*On Convergent Diffusions: The Densities and the Conditioned Processes**

THOMAS M. LIGGETT

Communicated by M. KAC

1. Introduction and notation. Let $X_n(t)$ be a sequence of birth and death processes or diffusions, and let $X(t)$ be a diffusion. In [6], Stone has given conditions which guarantee the weak convergence of X_n to X . Our first objective here is to show that the convergence of X_n to X automatically implies the uniform convergence of the transition densities of X_n to those of X on compact subsets of the interior of the state space of X . (Under a mild further condition, convergence also takes place at regular boundary points of X .) Ordinarily, one would expect this local theorem to require more stringent conditions on the processes than are required for their weak convergence. However, the path continuity of both birth and death processes and diffusions enables us to obtain the local theorem from the global one in this case. We understand that Stone has obtained unpublished related results through entirely different techniques.

Our second aim is to use this local theorem to show that the processes $(X_n(t) \mid X_n(1) = x_n)$ converge weakly to $(X(t) \mid X(1) = x)$ where $x_n \rightarrow x$ and where these conditioned processes are given their natural definition in terms of their transition densities. In [4] and [5], the author considered this problem for random walks which are in the domain of attraction of a stable process.

In the final section, we will show that the processes X_n converge weakly to X if and only if the finite dimensional distributions converge. In other words, the requirement that $\{X_n\}$ be relatively compact is automatically satisfied in this case. This, together with the results of Section 2, then implies that the convergence of the transition densities follows solely from the convergence of the transition functions.

For our purposes, a diffusion $X(t)$ will have as its state space an interval S whose closure is $[0, 1]$. The process is permitted to jump to the death point Δ , but we will assume that all points in $(0, 1)$ are regular. Let σ_a be the hitting

* Research supported in part by the National Science Foundation through its Graduate Fellowship program and by contract AF-AFOSR-69-1781 at the University of California, Los Angeles, California.

time of a . Then we take $0 \in S$ if and only if

$$\lim_{b \downarrow 0} \lim_{a \downarrow 0} E^a(e^{-\sigma_b}) = 1 = \lim_{a \downarrow 0} \lim_{b \downarrow 0} E^a(e^{-\sigma_b}).$$

$1 \in S$ if the corresponding condition is satisfied at 1. S^* will be the union of S and a subset of its boundaries. $0 \in S^* \setminus S$ if and only if

$$\lim_{b \downarrow 0} \lim_{a \downarrow 0} E^a(e^{-\sigma_b}) = 0,$$

with a similar condition determining whether or not $1 \in S^* \setminus S$. With these assumptions, $X(t)$ is a Feller process in the sense that the transition operators map the bounded continuous functions on S into themselves. The speed measure of $X(t)$ will be denoted by $m(dy)$.

Ito and McKean (see Section 4.11 of [2]) have shown that the transition functions of $X(t)$ have densities $p(t, x, y)$ with respect to the speed measure $m(dy)$. That is, if τ denotes the killing time of $X(t)$,

$$P^x(X(t) \in dy, \tau > t) = p(t, x, y)m(dy),$$

for $(t, x, y) \in (0, \infty) \times S \times S$ where $p(t, x, y)$ is strictly positive and continuous on $(0, \infty) \times S \times S$, is symmetric in x and y , and satisfies

$$p(t, x, y) = \int_S p(s, x, z)p(t-s, z, y)m(dz), \quad 0 < s < t; \quad x, y \in S.$$

Furthermore, $p(t, b, y) = 0$ for $y \in S$ and $b \in S^* \setminus S$ where $p(t, b, y)$ denotes the limit of $p(t, x, y)$ as $x \rightarrow b$ from within $(0, 1)$.

By a birth and death process $V(t)$, we will mean a Markov chain whose state space $S \subset [0, 1]$ has no limit points in $(0, 1)$ and does not contain any of its limit points, and which can make transitions only to adjacent states or to Δ . We will assume that all states communicate with each other and that there are no instantaneous states. Then there is a speed measure $m(dy)$ which concentrates on S and a function $p(t, x, y)$ defined for $t \geq 0$ and $x, y \in S$ so that

$$P^x(V(t) = y) = p(t, x, y)m\{y\},$$

$p(t, x, y) > 0$ for $t > 0$ and $x, y \in S$ and $p(t, x, y)$ is symmetric in x and y . For simplicity, we will assume that $0, 1 \in \bar{S}$.

In [3], Karlin and McGregor showed that if $V(t)$ is either a diffusion or a birth and death process, the transition function $P^x(V(t) \in E)$ is totally positive for each $t > 0$ as a function of x and E . In particular, if $x_1, x_2 \in S$ and $E_1, E_2 \subset S$ where $m(E_i) > 0$ are taken so that $x_1 < x_2$ and $E_1 < E_2$ (in the sense that $y_i \in E_i$ implies that $y_1 < y_2$), then

$$(1) \quad \frac{P^{x_2}(V(t) \in E_2)}{P^{x_1}(V(t) \in E_2)} \geq \frac{P^{x_2}(V(t) \in E_1)}{P^{x_1}(V(t) \in E_1)}.$$

It is this property of processes with continuous paths that will enable us to obtain the local limit theorem from the global one. By letting E_1 and E_2 shrink

to points in S , it is easy to check that inequality (1) holds also for the transition densities $p(t, x, y)$.

Throughout this paper, X will always denote a diffusion and X_n will denote either a diffusion or a birth and death process. We will say that $X_n \Rightarrow X$ if

- (2) whenever $x_n \in S_n$, $x \in S$ and $x_n \rightarrow x$ then $(X_n(t) \mid X_n(0) = x_n)$ converges weakly in Skorokhod's space $D[0, T]$ for each $T < \infty$ (see Chapter 3 of [1]) to $(X(t) \mid X(0) = x)$, and
 (3) there is a closed interval $A \subset (0, 1)$ so that $m_n(A) \rightarrow m(A)$ and $m(\partial A) = 0$.

Note that (3) is really no restriction at all. Since the speed measures are unique only up to a multiplicative constant, (3) serves only as a normalization and can always be achieved by proper choice of the measures $m_n(dy)$. As will be seen in the proof of Theorem 1, if $X_n \Rightarrow X$, then $m_n(dy)$ converges as a sequence of measures to $m(dy)$ on compact subsets of $(0, 1)$.

In studying the behavior of the densities $p_n(t, x_n, y_n)$, it will suffice to assume that $X_n \Rightarrow X$ as long as either x_n or y_n remains in a compact subset of $(0, 1)$ and the other remains in a compact subset of S^* . On the other hand, if for example both x_n and y_n approach a point in $S \setminus (0, 1)$, it will be necessary to make a further assumption which we will now state. If $a \in S \setminus (0, 1)$, let $\sigma(x(\cdot)) = \inf \{t > 0, x(t) = a\}$. It will be said that assumption $\mathfrak{A}$ holds at a if $(\sigma(X_n) \mid X_n(0) = u_n)$ converges in distribution to $(\sigma(X) \mid X(0) = u)$ whenever $u_n \in S_n$, $u \in S$ and $u_n \rightarrow u$. This condition will be assumed only when explicitly mentioned. While condition $\mathfrak{A}$ does not follow from $X_n \Rightarrow X$, it should be possible to check whether or not it holds whenever it is possible to verify that $X_n \Rightarrow X$. The remark that follows Theorem 1 explains this somewhat more fully.

2. The densities. The first step in obtaining the convergence of the densities is contained in the following lemma. Its proof depends in an essential way on the total positivity of the transition functions of the processes involved.

Lemma 1. Assume that $X_n \Rightarrow X$. Then if $t > 0$,

$$\frac{p_n(t, x_n, y_n)}{p_n(t, z_n, y_n)} \rightarrow \frac{p(t, x, y)}{p(t, z, y)}$$

whenever $x_n, y_n, z_n \in S_n$, $z \in S$, $x \in S^*$, $y \in (0, 1)$, $x_n \rightarrow x$, $y_n \rightarrow y$, and $z_n \rightarrow z$.

Proof. Let E and F be two closed intervals in $(0, 1)$ whose endpoints are continuity points of $m(dx)$ and which satisfy

$$E < \{y\} < F.$$

From the total positivity of the transition functions $P^n(X_n(t) \in U)$, it follows that for n so large that $E < \{y_n\} < F$,

$$(4) \quad \frac{P^{x_n}(X_n(t) \in F)}{P^{z_n}(X_n(t) \in F)} \leq \frac{p_n(t, x_n, y_n)}{p_n(t, z_n, y_n)} \leq \frac{P^{x_n}(X_n(t) \in E)}{P^{z_n}(X_n(t) \in E)}$$

if $x_n \leq z_n$, while if $x_n \geq z_n$ these inequalities are reversed.

Since $X_n \Rightarrow X$, if $x \in S$ we may let $n \rightarrow \infty$ to show that all limit points of the sequence

$$\frac{p_n(t, x_n, y_n)}{p_n(t, z_n, y_n)},$$

lie between

$$\frac{P^x(X(t) \in F)}{P^x(X(t) \in F)} \text{ and } \frac{P^x(X(t) \in E)}{P^x(X(t) \in E)}.$$

These two expressions approach

$$\frac{p(t, x, y)}{p(t, z, y)}$$

as E and F decrease to points, and then those points move toward y .

If $x \in S^* \setminus S$, we see from (4) that it will suffice to show that as $n \rightarrow \infty$,

$$(5) \quad P^{x_n}(X_n(t) \in E \cup F) \rightarrow 0.$$

Let $u_n \in S_n$ satisfy $u_n \rightarrow u$ where u lies between x and $E \cup F$. Then, since $X_n \Rightarrow X$ and the boundaries of $E \cup F$ are regular points for X ,

$$(6) \quad \overline{\lim}_{n \rightarrow \infty} P^{u_n}(X_n(s) \in E \cup F \text{ for some } s \leq t) \\ \leq P^u(X(s) \in E \cup F \text{ for some } s \leq t).$$

Since $x \in S^* \setminus S$, the right side of (6) tends to zero as $u \rightarrow x$. By the strong Markov property,

$$P^{x_n}(X_n(t) \in E \cup F) \leq P^{u_n}(X_n(s) \in E \cup F \text{ for some } s \leq t),$$

for sufficiently large n . So, (5) is established, and the proof of the lemma is complete.

Theorem 1. Assume that $X_n \Rightarrow X$. Then if $t > 0$,

$$(7) \quad p_n(t, x_n, y_n) \rightarrow p(t, x, y) \text{ whenever}$$

$$(8) \quad x_n, y_n \in S_n, \quad x \in (0, 1), \quad y \in S^*, \quad x_n \rightarrow x, \text{ and } y_n \rightarrow y.$$

If, in addition, condition $\mathcal{Q}$ is satisfied at x , (7) holds whenever

$$(9) \quad x_n, y_n \in S_n, \quad x \in S, \quad y \in S^*, \quad x_n \rightarrow x, \text{ and } y_n \rightarrow y.$$

Proof. Fix $t > 0$ and $\epsilon > 0$ so that $A \subset (\epsilon, 1 - \epsilon)$, and choose a sequence $u_n \in S_n$ which converges to $u \in (0, 1)$. Then

$$P^{u_n}(X_n(t) \in A) = \int_A p_n(t, u_n, w) m_n(dw) \\ \geq \inf_{w \in A \cap S_n} p_n(t, u_n, w) m_n(A).$$

Since $m_n(A) \rightarrow m(A) > 0$, there is a sequence $w_n \in S_n \cap A$ so that $\{p_n(t, u_n, w_n)\}$ is bounded in n . Using the same argument on the set $[\epsilon, 1 - \epsilon]$ together with the result of Lemma 1, it follows that $\{m_n[\epsilon, 1 - \epsilon]\}$ is bounded in n , so $m_n(dw)$ is a relatively compact sequence of measures on $[\epsilon, 1 - \epsilon]$.

Now, suppose we knew that $w_n \rightarrow w_0 \in (0, 1)$, $p_n(t, u_n, w_n) \rightarrow c$ and $m_n(dw)$ converged weakly to $\mu(dw)$ on $[\epsilon, 1 - \epsilon]$. Then, writing

$$\frac{p_n(t, x_n, y_n)}{p_n(t, u_n, w_n)} = \frac{p_n(t, x_n, y_n)}{p_n(t, x_n, w_n)} \cdot \frac{p_n(t, x_n, w_n)}{p_n(t, u_n, w_n)},$$

we see from Lemma 1 and the symmetry of the transition densities that

$$(10) \quad p_n(t, x_n, y_n) \rightarrow c \frac{p(t, x, y)}{p(t, u, w_0)},$$

whenever x_n, y_n, x and y satisfy (8). So, if $x_n \in S_n$, $x_n \rightarrow x \in (0, 1)$ and B is a measurable subset of $(\epsilon, 1 - \epsilon)$ which is a continuity set for both $m(dw)$ and $\mu(dw)$, we have

$$\begin{aligned} P^{x_n}(X_n(t) \in B) &= \int_B p_n(t, x_n, w) m_n(dw) \\ &\rightarrow \int_B c \frac{p(t, x, w)}{p(t, u, w_0)} \mu(dw). \end{aligned}$$

On the other hand, since $X_n \Rightarrow X$,

$$P^{x_n}(X_n(t) \in B) \rightarrow P^x(X(t) \in B) = \int_B p(t, x, w) m(dw).$$

Since this is true for every such B , it follows that

$$c\mu(dw) = p(t, u, w_0)m(dw)$$

on $(\epsilon, 1 - \epsilon)$. Since $\mu(A) = m(A) > 0$, we conclude that $p(t, u, w_0) = c$ and that $\mu = m$ on $(\epsilon, 1 - \epsilon)$. Now, from (10) it follows that

$$p_n(t, x_n, y_n) \rightarrow p(t, x, y),$$

whenever (8) is satisfied.

We have now shown that if (8) holds, every sequence of integers contains a subsequence for which (7) holds. So, the proof of the first part of the theorem is complete.

There are several cases to consider in the second part of the theorem, depending on whether x and y are the same boundary or different boundaries, and whether both boundaries are in S or not. The proof varies little from one case to another, so it will be carried out here only for the case in which $0 \in S$ and $x = y = 0$. The idea of the proof is to use a new sequence of birth and death processes or diffusions for which 0 is an interior point of the state space. The total positivity of these new processes will then be used in a manner similar to that of the proof of Lemma 1 to obtain the desired result.

So, we will now assume, in addition to the convergence $X_n \Rightarrow X$, that condition $\mathcal{A}$ is satisfied at $0 \in S$. As a consequence of these assumptions, it follows that if n is sufficiently large and X_n is a diffusion, then 0 is in S_n also. For each such n , there is a diffusion Y_n on $S_n \cup -S_n$ so that $(Y_n \mid Y_n(0) = x)$ and $(-Y_n \mid Y_n(0) = -x)$ have the same distributions for $x \in S_n$ and so that $|Y_n|$ is a version of X_n . Y_n can be constructed simply by writing down the speed measure, scale function, and killing measure of Y_n in terms of those of X_n , and then using the standard existence theorem. If X_n is a birth and death process and n is sufficiently large, it again follows from condition $\mathcal{A}$ that $0 \in S_n$, so there is a birth and death process Y_n on $S_n \cup -S_n$ with the same properties.

Now take $x_n, u_n \in S_n$ and $E, F \subset (0, 1)$ with $x_n < u_n < 1$, $E < F$, $m_n(E) > 0$, and $m_n(F) > 0$. (Here m_n still denotes the speed measure of X_n .) Using the total positivity of the transition functions of Y_n , we see that

$$\frac{P^{x_n}(Y_n(t) \in E \cup -E)}{P^{u_n}(Y_n(t) \in E \cup -E)} \leq \frac{P^{x_n}(Y_n(t) \in -F)}{P^{u_n}(Y_n(t) \in -F)}.$$

Using the symmetry of the process Y_n and the strong Markov property, we have

$$P^{x_n}(Y_n(t) \in -F) = \frac{1}{2} P^{x_n}(\sigma(X_n) \leq t, X_n(t) \in F),$$

where $\sigma(X_n)$ is again the hitting time of zero. So,

$$\frac{P^{x_n}(X_n(t) \in E)}{P^{u_n}(X_n(t) \in E)} \leq \frac{P^{x_n}(\sigma(X_n) \leq t, X_n(t) \in F)}{P^{u_n}(\sigma(X_n) \leq t, X_n(t) \in F)}.$$

If $y_n \in S_n$ and $\{y_n\} < F$, it then follows that

$$\frac{p_n(t, x_n, y_n)}{p_n(t, u_n, y_n)} \leq \frac{P^{x_n}(\sigma(X_n) \leq t, X_n(t) \in F)}{P^{u_n}(\sigma(X_n) \leq t, X_n(t) \in F)}.$$

So, if $x_n, y_n, u_n \in S_n$, $x_n \rightarrow 0$, $y_n \rightarrow 0$, $u_n \rightarrow u \in (0, 1)$, $t > 0$ and F is a continuity set of $m(dw)$ which is contained in a compact subset of $(0, 1)$ and satisfies $m(F) > 0$, we have

$$(11) \quad \overline{\lim}_{n \rightarrow \infty} p_n(t, x_n, y_n) \leq \frac{p(t, u, 0) P^0(X(t) \in F)}{\lim_{n \rightarrow \infty} P^{u_n}(\sigma(X_n) \leq t, X_n(t) \in F)}.$$

By condition $\mathcal{A}$,

$$P^{u_n}(\sigma(X_n) \leq t) \rightarrow P^u(\sigma(X) \leq t),$$

while if G is any closed subset of $[0, 1] \cup \{\Delta\}$, it follows from $X_n \Rightarrow X$ that

$$\overline{\lim}_{n \rightarrow \infty} P^{u_n}(\sigma(X_n) \leq t, X_n(t) \in G) \leq P^u(\sigma(X) \leq t, X(t) \in G).$$

So,

$$P^{u_n}(\sigma(X_n) \leq t, X_n(t) \in F) \rightarrow P^u(\sigma(X) \leq t, X(t) \in F).$$

As $u \downarrow 0$, $P^u(\sigma(X) \leq t) \rightarrow 1$ since $0 \notin S$. So, from (11) we may now conclude that

$$\overline{\lim}_{n \rightarrow \infty} p_n(t, x_n, y_n) \leq p(t, 0, 0).$$

The fact that

$$\lim_{n \rightarrow \infty} p_n(t, x_n, y_n) \geq p(t, 0, 0),$$

follows from the total positivity of the transition functions of the processes X_n themselves as in the proof of Lemma 1. So, the proof of the theorem is complete.

Remark. If condition $\mathcal{Q}$ does hold, it should not be difficult to verify. To see this, consider the case in which $0 \notin S$ and we are interested in condition $\mathcal{Q}$ at 0. Then, if $u_n \in S_n$, $(\sigma(X_n) \mid X_n(0) = u_n)$ and $(\sigma(Y_n) \mid Y_n(0) = u_n)$ have the same distribution, where Y_n is the process constructed in the proof of Theorem 1. So, if $Y_n \Rightarrow Y$, condition $\mathcal{Q}$ will be satisfied at 0. Furthermore, checking that $Y_n \Rightarrow Y$ should be no more difficult than checking that $X_n \Rightarrow X$, since the infinitesimal generator of Y_n can be written down easily in terms of the infinitesimal generator of X_n .

We conclude this section with an example which shows that condition $\mathcal{Q}$ is necessary for the second part of Theorem 1 to hold.

Example 1. Let X be Brownian motion on $[0, 1]$ with reflecting boundaries at both zero and one. Fix $t > 0$, and choose $\{\gamma_n\}$ so that

$$\gamma_n n^4 \exp \left\{ -n^4 \frac{\pi^2}{2} t \right\} \rightarrow \infty.$$

Let $\delta_n = n^{-4} \gamma_n^{-2}$. Then define X_n to be a diffusion on $[0, 1]$ with no killing, with speed measure

$$m_n(dx) = \begin{cases} \gamma_n dx & \text{on } (0, \delta_n), \\ dx & \text{on } (\delta_n, 1) \end{cases},$$

and with scale function $s_n(\cdot)$ which linearly interpolates the values

$$s_n(0) = -\gamma_n, \quad s_n(\delta_n) = 0 \quad \text{and} \quad s_n(1) = 1.$$

Let each boundary be reflecting. Then, while Theorem 1 of [6] does not apply immediately to this situation, it is not hard to use it to show that $X_n \Rightarrow X$. However, if $B(u)$ is a standard Brownian motion on R^1 ,

$$\begin{aligned} P^{\delta_n/2}(X_n(t) \leq \delta_n) &\geq P^0 \left(\max_{u \leq 1} |B(u)| < \frac{(\delta_n)^{1/2} \gamma_n}{2(t)^{1/2}} \right) \\ &\geq K \exp \left\{ -n^4 \frac{\pi^2}{2} t \right\}, \end{aligned}$$

where K is a constant which is independent of n (see page 80 of [1]).

Since $m_n(0, \delta_n) = \gamma_n \delta_n$,

$$\frac{P^{\delta_n/2}(X_n(t) \leq \delta_n)}{m_n(0, \delta_n)} \geq K \gamma_n n^4 \exp \left\{ -n^4 \frac{\pi^2}{2} t \right\} \rightarrow \infty.$$

So, there is a sequence $x_n \in (0, \delta_n)$ so that

$$p_n \left(t, \frac{\delta_n}{2}, x_n \right) \rightarrow \infty.$$

3. The conditioned processes. If $\{E_n\}$ is a sequence of intervals and $\{x_n\}$ is a sequence of points satisfying $m_n(E_n) > 0$ and $x_n \in S_n$ for each n , we can define processes $(Z_n(t), 0 \leq t \leq 1)$ as random elements of $D[0, 1]$ which satisfy

$$P(Z_n \in \mathfrak{M}) = P^{x_n}(X_n \in \mathfrak{M} \mid X_n(1) \in E_n)$$

for each measurable subset $\mathfrak{M}$ of $D[0, 1]$. We then have the following result, whose proof depends very strongly on the results of Section 2.

Theorem 2. Suppose that $X_n \Rightarrow X$, that $x_n \rightarrow x \in S$ and that for some $y \in (0, 1)$

$$(12) \quad \limsup_{n \rightarrow \infty} \{ |z - y|, z \in E_n \} = 0.$$

Then Z_n converges weakly in $D[0, 1]$ to a (temporally inhomogeneous) Markov process Z which has transition densities

$$(13) \quad q(s, t, u, v) = \frac{p(t - s, u, v)p(1 - t, v, y)}{p(1 - s, u, y)}, \quad u, v \in S, \quad 0 < s < t < 1,$$

with respect to the measure $m(dv)$ and which satisfies $Z(0) = x$ and $Z(1) = y$. This is also true for $y = 0$ or 1 if $y \in S$ and condition $\mathfrak{A}$ is satisfied at y .

Proof. The first step is to show that the finite dimensional distributions of Z_n converge to those of Z . At this stage, we are not assuming that Z is a random element of $D[0, 1]$, but rather that it is the collection of consistent finite dimensional distributions described by (13). Let $\{C_i, i = 1, \dots, k\}$ be k continuity sets of $m(dw)$, and choose t_i so that $0 < t_1 < \dots < t_k < 1$. Then we may write

$$P(Z_n(t_i) \in C_i, i = 1, \dots, k) = \int_{C_k} P^{x_n}(X_n(t_i) \in C_i, 1 \leq i \leq k-1, X_n(t_k) \in dz) \frac{P^{x_n}(X_n(1 - t_k) \in E_n)}{P^{x_n}(X_n(1) \in E_n)}.$$

By Theorem 1, if $z_n \in S_n$ and $z_n \rightarrow z \in S^*$, then

$$(14) \quad \frac{P^{x_n}(X_n(1 - t_k) \in E_n)}{P^{x_n}(X_n(1) \in E_n)} \rightarrow \frac{p(1 - t_k, z, y)}{p(1, x, y)}.$$

Furthermore, the total positivity of the transition functions of X_n implies that the left side of (14) is uniformly bounded even if $z \in [0, 1] \setminus S^*$. Since $X_n \Rightarrow X$, it now follows that

$$\begin{aligned}
P(Z_n(t_i) \in C_i, i = 1, \dots, k) &\rightarrow P(Z(t_i) \in C_i, i = 1, \dots, k) \\
&= \int_{C_k} P^x(X(t_i) \in C_i, 1 \leq i \leq k-1, X(t_k) \in dz) \frac{p(1-t_k, z, y)}{p(1, x, y)}.
\end{aligned}$$

It now remains to be proved that the sequence $\{Z_n\}$ is tight in $D[0, 1]$. This will be done by considering the behavior of $Z_n(s)$ in the intervals $[0, t]$ and $[t, 1]$ separately, where t will be taken near 1. Fix a $t \in (0, 1)$ and suppose that $\mathfrak{M}$ is a subset of $D[0, 1]$ which is measurable with respect to the smallest Borel field in $D[0, 1]$ with respect to which $x(s)$ is a measurable function for all $s \in [0, t]$. Then

$$\begin{aligned}
P(Z_n \in \mathfrak{M}) &= P^{x_n}(X_n \in \mathfrak{M} \mid X_n(1) \in E_n) \\
&= \int_{S_n} P^{x_n}(X_n \in \mathfrak{M}, X_n(t) \in dz) \frac{P^x(X_n(1-t) \in E_n)}{P^{x_n}(X_n(1) \in E_n)}.
\end{aligned}$$

By (14) and the comment immediately following it, there are constants $M(t)$ and $n(t)$ so that

$$\frac{P^x(X_n(1-t) \in E_n)}{P^{x_n}(X_n(1) \in E_n)} \leq M(t), \quad z \in S_n,$$

for $n \geq n(t)$. So, we have for $n \geq n(t)$

$$(15) \quad P(Z_n \in \mathfrak{M}) \leq M(t) P^{x_n}(X_n \in \mathfrak{M})$$

for any measurable $\mathfrak{M} \subset D[0, 1]$ which depends on times only up to time t .

The final step is to show that for each $\epsilon > 0$,

$$(16) \quad \lim_{t \uparrow 1} \lim_{n \rightarrow \infty} P(\sup_{t \leq s \leq 1} |Z_n(s) - y| \geq \epsilon) = 0.$$

Using the total positivity of the transition functions of X_n and the type of estimate used in the proof of Lemma 1, it is not difficult to show that for each fixed $\epsilon > 0$, there is an N so that

$$p_n(s, z, w)$$

is bounded for $n \geq N$, $s \in [0, 1]$, z and $w \in S_n$, $|z - y| \geq \epsilon$, and $|w - y| \leq \epsilon/2$. So, there is a K for which

$$\frac{P^x(X_n(s) \in E_n)}{P^{x_n}(X_n(1) \in E_n)} \leq K,$$

for all $z \in S_n$ which satisfy $|z - y| \geq \epsilon$ and for sufficiently large n . Let σ_n be the stopping time defined by

$$\sigma_n(X_n) = \inf \{s > 0 \mid |X_n(s) - y| \geq \epsilon\}.$$

Then, by conditioning on $\sigma_n(X_n)$ and using the strong Markov property, we see that

$$P^w(\sigma_n(X_n) \leq s, X_n(s) \in E_n) \leq K P^w(\sigma_n(X_n) \leq s) P^{x_n}(X_n(1) \in E_n),$$

for all $w \in S_n$ and $s \in (0, 1)$. Now

$$P(\sup_{t \leq s \leq 1} |Z_n(s) - y| \geq \epsilon) \leq P^{x_n}(|X_n(t) - y| \geq \epsilon/2 | X_n(1) \in E_n) \\ + \int_{|w-y| \leq \epsilon/2} P^{x_n}(X_n(t) \in dw) \frac{P^w(\sigma_\epsilon(X_n) \leq 1-t, X_n(1-t) \in E_n)}{P^{x_n}(X_n(1) \in E_n)}.$$

From this, the convergence of the finite dimensional distributions of Z_n , and the fact that $X_n \Rightarrow X$, we see that

$$\overline{\lim}_{n \rightarrow \infty} P(\sup_{t \leq s \leq 1} |Z_n(s) - y| \geq \epsilon) \leq P(|Z(t) - y| \geq \epsilon/2) \\ + K \int_{|w-y| \leq \epsilon/2} P^x(X(t) \in dw) P^w(\sigma_\epsilon(X) \leq 1-t).$$

Since $\lim_{t \uparrow 1} \sup_{|w-y| < \epsilon/2} P^w(\sigma_\epsilon(X) \leq 1-t) = 0$, (16) has been established.

Since $(X_n | X_n(0) = x_n)$ is tight in $D[0, 1]$, we can now use Theorem 15.3 of [1], together with (15) and (16), to show that the sequence $\{Z_n\}$ is tight also. The details of this argument can be found in the proof of Theorem 4 of [4]. So, by Theorem 15.1 of [1], Z_n converges weakly to Z .

If we let $X_n \equiv X$ (a diffusion) in Theorem 2 above, and we take $x, y \in S$, it follows from that theorem that there is a Markov process in $D[0, 1]$ which has transition densities

$$\frac{p(t-s, u, v)p(1-t, v, y)}{p(1-s, u, y)}, \quad u, v \in S, \quad 0 < s < t < 1,$$

with respect to $m(dv)$ and which equals x at $t = 0$ and y at $t = 1$ with probability one. We will call this process

$$(X | X(0) = x, X(1) = y).$$

In the birth and death case, the process

$$(X_n | X_n(0) = x_n, X_n(1) = y_n),$$

exists trivially in $D[0, 1]$ when $x_n, y_n \in S_n$, since now the conditioning event has positive probability. We can now state the following simple corollary to Theorem 2.

Corollary. Suppose that $X_n \Rightarrow X$ and that $x_n, y_n \in S_n$ satisfy $x_n \rightarrow x \in S$ and $y_n \rightarrow y \in (0, 1)$. Then $(X_n | X_n(0) = x_n, X_n(1) = y_n)$ converges weakly in $D[0, 1]$ to $(X | X(0) = x, X(1) = y)$. This is also true for $y = 0$ or 1 if $y \in S$ and condition α is satisfied at y .

Proof. Let $E_{n,m}$ be intervals satisfying $m_n(E_{n,m}) > 0$ and

$$\limsup_{m \rightarrow \infty} \sup_n \{|z - y_n|, z \in E_{n,m}\} = 0.$$

Then, by Theorem 2, $(X_n | X_n(0) = x_n, X_n(1) \in E_{n,m_n})$ converges weakly to

$(X \mid X(0) = x, X(1) = y)$ as $n \rightarrow \infty$ whenever $m_n \rightarrow \infty$. On the other hand, if X_n is a diffusion, Theorem 2 also says that $(X_n \mid X_n(0) = x_n, X_n(1) \in E_{n,m})$ converges weakly to $(X_n \mid X_n(0) = x_n, X_n(1) = y_n)$ as $m \rightarrow \infty$ for each n . If X_n is a birth and death process, this is immediate since the probability of the conditioning event is bounded below for each n . So, $(X_n \mid X_n(0) = x_n, X_n(1) = y_n)$ converges weakly to $(X \mid X(0) = x, X(1) = y)$.

4. Weak convergence. In this final section, we will show that the processes X_n converge weakly to X in $D[0, 1]$ if and only if the corresponding transition functions converge. By the convergence of the transition functions, which we will denote by $X_n \rightarrow X$, we will mean that if $t > 0$, $x_n \in S_n$, and $x_n \rightarrow x \in S$, then the measures $P^{x_n}(X_n(t) \in du)$ converge on R^1 to $P^x(X(t) \in du)$. We will need the following preparatory lemmas. The first states that the transition functions actually converge uniformly in t for bounded t .

Lemma 2. Assume that $X_n \rightarrow X$. If $x_n \in S_n$, $x_n \rightarrow x \in S$, and $t_n \rightarrow 0$, then for each $\epsilon > 0$

$$\lim_{n \rightarrow \infty} P^{x_n}(|X_n(t_n) - x| > \epsilon) = 0.$$

Proof. Let h be a continuous function supported by a compact subset of S which satisfies $h(x) = 1$ and $0 \leq h(y) < 1$ for all $y \neq x$. Define for each n and each $t \geq 0$

$$f_n(t) = E^{x_n}[h(X_n(t))],$$

and

$$f(t) = E^x[h(X(t))].$$

Then $|f_n(t)| \leq 1$ and $f_n(t) \rightarrow f(t)$ for each $t \geq 0$, and $f(0) = 1$. By the Markov property,

$$f_n(t + t_n) = E^{x_n} E^{X_n(t_n)}[h(X_n(t))].$$

Integrating this relation over the interval $0 \leq t \leq \delta$, we obtain

$$(17) \quad \int_{t_n}^{t_n + \delta} f_n(s) ds = E^{x_n} \int_0^\delta E^{X_n(t_n)} [h(X_n(t))] dt.$$

Let K be a compact subset of S so that $x \notin K$. Then

$$\frac{1}{\delta} \int_{t_n}^{t_n + \delta} f_n(s) ds \leq P^{x_n}(X_n(t_n) \notin K) + P^{x_n}(X_n(t_n) \in K) u_n(\delta),$$

where

$$u_n(\delta) = \sup_{y \in K \cap S_n} \frac{1}{\delta} \int_0^\delta E^y[h(X_n(t))] dt.$$

Since $X_n \rightarrow X$ and $h(y) < 1$ on K ,

$$\lim_{\delta \rightarrow 0} \lim_{n \rightarrow \infty} u_n(\delta) < 1,$$

and

$$\lim_{\delta \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{\delta} \int_{t_n}^{t_n + \delta} f_n(s) ds = 1.$$

It follows that $P^{x_n}(X_n(t_n) \in K) \rightarrow 0$. So, any vague limit of a subsequence of the measures $P^{x_n}(X_n(t_n) \in dy)$ must vanish off x . Hence, it suffices to show that these measures are tight. To do that, let $\mathcal{O}$ be a neighborhood of the support of h which has compact closure in S , and let τ denote the hitting time of $\mathcal{O}$. Then, if $u_n \in S_n \setminus \mathcal{O}$ and $s_n \rightarrow 0$,

$$E^{u_n}[h(X_n(s_n))] = \int_{S_n} \int_0^{s_n} P^{u_n}(\tau(X_n) \in ds, X_n(\tau) \in dv) E^v[h(X_n(s_n - s))].$$

By what we have just proved,

$$E^v[h(X_n(s_n - s))] \rightarrow 0$$

uniformly for all s and v which occur in this integrand. Hence,

$$E^{u_n}[h(X_n(s_n))] \rightarrow 0.$$

Using (17) again, we see that the measures $P^{x_n}(X_n(t_n) \in dy)$ are tight in S , thus concluding the proof of the lemma.

In what follows, let $\rho(x, y)$ be any metric on S which is equivalent to the Euclidean metric.

Lemma 3. Assume that $X_n \rightarrow X$. Given a compact subset K of S and $\epsilon > 0$, there is a $\delta > 0$ so that

$$P^x(\sup_{t \leq \delta} \rho(X_n(t), x) > \epsilon) \leq \epsilon,$$

for all $x \in S_n \cap K$ and all n .

Proof. For the proof, it suffices to write

$$\begin{aligned} P^x(\sup_{t \leq \delta} \rho(X_n(t), x) > \epsilon) &\leq P^x(\rho(X_n(\delta), x) > \epsilon/2) \\ &\quad + P^x(\rho(X_n(\delta), x) \leq \epsilon/2, \sup_{t \leq \delta} \rho(X_n(t), x) > \epsilon). \end{aligned}$$

The first term on the right can be made arbitrarily small by Lemma 2. The same lemma can also be used to make the second term arbitrarily small after conditioning on the stopping time

$$\inf \{t > 0 \mid \rho(X_n(t), x) > \epsilon\}$$

and applying the strong Markov property.

Lemma 4. Assume that $X_n \rightarrow X$ and take $x_n \in S_n$ so that $x_n \rightarrow x \in S$. Then given $\epsilon > 0$, there is a compact subset K of S so that for each n ,

$$P^{x_n}(X_n(t) \in S_n \setminus K \text{ for some } t \leq 1) < \epsilon.$$

Proof. Let Θ be an open set with compact closure in S which satisfies

$$P^x(X(t) \in S \setminus \Theta \text{ for some } t \leq 1) < \frac{\epsilon}{2}.$$

Take ρ in Lemma 3 to have the property that

$$K = \left\{ x \in S \mid \rho(x, \Theta) < \frac{\epsilon}{2} \right\}$$

is compact in S . Then, there is a $\delta > 0$ so that

$$P^x \left(\sup_{t \leq \delta} \rho(X_n(t), x) > \frac{\epsilon}{2} \right) < \frac{\epsilon}{2}$$

for all $x \in S_n \cap \Theta$. Let $t_0, \dots, t_k$ be numbers satisfying $0 = t_0 < \dots < t_k = 1$ and $|t_i - t_{i-1}| < \delta$ for each i . Then from the convergence of the finite dimensional distributions it follows that for sufficiently large n ,

$$P^{x_n}(X_n(t_i) \in S \setminus \Theta \text{ for some } i \leq k) < \epsilon/2.$$

Now, by conditioning on the hitting time of $S \setminus K$, we see that

$$P^{x_n}(X_n(t) \in S_n \setminus K \text{ for some } t \leq 1, X_n(t_i) \in \Theta \cup \{\Delta\} \text{ for every } i \leq k) < \frac{\epsilon}{2}.$$

So, the conclusion of the lemma is true for sufficiently large n . After making K larger, it will be true for all n .

Theorem 3. Assume that $X_n \rightarrow X$ and take $x_n \in S_n$ so that $x_n \rightarrow x \in S$. Then $(X_n \mid X_n(0) = x_n)$ converges weakly in $D[0, 1]$ to $(X \mid X(0) = x)$.

Proof. Since the processes are Markovian and the transition functions converge, it follows immediately that the finite dimensional distributions of X_n converge to those of X . So, it suffices to show that the processes X_n are tight in $D[0, 1]$. Since we have already obtained Lemma 4, we must show only that given $\epsilon > 0$, there is a $\delta > 0$ and an N so that for $n > N$, $P^{x_n}(\Lambda_n) \leq \epsilon$ where $\Lambda_n = \{\sup [|X_n(s) - X_n(t)|; s, t \leq 1, |s - t| \leq \delta, X_n(s) \in S_n, X_n(t) \in S_n] \geq \epsilon\}$.

By Lemma 4, we may choose a compact $K \subset S$ so that

$$P^{x_n}(X_n(t) \in S_n \setminus K \text{ for some } t \leq 1) < \frac{\epsilon}{2}.$$

Define the stopping times T_k^n by $T_0^n = 0$ and

$$T_k^n = \inf \left\{ t > T_{k-1}^n; |X_n(t) - X_n(T_{k-1}^n)| \geq \frac{\epsilon}{2} \right\},$$

for $k \geq 1$, where the death point Δ is interpreted as being infinitely far from points in the state space. Then

$$\begin{aligned}
P^{x_n}(\Lambda_n) &\leq \frac{\epsilon}{2} + P^{x_n}(\exists k \ni |T_k^n - T_{k-1}^n| \leq \delta, T_{k-1}^n \leq 1; X_n(t) \notin S_n \setminus K \forall t \leq 1) \\
&\leq \frac{\epsilon}{2} + \sum_{k=1}^{\infty} P^{x_n}(|T_k^n - T_{k-1}^n| \leq \delta, T_{k-1}^n \leq 1, X_n(T_{k-1}^n) \in K) \\
&= \frac{\epsilon}{2} + \sum_{k=1}^{\infty} E^{x_n}[T_{k-1}^n \leq 1, X_n(T_{k-1}^n) \in K, P^{X_n(T_{k-1}^n)}(T_1^n \leq \delta)].
\end{aligned}$$

From Lemma 3 and Chebyshev's Inequality, it follows that there is a constant η so that

$$E^x[\min(T_1^n, 1)] \geq \eta,$$

for all n and all $x \in S_n \cap K$. Using this lemma again, there is a $\delta \in (0, 1)$ so that

$$P^x(T_1^n \leq \delta) \leq \frac{\eta\epsilon}{4},$$

for all n and all $x \in S_n \cap K$. So,

$$P^{x_n}(\Lambda_n) \leq \frac{\epsilon}{2} + \frac{\eta\epsilon}{4} \sum_{k=1}^{\infty} P^{x_n}(T_{k-1}^n \leq 1, X_n(T_{k-1}^n) \in K).$$

Now, define random variables D_k^n by

$$D_k^n = \begin{cases} (T_k^n - T_{k-1}^n) & \text{if } T_k^n - T_{k-1}^n \leq 1, \\ 1 & \text{otherwise,} \end{cases}$$

and $R_k^n = D_1^n + \cdots + D_k^n$. By conditioning on $X_n(T_{k-1}^n)$, we now obtain

$$\begin{aligned}
P^{x_n}(T_{k-1}^n \leq 1, X_n(T_{k-1}^n) \in K) &\leq P^{x_n}(R_{k-1}^n \leq 1, X_n(T_{k-1}^n) \in K) \\
&\leq \frac{1}{\eta} E^{x_n}[D_k^n, R_{k-1}^n \leq 1].
\end{aligned}$$

However,

$$\sum_{k=1}^{\infty} E^{x_n}[D_k^n, R_{k-1}^n \leq 1] = \sum_{k=1}^{\infty} E^{x_n}[R_k^n, R_{k-1}^n \leq 1 < R_k^n] \leq 2.$$

So, $P^{x_n}(\Lambda_n) \leq \epsilon$ as required.

Remark. It should be noted that, while the results of Sections 2 and 3 of this paper depend strongly on the assumption that the state spaces are one dimensional, the results of Section 4 apply equally well to diffusion processes in higher dimensional Euclidean spaces. The proofs carry over to this more general setting with only minor changes.

The author would like to thank Professor C. Stone for helpful conversations concerning the results of Section 4.

REFERENCES

- [1] P. BILLINGSLEY, *Convergence of Probability Measures*, John Wiley and Sons, Inc., New York, 1968.

- [2] K. ITO & H. P. MCKEAN, *Diffusion Processes and their Sample Paths*, Springer-Verlag, Berlin, 1965.
- [3] S. KARLIN & J. MCGREGOR, Coincidence probabilities, *Pacific J. Math.*, **9** (1959) 1141–1164.
- [4] T. M. LIGGETT, An invariance principle for conditioned sums of independent random variables, *J. Math. Mech.*, **18** (1968) 559–570.
- [5] T. M. LIGGETT, Weak convergence of conditioned sums of independent random vectors. To appear in *Trans. A.M.S.*
- [6] C. STONE, Limit theorems for random walks, birth and death processes, and diffusion processes, *Illinois J. Math.*, **7** (1963) 638–660.

University of California, Los Angeles
Date communicated: NOVEMBER 7, 1969