# AN EXISTENCE THEOREM FOR SECOND ORDER PARABOLIC EQUATIONS(1)

BY

EDWARD NELSON

0. The Existence Theorem. *Let $a^{ij}$, $b^i$ for $1 \leqq i,\ j \leqq n$ be real valued functions of class $C^3$ defined on a connected open set $X$ in real Euclidean n-space, such that for each $x = (x_1, \cdots, x_n)$ in $X$, $a^{ij}(x)$ is a positive definite symmetric matrix. Let $B(X)$ be the Banach space of all real-valued bounded measurable functions $f$ on $X$, with the norm $\|f\| = \sup_{x \in X} |f(x)|$. Then there exists a family of linear transformations $P^t$ on $B(X)$, $0 \leqq t < \infty$, such that*

$$P^t P^s = P^{t+s}, \tag{1}$$

$$\|P^t f\| \leqq \|f\|, \qquad f \in B(X), \tag{2}$$

$$P^t f \geqq 0 \quad \text{if } f \geqq 0, \qquad f \in B(X) \tag{3}$$

*and such that whenever $f$ is of class $C^2$ vanishing outside a compact subset of $X$,*

$$\lim_{t \downarrow 0} t^{-1}(P^t f(x) - f(x)) = \sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2 f(x)}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b^i(x) \frac{\partial f(x)}{\partial x_i} \tag{4}$$

*uniformly for $x$ in $X$.*

Similar theorems have been proved under various restrictions on the behavior of the operator

$$A = \sum_{i,j=1}^{n} a^{ij} \frac{\partial^2}{\partial x_i \partial x_i} + \sum_{i=1}^{n} b^i \frac{\partial}{\partial x_i} \tag{5}$$

at the boundary. If $A$ is formally self-adjoint or if it is semibounded, Hilbert space methods have been used successfully in attacking the problem, e.g. [12; 15]. If $X$ is taken to be a compact manifold rather than a domain in Euclidean space, Yosida [14] has solved the problem by means of his theory of semigroups, and also in various cases when $X$ is not compact, e.g. [16]. Cases in which the behavior of $a^{ij}$ near infinity is restricted have been discussed also by Feller [3], Dressel [2], and S. Itô [6].

Our method relies heavily on the probabilistic interpretation of the Equation (4). As there is no added difficulty, we treat the case when $X$ is a differ-

Received October 19, 1956.

(1) This is in essence a portion of the author's thesis *On the operator theory of Markoff processes*, The University of Chicago, 1955. The author held a National Science Foundation Predoctoral Fellowship, and is greatly indebted to Professor I. E. Segal, under whose direction the thesis was written.

entiable manifold. It is to be hoped that there is a purely analytic method for solving the problem in sufficient generality. However, one compensation to the probabilist for the artificiality of our method is that in the process of proving the theorem we derive various interesting properties of the sample functions of the Markoff processes associated with the operator $A$.

The first section is entirely self-contained and probabilistic in content.

1. **Continuity of sample functions of stochastic processes.** Let $X$, $\mathcal{R}$ be a measurable space; i.e., a set $X$ with a $\sigma$-ring $\mathcal{R}$ of subsets of $X$. Let $\Omega$, $\mathcal{S}$, Pr be a probability space; i.e., a measurable space $\Omega$, $\mathcal{S}$ such that $\Omega \in \mathcal{S}$ and a measure Pr such that $\Pr(\Omega)=1$. By an $X$-valued *measurable function* on $\Omega$ we shall mean a function $\xi(\omega)$ defined for all $\omega \in \Omega$ and taking values in $X$ such that $\{\omega: \xi(\omega) \in E\}$ is in $\mathcal{S}$ whenever $E$ is in $\mathcal{R}$. By an $X$-valued *random variable* on $\Omega$ we shall mean an equivalence class of $X$-valued measurable functions, two such functions $\xi(\omega)$ and $\xi'(\omega)$ being considered equivalent whenever $\xi(\omega)=\xi'(\omega)$ a.e. Throughout this paper we will use the notation $\xi(\omega)$ for a measurable function. The symbol $\xi$ with no argument indicated will denote the corresponding random variable, the equivalence class of measurable function equal a.e. to $\xi(\omega)$.

This terminology differs from the standard language of probability theory according to which "random variable" is a synonym for "measurable function." We wish to put the emphasis on what we have called random variables for two reasons. Firstly, they are objects of greater intrinsic probabilistic interest. Secondly, nondenumberable families of them can be manipulated more easily than corresponding families of measurable functions. If $X$ is the extended real line (containing $-\infty$ and $\infty$) and $\mathcal{R}$ is the Borel sets of $X$, then an $X$-valued measurable function (respectively, random variable) will be called an *extended real-valued measurable function* (respectively, *random variable*). Both extended real-valued measurable functions and random variables form lattices under the formation of the maximum and minimum of two functions or random variables. The following lemma is known: (cf. [11]).

LEMMA 1. *The lattice $L$ of all extended real-valued random variables on a probability space $\Omega$, $\mathcal{S}$,* Pr *is complete. Any family $F$ contained in $L$ has a countable subfamily $F_0$ with the same least upper bound.*

**Proof.** By means of the transformation $\xi \rightarrow \pi^{-1}$ (arc tan $\xi$) we may assume that for all $\xi \in F$, ess sup $|\xi| \leqq 1$. Let $F'$ be the family of all least upper bounds of finite subfamilies of $F$. Let $\alpha$ be the least upper bound of the numbers $\int \xi d \Pr$, $\xi \in F'$. Then $-1 \leqq \alpha \leqq 1$. Since the real numbers are separable, there is an increasing sequence $\xi_n$ of elements of $F'$ such that $\text{l.u.b.}_n \int \xi_n d \Pr = \alpha$. For each $n$, choose an element $\xi_n(\omega)$ of the equivalence class $\xi_n$, let $\xi_\infty(\omega) = \text{l.u.b.}_n\, \xi_n(\omega)$, and let $\xi_\infty$ be the equivalence class of $\xi_\infty(\omega)$. Then if $\eta$ is an upper bound for $F$, and hence $F'$, then $\eta \geqq \xi_n$ for each $n$ and so $\eta \geqq \xi_\infty$. Conversely, suppose that $\xi$ in $F'$ were such that $\xi_\infty \geqq \xi$ did not hold. Then, if

$\xi(\omega)$ is any element of the equivalence class $\xi$, on some set $E$ of positive measure, $\xi(\omega) > \xi_\infty(\omega)$. Then l.u.b.$_n \int \max\{\xi_n(\omega), \xi(\omega)\} d\Pr(\omega) > \alpha$, which is a contradiction. Hence the lattice $L$ is complete. To obtain the countable family $F_0$ with the same least upper bound, notice that each $\xi_n$ is the least upper bound of a finite subfamily of $F$. Then we let $F_0$ consist of all elements of $F$ occurring in the representation of some $\xi_n$.

The lattice of all extended real-valued measurable functions is in general not complete, since the least upper bound of a nondenumerable family of measurable functions may fail to be measurable.

THEOREM 1. *Let $X$, $\mathcal{R}$ be a measurable space, $\rho$ a metric on $X$ such that $X$ is a complete metric space and all open sets are in $\mathcal{R}$. Let $\Omega$, $\mathcal{S}$, $\Pr$ be a probability space, $I$ a closed interval of the real line, $\xi_t(\omega)$ an $X$-valued measurable function on $\Omega$ for each $t \in I$. Then there exist $X$-valued measurable functions $\xi_t'(\omega)$ such that for each $t$ in $I$, $\xi_t'(\omega) = \xi_t(\omega)$ a.e. and $\xi_t'(\omega)$ is for each $\omega$ a continuous function from $I$ to $X$, if and only if the random variable defined by*

$$\eta = \underset{\epsilon > 0}{\text{g.l.b.}} \ \underset{|t-s|<\epsilon;\, t,s\in I}{\text{l.u.b.}} \ \rho(\xi_t, \xi_s) \tag{6}$$

*is equal to* 0.

REMARKS. We will call a family of $X$-valued random variables $\xi_t$, $t \in I$, on $\Omega$ on $X$-valued *stochastic process*. When, as in the theorem, the $\xi_t(\omega)$ can be changed on null sets to become a continuous function of $t$ we say that the stochastic process *has continuous sample functions*. In general there is no null set which will work for all $t$, but this is of no probabilistic interest.

**Proof.** For any family of random variables there is a countable subset with the same least upper bound. Hence for each integer $m$ there is a countable subset $K(m)$ of $I$ such that

$$\underset{|t-s|<1/m;\, t,s\in K(m)}{\text{l.u.b.}} \ \rho(\xi_t, \xi_s) = \underset{|t-s|<1/m;\, t,s\in I}{\text{l.u.b.}} \ \rho(\xi_t, \xi_s).$$

Let $K = \bigcup_m K(m) \cup Q$, where $Q$ is the set of rational points in $I$. Then $\eta = \text{g.l.b.}_m \, \text{l.u.b.}_{|t-s|<1/m;\, t,s\in K} \, \rho(\xi_t, \xi_s)$ since in the definition of $\eta$ the set of all $\epsilon > 0$ can be replaced by the set of all $1/m$.

Since $K$ is countable $\eta$ is represented by the measurable function $\eta(\omega) = \text{g.l.b.}_m \, \text{l.u.b.}_{|t-s|<1/m;\, t,s\in K} \, \rho(\xi_t(\omega), \rho_s(\omega))$. But $\eta(\omega) = 0$ a.e. if and only if the mapping $t \to \xi_t(\omega)$ is a uniformly continuous function from $K$ to $X$ for all $\omega$ except those in a null set $N$. Thus if $\eta(\omega) = 0$ a.e. we may define $\xi_t'(\omega)$ for each $\omega$ not in $N$ to be the unique (since $K$ was chosen dense in $I$) continuous extension from $K$ to $I$ of this mapping into the complete metric space $X$. For the $\omega$ in $N$, $\xi_t'(\omega)$ may be defined to have some constant value. Then $\xi_t'(\omega)$ is a measurable function of $\omega$, $\xi_t'(\omega) = \xi_t(\omega)$ a.e. for each $t$, and $\xi_t'(\omega)$ is a continuous function of $t$ for each $\omega$. Conversely, if such $\xi_t'(\omega)$ exist then the mapping $t \to \xi_t(\omega)$ must be uniformly continuous for $t$ in any countable set $K$ for a.e. $\omega$, so that $\eta(\omega) = 0$ a.e.

COROLLARY 1. *If the interval $I$ is infinite then the process has continuous sample functions if and only if $\eta$ defined as in the theorem for each compact subinterval $I'$ of $I$ is equal to* 0.

This follows immediately from the facts that the real line is the union of countably many compact intervals and that the union of countably many null sets is a null set.

We will need to apply Theorem 1 when the space $\Omega$ has certain special features.

COROLLARY 2. *Let $X$, $\rho$, $\Omega$, and* Pr *be as in Theorem* 1 *and assume further that $\Omega = \prod_{0 \leqq t < \infty} X$ is the Cartesian product space of all functions from $[0, \infty)$ to $X$ and that $\xi_t(\omega) = \omega(t)$ for $0 \leqq t < \infty$. Let $\mathcal{S}$ be the $\sigma$-ring generated by the $\xi_t$. Then* (6) *holds if and only if the $\sigma$-ring $\mathcal{S}$ and the measure* Pr *can be extended to a $\sigma$-ring $\mathcal{S}'$ and measure* Pr$'$ *in such a way that the set $\Delta$ of all $\omega$ in $\Omega$ which are continuous functions of $t$ is an element of $\mathcal{S}'$ and* $\Pr'(\Delta) = 1$.

**Proof.** The necessity of (6) follows immediately from Theorem 1. To prove the sufficiency of (6), let $\Gamma$ be any measurable subset of $\Omega$ containing $\Delta$. We need only show that $\Pr(\Gamma) = 1$. Since $\Gamma$ is in $\mathcal{S}$, there is a countable subset $K$ of $[0, \infty)$ such that $\Gamma$ is in the $\sigma$-ring generated by the $\xi_t$ with $t$ in $K$. Now $\omega$ is continuous if and only if

$$\operatorname*{g.l.b.}_{\epsilon > 0} \quad \operatorname*{l.u.b.}_{|t-s| < \epsilon;\, t,s \in I} \rho(\xi_t(\omega), \xi_s(\omega)) = 0$$

for each compact interval $I$ contained in $[0, \infty)$. Since $\Gamma$ contains the set of continuous functions and since it is measurable with respect to the $\xi_t$ with $t$ in $K$, $\Gamma$ contains

$$\left\{\omega : \operatorname*{g.l.b.}_{\epsilon > 0} \quad \operatorname*{l.u.b.}_{|t-s| < \epsilon;\, t,s \in I \cap K} \rho(\xi_t(\omega), \xi_s(\omega)) = 0\right\}.$$

Therefore, by Theorem 1, $\Pr(\Gamma) = 1$.

Let the $\xi_t$, $t \in I$, be an $X$-valued stochastic process. We say that they constitute a *Markoff process* in case

$$\Pr\left\{\xi_t \in E \,\middle|\, \xi_{t_1}, \cdots, \xi_{t_n}\right\} = \Pr\left\{\xi_t \in E \,\middle|\, \xi_{t_n}\right\} \tag{7}$$

whenever $t_1 < \cdots < t_n < t$ are in $I$ and $E \in \mathcal{R}$. The notation in (7) denotes the random variable which is the equivalence class of all functions which are the Radon-Nikodym derivatives occurring in the indicated conditional probabilities, cf. [1]. If $p(s, x; t, E)$ defined whenever $s < t$ are in $I$, $x \in X$, $E \in \mathcal{R}$ is such that each $p(s, x; t, \cdot)$ is a probability measure on $X$, $\mathcal{R}$ and each $p(s, \cdot; t, E)$ is a measurable function on $X$, and if

$$\Pr\left\{\xi_t \in E \,\middle|\, \xi_s\right\} = p(s, \xi_s; t, E), \qquad s < t \tag{8}$$

then $p(s, x; t, E)$ is called a *transition function* of the Markoff process.

THEOREM 2. *If $X$, $\mathcal{R}$, and $\rho$ are as in Theorem 1 and if the $\xi_t$ are the $X$-valued random variables of a Markoff process with transition function $p(s, x; t, E)$, then the process has continuous sample functions if for each $\beta > 0$,*

$$\lim_{t \downarrow s} (t - s)^{-1} p(s, x; t, cN_\beta(x)) = 0$$

*uniformly for $x$ in $X$ and $s$ in the closed interval $I$. (Here $N_\beta$ denotes $\beta$-neighborhood, $c$ denotes complement.)*

**Proof.** Given $\beta > 0$, $\epsilon > 0$, let $\delta > 0$ be such that for all $x$ in $X$, $s$ in $I$, we have $p(s, x; s+h, cN_\beta(x)) < h\epsilon$, for all $0 < h < \delta$. Let $s = h_0 < h_1 < \cdots < h_{k-1} < h_k = s + \delta$. Let

$$A = \left\{ \omega : \underset{0 \leqq i \leqq k}{\text{l.u.b.}}\ \rho(\xi_s(\omega), \xi_{h_i}(\omega)) \geqq 2\beta \right\},$$

$$B = \{\omega: \rho(\xi_s(\omega), \xi_{s+\delta}(\omega)) \geqq \beta\},$$

$$C_i = \{\omega: \rho(\xi_{h_i}(\omega), \xi_{s+\delta}(\omega)) \geqq \beta\},$$

$$D_i = \{\omega: \rho(\xi_s(\omega), \xi_{h_i}(\omega)) \geqq 2\beta\} \cap \bigcap_{j<i} \{\omega: \rho(\xi_s(\omega), \xi_{h_j}(\omega)) < 2\beta\}.$$

Then $\Pr(A) \leqq \Pr(B) + \sum_{i=1}^{k} \Pr(C_i \cap D_i)$, for if $\rho(\xi_s(\omega), \xi_{h_i}(\omega)) \geqq 2\beta$, then by the triangle inequality either $\rho(\xi_s(\omega), \xi_{s+\delta}(\omega)) \geqq \beta$ or $\rho(\xi_{h_i}(\omega), \xi_{s+\delta}(\omega)) \geqq \beta$. Therefore, since the $D_i$ are disjoint, we have

$$\Pr(A) \leqq \Pr(B) + \sum_{i=1}^{k} \sup_x p(h_i, x; s + \delta, cN_\beta(x)) \Pr(D_i)$$

$$\leqq \delta\epsilon + \delta\epsilon \sum_{i=1}^{k} \Pr(D_i) \leqq 2\delta\epsilon.$$

Since this is true for all finite sets of $h_i$ in the interval $[s, s+\delta]$ it follows that

$$\Pr\left( \underset{s \leqq h \leqq s+\delta}{\text{l.u.b.}}\ \rho(\xi_s, \xi_h) \geqq 2\beta \right) \leqq 2\delta\epsilon. \tag{9}$$

(That is, if we choose any representative $\theta$ of the equivalence class in question then $\Pr(\{\omega: \theta(\omega) \geqq 2\beta\}) \leqq 2\delta\epsilon$. We shall use notation of this type without further explanation in what follows.)

If $I = [a, b]$, consider the points $a + r(\delta/2)$, for all integers $r$ in the interval $[0, 2(b-a)/\delta]$. If $|t - s| < \delta/2$ there exists an $r$ such that $0 \leqq |t - (a + r(\delta/2))| < \delta/2$, $0 \leqq |s - (a + r(\delta/2))| < \delta/2$. It follows that

$$\Pr\left( \underset{|t-s|<\delta/2}{\text{l.u.b.}}\ \rho(\xi_t, \xi_s) \geqq 4\beta \right) \leqq \sum_r \Pr\left( \underset{h<\delta}{\text{l.u.b.}}\ \rho(\xi_{a+r(\delta/2)}, \xi_h) \geqq 2\beta \right)$$

$$\leqq \sum_r 2\delta\epsilon \leqq (2(b - a)/\delta) 2\delta\epsilon = 4(b - a)\epsilon.$$

That is, changing notation, for each $\beta > 0$, $\lim_{\delta\to 0}\Pr(\text{l.u.b.}_{|t-s|<\delta}\ \rho(\xi_t, \xi_s) \geqq \beta) = 0$, and from this it follows that $\text{g.l.b.}_{\delta>0}\ \text{l.u.b.}_{|t-s|<\delta}\ \rho(\xi_t, \xi_s) = 0$. By Theorem 1 the process has continuous sample functions.

COROLLARY. *If $U$ is an open subset of $X$, $K$ a compact subset of $U$, then*

$$\Pr\left\{\xi_{s+h} \notin U \text{ for some } h \in [0, t] \,\middle|\, \xi_s = x\right\} = o(t)$$

*uniformly for $x$ in $K$.*

**Proof.** The set $K$ is at a positive distance from $cU$, so this follows from the inequality (9).

We are now in a position to prove that the Markoff processes associated with parabolic differential equations have continuous sample functions.

THEOREM 3. *Let $X$ be a second countable differentiable manifold of class $C^k$, $\tilde{p}(s, x; t, E)$ the transition function of a Markoff process on the one point compactification $\tilde{X}$ of $X$, $A_s$ a family of differential operators with no zero order terms such that for all functions $f$ of class $C^k$ with compact support in $X$,*

$$\lim_{t\downarrow s} (t - s)^{-1}\left(\int f(y)\tilde{p}(s, x; t, dy) - f(x)\right) = A_s f(x)$$

*uniformly for $x$ in $\tilde{X}$ and $s$ in $I$, for all compact intervals $I$. Then the process has continuous sample functions on $\tilde{X}$.*

REMARKS. The point at infinity plays a double role in this situation. In the first place, it allows for the possibility of absorption. For instance, consider the process associated with $d^2/dx^2$ on $(0, 1)$ with boundary condition of absorption at 0 and 1. Then strictly speaking this is not a Markoff process on $(0, 1)$ itself since after a time the diffusing particle will be absorbed. We may make it into a bona-fide Markoff process by compactifying $(0, 1)$ by the point $\infty$ and saying that the particle is at $\infty$ when it has been absorbed. In the second place, even if we have an actual Markoff process on $X$ itself, we cannot expect the sample functions to be continuous in $X$ itself since they may reach the boundary at times. For instance, consider the process associated with $d^2/dx^2$ on $(0, 1)$ with reflecting barriers at 0 and 1. Then the sample functions will pass through both 0 and 1 infinitely often. Thus the general situation is for the sample functions to be continuous in $X$ so long as they are in $X$, and to converge to $\infty$ whenever $t$ converges to a value for which the value of the sample function is not in $X$. We state this more simply by saying that the sample functions are continuous on the one point compactification $\tilde{X}$ of $X$.

**Proof.** Since $X$ is assumed second countable $\tilde{X}$ is a second countable compact Hausdorff space, and so is metrizable. Let $\rho$ be a metric giving the topology of $\tilde{X}$. We wish to show that for all compact intervals $I$ and $\beta > 0$, $\lim_{t\downarrow s} (t-s)^{-1}\tilde{p}(s, x; t, cN_\beta(x)) = 0$ uniformly for $x$ in $\tilde{X}$ and $s$ in $I$.

Given $x$ in $\tilde{X}$, let $f$ be a positive function of class $C^k$ which is 0 on a neighborhood $U(x)$ of $x$ such that the closure $\overline{U}(x)$ of $U(x)$ is contained in $N_{\beta/2}(x)$, 1 on $cN_{\beta/2}(x)$, and such that either $f=0$ in a neighborhood of $\infty$ or $f=1$ in a neighborhood of $\infty$. The former may be achieved if $\infty \in \overline{U}(x)$, the latter if $\infty \notin \overline{U}(x)$. Then either $f$ has compact support in $X$ or is of the form $1-g$ where $g$ has compact support in $X$. In the first case $\lim_{t\downarrow s} (t-s)^{-1}(\int f(y)\tilde{p}(s, z; t, dy) - f(z)) = A_s f(z)$ and in the second case the same expression is equal to $A_s g(z)$, uniformly for $z$ in $\tilde{X}$ and $s$ in $I$ in both cases. In either case the expression is 0 in $U(x)$ since $f$ and $g$ are constant in $U(x)$ and $A_s$ has no zero order term. Since $f=0$ in $U(x)$, $\lim_{t\downarrow s} (t-s)^{-1}\int f(y)\tilde{p}(s, z; t, dy) = 0$ uniformly for $z$ in $U(x)$ and $s$ in $I$. But by the triangle inequality $\chi_{cN_\beta(z)} \leqq f$ for all $z$ in $U(x)$. (For if $\rho(w, z) \geqq \beta$, $\rho(z, x) \leqq \beta/2$ then $\rho(w, x) \geqq \beta/2$ and so $f(w)=1$.) Hence $\lim_{t\downarrow s} (t-s)^{-1}\tilde{p}(s, z; t, cN_\beta(z)) = 0$ uniformly for $z$ in $U(x)$ and $s$ in $I$. The theorem follows from Theorem 2 by taking a finite covering of $\tilde{X}$ by sets of the form $U(x)$.

Let us remark that since this proves the hypothesis of Theorem 2 the corollary to that theorem holds in this case, too.

Notice that the theorem makes no assumptions about the class of $X$ or regularity of the operators $A_s$. The only properties of $A_s$ that we have used are that if $f$ is constant in an open set $U$ and in the domain of $A_s$ then $A_s f = 0$ in $U$, and that the domain of $A_s$ contains sufficiently fine functions so that the Urysohn lemma is valid for functions in the domain of $A_s$. For this type of operator Theorem 3 is true for any second countable locally compact Hausdorff space. Heuristically also this is the reason for sample function continuity. The operators $A_s$ govern the transitions for infinitely short time: since they are local a particle can have come from only an infinitely short distance in an infinitely short time, and so travels continuously.

With our formulation of the problem, however, the assumption that the operators $A_s$ have no zero order term cannot be dropped. For let $A_s = A = -1$, $\tilde{p}(s, x; t, E) = e^{-(t-s)}\chi_E(x) + (1-e^{-(t-s)})\chi_E(\infty)$ on the one point compactification of the real line. A particle starting at $x$ in the associated Markoff process remains there for an exponential holding time and then jumps to $\infty$, where it remains. Therefore the sample functions of this Markoff process are not continuous.

Results quite similar to Theorems 2 and 3 have been obtained by Ray [9] and Kinney [7].

2. **Semigroups and sample functions.** Let $X$ be a locally compact Hausdorff space, $B(X)$ the Banach space of all real-valued bounded Borel-measurable functions, with the norm $\|f\| = \sup_{x\in X} |f(x)|$. Let $P^t$, $0 \leqq t < \infty$, be a family of operators on $B(X)$ satisfying (1), (2), and (3). Let $C(X)$ be the subset of $B(X)$ consisting of the continuous functions vanishing at infinity. For each $x$ in $X$, the mapping $f \to P^t f(x)$ for $f$ in $C(X)$ is a continuous, positive linear functional. By the Riesz-Markoff theorem, there is a regular positive

measure $p^t(x, \cdot)$ such that for all $f$ in $C(X)$,

$$P^t f(x) = \int f(y) p^t(x, dy). \tag{10}$$

By (2), $p^t(x, X) \leqq 1$. The semigroups $P^t$ we are going to construct will have the property that (10) holds for all $f$ in $B(X)$. We assume this to be true in the remainder of this section.

The function $p^t(x, \cdot)$ is called a *sub-transition function*. In case $p^t(x, X) = 1$ for all $t$ and $x$, it is called a *transition function*. In either case, the semigroup property (1) of the $P^t$ implies that for all Borel sets $E$,

$$p^{t+s}(x, E) = \int p^t(y, E) p^s(x, dy). \tag{11}$$

Let $\tilde{X}$ be the one-point compactification of $X$ by the point $\infty$. (In case $X$ itself is compact, $\infty$ is an isolated point adjoined to $X$.) If $p^t$ is a sub-transition function on $X$, let

$$\begin{aligned} \tilde{p}^t(x, E) &= p^t(x, E), && x \neq \infty, \ \infty \notin E, \\ \tilde{p}^t(x, \{\infty\}) &= 1 - p^t(x, X), && x \neq \infty, \\ \tilde{p}^t(\infty, E) &= \chi_E(\infty) \end{aligned} \tag{12}$$

where $\chi_E$ denotes the characteristic function of the set $E$. Then $\tilde{p}^t$ will be called the *transition function associated with* the sub-transition function $p^t$. The terminology is justified by the following lemma:

LEMMA 2. *The transition function $\tilde{p}^t$ associated with a sub-transition function $p^t$ is a transition function.*

**Proof.** Clearly, for each $x$ in $\tilde{X}$, $\tilde{p}^t(x, \cdot)$ is a regular positive measure on $\tilde{X}$ and $\tilde{p}^t(x, \tilde{X}) = 1$. It remains to verify (11) for the $\tilde{p}^t$. If $x \neq \infty$ and $\infty \notin E$, this is (11) for the $p^t$. If $x \neq \infty$, then

$$\begin{aligned} \int_{\tilde{X}} \tilde{p}^t(y, \{\infty\}) \tilde{p}^s(x, dy) &= \int_X (1 - p^t(y, X)) p^s(x, dy) + \tilde{p}^t(\infty, \{\infty\}) \tilde{p}^s(x, \{\infty\}) \\ &= p^s(x, X) - p^{t+s}(x, X) + 1 - p^s(x, X) = \tilde{p}^{t+s}(x, \{\infty\}). \end{aligned}$$

Finally, $\int_{\tilde{X}} \tilde{p}^t(y, E) \tilde{p}^s(\infty, dy) = \tilde{p}^t(\infty, E) = \chi_E(\infty) = \tilde{p}^{t+s}(\infty, E)$, concluding the proof.

If the $\xi_t$, $0 \leqq t < \infty$, on the probability space $\Omega$, $\mathcal{S}$, Pr are the random variables of an $\tilde{X}$-valued Markoff process with transition function $\tilde{p}(s, x; t, E) = \tilde{p}^{t-s}(x, E)$, we shall also call $\tilde{p}^t(x, E)$ its *transition function*. By a familiar technique due to Kolmogoroff [8], we may show that in the present case there actually is an $\tilde{X}$-valued Markoff process corresponding to a given sub-transition function on $X$.

LEMMA 3. *Let $\tilde{X}$ be a compact Hausdorff space, $\Omega = \prod_{0 \leq t < \infty} \tilde{X}$ the Cartesian product space of all functions from $[0, \infty)$ to $\tilde{X}$, $\mathcal{R}$ the ring generated by all sets of the form*

$$E(t) = \{\omega : \omega(t) \in E\} \tag{13}$$

*where $E$ is a Borel set in $\tilde{X}$. Let $\mathcal{R}(t_1, \cdots, t_n)$ be the subring of $\mathcal{R}$ generated by sets of the form $E(t_i)$, $i = 1, \cdots, n$. Finally, let* Pr *be a set function defined on $R$ such that the restriction of* Pr *to $\mathcal{R}(t_1, \cdots, t_n)$ is a regular probability measure for each n-tuple $t_1, \cdots, t_n$. Then* Pr *is a probability measure on $\mathcal{R}$, and so has an extension to a probability measure on the $\sigma$-ring $\mathcal{S}$ generated by $\mathcal{R}$.*

**Proof.** The proof in [4, §49, Theorem A] for the case of $\tilde{X}$ the unit interval holds without change for the general case.

COROLLARY. *If $\tilde{p}^t$ is a transition function on a compact Hausdorff space $\tilde{X}$, then there is a probability space $\Omega$, $\mathcal{S}$,* Pr *and an $\tilde{X}$-valued Markoff process $\xi_t$ having $\tilde{p}^t$ as its transition function.*

**Proof.** Let $\Omega$, $\mathcal{S}$, $\mathcal{R}(t_1, \cdots, t_n)$ be as above. Let $\nu$ be a probability measure on $\tilde{X}$. Let $E_1, \cdots, E_n$ be Borel sets in $\tilde{X}$, let $0 \leq t_1 < \cdots < t_n$, and define

$$\begin{aligned} &\Pr_\nu(E_1(t_1) \cap \cdots \cap E_n(t_n)) \\ &\quad = \int_{\tilde{X}} d\nu(x_0) \int_{E_1} p^{t_1}(x_0, dx_1) \cdots \int_{E_n} p^{t_n - t_{n-1}}(x_{n-1}, dx_n). \end{aligned} \tag{14}$$

This extends to a set function $\Pr_\nu$ on $R$ satisfying the hypotheses of Lemma 3. If we define $\xi_t(\omega) = \omega(t)$, then it is easily verified (cf. [1]) that the $\xi_t$ on $\Omega$, $\mathcal{S}$, $\Pr_\nu$ are a Markoff process with transition function $\tilde{p}^t$.

3. **Local diffusion.** Throughout the remainder of this paper, $X$ will be a second-countable differentiable manifold of class $C^3$, $A$ a differential operator such that for each $x$ in $X$ there is a coordinate system of class $C^3$ near $x$ with respect to which $A$ has the form (5), with the $a^{ij}$, $b^i$ of class $C^3$ and $a^{ij}$ a positive definite matrix. A semigroup of operators $P^t$ on the space $B(X)$ of bounded real-valued measurable functions satisfying (1), (2), (3), (4), and (10) will be called a *diffusion semigroup* or *A-diffusion semigroup* on $X$. A Markoff process on the one-point compactification $\tilde{X}$ of $X$ with transition function $\tilde{p}^t$ as in §2 will be called a *diffusion process* or *A-diffusion process*. We wish to show, then, that corresponding to any such operator $A$ there is an $A$-diffusion semigroup. In this section, we solve the problem locally, for coordinate neighborhoods, by means of the Hille-Yosida semigroup theory [5; 13]. Then it is shown that a diffusion process may be restricted to an open subset, and that if two of the $A$-diffusion processes already constructed are restricted to a common open set they agree. In the next section these local diffusions are pieced together to give a diffusion in the large.

THEOREM 4. *Let $B$ be the open unit sphere in Euclidean $n$-space, and let $A$ be given by* (5) *on an open set containing $\overline{B}$, such that $a^{ij}$, $b^i$ are of class $C^3$ and $a^{ij}$ is positive definite. Let the domain $D$ of $A$ consist of all functions $f$ in the Banach space $C(B)$ of all real-valued continuous functions vanishing on the boundary which can be extended to be of class $C^2$ on an open set containing $\overline{B}$ and such that $Af$ is in $C(B)$. Then the closure $\overline{A}$ of $A$ as an unbounded operator on $C(B)$ exists and is the infinitesimal generator of a semigroup of bounded operators $P^t$ converging strongly to the identity as $t$ approaches* 0.

LEMMA 4. *If $f \in D$ and $\lambda > 0$ then*

$$\sup_{x \in B} (f(x) - \lambda Af(x)) \geqq \sup_{x \in B} f(x),$$

$$\inf_{x \in B} (f(x) - \lambda Af(x)) \leqq \inf_{x \in B} f(x).$$

**Proof.** If $f$ has its maximum at $y$ then, because of the fact that $A$ is elliptic and has no zero order term, $Af(y) \leqq 0$ and $f(y) - \lambda Af(y) \geqq f(y) = \sup f(x)$. Similarly, if $f$ has its minimum at $z$ then $Af(z) \geqq 0$, and therefore $f(z) - \lambda Af(z) \leqq f(z) = \inf f(x)$. If $f$ has no maximum in $B$ then $f < 0$ on $B$, since $f$ vanishes on the boundary of $B$. Similarly, if $f$ has no minimum in $B$ then $f > 0$ on $B$. In either case the desired relation holds because $f - \lambda Af$ vanishes on the boundary of $B$.

LEMMA 5. *The closure $\overline{A}$ of $A$ exists.*

**Proof.** If the $f_n$ are in $D$ and $f_n$, $Af_n$ converge in $C(B)$ to 0, $h$ respectively we need to show that $h = 0$. Then we may unambiguously define $\overline{A}f$ to be $\lim Af_n$ whenever this limit exists with $\lim f_n = f$.

Let $k$ be a function of class $C^2$ with compact support in $B$, and let $dx$ denote $n$-dimensional Lebesgue measure. Then

$$\int_B h(x)k(x)dx = \lim_{n \to \infty} \int_B Af_n(x)k(x)dx = \lim_{n \to \infty} \int_B f_n(x)A'k(x)dx = 0.$$

Since this holds for all such $k$, $h = 0$. Here $A'$ is given by

$$A'f(x) = \sum_{i,j=1}^{n} (\partial^2/\partial x_i \partial x_j)(a^{ij}(x)f(x)) - \sum_{i=1}^{n} (\partial/\partial x_i)(b^i(x)f(x)).$$

LEMMA 6. *The range of $1 - \lambda A$ is dense in $C(B)$ for all $\lambda > 0$.*

**Proof.** Let $g$ be a function of class $C^3$ and compact support in $B$. Then if $f$ is a solution of class $C^3$ of the equation $(1 - \lambda A)f = g$ in an open set containing $\overline{B}$ it will have values of class $C^3$ on the boundary of $B$. In [10, p. 277] it is shown under weaker hypotheses that there is a solution $h$ of $(1 - \lambda A)h = 0$ such that $h = f$ on the boundary of $B$ and such that the first and second partial derivatives of $h$ are continuous on $\overline{B}$. Then $f - h$ is a function in the domain $D$ of $A$ such that $(1-\lambda A)(f-h)=g$. But the functions $g$ are dense in $C(B)$.

**Proof of Theorem** 4. By Yosida's theory of semigroups [13] these three lemmas imply that $\bar{A}$ is the infinitesimal generator of a semigroup of order preserving contraction operators $P^t$ on $C(B)$ such that $P^t$ converges strongly to the identity at $t=0$.

The proof on the whole follows Yosida's proof in [14] of a similar theorem for $B$ a compact manifold.

In what follows we shall always assume that the sample functions $\xi_t(\omega)$ of a diffusion process are chosen to be continuous, by Theorem 3.

THEOREM 5. *If $P^t$ is an $A$-diffusion semigroup on $X$ with sample functions $\xi_t(\omega)$ then for every open subset $U$ of $X$,*

$$(15)\qquad \overset{t}{p}_U(x, E) = \Pr_x(\{\omega: \xi_s(\omega) \in U \text{ for } 0 \leqq s \leqq t, \xi_t(\omega) \in E\})$$

*is the sub-transition function of an $A$-diffusion process on $U$. (Here $\Pr_x$ denotes $\Pr_\nu$ with $\nu$ the mass 1 concentrated at $x$.)*

**Proof.** If $f$ is a function of class $C^2$ with compact support in $U$ let $K$ be a compact subset of $U$ containing the support of $f$ in its interior. By the corollary to Theorem 2, $\Pr_x(\{\omega: \xi_s(\omega) \in U \text{ for } 0 \leqq s \leqq t\}) = 1+o(t)$ uniformly for $x$ in $K$. Therefore $p_U^t(x, E) = p^t(x, E)+o(t)$ uniformly for $x$ in $K$, $E$ a subset of $K$, and so $P_U^t f(x) = P^t f(x)+o(t)$, which implies that $\lim_{t\downarrow 0}(1/t)(P_U^t f(x)-f(x)) = Af(x)$ uniformly for $x$ in $K$. Here $(P_U^t f(x) = \int f(y)p_U^t(x, dy)$.

Again by the corollary to Theorem 2, since the support of $f$ is in the interior of $K$, $P_U^t f(x) = P^t f(x)+o(t)$, $\lim_{t\downarrow 0}(1/t)(P_U^t f(x)-f(x)) = 0 = Af(x)$ uniformly for $x$ in $U-K$. Therefore $p_U^t$ is the sub-transition function of an $A$-diffusion process on $U$.

We say that a set $N$ *separates* two sets if they lie in different components of $X-N$.

THEOREM 6. *If $p^t$ is the sub-transition function of a Markoff process with continuous sample functions $\xi_t(\omega)$ on $\tilde{X}$, if $N$ is an open subset of $\tilde{X}$ which separates a set $W$ from the support of a bounded measurable function $f$, then for all $x$ in $W$, $|P^t f(x)| \leqq \sup_{y\in N; s\leqq t}|P^s f(y)|$.*

**Proof.** For $m=1, \cdots, 2^n$ let

$$\Phi(n, m) = \{\omega: \xi_{tm/2^n}(\omega) \in N, \xi_{tk/2^n}(\omega) \notin N \text{ for } k < m\},$$

$$\Phi(n) = \bigcup_{m=1}^{2^n} \Phi(n, m),$$

$$\Phi = \bigcup_{n=1}^{\infty} \Phi(n).$$

Then since $\Phi(n, m)$ is in the ring generated by the $\xi_{t/2^n}, \cdots, \xi_{tm/2^n}$ we have

$$\left|\int_{\Phi(n,m)} f(\xi_t(\omega))d\,\Pr_x(\omega)\right|$$
$$= \left|\int_{\Phi(n,m)} E\{f(\xi_t) \mid \xi_{t/2^n}, \cdots, \xi_{tm/2^n}\}(\omega)d\,\Pr_x(\omega)\right.$$
$$= \left|\int_{\Phi(n,m)} E\{f(\xi_t) \mid \xi_{tm/2^n}\}(\omega)d\,\Pr_x(\omega)\right|$$
$$= \left|\int_{\Phi(n,m)} P^{t-(tm/2^n)}f(\xi_{tm/2^n}(\omega))d\,\Pr_x(\omega)\right| \leqq \Pr_x(\Phi(n,m)) \sup_{y\in N;s\leqq t} \left| P^s f(y)\right| .$$

Since the $\Phi(n, m)$ for fixed $n$ are disjoint and have $\Phi(n)$ as their union, $\left|\int_{\Phi(n)}f(\xi_t(\omega))d\,\Pr_x(\omega)\right| \leqq \sup_{y\in N;s\leqq t} \left|P^s f(y)\right|$. Since the $\Phi(n)$ are an increasing sequence of sets and have $\Phi$ as their union,

$$\left|\int_{\Phi} f(\xi_t(\omega))d\Pr_x(\omega)\right| \leqq \sup_{y\in N;s\leqq t} \left| P^s f(y)\right| .$$

Finally, if $\xi_0(\omega)$ is in $W$ and $\xi_t(\omega)$ is in the support of $f$ then $\omega$ must be in $\Phi$ by continuity, since $N$ separates $W$ from the support of $f$ and is open. Therefore for all $x$ in $W$, $\left|P^t f(x)\right| \leqq \left|\int_\Phi f(\xi_t(\omega))d\,\Pr_x(\omega)\right| \leqq \sup_{y\in N;s\leqq t}\left|P^s f(y)\right|$.

THEOREM 7. *If $B$ and $A$ are as in Theorem 4, $p^t$ is the sub-transition function given by Theorem 4, $U$ is the unit sphere of a coordinate system of class $C^3$ in $B$, so that $\overline{U}\subset B$, and $p_U^t$ is defined by (15) then $p_U^t$ is the sub-transition function of the $A$-diffusion process on $U$ given by Theorem 4. That is, $P_U^t$ is a strongly continuous semigroup on $C(U)$ with infinitesimal generator the closure of $A$ on the domain $D(U)$ of all functions $f$ in $C(U)$ which can be extended to be a function of class $C^2$ on an open set containing $\overline{U}$ and which is such that $Af$ is in $C(U)$.*

**Proof.** Let $D_0$ be the set of all functions of class $C^2$ with compact support contained in $U$. Let $\mathfrak{X}$ be the Banach space of all functions which are uniform limits of functions of the form $P_U^t f$ with $f$ in $D_0$. The norm on $\mathfrak{X}$ is the supremum norm. By Theorem 5, $p_U^t$ is the sub-transition function of a diffusion process on $U$. Therefore $P_U^t f$ converges to $f$ in $\mathfrak{X}$ as $t$ tends to 0. Since the $P_U^t$ are uniformly bounded in norm (they are all contraction operators) $P_U^t$ is a strongly continuous semigroup on $\mathfrak{X}$. Let $\tilde{A}$ be its infinitesimal generator. Then $\tilde{A}$ is some closed extension of the restriction of $A$ to $D_0$.

Let $f$ be in $D(U)$. We wish to show that $f$ is in the domain of $\tilde{A}$ and $\tilde{A}f=Af$.

Since $f$ is in $D(U)$ it may be extended to a function, also called $f$, of class $C^2$ with compact support in $X$.

Let $N_j$ be a decreasing sequence of open sets whose intersection is the boundary $\partial U$ of $U$. If for each $j$, $\Phi=\Phi(N_j)$ is defined as in Theorem 6, then as in the proof of Theorem 6, $\left|\int_{\Phi(N_j)}f(\xi_t(\omega))d\,\Pr_x(\omega)\right| \leqq \sup_{y\in N_j;s\leqq t}\left|P^s f(y)\right|$. If $\Psi$ is defined by $\Psi=\bigcap_j \Phi(N_j)$ then we have the inequality

$$\left| \int_{\Psi} f(\xi_t(\omega)) d \operatorname{Pr}_x(\omega) \right| \leqq \inf_j \sup_{y \in N_j; s \leqq t} \left| P^s f(y) \right| = \sup_{y \in \partial U; s \leqq t} \left| P^s f(y) \right| .$$

The last equality holds because $P^s$ is a strongly continuous semigroup on $C(X)$, so that $P^s f(y)$ is jointly continuous in $s$ and $y$.

Now $t^{-1}P^t f(y)$ converges to 0 uniformly for $y$ in $\partial U$ since $f$ is in the domain $D$ of Theorem 4 and $Af(y)=0$ for $y$ in $\partial U$. Therefore $t^{-1} \sup_{y\in\partial U; s\leqq t} \left|P^s f(y)\right| \leqq \sup_{y\in\partial U;s\leqq t} \left|s^{-1}P^s f(y)\right| \to 0$ as $t \downarrow 0$. Therefore $\lim_{t\downarrow 0} t^{-1}\int_\Psi f(\xi_t(\omega))d \operatorname{Pr}_x(\omega)=0$ uniformly in $x$.

If $\xi_0(\omega)\in U$ and $\xi_s(\omega)\notin U$ for some $s$ in the interval $[0, t]$ then $\omega$ is in $\Psi$ by continuity. Therefore for all $x$ in $U$, $t^{-1}(P^t f(x)-f(x))-t^{-1}(P^t_U f(x)-f(x)) =t^{-1}(P^t f(x)-P^t_U f(x))=t^{-1}\int_\Psi f(\xi_t(\omega))d \operatorname{Pr}_x(\omega)\to 0$ uniformly for $x$ in $U$. Hence $\lim_{t\downarrow 0} t^{-1}(P^t_U f(x)-f(x))=Af(x)$ uniformly for $x$ in $U$. That is, $f$ is in the domain in $\tilde{A}$ and $\tilde{A}f=Af$.

Let $\hat{A}$ be the closure of $A$ on $D(U)$, $\hat{P}^t$ the semigroup it generates (by Theorem 4), $R_\lambda=(1-\lambda\hat{A})^{-1}$ for $\lambda>0$. Since $\tilde{A}$ is a closed operator $\tilde{A}f=\hat{A}f$ for all $f$ in the domain of $\hat{A}$. In particular for all $g$ in $C(U)$, $\tilde{A}R_\lambda g=\hat{A}R_\lambda g$, $(1-\lambda\tilde{A})R_\lambda g=g$, and so $(1-\lambda\tilde{A})^{-1}g=R_\lambda g$ for all $g$ in $C(U)$. Since $P^t_U$ is a strongly continuous semigroup, $P^t_U g=\lim_{n\to\infty}(R_{t/n})^n g=\hat{P}^t g$, as is shown in [5, p. 236], for all $g$ in $C(U)$. Therefore $P^t_U$ maps $C(U)$ into itself, $P^t_U=\hat{P}^t$, and $\tilde{A}=\hat{A}$, concluding the proof.

## 4. Global diffusion.

THEOREM 8. *If $X$ and $A$ are as in* §3 (*i.e. $X$ is a second countable differentiable manifold of class $C^3$ and $A$ is a differential operator such that for each $x$ in $X$ there is a coordinate system of class $C^3$ near $x$ with respect to which $A$ has the form* (5) *with the $a^{ij}$, $b^i$ of class $C^3$ and $a^{ij}$ a positive definite matrix*) *then there is an $A$-diffusion semigroup on $X$. That is, there is a family of linear transformations $P^t$, $0\leqq t<\infty$, on the space of bounded measurable functions satisfying* (1), (2), (3), (4), *and* (10).

If $X$ is an open set in Euclidean space this is the existence theorem of §0. The idea of the proof is simple. We merely let the diffusing particle wander in each coordinate neighborhood according to the process established in Theorem 4. Theorem 7 shows that the choice of coordinate neighborhood about a given position does not matter. The unpleasant details follow.

**Proof.** Let $X$ be given a Riemann metric of class $C^3$. For each $x$ in $X$ if $\epsilon$ is a sufficiently small strictly positive number then $U=N_\epsilon(x)$ will be a convex coordinate neighborhood of class $C^3$ whose closure is contained in a larger coordinate neighborhood. Let $\mathfrak{U}$ be a countable covering of $X$ by such sets $U$. Then the intersection of any two elements of $\mathfrak{U}$ is either empty or is a convex open set, and hence an open cell.

If $U$ is in $\mathfrak{U}$ let $p^t_U$ be the sub-transition function on $U$ given by Theorem 4. Let $\tilde{U}=U\cup\{\infty\}$ be the one-point compactification of $U$. Then $\prod_{0\leqq t<\infty}\tilde{U}$

is, by means of the obvious identification, a subset of $\Omega=\prod_{0\leq t<\infty}\tilde{X}$ (although the topology on $\tilde{U}$ is not that induced by considering $\tilde{U}$ a subset of $\tilde{X}$). If $E$ is a Borel set in $\tilde{X}$, we will again define $E(t)$ by (13), for $0\leqq t<\infty$, and we will again set $\xi_t(\omega)=\omega(t)$ for all $\omega$ in $\Omega$ and $0\leqq t<\infty$. Let $\Pr(U, x; \cdot)$ be the measure on $\Omega$ defined by (14), with $p^t$ replaced by $p_U^t$ and $\nu$ the mass 1 concentrated at $x$, $x\in U$. By Theorem 3 and the second corollary to Theorem 1, $\Pr(U, x; \cdot)$ may be extended to the $\sigma$-ring $\mathcal{S}(\tilde{U})$ generated by the $E(t)$ with $E$ a Borel subset of $U$ and the set $\Delta(U)$ of $\omega$ in $\Omega$ which are continuous functions from $[0, \infty)$ to $\tilde{U}$, in such a way that $\Pr(U, x; \Delta(U))=1$.

If $0\leqq s<t$, let $\Delta(s, t)$ consist of all $\omega$ in $\Omega$ such that the restriction of $\omega$ to $[s, t]$ is a continuous function from $[s, t]$ to $X$. Let $\mathcal{S}(s, t)$ be the $\sigma$-ring generated by sets of the form $E(u)\cap\Delta(s, t)$ where $s\leqq u\leqq t$ and $E$ is a Borel set in $X$. If $U$ is an open subset of $X$, let

$$U(s, t) = \bigcap_{s\leqq u\leqq t} U(u) \cap \Delta(s, t). \tag{16}$$

Since the intersection in (16) may be replaced by the intersection over the (countable) set of rational $u$'s in the interval $[s, t]$, $U(s, t)$ is in $\mathcal{S}(\tilde{U})$ if $U$ is in $\mathfrak{U}$. Let $\mathcal{S}(U; s, t)$ be the $\sigma$-ring generated by sets of the form $E(u)\cap\Delta(s, t)$ where $E$ is a Borel subset of $U$, $U\in\mathfrak{U}$. Then $\mathcal{S}(U; s, t)\subset\mathcal{S}(\tilde{U})$, so that $\Pr(U, x; \Phi)$ is defined for all $\Phi$ in $\mathcal{S}(U; s, t)$.

Let $U_0, \cdots, U_{n-1}, U_n$ be in $\mathfrak{U}$ and let $0=t_0<\cdots<t_{n-1}<t_n<t_{n+1}=t$. Let $x\in U_0$, and define the measure $\Pr_x$ on $\mathcal{S}(U_0; 0, t_1)\times\cdots\times\mathcal{S}(U_{n-1}; t_{n-1}, t_n)\times\mathcal{S}(U_n; t_n, t)$ (which, by means of the natural identification, is a sub-$\sigma$-ring of $\mathcal{S}(0, t)$) by

$$\begin{aligned}\Pr_x(\Phi_0\times\cdots\times\Phi_{n-1}\times\Phi_{n-1})\\ &=\int_{\Phi_0}\Pr(U_0, x; d\omega_0)\cdots\int_{\Phi_{n-1}}\Pr(U_{n-1}, \xi_{t_{n-1}}(\omega_{n-2}); d\omega_{n-1})\\ &\quad\cdot\int_{\Phi_n}\Pr(U_n, \xi_{t_n}(\omega_{n-1}); d\omega_n)\end{aligned} \tag{17}$$

whenever $\Phi_i\in\mathcal{S}(U_i; t_i, t_{i+1})$ for $i=0, \cdots, n$.

First we show that if $\Phi\in\mathcal{S}(U_0; 0, t_1)\times\cdots\times\mathcal{S}(U_n; t_n, t)$ and in addition $\Phi\in\mathcal{S}(V_0; 0, s_1)\times\cdots\times\mathcal{S}(V_m; s_m, s)$ (where $V_0, \cdots, V_m$ are in $\mathfrak{U}$, $x\in V_0$, and $0<s_1<\cdots<s_m<s$) then the two definitions of $\Pr_x(\Phi)$ agree. It suffices to show that if $\Phi\in\mathcal{S}(U; 0, u)\cap\mathcal{S}(V; 0, u)$ then $\Pr(U, y; \Phi)=\Pr(V, y; \Phi)$ for all $y$ in $U\cap V$. Since $\Pr(U; y; \cdot)$ and $\Pr(V, y; \cdot)$ are induced by $p_U^u$ and $p_V^u$ respectively, it is enough to show that

$$\Pr(U, y; (U\cap V)(0, u)\cap E(u)) = \Pr(V, y; (U\cap V)(0, u)\cap E(u)) \tag{18}$$

whenever $y\in U\cap V$ and $E$ is a Borel set contained in $U\cap V$. Now, by definition of $\mathfrak{U}$, $U\cap V$ is an open cell, so that there is an increasing sequence of coordinate neighborhoods $W_n$ of class $C^3$ such that $\overline{W}_n \subset U \cap V$ and $\bigcup_n W_n = U \cap V$. Let $\Psi_n = W_n(0, u) \cap (E \cap W_n)(u)$. Applying Theorem 7 twice, we have that the left hand side of (18) is equal to

$$\sup_n \Pr(U, y; \Psi_n) = \sup_n \overset{u}{p}_{W_n}(y, E \cap W_n) = \sup_n \Pr(V, y; \Psi_n),$$

which is the right hand side of (18).

Now $\mathcal{S}(U_0;0,t_1) \times \cdots \times \mathcal{S}(U_n;t_n,t)$ is the $\sigma$-ring of subsets of $U_0(0, t_1) \cap \cdots \cap U_n(t_n, t)$ which are in $\mathcal{S}(0, t)$. We have just shown that $\Pr_x(\Phi)$ is independent of the set $U_0(0, t_1) \cap \cdots \cap U_n(t_n, t)$ containing $\Phi$. Furthermore, $\Delta(0, t)$ is the union of a countable number of sets of this form: we take the union for all finite ordered sets $\{U_0, \cdots, U_n\}$ of sets in $\mathfrak{U}$ and all rational $t_1 < \cdots < t_n < t$. Hence $\Pr_x$ may be extended to be a measure on $\mathcal{S}(0, t)$. Now we shall define

$$p^t(x, E) = \Pr_x (E(t) \cap \Delta(0, t)) \tag{19}$$

for all $0 \leq t < \infty$, $x \in X$, and Borel sets $E \subset X$. If $f$ is a bounded Borel-measurable function on $X$, let

$$P^t f(x) = \int f(\xi_t(\omega)) d \Pr{}_x(\omega). \tag{20}$$

Then (10) holds by definition, (3) holds since $\Pr_x$ is a positive measure, and (1) holds by (17). Also, $\Pr_x(\Delta(0, t)) \leq 1$ since $\Pr_x(\Delta(0, t)) \leq p^t_{U_0}(x, U_0) + p^t_{U_0}(x, \{\infty\}) = 1$ by (17), so that (2) holds. It remains to prove (4).

Let $f$ be a function of class $C^2$ with compact support contained in some $U$ in $\mathfrak{U}$. If we can show that (4) holds for such an $f$, it follows by taking a $C^2$ partition of unity that (4) holds for all functions of class $C^2$ with compact support in $X$. But by the corollary to Theorem 2, if $K$ is a compact subset of $U$ containing the support of $f$ in its interior then $\Pr(u, x;\ U(0, t)) = 1 + o(t)$ uniformly for $x$ in $K$. By definition, $p^t(x, E) \geq \Pr(U, x;\ U(0, t)) \cap E(t)$ so that $p^t(x, E) = p^t_U(x, E) + o(t)$ uniformly for $x$ in $K$ and $E \subset U$. Therefore $\int f(y) p^t(x,dy) - \int f(y) p^t_U(x,dy) = o(t)$, $\lim_{t \downarrow 0} t^{-1}(P^t f(x) - f(x)) = Af(x)$ uniformly for $x$ in $K$. Now let $N$ be an open subset of $K$ separating the support of $f$ from $X - K$. By Theorem 6, $|P^t f(x)| \leq \sup_{y \in N; s \leq t} |P^s f(y)|$ for all $x$ in $X - K$. Since $s^{-1} P^s f(y)$ converges to 0 uniformly for $y$ in $N$, this implies that $t^{-1} P^t f(x)$ converges to 0 uniformly for $x$ in $X - K$. Therefore

$$\lim_{t \downarrow 0} t^{-1}(P^t f(x) - f(x)) = Af(x)$$

uniformly for all $x$ in $X$, and the theorem is proved.

## References

1. J. L. Doob, *Stochastic processes*, New York, 1953.

2. F. G. Dressel, *The fundamental solution of the parabolic equation* I and II, Duke Math J. vol. 7 (1940) pp. 186–203 and vol. 13 (1946) pp. 61–70.

3. W. Feller, *Zur Theorie der stochastischen Prozesse*, Math. Ann. vol. 113 (1936) pp. 113–160.

4. P. R. Halmos, *Measure theory*, New York, 1950.

5. E. Hille, *Functional analysis and semi-groups*, Amer. Math. Soc. Colloquium Publications, vol. 31, New York, 1948.

6. S. Itô, *The fundamental solution of the parabolic equation in a differentiable manifold*, Osaka Mathematical Journal vol. 5 (1953) pp. 75–92.

7. J. R. Kinney, *Continuity properties of sample functions of Markov processes*, Trans. Amer. Math. Soc. vol. 82 (1956) pp. 452–493.

8. A. Kolmogoroff, *Grundbegriffe der Wahrscheinlichskeitrechnung*, Ergebnisse der Mathematik, vol. 2, 1933.

9. D. Ray, *Stationary Markov processes with continuous paths*, Trans. Amer. Math. Soc. vol. 82 (1956) pp. 452–493.

10. J. Schauder, *Über lineare elliptische Differentialgleichungen zweiter Ordnung*, Math. Zeit. vol. 38 (1933) p. 257.

11. I. E. Segal, *Equivalences of measure spaces*, Amer. J. Math. vol. 83 (1951) pp. 275–313.

12. D. C. Spencer, *Heat conduction on arbitrary Riemannian manifolds*, Proc. Nat. Acad. Sci. U.S.A. vol. 39 (1953) pp. 327–330.

13. K. Yosida, *On the differentiability and the representation of one-parameter semigroup of linear operators*, Journal of the Mathematical Society of Japan vol. 1 (1949) p. 15.

14. ———, *Integrability of the backward diffusion equation in a compact Riemannian space*, Nagoya Mathematical Journal vol. 3 (1951) p. 1.

15. ———, *An ergodic theorem associated with harmonic integrals*, Proceedings of the Japanese Academy vol. 27 (1951) pp. 540–543.

16. ———, *On the fundamental solution of the parabolic equation in a Riemannian space*, Osaka Mathematical Journal vol. 5 (1953) pp. 65–74.

Institute for Advanced Study,
Princeton, N. J.