The Annals of Mathematical Statistics
1971, Vol. 42, No. 3, 991-1002

COMPARISON OF SEMI-MARKOV AND MARKOV PROCESSES[1]

BY THOMAS G. KURTZ
University of Wisconsin

Conditions are given under which a semi-Markov process $Z(t)$ can be obtained from a Markov process $Y(t)$ by a time change (i.e. $Z(t) = Y(\gamma(t))$). Estimates are given for $P\{\sup_{s \leqq t} |s - \gamma(s)| > \varepsilon\}$ and the construction is used to give conditions under which a sequence of semi-Markov processes will have the same convergence properties as the corresponding sequence of Markov processes.

1. Introduction. In the study of jump Markov processes, the fact that the waiting time between jumps is exponentially distributed ordinarily plays a major role. However, if the process proceeds by a series of small, rapid jumps it is reasonable to expect (from law of large number considerations) that the importance of the distributions of the waiting times is minimized at least as far as they affect the finite dimensional distributions of the process. We will demonstrate that this is in fact the case, by showing under reasonable assumptions that a semi-Markov process (roughly, a process that behaves like a Markov process except that the waiting times are not exponentially distributed) is "close" to a corresponding Markov process provided the expected waiting time in each state is small.

We consider semi-Markov and Markov processes of the following type: let $X(0), X(1), \cdots$ be a stationary discrete parameter Markov process with state space $(E, \mathscr{B})$ and transition function

$$\mu(x, \Gamma) = P\{X(k+1) \in \Gamma \mid X(k) = x\}, \qquad x \in E, \Gamma \in \mathscr{B},$$

and let $\tau_0, \tau_1, \cdots$ be nonnegative random variables satisfying

$$P\{\tau_k \leqq t \mid X(0), X(1), \cdots, X(k), X(k+1), \tau_0, \tau_1, \cdots, \tau_{k-1}\} = F(t, X(k), X(k+1)).$$

In addition, we suppose

$$0 < h(x, z) \equiv E(\tau_k \mid X(k) = x, X(k+1) = z) < \infty \tag{1.1}$$

and

$$0 < h(x) \equiv E(\tau_k \mid X(k) = x) < \infty. \tag{1.2}$$

We define a semi-Markov process $Z(t)$ by

$$Z(t) = X(k) \quad \text{if} \quad \textstyle\sum_{l=0}^{k-1} \tau_l \leqq t < \sum_{l=0}^{k} \tau_l. \tag{1.3}$$

Received February 13, 1970; revised December 2, 1970.

[1] Research supported in part by the NIH at the University of Wisconsin.

Of course $Z(t)$ is defined only for

$$t < \sum_{l=0}^{\infty} \tau_l.$$

Let $\Delta_0, \Delta_1, \cdots$ be independent exponentially distributed random variables with $E(\Delta_i) = 1$ which are in addition independent of $X(0), X(1), \cdots$. We note that $h(X(k))\Delta_k$ satisfies

$$P(h(X(k))\Delta_k \leqq t \mid X(0), \cdots, X(k), \Delta_0, \cdots, \Delta_{k-1}) = 1-\exp[-t/h(X(k))],$$

and we define a Markov process $Y(t)$ by

$$(1.4) \qquad Y(t) = X(k) \quad \text{if} \quad \sum_{l=0}^{k-1} h(X(l))\Delta_l \leqq t < \sum_{l=0}^{k} h(X(l))\Delta_l.$$

As before, $Y(t)$ is defined only for

$$t < \sum_{l=0}^{\infty} h(X(l))\Delta_l.$$

We are interested in the relationship between semi-Markov processes of the type defined in (1.3) and Markov processes of the type defined in (1.4).

More precisely, let $(\Omega, \mathscr{F})$ be a measurable space, $\mathscr{F}_k$, $k = 0, 1, 2, \cdots$ σ-algebras with $\mathscr{F}_k \subset \mathscr{F}_{k+1} \subset \mathscr{F}$ and let $P_{x,z}$, $x, z \in E$ be probability measures on $\mathscr{F}$. Let $X(0), X(1), \cdots$ be E-valued random variables and suppose

$$P_{x,z}\{X(0) = x, X(1) = z\} = 1,$$

$$P_{x,z}\{X(k) \in \Gamma_0, X(k+1) \in \Gamma_1, \cdots, X(k+n) \in \Gamma_n \mid \mathscr{F}_{k+1}\}$$
$$= P_{X(k),X(k+1)}\{X(0) \in \Gamma_0, X(1) \in \Gamma_1, \cdots, X(n) \in \Gamma_n\}$$

and

$$P_{x,z}\{X(k+1) \in \Gamma \mid \mathscr{F}_k\} = \mu(X(k), \Gamma) \quad \text{for} \quad k \geqq 1.$$

Defining $P_x(A) = \int_E P_{x,z}(A)\mu(x, dz)$ for all $A \in \mathscr{F}$, it follows that

$$P_x\{X(0) = x\} = 1$$

and

$$P_x\{X(k) \in \Gamma_0, X(k+1) \in \Gamma_1, \cdots, X(k+n) \in \Gamma_n \mid \mathscr{F}_k\}$$
$$= P_{X(k)}\{X(0) \in \Gamma_0, X(1) \in \Gamma_1, \cdots, X(n) \in \Gamma_n\}.$$

Suppose there are random variables $\Delta_0, \Delta_1, \cdots$ defined on $(\Omega, \mathscr{F})$ such that

$$P_{x,z}\{\Delta_l > t\} = e^{-t}$$

and $\Delta_0, \Delta_1, \cdots$ are mutually independent and independent of $X(0), X(1), \cdots$ for all of the probability measures $P_{x,z}$.

Let $m(t)$ satisfy

$$\sum_{l=0}^{m(t)-1} h(X(l))\Delta_l \leqq t < \sum_{l=0}^{m(t)} h(X(l))\Delta_l$$

and define

$$Y(t) = X(m(t))$$

and

$$\xi(t) = (X(m(t)), X(m(t)+1)).$$

Let
$$\mathcal{N}_t = \cap_{x,z}\, \sigma(m(t), \Delta_0, \cdots, \Delta_{m(t)-1}, X(0), \cdots, X(m(t)))|^{P_{x,z}}$$
and
$$\mathcal{M}_t = \cap_{x,z}\, \sigma(m(t), \Delta_0, \cdots, \Delta_{m(t)-1}, X(0), \cdots, X(m(t)), X(m(t)+1))|^{P_{x,z}}$$
where $\sigma(\ )|^{P_{x,z}}$ denotes the completion of the σ-algebra with respect to $P_{x,z}$. We have
$$P_x\{Y(t+s) \in \Gamma \mid \mathcal{N}_t\} = P_{Y(t)}\{Y(s) \in \Gamma\}$$
and
$$P_{x,z}\{\xi(t+s) \in \Gamma_1 \times \Gamma_2 \mid \mathcal{M}_t\} = P_{\xi(t)}\{\xi(s) \in \Gamma_1 \times \Gamma_2\}.$$
That is, $Y(t)$ is a Markov process with state space $(E, \mathcal{B})$ and $\xi(t)$ is a Markov process with state space $(E \times E, \mathcal{B} \times \mathcal{B})$.

In Section 2 we show that we can define a version of $Z(t)$ on $(\Omega, \mathcal{F}, P_\xi)$ for all $\xi \in E \times E$ and hence on $(\Omega, \mathcal{F}, P_x)$ for all $x \in E$, and that in fact the version can be written as
$$Z(t) = Y(\gamma(t))$$
where $\gamma(t)$, while not in general a stopping time for $\{\mathcal{N}_t\}$ is a stopping time for $\{\mathcal{M}_t\}$.

In Section 3, we examine the relationship between t and $\gamma(t)$, which allows us, in Section 4, to draw a number of conclusions about the behavior of sequences of semi-Markov processes based on knowledge of the corresponding Markov processes.

2. The time change. We will need the following lemma.

LEMMA 2.1. *Let Δ be an exponentially distributed random variable with $E(\Delta) = 1$, and let $F(t)$ be a right-continuous distribution function. Then there exists an increasing right continuous function $G(u)$ such that*
$$P\{G(\Delta) \leqq t\} = F(t).$$

PROOF. Take
$$G(u) = \inf\{s: -\log(1-F(s)) > u\}.$$
Then
$$\{\Delta < -\log(1-F(t))\} \subset \{G(\Delta\} \leqq t\} \subset \{\Delta \leqq -\log(1-F(t))\},$$
and
$$P\{\Delta < -\log(1-F(t))\} = P\{\Delta \leqq -\log(1-F(t))\} = F(t).$$

By Lemma 2.1 there exist functions $G(u, x, z)$ such that
$$P\{G(\Delta, x, z) \leqq t\} = F(th(x, z), x, z).$$
If we define
$$\tau_k = h(X(k), X(k+1))G(\Delta_k, X(k), X(k+1)),$$
we see that
$$P_\xi\{\tau_k \leqq t \mid X(0), X(1), \cdots, X(k), X(k+1), \tau_0, \tau_1, \cdots, \tau_{k-1}\}$$
$$= F(t, X(k), X(k+1)),$$

and hence we can define a version of $Z(t)$ on the probability space $(\Omega, \mathscr{F}, P_\xi)$.

We also note that

$$E_\xi(G(\Delta_k, X(k), X(k+1)) \mid X(k), X(k+1)) = 1$$

where E_ξ denotes the expectation with respect to P_ξ.

Let $l(t)$ and $m(t)$ be the integers that satisfy

$$\sum_{k=0}^{l(t)-1} h(X(k), X(k+1))G(\Delta_k, X(k), X(k+1)) \leqq t < \sum_{k=0}^{l(t)} h(X(k), X(k+1)) \cdot G(\Delta_k, X(k), X(k+1)),$$

and

$$\sum_{k=0}^{m(t)-1} h(X(k))\Delta_k \leqq t < \sum_{k=0}^{m(t)} h(X(k))\Delta_k.$$

We note that $Z(t) = X(l(t))$ and $Y(t) = X(m(t))$.

We want to prove the following:

THEOREM 2.2. *The semi-Markov process* $Z(t)$ *may be represented as* $Z(t) = Y(\gamma(t))$ *where, for each* t, $\gamma(t)$ *is a stopping time for* $\{\mathscr{M}_t\}$.

PROOF. We want to find $\gamma(t)$ such that $l(t) = m(\gamma(t))$ and $\{\gamma(t) \leqq u\} \in \mathscr{M}_u$ for all $u \geqq 0$. In order that $l(t) = m(\gamma(t))$ we must have

$$\gamma(t) \in \mathscr{S}_t \equiv \{s: \sum_{k=0}^{m(s)-1} \tau_k \leqq t < \sum_{k=0}^{m(s)} \tau_k\}.$$

Let

$$\eta_u = (u - \sum_{k=0}^{m(u)-1} h(X(k))\Delta_k)/h(X(m(u))).$$

By the definition of $m(u)$, $\Delta_{m(u)} > \eta_u$. Let

$$\zeta_u = h(X(m(u)), X(m(u)+1))G(\eta_u, X(m(u)), X(m(u)+1)).$$

Since G need not be strictly increasing we can only say

$$\tau_{m(u)} \geqq \zeta_u.$$

Let

$$\hat{\mathscr{S}}_t = \{s: \sum_{k=0}^{m(s)-1} \tau_k \leqq t < \sum_{k=0}^{m(s)-1} \tau_k + \zeta_s\}.$$

We have $\hat{\mathscr{S}}_t \subset \mathscr{S}_t$ and, since $\sum_{k=0}^{m(s)-1} \tau_k$ and ζ_s are $\mathscr{M}_s$ measurable

$$\{\hat{\mathscr{S}}_t \cap [0, u] \neq \varnothing\} \in \mathscr{M}_u.$$

Define

$$\begin{aligned} \gamma(t) &= \inf\{s \in \hat{\mathscr{S}}_t\} \quad &\text{if } \hat{\mathscr{S}}_t \neq \varnothing \\ &= \inf\{s \in \mathscr{S}_t\} \quad &\text{if } \hat{\mathscr{S}}_t = \varnothing. \end{aligned}$$

Observe that $\hat{\mathscr{S}}_t = \varnothing$ implies $\Delta_{l(t)}$ is a discontinuity point of $G(\cdot, X(l(t)), X(l(t)+1))$. Consequently the event $\{\hat{\mathscr{S}}_t = \varnothing\}$ has probability zero for all P_ξ and is in $\mathscr{M}_u$. Therefore

$$\begin{aligned} \{\gamma(t) \leqq u\} &= \{\gamma(t) \leqq u, \hat{\mathscr{S}}_t \neq \varnothing\} \cup \{\gamma(t) \leqq u, \hat{\mathscr{S}}_t = \varnothing\} \\ &= \{\hat{\mathscr{S}}_t \cap [0, u] \neq \varnothing\} \cup \{\gamma(t) \leqq u, \hat{\mathscr{S}}_t = \varnothing\} \in \mathscr{M}_u, \end{aligned}$$

and $\gamma(t)$ is a stopping time for $\{\mathscr{M}_t\}$.

3. Comparison of t and $\gamma(t)$. We know that

$$\sum_{k=0}^{l(t)-1} h(X(k))\Delta_k \leqq \gamma(t) < \sum_{k=0}^{l(t)} h(X(k))\Delta_k \tag{3.1}$$

and

$$\sum_{k=0}^{l(t)-1} \tau_k \leqq t < \sum_{k=0}^{l(t)} \tau_k. \tag{3.2}$$

We want to find a bound on $P_\xi\{\sup_{s \leqq t} |\gamma(s)-s| > \delta\}$ for $\delta > 0$. We will do this by comparing the sums in (3.1) and (3.2) to sums of $h(X(k))$. First we will prove the following lemma.

LEMMA 3.3. *Let a_k and Z_k be random variables on $(\Omega, \mathscr{F}, P_\xi)$ with $a_k \geqq 0$. Let $\{\mathscr{G}_k\}$ be an increasing family of σ-algebras, $\mathscr{G}_k \subset \mathscr{F}$, such that $a_0, \cdots, a_{k+1}$ and $Z_0, \cdots, Z_k$ are $\mathscr{G}_k$-measurable. Suppose there exist a constant η and a function $F(t)$ satisfying*

$$\lim_{t\to\infty} F(t) = 0$$

and

$$\left|\int_0^\infty t\, dF(t)\right| < \infty$$

such that

$$P\{|Z_{k+1}| > t \mid \mathscr{G}_k\} \leqq F(t) \text{ a.s.}$$

and

$$a_k \leqq \eta \text{ a.s.}$$

and that

$$E(Z_{k+1} \mid \mathscr{G}_k) = 0.$$

If there exists an M such that $\sum_{k=1}^n a_k \leqq M$, then

$$P\{\sup_{k\leqq n} |\textstyle\sum_{l=1}^k a_l Z_l| > 2\varepsilon\} \leqq c/1-c$$

where

$$c = MC(\varepsilon, \eta, F)$$

and $C(\varepsilon, \eta, F)$ is a function depending only on ε, η, and F with

$$\lim_{\eta\to 0} C(\varepsilon, \eta, F) = 0$$

for fixed ε and F.

PROOF. The lemma follows from a lemma of Skorokhod (see Breiman [2], page 45 for a proof in the case of sums of independent random variables that can easily be generalized to the present case) provided we can show that

$$\sup_{k\leqq n} P\{|\textstyle\sum_{l=k+1}^n a_l Z_l| > \varepsilon \mid \mathscr{G}_k\} \leqq c \quad \text{a.s.}$$

To see this consider

$$|E(\exp\{i\theta \textstyle\sum_{l=k+1}^n a_l Z_l\} - 1 \mid \mathscr{G}_k)| \leqq E(\sum_{l=k+1}^n |E(\exp\{i\theta a_l Z_l\} - 1 \mid \mathscr{G}_{l-1})| \mid \mathscr{G}_k).$$

Since $E(Z_l \mid \mathscr{G}_{l-1}) = 0$ and a_l is $\mathscr{G}_{l-1}$-measurable

$$\begin{aligned} &|E(\exp\{i\theta a_l Z_l\} - 1 \mid \mathscr{G}_{l-1})| \\ &\quad = |E(\exp\{i\theta a_l Z_l\} - i\theta a_l Z_l - 1 \mid \mathscr{G}_{l-1})| \\ &\quad \leqq K_1 E(|\theta a_l Z_l|; |Z_l| > 1/a_l \mid \mathscr{G}_{l-1}) + K_2 E(\theta^2 a_l^2 Z_l^2; |Z_l| \leqq 1/a_l \mid \mathscr{G}_{l-1}), \end{aligned}$$

where

$$K_1 = \sup_x \left|\frac{e^{ix}-ix-1}{x}\right|$$

and

$$K_2 = \sup_x \left|\frac{e^{ix}-ix-1}{x^2}\right|.$$

It follows that

$$|E(\exp\{i\theta a_l Z_l\}-1 \,|\mathcal{G}_{l-1})| \leqq K_1|\theta|a_l|\int_{1/\eta}^{\infty} u\, dF(u)| + K_2\theta^2 a_l \sup_{a\leqq\eta}\int_0^{1/a} au^2\, dF(u)$$
$$\equiv a_l K(\theta,\eta,F).$$

Note that

$$\lim_{\eta\to 0} K(\theta, \eta, F) = 0.$$

This gives

$$|E(\exp\{i\theta \textstyle\sum_{l=k+1}^{n} a_l Z_l\}-1 \,|\, \mathcal{G}_k)| \leqq E(\sum_{l=k+1}^{n} a_l K(\theta, \eta, F)) \leqq MK(\theta, \eta, F).$$

The rest follows from standard inequalities involving characteristic functions.

THEOREM 3.4. *Let $F(t)$ be a function satisfying*

(3.5) $\quad F(t) \geqq \sup_x \max (P_x\{|\tau_0/h(x)-1| > t\}, P_x\{|\tau_0/h(x)-\Delta_0| > t\}),$

and let

$$\eta = \sup_x h(x).$$

Suppose $\eta < \infty$ and

$$|\textstyle\int_0^{\infty} t\, dF(t)| < \infty.$$

Then there exists a function $B(\varepsilon, \eta, t, F)$ such that

$$P_\xi\{\sup_{s\leqq t} |\gamma(s)-s| > 3\varepsilon\} < B(\varepsilon, \eta, t, F)$$

and

$$\lim_{\eta\to 0} B(\varepsilon, \eta, t, F) = 0.$$

PROOF. Let $n(t)$ satisfy

$$\textstyle\sum_{k=0}^{n(t)-1} h(X(k)) \leqq t < \sum_{k=0}^{n(t)} h(X(k)).$$

Define

$$\mathcal{G}_k = \sigma(\Delta_0, \cdots, \Delta_k, X(0), \cdots, X(k+1))$$
$$a_k = h(X(k)) \qquad k \leqq n(t)$$
$$= 0 \qquad k > n(t),$$
$$Z_k = \tau_k/h(X(k))-1.$$

Noting that

$$\textstyle\sum_{k=0}^{n} a_k \leqq t+\eta,$$

we apply Lemma 3.3 and obtain

$$(3.6)\qquad P_\xi\{\sup_{k\leqq n(t)} |\textstyle\sum_{l=0}^{k}(\tau_l - h(X(l)))| > 2\varepsilon\} \leqq \frac{(t+\eta)C(\varepsilon,\eta,F)}{1-(t+\eta)C(\varepsilon,\eta,F)}.$$

Keeping $\mathscr{G}_k$ and a_k the same, but defining

$$Z_k = \tau_k/h(X(k)) - \Delta_k$$

we obtain

$$(3.7)\qquad P_\xi\{\sup_{k\leqq n(t)} |\textstyle\sum_{l=0}^{k}(\tau_k - h(X(k))\Delta_k)| > 2\varepsilon\} \leqq \frac{(t+\eta)C(\varepsilon,\eta,F)}{1-(t+\eta)C(\varepsilon,\eta,F)}.$$

Observe that

$$\begin{aligned}\{l(t) > n(t+2\varepsilon)\} &= \{\textstyle\sum_{k=0}^{n(t+2\varepsilon)} \tau_k \leqq t\}\\ &= \{\textstyle\sum_{k=0}^{n(t+2\varepsilon)}(\tau_k - h(X(k))) \leqq t - \sum_{k=0}^{n(t+2\varepsilon)} h(X(k))\}\\ &\subset \{\textstyle\sum_{k=0}^{n(i+2\varepsilon)}(\tau_k - h(X(k))) < -2\varepsilon\}.\end{aligned}$$

Consequently by (3.6)

$$(3.8)\qquad P_\xi\{l(t) > n(t+2\varepsilon)\} \leqq \frac{(t+2\varepsilon+\eta)C(\varepsilon,\eta,F)}{1-(t+2\varepsilon+\eta)C(\varepsilon,\eta,F)}.$$

A small amount of calculus yields

$$(3.9)\qquad \begin{aligned} &P_\xi\{\max_{k\leqq n(t)} h(X(k))\Delta_k > \varepsilon\}\\ &\quad= E_\xi(1 - \Pi_{k\leqq n(t)}(1-\exp\{-\varepsilon/h(X(k))\}))\\ &\quad\leqq 1-\exp\{a(\varepsilon,\eta)\ln(1-\exp\{-a(\varepsilon,\eta)\})(t+\eta)/\varepsilon\}\\ &\quad\equiv D(\varepsilon,\eta,t),\end{aligned}$$

where $a(\varepsilon, \eta) = \max(\varepsilon/\eta, \ln 2)$. We note that

$$\lim_{x\to\infty} x\ln(1-e^{-x}) = 0,$$

and hence

$$\lim_{\eta\to 0} D(\varepsilon,\eta,t) = 0.$$

Consider the event

$$\{\sup_{k\leqq n(t+2\varepsilon)} |\textstyle\sum_{l=0}^{k}(\tau_k - h(X(k))\Delta_k)| \leqq 2\varepsilon\}\cap\{l(t) \leqq n(t+2\varepsilon)\}$$
$$\cap\{\max_{k\leqq n(t+2\varepsilon)} h(X(k))\Delta_k \leqq \varepsilon\}.$$

Using (3.2), we see that in the above event

$$\textstyle\sum_{k=0}^{l(s)-1} h(X(k))\Delta_k - 2\varepsilon \leqq t < \sum_{k=0}^{l(s)} h(X(k))\Delta_k + 2\varepsilon, \qquad \text{all}\quad s \leqq t.$$

Comparing this with (3.1) and using the fact that $h(X(l(s)))\Delta_{l(s)} \leqq \varepsilon$ we have

$$\sup_{s\leqq t} |\gamma(s) - s| \leqq 3\varepsilon.$$

Therefore combining (3.7), (3.8) and (3.9) we have

$$P_\xi\{\sup_{s\leqq t}|\gamma(s)-s| > 3\varepsilon\} \leqq 2\,\frac{(t+2\varepsilon+\eta)C(\varepsilon,\eta,F)}{1-(t+2\varepsilon+\eta)C(\varepsilon,\eta,F)}+D(\varepsilon,\eta,t+2\varepsilon)$$
$$\equiv B(\varepsilon,\eta,t,F).$$

4. Comparison of $Y(t)$ and $Z(t)$. It is easy to check that if

$$P\{\tau_k \leqq t \mid X(0),\cdots,X(k),X(k+1),\tau_1,\cdots,\tau_{k-1}\} = F(t,X(k)), \tag{4.1}$$

(that is τ_k depends only on $X(k)$, not $X(k+1)$), then $\gamma(t)$ is a stopping time for Y. If in addition $E_y(\gamma(t)) < \infty$ then by Dynkin's formula

$$E_y(f(Y(\gamma(t))))-f(y) = E_y\left(\int_0^{\gamma(t)} \tilde{A}f(Y(s))\,ds\right)$$

for every $f\in\mathscr{D}(\tilde{A}_Y)$, and hence

$$E_y(f(Z(t)))-E_y(f(Y(t))) = E_y\left(\int_t^{\gamma(t)} \tilde{A}f(Y(s))\,ds\right). \tag{4.2}$$

REMARK. If $f\in\mathscr{D}(\tilde{A}_Y)$ then

$$\tilde{A}_Yf(x) = \int_E (f(z)-f(x))\mu(x,dz)/h(x).$$

If

$$P\left\{\sum_{k=0}^\infty h(X(k))\Delta_k = \infty\right\} = 1$$

then

$$\mathscr{D}(\tilde{A}) = \left\{f: f \text{ bounded, measurable, } \sup_x \left|\int_E (f(z)-f(x))\mu(x,dz)/h(x)\right| < \infty\right\}.$$

This expression gives one means of comparing the behavior of a sequence of semi-Markov processes with that of a sequence of Markov processes. The results of the last section can be applied using the following:

THEOREM 4.3. *Assuming* (4.1), *for every* $f\in\mathscr{D}(\tilde{A}_Y)$ *and every* $\varepsilon > 0$,

$$\sup_{s\leqq t}|E_y(f(Z(s)))-E_y(f(Y(s)))|$$
$$\leqq (2\|f\|+t\|\tilde{A}_Yf\|)P_y\{\sup_{s\leqq t}|\gamma(s)-s| > \varepsilon\}+\varepsilon\|\tilde{A}_Yf\|.$$

PROOF. In order to avoid the assumption $E_y(\gamma(s)) < \infty$, let $\tau = \gamma(s)\wedge(s+\varepsilon)$. Then

$$|E_y(f(Z(s)))-E_y(f(Y(s)))|$$
$$\leqq 2\|f\|P_y\{\gamma(s)-s > \varepsilon\}+|E_y(f(Y(\tau)))-E_y(f(Y(s)))|$$
$$= 2\|f\|P_y\{\gamma(s)-s > \varepsilon\}+\left|E_y\left(\int_s^\tau \tilde{A}_Yf(Y(u))\,du\right)\right|$$
$$\leqq 2\|f\|P_y\{\gamma(s)-s > \varepsilon\}+\varepsilon\|\tilde{A}_Yf\|+s\|\tilde{A}_Y\|P\{s-\gamma(s) > \varepsilon\}.$$

In the general case we have the following theorem relating the behavior of a sequence of semi-Markov processes to the corresponding sequence of Markov processes.

THEOREM 4.4. *Let* $\{Z_n(t)\}$ *be a sequence of semi-Markov processes of the type defined in* (1.3) *taking values in a separable, locally compact, Hausdorff space* E, *and let* $\{Y_n(t)\}$ *be the related pure jump Markov processes. Let* $h_n(x)$ *be the waiting time expectation function and* $\tilde{A}_n$ *the weak infinitesimal operator for* Y_n. *Suppose there exists a single function* $F(t)$ *satisfying the conditions of Theorem* 3.4 *for all pairs* $(Y_n(t), Z_n(t))$ *and that*

$$\lim_{n\to\infty} \eta_n \equiv \lim_{n\to\infty} \sup_x h_n(x) = 0.$$

Then the following hold:

(a) *If the sequence* $\{Y_n\}$ *is tight in the Skorokhod topology on* $D[0, T]$ *for every* $T > 0$ *then* $\{Z_n\}$ *is also tight.* ($D[0, T]$ *is the space of E-valued right continuous functions on* $[0, T]$ *having left-hand limits. See Billingsley* [1] *or Parthasarathy* [4].)

If, in addition $P_y\{\tau_k^n > 0\} = 1$ *for all* k *and* n, *where* τ_k^n *is the* kth *waiting time for* Z_n, *then tightness of* $\{Z_n\}$ *in* $D[0, T]$ *for all* $T > 0$ *implies tightness for* $\{Y_n\}$

(b) *Let* $\rho(x, y)$ *be a metric on* E. *Suppose for all* $\varepsilon, \delta > 0$ *and* $x \in E$ *there are* $f_{n,x} \in \mathscr{D}(\tilde{A}_n)$ *such that*

$$-\delta \leqq f_{n,x}(y) \leqq 1 \qquad \text{for all } \; y,$$

$$f_{n,x}(y) \geqq 1-\delta \qquad \text{for } \; \rho(x, y) < \varepsilon/4,$$

$$f_{n,x}(y) \leqq \delta \qquad \text{for } \; \rho(x, y) \geqq \varepsilon/2,$$

and that

$$\sup_n \|\tilde{A}_n f_{n,x}\| < \infty.$$

Then, if the finite dimensional distributions (f.d.d.) *of* $\{Y_n\}$ *converge weakly to the* f.d.d. *of a process* Y, *the* f.d.d. *of* Z_n *converge to the* f.d.d. *of* Y.

REMARK. We observe that the conditions of part (b) are satisfied if the closure of $\{f: \sup_n \|\tilde{A}_n f\| < \infty\}$ contains the continuous functions with compact support.

PROOF. (a) A sequence of stochastic processes $\{W_n\}$ is tight in $D(0, T]$ if and only if

(i) for every $\eta > 0$ there is a compact set K such that $P_y\{W_n(t) \in K \text{ all } t \leqq T\} > 1-\eta$, for every $n \geqq 1$, and

(ii) for all $\varepsilon > 0$ and $\eta > 0$ there is a $\delta > 0$ and an integer N such that

$$P_y\{\omega'(\delta, T, W_n) \geqq \varepsilon\} \leqq \eta \quad \text{for every} \quad n \geqq N,$$

where $\omega'(\delta, T, y(\cdot))$ is defined as follows:

Let $S_\delta(T) = \{\{t_i\}: 0 = t_0 < t_1 < \cdots < t_k = T, t_{i+1}-t_i > \delta\}$. Then

$$\omega'(\delta, T, y(\cdot)) = \inf_{\{t_i\} \in S_\delta(T)} \max_i \sup_{t_i \leqq s,\, t \leqq t_{i+1}} \rho(y(s), y(t)).$$

Let $\gamma_n(t)$ be the stopping time satisfying $Z_n(t) = Y_n(\gamma_n(t))$. Theorem 3.4 implies for every $\varepsilon > 0$

$$\lim_{n\to\infty} P\{\sup_{t\leqq T} |\gamma_n(t)-t| > 3\varepsilon\} \leqq \lim_{n\to\infty} B(\varepsilon, \eta_n, T, F) = 0. \tag{4.5}$$

Consequently, if condition (i) holds for $\{Y_n\}$ and every $T > 0$ then it holds for $\{Z_n\}$ and every $T > 0$. Since $Z_n(t) = Y_n(\gamma_n(t))$ it follows that

$$\omega'(\delta, T, Z_n) \leqq \omega'(3\delta, T+\delta, Y_n)$$

on the set $\{\sup_{s \leqq T+\delta} |\gamma_n(s)-s| \leqq \delta\}$. Consequently, (4.5) implies that if condition (ii) holds for $\{Y_n\}$ and every $T > 0$ then it holds for $\{Z_n\}$ and every $T > 0$.

Assume now that $P_y\{\tau_k^n > 0\} = 1$ for all k. This is the same as assuming that Z_n has the same sequence of states as Y_n. It follows from the definition of $\gamma_n(t)$ and the assumption that $P_y\{\tau_k^n > 0\} = 1$, that

$$\gamma_n(\textstyle\sum_{k=0}^{l} \tau_k^n) = \sum_{k=0}^{l} h_n(X_n(k))\Delta_k. \tag{4.6}$$

Defining

$$\hat{\gamma}_n(s) = \inf\{t : \gamma_n(t) \geqq s\},$$

it follows from (4.6) that

$$\hat{\gamma}_n(\textstyle\sum_{k=0}^{l} h_n(X_n(k))\Delta_k) = \sum_{k=0}^{l} \tau_k^n$$

and since $\hat{\gamma}_n(s)$ is monotone

$$Y_n(s) = Z_n(\hat{\gamma}_n(s)). \tag{4.7}$$

Noting that $\sup_{s \leqq T+2\delta} |\gamma_n(s)-s| \leqq \delta$ implies $\sup_{s \leqq T+\delta} |\hat{\gamma}_n(s)-s| \leqq \delta$, we have

$$\omega'(\delta, T, Y_n) \leqq \omega'(3\delta, T+\delta, Z_n)$$

on the set $\{\sup_{s \leqq T+2\delta} |\gamma_n(s)-s| \leqq \delta\}$, and hence if condition (ii) holds for $\{Z_n\}$ and every $T > 0$ then it holds for $\{Y_n\}$ and every $T > 0$.

(b) We first prove the following lemma:

LEMMA 4.8. *Under the conditions of part* (b)

$$\lim_{\eta \to 0} \limsup_{n \to \infty} \sup_t \sup_y P_y\{Y_n(t) \in K, \quad \sup_{t \leqq s \leqq t+\eta} \rho(Y_n(t), Y_n(s)) > \varepsilon\} = 0$$

for every compact set K.

PROOF. Given $\varepsilon, \delta > 0$ and $x \in E$, let $\tau = \inf\{t : \rho(Y_n(0), Y_n(t)) > \varepsilon\}$. By Dynkin's formula we have

$$f_{n,x}(y) - E_y(f_{n,x}(Y_n(\tau \wedge t))) = -E_y(\textstyle\int_0^{\tau \wedge t} \tilde{A}_n f_{n,x}(Y_n(s))\, ds).$$

Hence for $\rho(x, y) < \varepsilon/4$

$$-E_y(\textstyle\int_0^{\tau \wedge t} \tilde{A}_n f_{n,x}(Y_n(s))\, ds) \geqq 1-\delta-P_y\{\tau > t\}-\delta P_y\{\tau \leqq t\}.$$

Therefore

$$P_y\{\tau \leqq t\} \leqq \frac{t\|\tilde{A}f_{n,x}\| + \delta}{1-\delta}$$

for every y such that $\rho(y, x) < \varepsilon/4$. Since any compact set can be covered by a finite number of spheres of radius $\varepsilon/4$, for every compact set K and any $\delta > 0$ there is a constant $M(K, \delta)$ such that

$$\sup_{y \in K} P_y\{\tau \leqq t\} \leqq \frac{tM(K, \delta)+\delta}{1-\delta}$$

The lemma follows, using the Markov property.

A similar argument gives

LEMMA 4.9. *Under the conditions of part* (b)

$$\lim_{\eta \to 0} \limsup_{n \to \infty} \sup_y P_y\{\inf_{x \in K} \rho(Y_n(t-\eta), x) > \varepsilon,\ Y_n(t) \in K\} = 0,$$

for every compact set K.

PROOF. Suppose $\rho(x, y) > \varepsilon$ and let

$$\tau_x = \inf\{t : \rho(x,\ Y_n(t)) < \varepsilon/4\}.$$

Observe that

$$(1-\delta)P_y\{\tau_x \leqq t\} \leqq E_y(f_{n,x}(Y(t \wedge \tau_x)))+\delta \leqq \delta + t\|\tilde{A}_n f_{n,x}\|.$$

Since K is compact there are $x_1, \cdots, x_k \in K$ with $K \subset \cup_{i=1}^k \{z : \rho(x_i, z) < \varepsilon/4\}$. We note that

$$\begin{aligned} P_y\{\inf_{x \in K} \rho(Y_n(t-\eta), x) > \varepsilon,\ Y_n(t) \in K\} &\leqq \sum_{i=1}^k P_y\{\inf_{x \in K} \rho(Y_n(t-\eta), x) > \varepsilon,\ \tau_{x_i} \leqq \eta\} \\ &\leqq \sum_{i=1}^k \frac{\delta + \eta\|\tilde{A}_n f_{n,x_i}\|}{1-\delta}, \end{aligned}$$

and the lemma follows.

Combining the results of these lemmas with (4.5) gives

$$\lim_{n \to \infty} P_y\{Y_n(t) \in K,\ \rho(Y_n(t), Z_n(t)) > \varepsilon\} = 0, \tag{4.10}$$

and part (b) follows.

REMARK. Theorem 4.4 extends much of the work that has been done on limits of sequences of pure jump Markov processes to sequences of semi-Markov processes. For examples of the work done on Markov processes see Trotter [6], Stone [5], or Kurtz [3]. In each of these papers sequences of jump Markov processes (either discrete or continuous parameter) are shown to converge to diffusion processes (in the case of [3] to solutions of ordinary differential equations). Theorem 4.4 shows that altering the shape of the waiting time distributions of the approximating processes (within the restrictions of Theorem 3.4) does not affect the nature of the convergence.

REFERENCES

[1] BILLINGSLEY, PATRICK (1968). *Convergence of Probability Measures.* Wiley, New York.
[2] BREIMAN, LEO (1968). *Probability.* Addison-Wesley, Reading, Mass.
[3] KURTZ, THOMAS, G. (1970). Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Probability* **7** 49–58.
[4] PARTHASARATHY, K. R. (1967). *Probability Measures on Metric Spaces.* Academic Press, New York.
[5] STONE, CHARLES (1963). Limit theorems for random walks, birth and death processes, and diffusion processes. *Illinois J. Math.* **7** 638–660.
[6] TROTTER, H. F. (1958). Approximation of semigroups of operators. *Pacific J. Math.* **8** 887–919.