J. Appl. Prob. **46,** 812–826 (2009)
Printed in England

MARKOV CHAINS CONDITIONED NEVER TO WAIT TOO LONG AT THE ORIGIN

SAUL JACKA,* *University of Warwick*

Abstract

Motivated by Feller's coin-tossing problem, we consider the problem of conditioning an irreducible Markov chain never to wait too long at 0. Denoting by τ the first time that the chain, X, waits for at least one unit of time at the origin, we consider conditioning the chain on the event $(\tau > T)$. We show that there is a weak limit as $T \to \infty$ in the cases where either the state space is finite or X is transient. We give sufficient conditions for the existence of a weak limit in other cases and show that we have vague convergence to a defective limit if the time to hit zero has a lighter tail than τ and τ is subexponential.

Keywords: Subexponential tail; evanescent process; Feller's coin-tossing constants; hitting probabilities; conditioned process

2000 Mathematics Subject Classification: Primary 60J27
Secondary 60J80; 60J50; 60B10

1. Introduction and notation

1.1. Introduction

Feller [4, Section XIII.7] showed that if $p_n^{(k)}$ is the probability that there is no run of heads of length k or more in n tosses of a fair coin, then, for a suitable positive constant c_k,

$$p_n^{(k)} \sim c_k s_k^{n+1},$$

where s_k is the largest real root in (0,1) of the equation

$$x^k - \sum_{j=0}^{k-1} 2^{-(j+1)} x^{k-1-j} = 0. \tag{1.1}$$

More generally, if the probability of a head is $p = 1 - q$, then the same asympotic formula is valid, with (1.1) modified to become

$$x^k - q \sum_{j=0}^{k-1} p^j x^{k-1-j} = 0,$$

and $c_k = (s_k - p)/q((k+1)s_k - k)$.

The continuous-time analogue of this question is to seek the asymptotic behaviour of the probability that Y, a Poisson process with rate r, has no interjump time exceeding one unit by

Received 1 September 2008; revision received 20 June 2009.
* Postal address: Department of Statistics, University of Warwick, Coventry CV4 7AL, UK.
Email address: s.d.jacka@warwick.ac.uk
I am grateful to an anonymous referee for many helpful suggestions on improving the presentation of this paper.

time T. It follows, essentially from Theorem 1.2, that, denoting by τ_Y the first time that Y waits to jump longer than one unit of time,

$$\mathrm{P}(\tau_Y > t) \sim c_r \mathrm{e}^{-\phi_r t}$$

for a suitable constant c_r, where $\phi_r = 1$ if $r = 1$, otherwise ϕ_r is the root (other than r itself) of the equation

$$x\mathrm{e}^{-x} = r\mathrm{e}^{-r}. \tag{1.2}$$

A natural extension is then to seek the tail behaviour of the distribution of $\tau \equiv \tau_X$, the first time that a Markov chain, X, waits longer than one unit of time at a distinguished state, 0. In general, there has also been much interest (see [1], [2], [3], [6], [7], [8], [9], [10], [11], [12], [13], [14], and [15]) in conditioning an evanescent Markov process X on its survival time being increasingly large and in seeing whether a weak limit exists.

1.2. Notation

We consider a continuous-time Markov chain X on a countable state space S, with a distinguished state ∂. We denote $S \setminus \{\partial\}$ by C. For convenience, and without loss of generality, we assume henceforth that $S = \mathbb{Z}^+$ or $S = \{0, \dots, n\}$ and $\partial = 0$ so that $C = \mathbb{N}$ or $C = \{1, \dots, n\}$.

We assume that X is irreducible, and nonexplosive. We denote the transition semigroup of X by $\{P(t);\, t \geq 0\}$ and its Q-matrix by $Q = (q_{ij})$, and set $q_i = -q_{ii}$. We define the process $\tilde{X}$ as X killed on first hitting 0, and we shall usually assume that $\tilde{X}$ is also irreducible on C. We denote the substochastic semigroup for $\tilde{X}$ by $(\tilde{P}(t))_{t\geq 0}$. We first define the return and departure epochs as follows:

$$S_0 = \inf\{t \geq 0 \colon X_t = 0\} \quad \text{and} \quad T_0 = \inf\{t \geq 0 \colon X_t \neq 0\},$$

and then, for $n \geq 1$,

$$T_n = \inf\{t \geq S_n \colon X_t \neq 0\}$$

and

$$S_n = \inf\{t \geq T_{n-1} \colon X_t = 0\}.$$

We then define the successive holding and return times $H^n_{n\geq 0}$ and $R^n_{n\geq 0}$ by

$$H^0 = T_0 \quad \text{and} \quad R^0 = S_0,$$

and

$$H^n = T_n - S_n \quad \text{and} \quad R^n = S_{n+1} - T_n \quad \text{for } n \geq 1.$$

Then we define the current holding time as follows:

$$H_t = t - S_n \quad \text{if} \quad S_n \leq t < T_n \quad \text{for some } n.$$

It will be convenient in what follows to define

$$H_t = \varnothing \quad \text{if} \quad X_t \neq 0.$$

We denote the first time that X waits in 0 for time 1 by τ, i.e.

$$\tau = \inf\{t \colon H_t \geq 1\},$$

https://doi.org/10.1239/jap/1253279853 Published online by Cambridge University Press

and denote the process X killed at time τ by $\hat{X}$. We denote the state space augmented by the current holding time in 0 by $\hat{S} := C \cup \{\{0\} \times [0, 1)\}$. By a slight abuse of notation, we denote the (substochastic) Markov chain $(\hat{X}_t, H_t)$ on the state space $\hat{S}$ by $\hat{X}$ also. (Note that if $\hat{X}_t \neq 0$ then $H_t = \varnothing$, so $(\hat{X}_t, H_t) = \hat{X}_t$.) The associated semigroup is denoted $(\hat{P}(t))_{t \geq 0}$. Throughout the rest of the paper, we denote by P_i the probability on Skorokhod path space $D(S, [0, \infty))$, conditional on $\hat{X}_0 = i$, and the corresponding filtration by $(\mathcal{F}_t)_{t\geq 0}$. Finally, we denote a typical hitting time of 0 from state i by $\tau_0^{(i)}$ and its density by ρ_i. We denote the density of a typical return time, R^1, by ρ.

1.3. Convergence/decay parameters for evanescent chains

We recall (see, for example, [7]) that, if X^* is a Markov chain on C, with substochastic transition semigroup P^* and Q-matrix $Q^* = (q^*_{ij})_{(i,j)\in C\times C}$, then X^* is said to be evanescent if it is irreducible and dies with probability 1. In this case we define

$$\alpha_{X^*} = \alpha = \inf\left\{\lambda \geq 0 \colon \int_0^\infty P^*_{ij}(t)\mathrm{e}^{\lambda t}\,\mathrm{d}t = \infty\right\}$$

for any $i, j \in C$, and (see, for example, [17]) X^* is classified as α-recurrent or α-transient depending on whether $\int_0^\infty P^*_{ij}(t)\mathrm{e}^{\alpha t}\,\mathrm{d}t = \infty$ or is finite. Moreover, X^* is α-recurrent if and only if $\int_0^\infty f^*_{ii}(t)\mathrm{e}^{\alpha t}\,\mathrm{d}t = 1$, where f^*_{ii} is the defective density of the first return time to i (starting in i).

In the α-recurrent case, X^* is α-positive recurrent if

$$\int_0^\infty t f^*_{ii}(t)\mathrm{e}^{\alpha t}\,\mathrm{d}t < \infty,$$

otherwise X^* is α-null recurrent. Defining $q^*_i = -q^*_{ii}$, it is easy to see that $\alpha < q^*_i$ for all $i \in \mathbb{N}$ and, hence,

$$0 \leq \alpha \leq \inf_i q^*_i.$$

Thus, α measures the rate of decay of transition probabilities (in C). There is a second decay parameter, μ^*, which measures the rate of dying.

We define τ^* as the death time of X^*, define $s^*_i(t) = \sum_j P^*_{ij}(t) = \mathrm{P}_i(\tau^* > t)$, and set

$$\mu^* = \inf\left\{\lambda \colon \int_0^\infty s^*_i(t)\mathrm{e}^{\lambda t}\,\mathrm{d}t = \infty\right\}.$$

Note that μ^* is independent of i by the usual irreducibility argument; moreover, since $1 \geq s^*_i(t) \geq P^*_{ii}(t)$, it follows that

$$0 \leq \mu^* \leq \alpha^*.$$

Note that in our current setting, we shall take $X^* = \tilde{X}$ and write $\tau^* = \tau_0$, the first hitting time of 0. We shall denote the rate of hitting 0, which is the death rate for X^*, by μ^C and α^* by α^C, and we shall denote the survival probabilities for $\tilde{X}$ by s^C, so that $s^C_i(t) = \mathrm{P}_i(\tau_0 > t)$.

1.4. Doob h-transforms

Recall (see, for example, [19, Section III.49]) that we may form the h-transform of a substochastic Markovian semigroup on S, $(P(t))_{t\geq 0}$, if $h \colon S \to \mathbb{R}^+$ is P-superharmonic (i.e. $[P(t)h](x) \leq h(x)$ for all $x \in S$ and all $t \geq 0$). The h-transform of P, P^h, is specified by its transition kernel, which is given by

$$P^h(x, \mathrm{d}y; t) := \frac{h(y)}{h(x)} P(x, \mathrm{d}y; t),$$

so that if we consider the corresponding substochastic measures on path space, P_x and P^h_x (conditional on $X_0 = x$), then

$$\frac{\mathrm{dP}^h_x}{\mathrm{dP}_x}\bigg|_{\mathcal{F}_t} = h(X_t)$$

and P^h forms another substochastic Markovian semigroup. If h is actually *space–time* P-superharmonic then appropriate changes need to be made to these definitions. In particular, if $h(x,t) = \mathrm{e}^{\phi t} h_x$ then

$$\frac{\mathrm{dP}^h_x}{\mathrm{dP}_x}\bigg|_{\mathcal{F}_t} = \mathrm{e}^{\phi t} h_{X_t}.$$

As shown in [7], in general, when a weak limit or a vague limit exists for the problem of interest, it must be a Doob-h-transform of the original process, with the state augmented by the current waiting time in state 0 in the case we study here.

1.5. Main results

Denoting by $\hat{P}(t)$ the substochastic transition semigroup for $\hat{X}$, we define

$$s_i(t) := \mathrm{P}_i(\tau > t) = \hat{P}(i, \hat{S}; t) \quad \text{for } i \in \hat{S}.$$

Our first result is as follows.

Theorem 1.1. *Suppose that X is transient. Denote P_i(X never hits 0) by β_i, and define $\Delta = \sum_{j \in C} q_{0,j}\beta_j / q_0$. Set*

$$p_{(0,0)} := p_0 = \frac{(1-\mathrm{e}^{-q_0})\Delta}{\mathrm{e}^{-q_0} + (1-\mathrm{e}^{-q_0})\Delta}, \tag{1.3}$$

$$p_{(0,u)} = \frac{1-\mathrm{e}^{-q_0(1-u)}}{1-\mathrm{e}^{-q_0}} p_0, \tag{1.4}$$

and

$$p_i = \beta_i + (1-\beta_i) p_0 \quad \textit{for } i \in C. \tag{1.5}$$

Then

$$s_i(t) \to p_i \quad \textit{as } t \to \infty \textit{ for all } i \in \hat{S}.$$

Hence, if we condition X on $\tau = \infty$, we obtain a new Markov process, X^∞, on $\hat{S}$ with honest semigroup P^∞ given by

$$P^\infty_{i,j}(t) = \frac{p_j}{p_i} \hat{P}_{i,j}(t) \quad \textit{for } j \in C \tag{1.6}$$

and

$$P^\infty_i((0,\,\mathrm{d}u); t) = \frac{p_{(0,u)}}{p_i} \hat{P}_i((0,\,\mathrm{d}u); t), \tag{1.7}$$

so that X^∞ looks like a Markov chain with Q-matrix given by $q^\infty_{i,j} = (p_j/p_i) q_{i,j}$ on C, whilst X^∞ has a holding time in 0 with density d given by

$$d(t) = \frac{\mathrm{e}^{-q_0 t}}{\int_0^1 \mathrm{e}^{-q_0 s}\,\mathrm{d}s} \mathbf{1}_{\{t<1\}}$$

and a time-homogeneous jump probability out of state 0 to state j of $q_{0,j} p_j / q_0 p_0$.

In the case where X is recurrent, it is clear that $s_i(t) \to 0$ as $t \to \infty$ for each $i \in \hat{S}$.

Now let $W := H^1 + R^1$ (so that W is the first return time of X to 0 from 0), and let g be the (defective) density of $W\,\mathbf{1}_{\{H^1<1\}}$ on $(0, \infty)$. Our first result under these conditions is as follows. It is a generalisation to our more complex setting of Seneta and Vere-Jones' [17] result in the α-positive case.

Theorem 1.2. *Let*

$$I(\lambda) := \int_0^\infty \mathrm{e}^{\lambda t} g(t)\,\mathrm{d}t = \mathrm{E}\mathrm{e}^{\lambda W}\,\mathbf{1}_{\{H^1<1\}}.$$

Then if

$$\text{there exists a } \phi \text{ such that } I(\phi) = 1, \text{ and } I'(\phi-) < \infty, \tag{1.8}$$

then, for each $i \in \hat{S}$,

$$\mathrm{e}^{\phi t} s_i(t) \to p_i > 0 \quad \text{as } t \to \infty,$$

where the function p *is given by*

$$p_{(0,0)} = \frac{\mathrm{e}^{\phi - q_0}}{\phi I'(\phi-)} =: \kappa, \tag{1.9}$$

$$p_{(0,u)} = \frac{\int_0^{1-u} \mathrm{e}^{(\phi-q_0)s}\,\mathrm{d}s}{\int_0^1 \mathrm{e}^{(\phi-q_0)s}\,\mathrm{d}s}\kappa, \tag{1.10}$$

and, for $i \neq 0$,

$$p_i = F_{i,0}(\phi)\kappa, \tag{1.11}$$

where

$$F_{i,0}(\lambda) := \mathrm{E}\mathrm{e}^{\lambda \tau_0^{(i)}} = \int_0^\infty \mathrm{e}^{\lambda t} \rho_i(t)\,\mathrm{d}t.$$

The following simple condition ensures that condition (1.8) holds.

Lemma 1.1. *Suppose that* $\tilde{X}$ *is* α*-recurrent, and that both* $N_0 := \{i : q_{i,0} > 0\}$ *and* $N_0^* := \{i : q_{0,i} > 0\}$ *are finite. Then (1.8) holds.*

Corollary 1.1. *Let* X^T *denote the chain on* $\hat{S}$ *obtained by conditioning* $\hat{X}$ *on the event* $(\tau > T)$*. Then, if condition (1.8) holds, for each* $s > 0$*, the restriction of the law of* X^T *to* $\mathcal{F}_s$ *converges weakly as* $T \to \infty$ *to that of* X^∞ *restricted to* $\mathcal{F}_s$*, where the transition semigroup of* X^∞ *is given by (1.6) and (1.7).*

In the case where $I(\phi) < 1$ or $I'(\phi-) = \infty$, Theorems 3.1, 3.2, and 3.3, below, (may) apply, giving some sufficient conditions for weak or vague convergence to take place. In Theorem 3.4 and Corollary 3.1, below, we give an application to the case of a recurrent birth-and-death process conditioned not to wait too long in state 0.

2. Proofs of the transient and α-positive cases

To prove Theorem 1.1 is straightforward.

Proof of Theorem 1.1. It is trivial to establish the equations

$$p_i = \beta_i + (1 - \beta_i) p_0$$

and

$$p_0 = (1 - \mathrm{e}^{-q_0}) \sum_{j \in C} \frac{q_{0,j}}{q_0} p_i.$$

Equations (1.3)–(1.5) follow immediately. Then the conditioning result follows straightforwardly.

Example 2.1. We take a transient nearest-neighbour random walk with reflection at 0 and an up-jump rate of b and a down-jump rate of d. Note that $1 - \beta$ is the minimal positive solution to $P(t)h = h$ with $h(0) = 1$, and that $1 - \beta_i = (d/b)^i$.

The main tool in the proof of Theorem 1.2 is the renewal theorem.

Proof of Theorem 1.2. Recall that state $(0, u)$ denotes that the killed chain is at 0 and its current holding time is u. First note that $s_{(0,0)}$ satisfies the renewal equation

$$s_{(0,0)}(t) = \left(1 - \int_0^\infty g(u)\,\mathrm{d}u\right) \mathbf{1}_{\{t<1\}} + \int_t^\infty g(u)\,\mathrm{d}u + \int_0^t g(u) s_{(0,0)}(t-u)\,\mathrm{d}u. \tag{2.1}$$

If we define

$$f(t) = \mathrm{e}^{\phi t} s_{(0,0)}(t),$$

it follows immediately from (2.1) that

$$f(t) = \mathrm{e}^{\phi t}\left(\left(1 - \int_0^\infty g(u)\,\mathrm{d}u\right)\mathbf{1}_{\{t<1\}} + \int_t^\infty g(u)\,\mathrm{d}u\right) + \int_0^t \tilde{g}(u) f(t-u)\,\mathrm{d}u,$$

where $\tilde{g}(t) := \mathrm{e}^{\phi t} g(t)$. Now, it is easy to check that the conditions of Feller's alternative formulation of the renewal theorem (see [5, Section XI.1]) are satisfied, so we conclude that

$$f(t) \to \mu^{-1} \int_0^\infty \mathrm{e}^{\phi t}\left(\left(\left(1 - \int_0^\infty g(u)\,\mathrm{d}u\right)\mathbf{1}_{\{t<1\}}\right) + \int_t^\infty g(u)\,\mathrm{d}u\right)\mathrm{d}t \quad \text{as } t \to \infty,$$

where

$$\mu = \int_0^\infty t\tilde{g}(t)\,\mathrm{d}t = I'(\phi-).$$

It is trivial to establish, by changing the order of integration, that

$$\begin{aligned}
\int_0^\infty \mathrm{e}^{\phi t} \int_t^\infty g(u)\,\mathrm{d}u\,\mathrm{d}t &= \int_0^\infty g(u) \int_0^u \mathrm{e}^{\phi t}\,\mathrm{d}t\,\mathrm{d}u \\
&= \frac{I(\phi) - \int_0^\infty g(u)\,\mathrm{d}u}{\phi} \\
&= \frac{1 - \int_0^1 q_0 \mathrm{e}^{-q_0 u}\,\mathrm{d}u}{\phi} \\
&= \frac{\mathrm{e}^{-q_0}}{\phi},
\end{aligned}$$

and, hence, (1.9) follows.

https://doi.org/10.1239/jap/1253279853 Published online by Cambridge University Press

To establish (1.11), note that (by conditioning on the time of the first hit of 0),

$$s_i(t) = \int_t^\infty \rho_i(u)\,\mathrm{d}u + \int_0^t \rho_i(u) s_{(0,0)}(t-u)\,\mathrm{d}u,$$

and so, denoting $\mathrm{e}^{\phi t} s_i(t)$ by $f_i(t)$, we obtain

$$f_i(t) = \mathrm{e}^{\phi t} \int_t^\infty \rho_i(u)\,\mathrm{d}u + \int_0^t \tilde{\rho}_i(u) f(t-u)\,\mathrm{d}u,$$

where $\tilde{\rho}_i(t) := \mathrm{e}^{\phi t}\rho_i(t)$. Now f is continuous and converges to κ so, by the dominated convergence theorem,

$$\int_0^t \tilde{\rho}_i(u) f(t-u)\,\mathrm{d}u \to \int_0^\infty \kappa \tilde{\rho}_i(u)\,\mathrm{d}u = \kappa F_{i,0}(\phi) \quad \text{as } t \to \infty.$$

Moreover, since $F_{i,0}(\phi) = \int_0^\infty \tilde{\rho}_i(u)\,\mathrm{d}u < \infty$, it follows that

$$\mathrm{e}^{\phi t} \int_t^\infty \rho_i(u)\,\mathrm{d}u \le \int_t^\infty \tilde{\rho}_i(u)\,\mathrm{d}u \to 0 \quad \text{as } t \to \infty,$$

and, hence,

$$f_i(t) \to \kappa F_{i,0}(\phi) \quad \text{as } t \to \infty,$$

as required.

To establish (1.10), observe that

$$s_{(0,u)}(t) = \mathrm{e}^{-q_0 t}\,\mathbf{1}_{\{t<1-u\}} + \int_0^t \int_0^{(1-u)\wedge v} q_0 \mathrm{e}^{-q_0 v} \rho(w-v) s_{(0,0)}(t-w)\,\mathrm{d}v\,\mathrm{d}w,$$

and, hence,

$$\begin{aligned} f_{(0,u)}(t) &:= \mathrm{e}^{\phi t} s_{(0,u)}(t) \\ &= \mathrm{e}^{(\phi-q_0)t}\,\mathbf{1}_{\{t<1-u\}} + \int_{w=0}^t \int_{v=0}^{(1-u)\wedge w} q_0 \mathrm{e}^{(\phi-q_0)v} \tilde{\rho}(w-v) f(t-w)\,\mathrm{d}v\,\mathrm{d}w, \end{aligned}$$

and, hence, by the dominated convergence theorem,

$$\begin{aligned} f_{(0,u)}(t) &\to \kappa \int_0^\infty \int_0^{w\wedge(1-u)} q_0 \mathrm{e}^{(\phi-q_0)v} \tilde{\rho}(w-v)\,\mathrm{d}v\,\mathrm{d}w \quad (t \to \infty) \\ &= \kappa \int_0^\infty \tilde{\rho}(t)\,\mathrm{d}t \int_0^{1-u} q_0 \mathrm{e}^{(\phi-q_0)v}\,\mathrm{d}v \\ &= \kappa \frac{\int_0^{1-u} \mathrm{e}^{(\phi-q_0)s}\,\mathrm{d}s}{\int_0^1 \mathrm{e}^{(\phi-q_0)s}\,\mathrm{d}s}, \end{aligned}$$

as required.

https://doi.org/10.1239/jap/1253279853 Published online by Cambridge University Press

Remark 2.1. Note that the case mentioned in the introduction, where Y is a Poisson(r) process and we let τ_Y be the first time that an interjump time is one or larger, can be addressed using the proof of Theorem 1.2. In this case, if we consider that the chain 'returns directly to 0' at each jump time Y then

$$I(\lambda) = \int_0^1 r\mathrm{e}^{(\lambda-r)t}\,\mathrm{d}t,$$

and so ϕ satisfies $r(\mathrm{e}^{\phi-r}-1)/(\phi-r)=1$, which establishes (1.2), and

$$\mathrm{e}^{\phi t}\mathrm{P}(\tau > t) \to \frac{\mathrm{e}^{\phi-r}}{\phi I'(\phi-)} = \frac{\phi-r}{r(\phi-1)} \quad \text{as } t\to\infty$$

for $r \neq 1$. The case in which $r=1$ gives $\phi=1$ and $c_1=2$.

Proof of Lemma 1.1. It follows from Theorem 3.3.2 of [8] that if N_0 is finite then $\alpha^C=\mu^C$. Now, since $\tilde{X}$ is α-recurrent, it follows that

$$\int_0^\infty \mathrm{e}^{\lambda t}\tilde{P}_{ii}(t)\,\mathrm{d}t < \infty \quad \text{if and only if } \lambda < \alpha^C.$$

Since $s_i^C(t) \geq \tilde{P}_{ii}(t)$, it follows that

$$\int_0^\infty \mathrm{e}^{\lambda t}s_i^C(t)\,\mathrm{d}t = \infty \quad \text{if } \lambda \geq \alpha^C.$$

Conversely, since $\alpha^C=\mu^C$, we see that

$$\int_0^\infty \mathrm{e}^{\lambda t}s_i^C(t)\,\mathrm{d}t < \infty \quad \text{if } \lambda < \alpha^C,$$

and so we conclude that

$$\int_0^\infty \mathrm{e}^{\lambda t}s_i^C(t)\,\mathrm{d}t < \infty \quad \text{if and only if } \lambda < \alpha^C.$$

Now

$$\begin{aligned} I(\lambda) &= \int_0^\infty \mathrm{e}^{\lambda t}g(t)\,\mathrm{d}t \\ &= \int_0^1 \mathrm{e}^{(\lambda-q_0)t}\,\mathrm{d}t\biggl(\sum_{i\in N_0^*} q_{0,i}F_{i,0}(\lambda)\biggr) \\ &= \int_0^1 \mathrm{e}^{(\lambda-q_0)t}\,\mathrm{d}t\biggl(\sum_{i\in N_0^*} q_{0,i}\biggl(\frac{F_{i,0}(\lambda)-1}{\lambda}\biggr)\biggr). \end{aligned}$$

Now $F_{i,0}(\lambda)<\infty$ if and only if $\lambda<\alpha^C$ and so, since N_0^* is finite by assumption, $I(\lambda)<\infty$ if and only if $\lambda<\alpha^C$. It now follows trivially that $\phi<\alpha^C$ and that (1.8) is satisfied.

Proof of Corollary 1.1. This follows immediately from Theorem 1.2 and Theorem 4.1.1 of [7] provided that we can show that h, given by $h\colon (i,t)\mapsto \mathrm{e}^{\phi t}p_i$, is $\hat{P}$-harmonic. This is easy to check by considering the chain at the epochs when it leaves and returns to 0, i.e. we show that, defining σ as the first exit time from 0, $\mathrm{E}_{(0,u)}h(\hat{X}_{t\wedge\sigma}, t\wedge\sigma) = h((0,u),0)$ and $\mathrm{E}_i h(\hat{X}_{t\wedge\tau_0}, t\wedge\tau_0) = h(i,0)$ for $i\in C$. This is sufficient since $\hat{X}$ is nonexplosive.

3. The α-transient case

We now seek to consider the α-transient case. In particular, we shall focus on the case where $\phi = 0$. This is not so specific as one might think since one can (at the cost of a slight extra difficulty) reduce the general case to that where $\phi = 0$.

3.1. Reducing to the case where $\phi = 0$

We discuss briefly how to transform the problem to this case.

The essential technique is to note that if, for any $\lambda \le \phi$, we h-transform $\hat{P}$ using the space–time $\hat{P}$-superharmonic function h^λ given by

$$h^\lambda(i,t) = F_{i,0}(\lambda)\mathrm{e}^{\lambda t} \quad \text{for } i \in C$$

and

$$h^\lambda((0,u),t) = \left(1 - I(\lambda)\frac{J^\lambda(u)}{J^\lambda(1)}\right)\mathrm{e}^{-(\lambda-q_0)u}\mathrm{e}^{\lambda t} \quad \text{for } u \in [0,1),$$

where

$$J^\lambda(x) := \int_0^x \mathrm{e}^{(\lambda-q_0)v}\,\mathrm{d}v,$$

then we obtain a new chain $\overline{X}$ on $\hat{S}$, with $\phi_{\overline{X}} = \phi - \lambda$ and satisfying $g_{\overline{X}}(t) = \mathrm{e}^{\lambda t} g(t)$, which dies only from state $(0, 1-)$.

The proof of this result uses the standard result that h^λ is space–time harmonic for $\hat{P}$ off $\{0\} \times [0,1)$, while, since $I(\lambda) < 1$, it is easy to see that h^λ is superharmonic on $\{0\} \times [0,1)$, by conditioning on the time of first exit from 0. Now it is easy to check that $\overline{X}$ dies only from state $(0, 1-)$ and dies on a visit to 0 with probability $1 - I(\lambda)$, so the result follows immediately.

Remark 3.1. Note that, in the α-null recurrent case, where $I(\phi) = 1$ but $I'(\phi-) = \infty$, the above transform produces a null recurrent h-transform when $\lambda = \phi$, whereas the transform is still evanescent in the α-transient case.

It will follow from l'Hôpital's theorem in the α-transient cases that if ψ_i denotes the density (on $(1,\infty)$) of τ when starting from state i, then, if $\psi_i(t-v)/\psi_j(t)$ has a limit as $t \to \infty$, it is the common limit of

$$\frac{s_i(t-v)}{s_j(t)} = \frac{\int_{t-v}^\infty \psi_i(u)\,\mathrm{d}u}{\int_t^\infty \psi_j(u)\,\mathrm{d}u} \quad \text{and} \quad \frac{h_i^\phi}{h_j^\phi}\frac{s_i^{h\phi}(t-v)}{s_j^{h\phi}(t)} = \frac{\int_{t-v}^\infty \mathrm{e}^{\phi u}\psi_i(u)\,\mathrm{d}u}{\int_t^\infty \mathrm{e}^{\phi u}\psi_j(u)\,\mathrm{d}u}.$$

In the α-null recurrent case we see that this is not of much help. It is not hard to generalise Lemma 3.3.3 of [15] to prove that in this case $(i,t) \mapsto \mathrm{e}^{\phi t}h_i^\phi$ is the unique $\hat{P}$-superharmonic function of the form $\mathrm{e}^{\lambda t}k_i$ and so gives the only possible weak or vague limit.

3.2. Heavy and subexponential tails

All the results quoted in this subsection, apart from the last, are taken from [18].

Recall first that a random variable (normally taking values in $\mathbb{R}^+$) Z, with distribution function F_Z, is said to be *heavy tailed*, or to have a heavy tail, if

$$\frac{\overline{F_Z}(t+s)}{\overline{F_Z}(t)} \to 1 \quad \text{as } t \to \infty \text{ for all } s \ge 0,$$

where $\overline{F_Z} := 1 - F_Z$ is the complementary distribution function.

Denoting the n-fold convolution of F_Z by F_Z^n, Z is said to have a *subexponential tail*, or just to be subexponential, if

$$\frac{\overline{F_Z^n}(t)}{\overline{F_Z}(t)} \to n \quad \text{as } t \to \infty \text{ for all } n, \tag{3.1}$$

and (3.1) holds if and only if

$$\limsup_{t\to\infty} \frac{\overline{F_Z^n}(t)}{\overline{F_Z}(t)} \le n \quad \text{for some } n \ge 2. \tag{3.2}$$

A subexponential random variable always has a heavy tail.

Two random variables, X and Y, are said to have comparable tails, or to be tail equivalent, if

$$\overline{F_Y}(t) \sim c\overline{F_X}(t) \quad \text{for some } c > 0.$$

The random variable Y is said to have a lighter tail than X if

$$\frac{\overline{F_Y}(t)}{\overline{F_X}(t)} \to 0 \quad \text{as } t \to \infty.$$

Lemma 3.1. *If X and Y are independent, Y is lighter tailed than X, and X has a subexponential tail, then $X + Y$ has a subexponential tail and*

$$\overline{F_{X+Y}}(t) \sim \overline{F_X}(t).$$

Lemma 3.2. *If X and Y are independent, subexponential, and tail equivalent with*

$$\overline{F_Y}(t) \sim c\overline{F_X}(t),$$

then $X + Y$ is subexponential and

$$\overline{F_{X+Y}}(t) \sim (1 + c)\overline{F_X}(t).$$

This generalises to the following random case.

Lemma 3.3. *Suppose that $X_1, \dots$ are independent and identically distributed with common distribution function F, which is subexponential, and that N is an independent geometric random variable. Then, if*

$$S := \sum_1^N X_i,$$

S is subexponential and

$$\overline{F_S}(t) \sim (\mathrm{E}N)\overline{F_X}(t).$$

Finally, we have the following lemma.

Lemma 3.4. *Suppose that $X_1, \dots$ are independent and tail equivalent with*

$$F_{X_i} := F_i,$$

and that J is an independent random variable taking values in $\mathbb{N}$. Let

$$Y = X_J$$

(so that Y is a mixture of the X_is), and denote its distribution function by F (so $F(t) = \sum_{i\in\mathbb{N}} \mathrm{P}(J = i)F_i(t)$).

Now suppose that

$$\overline{F_i}(t) \sim a_i \overline{F_1}(t).$$

If the collection $\{\overline{F_J}(t)/\overline{F_1}(t);\, t \geq 0\}$ is uniformly integrable (u.i.) then

$$\overline{F}(t) \sim (\mathrm{E}a_J)\overline{F_1}(t). \tag{3.3}$$

In particular, if J is a bounded random variable then (3.3) holds.

Proof. It follows from the assumptions that

$$\frac{\overline{F_J}(t)}{\overline{F_1}(t)} \to a_J \quad \text{almost surely as } t \to \infty.$$

Thus, if the collection is u.i. then convergence is also in L^1 and so, since $\mathrm{E}\overline{F_J}(t) = \overline{F}(t)$, we see that

$$\frac{\overline{F}(t)}{\overline{F_1}(t)} \to \mathrm{E}a_J \quad \text{as } t \to \infty.$$

In particular, if $J \leq n$ almost surely then

$$\limsup_{t\to\infty} \frac{\overline{F_J}(t)}{\overline{F_1}(t)} \leq \max_{1\leq i\leq n} a_i \quad \text{almost surely,}$$

and so the collection is indeed u.i.

3.3. Results for heavy tails

Suppose first that $0 = \phi < \mu^C$.

Theorem 3.1. *If $0 = \phi < \mu^C$ and τ is subexponential, then $s_i(t - v)/s_j(t) \to 1$ as $t \to \infty$ for all $v \geq 0$ and $s_{(0,u)}(t - v)/s_{(0,0)}(t) \to (1 - \mathrm{e}^{-q_0(1-u)})/(1 - \mathrm{e}^{-q_0})$ as $t \to \infty$.*

Proof. Note first that, since $\mu^C > 0$, $\mathrm{P}_i(\tau_0 > t) \leq k_i \mathrm{e}^{-\mu^C t/2}$, so that $\tau_0^{(i)}$ has a lighter tail than τ and, by Lemma 3.1,

$$s_i(t - v) = \mathrm{P}_{(0,0)}(\tau_0^{(i)} + \tau > t - v) \sim \mathrm{P}_{(0,0)}(\tau > t - v) \sim \mathrm{P}_{(0,0)}(\tau > t) = s_{(0,0)}(t).$$

Similarly,

$$\begin{aligned} s_{(0,u)}(t - v) &= \int_0^{1-u} q_0 \mathrm{e}^{-q_0 w} \mathrm{P}(R^1 + \tau > t - v - w)\, \mathrm{d}w \\ &\sim (1 - \mathrm{e}^{-q_0(1-u)})\mathrm{P}(R^1 + \tau > t), \end{aligned}$$

and so $s_{(0,u)}(t - v)/s_{(0,0)}(t)$ converges to the desired limit.

It is easy to see that h, defined by $h_i = 1$ for $i \in C$ and $h_{(0,u)} = (1 - \mathrm{e}^{-q_0(1-u)})/(1 - \mathrm{e}^{-q_0})$, is strictly $\hat{P}$-superharmonic and is harmonic on C; Theorem 3.2, below, then follows easily from a mild adaptation of Theorem 4.1.1 of [7].

Theorem 3.2. *Under the conditions of Theorem 3.1, the restriction of the law of $\tilde{X}^T$ to $\mathcal{F}_s$ converges* vaguely *to that of X^∞ restricted to $\mathcal{F}_s$, where P^∞ is the (substochastic) h-transform of $\tilde{P}$ (which dies from state $(0, u)$ with hazard rate $\lambda(u) = q_0\mathrm{e}^{-q_0}/(1 - \mathrm{e}^{-q_0(1-u)})$).*

Example 3.1. Consider the case where $\sum_{j\in C} q_{0,j} F_{j,0}(\lambda) = \infty$ for all $\lambda > 0$ but $\mu^C > 0$. For example, we may take the nearest-neighbour random walk on $\mathbb{N}$ with up-jump rate b and down-jump rate d (with $b < d$) and then set

$$q_0 = 1, \qquad q_{0,i} = \frac{6}{\pi^2 i^2} \quad \text{for } i \in \mathbb{N}.$$

It is well known that

$$\mu^C = b + d - 2\sqrt{bd}$$

and

$$F_{i,0}(\lambda) = \gamma_\lambda^i,$$

where

$$\gamma_\lambda = \frac{b + d - \lambda - \sqrt{(b + d - \lambda)^2 - 4bd}}{2b} > 1 \quad \text{for } 0 < \lambda \le \mu^C.$$

So, for any $\lambda > 0$, $\sum_{i\in\mathbb{N}} q_{0,i} F_{i,0}(\lambda) = \mathrm{E}\mathrm{e}^{\lambda R^1} = \infty$ and, hence, $\phi = 0$.

Now we consider the case where $\mu^C = 0$ (and, hence, $\phi = 0$ also).

Theorem 3.3. *Denote by $\tau^{(i)}$ a generic random variable having the distribution of τ conditional on $X_0 = i$. Suppose that the $\tau^{(i)}$s have comparable heavy tails, so that*

$$\mathrm{P}(\tau^{(i)} > t) = \mathrm{P}_i(\tau > t) \sim c_i \mathrm{P}(\tau^{(0)} > t) = c_i \mathrm{P}_{(0,0)}(\tau > t)$$

and $\mathrm{P}_i(\tau > t + s)/\mathrm{P}_i(\tau > t) \to 1$ as $t \to \infty$. Then, defining

$$h_i = c_i \quad \textit{for } i \in S$$

and

$$h_{(0,u)} = \frac{1 - \mathrm{e}^{-q_0(1-u)}}{1 - \mathrm{e}^{-q_0}},$$

$$\frac{s_j(t - v)}{s_i(t)} \to \frac{c_j}{c_i} \quad \textit{as } t \to \infty \textit{ for all } v \ge 0 \textit{ and all } i, j \in \hat{S}. \tag{3.4}$$

In particular, if the $\tau_0^{(i)}$s have comparable subexponential tails, with

$$\mathrm{P}(\tau_0^{(i)} > t) = \mathrm{P}_i(\tau_0 > t) \sim a_i \mathrm{P}(\tau_0^{(1)} > t) = a_i \mathrm{P}_1(\tau_0 > t)$$

and

$$q_{0,i} = 0 \quad \textit{for } i > n,$$

then, defining $a_0 = 0$, $m = \sum_{i\in C} q_{0,i} a_i / q_0$,

$$h_i = 1 + \frac{a_i}{(\mathrm{e}^{q_0} - 1)m} \quad \textit{for } i \in S,$$

and

$$h_{(0,u)} = \frac{1 - \mathrm{e}^{-q_0(1-u)}}{1 - \mathrm{e}^{-q_0}},$$

we have

$$\frac{s_j(t - v)}{s_i(t)} \to \frac{h_j}{h_i} \quad \textit{as } t \to \infty \textit{ for all } v \ge 0 \textit{ and all } i, j \in \hat{S}.$$

In general, the vector a must be $\tilde{P}$-superharmonic. If a is $\tilde{P}$-harmonic then h is $\hat{P}$-harmonic, so that, in this case, the restriction of the law of $\tilde{X}^T$ to $\mathcal{F}_s$ converges weakly to that of X^∞ restricted to $\mathcal{F}_s$, where P^∞ is the (stochastic) h-transform of $\tilde{P}$.

https://doi.org/10.1239/jap/1253279853 Published online by Cambridge University Press

Proof. The first claim is essentially a restatement of the conditions for convergence in (3.4). To prove the second statement, first note that we may write

$$\tau^{(i)} = \tau_0^{(i)} + 1 + \sum_{n=1}^{N} (\tilde{H}^n + R^n),$$

where $(\tilde{H}^n)_{n\geq 1}$ is a sequence of independent and identically distributed random variables with distribution equal to that of the holding time in 0 conditioned on its lying in (0,1), N is a geometric(e^{-q_0}) random variable and the R^ns are as in Section 2 and are all independent.

Now each R^n is a mixture of $\tau_0^{(i)}$s, so, by Lemma 3.4,

$$\mathrm{P}(R^n \geq t) = \sum_{i\in C} \frac{q_{0,i}}{q_0} \mathrm{P}(\tau_0^{(i)} \geq t) \sim \sum_{i\in C} \frac{q_{0,i}}{q_0} a_i \mathrm{P}(\tau_0^{(1)} \geq t) = m\mathrm{P}(\tau_0^{(1)} \geq t).$$

Now it follows from Lemma 3.1 that $(\tilde{H}^n + R^n)$ is tail equivalent to R^n and is subexponential. Then we deduce, from Lemmas 3.2 and 3.3, that

$$\mathrm{P}(\tau^{(i)} > t) \sim (a_i + m(\mathrm{e}^{q_0} - 1))\mathrm{P}(\tau_0^{(1)} \geq t) = m(\mathrm{e}^{q_0} - 1)h_i \mathrm{P}(\tau_0^{(1)} \geq t).$$

The last statement follows from the fact that $\tilde{X}$ is nonexplosive, and it is then easy to check (by considering the chain at the epochs when it leaves and returns to 0) that h is then $\hat{P}$-harmonic if a is $\tilde{P}$-harmonic

Theorem 3.4. *Suppose that $\tilde{X}$ is a recurrent birth-and-death process on $\mathbb{Z}^+$ and that, for some i, $\tau_0^{(i)}$ is subexponential. Then $\mathrm{P}(\tau_0^{(j)} > t) \sim (\beta_j/\beta_i)\mathrm{P}(\tau_0^{(i)} > t)$, where β is the unique $\tilde{P}$ harmonic function on $\mathbb{N}$ with $\beta_1 = 1$.*

Proof. Note that, since $\tau_0^{(i)}$ is subexponential, it follows that $\mu^C = 0$ and, hence, by Theorem 5.1.1 of [8], there is a unique $\tilde{P}$-harmonic β. It follows that, for any n, σ_n, the first exit time of X from the set $\{1, \dots, n-1\}$, has an exponential tail (i.e. its tail decreases to 0 at an exponential rate) and the exit is to n with probability β_i/β_n if X starts in i.

It follows that, for each $j \leq i$,

$$\mathrm{P}(\tau_0^{(j)} > t) \sim \frac{\beta_j}{\beta_i} \mathrm{P}(\tau_0^{(i)} > t).$$

Similarly, for $i < n$, $\tau_0^{(i)} = \sigma_n + \mathbf{1}_A\, \tau_0^{(n)}$, where $A =$(X exits $\{1, \dots, n-1\}$ to n), so that

$$\mathrm{P}(\tau_0^{(i)} > t) \sim \mathrm{P}(A)\mathrm{P}(\tau_0^{(n)} > t) = \frac{\beta_i}{\beta_n} \mathrm{P}(\tau_0^{(n)} > t).$$

The following corollary is an immediate consequence of Theorems 3.3 and 3.4.

Corollary 3.1. *If $\tilde{X}$ is a birth-and-death process on $\mathbb{Z}^+$ and, for some i, $\tau_0^{(i)}$ is subexponential, and, for some n, $q_{0,j} = 0$ for $j > n$, then the conclusion of Theorem 3.3 holds.*

Remark 3.2. If $\mu^C = 0$ and the process conditioned on not hitting 0 until time T converges vaguely, then the $\tau_0^{(i)}$s must have comparable *heavy* tails. If, in fact, the convergence is weak (i.e. to an honest process) then the vector a must be harmonic for $\tilde{P}$.

Remark 3.3. Suppose that X is a birth-and-death process, with birth rates b_i equal to the corresponding death rates. If the rates are decreasing in i then $\tau_0^{(1)}$ is subexponential.

To see this, first observe that, by conditioning on the first jump we obtain

$$\mathrm{P}(\tau_0^{(1)} > t) = \tfrac{1}{2}\mathrm{P}(E_1 > t) + \tfrac{1}{2}\mathrm{P}(E_1 + \tau_0^{(2)} > t),$$

where E_1 is the first waiting time in state 1. Now, since

$$\tau_0^{(2)} = \tau_1^{(2)} + \tau_0^{(1)}$$

and since $\tau_1^{(2)}$ stochastically dominates $\tau_0^{(1)}$, we obtain the desired result that

$$\limsup_{t\to\infty} \frac{\overline{F^{(2)}}(t)}{\overline{F}(t)} \le 2,$$

where F is the distribution function of $\tau_0^{(1)}$. The result now follows by (3.2).

4. Some concluding remarks

Sigman [18] provided some conditions which ensure that a random variable has a subexponential tail.

Many obvious examples exist of the α-recurrent case. We have exhibited a few examples in the α-transient case, always assuming that C is irreducible. If it is not then in principle we can divide C into communicating classes $\{C_l : l \in L\}$, where L is some countable or finite index set. It is easy to show that

$$\phi \le \inf_{l\in L} \mu^{C_l}.$$

By adapting the proof of Theorem 3.1, it is easy to see that if τ is subexponential, but $\mu^{C_l} > 0$ for some $l \in L$, then $s_i(t-v)/s_j(t) \to 1$ as $t \to \infty$ for $i, j \in C_l \cup \{\{0\} \times [0, 1)$, and so, as in Theorem 3.2, weak convergence of the conditioned chains is not possible if each $\mu^{C_l} > 0$. Conversely, if $\min_{l\in L} \mu^{C_l} = \mu^{C_{l^*}}$ and X restricted to C_{l^*} is α-recurrent, then $\phi = \mu^{C_{l^*}}$ and a suitably adapted version of Theorem 1.2 and Corollary 1.1 will apply.

References

[1] Doney, R. A. and Bertoin, J. (1996). Some asymptotic results for transient random walks. *Adv. App. Prob.* **28,** 207–226.

[2] Doney, R. A. and Bertoin, J. (1997). Spitzer's condition for random walks and Lévy processes. *Ann. Inst. H. Poincaré Prob. Statist.* **33,** 167–178.

[3] Doney, R. A. and Chaumont, L. (2005). On Lévy processes conditioned to stay positive. *Electron. J. Prob.* **10,** 948–961.

[4] Feller, W. (1968). *An Introduction To Probability Theory and Its Applications*, Vol. I, 3rd edn. John Wiley, New York.

[5] Feller, W. (1971). *An Introduction To Probability Theory and Its Applications*, Vol. II, 2nd edn. John Wiley, New York.

[6] Ferrari, P. A., Kesten, H., Martinez, S. and Picco, P. (1995). Existence of quasi-stationary distributions. A renewal dynamical approach. *Ann. Prob.* **23,** 501–521.

[7] Jacka, S. D. and Roberts, G. O. (1994). Strong forms of weak convergence. *Stoch. Process. Appl.* **67,** 41–53.

[8] Jacka, S. D. and Roberts, G. O. (1995). Weak convergence of conditioned processes on a countable state space. *J. Appl. Prob.* **32,** 902–916.

[9] Jacka, S. and Warren, J. (2002). Examples of convergence and non-convergence of Markov chains conditioned not to die. *Electron. J. Prob.* **7,** 22pp.

[10] Jacka, S., Lazic, Z. and Warren, J. (2005). Conditioning an additive functional of a Markov chain to stay nonnegative. I. Survival for a long time. *Adv. Appl. Prob.* **37,** 1015–1034.

[11] JACKA, S., LAZIC, Z. AND WARREN, J. (2005). Conditioning an additive functional of a Markov chain to stay nonnegative. II. Hitting a high level. *Adv. Appl. Prob.* **37,** 1035–1055.
[12] KESTEN, H. (1995). A ratio limit theorem for (sub) Markov chains on $\{1, 2, \dots\}$ with bounded jumps. *Adv. Appl. Prob.* **27,** 652–691.
[13] KYPRIANOU, A. E. AND PALMOWSKI, Z. (2006). Quasi-stationary distributions for Lévy processes. *Bernoulli* **12,** 571–581.
[14] ROBERTS, G. O. AND JACKA, S. D. (1994). Weak convergence of conditioned birth and death processes. *J. Appl. Prob.* **31,** 90–100.
[15] ROBERTS, G. O., JACKA, S. D. AND POLLETT, P. K. (1997). Non-explosivity of limits of conditioned birth and death processes. *J. Appl. Prob.* **34,** 35–45.
[16] SENETA, E. (1981). *Nonnegative Matrices and Markov Chains*. Springer, New York.
[17] SENETA, E. AND VERE-JONES, D. (1966). On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J. Appl. Prob.* **3**, 403–434.
[18] SIGMAN, K. (1999). Appendix: a primer on heavy-tailed distributions. Queues with heavy-tailed distributions. *Queueing Systems* **33,** 261–275.
[19] WILLIAMS, D. (1979). *Diffusions, Markov Processes, and Martingales*, Vol. I. John Wiley, Chichester.