

STATIONARY MARKOV PROCESSES WITH CONTINUOUS PATHS⁽¹⁾

BY
DANIEL RAY

I. Introduction. Let $P(x, t, E)$ be a stationary transition probability function on the real line. It is well known [IV; V] that for each probability measure μ , the function P determines a Markov process $\{\Omega, \mathfrak{F}, \Pr\}$ with initial distribution μ . That is, a space Ω of paths $\omega = \{x(\tau, \omega) | \tau \geq 0\}$ can be chosen, together with a field \mathfrak{F} of subsets of Ω containing all sets of the form $\{\omega | a < x(t, \omega) < b\}$, and a measure $\Pr\{\cdot\}$ on \mathfrak{F} , such that

$$(1.1) \quad \Pr \{a < x(t, \omega) < b\} = \int P(x, t, (a, b))\mu(dx)$$

for $t > 0$, a, b real. (Here and throughout the paper we suppress the indication of the range of an integration, if that range is the real line.)

For each $t > 0$,

$$(1.2) \quad \mu \rightarrow \mu T_t: \mu T_t(E) = \int P(x, t, E)\mu(dx)$$

defines a positivity and norm preserving operator on the space of bounded additive Borel set functions. The hypothesis that the transition probabilities are stationary is reflected in the fact that the operators $\{T_t | t > 0\}$ form a semi-group: $T_{t+s} = T_t T_s$.

Among the semi-groups defined in this way, those satisfying the following two conditions have been called *diffusion processes* by Feller [VIII]:

A_0 : For every bounded continuous real function f and for every $t > 0$,

$$(1.3) \quad T_t f(x) = E\{f(x(t)) | x\} = \int f(y)P(x, t, dy)$$

defines a continuous function of the real variable x .

B_0 : The transition probabilities are of LOCAL CHARACTER: for every x and for every open interval I containing x ,

$$(1.4) \quad \lim_{t \rightarrow 0} \frac{1}{t} P(x, t, I') = 0,$$

Received by the editors February 21, 1956.

⁽¹⁾ Supported by the U. S. Army Ordnance, Office of Ordnance Research, Contract No. DA-36-034-ORD-1296.

where I' is the complement of I .

Feller has shown [VIII; IX] that under these conditions the semi-group defined by (1.3) is generated by a generalized differential operator.

Feller has suggested that for transition probabilities satisfying A_0 , the additional hypothesis B_0 is equivalent to the proposition that for every initial distribution a Markov process can be formed from the given transition probabilities with paths whose only discontinuities are jumps from $\pm \infty$.

This paper is devoted to an examination of the relation of the continuity properties of paths of Markov processes to properties similar to A_0 and B_0 . It will be proved in particular that Feller's conjecture is indeed true. Moreover, for a stationary process whose paths are continuous in the above sense, the hypothesis A_0 implies a strong Markov property: the process starts anew at the first passage of a given point, without memory, but with the same transition probabilities. This property is of such importance in the applications of Markov processes to problems in analysis [VI; VII; XIV] that we suggest the name *diffusion process* for a Markov process with paths continuous except for jumps from $\pm \infty$ and satisfying the strong Markov property. The terminology is essentially consistent with Feller's, for, as we shall prove, the strong Markov property is actually equivalent to a regularity condition on the transition probabilities only slightly weaker than A_0 .

In these terms, then, we have the following situation concerning the continuity of paths: A stationary Markov process whose paths are continuous except for jumps from $\pm \infty$, and which also has the strong Markov property, has transition probabilities which are of local character uniformly on bounded intervals: For each $\delta > 0$, $a > 0$, if we set $I = (x - \delta, x + \delta)$, then (1.4) holds uniformly for $|x| \leq a$. But if the process does not have the strong Markov property, then nothing can be said about condition B_0 , while in no sense is A_0 satisfied. Conversely, if given transition probabilities are of local character uniformly on bounded intervals, then for every initial distribution a process may be constructed whose paths are continuous except for jumps from $\pm \infty$. Another sufficient condition for this construction is that the transition probabilities satisfy A_0 and B_0 ; in this case the constructed process will have the strong Markov property and all its consequences for the transition probabilities will follow.

In §II is contained a precise discussion of the concepts mentioned above, and in §III a detailed statement and discussion of our results. §IV is devoted to examples which illustrate various of the concepts and results. Proofs are given in the last two sections.

The author wishes to express his gratitude to Professor Feller, whose constant and patient interest were invaluable, and also to Professors Gilbert Hunt and K. Ito, and to Dr. H. McKean for their helpful suggestions.

II. A probabilistic definition of diffusion. The restriction to processes with transition probabilities is of course essential to our problem. At the

same time the analytic basis afforded by transition probabilities makes possible some simplification in the concept of diffusion and in the statement of our results. It is not out of order therefore to include here the definitions and known results from which we will proceed, stating them within the framework of the problem. In particular, we shall avoid an explicit definition of a Markov process [IV; V], with which we assume the reader to be familiar. Instead, we will describe certain processes which are constructed from stationary transition probabilities, and discuss the continuity properties of the paths and the strong Markov property in terms of the transition probabilities.

DEFINITION I. *Stationary transition probabilities on the real line R are given by a function $P(x, t, E)$ of the variables x in $R, t \geq 0, E$ a Borel subset of R , with the properties:*

1°: *For each x in R and $t \geq 0, P(x, t, \cdot)$ is a Borel measure on R with $P(x, t, R) = 1$.*

2°: *For each $t \geq 0$ and Borel subset E of $R, P(\cdot, t, E)$ is a Borel measurable function on R .*

3°: *For $t \geq 0, s \geq 0, x$ in R , and E a Borel subset in R ,*

$$(2.1) \quad P(x, t + s, E) = \int P(y, t, E)P(x, s, dy).$$

Our results will be stated in terms of not a single process constructed from given transition probabilities and initial distribution, but the collection of processes corresponding to the different initial distributions. These are determined by the distributions $\{\mu_x | x \in R\}$ which give probability one to the various points of the real line R . Thus for the purposes of this paper, we shall define a diffusion process in terms of a collection of spaces $\Omega_x, x \in R$, consisting of paths with initial point x . The desired continuity properties of the paths are given by

DEFINITION II. *For each point x in the real line R, Ω_x is to denote the space of paths $\omega = x(\cdot, \omega)$ in the extended real line $R^+ : -\infty \leq x(\tau, \omega) \leq +\infty, \tau \geq 0$, which have the properties:*

1°: $x(0, \omega) = x$;

2°: *For every compact subset F of R , for every $t_0 > 0$, and for every $\delta > 0$, there is $\Delta > 0$ so that $|x(t, \omega) - x(t', \omega)| \leq \delta$ whenever $0 \leq t < t' < t + \Delta \leq t_0 + \Delta, x(t, \omega) \in F$.*

In each space Ω_x , a cylinder set is determined by a finite set of parameter values $\tau_1, \dots, \tau_n, 0 < \tau_1 < \dots < \tau_n$, and corresponding Borel sets E_1, \dots, E_n in R^+ as

$$\{\omega \in \Omega_x | x(\tau_j, \omega) \in E_j, j = 1, \dots, n\}.$$

Suppose for each x a measure $\Pr\{\cdot | x\}$ is defined on the Borel field of subsets of Ω_x generated by the cylinder sets, and satisfies

1°: $\Pr\{\Omega_x | x\} = 1$;

2°: *The function*

$$P(x, t, E) = \Pr \{x(t, \omega) \in E \mid x\}$$

for x in R , $t \geq 0$, E a Borel subset of R , defines stationary transition probabilities on R .

3°: *For every cylinder set in Ω_x ,*

$$\begin{aligned} &\Pr \{x(\tau_j, \omega) \in E_j, j = 1, \dots, n \mid x\} \\ &= \int_{E_1} \dots \int_{E_n} P(x, \tau_1, dy_1)P(y_1, \tau_2 - \tau_1, dy_2) \dots P(y_{n-1}, \tau_n - \tau_{n-1}, dy_n). \end{aligned}$$

Then each pair $\{\Omega_x, \Pr \{\cdot \mid x\}\}$ forms a Markov process [V].

In this case we will say that the collection of Markov processes is determined from stationary transition probabilities.

The integral of a function Φ defined on Ω_x will be written as $E\{\Phi(\omega) \mid x\}$; the integral over a measurable subset Λ of Ω_x will be $E\{\Phi(\omega); \omega \in \Lambda \mid x\}$. We shall often suppress the variable ω in such expressions. Finally, we suppose each measure $\Pr \{\cdot \mid x\}$ has been completed on Ω_x : that is, $\Pr \{\Lambda \mid x\} = 0$ has been defined for every set Λ contained in a measurable set of measure zero.

We turn now to a precise description of the strong Markov property, or first passage time relation. The definition of the first passage times causes no difficulty. Let x be a point in R ; let $(\Omega_x, \Pr \{\cdot \mid x\})$ be a Markov process on R , whose paths have the properties of Definition II. For every interval I in R containing x , set

$$(2.2) \quad \mathfrak{J}_I(\omega) = \mathfrak{J}_I(\omega, x) = \inf \{ \tau \geq 0 \mid x(\tau, \omega) \notin I \}$$

for each ω in Ω_x . \mathfrak{J}_I is certainly a measurable function on Ω_x , for, because of the continuity properties of each path ω in Ω_x , \mathfrak{J}_I can be defined in terms of a countable number of characteristic functions of cylinder sets:

$$\begin{aligned} \mathfrak{J}_I(\omega) &= \sup \{ t \geq 0 \mid x(\tau, \omega) \in I, 0 \leq \tau \leq t \} \\ &= \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \sup_{k=0,1,\dots} \left\{ \frac{k}{n} \mid x\left(\frac{j}{n}\right) \in I_m, j = 0, 1, \dots, k \right\}, \end{aligned}$$

where I_m is an increasing sequence of closed intervals whose union is I .

Again, since each path ω in Ω_x is continuous for $\tau \leq \mathfrak{J}_I(\omega)$, if I is bounded, $x(\mathfrak{J}_I(\omega), \omega)$ can only be one of the two end points a_1 and a_2 of I . The subspace $\{\omega \mid \mathfrak{J}_I(\omega) \leq t\}$ consists of the two disjoint measurable sets $\{\omega \mid \mathfrak{J}_I \leq t, x(\mathfrak{J}_I) = a_j\}, j = 1, 2$. Thus the distribution function of the first passage time satisfies

$$\begin{aligned} (2.3) \quad \Pr \{ \mathfrak{J}_I \leq t \mid x \} &= \Pr \{ \mathfrak{J}_I \leq t, x(\mathfrak{J}_I) = a_1 \mid x \} \\ &\quad + \Pr \{ \mathfrak{J}_I \leq t, x(\mathfrak{J}_I) = a_2 \mid x \} \\ &= P_0(t, x; I, a_1) + P_0(t, x; I, a_2). \end{aligned}$$

We shall call a point y of R *accessible from the right (left)* with respect to the process $(\Omega_x, \Pr \{ \cdot | x \})$ if there is a bounded open interval I whose left (right) end point is y , for which $P_0(t, x; I, y)$ does not vanish for all $t > 0$.

The strong Markov property which we have discussed in §I states that whenever $\Lambda = \{ \omega | \mathfrak{J}_I < \infty, x(\mathfrak{J}_I) = y \}$ has positive probability,

$$x_y(\tau, \omega) = x(\mathfrak{J}_I(\omega) + \tau, \omega), \quad \tau \geq 0, \omega \in \Lambda,$$

should define random variables independent of \mathfrak{J}_I on the sample space Λ , forming a new Markov process. Hunt [XIII] has shown that if the process in question has independent increments, then $x(\mathfrak{J}(\omega) + \tau, \omega)$, $\tau \geq 0$, defines a Markov process independent of $\mathfrak{J}(\omega)$ for a wide class of stopping methods which includes stopping at the first passage of a given point. However, it is unfortunately not true even that every process with continuous paths has this useful property. That $x_y(\tau, \omega)$ and $\mathfrak{J}_I(\omega)$ be independent is by no means equivalent to the Markov property. For the latter restricts the probability measure, conditioned by the values of the path ω up to a fixed time, only for almost every ω . For most processes of interest, the probability is zero that the first passage of a given point occurs at a given time, and the Markov property gives no information in this case. We illustrate this situation in Example V of §IV.

For the purposes of this paper, we formulate the strong Markov property for linear processes in terms of transition probabilities: namely, if $I = (a_1, a_2)$ is a bounded open interval containing x , if f is a bounded Borel measurable function on R , and if $t > 0$, then

$$\begin{aligned} E\{f(x(t)) | x\} &= E\{f(x(t)); \mathfrak{J}_I > t | x\} \\ (2.4) \quad &+ \int_0^t E\{f(x(t-t')) | a_1\} d_t P_0(t', x; I, a_1) \\ &+ \int_0^t E\{f(x(t-t')) | a_2\} d_t P_0(t', x; I, a_2). \end{aligned}$$

We shall say that the processes $\{\Omega_x, \Pr \{ \cdot | x \} | x \in R\}$ satisfy the *first passage time relation* if (2.4) holds for every choice of f, t, x , and I [II; X].

It seems to us both natural and useful to give the name “linear diffusion” to a collection of Markov processes on the line with these properties. Formally, then:

DEFINITION III. *A collection $\{\Omega_x, \Pr \{ \cdot | x \} | x \in R\}$ of Markov processes on the line determined from stationary transition probabilities will be called a DIFFUSION PROCESS if the space Ω_x is given for each x in R by Definition II and if the processes satisfy the first passage time relation (2.4).*

III. Results of this paper. In terms of the concepts just described, we may now state and discuss our results in detail.

THEOREM I. *Let $\{\Omega_x, \Pr \{ \cdot | x \} | x \in R \}$ be a collection of Markov processes determined from stationary transition probabilities on the real line R , with the path spaces $\{\Omega_x | x \in R\}$ described in Definition II. For each point x in R , let $P_0(t, x; I, a_j)$ be the distribution function of the first passage time from the bounded open interval I by way of its end point a_j , as defined in (2.3).*

A. The processes are connected by the first passage time relation (2.4) if for every bounded continuous function f on R and for every $s > 0$, the Laplace transform

$$(3.1) \quad R_s f(x) = \int_0^\infty E\{f(x(t)) | x\} e^{-st} dt, \quad x \in R,$$

is right (left) continuous at every point which is right (left) accessible with respect to at least one process $(\Omega_x, \Pr \{ \cdot | x \})$.

B. Conversely, if the processes are connected by the first passage time relation (2.4), then for every x in R , for every bounded Borel measurable function f on R for which $f(x+)(f(x-))$ exists, and for every $s > 0$, the Laplace transform $R_s f$ has right (left) limits at x . Moreover, the same statement holds if $x = \pm \infty$. Except for at most a countable number of points, independent of f and s , $R_s f$ is continuous at every point x where f is continuous; the exceptional points are all inaccessible from either the right or the left for each process $(\Omega_x, \Pr \{ \cdot | x \})$.

Besides having these regularity properties, the transition probabilities of a diffusion process are of local character uniformly on bounded intervals:

COROLLARY I.1. *Let $\{\Omega_x, \Pr \{ \cdot | x \} | x \in R\}$ be a diffusion process on the real line, in the sense of Definition III. Then for every $\delta > 0$, and for every bounded interval F in R ,*

$$(3.2) \quad \lim_{t \rightarrow 0} \frac{1}{t} \Pr \{ |x(t) - x| > \delta | x \} = 0$$

uniformly for x in F .

But even for a process whose paths are all continuous, the transition probabilities are not necessarily of local character unless the first passage time relation also holds. This is seen in the same Example V of the next section which illustrates the possible lack of the first passage time relation. That Corollary I.1 is not true except in the one-dimensional case is seen in Example II.

From the point of view of applications, it is both fortunate and natural that the first passage time relation is equivalent to a fairly simple regularity condition on the transition probabilities, for processes on the real line whose paths are continuous except at the boundary. In particular, Hypothesis A_0 given in the Introduction is sufficient to imply (2.4) for such processes; and in this case, a considerably stronger statement holds than the pointwise local character condition B_0 :

COROLLARY I.2. Let $\{\Omega_x, \Pr \{ \cdot | x \} | x \in R\}$ be a collection of Markov processes determined from transition probabilities on the real line R , with the path spaces $\{\Omega_x | x \in R\}$ described in Definition II. Suppose for every bounded continuous real function f on R ,

$$(3.3) \quad T_t f(x) = E\{f(x(t)) | x\}$$

defines a continuous function on R . Then the collection is a diffusion process, and the property (3.2) of local character holds uniformly on bounded intervals.

It is possible of course that the transition probabilities of a diffusion process will not satisfy Hypothesis A_0 . Three types of exceptional points are illustrated in Example III. Moreover, Example IV will show that even if there are no exceptional points, the semi-group (3.3) need not preserve continuity. But in this case there is always a subspace of continuous functions, described in terms of the behavior of the process at the boundary points, on which the semi-group acts.

COROLLARY I.3. Let $\{\Omega_x, \Pr \{ \cdot | x \} | x \in R\}$ be a diffusion process on the real line R with no exceptional points: for every bounded continuous function f on R , the resolvent $R_s f$, defined by (3.1) is a continuous function on R .

Let C^+ be the space of continuous functions on R for which $f(+\infty) = \lim_{x \rightarrow +\infty} f(x)$ and $f(-\infty) = \lim_{x \rightarrow -\infty} f(x)$ exist. C^+ is thus the space of continuous functions on the extended real line $R^+ = \{x | -\infty \leq x \leq +\infty\}$.

There exist two regular measures, $P(\cdot | \pm \infty)$, on R^+ such that for f in C^+ , $s > 0$,

$$(3.4) \quad R_s f(\pm \infty) = \lim_{x \rightarrow \pm \infty} R_s f(x) = \int_{R^+} R_s f(y) P(dy | \pm \infty).$$

Let C_0^+ be the subspace of C^+ of functions f for which the two relations

$$f(\pm \infty) = \int_{R^+} f(y) P(dy | \pm \infty)$$

hold. If $t > 0$ and f is in C_0^+ , then the function

$$(3.5) \quad T_t f(x) = E\{f(x(t)) | x\}$$

is in C_0^+ . The operators $\{T_t | t > 0\}$ form a semi-group which is strongly continuous for $t \geq 0$ in the uniform norm in C_0^+ :

$$\lim_{t \rightarrow 0} \text{Max}_{x \in R^+} | E\{f(x(t)) - f(x(t')) | x\} | = 0.$$

Of course, the result (3.2) of uniform local character remains true.

If there are exceptional points, then essentially the same result holds: there is a separable space of functions on which the semi-group acts. But we will not give this more complicated result a formal status.

Finally, we mention that for the semi-group defined by Corollary I.3, all the results of Feller [VIII; IX] concerning the infinitesimal generator remain true. In effect, the process is given as the solution of a generalized differential equation with certain boundary conditions.

The results in the last two paragraphs were announced, while this paper was in preparation, by Dynkin [XIX; XX], in the case of linear diffusion processes whose paths are always continuous. For such processes, Dynkin states that the domain of the infinitesimal generator consists of functions whose only discontinuities are jumps (part B of our Theorem I), and derives the form of the generator in probabilistic terms. Thus he obtains the generalization of Corollary I.3 just described; since Dynkin considers only processes whose paths are always continuous, however, the boundary conditions are trivial: $C_0^+ = C^+$. Dynkin does not mention that there are only countably many exceptional points, nor that the transition probabilities of a diffusion process on the line are of local character.

We turn now to the investigation of conditions under which given transition probabilities determine processes whose paths are all continuous except for jumps from the boundary. The essential hypothesis in past results [III; X; XV; XVIII] of this type has been that the transition probabilities be of uniform local character. The most general conditions are those of Kinney [XV], who proved that the paths of a (not necessarily stationary) Markov process are almost all continuous when they are bounded, if the transition probabilities are of local character uniformly on bounded sets, and in addition satisfy one of a number of conditions restricting their behavior near the boundary. We shall prove that the latter restrictions can be removed.

But that the local character be uniform on compact sets is a more essential restriction. It is seen in Example I of the next section that the hypothesis B_0 of pointwise local character is not sufficient in itself to ensure the desired continuity properties. However, the use of techniques of the theory of semi-groups enables us to replace the uniformity condition by hypothesis A_0 . In the light of Feller's work this seems a much more natural condition.

The use of semi-group methods of course confines the proof to the stationary case. On the other hand, it is only for convenience and consistency in notation that we restrict the space of the process to be the real line. With trivial changes our proof, like that of Kinney, applies to transition probabilities on a locally compact metric space.

THEOREM II. *Let $P(x, t, E)$ be stationary transition probabilities on the real line R . Then either of the two following conditions is sufficient that there be a collection of Markov processes $\{\Omega_x, \Pr \{ \cdot | x \} | x \in R \}$ determined from the given transition probabilities, with each path space $\Omega_x, x \in R$, as described in Definition II.*

A. 1°: For every bounded continuous function f on R and for every $t > 0$, $\int f(y)P(x, t, dy)$ is a continuous function of x in R ; and

2°: For every x in R and for every $\delta > 0$,

$$(3.6) \quad \lim_{t \rightarrow 0} \frac{1}{t} P(x, t, I'(x, \delta)) = 0,$$

where $I'(x, \delta)$ is the complement of the open interval $I(x, \delta) = (x - \delta, x + \delta)$.

B. For every $\delta > 0$ and for every bounded interval F in R ,

$$\lim_{t \rightarrow 0} \frac{1}{t} P(x, t, I'(x, \delta)) = 0$$

uniformly for x in F .

Combining Theorem II with the corollaries to Theorem I, we have a rather interesting result concerning transition probabilities satisfying Feller's conditions A_0 and B_0 .

COROLLARY II.1. Let $P(x, t, E)$ be stationary transition probabilities on the real line R which satisfy the hypotheses A of Theorem II. Then

1°: The condition (3.6) of local character holds uniformly on every bounded interval of R .

2°: There are two regular measures, $P(\cdot | \pm \infty)$, on the extended real line R^+ such that if C_0^+ denotes the space of functions f continuous on R^+ and satisfying

$$f(\pm \infty) = \int_{R^+} f(y)P(dy | \pm \infty),$$

then whenever f is in C_0^+ ,

$$T_t f(x) = \int f(y)P(x, t, dy)$$

defines, for each $t > 0$, a function in C_0^+ ; and the operators $\{T_t | t > 0\}$ form a strongly continuous semi-group in the uniform norm on C_0^+ .

We must mention, finally, one essential restriction which has been assumed throughout the paper. This is that, although the paths of a process constructed from given transition probabilities are allowed to range over the extended line R^+ , the range of the transition probabilities is restricted to R . The effect is that for every process which we consider,

$$\Pr \{x(t) \in R\} = 1$$

for each $t > 0$, although it is possible that

$$\Pr \{x(\tau) \in R, 0 \leq \tau \leq t\} < 1$$

for all $t > 0$. In this way our results do not apply to a wide range of interesting diffusion processes [VII] in which the path is allowed to remain at the boundary points or wander arbitrarily within the boundary. Actually, this restriction is only a matter of convenience in the collection of results under Theorem I. But it is essential to the proof of Theorem II which we shall give in §VI, for reasons which we shall explain at that time.

IV. Examples. We illustrate first the fact that the condition (3.6) of local character is not sufficient in itself that a process with continuous paths may be constructed from given transition probabilities.

EXAMPLE I. Define transition probabilities on the line by

$$\begin{aligned} P(x, t, E) &= c_E(x + t), & x \geq 0, t \geq 0; \\ P(x, t, E) &= c_E(x + t), & x < 0, x + t < 0, \\ P(x, t, E) &= c_E(x + t + 1), & x < 0, x + t \geq 0, \end{aligned}$$

where $c_E(x)$ is the characteristic function of the Borel set $E: c_E(x) = 1, x \in E; c_E(x) = 0, x \notin E$. The usual properties of transition probabilities are satisfied. The process determined by these consists of paths which with probability one move with unit velocity to the right along the real line, but which, on approaching the origin from the left, never reach it, but instead jump a unit distance. The paths are not continuous with probability one, but the transition probabilities are of local character at each point.

Example II shows that the result of Corollary I.1 is restricted to processes on the real line. Even in the plane, a process may have continuous paths, and satisfy the strong Markov property discussed in the Introduction, without having transition probabilities of local character. Whether or not the addition of Hypothesis A_0 is sufficient to insure local character remains an open problem.

EXAMPLE II. Define transition probabilities in the plane in polar coordinates by

$$\begin{aligned} P((r, \theta), t, E) &= c_E((r + t/|\theta|, \theta)), & r > 0, 0 < |\theta| \leq \pi, t \geq 0; \\ P((r, 0), t, E) &= c_E((r + t, 0)), & r > 0, t \geq 0; \\ P(0, t, E) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} c_E((t/|\theta|, \theta)) d\theta, & t \geq 0. \end{aligned}$$

A process is determined whose paths are in this case all continuous. Essentially the paths start at the origin by choosing a direction $\theta, -\pi < \theta \leq \pi$, at random, and translating outward along the ray of that direction with velocity $1/|\theta|$; if the direction $\theta = 0$ is chosen, the path moves with unit velocity, although this event of probability zero cannot affect the character of the process. Although the paths are all continuous, the transition probabilities are not of local character: it is easy to calculate that $P(0, t, \{r \geq 1\}) = t/\pi$.

EXAMPLE III. The diffusion process to be constructed here exhibits three types of exceptional points. The process is determined by the following motion: if the diffusing particle starts at $x = -1$, it flips a fair coin; if heads is up, the particle moves off to the left with speed one; if tails, it moves to the right with speed one until it reaches $x = 0$, where it remains forever. If the particle starts at a point just to the right of $x = 0$, it moves with speed one to the right, pausing at $x = 1$ for a random length of time $T(\omega)$ whose distribution is given by $\Pr \{T(\omega) \leq t\} = 1 - \exp(-t)$, and then renewing its motion.

Analytically, the component processes are given by:

1°: If $x(0, \omega) < -1$, $x(\tau, \omega) = x(0, \omega) - \tau$ with probability one.

2°: If $x(0, \omega) = -1$, $x(\tau, \omega) = -1 - \tau$ or $x(\tau, \omega) = \min(-1 + \tau, 0)$, each with equal probability.

3°: If $-1 < x(0, \omega) \leq 0$, $x(\tau, \omega) = \min(x(0, \omega) + \tau, 0)$ with probability one.

4°: If $0 < x(0, \omega) \leq 1$, $x(\tau, \omega) = \min(x(0, \omega) + \tau, \max(x(0, \omega) + \tau - T(\omega), 1))$, with the probability distribution determined from that of $T(\omega)$.

5°: If $1 < x(0, \omega)$, $x(\tau, \omega) = x(0, \omega) + \tau$ with probability one.

The points $x = -1, 0, 1$ are exceptional points for this diffusion process. For instance, if $f(x) = x$, the resolvent

$$R_s f(x) = \int_0^\infty e^{-s t} E\{x(t) \mid x\} dt$$

is discontinuous from the right at $x = 0$ and $x = 1$, and is discontinuous from both right and left at $x = -1$. We might call the point $x = -1$ a source, the point $x = 0$ a sink, and $x = 1$ a waiting point.

EXAMPLE IV. This example illustrates the fact that the resolvent $R_s f(x)$ of the semi-group defined by a diffusion process may preserve the continuity of the function f , while the semi-group fails to do so except on a subspace of continuous functions.

A particle starting at the point x translates to the left with velocity increasing exponentially: $(d/dt)x(t) = -\exp(-x(t))$. The particle reaches the boundary point $-\infty$ at time $t = e^x$, and then jumps to the origin and tries again. The process is completely deterministic, and so needs no further description.

The resolvent is given by

$$R_s f(x) = e^{-s e^x} \int_{-\infty}^x e^{s e^y} f(y) e^y dy + \frac{e^{-s(e^x+1)}}{1 - e^{-s}} \int_{-\infty}^0 e^{s e^y} f(y) e^y dy,$$

and defines a continuous function $R_s f$ when f is continuous. There are no exceptional points. However, the transform

$$T_t f(x) = E\{f(x(t)) \mid x\} = \begin{cases} f(\log(e^x - t)), & t < e^x \\ f(\log(e^x + k - t)), & k - 1 + e^x \leq t < k + e^x, \end{cases}$$

for f continuous, is continuous neither in t nor in x , in general. But the sub-space of continuous functions satisfying

$$\lim_{x \rightarrow -\infty} f(x) = f(0)$$

is transformed into itself by each operator T_t , and the semi-group thus defined is strongly continuous for $t \geq 0$.

Finally, we give a more sophisticated example to show that the strong Markov property may fail for a Markov process with continuous paths, and that if so, the transition probabilities may not be of local character. The example is inspired by a construction of Paul Lévy [XVII, II. 10, 4: p. 366].

EXAMPLE V. For each prime $p \geq 2$, let r_p denote a variable point of the sequence $\{kp^{-n} \mid k=0, 1, \dots, p^n-1; n=1, 2, \dots\}$. For each p we define a space $\Omega^{(p)}$ of paths $\{\omega = x(\tau, \omega) \mid \tau \geq 0\}$ each of which is the inverse of a Cantor-type function, proceeding monotonically from 0 to 1, moving only by jumps, and pausing at each point r_p for a time $T(r_p, \omega)$. $\Omega^{(p)}$ can be represented as the space of the random variables $T(r_p, \omega)$, a countably dimensioned product space. A Borel measure is determined on $\Omega^{(p)}$ by requiring that the waiting times be independent, and that, if the irreducible form of r_p is kp^{-n} , $T(r_p, \omega)$ have an exponential distribution with lifetime $\mu(r_p) = p^{-2n}$: $\Pr^{(p)} \{T(r_p, \omega) \leq t\} = 1 - \exp(-t/\mu(r_p))$. If $r_p = 0$, we set $\mu(r_p) = 1/p$. With probability one, $\sum T(r_p, \omega)$ converges, and the function

$$\tau(x, \omega) = \sum_{r_p \leq x} T(r_p, \omega)$$

is bounded and strictly increasing on the interval $0 \leq x \leq 1$. The corresponding path $x(\tau, \omega)$ in $\Omega^{(p)}$ is the inverse of this function:

$$\begin{aligned} x(\tau, \omega) &= x, & \sum_{r_p < x} T(r_p, \omega) &\leq \tau \leq \sum_{r_p \leq x} T(r_p, \omega), \\ x(\tau, \omega) &= 1, & \tau(1, \omega) &= \sum_{r_p < 1} T(r_p, \omega) \leq \tau. \end{aligned}$$

It is clear that almost every path thus constructed is continuous for $\tau \geq 0$, and that with the measure $\Pr^{(p)}\{\cdot\}$ determined by the waiting times, we have a stationary Markov process on $[0, 1]$.

Consider now the space of paths

$$\Omega = \bigcup_p \Omega^{(p)},$$

with the measure

$$\Pr \{ \cdot \} = \sum_p C_p \Pr^{(p)} \{ \cdot \cap \Omega^{(p)} \},$$

where $C_p = 1/p' - 1/p$, p' being the largest prime less than p , or if $p = 2$, $p' = 1$. We state that these form a stationary Markov process on $[0, 1]$: the condi-

tional probability of $x(t+s, \omega)$ being in a Borel set E , given $x(\tau, \omega)$ for $\tau \leq t$, is a function only of E , s , and $x(t, \omega)$ for almost every ω . Indeed for $t \geq 0$, the value of $x(t, \omega)$ is with probability one in the sequence $\{r_p\}$ whenever ω is in $\Omega^{(p)}$; and since these sequences for different primes p have only 0 in common, if $x(t, \omega) = r_p > 0$, then ω is in $\Omega^{(p)}$ with probability one, and the distribution of ω is given by $\text{Pr}^{(p)}$. In this case the Markov property is a consequence of the independence of the waiting times used to construct this measure. If on the other hand $x(t, \omega) = 0$, the conditional probability in question is simply $\text{Pr} \{x(s, \omega) \in E\}$.

The process $\{\Omega, \text{Pr} \{ \cdot \} \}$ does not have the strong Markov property. In particular, let I be the open interval $[0, 1/2)$ of $[0, 1]$. If $\omega \in \Omega^{(p)}$, then $\mathfrak{J}_I(\omega) = \sum_{r_p < 1/2} T(r_p, \omega)$, while $x(t + \mathfrak{J}_I(\omega), \omega)$ depends only on $T(r_p, \omega)$ for $r_p \geq 1/2$. Hence for ω in $\Omega^{(p)}$, $\mathfrak{J}_I(\omega)$ is independent of $x(t + \mathfrak{J}_I(\omega), \omega)$. But then clearly $\mathfrak{J}_I(\omega)$ is independent of $x(t + \mathfrak{J}_I(\omega), \omega)$ when ω ranges over all of Ω if and only if \mathfrak{J}_I is independent of the choice of the component $\Omega^{(p)}$; and this last statement is certainly not true.

Intuitively, the process is very similar to Example II. For the paths start at the origin by choosing a prime p at random, and then proceeding along the "direction" $\{r_p\}$. Although the path passes through every point in $[0, 1]$, at a fixed time t it is almost surely in $\{r_p\}$, and because of the Markov property does not remember any other values. Hence the processes for different primes p are completely independent. Of course the speed of the path depends on the choice of p ; if \mathfrak{J}_I is small, then it is more likely that a large value of p has been chosen, and that the path will continue with a high speed. Hence the Markov property holds, but not the strong Markov property.

The similarity with Example II suggests that this process has been constructed also to show that a linear Markov process with continuous paths need not have transition probabilities of local character, and this is indeed true. In fact if $\lim_{t \rightarrow 0} 1/t \text{Pr} \{x(t) = 1\} = 0$, then the Laplace-Stieltjes transform

$$\phi(s) = \int_0^\infty e^{-st} d_t \text{Pr} \{x(t) = 1\}$$

would satisfy $\lim_{s \rightarrow \infty} s\phi(s) = 0$. But $\phi(s)$ can be expressed explicitly as

$$\begin{aligned} \phi(s) &= \sum_p C_p (1 + p^{-1}s)^{-1} \prod_{n=1}^\infty (1 + p^{-2n}s)^{-(p^n - p^{n-1})} \\ &\geq \sum_p C_p e^{-p^{-1}s} \prod_{n=1}^\infty e^{-p^{-2n}s(p^n - p^{n-1})} \\ &= \sum_p C_p e^{-2s/p} \geq e^{-2} \sum_{r \geq s} c_p \geq e^{-2}/s. \end{aligned}$$

We complete this example by showing that the process has transition probabilities. These are uniquely defined by the process for every initial point in one of the sequences $\{r_p\}$ for p prime. For instance, the Laplace-Stieltjes transform of $P(r_p, t, [x, 1])$ is

$$\prod_{r_p \leq r_{p'} < x} (1 + s\mu(r_{p'}))^{-1}, \quad s \geq 0, r_p > 0,$$

or

$$\sum_p C_p \prod_{0 \leq r_{p'} < x} (1 + s\mu(r_{p'}))^{-1}, \quad s \geq 0, r_p = 0.$$

It is quite clear that transition probabilities may be defined for x not in any sequence $\{r_p\}$ in a completely arbitrary fashion, as long as they satisfy the conditions of Definition I. For instance, we may set for such x

$$P(x, t, E) = C_E(x) = 1, \quad x \in E, \\ = 0, \quad x \notin E,$$

for $t \geq 0, E \subset R$. That this choice and the above form stationary transition probabilities on the real line R is trivial to verify.

V. The transition probabilities of a diffusion process. The proofs of Theorem I and its corollaries will be given in this section. Throughout the section we assume a given collection of Markov processes $\{\Omega_x, \Pr \{\cdot | x\} | x \in R\}$ determined from stationary transition probabilities on the real line R , with the path spaces $\{\Omega_x | x \in R\}$ as described in Definition II.

We shall divide the reasoning into a series of lemmas, but we precede these with some preliminary results concerning processes of the type considered.

We note first that the integrals involved in (2.4) and (3.1) exist. In fact, when x is in R and f is a bounded continuous function on $R, E\{f(x(t)) | x\}$ is a right continuous function of $t \geq 0$. To show this, and only for this purpose, we use the fact that for every $t \geq 0, \Pr \{x(t) \in R | x\} = 1$. Because of the continuity properties of the paths, if y is in $R, \lim_{\Delta \rightarrow 0} E\{f(x(\Delta)) | y\} = f(y)$. Hence

$$\begin{aligned} \lim_{\Delta \rightarrow 0} E\{f(x(t + \Delta)) | x\} &= \lim_{\Delta \rightarrow 0} \int E\{f(x(\Delta)) | y\} \Pr \{x(t) \in dy | x\} \\ &= \int f(y) \Pr \{x(t) \in dy | x\} \\ &= E\{f(x(t)) | x\}. \end{aligned}$$

Since continuous functions are dense among the bounded Borel measurable functions on $R, (2.4)$ and (3.1) have a sense when f is a function in the latter space; moreover, the validity of the first passage time relation (2.4) for bounded continuous functions implies (2.4) for all bounded Borel measurable

functions. Finally, because of the right continuity of $E\{f(x(t))|x\}$ when f is continuous, (2.4) is equivalent to the proposition that whenever x is a point of R , $I = (a_1, a_2)$ is a bounded open interval containing x , f is a bounded continuous function on R , and $s > 0$, then

$$(5.1) \quad R_s f(x) = \int_0^\infty e^{-st} E\{f(x(t)); \mathfrak{J}_T > t | x\} dt \\ + R_s f(a_1) \phi_s(x; I, a_1) + R_s f(a_2) \phi_s(x; I, a_2),$$

where $R_s f$, defined by (3.1), is the Laplace transform of $E\{f(x(t))|\cdot\}$, and

$$(5.2) \quad \phi_s(x; I, a_j) = \int_0^\infty e^{-st} d_t P_0(t, x; I, a_j).$$

Our method of treating the first passage time relation requires some computation; this can be made somewhat less tedious by the use of a technique of Kac [XIV]. For every sufficiently small positive δ , let $V_\delta^{(1)}$ and $V_\delta^{(2)}$ be the characteristic functions of the open intervals $(a_1, a_1 + \delta)$ and $(a_2 - \delta, a_2)$, respectively; let $V_\delta = V_\delta^{(1)} + V_\delta^{(2)}$. For each ω in Ω_x , the functions $V_\delta(x(\tau, \omega))$, $V_\delta^{(j)}(x(\tau, \omega))$, $j = 1, 2$, are because of the continuity property integrable over every τ -interval. We claim that

$$(5.3) \quad \lim_{\delta \rightarrow 0} \lim_{u \rightarrow \infty} u \int_0^t \exp \left\{ -u \int_0^{t'} V_\delta(x(\tau, \omega)) d\tau \right\} V_\delta(x(t', \omega)) dt' \\ = \begin{cases} 1, & \mathfrak{J}_T(\omega) \leq t, \\ 0, & \mathfrak{J}_T(\omega) > t. \end{cases}$$

For

$$u \int_0^t \exp \left\{ -u \int_0^{t'} V_\delta(x(\tau)) d\tau \right\} V_\delta(x(t')) dt' = 1 - \exp \left\{ -u \int_0^t V_\delta(x(\tau)) d\tau \right\}.$$

If $\mathfrak{J}_T(\omega) \leq t$, then for every $\delta > 0$, the path ω spends a positive time η in the set $(a_1, a_1 + \delta) \cup (a_2 - \delta, a_2)$; for such a path,

$$\lim_{u \rightarrow \infty} \exp \left\{ -u \int_0^t V_\delta(x(\tau)) d\tau \right\} \leq \lim_{u \rightarrow \infty} e^{-u\eta} = 0$$

for each $\delta > 0$, so that (5.3) holds. If on the other hand $\mathfrak{J}_T(\omega) > t$, then there is $\delta_0 > 0$ such that

$$a_1 + \delta_0 \leq x(\tau, \omega) \leq a_2 - \delta_0 \text{ for } 0 \leq \tau \leq t;$$

if $\delta < \delta_0$, $\int_0^t V_\delta(x(\tau, \omega)) d\tau = 0$ for every u , implying (5.3) in this case as well.

We have likewise, for $j = 1, 2$,

$$(5.4) \quad \lim_{\delta \rightarrow 0} \lim_{u \rightarrow \infty} u \int_0^t \exp \left\{ -u \int_0^{t'} V_\delta(x(\tau, \omega)) d\tau \right\} V_\delta^{(j)}(x(t', \omega)) dt' \\ = \begin{cases} 1, & \mathfrak{J}_I(\omega) \leq t, x(\mathfrak{J}_I) = a_j, \\ 0, & \text{otherwise.} \end{cases}$$

The sum of the functionals defined in (5.4) for $j=1, 2$ is the functional of (5.3). Thus (5.4) clearly follows if we show that the functional is not less than one for a path ω for which $\mathfrak{J}_I(\omega) \leq t$ and $x(\mathfrak{J}_I(\omega), \omega) = a_j$. For such a path there is $\delta_0 > 0$ such that $V_\delta^{(j)}(x(\tau, \omega)) = V_\delta(x(\tau, \omega))$ whenever $\delta < \delta_0$ and $0 \leq \tau \leq \mathfrak{J}_I(\omega)$. Thus

$$u \int_0^t \exp \left\{ -u \int_0^{t'} V_\delta(x(\tau)) d\tau \right\} V_\delta^{(j)}(x(t')) dt' \\ \geq u \int_0^{\mathfrak{J}_I} \exp \left\{ -u \int_0^{t'} V_\delta(x(\tau)) d\tau \right\} V_\delta^{(j)}(x(t')) dt' \\ = u \int_0^{\mathfrak{J}_I} \exp \left\{ -u \int_0^{t'} V_\delta(x(\tau)) d\tau \right\} V_\delta(x(t')) dt' \\ = 1 - \exp \left\{ -u \int_0^{\mathfrak{J}_I} V_\delta(x(\tau)) d\tau \right\}.$$

But certainly $\mathfrak{J}_I(\omega) \leq t$ implies $\int_0^{\mathfrak{J}_I} V_\delta(x(\tau, \omega)) d\tau > 0$, whence (5.4) follows as above.

Turning now to the proof of Theorem I, we restate Part A as

LEMMA V.1. *If for every bounded continuous function f on R , and for every $s > 0$, the resolvent $R_s f$ is right (left) continuous at every point which is right (left) accessible with respect to at least one of the given processes*

$$\{\Omega_x, \text{Pr} \{ \cdot | x \} | x \in R\},$$

then the first passage time relation (2.4) holds.

In view of (5.1), it is sufficient to prove that

$$(5.5) \quad \int_0^\infty e^{-st} E\{f(x(t)); \mathfrak{J}_I \leq t, x(\mathfrak{J}_I) = a_j | x\} dt \\ = R_s f(a_j) \phi_s(x; I, a_j), \quad j = 1, 2,$$

holds under the given hypothesis, whenever f is a bounded continuous function on R , and $s > 0$. We give the proof only for $j=1$, obvious changes providing for the case $j=2$.

If a_1 is inaccessible from the right with respect to the process with initial point x , then certainly both sides of (5.5) vanish. If not, then according to

the hypothesis, $R_\delta f$ is right continuous at the point a_1 .

In this case, (5.4) and the Markov property imply

$$\begin{aligned}
 & \int_0^\infty e^{-st} E \{ f(x(t)); \exists t \leq t, x(\exists t) = a_1 | x \} dt \\
 &= \int_0^\infty e^{-st} dt \lim_{\delta \rightarrow 0} \lim_{u \rightarrow \infty} E \left\{ f(x(t)) \right. \\
 &\quad \left. \cdot u \int_0^t \exp \left\{ -u \int_0^{t'} V_\delta(x(\tau)) d\tau \right\} V_\delta^{(1)}(x(t')) dt' | x \right\} \\
 &= \lim_{\delta \rightarrow 0} \lim_{u \rightarrow \infty} u \int_0^\infty e^{-st} dt \int_0^t dt' E \left\{ f(x(t)) \right. \\
 &\quad \left. \cdot \exp \left\{ -u \int_0^{t'} V_\delta(x(\tau)) d\tau \right\} V_\delta^{(1)}(x(t')) | x \right\} \\
 &= \lim_{\delta \rightarrow 0} \lim_{u \rightarrow \infty} u \int_0^\infty e^{-st'} dt' \int_{t'}^\infty e^{-s(t-t')} dt E \left\{ E \{ f(x(t-t')) | x(t') \} \right. \\
 &\quad \left. \cdot \exp \left\{ -u \int_0^{t'} V_\delta(x(\tau)) d\tau \right\} V_\delta^{(1)}(x(t')) | x \right\} \\
 &= \lim_{\delta \rightarrow 0} \lim_{u \rightarrow \infty} u \int_0^\infty e^{-st'} dt' E \left\{ R_\delta f(x(t')) \right. \\
 &\quad \left. \cdot \exp \left\{ -u \int_0^{t'} V_\delta(x(\tau)) d\tau \right\} V_\delta^{(1)}(x(t')) | x \right\}.
 \end{aligned}$$

Now

$$\begin{aligned}
 \phi_\delta(x; I, a_1) &= \int_0^\infty e^{-st} d_t P_0(t, x; I, a_1) \\
 &= s \int_0^\infty e^{-st} P_0(t, x; I, a_1) dt \\
 &= s \int_0^\infty e^{-st} dt E \left\{ \lim_{\delta \rightarrow 0} \lim_{u \rightarrow \infty} u \int_0^t \exp \left\{ -u \int_0^{t'} V_\delta(x(\tau)) d\tau \right\} \right. \\
 &\quad \left. \cdot V_\delta^{(1)}(x(t')) dt' | x \right\} \\
 &= \lim_{\delta \rightarrow 0} \lim_{u \rightarrow \infty} u \int_0^\infty e^{-st} dt E \left\{ \exp \left\{ -u \int_0^t V_\delta(x(\tau)) d\tau \right\} V_\delta^{(1)}(x(t)) | x \right\}.
 \end{aligned}$$

Thus

$$\int_0^\infty e^{-st} dt E\{f(x(t)); \mathfrak{J}_I \leq t, x(\mathfrak{J}_I) = a_1 | x\}$$

differs from $R_s f(a_1)\phi_s(x; I, a_1)$ by less than

$$\begin{aligned} \lim_{\delta \rightarrow 0} \lim_{u \rightarrow \infty} u \int_0^\infty e^{-st} dt E \left\{ \left| R_s f(x(t)) - R_s f(a_1) \right| \right. \\ \left. \cdot \exp \left\{ -u \int_0^t V_\delta(x(\tau)) d\tau \right\} V_\delta^{(1)}(x(t)) \mid x \right\} \\ \leq \lim_{\delta \rightarrow 0} \sup_{a_1 \leq y \leq a_1 + \delta} \left| R_s f(y) - R_s f(a_1) \right| \\ = 0, \end{aligned}$$

since $V_\delta^{(1)}$ vanishes outside $(a_1, a_1 + \delta)$ and since $R_s f$ is right continuous at a_1 .

The proof of Part B of Theorem I is based on the following fundamental relations for the distribution functions of the first passage times for a diffusion process:

LEMMA V.2. *Suppose the given processes $\{\Omega_x, \Pr \{ \cdot | x \} | x \in R\}$ satisfy the first passage time relation (2.4). Let $\phi_s(x; I, a_j)$, given by (5.2), be the Laplace-Stieltjes transform of the first passage time distribution function. Then*

$$(5.6) \quad \phi_s(x; I, a_1) = \phi_s(x; J, y)\phi_s(y; I, a_1),$$

$$(5.7) \quad \phi_s(x; I, a_2) = \phi_s(x; J, a_2) + \phi_s(x; J, y)\phi_s(y, I, a_2),$$

where $I = (a_1, a_2)$, $J = (y, a_2)$, with $-\infty < a_1 < y < x < a_2 < +\infty$, and $s > 0$. The same relation holds when $-\infty < a_2 < x < y < a_1 < +\infty$.

Preliminary to the proof of Lemma V.2 we show that (2.4) implies

$$(5.8) \quad \begin{aligned} E\{f(x(t)); \mathfrak{J}_J \leq t < \mathfrak{J}_I, x(\mathfrak{J}_J) = y | x\} \\ = \int_0^t E\{f(x(t-t')); \mathfrak{J}_I > t-t' | y\} d_t P_0(t' | x; J, y) \end{aligned}$$

whenever f is bounded and Borel measurable on R , and x, y, I, J are as above. By (5.3) and the equation following (5.3),

$$\begin{aligned} E\{f(x(t)); \mathfrak{J}_J \leq t < \mathfrak{J}_I, x(\mathfrak{J}_J) = y | x\} \\ = \lim_{\delta \rightarrow 0} \lim_{u \rightarrow \infty} E \left\{ f(x(t)) \exp \left\{ -u \int_0^t V_\delta(x(\tau)) d\tau \right\}; \mathfrak{J}_J \leq t, x(\mathfrak{J}_J) = y | x \right\} \\ = \lim_{\delta \rightarrow 0} \lim_{u \rightarrow \infty} E \left\{ f(x(t)) \exp \left\{ -u \int_{\mathfrak{J}_J}^t V_\delta(x(\tau)) d\tau \right\}; \mathfrak{J}_J \leq t, x(\mathfrak{J}_J) = y | x \right\}, \end{aligned}$$

since for each path ω , $x(\mathfrak{J}_J) = y$ implies there is δ with

$$\int_0^{\mathfrak{J}_J} V_\delta(x(\tau)) d\tau = 0.$$

Now for each u and δ , the expectation above is

$$\begin{aligned} & E \left\{ f(x(t)) \sum_{n=0}^{\infty} (-u)^n \frac{1}{n!} \left\{ \int_{\mathfrak{J}_J}^t V_\delta(x(\tau)) d\tau \right\}^n ; \mathfrak{J}_J \leq t, x(\mathfrak{J}_J) = y \mid x \right\} \\ &= \sum_{n=0}^{\infty} (-u)^n E \left\{ f(x(t)) \int_{\mathfrak{J}_J}^t dt_1 \int_{t_1}^t dt_2 \cdots \int_{t_{n-1}}^t dt_n \right. \\ &\quad \left. \cdot V_\delta(x(t_1)) \cdots V_\delta(x(t_n)); \mathfrak{J}_J \leq t, x(\mathfrak{J}_J) = y \mid x \right\} \\ &= \sum_{n=0}^{\infty} (-u)^n \int_0^t dt_1 \int_{t_1}^t dt_2 \cdots \int_{t_{n-1}}^t dt_n \\ &\quad \cdot E \{ V_\delta(x(t_1)) \cdots V_\delta(x(t_n)) f(x(t)); \mathfrak{J}_J \leq t_1, x(\mathfrak{J}_J) = y \mid x \} \\ &= \sum_{n=0}^{\infty} (-u)^n \int_0^t dt_1 \int_{t_1}^t dt_2 \cdots \int_{t_{n-1}}^t dt_n \int_0^{t_1} d_t' P_0(t', x; J, y) \\ &\quad \cdot E \{ V_\delta(x(t_1 - t')) \cdots V_\delta(x(t_n - t')) f(x(t - t')) \mid y \} \end{aligned}$$

by (2.4). Interchanging integrations and reversing the above steps, the last expression is seen to equal

$$\int_0^t E \left\{ f(x(t - t')) \exp \left\{ -u \int_0^{t-t'} V_\delta(x(\tau)) d\tau \right\} \mid y \right\} d_t' P_0(t', x; J, y)$$

and (5.8) follows if we let u tend to ∞ and δ tend to 0.

Applying (5.1) and (5.8),

$$\begin{aligned} & R_\delta f(a_1) \phi_\delta(x; I, a_1) + R_\delta f(a_2) \phi_\delta(x; I, a_2) \\ &= R_\delta f(x) - \int_0^\infty e^{-st} E \{ f(x(t)); \mathfrak{J}_I > t \mid x \} dt \\ &= R_\delta f(x) - \int_0^\infty e^{-st} E \{ f(x(t)); \mathfrak{J}_J > t \mid x \} dt \\ &\quad - \int_0^\infty e^{-st} E \{ f(x(t)); \mathfrak{J}_J \leq t < \mathfrak{J}_I; x(\mathfrak{J}_J) = y \mid x \} dt \\ &= R_\delta f(y) \phi_\delta(x; J, y) + R_\delta f(a_2) \phi_\delta(x; J, a_2) \\ &\quad - \int_0^\infty e^{-st} E \{ f(x(t)); \mathfrak{J}_I > t \mid y \} dt \phi_\delta(x; J, y) \\ &= R_\delta f(a_1) \phi_\delta(y; I, a_1) \phi_\delta(x; J, y) + R_\delta f(a_2) \phi_\delta(y; I, a_2) \phi_\delta(x; J, y) \\ &\quad + R_\delta f(a_2) \phi_\delta(x; J, a_2) \end{aligned}$$

whenever f is a bounded Borel measurable function on R and $s > 0$. Let f_j be the characteristic function of a small open interval about the point $a_j, j = 1, 2$. The path continuity clearly implies that $\lim_{s \rightarrow \infty} sR_s f_i(a_j) = \delta_{ij}$; for sufficiently large s , then, the determinant $|R_s f_i(a_j)|$ is not zero. But this and the above identity imply (5.6) and (5.7) for all sufficiently large values of s . Since each function involved in (5.6) and (5.7) is a Laplace-Stieltjes transform, (5.6) and (5.7) must then hold for all $s > 0$.

We are now in a position to discuss the continuity properties of the resolvent $R_s f$, and thus complete the proof of Theorem I.

LEMMA V.3. *Suppose the given collection of processes satisfies the first passage time relation. Then the resolvent $R_s f$, defined by (3.1), satisfies for each $s > 0$:*

1°: $R_s f(x+)(R_s f(x-))$ exists at every point x at which $f(x+)(f(x-))$ exists.

2°: $\lim_{x \rightarrow +\infty} R_s f(x)(\lim_{x \rightarrow -\infty} R_s f(x))$ exists if $\lim_{x \rightarrow +\infty} f(x)(\lim_{x \rightarrow -\infty} f(x))$ exists.

3°: *If f is continuous at x , then so is $R_s f$, unless x is in an exceptional set which is at most countable, which is independent of s and f , and which consists only of points inaccessible from either left or right with respect to every process.*

Since $0 \leq \phi_s(x; J, y) \leq 1$, (5.6) implies that $\phi_s(x; I, a_1)$ is a decreasing function of x in I . Thus the right and left limits

$$\phi_s(x+; I, a_1) = \lim_{y \searrow x} \phi_s(y; I, a_1),$$

$$\phi_s(x-; I, a_1) = \lim_{y \nearrow x} \phi_s(y; I, a_1)$$

exist for each x in $[a_1, a_2]$ and $(a_1, a_2]$ respectively, and the points of discontinuity of $\phi_s(\cdot; I, a_1)$ in I are at most countable in number. So likewise are the discontinuities of $\phi_s(\cdot; I, a_2)$, because of the obvious duality.

Each of these points of discontinuity is inaccessible from either the left or right with respect to each process. For by (5.6), and since $\lim_{y \searrow a_1} \phi_s(x; J, y) = \phi_s(x; I, a_1)$,

$$\begin{aligned} \phi_s(a_1+; I, a_1) &= \lim_{x \searrow a_1} \phi_s(x; I, a_1) \\ &= \lim_{x \searrow a_1} \lim_{y \searrow a_1} (\phi_s(x; J, y)\phi_s(y; I, a_1)) \\ &= \lim_{x \searrow a_1} (\phi_s(x; I, a_1)\phi_s(a_1+; I, a_1)) \\ &= (\phi_s(a_1+; I, a_1))^2. \end{aligned}$$

Hence for each $s > 0, \phi_s(a_1+; I, a_1) = 0$ or 1 . In the former case, by (5.6), $\phi_s(x; I, a_1) = 0$ for all x in I , implying $P_0(t, x; I, a_1) = 0$ for all $t > 0$. It is not hard to prove that then $P_0(t, x; (a_1, b), a_1) = 0$ for all $t > 0, b > x > a_1$. In other

words, for each $s > 0$ and interval $I = (a_1, a_2)$, either $\phi_s(a_1+; I, a_1) = 1$ or a_1 is inaccessible from the right with respect to each process.

Now we infer from (5.6) and (5.7) that

$$\begin{aligned} \phi_s(y+; I, a_1) &= \phi_s(y+; (y, a_2), y)\phi_s(y; I, a_1), \\ \phi_s(y+; I, a_2) &= \phi_s(y+, (y, a_2), a_2) + \phi_s(y+; (y, a_2), y)\phi_s(y; I, a_2) \end{aligned}$$

whenever $y \in I = (a_1, a_2)$ and $s > 0$. If y is accessible from the right, then $\phi_s(y+; (y, a_2), y) = 1$, $\phi_s(y+; (y, a_2), a_2) = 0$, implying

$$\begin{aligned} \phi_s(y+; I, a_1) &= \phi_s(y; I, a_1), \\ \phi_s(y+; I, a_2) &= \phi_s(y; I, a_2). \end{aligned}$$

A dual argument shows that whenever y is accessible from the left, $\phi_s(\cdot; I, a_j)$ is left continuous at y , $j = 1, 2$.

Since $\phi_s(x; I, a_j)$ is a continuous decreasing function of $s > 0$, it is clear that continuity of $\phi_s(\cdot; I, a_j)$ at a point x of I for rational $s > 0$ implies continuity at x for every $s > 0$. Hence unless x belongs to at most a countable exceptional set, $\phi_s(\cdot; I, a_j)$ is continuous at x for every open interval I with rational end points containing x , for each end point a_j of I , and for every $s > 0$. We have shown above that every exceptional point is inaccessible.

If x is not an exceptional point, and if f is a bounded Borel measurable function on R which is continuous at x then for every $s > 0$, $R_s f$ is continuous at x : We may assume $f(x) = 0$; otherwise we subtract a suitable multiple of the unit function from f , which certainly does not change the continuity properties of $R_s f$. Given $\epsilon > 0$, there is an open interval $I = (a_1, a_2)$, with rational end points and containing x , in which $|f(y)| < \epsilon$. This implies that for y in I and $s > 0$,

$$\left| \int_0^\infty e^{-st} E\{f(x(t)); \mathfrak{F}_t > t \mid y\} dt \right| < \epsilon/s.$$

But by (5.1)

$$\begin{aligned} R_s f(y) &- \int_0^\infty e^{-st} E\{f(x(t)); \mathfrak{F}_t > t \mid y\} dt \\ &= \int_0^\infty e^{-st} E\{f(x(t)); \mathfrak{F}_t \leq t \mid y\} dt \\ &= R_s f(a_1)\phi_s(y; I, a_1) + R_s f(a_2)\phi_s(y; I, a_2), \end{aligned}$$

and by hypothesis the right-hand side is continuous at x . Therefore

$$\limsup_{y \rightarrow x} |R_s f(y) - R_s f(x)| < \epsilon/s;$$

since ϵ is arbitrary, $R_s f$ is continuous at x .

Similarly, whether x is exceptional or not, $R_s f(x+)$ and $R_s f(x-)$ exist for every $s > 0$ and for every bounded Borel measurable function f for which $f(x+)$ and $f(x-)$ exist. We discuss only the right-hand limit. Again, assuming $f(x+) = 0$, there is an interval $I = (x, a)$ in which $|f(y)| < \epsilon$. Let $\{a_n\}$ be a sequence of points of I decreasing to x , for which

$$R_s f(a_n) < \liminf_{y \searrow x} R_s f(y) + \epsilon/s.$$

Since the interval $I_n = (a_{n+1}, a_n)$ is contained in I ,

$$\left| \int_0^\infty e^{-st} E \{ f(x(t)); \exists I_n > t \mid y \} dt \right| < \epsilon/s$$

for y in I_n . Then

$$\begin{aligned} R_s f(y) &\leq R_s f(a_{n+1}) \phi_s(y; I_n, a_{n+1}) + R_s f(a_n) \phi_s(y; I_n, a_n) + \epsilon/s \\ &\leq \liminf_{y \searrow x} R_s f(y) + 2\epsilon/s \end{aligned}$$

since $\phi_s(y; I_n, a_{n+1}) + \phi_s(y; I_n, a_n) \leq 1$. If $x < y < a_1$, either y is in I_n for some n , or $y \in \{a_n\}$. Thus

$$\limsup_{y \searrow x} R_s f(y) \leq \liminf_{y \searrow x} R_s f(y) + 2\epsilon/s,$$

and, since ϵ is arbitrary, this implies that $R_s f(x+)$ exists. Observe that there is no need to restrict x to be finite in the above proof: we have shown in addition that for every bounded Borel measurable function f such that $\lim_{x \rightarrow -\infty} f(x) \lim_{x \rightarrow +\infty} f(x)$ exists, and for every $s \geq 0$, $\lim_{x \rightarrow -\infty} R_s f(x)$ ($\lim_{x \rightarrow +\infty} R_s f(x)$) exists.

The combination of Lemmas V.2 and V.3 yields the second part of Theorem I. Corollary I.1 is next on the list, and follows from the first passage time relation and the continuity of the path functions. Again, we restate the result as a lemma.

LEMMA V.4. *If the given collection $\{\Omega_x, \Pr \{ \cdot \mid x \} \mid x \in R\}$ is a diffusion process, then for every $\delta > 0$*

$$\lim_{t \rightarrow 0} \frac{1}{t} \Pr \{ |x(t) - x| > \delta \mid x \} = 0$$

uniformly for x in each bounded interval.

It is sufficient to prove that if $\delta > 0$ and x in R are given, and if I is the interval $(x - 4\delta, x + 4\delta)$, then

$$\lim_{t \rightarrow 0} \frac{1}{t} P_0(t, y; I, x + 4\delta) = 0$$

uniformly for $|y - x| \leq \delta$. For again a trivial change yields the same result for $P_0(t, y; I, x - 4\delta)$; and then the inequality

$$\begin{aligned} \Pr \{ |x(t) - x| \geq 4\delta \mid y \} &\leq \Pr \{ \mathfrak{J}_I \leq t \mid y \} \\ &\leq P_0(t, y; I, x - 4\delta) + P_0(t, y; I, x + 4\delta) \end{aligned}$$

combined with these results implies V.4.

To this effect observe that because of the continuity properties of the paths, we may choose t_0 so that $t' \leq t_0$ implies

$$\Pr \{ \mathfrak{J}_J > t' \mid x' \} \geq 1/2$$

where $x' = x + 2\delta$ and J is the interval $(x' - \delta, x' + \delta)$. For the same reason, if $\epsilon > 0$ is given, there is t such that $t' \leq t$ implies

$$\begin{aligned} \epsilon &> \Pr \{ \text{For some } \tau, \tau', 0 < \tau < \tau' < \tau + t' < t_0 + t' : \\ &\quad |x(\tau) - x'| < \delta, |x(\tau') - x| \geq 4\delta \mid x' \} \\ &\geq \sum_{0 \leq k t' < t_0} \Pr \{ \mathfrak{J}_J > k t', k t' < \mathfrak{J}_I \leq (k + 1) t' \mid x' \} \\ &\geq \sum_{0 \leq k t' < t_0} \int_{|x' - z| < \delta} \Pr \{ \mathfrak{J}_J > k t', x(k t') \in dz \mid x' \} P_0(t', z; I, x + 4\delta). \end{aligned}$$

But if $|x' - z| < \delta$, then $|x - y| < \delta$ implies $y < z$; and it clearly follows from (5.6) that

$$P_0(t', z; I, x + 4\delta) \geq P_0(t', y; I, x + 4\delta)$$

for such y . Hence using the first inequality in this paragraph, if $|x - y| < \delta$,

$$\begin{aligned} \epsilon &> \sum_{0 \leq k t' < t_0} \Pr \{ \mathfrak{J}_J > k t' \mid x' \} P_0(t', y; I, x + 4\delta) \\ &\geq \sum_{0 \leq k t' < t_0} \frac{1}{2} P_0(t', y; I, x + 4\delta) \\ &\geq \frac{t_0}{2 t'} P_0(t', y; I, x + 4\delta). \end{aligned}$$

We have thus

$$P_0(t', y; I, x + 4\delta) < \frac{2}{t_0} \epsilon t'$$

uniformly for $|x - y| < \delta, t' \leq t$.

It is certainly unnecessary to give a formal proof of Corollary I.2: the hypotheses of part A of Theorem I are almost trivially verified if the transition probabilities satisfy A_0 , and so the results of part B of Theorem I and Corollary I.1 hold for such processes.

However, a somewhat more detailed discussion is needed concerning the operators

$$T_t f(x) = E\{f(x(t)) \mid x\}$$

defined by a diffusion process. We have developed properties of the resolvent operator R_s , and from these can be deduced properties of the operators T_t by means of standard results from the theory of semi-groups. But to apply such arguments, we must restrict the domain of the operators T_t to be the closure of the range of the resolvent. This space is described in the following lemma; for simplicity, we consider only the case of processes without exceptional points.

LEMMA V.5. *Let $\{\Omega_x, \text{Pr}\{\cdot \mid x\} \mid x \in R\}$ be a diffusion process on the real line R which has no exceptional points. Let C^+ be the space of functions defined and continuous on the extended real line $R^+ = \{x \mid -\infty \leq x \leq \infty\}$.*

Then there exist two regular measures, $P(\cdot \mid \pm \infty)$, on R^+ such that for f in C^+ and $s > 0$,

$$R_s f(\pm \infty) = \int_{R^+} R_s f(y) P(dy \mid \pm \infty).$$

Let C_0^+ be the subspace of C^+ of functions f for which the two relations

$$f(\pm \infty) = \int_{R^+} f(y) P(dy \mid \pm \infty)$$

hold. Then the set $\{R_s f \mid f \in C_0^+, s > 0\}$ is dense in C_0^+ in the uniform norm.

From Lemma V.3, we know that $R_s f$ is in C^+ if f is in C^+ , since there are no exceptional points. Hence in particular, $sR_s f(+\infty) = \lim_{x \rightarrow +\infty} sR_s f(x)$ exists. Now it is easy to verify that

$$\left| sR_s f(+\infty) \right| \leq \text{Max}_{x \in R^+} |f(x)|$$

for f in C^+ . Thus $sR_s f(+\infty)$ defines a positive linear functional on the space of continuous functions on the compact set R^+ ; and by the Riesz representation theorem, there is a regular measure $P_s(\cdot \mid +\infty)$ on R^+ such that for f in C^+ ,

$$sR_s f(+\infty) = \int_{R^+} f(y) P_s(dy \mid +\infty).$$

For each $s > 0$, $P_s(R^+ \mid +\infty) = 1$; Helly's theorem then implies that a regular measure $P(\cdot \mid +\infty)$ exists and a sequence $\{s_n\}$ tending to $+\infty$ such that for each f in C^+ ,

$$(5.9) \quad \lim_{n \rightarrow \infty} s_n R_{s_n} f(+\infty) = \int_{R^+} f(y) P(dy | +\infty).$$

By a simple computation, one may verify the resolvent equation

$$(5.10) \quad R_s f(x) - R_{s'} f(x) = (s' - s) R_{s'} (R_s f)(x)$$

for all f in C^+ , $s, s' > 0$, and x in R^+ . If now for $s > 0$, we set $x = +\infty$, and allow s' to become infinite through the sequence $\{s_n\}$, then (5.9) and the fact that

$$\lim_{s' \rightarrow \infty} R_{s'} f(+\infty) = 0$$

imply

$$\begin{aligned} R_s f(+\infty) &= \lim_{n \rightarrow \infty} (R_s f(+\infty) - R_{s_n} f(+\infty)) \\ &= \lim_{n \rightarrow \infty} (s_n - s) R_{s_n} (R_s f)(+\infty) \\ &= \lim_{n \rightarrow \infty} s_n R_{s_n} (R_s f)(+\infty) \\ &= \int_{R^+} R_s f(y) P(dy | +\infty). \end{aligned}$$

Of course, a similar result holds for the boundary point $-\infty$. In fact, it is easy to see that the same sequence $\{s_n\}$ may be chosen for (5.9) in both cases.

The set $\{R_s f | f \in C_0^+, s > 0\}$ is a linear subspace of C_0^+ , as one can easily verify with the use of (5.10). By a theorem of Riesz, this subspace is dense in C_0^+ if for every regular measure μ on R^+ satisfying

$$\int_{R^+} R_s f(y) \mu(dy) = 0$$

for all f in C_0^+ , $s > 0$, we have

$$\int_{R^+} f(y) \mu(dy) = 0$$

for all f in C_0^+ . Let μ be such a measure, and f a function in C_0^+ . It clearly follows from the continuity properties of the paths in Ω_x , $x \in R$, that

$$\lim_{t \rightarrow 0} E\{f(x(t)) | x\} = f(x);$$

hence

$$\lim_{s \rightarrow \infty} s R_s f(x) = f(x), \quad x \in R.$$

Again, by (5.9) and the preceding paragraph,

$$\begin{aligned} \lim_{n \rightarrow \infty} s_n R_{s_n} f(\pm \infty) &= \int_{R^+} f(y) P(dy | \pm \infty) \\ &= f(\pm \infty), \end{aligned}$$

since f is in C_0^+ . Hence

$$\begin{aligned} 0 &= \lim_{n \rightarrow \infty} s_n \int_{R^+} R_{s_n} f(y) \mu(dy) \\ &= \int_{R^+} \lim_{n \rightarrow \infty} s_n R_{s_n} f(y) \mu(dy) \\ &= \int_{R^+} f(y) \mu(dy); \end{aligned} \qquad \text{q.e.d.}$$

Finally, the techniques used to complete the proof of Corollary I.3 are completely standard. But since they are not at all involved, we repeat them here for the sake of completeness.

LEMMA V.6. *Assuming the hypothesis and notation of Lemma V.5, if f is in the space C_0^+ , then the function*

$$T_t f(x) = E\{f(x(t)) | x\}$$

is in C_0^+ . The operators T_t on C_0^+ form a semi-group, and for f in C_0^+ ,

$$\lim_{t-t' \rightarrow 0, t, t' \geq 0} \text{Max}_{x \in R^+} |T_t f(x) - T_{t'} f(x)| = 0.$$

We prove first the standard result [XII, p. 230] that

$$(5.11) \quad \lim_{t-t' \rightarrow 0, t, t' \geq 0} \text{Max}_{x \in R^+} |E\{R_s f(x(t)) - R_s f(x(t')) | x\}| = 0$$

if f is in C_0^+ and $s > 0$. For by the Markov property

$$\begin{aligned} &E\{R_s f(x(t)) - R_s f(x(t')) | x\} \\ &= \int_0^\infty e^{-s\tau} E\{f(x(t + \tau)) | x\} d\tau - \int_0^\infty e^{-s\tau} E\{f(x(t' + \tau)) | x\} d\tau \\ &= e^{st} \int_t^\infty e^{-s\tau} E\{f(x(\tau)) | x\} d\tau - e^{st'} \int_{t'}^\infty e^{-s\tau} E\{f(x(\tau)) | x\} d\tau \\ &= (e^{st} - e^{st'}) \int_t^\infty e^{-s\tau} E\{f(x(\tau)) | x\} d\tau + e^{st'} \int_t^{t'} e^{-s\tau} E\{f(x(\tau)) | x\} d\tau; \end{aligned}$$

and this expression certainly tends to zero uniformly for x in R as $t-t'$ tends to zero, since f is bounded.

Now suppose $f = R_{s_0} g$ for some g in C_0^+ and $s_0 > 0$. We apply to

$$R_s f(x) = \int_0^\infty e^{-st} E\{f(x(t)) \mid x\} dt$$

the well known inversion formula [e.g. XII, p. 210],

$$(5.12) \quad E\{f(x(t)) \mid x\} = \lim_{s \rightarrow \infty} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n!} n s e^{n s t} R_{n s} f(x),$$

which is valid for all $t > 0$, since $E\{f(x(t)) \mid x\}$ is by (5.11) a continuous function of $t > 0$. In fact

$$\begin{aligned} & \left| \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n!} n s e^{n s t} R_{n s} f(x) - E\{f(x(t)) \mid x\} \right| \\ & \leq s \int_0^\infty e^{s(t-t')} e^{-e^{s(t-t')}} \cdot |E\{f(x(t')) - f(x(t)) \mid x\}| dt' + e^{-e^{s t}} |E\{f(x(t)) \mid x\}| \\ & \leq \text{Max}_{t-t' \leq \delta} |E\{f(x(t')) - f(x(t)) \mid x\}| \\ & \quad + 2 \text{Max}_{x \in R^+} |f(x)| (e^{-e^{s \delta}} + (1 - e^{-e^{-s \delta}}) + e^{-e^{s t}}) \end{aligned}$$

for $\delta > 0$. Now since $f = R_{s_0} g$ for g in C_0^+ , $s_0 > 0$, by (5.11) the first term of the right-hand side above can be made less than an arbitrary positive number ϵ , uniformly for x in R , by choosing δ sufficiently small. Since the second term then is less than ϵ for sufficiently large s , uniformly in x , the result is that (5.12) holds in the sense of uniform convergence on R^+ . But each term in the sum in (5.12) is in C_0^+ and C_0^+ is closed under uniform convergence on R^+ . Hence finally, $E\{f(x(t)) \mid x\} = T_t f(x)$ is in C_0^+ whenever $f = R_{s_0} g$ for some g in C_0^+ and $s_0 > 0$.

Now the functions $\{R_s g \mid g \in C_0^+, s > 0\}$ are dense in C_0^+ , and the operator T_t is continuous on C_0^+ . Thus $T_t f$ is in C_0^+ whenever f is in C_0^+ and $t \geq 0$. The semi-group property for the operators $\{T_t \mid t > 0\}$ is of course equivalent to (2.1). Finally the strong continuity of the semi-group follows from the strong continuity on the dense set $\{R_s g\}$. For given a function f in C_0^+ and $\epsilon > 0$, there is a function g in C_0^+ and $s > 0$ so that

$$\text{Max}_{x \in R^+} |f(x) - R_s g(x)| < \epsilon/3.$$

Hence also

$$\text{Max}_{x \in R^+} |E\{f(x(t)) - R_s g(x(t)) \mid x\}| < \epsilon/3$$

for all $t \geq 0$. By (5.11) there is $\delta > 0$ so that $t, t' \geq 0$, $|t - t'| < \delta$ imply

$$\text{Max}_{x \in R^+} |E\{R_s g(x(t)) - R_s g(x(t')) \mid x\}| < \epsilon/3.$$

These inequalities then imply that for $t, t' \geq 0, |t - t'| < \delta,$

$$\text{Max}_{x \in R^+} | E \{ f(x(t)) - f(x(t')) \mid x \} | < \epsilon,$$

which completes the proof of Lemma V.6 and thus of Corollary I.3.

VI. Construction of diffusion processes. In this final section we give the proof of Theorem II. As before, we divide the proof into a series of lemmas. The first two of these show that the second hypothesis in Theorem II, that the given transition probabilities be of local character uniformly on bounded intervals, is sufficient for a collection of Markov processes $\{ \Omega_x, \text{Pr} \{ \cdot \mid x \} \mid x \in R \}$ with the path spaces described in Definition II to be constructed from them by standard methods.

The remainder of this section is devoted to showing that hypotheses A_0 and B_0 together are likewise sufficient. This proof is considerably more complicated; but it is in outline identical with the first proof, which may serve to clarify the steps of the second.

Again, we precede the proofs proper with some preliminary generalities, in order to reduce to a more manageable form the problem of constructing Markov processes from given transition probabilities. This reduction is well-known, but since the space of the desired process is prescribed, we may use a procedure which is perhaps a little more direct than the general methods developed by Doob [IV; V].

For each x in $R,$ let Ω_x be the space of paths described in Definition II. For each cylinder set, define

$$(6.1) \quad \begin{aligned} & \text{Pr} \{ x(\tau_j) \in E_j, j = 1, \dots, n \mid x \} \\ &= \int_{E_1} \dots \int_{E_n} P(x, \tau_1, dy_1) P(y_1, \tau_2 - \tau_1, dy_2) \dots P(y_{n-1}, \tau_n - \tau_{n-1}, dy_n); \end{aligned}$$

define $\text{Pr} \{ \Lambda \mid x \}$ by additivity for every set Λ which is a finite union of cylinder sets. In order that the definition of $\text{Pr} \{ \cdot \mid x \}$ may be extended to the Borel field generated by the cylinder sets it is necessary and sufficient that whenever $\{ \Lambda_n \}$ is a decreasing sequence of sets, each a finite union of cylinder sets, with $\cap \Lambda_n$ empty, then $\lim_{n \rightarrow \infty} \text{Pr} \{ \Lambda_n \mid x \} = 0$ [XVI]. If this condition is fulfilled, then certainly the desired construction will be complete, since the necessary properties of the measure follow from (6.1).

Now if such a sequence of sets is presented, each set Λ_n is defined by a finite set S_n of parameter values in the sense that if $\omega \in \Lambda_n, x(\tau, \omega) = x(\tau, \omega')$ for all τ in $S_n,$ then $\omega' \in \Lambda_n.$ Let $S = \cup S_n.$ Let $\Omega_x^{(S)}$ denote the space of all paths, with no continuity restriction, from the sequence S to the extended real line. By Kolmogorov's extension theorem [XVI], the assignment (6.1) determines a Borel measure $\text{Pr}^{(S)} \{ \cdot \mid x \}$ on $\Omega_x^{(S)}.$ To each set Λ_n in Ω_x depending only on S corresponds naturally a set $\Lambda_n^{(S)}$ in $\Omega_x^{(S)}$ such that $\text{Pr} \{ \Lambda_n \mid x \} = \text{Pr}^{(S)} \{ \Lambda_n^{(S)} \mid x \};$ moreover, $\cap \Lambda_n$ is empty in Ω_x if and only if no path in

$\cap \Lambda_n^{(S)}$ is the restriction to S of a path in Ω_x . That is, for every path ω in $\cap \Lambda_n^{(S)}$ there are $t_0 > 0, \delta > 0$, and a compact subset F of R such that for every $\Delta > 0$, parameter values t, t' may be found in S with $0 \leq t < t' < t + \Delta < t_0$ for which $x(t, \omega) \in F, |x(t', \omega) - x(t, \omega)| > \delta$.

Thus the assignment (6.1) defined by the transition probabilities determines a process with the path space Ω_x for each x in R if and only if for every sequence S of parameter values, and for every $t_0 > 0, \delta > 0$, and compact subset F of R , the measure $\text{Pr}^{(S)} \{ \cdot | x \}$ satisfies

$$(6.2) \quad \lim_{\Delta \rightarrow 0} \text{Pr}^{(S)} \{ \text{For some } t, t' \text{ in } S, 0 \leq t < t' < t + \Delta < t_0: x(t, \omega) \in F, \\ |x(t', \omega) - x(t, \omega)| > \delta | x \} = 0.$$

This expression is, at least in theory, calculable from the given transition probabilities. We observe finally that (6.2) is just the condition that every separable process constructed from the transition probabilities have paths satisfying Definition II with probability one [IV; V].

For the proof of Theorem II, then, it is sufficient to show that (6.2) is satisfied for a fixed but arbitrary sequence S of parameter values which we may assume is dense in the half line $\{t \geq 0\}$ and is closed under addition and subtraction. Since the description of sets and functions will explicitly involve S , it will cause no confusion to drop the preceding notation, writing $\Omega_x^{(S)} = \Omega_x$ and $\text{Pr}^{(S)} = \text{Pr}$.

LEMMA VI.1. *If given stationary transition probabilities are of local character uniformly on bounded intervals: i.e., if for every $\delta > 0, a > 0$,*

$$\lim_{t \rightarrow 0} \frac{1}{t} P(x, t, I'(x, \delta)) = 0$$

uniformly for $|x| \leq a$, then the measure determined by (6.1) satisfies: for positive numbers a, δ , there corresponds $t_1 > 0$ such that

$$\text{Pr} \left\{ \text{Sup}_{\tau \in S, \tau \leq t} |x(\tau) - x| > \delta | x \right\} \leq 2 \text{Pr} \{ |x(t) - x| > \delta/2 | x \}$$

whenever $t \leq t_1, |x| \leq a$.

Since the set function $\text{Pr} \{ \cdot | x \}$ is a measure on $\Omega_x^{(S)}, x \in R$, and since S is a dense sequence in the half line,

$$\text{Pr} \left\{ \text{Sup}_{\tau \in S, \tau \leq t} |x(\tau) - x| > \delta | x \right\} \\ = \lim_{\Delta \rightarrow 0} \text{Sup}_{\tau_1, \dots, \tau_n \in S, \tau_k - \Delta < \tau_{k-1} < \tau_k \leq t} \text{Pr} \left\{ \text{Sup}_{1 \leq k \leq n} |x(\tau_k) - x| > \delta | x \right\}.$$

Now for every choice of τ_1, \dots, τ_n , setting $\tau_0 = 0$,

$$\begin{aligned} & \Pr \left\{ \text{Sup}_{1 \leq k \leq n} |x(\tau_k) - x| > \delta | x \right\} \\ &= \sum_{k=1}^n \Pr \left\{ \text{Sup}_{0 \leq j < k} |x(\tau_j) - x| \leq \delta, |x(\tau_k) - x| > \delta | x \right\} \\ &= \sum_{k=1}^n \Pr \left\{ \text{Sup}_{0 \leq j < k} |x(\tau_j) - x| \leq \delta, |x(\tau_k) - x| > 2\delta | x \right\} \\ &+ \sum_{k=1}^n \Pr \left\{ \text{Sup}_{0 \leq j < k} |x(\tau_j) - x| \leq \delta, \delta < |x(\tau_k) - x| \leq 2\delta | x \right\}. \end{aligned}$$

The first term equals

$$\begin{aligned} & \sum_{k=1}^n \int_{I(x, \delta)} \Pr \{ |x(\tau_k - \tau_{k-1}) - x| > 2\delta | y \} \\ & \quad \cdot \Pr \left\{ \text{Sup}_{0 \leq j < k} |x(\tau_j) - x| \leq \delta, x(\tau_{k-1}) \in dy | x \right\} \\ & \leq \sum_{k=1}^n \text{Sup}_{y \in I(x, \delta)} \Pr \{ |x(\tau_k - \tau_{k-1}) - y| > \delta | y \} \\ & \leq \sum_{k=1}^n (\tau_k - \tau_{k-1}) \text{Sup}_{y \in I(x, \delta)} \text{Sup}_{0 < \tau \leq \Delta} \frac{1}{\tau} P(y, \tau, I'(y, \delta)) \\ & \leq t \text{Sup}_{y \in I(x, \delta)} \text{Sup}_{0 < \tau \leq \Delta} \frac{1}{\tau} P(y, \tau, I'(y, \delta)), \end{aligned}$$

if $\tau_k - \Delta < \tau_{k-1} < \tau_k \leq t, k = 1, \dots, n$; and by hypothesis the last expression tends to zero with Δ . We find a bound for the second term above by writing

$$\begin{aligned} & \Pr \{ |x(t) - x| > \delta/2 | x \} \\ & \geq \Pr \left\{ |x(t) - x| > \delta/2; \text{ For some } k = 1, \dots, n: \right. \\ & \quad \left. \text{Sup}_{0 \leq j < k} |x(\tau_j) - x| \leq \delta, \delta < |x(\tau_k) - x| \leq 2\delta | x \right\} \\ & = \sum_{k=1}^n \int_{\delta < |x-y| \leq 2\delta} \Pr \{ |x(t - \tau_k) - x| > \delta/2 | y \} \\ & \quad \cdot \Pr \left\{ \text{Sup}_{0 \leq j < k} |x(\tau_j) - x| \leq \delta, x(\tau_k) \in dy | x \right\} \\ & \geq \sum_{k=1}^n \Pr \left\{ \text{Sup}_{0 \leq j < k} |x(\tau_j) - x| \leq \delta, \delta < |x(\tau_k) - x| \leq 2\delta | x \right\} \\ & \quad \text{Inf}_{\delta < |x-y| \leq 2\delta} \text{Inf}_{0 \leq \tau \leq t} \Pr \{ |x(\tau) - y| \leq \delta/2 | y \}. \end{aligned}$$

Now because of the hypothesis of uniform local character, we may certainly choose t_1 so that $t \leq t_1$ implies

$$\inf_{\delta < |x-y| \leq 2\delta} \inf_{0 \leq \tau \leq t} \Pr \{ |x(\tau) - y| \leq \delta/2 \mid y \} \geq \frac{1}{2}.$$

In this case

$$\begin{aligned} \sum_{k=1}^n \Pr \left\{ \sup_{0 \leq j < k} |x(\tau_j) - x| \leq \delta, \delta < |x(\tau_k) - x| \leq 2\delta \mid x \right\} \\ \leq 2 \Pr \{ |x(t) - x| > \delta/2 \mid x \}. \end{aligned}$$

Since the two bounds are independent of the particular choice of τ_1, \dots, τ_n , they yield, together with the relation at the beginning of the proof, the inequality VI.1.

LEMMA VI.2. *If the given transition probabilities are of local character uniformly on bounded intervals, then for positive numbers t_0, a, δ , and for x in R*

$$\lim_{\Delta \rightarrow 0} \Pr \{ \text{For some } t, t' \text{ in } S, 0 \leq t < t' < t + \Delta < t_0:$$

$$|x(t)| \leq a, |x(t') - x(t)| > \delta \mid x \} = 0.$$

Again we write the expression in VI.2 as

$$\begin{aligned} \lim_{\Delta \rightarrow 0} \sup_{\tau_1, \dots, \tau_n \in S, \tau_k - \Delta < \tau_{k-1} < \tau_k < t_0} \Pr \left\{ \text{For some } k = 1, \dots, n: |x(\tau_k)| \leq a; \right. \\ \left. \sup_{\tau_k < \tau \leq \tau_{k+1}, \tau \in S} |x(\tau) - x(\tau_k)| > \delta \mid x \right\}. \end{aligned}$$

For each choice of τ_1, \dots, τ_n , we have

$$\begin{aligned} \Pr \left\{ \text{For some } k = 1, \dots, n: |x(\tau_k)| \leq a; \sup_{\tau_k < \tau \leq \tau_{k+1}, \tau \in S} |x(\tau) - x(\tau_k)| > \delta \mid x \right\} \\ \leq \sum_{k=1}^{n-1} \int_{|y| \leq a} \Pr \left\{ \sup_{0 \leq \tau \leq \tau_{k+1} - \tau_k, \tau \in S} |x(\tau) - y| > \delta \mid y \right\} \Pr \{ x(\tau_k) \in dy \mid x \} \\ \leq 2 \sum_{k=1}^{n-1} (\tau_{k+1} - \tau_k) \sup_{|y| \leq a} \frac{1}{\tau_{k+1} - \tau_k} P(y, \tau_{k+1} - \tau_k, I'(y, \delta/2)) \\ \leq 2t_0 \sup_{0 \leq \tau \leq \Delta} \sup_{|y| \leq a} \frac{1}{\tau} P(y, \tau, I'(y, \delta/2)). \end{aligned}$$

Hence the hypothesis of local character uniformly for $|y| \leq a$ implies VI.2.

Finally, in view of (6.2), Lemma VI.2 implies the result of Theorem II under the hypothesis B. The proof has been arranged so that the steps may correspond to those of the remainder of the section, where we assume through-

out that stationary transition probabilities are given satisfying:

A₀: For every bounded continuous function f on R and for every $t > 0$, $\int f(y)P(x, t, dy)$ is a continuous function of x in R ; and

B₀: For every x in R and $\delta > 0$,

$$\lim_{t \rightarrow 0} \frac{1}{t} P(x, t, I'(x, \delta)) = 0.$$

LEMMA VI.3. *Suppose given transition probabilities, satisfying A₀ and B₀, are used to construct a measure $\Pr\{\cdot | x\}$ on the space $\Omega_x = \Omega_x^{(S)}$, according to (6.1). Then for positive numbers a, b, δ , there corresponds $t_1 > 0$ such that*

$$\Pr \{ \text{For some } \tau \text{ in } S, 0 \leq \tau \leq t: |x(\tau) - x| > \delta, |x(\tau)| \leq b |x\} \\ \leq 2 \Pr \{ |x(t) - x| > \delta/2 | x\}$$

for $t \leq t_1, |x| \leq a$.

For let x be a point in R ; let f be a continuous function on R such that $0 \leq f \leq 1, f(y) = 0$ for $|x - y| \leq 2\delta/3$, and $f(y) = 1$ for $|x - y| \geq 5\delta/6$. The Laplace transform

$$R_s f(y) = \int_0^\infty e^{-st} \int f(z)P(y, t, dz)dt$$

is for each $s > 0$ a bounded continuous function on R and satisfies

$$\lim_{s \rightarrow \infty} sR_s f(y) = f(y), \quad y \in R,$$

$$\lim_{t \rightarrow 0} \text{Max}_{y \in R} \left| \int R_s f(z)P(y, t, dz) - R_s f(y) \right| = 0.$$

The first equation above clearly follows from the local character condition; the latter equation is a standard result from the theory of semi-groups which we have already used (cf. (5.11)). From the first of these equations and the continuity of $R_s f$ it follows [XI] that given $b > 0$, there are positive numbers s_1, \dots, s_n and c_1, \dots, c_n , with $\sum c_k = 1$, such that

$$\text{Max}_{|y| \leq b} \left| \sum_{k=1}^n c_k s_k R_{s_k} f(y) - f(y) \right| < \frac{1}{6}.$$

Denoting the function $\sum c_k s_k R_{s_k} f$ by g , we have for $|y| \leq b$

$$\Pr \{ |x(t) - x| \geq 2\delta/3 | y\} \geq \int g(z)P(y, t, dz) - \frac{1}{6}.$$

But since for each $s > 0, \int R_s f(z)P(y, t, dz)$ tends to $R_s f(y)$ uniformly for y in R as t tends to 0, the same is true of g . Hence there is $t_1(x) > 0$ such that

the paths of the process are continuous with probability one at $t=0$, except for the possibility of extremely large jumps. Note that if (6.3) held for all y in R with $|x-y| \geq \delta$, then the latter possibility would be excluded. This is one hypothesis and its result as given by Kinney [XV]; the proof of VI.3 occurs also in a paper by Ambrøse [I]. That is, the probability that the path has a large jump can be bounded in terms of the probability of return. In general it seems impossible to bound this probability without assuming a condition on the probability of return, such as Kinney's, or on the uniformity of local character.

But before resolving this difficulty, we must develop the condition B_0 of local character into a more powerful tool by the use of the regularity condition A_0 . Namely, for each positive integer k , for each x in R , and for $\delta' > 0$, the hypothesis B_0 implies the existence of a parameter value τ in S , with $0 < \tau < 2^{-k}$, such that $P(x, \tau, I'(x, \delta'/3)) < 2^{-k}\tau$. Again let f be a continuous function on R with $0 \leq f \leq 1$, $f(y) = 0$ for $|y-x| \leq \delta'/3$, $f(y) = 1$ for $|y-x| \geq 2\delta'/3$. Then $\int f(z)P(x, \tau, dz) \leq P(x, \tau, I'(x, \delta'/3)) < 2^{-k}\tau$; and since under hypothesis A_0 the left side is a continuous function of x , $P(y, \tau, I'(y, \delta')) \leq P(y, \tau, I'(x, 2\delta'/3)) \leq \int f(z)P(x, \tau, dz) < 2^{-k}\tau$ for y in some neighborhood U of x . Since x is arbitrary, there are a finite number of points x for which the associated neighborhoods cover the closed interval $[-b, b]$. Thus

LEMMA VI.4. *Under the hypothesis of Lemma VI.3, for positive numbers b, δ' and for a positive integer k , there is a finite valued function $\tau^{(k)}(x)$ on $[-b, b]$ such that*

$$P(x, \tau^{(k)}(x), I'(x, \delta')) < 2^{-k}\tau^{(k)}(x),$$

$$\tau^{(k)}(x) \in S, \quad 0 < \tau^{(k)}(x) < 2^{-k}$$

for $|x| \leq b$.

We set $\Delta_k = \inf_{|x| \leq b} \tau^{(k)}(x)$, and $\tau^{(k)}(x) = \Delta_k$ for $|x| > b$. For each x in R and each path ω in Ω_x , define by induction

$$s_0^{(k)}(\omega) = 0,$$

$$s_{n+1}^{(k)}(\omega) = s_n^{(k)}(\omega) + \tau^{(k)}(x(s_n^{(k)}(\omega), \omega)).$$

With these tools we can now prove that the process we are constructing has paths which are almost all continuous at $t=0$:

LEMMA VI.5. *If the given transition probabilities satisfy conditions A_0 and B_0 , then for positive numbers a, δ there is $t_1 > 0$ such that*

$$\Pr \left\{ \sup_{\tau \in S, \tau \leq t} |x(\tau) - x| > \delta \mid x \right\} \leq 3 \Pr \left\{ |x(t) - x| > \frac{\delta}{2} \mid x \right\}$$

whenever $|x| \leq a$ and $t \leq t_1$.

For take $b = a + 2\delta$ and choose t_1 according to VI.3. Then for each positive integer k_0 , using the preceding construction, we have

$$\begin{aligned}
 & \Pr \left\{ \text{Sup}_{\tau \in S, \tau \leq t} |x(\tau) - x| > \delta | x \right\} \\
 (6.4) \quad & \leq \Pr \left\{ \text{For some } \tau \leq t \text{ in } S: |x(\tau) - x| > \delta, |x(\tau)| \leq b | x \right\} \\
 & + \Pr \left\{ \text{For integers } k > k_0, n > 0; s_n^{(k)}(\omega) \leq t, \right. \\
 & \qquad \qquad \qquad \left. |x(s_{n-1}^{(k)}) - x| \leq \delta, |x(s_n^{(k)})| > b | x \right\} \\
 & + \Pr \left\{ \text{Sup}_{\tau \in S, \tau < t} |x(\tau)| > b; \text{Sup}_{k > k_0, s_n^{(k)} \leq t} |x(s_n^{(k)}) - x| \leq \delta | x \right\} \\
 & + \Pr \left\{ |x(t)| > b | x \right\}.
 \end{aligned}$$

By VI.3 the first term on the right is less than $2 \Pr \{ |x(t) - x| > (\delta/2) | x \}$, if $t \leq t_1$, $|x| \leq a$. Likewise the last term is certainly less than $\Pr \{ |x(t) - x| > \delta/2 | x \}$ because of the choice of b . Denoting the values of $\tau^{(k)}(x)$ by τ_1, \dots, τ_μ , the second term is not greater than

$$\begin{aligned}
 & \sum_{k=k_0+1}^{\infty} \sum_{n=1}^{\infty} \Pr \left\{ |x(s_{n-1}^{(k)}) - x| \leq \delta, s_{n-1}^{(k)} + \tau^{(k)}(x(s_{n-1}^{(k)})) \leq t; \right. \\
 & \qquad \qquad \qquad \left. |x(s_n^{(k)})| > b | x \right\} \\
 & \leq \sum_{k=k_0+1}^{\infty} \sum_{n=1}^{\infty} \sum_{\nu=1}^{\mu} \int_{\tau^{(k)}(y)=\tau_\nu} \Pr \left\{ |x(s_{n-1}^{(k)}) - x| \leq \delta, s_{n-1}^{(k)} + \tau_\nu \leq t, \right. \\
 & \qquad \qquad \qquad \left. x(s_{n-1}^{(k)}) \in dy | x \right\} \Pr \left\{ |x(\tau_\nu) - y| \geq \delta | y \right\} \\
 & \leq \sum_{k=k_0+1}^{\infty} \sum_{n=1}^{\infty} \sum_{\nu=1}^{\mu} 2^{-k} \tau_\nu \Pr \left\{ |x(s_{n-1}^{(k)}) - x| \leq \delta, \right. \\
 & \qquad \qquad \qquad \left. \tau^{(k)}(x(s_{n-1}^{(k)})) = \tau_\nu, s_{n-1}^{(k)} + \tau_\nu \leq t | x \right\} \\
 & \leq \sum_{k=k_0+1}^{\infty} \sum_{n=1}^{\infty} 2^{-k} E \left\{ \tau^{(k)}(x(s_{n-1}^{(k)})); |x(s_{n-1}^{(k)}) - x| \leq \delta, s_n^{(k)} \leq t | x \right\} \\
 & \leq \sum_{k=k_0+1}^{\infty} 2^{-k} \sum_{n=1}^{\infty} E \left\{ s_n^{(k)} - s_{n-1}^{(k)}; s_n^{(k)} \leq t | x \right\} \\
 & \leq \sum_{k=k_0+1}^{\infty} 2^{-k} E \left\{ \sum_{s_n^{(k)} \leq t} s_n^{(k)} - s_{n-1}^{(k)} | x \right\} \\
 & \leq \sum_{k=k_0+1}^{\infty} 2^{-k} t = 2^{-k_0} t.
 \end{aligned}$$

Finally we show that the third term in (6.4) is zero. This term is less than the sum

$$\sum_{\tau \in S, \tau < t} \Pr \left\{ |x(\tau)| > b; \sup_{k > k_0, s_n^{(k)} \leq t} |x(s_n^{(k)}) - x| \leq \delta |x \right\},$$

each term of which will be seen to vanish. For if $\tau \in S, \tau < t$, the corresponding term is less than

$$\begin{aligned} \Pr \{ |x(\tau)| > b; \text{For some } \tau' < \epsilon \text{ in } S: |x(\tau + \tau') - x| \leq \delta |x \} \\ \leq \Pr \{ |x(\tau)| > b; \text{For some } \tau' < \epsilon \text{ in } S: \\ |x(\tau + \tau') - x(\tau)| > \delta, |x(\tau + \tau')| \leq b |x \} \\ \leq \int \Pr \{ \text{For some } \tau' < \epsilon \text{ in } S: |x(\tau') - y| > \delta, \\ |x(\tau')| \leq b |y \} P(x, \tau, dy) \end{aligned}$$

for every $\epsilon > 0$, since by construction the sequence $\{s_n^{(k)} | k > k_0, n > 0\}$ is dense in the interval (τ, t) , and since S is closed under addition and subtraction. Now VI.3 and B_0 clearly imply that for each y in R ,

$$\lim_{\epsilon \rightarrow 0} \Pr \{ \text{For some } \tau' < \epsilon \text{ in } S: |x(\tau') - y| > \delta, |x(\tau')| \leq b |y \} = 0;$$

and so

$$\lim_{\epsilon \rightarrow 0} \Pr \{ |x(\tau)| > b; \text{For some } \tau' < \epsilon \text{ in } S: |x(\tau + \tau') - x| \leq \delta |x \} = 0,$$

implying that the third term in (6.4) vanishes.

Combining the bounds for the terms in (6.4),

$$\Pr \left\{ \sup_{\tau \in S, \tau \leq t} |x(\tau) - x| > \delta |x \right\} \leq 3 \Pr \{ |x(t) - x| > \delta/2 |x \} + 2^{-k_0 t}.$$

Since k_0 is arbitrary, VI.5 follows.

It seems very difficult to prove VI.5 without using the hypothesis that for every $t > 0, \Pr \{x(t) \in R | x\} = 1$. For this allows the result of Lemma VI.3 to be applied to the third term in (6.4); and we have no other techniques available to treat the third term. Now the use we have made of this hypothesis is essentially as a condition on return: intuitively, the probability of return from large finite values of the path is bounded by VI.3, while the probability of return from $\pm \infty$ is zero, since the path has these values on the sequence S only with probability zero. Whether other techniques will yield VI.5 without such restrictions is not apparent.

We are not ready to apply VI.5 to the proof of (6.2), in analogy to Lemma VI.2. Two more tools are needed before the final arguments can be made.

LEMMA VI.6. *Under the same hypotheses, if $t_0 > 0$, $a > 0$, $\delta > 0$, and $x \in R$, then*

$$\Pr \left\{ \text{For some } t \leq t_0 \text{ in } S: |x(t)| \leq a, \limsup_{\tau \searrow 0, \tau \in S} |x(t + \tau) - x(t)| > \delta | x \right\} = 0.$$

For the left side is not greater than

$$\begin{aligned} & \sum_{t \in S, t \leq t_0} \Pr \left\{ |x(t)| \leq a; \text{ For every } \Delta > 0, \text{ Sup}_{\tau \in S, \tau \leq \Delta} |x(t + \tau) - x(t)| > \delta | x \right\} \\ &= \sum_{t \in S, t \leq t_0} \lim_{\Delta \searrow 0} \Pr \left\{ |x(t)| \leq a, \text{ Sup}_{\tau \in S, \tau \leq \Delta} |x(t + \tau) - x(t)| > \delta | x \right\} \\ &= \sum_{t \in S, t \leq t_0} \int_{|y| \leq a} \lim_{\Delta \searrow 0} \Pr \left\{ \text{Sup}_{\tau \in S, \tau \leq \Delta} |x(\tau) - y| > \delta | y \right\} P(x, t, dy) \\ &= 0, \end{aligned}$$

since by VI.5 and B_0 , the integrand in each of the countably many terms is zero.

LEMMA VI.7. *Let a, t_0, δ be given positive numbers. For each x in R , and for each path ω in Ω_x , define $N(\omega; x, t_0)$ to be the largest number N of pairs $0 \leq \tau_1 < \tau'_1 < \tau_2 < \dots < \tau_N < \tau'_N \leq t_0$ which can be found in S satisfying $|x(\tau_k)| \leq a, |x(\tau'_k) - x(\tau_k)| > \delta, k = 1, \dots, N$. If no such parameter values exist, $N(\omega; x, t_0) = 0$. That is, $N(\omega; x, t_0)$ is the number of displacements on $S \cap [0, t_0]$ of a distance more than δ from points in $[-a, a]$. Then*

$$\Pr \{ N(\omega; x, t_0) < \infty | x \} = 1.$$

It is clear that if $\omega \in \Omega_x, t, t' \in S, N(\omega; x, t + t') \leq N(\omega; x, t) + N(\omega; x(t, \omega), t') + 1$. In view of this it is sufficient to prove that for some $t > 0$, and for every x in $R, N(\omega; x, t)$ is finite with probability one. For if so, then for every $t_0 > 0, x$ in $R, \Pr \{ N(\omega; x, t_0) = \infty | x \} \leq \sum_{nt < t_0} \int \Pr \{ N(\omega; y, t) = \infty | y \} \cdot P(x, nt, dy) = 0$.

To this effect we observe that

$$(6.5) \quad \lim_{t \searrow 0} \text{Sup}_{|z| \leq a} \Pr \left\{ \text{Sup}_{\tau \in S, \tau \leq t} |x(\tau) - x| > \delta | x \right\} = 0.$$

For if not, there is $\epsilon > 0$, a sequence of points $\{x_n\}$ in $[-a, a]$, and a sequence $\{t_n\}$ of parameter values with $t_n \searrow 0$, such that

$$\Pr \left\{ \text{Sup}_{\tau \in S, \tau \leq t_n} |x(\tau) - x_n| > \delta | x_n \right\} \geq \epsilon.$$

Since the interval $[-a, a]$ is compact, we may assume that there is a point x_0 to which $\{x_n\}$ converges. Let f be a continuous function on R such that $0 \leq f \leq 1$, $f(y) = 1$ for $|x_0 - y| \geq \delta/4$, $f(x_0) = 0$. Since $E\{f(x(t)) | y\}$ is a continuous function of y in R and since $\lim_{t \rightarrow 0} E\{f(x(t)) | x_0\} = 0$, certainly there is t such that $E\{f(x(t)) | x_n\} < \epsilon/3$ for n sufficiently large. But using VI.5,

$$\begin{aligned} \Pr \left\{ \text{Sup}_{\tau \in S, \tau \leq t_n} |x(\tau) - x_n| > \delta | x_n \right\} &\leq \Pr \left\{ \text{Sup}_{\tau \in S, \tau \leq t} |x(\tau) - x_n| > \delta | x_n \right\} \\ &\leq 3 \Pr \left\{ |x(t) - x_n| > \delta/2 | x_n \right\} \\ &\leq 3E\{f(x(t)) | x_n\} < \epsilon \end{aligned}$$

if n is so large that $|x_0 - x_n| < \delta/4$, $t_n \leq t \leq t_1$. The contradiction establishes (6.5).

Now choose t so small that

$$\text{Sup}_{|x| \leq a} \Pr \left\{ \text{Sup}_{\tau \in S, \tau \leq t} |x(\tau) - x| > \delta | x \right\} \leq 1/2.$$

Let $0 \leq \tau_1 < \dots < \tau_n \leq t$ be an arbitrary finite set of parameter values in $S \cap [0, t]$. Define $N'(\omega; x, \tau_1, \dots, \tau_n)$ to be the number of displacements of the path ω in Ω_x of a distance more than δ from a point in $[-a, a]$ on the set $\{\tau_1, \dots, \tau_n\}$. Then whether or not $|x| \leq a$,

$$\begin{aligned} \Pr \{N'(\omega; x, \tau_1, \dots, \tau_n) \geq 1 | x\} &= \Pr \left\{ \text{For some integer } j < n: |x(\tau_j)| \leq a, \text{Max}_{j < i \leq n} |x(\tau_i) - x(\tau_j)| > \delta | x \right\} \\ &\leq \sum_{j=1}^n \Pr \left\{ |x(\tau_i)| > a, i < j; |x(\tau_j)| \leq a; \right. \\ &\qquad \qquad \qquad \left. \text{Sup}_{\tau \in S, \tau_j < \tau \leq t} |x(\tau) - x(\tau_j)| > \delta | x \right\} \\ &= \sum_{j=1}^n \int_{|y| \leq a} \Pr \left\{ |x(\tau_i)| > a, i < j; x(\tau_j) \in dy | x \right\} \\ &\qquad \qquad \qquad \cdot \Pr \left\{ \text{Sup}_{\tau \in S, \tau \leq t - \tau_j} |x(\tau) - y| > \delta | y \right\} \\ &\leq 1/2 \sum_{j=1}^n \Pr \left\{ |x(\tau_i)| > a, i < j; |x(\tau_j)| \leq a | x \right\} \\ &\leq 1/2; \end{aligned}$$

and this bound is independent of the set τ_1, \dots, τ_n . If $k > 1$,

$$\begin{aligned}
 & \Pr \{ N'(\omega; x, \tau_1, \dots, \tau_n) \geq k \mid x \} \\
 &= \Pr \{ \text{For some integer } j < n: N'(\omega; x, \tau_1, \dots, \tau_j) = k - 1; \\
 & \qquad \qquad \qquad N'(\omega; x(\tau_j), \tau_{j+1}, \dots, \tau_n) \geq 1 \mid x \} \\
 &= \sum_{j=1}^n \int \Pr \{ N'(\omega; x, \tau_1, \dots, \tau_i) < k - 1, i < j; \\
 & \qquad \qquad \qquad N'(\omega; x, \tau_1, \dots, \tau_j) = k - 1, x(\tau_j) \in dy \mid x \} \\
 & \quad \cdot \Pr \{ N'(\omega; y, \tau_{j+1} - \tau_j, \dots, \tau_n - \tau_j) \geq 1 \mid y \} \\
 & \leq 1/2 \Pr \{ N'(\omega; x, \tau_1, \dots, \tau_n) \geq k - 1 \mid x \}.
 \end{aligned}$$

These relations imply

$$\Pr \{ N'(\omega; x, \tau_1, \dots, \tau_n) \geq k \mid x \} \leq 2^{-k}, k \geq 1.$$

But again since S is countable, the least upper bound of

$$\Pr \{ N'(\omega; x, \tau_1, \dots, \tau_n) \geq k \mid x \}$$

for all finite sets $0 \leq \tau_1 < \dots < \tau_n \leq t$ of parameter values in $S \cap [0, t]$ is equal to $\Pr \{ N(\omega; x, t) \geq k \mid x \}$; and thus the above bound holds for the latter expression. Hence $N(\omega; x, t)$ is finite with probability one for sufficiently small t , and as we observed previously, this implies the same statement for all t .

Combining the last four lemmas, we can now verify that (6.2) holds, and that hence a diffusion process can be constructed from the given transition probabilities.

LEMMA VI.8. *If stationary transition probabilities are given satisfying hypotheses A_0 and B_0 , then the measure determined by (6.1) satisfies, for all $t_0 > 0, \delta > 0, a > 0$,*

$$\lim_{\Delta \rightarrow 0} \Pr \{ \text{For some } t, t' \text{ in } S, 0 \leq t < t' < t + \Delta < t_0: \\
 |x(t)| \leq a, |x(t') - x(t)| > \delta \mid x \} = 0.$$

Let $\{s_n^{(k)}\}$ be the sequence constructed in Lemma VI.4 for $b = a + 2\delta, \delta' = \delta/3$. Given $\epsilon > 0$ we apply the preceding lemmas to prove first that a positive integer k_0 may be chosen so that for $k > k_0$, the probability is less than $\epsilon/2$ that there are an integer n with $s_n^{(k)} \leq t_0, |x(s_n^{(k)})| > a + \delta, a$ t in S with $s_n^{(k)} < t < s_{n+1}^{(k)}, |x(t)| \leq a$, and a τ in S with $t < \tau \leq s_{n+1}^{(k)}, |x(\tau) - x(t)| > 2\delta/3$.

For suppose, given ω , that there is a set k of quadruples (k, n, t, τ) , consisting of infinitely many integers k , to each corresponding an integer n with $s_n^{(k)}(\omega) \leq t_0, |x(s_n^{(k)})| > a + \delta, a$ value t in S with $s_n^{(k)} < t < s_{n+1}^{(k)}, |x(t)| \leq a$, and a value τ in S with $t < \tau \leq s_{n+1}^{(k)}, |x(\tau) - x(t)| > 2\delta/3$. Because of VI.6, unless ω is in a fixed set of probability zero, for each (k, n, t, τ) in K , there is $\eta_k > 0$ such that $\text{Sup}_{\tau \in S, \tau \leq \eta_k} |x(t + \tau) - x(t)| \leq \delta/3$. We may clearly assume for the

given set K that $k' > k$ implies $2^{-k'} < \eta_k$. We may also assume that the set K satisfies either: $k' > k$ implies $s_n^{(k')} > t$; or: $k' > k$ implies $s_n^{(k')} < t$. In the first case, the intervals $[t, s_n^{(k)}]$, where $(k, n, t, \tau), (k', n', t', \tau')$ are quadruples in K with k' the successor of k , are disjoint, and these pairs $t, s_n^{(k)}$ of parameter values define infinitely many displacements of the path ω of a distance more than $\delta/3$ from a point in $[-a, a]$. In the second case, since clearly then $k' > k$ implies $s_{n+1}^{(k')} < t + \eta_k$, the intervals $[t + \eta_k, \tau]$ corresponding to distinct integers k are disjoint, and these likewise define infinitely many displacements of the same type. But the probability that the path ω has infinitely many such displacements is zero, by VI.7; thus the probability is zero that the path has the property described in the preceding paragraph for infinitely many integers k . Hence k_0 may be chosen so that for $k > k_0$, the probability of this event is less than $\epsilon/2$.

Now let k be a sufficiently large positive integer. We have for $\Delta < \Delta_k$, as given in VI.4,

$$\begin{aligned} & \Pr \left\{ \text{For some } t, t' \text{ in } S, 0 \leq t < t' < t + \Delta < t_0: \right. \\ & \qquad \qquad \qquad \left. |x(t)| \leq a, |x(t') - x(t)| > \delta |x| \right\} \\ & \leq \Pr \left\{ \text{For some } t \leq t_0 \text{ in } S: |x(t)| \leq a, \right. \\ & \qquad \qquad \qquad \left. \sup_{\tau \in S, \tau \leq \Delta} |x(t + \tau) - x(t)| > \delta |x| \right\} \\ (6.6) \quad & \leq \Pr \left\{ \text{For some integer } n, s_n^{(k)} \leq t_0: |x(s_n^{(k)})| \leq a + 2\delta, \right. \\ & \qquad \qquad \qquad \left. \sup_{\tau \in S, s_n^{(k)} < \tau \leq s_{n+1}^{(k)}} |x(s_n^{(k)}) - x(\tau)| > \delta/3 |x| \right\} \\ & + \Pr \left\{ \text{For some integer } n, s_n^{(k)} \leq t_0: |x(s_n^{(k)})| > a + \delta; \right. \\ & \qquad \qquad \qquad \left. \text{for some } t \text{ in } S, s_n^{(k)} < t < s_{n+1}^{(k)}: |x(t)| \leq a, \right. \\ & \qquad \qquad \qquad \left. \sup_{\tau \in S, t < \tau \leq s_{n+1}^{(k)}} |x(\tau) - x(t)| > 2\delta/3 |x| \right\}. \end{aligned}$$

For suppose $|x(t)| \leq a, t \in S$, and $\sup_{\tau \in S, \tau \leq \Delta} |x(t + \tau) - x(t)| > \delta$; let n be such that $s_n^{(k)} \leq t < s_{n+1}^{(k)}$. If $|x(s_n^{(k)})| \leq a + \delta$, then either

$$\sup_{\gamma \in S, s_n^{(k)} < \gamma \leq s_{n+1}^{(k)}} |x(s_n^{(k)}) - x(\gamma)| > \delta/3$$

or $|x(s_{n+1}^{(k)})| \leq a + 2\delta$,

$$\sup_{\gamma \in S, s_{n+1}^{(k)} \leq \gamma \leq s_{n+2}^{(k)}} |x(s_{n+1}^{(k)}) - x(\gamma)| > \delta/3.$$

If on the other hand, $|x(s_n^{(k)})| > a + \delta$, then either

$$\text{Sup}_{\nu \in S, t < \tau_{s_{n+1}^{(k)}}} |x(\tau) - x(t)| > 2\delta/3,$$

or $|x(s_{n+1}^{(k)})| \leq a + 2\delta$,

$$\text{Sup}_{\tau \in S, s_{n+1}^{(k)} < \tau \leq s_{n+1}^{(k)}} |x(s_{n+1}^{(k)}) - x(\tau)| > \delta/3.$$

The last term in (6.6) is less than $\epsilon/2$, because of the paragraph before last. As for the first term

$$\begin{aligned} & \Pr \left\{ \text{For some integer } n; s_n^{(k)} \leq t_0, |x(s_n^{(k)})| \leq a + 2\delta, \right. \\ & \qquad \qquad \qquad \left. \text{Sup}_{\tau \in S, s_n^{(k)} < \tau \leq s_{n+1}^{(k)}} |x(s_n^{(k)}) - x(\tau)| > \delta/3 \mid x \right\} \\ & \leq \sum_{n=0}^{\infty} \sum_{\nu=1}^{\mu} \int_{\tau^{(k)}(y)=\tau_{\nu}} \Pr \left\{ |x(s_n^{(k)})| \leq a + 2\delta, s_n^{(k)} \leq t_0, x(s_n^{(k)}) \in dy \mid x \right\} \\ & \qquad \Pr \left\{ \text{Sup}_{\tau \in S, \tau \leq \tau_{\nu}} |x(\tau) - y| > \delta/3 \mid y \right\} \\ & \leq \sum_{n=0}^{\infty} \sum_{\nu=1}^{\mu} \int_{\tau^{(k)}(y)=\tau_{\nu}} \Pr \left\{ |x(s_n^{(k)})| \leq a + 2\delta, s_n^{(k)} \leq t_0, x(s_n^{(k)}) \in dy \mid x \right\} \\ & \qquad 3 \Pr \left\{ |x(\tau_{\nu}) - y| > \delta/6 \mid y \right\} \\ & \leq \sum_{n=0}^{\infty} \sum_{\nu=1}^{\mu} \int_{\tau^{(k)}(y)=\tau_{\nu}} 3 \cdot 2^{-k} \tau_{\nu} \Pr \left\{ |x(s_n^{(k)})| \leq a + 2\delta, \right. \\ & \qquad \qquad \qquad \left. s_n^{(k)} \leq t_0, x(s_n^{(k)}) \in dy \mid x \right\} \\ & \leq \sum_{n=0}^{\infty} 3 \cdot 2^{-k} E \left\{ \tau^{(k)}(x(s_n^{(k)})); |x(s_n^{(k)})| \leq a + 2\delta, s_n^{(k)} \leq t_0 \mid x \right\} \\ & \leq 3 \cdot 2^{-k} \sum_{n=0}^{\infty} E \left\{ s_{n+1}^{(k)} - s_n^{(k)}; s_n^{(k)} \leq t_0 \mid x \right\} \\ & \leq 3 \cdot 2^{-k} (t_0 + 2^{-k}) \\ & < \epsilon/2, \end{aligned}$$

if k is sufficiently large. Thus (6.2) is verified, and the proof of Theorem II is finally complete.

REFERENCES

I. W. Ambrose, *On measurable stochastic processes*, Trans. Amer. Math. Soc. vol. 47 (1940) pp. 66-79.

II. D. A. Darling and A. J. F. Siegert, *The first passage problem for a continuous Markoff process*, Ann. Math. Statist. vol. 24 (1953) pp. 624-639.

III. W. Doblin, *Sur l'équation de Kolmogoroff*, C. R. Acad. Sci. Paris vol. 207 (1938) pp. 705-707.

IV. J. L. Doob, *Stochastic processes depending on a continuous parameter*, Trans. Amer. Math. Soc. vol. 42 (1937) pp. 107-140.

V. ———, *Stochastic processes*, New York, 1953.

VI. ———, *Semimartingales and subharmonic functions*, Trans. Amer. Math. Soc. vol. 77 (1954) pp. 86-121.

VII. W. Feller, *Diffusion processes in one dimension*, Trans. Amer. Math. Soc. vol. 77 (1954) pp. 1-31.

VIII. ———, *The general diffusion operator and positivity preserving semi-groups in one dimension*, Ann. of Math. vol. 60 (1954) pp. 417-436.

IX. ———, *On second order differential operators*, Ann. of Math. vol. 61 (1955) pp. 90-105.

X. R. Fortet, *Les fonctions aléatoires du type de Markoff associées à certaines équations linéaires aux dérivées partielles du type paraboliques*, J. Math. Pures Appl. vol. 22 (1943) pp. 177-243.

XI. D. C. Gillespie and W. A. Hurwitz, *On sequences of continuous functions having continuous limits*, Trans. Amer. Math. Soc. vol. 32 (1930) pp. 527-543.

XII. E. Hille, *Semi-groups and functional analysis*, American Mathematical Society Colloquium Publications 31, New York, 1948.

XIII. G. A. Hunt, *Some theorems concerning Brownian motion*, Trans. Amer. Math. Soc. vol. 81 (1956) pp. 294-319.

XIV. M. Kac, *On some connections between probability theory and differential and integral equations*, Proc. Second Berkeley Symposium on Mathematical Statistics and Probability, 1951, pp. 189-215.

XV. J. R. Kinney, *Continuity properties of sample functions of Markov processes*, Trans. Amer. Math. Soc. vol. 74 (1953) pp. 280-302.

XVI. A. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Ergebnisse der Math. vol. 2, Berlin, 1933.

XVII. P. Lévy, *Systèmes markoviens et stationnaires*, Ann. École Norm. (3) vol. 68 (1951) pp. 327-381.

XVIII. E. B. Dynkin, *Criteria of continuity for trajectories of a Markov random process*, Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya vol. 16 (1952) pp. 563-572.

XIX. ———, *Infinitesimal operators of Markov processes*, Doklady Akademii Nauk SSSR. vol. 105 (1955) pp. 206-209.

XX. ———, *Continuous stationary Markov processes*, Doklady Akademii Nauk SSSR. vol. 105 (1955) pp. 405-408.

PRINCETON UNIVERSITY,
PRINCETON, N. J.