BULL. AUSTRAL. MATH. SOC. 60J05, 60K25
VOL. 17 (1977), 97-107.

Hitting times of Markov chains, with application to state-dependent queues

R.L. Tweedie

We present in this note a useful extension of the criteria given in a recent paper [*Advances in Appl. Probability* 8 (1976), 737-771] for the finiteness of hitting times and mean hitting times of a Markov chain on sets in its (general) state space. We illustrate our results by giving conditions for the finiteness of the mean number of customers in the busy period of a queue in which both the service-times and the arrival process may depend on the waiting time in the queue. Such conditions also suffice for the embedded waiting time chain to have a unique stationary distribution.

1. Finiteness criteria for hitting times

We give our results for Markov chains on a general state space X, with a σ-field F of subsets of X, since the proofs require no other structure. Let $\{X_n\}$ be such a chain, with temporally homogeneous transition probabilities

$$P^n(x, A) = \Pr\left(X_n \in A \mid X_0 = x\right) , \quad A \in F , \quad x \in X ;$$

we assume $P^n(x, \cdot)$ is a measure on F for each $x \in X$, and $P^n(\cdot, A)$ is a measurable function on X for each $A \in F$. We let Δ denote a particular set in F, and let

$$N = \inf\left(n > 0 : X_n \in \Delta\right)$$

Received 16 February 1977. The author is grateful to Dr Mark Westcott for emphasising the desirability of the generalisation of Theorem 1 contained in Theorem 2, and for much useful discussion on the paper.

denote the number of steps until $\{X_n\}$ enters Δ. In [10], Theorems 6.1 and 10.1, the following results are proved.

THEOREM 1. *(i) If there exists a non-negative measurable function* g *and an* $\varepsilon > 0$ *such that*

$$\int_X P(x, dy)g(y) \le g(x) - \varepsilon \tag{1}$$

for all $x \in \Delta^c$, *then*

$$E(N \mid X_0 = x) \le g(x)/\varepsilon , \quad x \in \Delta^c , \tag{2}$$

$$E(N \mid X_0 = x) \le 1 + \int_X P(x, dy)g(y)/\varepsilon , \quad x \in \Delta . \tag{3}$$

(ii) If there exists a non-negative measurable function g *such that*

$$g(x) > \sup_{y \in \Delta} g(y) , \quad x \in \Delta^c , \tag{4}$$

$$\Pr\left(\limsup g(X_n) = \infty \mid X_0 = x\right) \equiv 1 , \tag{5}$$

and

$$\int_X P(x, dy)g(y) \le g(x) \tag{6}$$

for all $x \in \Delta^c$, *then*

$$\Pr(N < \infty \mid X_0 = x) \equiv 1 . \tag{7}$$

It is often the case in practice that the "negative mean drift" conditions (1) or (6) can only be shown to hold for x outside some larger set Γ containing Δ, and yet it is reasonable that if Γ is reached in a finite (or finite mean) time, then the same is true for Δ. A typical case is when $X = \{0, 1, \ldots\}$ and $\Delta = \{0\}$, but (1) or (6) only hold for x outside $\{0, 1, \ldots, j\}$ where j is sufficiently large. For this case (when also $\{X_n\}$ is assumed irreducible), Foster [2] first proved Theorem 1 with $\Delta = \{0\}$, whilst Pakes [6] gave auxiliary conditions for the above type of generalization. This countable case is a simple corollary of our results below. (I am grateful to Dr Pakes for pointing out to me that

https://doi.org/10.1017/S0004972700025508 Published online by Cambridge University Press

Foster himself gave the more general result, in the discussion of [3]; it is surprising that in [2] Foster only presented the form with Δ .)

The extension of Theorem 1 which we prove is

THEOREM 2. *(i) Suppose* $\Delta \subseteq \Gamma$ *and* g *is a non-negative function such that*

$$\int_X P(x, dy)g(y) \le g(x) - \varepsilon , \quad x \in \Gamma^c . \tag{8}$$

If further

$$\sup_{x\in\Gamma\backslash\Delta} \left[\int_X P(x, dy)g(y)-g(x)\right] = B < \infty , \tag{9}$$

and for some $n > 0$,

$$\inf_{x\in\Gamma\backslash\Delta} \sum_1^n P^m(x, \Delta) = \delta > 0 , \tag{10}$$

then for some constant $M \ge 0$,

$$E\left(N \mid X_0 = x\right) \le [M+g(x)]/\varepsilon , \quad x \in \Delta^c . \tag{11}$$

(ii) Suppose $\Delta \subseteq \Gamma$, *and* (4)-(6) *hold for* Γ *rather than* Δ . *If further*

$$\inf_{x\in\Gamma\backslash\Delta} P\left(N < \infty \mid X_0 = x\right) = \lambda > 0 , \tag{12}$$

then (7) *holds. A sufficient condition for* (12) *is that, for some* $\theta < 1$,

$$\inf_{x\in\Gamma\backslash\Delta} \sum_1^\infty P^m(x, \Delta)\theta^m = \lambda' > 0 , \tag{13}$$

so that in particular (10) *suffices for* (12) *to hold.*

Proof. *(i)* We show that (8)-(10) enable us to construct a function g^* which satisfies (1), and is bounded above by $M + g$ for some constant M . The result then follows from Theorem 1 *(i)*.

We will assume that for $x \in \Delta$, $P(x, \{x\}) = 1$: this affects neither the hypothesis nor the conclusion of the theorem. From (10), for every $x \in \Gamma\backslash\Delta$ there exists m with $1 \le m \le n$ such that $P^m(x, \Delta) \ge \delta/n$, and hence, since Δ is absorbing, $P^n(x, \Delta) \ge \delta/n$. Let

$\beta < 1$ be small enough that $n\beta \le \delta/n$, and put $\Gamma_0 = \Delta$ and for $k = 1, \ldots, n$,

$$\Gamma_k = \{y : P^k(y, \Delta) \ge k\beta\} .$$

Note that $\Gamma \backslash \Delta \subseteq \Gamma_n$. Now by construction, if $x \in \Gamma_{k+1}$, $k = 1, 2, \ldots, n-1$, we have

$$(k+1)\beta \le P^{k+1}(x, \Delta) = \int_{\Gamma_k} P(x, dy)P^k(y, \Delta) + \int_{\Gamma_k^c} P(x, dy)P^k(y, \Delta)$$

$$\le P(x, \Gamma_k) + k\beta ,$$

so that for $k = 0, 1, \ldots, n-1$,

$$(14) \qquad P(x, \Gamma_k) \ge \beta , \quad x \in \Gamma_{k+1} .$$

Put $M = n(B+\varepsilon)/\beta^n$, where B, ε are as in (8) and (9); and define g^* by setting $\Gamma_k^* = \Gamma_k \backslash \left[\bigcup_{j=0}^{k-1} \Gamma_j\right]$, $k = 1, \ldots, n$, and putting

$$g^*(x) = \begin{cases} g(x) , & x \in \Delta , \\ g(x) + M - [n-k][B+\varepsilon]/\beta^{n-k} , & x \in \Gamma_k^* , \quad k = 1, \ldots, n , \\ g(x) + M , & x \in \left[\bigcup_0^n \Gamma_j^*\right]^c . \end{cases}$$

With this construction, we have for any $x \in \Gamma_k^*$, $k = 1, \ldots, n$, from (9) and (14),

$$(15) \quad \int_X P(x, dy)g^*(y) - g^*(x)$$

$$\le \int_X P(x, dy)[g(y)+M] - P(x, \Gamma_{k-1})(n-k+1)[B+\varepsilon]/\beta^{n-k+1}$$

$$- g(x) - M + (n-k)[B+\varepsilon]/\beta^{n-k}$$

$$\le B - [B+\varepsilon]/\beta^{n-k} \le -\varepsilon .$$

Since $\Gamma \backslash \Delta \subseteq \bigcup_k \Gamma_k^*$, for $x \in \left[\bigcup \Gamma_j^*\right]^c$, we have (8) holding, so that

https://doi.org/10.1017/S0004972700025508 Published online by Cambridge University Press

$$\int_X P(x, dy)g^*(y) \le \int_X P(x, dy)[g(y)+M] \le g(x) - \varepsilon + M = g^*(x) - \varepsilon . \tag{16}$$

From (15) and (16), we see that the function g^* satisfies (1) for every $x \notin \Delta$, and the theorem holds.

(ii) Again, let us consider Δ to be absorbing. From (4)-(6) we have that

$$\Pr\left(X_n \in \Gamma, \text{ some } n > 0 \mid X_0 = x\right) \equiv 1 ;$$

this together with (12), implies, for $x \in \Delta^c$,

$$\inf_{x\in\Delta^c} P\left(N < \infty \mid X_0 = x\right) = \lambda > 0 . \tag{17}$$

From (17), then, we have that

$$\inf_{x\in X} P\left(N < \infty \mid X_0 = x\right) = \lambda > 0 ,$$

since $P\left(N = 1 \mid X_0 = x\right) = 1$, $x \in \Delta$; now from Proposition 1.5.1 of [5], it follows that for all $x \in X$,

$$\Pr\left(X_n \in \Delta \text{ infinitely often} \mid X_0 = x\right) = \Pr\left(X_n \in X \text{ infinitely often} \mid X_0 = x\right) \equiv 1 ,$$

and so, in particular, (7) holds. That (13) implies (12) is shown in (1.5) of [8].

REMARKS. (i) Clearly if $\Gamma\backslash\Delta$ is finite, as it is in the countable state space extension of Foster's result mentioned earlier, then (10) holds when $\sum_1^\infty P^m(x, \Delta) > 0$ for each $x \in \Gamma\backslash\Delta$, and so in particular when $\{X_n\}$ is irreducible. In this case, (9) is also equivalent to

$$\int_X P(x, dy)g(y) < \infty , \quad x \in \Gamma\backslash\Delta .$$

Hence Pakes' results [6] follow from ours.

(ii) If Δ is a single point, then the assumptions of either part of

Theorem 2 ensure that the chain $\{X_n\}$ is ϕ-irreducible, with ϕ concentrated at Δ (*cf.* [5]). It then follows that, if (10) holds, Γ is a *status set* for $\{X_n\}$ (Proposition 5.2 of [*10*]), and so, as shown in [*10*], (8) and (9) suffice to ensure that

(a) a stationary distribution π exists for $\{X_n\}$; and

(b) for π-almost all x , $E(N \mid X_0 = x) < \infty$.

However, (b) may be considerably weaker than the conclusion of Theorem 2 *(i)*: consider, for example, the case when Δ actually *is* absorbing, so that π is concentrated at Δ and (b) implies nothing about the mean hitting times on this absorbing set from points in Δ^c .

(iii) To provide counter examples to possible weakenings of the hypotheses of Theorem 2, consider a chain where $\Gamma\backslash\Delta$ contains a countably infinite number of points x_j , $j = 1, 2, \ldots$, and

$$P\left(x_j, \{x_{j+1}\}\right) = \alpha_j = 1 - P\left(x_j, \Delta\right) .$$

Even if (8), (9) hold, then $E\left(N \mid X_0 = x\right)$ can be infinite for $x \in \Gamma\backslash\Delta$ if $\sum j\left[1-\alpha_j\right] = \infty$; if this happens then (10) of course fails.

Similarly, if $\prod \alpha_j > 0$, the hypothesis and conclusion of Theorem 2 *(ii)* both fail; in this case, also, if Δ is absorbing $\sum P^n(x, \Delta) = \infty$, which indicates that $\theta = 1$ cannot be allowed in (13).

The condition (9) cannot be weakened, in general (unless $\Gamma\backslash\Delta$ is finite), to

$$(18) \qquad \int_X P(x, dy)g(y) < \infty , \quad x \in \Gamma\backslash\Delta ;$$

for an example where (18) and (8) hold but the mean hitting times on Γ itself are infinite, see Section 6 of [9].

2. State-dependent queues

In this section we apply our results to a very general state-dependent queueing model. We consider a system where a customer, arriving to find a

https://doi.org/10.1017/S0004972700025508 Published online by Cambridge University Press

waiting time w before he is served, has a service time with a distribution S_w depending on w, with mean μ_w; and where also, if a customer arrives to find a waiting time w, then the time T until the next customer arrives has a distribution depending on the new virtual waiting time after his service time has been added to w; if this service time is s, then we write

$$\Pr(T \leq x \mid w+s) = F(x \mid w+s) .$$

This allows for a wide variety of input mechanisms: for example, if the input is a Poisson process with rate λ, and if a customer arriving to find a waiting time w turns away with probability $p(w)$, then for $x < w + s$,

$$F(x \mid w+s) = \exp\left\{-\int_{w+s-x}^{w+s} \lambda[1-p(u)]du\right\} .$$

This illustrates the fact that allowing the interarrival time to depend on the virtual waiting time immediately after the last customer arrives also allows the interarrival time to depend on all the waiting times after that point; we do, however, assume that the interarrival times are independent of the past before the last customer arrived, and the service time of a customer is independent of the time he takes to arrive, and all other service times, once w is prescribed.

When the interarrival time process is independent of the waiting time process, such queues have been considered in [7], [1], [9], and [4]. In our application of Theorem 2, we find conditions for the finiteness of the mean number of customers in the busy period of such a queue; these lead to conditions for the embedded waiting time chain to have a stationary distribution.

We let X_n denote the waiting time in the queue immediately before the arrival of the nth customer. The queue-size at this time is no longer a Markov chain, since we have dependence of the system on waiting times; however, $\{X_n\}$ is a Markov chain on $([0, \infty), \mathcal{B}^+)$, where $\mathcal{B}^+$ denotes the Borel subsets of $[0, \infty)$; and we can exploit the fact that, if $\Delta = \{0\}$ and $N = \inf(n > 0 : X_n \in \Delta)$ as in the previous sections, then the mean number of customers in a busy period is merely

$E(N \mid X_0 = 0)$.

THEOREM 3. *(a) The mean number of customers given* $X_0 = x > 0$ *in the above queueing model satisfies*

$$E(N \mid X_0 = x) \le a + bx \ , \quad x > 0 \ ,$$

for some $a \ge 0$, $b \ge 0$, *provided*

(i) $$\limsup_{w\to\infty} \left[\mu_w - \int_0^\infty \int_0^{w+y} t dF(t \mid w+y) dS_w(y)\right] < 0 \ ;$$

(ii) for any $K > 0$,

$$\sup_{0<w\le K} \mu_w = \bar{\mu}(K) < \infty \ ;$$

(iii) for any $K > 0$, *there exists* $\delta(K) > 0$ *and* $\varepsilon(K) > 0$ *such that for every* $x \in [0, K]$, *either*

$$F\left(\bar{\mu}(K)+\delta(K) \mid x\right) < 1 - \varepsilon(K) \ ,$$

or

$$F(x \mid x) < 1 - \varepsilon(K) \ .$$

(b) Consequently the mean number of customers in a busy period satisfies $E(N \mid X_0 = 0) \le [1+a] + b\mu_0$, *and so is finite if* $\mu_0 < \infty$.

Proof. Let $\{P(x, A)\}$ be the transition law of the waiting time chain $\{X_n\}$. From *(i)*, there exists a $K_0 > 0$ and an $\varepsilon > 0$ such that for $x > K_0$,

$$\int_0^\infty P(x, dy)y - x \le \mu_x - \int_0^\infty \int_0^{w+y} t dF(t \mid w+y) dS_w(y) \le -\varepsilon \ ;$$

hence (8) holds with $g(x) = x$ and $\Gamma = \left(0, K_0\right]$. From *(ii)* at K_0 , (9) holds with $\Delta = \{0\}$. Also from *(ii)*, for $w \in \left(0, K_0\right]$, there must be a $\beta > 0$ such that for each w ,

$$S_w\left(K_0+\bar{\mu}\left(K_0\right)\right) \ge \beta \ ;$$

that is, with probability bounded away from zero the jumps out of $\left(0, K_0\right]$

stay below $K_0 + \bar{\mu}(K_0) = K_1$, say. Now applying *(iii)* for $K = K_1$, we find that for each $x \in (0, K_1]$, there is probability, bounded from zero by $\varepsilon(K_1)$, either of no jump until zero is reached, or of waiting at least $\bar{\mu}(K_1) + \delta(K_1)$ before the next jump; that is, of the waiting time decreasing again by at least $\bar{\mu}(K_1) + \delta(K_1)$ before the next jump. Since $\bar{\mu}$ is an increasing function, there is thus probability bounded from zero by $\beta\varepsilon(K_1)$ that $X_{n+1} \le X_n - \delta(K_1)$, given $X_n \in [\delta(K_1), K_0]$. Repeating this $K_0/\delta(K_1)$ times, we see that by choosing $n = K_0/\delta(K_1)$, we have from *(ii)* and *(iii)* that (10) holds. Hence for $x > 0$, the mean of the hitting time N on zero from x is bounded above by $a + bx$, for some $a \ge 0$, $b \ge 0$, from Theorem 2 *(i)*.

Now finally, we have that the mean number of customers in a busy period is given by

$$E(N \mid X_0 = 0) = 1 + \int_0^\infty E(N \mid X_0 = y)\,dS_0(y)$$
$$\le 1 + a + b\mu_0 ,$$

and the theorem is proved.

REMARKS. (i) In the case of an independent renewal input process, as studied in the references cited above, the condition *(iii)* of the theorem is not a great improvement on the condition that the (independent and identically distributed) interarrival times T be unbounded; that is, $F(x) < 1$ for all x . However, in one case it is a distinct improvement: this occurs when

$$\sup_{0<w<\infty} \mu_w = \bar{\mu}(\infty) < \infty ,$$

for then *(iii)* gives our results provided for some $\delta > 0$,

$$F(x) < 1 , \quad x < \bar{\mu}(\infty) + \delta .$$

(ii) In [11], Theorem 2 is applied to continuous time processes. There, we derive extra conditions under which the mean length of a busy period (in contrast to the mean number of customers) can be shown to be finite.

Finally, we note that since Δ is a single point, as mentioned in Remark (ii) after the proof of Theorem 2 we can also show

THEOREM 4. *If the conditions of Theorem* 3 *hold, then the waiting time chain* $\{X_n\}$ *admits a unique stationary distribution* π *; and provided, say, for every* $x \leq \mu_0$ *, there is* $\delta > 0$ *,* $\varepsilon > 0$ *such that*

$$F\left(\mu_0+\delta \mid x\right) < 1 - \varepsilon , \tag{24}$$

then for any initial distribution λ *of waiting times,*

$$\sup_{A \in B^+} \left| \int_0^\infty \lambda(dy) P^n(y, A) - \pi(A) \right| \to 0 , \quad n \to \infty . \tag{25}$$

If (24) *fails then at least*

$$\sup_{A \in B^+} \left| \int_0^\infty \lambda(dy) \left[\frac{1}{n} \sum_1^n P^m(y, A)\right] - \pi(A) \right| \to 0 , \quad n \to \infty . \tag{26}$$

Proof. From Theorem 3, if ϕ is a measure concentrated at $\{0\}$, the chain is ϕ-recurrent [5], and thus has a unique invariant measure π , which must be finite since $E\left(N \mid X_0 = x\right) < \infty$, $x > 0$, and $E\left(N \mid X_0 = 0\right) < \infty$ (see [*10*] or [*4*]). From Theorem 1.7.1 of [5], then, (26) holds, and can be strengthened to (25) if the chain is aperiodic. The condition (24) ensures (as in the proof of Theorem 3), that $P(0, \{0\}) > 0$, which is enough for aperiodicity.

Even when the arrival process is an independent renewal process, Theorem 4 improves the results of [*4*], which demanded either

(a) that the service times be continuously dependent on waiting times in some way; or

(b) that the service times be deterministic, but T have a density not concentrated on a finite interval.

Some of the cases when (a) holds may not be covered by Theorem 4, but (b) is superseded by the present approach.

References

[1] J.R. Callahan, "A queue with waiting time dependent service times", *Naval Res. Logist. Quart.* 20 (1973), 321-324.

[2] F.G. Foster, "On the stochastic matrices associated with certain queueing processes", *Ann. Math. Statist.* 24 (1953), 355-360.

[3] David G. Kendall, "Some problems in the theory of queues", *J. Roy. Statist. Soc. Ser. B* 13 (1951), 151-185.

[4] G.M. Laslett, D.B. Pollard, R.L. Tweedie, "Techniques for establishing ergodic and recurrence properties of continuous-valued Markov chains", submitted.

[5] Steven Orey, *Lecture notes on limit theorems for Markov chain transition probabilities* (Van Nostrand Reinhold, London, New York, Cincinnati, Toronto, Melbourne, 1971).

[6] A.G. Pakes, "Some conditions for ergodicity and recurrence of Markov chains", *Operations Res.* 17 (1969), 1058-1061.

[7] M. Posner, "Single-server queues with service time dependent on waiting time", *Operations Res.* 21 (1973), 610-616.

[8] Pekka Tuominen and Richard L. Tweedie, "Markov chains with continuous components", submitted.

[9] Richard L. Tweedie, "Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space", *Stochastic Processes Appl.* 3 (1975), 385-403.

[10] R.L. Tweedie, "Criteria for classifying general Markov chains", *Advances in Appl. Probability* 8 (1976), 737-771.

[11] R.L. Tweedie, M. Westcott, "First-passage times in skip-free processes", submitted.

Division of Mathematics and Statistics,
CSIRO,
Canberra, ACT.