) \to 0$ as $\varphi_i \to \varphi_i^0$ (and therefore as $x \to x_0$). Inserting this expression into (19), we obtain

$$
\mathfrak{A}F(\varphi_1, \cdots, \varphi_n) = \lim_{\delta \to 0} \left\{ \sum_i \frac{\partial F}{\partial \varphi_i} \int_{U'} \Delta_i \Pi_U(x_0, dx) \right\} + \sum_{i,j} \frac{\partial^2 F}{\partial \varphi_i \partial \varphi_j} \int_{U'} \Delta_i \Delta_j \Pi_U(x_0, dx) + \sum_{i,j} \frac{\int_{U'} \varepsilon_{ij} \Delta_i \Delta_j \Pi_U(x_0, dx)}{c_U(x)}.
$$

By assumption, the first term converges to $\sum_i (\frac{\partial F}{\partial \varphi_i}) a_i$, and the second to $\sum_{i,j} (\frac{\partial^2 F}{\partial \varphi_i \partial \varphi_j}) b_{ij}$. The third term converges to zero. To see this, we write

$$
\left| \frac{1}{c_U(x)} \int_{U'} \varepsilon_{ij} \Delta_i \Delta_j \Pi_U(x_0, dx) \right| \leq \max_{x \in U} |\varepsilon_{ij}| \int_{U'} \Delta_i^2 \Pi_U(x_0, dx)
$$

and note that of the three terms on the right side the first approaches zero, while the second two remain bounded. Thus we have proved property 3 in the case when $\frac{\partial^2 F}{\partial \varphi_i \partial \varphi_j} \neq 0$ ($i, j = 1, 2, \cdots, n$). If, however, some of the derivatives are equal to zero, we take a new function

$$
F_{\varepsilon} = F + \varepsilon \left[ \sum_{i=1}^{n} (\varphi_i - \varphi_i^0)^2 \right].
$$

Obviously

$$
\frac{\partial^2 F_{\varepsilon}}{\partial \varphi_i \partial \varphi_j} = \frac{\partial^2 F}{\partial \varphi_i \partial \varphi_j} + 2\varepsilon \quad (i, j = 1, 2, \cdots, n)
$$

and for sufficiently small $\varepsilon$, all these derivatives are different from zero at the point $(\varphi_1^0, \cdots, \varphi_n^0)$. Therefore property 3 is applicable to the function $F_{\varepsilon}$. Noting that

$$\mathfrak{A}F_{\varepsilon} = \mathfrak{A}F + \varepsilon \sum_{i,j} a_{ij}$$
and going to the limit as $\varepsilon \downarrow 0$, we conclude that property 3 is valid also for the function $F$.

In order to prove that the matrix $b_{ij}$ is nonnegative, we need only remark that

$$\sum_{i,j} b_{ij} \lambda_i \lambda_j = \mathbb{A} \mathbb{h}(x_0),$$

where $\mathbb{h}(x) = [\sum_{i=1}^{n} \lambda_i (q_i(x) - q_i(x_0))]^2$ takes on its maximum value at $x_0$.

The properties we have proved for $\mathbb{A}$ give us reason to consider this operator as a natural generalization of an elliptic differential operator of order 2. In particular, property 3 shows that if in the neighborhood of some point $x_0$ a coordinate system can be introduced so that the coordinates themselves and their products in pairs belong to the domain of $\mathbb{A}$, then for any function of these coordinates with two continuous derivatives the operator $\mathbb{A}$ is equal to a certain classical elliptic differential operator of order 2 (which in some cases may degenerate to an operator of order 1 or even 0).

11. Let us consider an example. Let $x(t)$ be a continuous process in an $n$-dimensional Euclidean space, invariant with respect to all motions. Let $S_r(x)$ be the sphere and $U_r(x)$ the open sphere of radius $r$ with center at $x$. We write $\tau_r = \tau_{U_r(x)}$. Obviously $c_{\tau_r(x)} = c(r)$ is independent of $x$, and the point $x(\tau_r)$ is uniformly distributed on $S_r(x)$. (19) for $\mathbb{A}$ then becomes

$$\mathbb{A} f(x) = \lim_{r \to 0} \frac{r^2}{c(r)} \int_{S_r(x)} \frac{[f(y) - f(x)]d\sigma}{r^2},$$

where the integration is taken over the uniform distribution on $S_r(x)$ (the measure of the whole sphere is taken as unity). We set

$$\mathbb{A}_1 f(x) = \lim_{r \to 0} \int_{S_r(x)} \frac{[f(y) - f(x)]d\sigma}{r^2}$$

and let $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$. We then have

$$\mathbb{A}_1 (y_i - x_i) = 0, \quad \mathbb{A}_1 (y_j - x_j)(y_i - x_i) = \frac{1}{n} \delta_{ij}.$$

For any function $f$ with two continuous derivatives, therefore,

$$\mathbb{A}_1 f = \frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}.$$

Turning now to (21), we see that if $\lim r^2/c(r) = k$, then the operator $\mathbb{A}$ acting on any function with two continuous derivatives differs from the Laplacian only by the factor $k/n$. If $r^2/c(r)$ has no limit (or is infinite) the intersection of the domain of $\mathbb{A}$ with the set of all functions with two continuous derivatives is the set of harmonic functions. A more detailed analysis shows that the second case cannot occur. The first case is realized for a many-dimensional Wiener process.

12. Let us now consider the one-dimensional case in more detail. Let $\mathcal{E}$ be some interval (open or closed, bounded or not) on the real line. We shall say that the point $x_0$ is a right transition point (or left) if there is some $t$ for which
\( \mathbb{P}_{x_0}(x(t) > x_0) > 0 \) (or \( \mathbb{P}_{x_0}(x(t) < x_0) < 0 \)). If the point \( x_0 \) is neither a right or a left transition point, then it is an absorption point and \( \mathbb{A}(x_0) = 0 \). In all other cases, Lemma 4 of Section 3 shows that there exists a neighborhood \( U_0 \) of the point \( x_0 \) such that \( m(x) = M_x \tau_{U_0} < \infty \) for all \( x \in U_0 \). (20) becomes

\[
(22) \quad \mathbb{A}(x) = \lim_{x \to U_0} \frac{\rho_2[f(x_2) - f(x)] + \rho_1[f(x_1) - f(x)]}{\rho_2[m(x_2) - m(x)] + \rho_1[m(x_1) - m(x)]}
\]

where

\[ \rho_i = \mathbb{P}_x(x(\tau_i) = x_i) = \Pi_i(x, x_i) \quad (U = (x_1, x_2), \ i = 1, 2). \]

If the point \( x_0 \) is a right, but not a left transition point, then \( \rho_2 = 1, \rho_1 = 0 \) and

\[
(23) \quad \mathbb{A}(x) = \lim_{x \to x_0} \frac{f(x_2) - f(x_0)}{m(x_2) - m(x_0)} = -D^+_m(x_0).
\]

If the point \( x_0 \) is a left but not a right transition point, then \( \rho_1 = 1, \rho_2 = 0 \) and

\[
(24) \quad \mathbb{A}(x_0) = \lim_{x_1 \to x_2 \to 0} \frac{f(x_1) - f(x_0)}{m(x_1) - m(x_0)} = -D^-_m(x_0).
\]

Points which are both right and left transition points shall be called \textit{regular}. Before studying the form of \( \mathbb{A} \) at regular points, we shall prove several lemmas.

13. We write \( \tau_a = \inf_x(x(t) = a) \) (This function is measurable, since if \( x(0, \omega) < a \), then \( \tau_a = \tau_{(-\infty, a)} \) and if \( x(0, \omega) > a \), then \( \tau_a = \tau_{(a, \infty)} \)). Further, we write

\[ \rho(x, a) = \mathbb{P}_x(\tau_a < \infty); \quad \rho(x, a, b) = \mathbb{P}_x(\tau_a < \tau_b). \]

We may, if we wish, think of \( \rho(x, a) \) as the probability that a particle leaving the point \( x \) will at some time reach the point \( a \), and of \( \rho(x, a, b) \) as the probability that it will reach \( a \) before \( b \).

**Lemma 5.** If \( a < x < y < b \) or \( a > x > y > b \), then

\[
(25) \quad \rho(y, a) = \rho(y, x) \rho(x, a),
\]

\[
(26) \quad \rho(y, a, b) = \rho(y, x, b) \rho(x, a, b).
\]

**Proof.** We have

\[
\{x(0) = y; \tau_a < \infty\} = \{x(0) = y; \tau_x < \infty\} \cap \{\theta_{\tau_x}\tau_a < \infty\}
\]

and in view of (3)

\[
\mathbb{P}_y(x(0) = y, \tau_a < \infty) = \int_{\tau_x < \infty} \mathbb{P}_y(\tau_a < \infty) \mathbb{P}(d\omega).
\]

In view of the fact that \( \mathbb{P}_y(x(0) = y) = 1 \), this formula may be rewritten

\[
\mathbb{P}_y(\tau_a < \infty) = \int_{\tau_x < \infty} \mathbb{P}_y(\tau_a < \infty) \mathbb{P}(d\omega).
\]

The left side is \( \rho(y, a) \). Since \( x(\tau_x) = x \), the right side is \( \mathbb{P}_x(\tau_a < \infty) \mathbb{P}_y(\tau_x < \infty) = \rho(x, a) \rho(y, x) \), which proves (25); (26) is proved similarly.

**Lemma 6.** If \( \rho(a, b) > 0 \), then \( M_x \tau_{(a, b)} < \infty \) for all \( x \in (a, b) \).

**Proof.** From the relation

\[
\theta_{\tau_x}\{\tau_b > t\} \subset \{\tau_b(\omega) > t\}
\]

or

\[ * \quad \text{If } x(t, \omega) \neq a \text{ for all } t, \text{ then } \tau_x(\omega) = \infty. \]
and (3), we have, bearing in mind the fact that \( x(\tau_a) = x \),
\[
\mathbb{P}_x(\tau_b \geq t) \geq \mathbb{P}_a(\theta(\tau_a) > t) = M_x \mathbb{P}_x(\tau_b > t) = \mathbb{P}_x(\tau_b > t).
\]
Since \( \tau_{a,b} = \min(\tau_a, \tau_b) \), we have
\[
\mathbb{P}_x(\tau_{a,b} > t) \leq \mathbb{P}_x(\tau_b > t) \leq \mathbb{P}_a(\tau_b > t).
\]
As \( t \to \infty \), the right side approaches \( \mathbb{P}_a(\tau_b = \infty) = 1 - \rho(a,b) \), and since \( \rho(a,b) > 0 \), there is some \( t \) for which \( \mathbb{P}_a(\tau_b > t) = 1 - \alpha < 1 \). Then \( \mathbb{P}_x(\tau_{a,b} > t) < 1 - \alpha \) for all \( x \in (a,b) \). In view of Lemma 2, it follows from this that \( M_x \tau_{a,b} < \infty \) for all \( x \in (a,b) \).

**Lemma 7.** Assume that \( a < x < b \) or \( a > x > b \). If \( \rho(x, a) = 0 \), then \( \rho(x, b) = 0 \).

**Proof.** To be specific, let us assume that \( a < x < b \). We set \( z \in I_a \), if \( z < a + (x-a)/3 \) and \( z \in I_b \), if \( z > b - (b-x)/3 \). We shall say that the path function \( x(t) \) passes from the set \( I_a \) to the set \( I_b \) exactly \( n \) times before the time \( T \) if there exist \( 0 \leq s_1 < t_1 < s_2 < t_2 < \cdots < s_n < t_n \leq T \) such that \( x(s_i) \in I_a \), \( x(t_i) \in I_b \) \( (i = 1, 2, \cdots, n) \), but if there exists no similar set of \( 2(n+1) \) points. We write \( \omega \in A \), if \( x(0, \omega) = x \) and \( \tau_n(\omega) < \infty \). We write \( \omega \in A_n \), if \( \omega \in A \) and if before the time \( \tau_n(\omega) \) the path function \( x(t, \omega) \) passes exactly \( n \) times from \( I_a \) to \( I_b \). We denote by \( \tau_n(\omega) \) the lower bound of values of \( T \) such that \( x(T, \omega) = x \) and the path function \( x(t, \omega) \) passes exactly \( n \) times from \( I_a \) to \( I_b \) before the time \( T \) (and if no such values of \( T \) exist, we write \( \tau_n(\omega) = \infty \)). We note that \( A = \bigcup_{n=0}^{\infty} A_n \) and \( A_n \subset \theta(\tau_n, A_0) \).

Since the random quantity \( \tau_n(\omega) \) is independent of the future, we have according to (3)
\[
\mathbb{P}_x(\theta(\tau_n, A_0) = M_x \mathbb{P}_x(\tau_n, A_0).
\]
But \( x(\tau_n(\omega)) = x \) and by assumption \( \mathbb{P}_x(A_0) = 0 \). Therefore \( \mathbb{P}_x(\theta(\tau_n, A_0) = 0, \mathbb{P}_x(A_n) = 0, \) and \( \mathbb{P}_x(A) = 0 \).

**Lemma 8.** Let \( \rho(a, b) > 0 \). If \( a < b \), then \( \lim_{a \to b} \rho(x, b, a) = 1 \). If \( a > b \), then \( \lim_{a \to b} \rho(x, b, a) = 1 \).

**Proof.** To be specific, let us assume that \( a < b \) and that \( a < x < x_1 < x_2 < \cdots < x_n < \cdots < b \), where \( x_n \to b \). If \( x(0, \omega) = x \), then
\[
(27) \quad \tau_{x_1} < \tau_{x_2} < \cdots < \tau_{x_n} < \tau_b.
\]
Let \( A_n = \{ x_{\tau_a} < \tau_a \} \); \( A = \{ \tau_b < \tau_a \} \). Obviously \( A \subseteq \bigcap_{n=1}^{\infty} A_n \). Let us assume that \( \omega \in (\bigcap_{n=1}^{\infty} A_n) - A \). Then \( x(t, \omega) \neq 0 \) with \( t \leq \tau_{x_n}(\omega) \) for all \( n \), and therefore with \( t \leq \tau(\omega) = \lim_{n \to \infty} \tau_{x_n}(\omega) \). If it were true that \( \tau(\omega) < \infty \), the relation \( x(\tau_{x_n}) = x_n \) and the fact that \( x(t) \) is a continuous function would give \( x(\tau) = b \) or \( \omega \in A \). Therefore \( \tau(\omega) = \infty \). But then \( \tau_n(\omega) = \infty \) and from (27) we have \( \tau_n(\omega) = \infty \). Thus \( \tau_{a,b} = \min(\tau_a, \tau_b) = \infty \). According to Lemma 6 the probability of this vanishes. Thus
\[
\mathbb{P}_x(\bigcap_{n=1}^{\infty} A_n - A) = 0.
\]
We therefore find that \( \mathbb{P}_x(A) = \lim_{n \to \infty} \mathbb{P}_x(A_n) \), or
\[
(28) \quad \rho(x, x_n, a) \to \rho(x, b, a).
\]
But according to (26)

\[ p(x, b, a) = p(x_n, b, a)p(x_n, b, a). \tag{29} \]

It follows from Lemma 7 that \( p(x, b, a) \neq 0 \) (since otherwise we would have \( p(x, b) = 0 \) and \( p(a, b) = p(a, x)p(x, b) = 0 \)). Equations (28) and (29) therefore lead to the fact that \( p(x_n, b, a) \to 1 \).

**Lemma 9.** Let \( a < b \). If \( p(a, b) > 0 \), all the points of \([a, b]\) are right transition points. Conversely, if all points of \([a, b]\) are right transition points, then \( p(a, b) > 0 \).

**Proof.** We note first that a point \( x_0 \) is a right transition point if and only if \( p(x_0, x) = 0 \) for all \( x > x_0 \). Let us assume that \( a \leq x_0 < b \) and that \( x_0 \) is not a right transition point. Then \( p(a, b) = p(a, x_0)p(x_0, b) = 0 \). This proves the first part of the lemma.

Let us now show that if \( x_0 \) is a right transition point there exist \( x_1 < x_0 < x_2 \) such that \( p(x_1, x_2) > 0 \). Indeed, by the definition of a right transition point \( P_{x_0}\{x(t) > x_0\} > 0 \) for some value of \( t \). Thus there exists an \( x_2 > x_0 \) such that \( P_{x_0}\{x(t) > x_2\} > 0 \). According to Lemma 3 there exists an \( \epsilon > 0 \) such that \( P_x\{x(t) > x_2\} > 0 \) for all \( x \in (x_0 - \epsilon, x_0 + \epsilon) \). We may now choose any \( x_1 \in (x_0 - \epsilon, x_0) \). Then \( p(x_1, x_2) > 0 \).

Let us assume that all the points of \([a, b]\) are right transition points. Then for each point \( x_0 \in [a, b] \) we can construct an interval \((x_1, x_2)\) containing \( x_0 \) and such that \( p(x_1, x_2) > 0 \). Of these intervals we can choose a finite set \((x_1^{(0)}, x_2^{(0)})\), which cover the segment \([a, b]\). Since \( p(x_1^{(0)}, x_2^{(0)}) > 0 \), it follows from (25) that \( p(a, b) > 0 \).

**14.** We shall say that a process \( x(t, \omega) \) is regular on the segment \([a, b]\) if \( p(a, b) > 0 \) and \( p(b, a) > 0 \). According to Lemma 9 a necessary condition for this is that all the points of \((a, b)\) be regular, and a sufficient condition is that all the points of \([a, b]\) be regular.

**Lemma 10.** If a continuous Feller process is regular on the segment \([a, b]\), the function

\[ p(x) = p(x, b, a) \]

is continuous, monotonically increasing, and satisfies the conditions

\[ \lim_{x \to a^+} p(x) = 0, \quad \lim_{x \to b^-} p(x) = 1. \tag{30} \]

For all \( a < x_1 < x < x_2 < b \) we have

\[ p(x, x_1, x_2) = \frac{p(x_2) - p(x)}{p(x_2) - p(x_1)}. \tag{31} \]

**Proof.** Let \( a < x < y < b \). From (26), we have

\[ p(x, b, a) = p(x, y, a)p(y, b, a) \]

or, which is the same thing,

\[ p(x) = p(x, y, a)p(y). \tag{32} \]

Therefore \( p(x) \leq p(y) \) and if \( p(x) = p(y) \) either \( p(x, b, a) = 0 \) or \( p(x, y, a) = 1 \). From this we have \( p(x, a, y) = 1 - p(x, a, y) = 0 \). In view of the fact that
\( \phi(x, a, y) = 0 \), it follows from Lemma 7 that \( \phi(x, a) = 0 \) and from (25) that 
\[ \phi(b, a) = \phi(b, x) \phi(x, a) = 0. \]
This is in contradiction with the assumption. From \( \phi(x, b, a) = 0 \) it follows by Lemma 7 that \( \phi(x, b) = 0 \) and by (25) that 
\[ \phi(a, b) = \phi(a, x) \phi(x, b) = 0, \]
and this is also in contradiction with the assumption. We have thus shown that \( \phi(x) \) increases monotonically.

Equations (30) follow from Lemma 8 (the first of these equations is proved by applying Lemma 8 to the function \( \phi(x, a, b) = 1 - \phi(x, b, a) \)). By the same Lemma 8, we have 
\[ \lim_{x \to y^-} \phi(x, y, a) = 1 \quad (a < x < y), \]
and from (32) it is seen that 
\[ \lim_{x \to y^-} \phi(x) = \phi(y). \]
This proves that \( \phi(x) \) is continuous on the right. It can be proved similarly that \( \phi(x, a, b) = 1 - \phi(x) \) is continuous on the left. Therefore \( \phi(x) \) is continuous.

We note, finally, that in view of (26)
\[ \phi(x, x_1, x_2) = \frac{\phi(x, a, x_2)}{\phi(x_1, a, x_2)} = \frac{1 - \phi(x, x_2, a)}{1 - \phi(x_1, x_2, a)}. \]
On the other hand, from (32)
\[ \phi(x, x_2, a) = \frac{\phi(x)}{\phi(x_2)}, \quad \phi(x_1, x_2, a) = \frac{\phi(x_1)}{\phi(x_2)}. \]
From (33) and (34) we obtain (31).

Inserting \( \phi_1 = \phi(x, x_1, x_2) \) and \( \phi_2 = 1 - \phi \) from (31) into (22), we obtain the following expression for \( \mathcal{M} \) at all points \( x \) within a regular interval \([a, b] \):
\[ \mathcal{M}(x) = -\lim_{x \to y^-} \frac{f(x_2) - f(x)}{m(x_2) - m(x)} - \frac{f(x_1) - f(x)}{m(x_1) - m(x)} \frac{\phi(x_2) - \phi(x)}{\phi(x_1) - \phi(x)}. \]

15. We shall now show that the limit on the right side of (35) defines the generalized second derivative of a function \( f \) with respect to two monotonic functions. To do this we shall need the following lemma.

**Lemma 11.** Let \( x(t) \) be a continuous Feller process regular on the segment \([a, b] \) and let \( \phi(x) = \phi(x, b, a) \) and \( m(x) = M_x \tau_{(a,b)}. \) Then \( m(x) \) is continuous on \([a, b], \) the derivative 
\[ D^+ m(x) = \lim_{y \to x^+} \frac{m(y) - m(x)}{\phi(y) - \phi(x)} \]
e exists for all \( x \in (a, b) \) and is continuous on the right and monotonically decreasing, and the derivative 
\[ D^m m(x) = \lim_{y \to x^+} \frac{m(y) - m(x)}{\phi(y) - \phi(x)} \]
e exists for all \( x \in (a, b) \) and is continuous on the left and monotonically decreasing. At any point \( x \in (a, b) \) where at least one of the functions \( D^+ m(x) \) or \( D^m m(x) \) is continuous, we have \( D^+ m(x) = D^m m(x). \)
Proof. In view of Lemma 10, there exists an inverse function of $\dot{p}(x)$, which we shall call $q(z)$. Let us set $\tilde{m}(z) = -m[q(z)]$. Let $0 \leq z_1 < z < z_2 \leq 1$ and $x_1 = q(z_1)$, $x = q(z)$, $x_2 = q(z_2)$. Then $a \leq x_1 < x < x_2 \leq b$. According to Lemma 1,

$$m(x) - \dot{p}(x, x_1, x_2)m(x_1) - \dot{p}(x, x_2, x_1)m(x_2) = M_x \tau_{x_1, x_2} > 0.$$ 

According to (30) this implies that

$$m(x) > \frac{\dot{p}(x_2) - \dot{p}(x)}{\dot{p}(x_2) - \dot{p}(x_1)} m(x_1) + \frac{\dot{p}(x) - \dot{p}(x_1)}{\dot{p}(x_2) - \dot{p}(x_1)} m(x_2)$$

or

$$\tilde{m}(z) < \frac{z_2 - z}{z_2 - z_1} \tilde{m}(z_1) + \frac{z - z_1}{z_2 - z_1} \tilde{m}(z_2).$$

The function $\tilde{m}(z)$ is therefore convex on $[a, b]$. Further, it is nonpositive and is therefore bounded from above. As is well known (see, for instance, Hardy, Littlewood, and Polya [13], paragraph 3.18, pages 113–117), such a function is continuous, and its left and right derivatives exist everywhere, are increasing functions, and are equal at all points where at least one of them is continuous. If $D^- z \tilde{m}(z)$ were not continuous on the left at $z_0$, there would exist $z_1 < z_0$ and $e > 0$ such that for all $z \in (z_1, z_0)$

$$D^- z \tilde{m}(z) < D^- z \tilde{m}(z_0) - e.$$ 

This would imply (see, for instance, de la Vallée Poussin [14], paragraph 112) that for all $z \in (z_1, z_0)$, we would have

$$\frac{\tilde{m}(z) - \tilde{m}(z_0)}{z - z_0} < D^- \tilde{m}(z_0) - e,$$

which is clearly impossible. Thus $D^- z \tilde{m}(z)$ is continuous on the left. It can be shown in the same way that $D^+ \tilde{m}(z)$ is continuous on the right. In order now to prove the lemma it is sufficient to go from $\tilde{m}(z)$ to the function $m(x) = -\tilde{m}[\dot{p}(x)]$ and to note that $D^+ m(x) = -D^- z \tilde{m}(z)$ and $D^- m(x) = -D^+ \tilde{m}(z)$.

**Theorem 6.** Let $x(t)$ be a continuous Feller process regular on the segment $[a, b]$. Let $\dot{p}(x) = \dot{p}(x, b, a)$, $m(x) = M_x \tau_{(a, b)}$, $n_+(x) = -D^+ m(x)$, $n_-(x) = -D^- m(x)$.

If $f \in D_p$, then for any $x \in (a, b)$ the derivative $D^+_p f(x)$ exists and is continuous on the right, and the derivative $D^-_p f(x)$ exists and is continuous on the left. Further $D^+_p f(x) = D^-_p f(x)$ at all points $x$ at which $n_+(x) = n_-(x)$. In the interval $(a, b)$ we have

$$D^+_p f(x) = D^-_p f(x) = D^-_n D^-_p f(x).$$

**Remark.** Let $N$ be the set of all points of continuity $n_+(x)$. According to Lemma 11 if $x \in N$, then $n_+(x) = n_-(x)$ and we may denote the common value of these functions by $n(x)$. Then (30) can be written

$$n(x) = \begin{cases} D_n D_p f(x) & (x \in N), \\ D_p f(x+0) - D_p f(x-0) & (x \notin N). \end{cases}$$
It is just in this form that Feller calculated the differential operator [11, 12] (without explaining the probability theory meaning of the functions $\rho$ and $n$).

PROOF. According to (35) and Lemma 11 we have

$$\lim_{x \to x^+} \left[ \frac{f(x_2) - f(x)}{\rho(x_2) - \rho(x)} - \frac{f(x_1) - f(x)}{\rho(x_1) - \rho(x)} \right] = [n_+(x) - n_-(x)] \mathcal{A}f(x).$$

It follows from this that $D^+_\rho f$ and $D^-_\rho f$ exist and coincide if $n_+(x) = n_-(x)$. Let us write

$$F(y) = f(y) - f(x) - D^-_\rho f(x) [\rho(y) - \rho(x)],$$

$$G(y) = -m(y) + m(x) - n_-(x) [\rho(y) - \rho(x)].$$

We have $G(x) = 0$ and if $y > x$ then $D^-_\rho G(y) = n_-(y) - n_-(x) > 0$, $D^+_\rho G_+(y) = n_+(y) - n_-(x) > 0$ which means that $G(y) > 0$. Going to the limit $x_1 \to x - 0$ in (35), we obtain

$$\mathcal{A}f(x) = \lim_{y \to x^+} \frac{F(y)}{G(y)}.$$  

The rest of the proof consists in showing that one can use l'Hospital's rule to calculate the limit of (38) (compare Feller [12], the end of Paragraph 7).

Let us assume that $\lim_{y \to x^+} D^+_\rho F(y)/D^+_\rho G(y)$ does not exist, and let

$$\lim_{y \to x^+} \frac{D^+_\rho F(y)}{D^+_\rho G(y)} < c < \lim_{y \to x^+} \frac{D^+_\rho F(y)}{D^+_\rho G(y)}.$$  

We set $\varphi(y) = F(y) - cG(y)$. Since $\mathcal{A}$ annihilates constants and the function $\rho$, whereas it transforms $m$ into $-1$, we have

$$\mathcal{A}\varphi = \mathcal{A}f + c.$$  

Further, in view of (39) we know that for all $x_2 > x$, the function $D^+_\rho \varphi = D^+_\rho f - cD^+_\rho G$ changes sign in the interval $(x, x_2)$. Therefore there exist sequences $y_n \to x^+0$ and $y'_n \to x^+0$ such that at the points $y_n$ the function $\varphi(y)$ takes on relative maximum values, and at the points $y'_n$ relative minimum values. In view of property 2 of paragraph 10 $\mathcal{A}\varphi(y_n) \leq 0$, $\mathcal{A}\varphi(y'_n) \geq 0$ and therefore $\mathcal{A}\varphi(y_n) \leq -c$, $\mathcal{A}\varphi(y'_n) \geq -c$. From this we obtain $\mathcal{A}f(x) = -c$. This should be true for any constant $c$ satisfying the inequality (39), but this is obviously impossible. We have therefore proved that

$$\lim_{y \to x^+} \frac{D^+_\rho F(y)}{D^+_\rho G(y)} = \lim_{y \to x^+} \frac{D^+_\rho f(y) - D^+_\rho f(x)}{n_+(y) - n_-(x)}$$

exists. Similarly, we can prove the existence of

$$\lim_{y \to x^+} \frac{D^-_\rho F(y)}{D^-_\rho G(y)} = \lim_{y \to x^+} \frac{D^-_\rho f(y) - D^-_\rho f(x)}{n_-(y) - n_-(x)}.$$  

Since $D^-_\rho f(y) = D^+_\rho f(y)$ on the everywhere dense set $N$, (40) and (41) coincide.

Let us now write

$$\psi(y) = F(y)G(x_2) - G(y)F(x_2) \quad (x \leq y \leq x_2).$$

Since $\psi(x) = \psi(x_2) = 0$, either the maximum or minimum value of $\psi(y)$ on $[x, x_2]$ is attained at some internal point $y'$ of the interval. If it is the maximum,
then \( D^+_\varphi(y') \geq 0 \geq D^-\varphi(y') \), from which we obtain

\[
\frac{D^+_\varphi F(y')}{D^+_\varphi G(y')} \leq \frac{F(x_2)}{G(x_2)} \leq \frac{D^-\varphi F(y')}{D^-\varphi G(y')}.
\]

If it is the minimum, we obtain the opposite inequality. Now let \( x_2 \to x+0 \). Then \( y' \to x+0 \). From what has been proved, the left and right sides of (42) approach a single limit. According to (38) the expression in the middle approaches \( \Psi f(x) \). Therefore

\[
\Psi f(x) = \lim_{y' \to x+0} \frac{D^+_\varphi f(y') - D^-\varphi f(x)}{n_+(y') - n_-(x)} = \lim_{y' \to x+0} \frac{D^-\varphi f(y') - D^-\varphi f(x)}{n_-(y) - n_-(x)}.
\]

This proves the first of (35). Furthermore, the continuity of \( n_+ \) on the right implies, with (43) that

\[
D^+_\varphi f(x+0) - D^-\varphi f(x) = [n_+(x) - n_-(x)] \Psi f(x).
\]

Comparing this expression with (37), we obtain

\[
D^+_\varphi f(x+0) = D^+_\varphi f(x),
\]

which means that \( D^+_\varphi f \) is continuous on the right.

The proof of (36) and the continuity of \( D^-\varphi f(x) \) on the left is similar.

5. Some Discontinuous Processes

16. Let us assume that the phase space \( \mathcal{E} \) is finite or denumerable and that its topology is discrete, meaning that all of its points are isolated. We may assume without loss of generality that \( \mathcal{E} \) is a set of integers or a subset of this set. The space \( C \) (like \( B \)) consists in our case of all bounded functions \( f(i) \). Therefore every Markov process with a discrete set of states is a Feller process.

Let us set \( \phi_{ij}(t) = P_i\{x(t) = j\} \). Equation (5) can be written

\[
T_t f(i) = \sum_j \phi_{ij}(t) f(j).
\]

As is well known (see, for instance Kolmogorov [10]), the limits

\[
a_i = -a_{ii} = \lim_{t \to 0} \frac{1 - \phi_{ii}(t)}{t},
\]

\[
a_{ij} = \lim_{t \to 0} \frac{\phi_{ij}(t)}{t} \quad (i \neq j),
\]

always exist, and are such that

\[
0 \leq a_{ij} < +\infty \quad (i \neq j),
\]

\[
0 \leq a_i \leq +\infty,
\]

\[
\sum_{j \neq 0} a_{ij} \leq a_i.
\]

The transition probability \( \phi_{ij}(t) \) uniquely defines a Markov process continuous on the right. A necessary and sufficient condition that there exist a Markov process continuous on the right with transition probabilities \( \phi_{ij}(t) \) is that

\[
a_i < +\infty, \sum_{j \neq i} a_{ij} = a_i.
\]

The probability-theory meaning of \( a_i \) and \( a_{ij} \) can be stated as follows. If at the initial time a particle is at a point \( i \), it remains at this point for a random time interval \( \gamma \) with \( P_i(\gamma > t) = e^{-a_i t} \). After this it under-
goes a transition to some other point, the transition probability to \( j \) being equal to \( a_{ij}/a_i \).

Any discrete Markov process continuous on the right, since it is a Feller process, is a strong Markov process. Let us calculate \( \mathcal{A} \) for such a process. Every sufficiently small neighborhood of a point \( i \) consists of this point alone. For such a neighborhood, therefore, \( \tau_u = \gamma, \ P_i(\gamma = j) = P_i(\gamma = i) = a_{ij}/a_i \) and

\[
\mathcal{A} f(i) = \frac{M_i f(\gamma) - f(i)}{M_i \gamma_i} = \sum_{i \neq i} \frac{a_{ij} f(j) - f(i)}{1/a_i} = \sum_j a_{ij} f(j).
\]

The domain \( D_{\mathcal{A}} \) consists of all possible bounded functions \( f(j) \) for which \( \mathcal{A} f(i) \) is also bounded.

If \( \mathcal{A} \) is finite, we find from Theorem 5 that \( A = \tilde{A} = \mathcal{A} \). If \( \mathcal{A} \) is denumerable, we can conclude from Theorem 4 only that \( A \subseteq \tilde{A} \subseteq \mathcal{A} \).

17. Let us consider in greater detail the probability description of the course of the process.

A particle which enters from the point \( i = i_1 \) remains at this point for some interval of time \( \gamma_1 \), then jumps to some randomly chosen point \( i_2 \), remains there for a time \( \gamma_2 \), then jumps to some other point \( i_3 \) etc. The distributions of the random variables \( \gamma_1, \gamma_2, \cdots \) and the random points \( i_2, i_3, \cdots \), as has been mentioned above, are completely determined by the \( a_{ij} \) matrix. This \( a_{ij} \) matrix therefore uniquely determines the motion of the particle up to the time \( \eta = \sum_{i=1}^{\infty} \gamma_i \), and if \( \eta = \infty \) with probability unity, the process is uniquely determined by \( \mathcal{A} \). This operator \( \mathcal{A} \) however, gives no information on the behavior of the path function after time \( \eta \). Therefore if \( P(\eta < \infty) > 0 \), there may exist many different processes belonging to the same operator \( \mathcal{A} \). Such processes actually exist. It was shown by Doob [3], for instance, that for a given matrix \( a_{ij} \), it is possible to construct a process with any arbitrarily prescribed probability distribution \( x(\eta) \). This shows clearly, parenthetically, why for some processes with denumerable sets of states we obtain the strict inclusion \( \tilde{A} \subseteq \mathcal{A} \).

18. What has been said about processes with discrete sets of states is not difficult to extend to so-called purely discontinuous processes in an arbitrary metric space \( \mathcal{E} \). A purely discontinuous process is characterized by the fact that a particle which at some time is at some point \( x \) remains there with probability unity for a finite time \( \gamma \). The probability distribution for \( \gamma \) can be only exponential, that is \( P_x(\gamma > t) = e^{-a(x)t} \). Let us assume that this is a Feller process discontinuous on the right, and let us calculate its operator \( \mathcal{A} \). If \( d(U) \rightarrow 0 \), then \( \tau_u \rightarrow \gamma + 0 \). Therefore \( M_x(\tau_u) \rightarrow M_x(\gamma) = 1/a(x), x(\tau_u) \rightarrow x(\gamma) \), and for all \( f \in C \) we have

\[
f[x(\tau_u)] \rightarrow f[x(\gamma)], \quad M_x f[x(\tau_u)] \rightarrow M_x f[x(\gamma)] = \int_\mathcal{E} f(y) \Pi(x, dy),
\]

where \( \Pi(\Gamma) = P_x(x(\gamma) \in \Gamma) \).

Inserting the limits obtained into (17), we arrive at

\[
\mathcal{A} f(x) = a(x) \left[ \int_\mathcal{E} f(y) \Pi(x, dy) f(x) \right].
\]
19. As an example of more general discontinuous processes let us consider a line with a random process discontinuous on the right and with positive independent increments, which is given by
\[ M_e e^{-\lambda[x(t)-\varepsilon]} = e^{-ct\lambda\varepsilon} \quad (0 < \lambda < 1, \ c > 0). \]
This is a Feller process. We write \( U_e = (x-\varepsilon, x+\varepsilon), \ \tau_e = \tau_{U_e} \). According to a previous work [4] (Section 6, Theorem 6) the density distribution of the random quantity \( \tau_e - x - \varepsilon \) is given by
\[ p(y) = \begin{cases} \sin \pi \frac{\varepsilon}{y} & (y > 0), \\ \frac{\pi}{(y+\varepsilon)y^2} & (y \leq 0). \end{cases} \]
In the same reference it was shown that \( M_{\tau_e} = e^{\varepsilon c}/\Gamma(1-\alpha) \). Inserting these values into (17), we find that \( \mathfrak{A} \) is defined by
\[ \mathfrak{A} F(x) = \lim_{\varepsilon \to +0} \frac{1}{c \Gamma(\alpha)} \int_0^\infty \frac{f(x+z)-f(x)}{z^{1+\alpha}} \, dz. \]
From this we have
\[ \mathfrak{A} f(x) = \frac{1}{c \Gamma(\alpha)} \int_0^\infty \frac{f(x+z)-f(x)}{z^{1+\alpha}} \, dz \]
for all \( f(x) \) for which the integral on the right side converges absolutely.

REFERENCES


\(^{10}\) The probability distributions of the increments are stable distributions with indices \( \alpha, \beta = -1 \) (in the notation of Gnedenko and Kolmogorov [1]).
INFINITESIMAL OPERATORS OF MARKOV PROCESSES

E. B. DYNKIN (MOSCOW)

(Summary)

In 1931 A. Kolmogorov showed [9] that a wide class of one dimensional Markov processes can be described by the differential equation

\[ \frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + b \frac{\partial u}{\partial x}. \]

Are there one-dimensional Markov processes governed by equations of the type

\[ \frac{\partial u}{\partial t} = \mathfrak{A} u, \]

where \( \mathfrak{A} \) is a differential operator of an order higher than 2? This question remained unsettled until 1954—1955 when it was completely solved by W. Feller. Feller showed that every one-dimensional Markov process with a continuous path function is described by (2), where \( \mathfrak{A} \) is the generalized second derivative.

The purely analytical method of Feller is essentially connected with the one-dimensional character of the problem. It is very difficult to extend this method to the case of two (and more) dimensions.

In this paper, a method is developed which can be applied to \( n \) dimensions as well as to one dimension. It is based on considering the probability properties of path functions. We prove by this method that every continuous Markov process in a space of arbitrary dimension is governed by (2), where \( \mathfrak{A} \) is a generalized elliptical differential operator of the second order. Our method can be applied not only to continuous Markov processes, but also to any Markov processes with path functions continuous on the right.

In Feller's papers it is assumed that the semi-group associated with the Markov process under consideration takes continuous functions into continuous ones. The class of processes specified by this condition is very important. We call them Feller processes. However, our method can be applied to a wider class of processes which we call strong Markov processes. By the conventional definition of a Markov process it is required that the development of the process after the time \( t \) does not depend on its previous history if the position at the time \( t \) is known. We say that a process is a strong Markov process if this condition is fulfilled not only for any constant \( t \), but also for a random \( t \) which “does not depend on the future”. All Feller processes are strong Markov processes, however, the converse is not true.

In the introductory § 1 a brief exposition of basic facts about Markov and strong Markov processes and associated semi-groups of operators is given. In this we use the results of [5] and [6].

In § 2 the main method of calculating infinitesimal operators corresponding to Markov processes is developed.

Feller processes are treated in § 3. The operator \( \mathfrak{A} \), which is constructed in this section, plays the same role in the general case as the differential operator in the right-hand side of (1) for diffusion processes, or the matrix of the densities of transition probabilities for processes with denumerably many states.

In order to determine a Markov process, some lateral conditions (boundary conditions in the case of diffusion processes, probability distribution at the moment of condensation of jumps in the case of a process with denumerably many states) are needed in addition to the operator \( \mathfrak{A} \). In terms of semi-groups this means that generally the infinitesimal operator \( A \) of the Markov process is a contraction of the operator \( \mathfrak{A} \) (see Theorem 4). However, Theorem 5 shows that \( A = \mathfrak{A} \) if the state space is compact.

In § 4 Feller processes with continuous path functions are treated. It is shown that in this case the operator \( \mathfrak{A} \) is a natural generalization of an elliptical differential operator of the second order. A more detailed description of one-dimensional continuous Feller processes is given, and it is proved that the operator \( \mathfrak{A} \) in this case is identical to Feller’s generalized second derivative. At the same time the probability meaning of the two functions, which appear in the definition of this derivative, comes to light.

In § 5 some discontinuous processes are considered.

Some results contained in this paper have been published in [7] without detailed proofs.