MARKOV CHAINS CONDITIONED NEVER TO WAIT TOO LONG AT THE ORIGIN

SAUL JACKA

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 there is a weak limit as $T \to \infty$ in the cases where either the statespace 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.

August 29, 2008

1. Introduction and notation

1.1. **Introduction.** Feller (in Section XIII.7 of [4]) 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 equation (1.1) modified to become

$$x^k - q \sum_{j=0}^{k-1} p^j x^{k-1-j} = 0, \tag{1.2}$$

and $c_k = \frac{s_k - p}{q((k+1)s_k - k)}$.

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 inter-jump time exceeding one unit by 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,

$$\mathbb{P}(\tau_Y > t) \sim c_r e^{-\phi_r t}, \tag{1.3}$$

Key words: SUBEXPONENTIAL TAIL; EVANESCENT PROCESS; FELLER'S COIN-TOSSING CONSTANTS; HITTING PROBABILITIES; CONDITIONED PROCESS .

AMS 2000 subject classifications: Primary 60J27; secondary 60J80, 60J50, 60B10.

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

$$xe^{-x} = re^{-r}. \tag{1.4}$$

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 [7], [8], [11], [1], [2], [3], [6], [12], [9], [15], [10], [14], [13]) 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 $S = \{0, \ldots, n\}$ and $\partial = 0$ so that $C = \mathbb{N}$ or $C = \{1, \ldots, n\}$.

We assume that X is irreducible, and non-explosive. We denote the transition semi-group of X by $\{P(t); t \geq 0\}$ and its Q-matrix by Q. 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 denote the successive holding times in state 0 by $(H^n)_{n\geq 0}$ and the successive return times to state 0 by $(R^n)_{n\geq 0}$, with the convention that $H^0 = 0$ if $X_0 \neq 0$ and $R^0 = 0$ if $X_0 = 0$. From time to time it will be convenient to refer to the current holding time, so we define $H_t = t - H^0 + R^0 + \ldots H^{n-1} + R^{n-1}$ if $X_t = 0$ and $H^0 + R^0 + \ldots H^{n-1} + R^{n-1} \leq t \leq H^0 + R^0 + \ldots H^{n-1} + R^{n-1} + H^n$ and $H_t = \emptyset$ otherwise. We denote the first time that X waits in 0 for time 1 by τ and denote X killed at time τ by $\hat{X}$. We denote the statespace augmented by the current holding time in 0 by $\hat{S} \overset{\text{def}}{=} C \cup \{\{0\} \times [0, 1)\}$. By a slight abuse of notation, we denote the (substochastic) Markov chain $(\hat{X}_t, H_t)$ on the statespace $\hat{S}$ by $\hat{X}$ also. The associated semigroup is denoted $\{\hat{P}(t); t \geq 0\}$. Throughout the rest of the paper we denote by $\mathbb{P}_i$ the probability on Skorokhod pathspace $D(S, [0, \infty))$, 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 [9]) that, if X^* is a Markov chain on C, with substochastic transition semi-group P^* and Q- matrix Q^*, then X^* is said to be evanescent if it is irreducible and dies with probability one. In that case, we define

$$\alpha_{X^*} = \alpha = \inf\{\lambda \geq 0 : \int_0^\infty P^*_{ij}(t) e^{\lambda t} dt = \infty\}$$

for any $i, j \in C$, and (see, for example, Seneta and Vere-Jones [17]) X^* is classified as α-recurrent or α-transient depending on whether $\int_0^\infty P^*_{ij}(t) e^{\alpha t} dt = \infty$ or is finite. Moreover, X^* is α-recurrent if and only if $\int_0^\infty f^*_{ii}(t) e^{\alpha t} dt = 1$, 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

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

otherwise X^* is α-null recurrent 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^* and we define $s_i^*(t) = \sum_j P_{ij}^*(t) = \mathbb{P}(\tau^* > t)$ and set

$$\mu^* = \inf\{\lambda : \int_0^\infty s_i^*(t) e^{\lambda t} dt = \infty\}.$$

Notice 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 the survival probabilities for $\tilde{X}$ as s^C, so that $s_i^C(t) = \mathbb{P}_i(\tau_0 > t)$.

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

$$P^h(x, dy; t) \stackrel{\text{def}}{=} \frac{h(y)}{h(x)} P(x, dy; t),$$

so that if we consider the corresponding substochastic measures on path- space, $\mathbb{P}_x$ and $\mathbb{P}_x^h$ (conditional on $X_0 = x$) then

$$\frac{d\mathbb{P}_x^h}{d\mathbb{P}_x}\Big|_{\mathcal{F}_t} = h(X_t)$$

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 $h(x,t) = e^{\phi t} h_x$ then

$$\frac{d\mathbb{P}_x^h}{d\mathbb{P}_x}\Big|_{\mathcal{F}_t} = e^{\phi t} h_{X_t}.$$

As shown in [9], 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.** We define

$$s_i(t) \stackrel{\text{def}}{=} \mathbb{P}_i(\tau > t) = \hat{P}(i, \hat{S}; t).$$

Our first result is

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

$$p_{(0,0)} \stackrel{\text{def}}{=} p_0 = \frac{(1-e^{-q_0})\Delta}{e^{-q_0}+(1-e^{-q_0})\Delta}, \tag{1.5}$$

$$p_{(0,u)} = \frac{1-e^{-q_0(1-u)}}{1-e^{-q_0}} p_0, \tag{1.6}$$

and

$$p_i = \beta_i + (1-\beta_i)p_0 \text{ for } i \in C. \tag{1.7}$$

Then

$$s_i(t) \stackrel{t\to\infty}{\longrightarrow} p_i \text{ for all } i \in \hat{S}. \tag{1.8}$$

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

$$P^\infty_{i,j}(t) = \frac{p_j}{p_i}\hat{P}_{i,j}(t) \text{ for } j \in C \tag{1.9}$$

and

$$P^\infty_i((0,du);t) = \frac{p_{(0,u)}}{p_i}\hat{P}_i((0,du);t), \tag{1.10}$$

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

$$d(t) = \frac{e^{-q_0 t}}{\int_0^1 e^{-q_0 s} ds}$$

and a jump probability out of state 0 to state j of $\frac{q_{0,j}p_j}{q_0 p_0}$ (independent of the holding time).

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

Now let $W \stackrel{\text{def}}{=} H^1 + R^1$ (so that W is the first return time of X to 0 from 0) and let g be the (defective) density of $W1_{(H^1<1)}$ 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' result in the α-positive case.

Theorem 1.2. *Let*

$$I(\lambda) \stackrel{\text{def}}{=} \int_0^\infty e^{\lambda t} g(t)dt = \mathbb{E}e^{\lambda W} 1_{(H^1<1)},$$

then if

$$\text{there exists a } \phi \text{ such that } I(\phi) = 1, \text{ and } I'(\phi-) < \infty, \tag{1.11}$$

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

$$e^{\phi t} s_i(t) \stackrel{t\to\infty}{\longrightarrow} p_i > 0,$$

for a suitable function p. The function p is now given by

$$p_{(0,0)} = \frac{e^{\phi - q_0}}{\phi I'(\phi-)} \stackrel{\text{def}}{=} \kappa, \tag{1.12}$$

$$p_{(0,u)} = \frac{\int_0^{1-u} e^{(\phi-q_0)s} ds}{\int_0^1 e^{(\phi-q_0)s} ds} \kappa, \tag{1.13}$$

and

$$p_i = F_{i,0}(\phi)\kappa, \tag{1.14}$$

where

$$F_{i,0}(\lambda) \stackrel{\text{def}}{=} \mathbb{E}e^{\lambda\tau_0^{(i)}} = \int_0^\infty e^{\lambda t}\rho_i(t)dt.$$

The following simple condition ensures that condition (1.11) holds.

Lemma 1.3. *Suppose that $\tilde{X}$ is α-recurrent, and both $N_0 \stackrel{\text{def}}{=} \{i : q_{i,0} > 0\}$ and $N_0^* \stackrel{\text{def}}{=} \{i : q_{0,i} > 0\}$ are finite, then (1.11) holds.*

Corollary 1.4. *Let X^T denote the chain on $\hat{S}$ obtained by conditioning $\hat{X}$ on the event $(\tau > T)$, then, if condition (1.11) holds, for each $s > 0$, the restriction of the law of X^T to $\mathcal{F}_s$ converges weakly to that of X^∞ restricted to $\mathcal{F}_s$, where the transition semigroup of X^∞ is given by equations (1.9) and (1.10).*

In the case where $I(\phi) < 1$ or $I'(\phi-) = \infty$, Theorems 3.6, 3.7 and 3.9 (may) apply, giving some sufficient conditions for weak or vague convergence to take place. In Theorem 3.10 and Corollary 3.11, we give an application to the case of a recurrent birth and death process conditioned not to wait too long in state 0.

2. Proof 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

$$p_i = \beta_i + (1-\beta_i)p_0 \tag{2.1}$$

and

$$p_0 = (1 - e^{-q_0}) \sum_{j\in C} \frac{q_{0,j}}{q_0} p_i. \tag{2.2}$$

Equations (1.5)-(1.7) follow immediately. Then the conditioning result follows straightforwardly $\square$

Example 2.1. *We take a transient nearest-neighbour random walk with reflection at 0 and with up-jump rate of b and 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 $1 - \beta_i = (\frac{d}{b})^i$.*

The main tool in the proof of Theorem 1.2 is the Renewal Theorem.

Proof of Theorem 1.2 First note that $s_{(0,0)}$ satisfies the renewal equation

$$s_{(0,0)}(t) = \big(1 - \int_0^\infty g(u)du\big)1_{(t<1)} + \int_t^\infty g(u)du + \int_0^t g(u)s_{(0,0)}(t-u)du. \tag{2.3}$$

If we define

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

it follows immediately from (2.3) that

$$f(t) = e^{\phi t}\big((1 - \int_0^\infty g(u)du)1_{(t<1)} + \int_t^\infty g(u)du\big) + \int_0^t \tilde{g}(u)f(t-u)du, \tag{2.4}$$

where $\tilde{g}(t) \stackrel{\text{def}}{=} e^{\phi t}g(t)$. Now, it is easy to check that the conditions of Feller's alternative formulation of the Renewal Theorem (see XI.1 of [5], p.363) are satisfied, so we conclude that

$$f(t) \stackrel{t\to\infty}{\longrightarrow} \mu^{-1}\int_0^\infty e^{\phi t}((1 - \int_0^\infty g(u)du)1_{(t<1)}) + \int_t^\infty g(u)du)dt, \tag{2.5}$$

where

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

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

$$\int_0^\infty e^{\phi t}\int_t^\infty g(u)dudt = \int_0^\infty g(u)\int_0^u e^{\phi t}dtdu = \frac{I(\phi) - \int_0^\infty g(u)du}{\phi} = \frac{1 - \int_0^1 q_0 e^{-q_0 u}du}{\phi} = \frac{e^{-q_0}}{\phi},$$

and hence (1.12) follows.

To establish (1.14), notice that (by conditioning on the time of the first hit of 0),

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

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

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

where $\tilde{\rho}_i(t) \stackrel{\text{def}}{=} e^{\phi t}\rho_i(t)$. Now f is continuous and converges to κ so, by the Dominated Convergence Theorem,

$$\int_0^t \tilde{\rho}_i(u)f(t-u)du \stackrel{t\to\infty}{\longrightarrow} \int_0^\infty \kappa\tilde{\rho}_i(u)du = \kappa F_{i,0}(\phi).$$

Moreover, since $F_{i,0}(\phi) = \int_0^\infty \tilde{\rho}_i(u)du < \infty$, it follows that $e^{\phi t}\int_t^\infty \rho_i(u)du \le \int_t^\infty \tilde{\rho}_i(u)du \stackrel{t\to\infty}{\longrightarrow} 0$, and hence

$$f_i(t) \stackrel{t\to\infty}{\longrightarrow} \kappa F_{i,0}(\phi)$$

as required.

To establish (1.13), observe that

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

and hence

$$f_{(0,u)}(t) \stackrel{\text{def}}{=} e^{\phi t}s_{(0,u)}(t) = e^{(\phi-q_0)t}1_{(t<1-u)} + \int_{w=0}^t\int_{v=0}^{(1-u)\wedge w} q_0 e^{(\phi-q_0)v}\tilde{\rho}(w-v)f(t-w)dvdw,$$

and hence, by the Dominated Convergence Theorem,

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

as required ∎

Remark 2.2. *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*

$$I(\lambda) = \int_0^1 re^{(\lambda - r)t} dt,$$

and so ϕ satisfies $r\frac{e^{\phi-r}-1}{\phi-r} = 1$ which establishes (1.4), and

$$e^{\phi t}\mathbb{P}(\tau > t) \overset{t\to\infty}{\longrightarrow} \frac{e^{\phi-r}}{\phi I'(\phi-)} = \frac{\phi - r}{r(\phi-1)},$$

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

Now we give the

Proof of Lemma 1.3 It follows from Theorem 3.3.2 of [10] that if N_0 is finite then $\alpha^C = \mu^C$. Now since $\tilde{X}$ is α-recurrent it follows that

$$\int_0^\infty e^{\lambda t}\tilde{P}_{ii}(t)dt < \infty \text{ iff } \lambda < \alpha^C,$$

and so one easily deduces (since $s_i^C(t) \geq \tilde{P}_{ii}(t)$) that

$$\int_0^\infty e^{\lambda t} s_i^C(t)dt < \infty \text{ iff } \lambda < \alpha^C.$$

Now

$$I(\lambda) = \int_0^\infty e^{\lambda t} g(t)dt = \int_0^1 e^{(\lambda - q_0)t}dt\big(\sum_{i\in N_0^*} q_{0,i}F_{i,0}(\lambda)\big) = \int_0^1 e^{(\lambda-q_0)t}dt\big(\sum_{i\in N_0^*} q_{0,i}(\frac{F_{i,0}(\lambda)-1}{\lambda})\big),$$

and so $I(\lambda) < \infty$ iff $\lambda < \alpha^C$. It now follows trivially that $\phi < \alpha^C$ and that (1.9) is satisfied □

Now we give the

Proof of Corollary 1.4 This follows immediately from Theorem 1.2 and Theorem 4.1.1 of [9] provided that we can show that h, given by $h : (i,t) \mapsto e^{\phi t}p_i$, 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, $\mathbb{E}_{(0,u)}h(\hat{X}_{t \min \sigma}, t\wedge\sigma) = h((0,u),0)$ and $\mathbb{E}_i h(\hat{X}_{t\wedge\tau_0}, t\wedge\tau_0) = h(i,0)$ for $i \in C$. This is sufficient since $\hat{X}$ is non-explosive □

3. The α-transient case

We seek now 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

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

and

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

where

$$J^\lambda(x) \stackrel{\text{def}}{=} \int_0^x e^{(\lambda-q_0)v}dv,$$

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

Proof. It is a 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 transform above 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 $\frac{\psi_i(t-v)}{\psi_j(t)}$ has a limit as $t \to \infty$ then it is the common limit of $\frac{s_i(t-v)}{s_j(t)} = \frac{\int_{t-v}^\infty \psi_i(u)du}{\int_t^\infty \psi_j(u)du}$ and $\frac{h_i^\phi}{h_j^\phi}\frac{s_i^{h^\phi}(t-v)}{s_j^{h^\phi}(t)} = \frac{\int_{t-v}^\infty e^{\phi u}\psi_i(u)du}{\int_t^\infty e^{\phi u}\psi_j(u)du}$.

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 $(i,t) \mapsto e^{\phi t}h_i^\phi$ is the unique $\hat{P}$-superharmonic function of the form $e^{\lambda t}k_i$ 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 Sigman [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

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

where $\overline{F_Z} \stackrel{\text{def}}{=} 1 - F_Z$, 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

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

and (3.1) holds iff

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

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

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

for some $c > 0$.

Y is said to have a lighter tail than X if

$$\frac{\overline{F_Y}(t)}{\overline{F_X}(t)} \overset{t\to\infty}{\longrightarrow} 0.$$

Lemma 3.2. *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 3.3. *If X and Y are independent and 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 3.4. *Suppose that N is a geometric r.v. and $X_1, \ldots$ are iid with common d.f F which is subexponential, then if*

$$S \overset{\text{def}}{=} \sum_1^N X_i,$$

then S is subexponential and

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

Finally, we have the following

Lemma 3.5. *Suppose that $X_1, \ldots$ are independent and tail-equivalent with*

$$F_{X_i} \overset{\text{def}}{=} F_i,$$

and 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 $F(t) = \sum_{i\in\mathbb{N}} \mathbb{P}(J = i)F_i(t)$).

Now suppose that

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

if the collection $\{\frac{\overline{F_J}(t)}{\overline{F_1}(t)};\ t \ge 0\}$ are uniformly integrable then,

$$\overline{F}(t) \sim (\mathbb{E}c_J)\overline{F_1}(t). \tag{3.3}$$

In particular, if J is a bounded r.v. then (3.3) holds.

Proof. It follows from the assumptions that

$$\frac{\overline{F_J}(t)}{\overline{F_1}(t)} \overset{t\to\infty}{\longrightarrow} c_J \text{ a.s.}$$

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

$$\frac{\overline{F}(t)}{\overline{F_1}(t)} \overset{t\to\infty}{\longrightarrow} \mathbb{E}c_J.$$

In particular, if $J \le n$ a.s. then

$$\limsup_{t\to\infty} \frac{\overline{F_J}(t)}{\overline{F_{X_1}}(t)} \le \max_{1\le i\le n} c_i \text{ a.s.},$$

and so the collection is indeed u.i. □

3.3. **Results for heavy tails.** Suppose first that $0 = \phi < \mu_C$.

Theorem 3.6. *If $0 = \phi < \mu_C$ and τ is subexponential, then $\frac{s_i(t-v)}{s_j(t)} \overset{t\to\infty}{\longrightarrow} 1$ for all $v \ge 0$ and $s_{(0,u)}(t-v)/s_{(0,0)}(t) \overset{t\to\infty}{\longrightarrow} \frac{1-e^{-q_0(1-u)}}{1-e^{-q_0}}$.*

Proof. Notice first that, since $\mu_C > 0$, $\mathbb{P}_i(\tau_0 > t) \le k_i e^{-\mu_C t/2}$, so that τ_0^i has a lighter tail than τ so, by Lemma 3.2,

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

Similarly,

$$s_{(0,u)}(t-v) = \int_0^{1-u} q_0 e^{-q_0 w}\mathbb{P}(R^1 + \tau > t - v - w)dw \sim (1 - e^{-q_0(1-u)})\mathbb{P}(R^1 + \tau > t)$$

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 $h_{(0,u)} = \frac{1-e^{-q_0(1-u)}}{1-e^{-q_0}}$ is strictly $\hat{P}$- superharmonic and is harmonic on C: the following theorem then follows easily from a mild adaptation of Theorem 4.1.1 of [9].

Theorem 3.7. *Under the conditions of Theorem 3.6, 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 $\lambda(u) = \frac{q_0 e^{-q_0}}{1-e^{-q_0(1-u)}}$).*

Example 3.8. *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*

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

It is well-known that

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

and

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

where

$$\gamma_\lambda = \frac{b + d - \lambda - \sqrt{(b+d-\lambda)^2 - 4bd}}{2b} > 1$$

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

Now we consider the case where $\mu_C = 0$ (and hence $\phi = 0$ also).

Theorem 3.9. *Suppose that $\tau^{(i)}$ have comparable heavy tails, so that $\mathbb{P}(\tau^{(i)} > t) = \mathbb{P}_i(\tau > t) \sim c_i \mathbb{P}(\tau^{(0)} > t) = \mathbb{P}_{(0,0)}(\tau > t)$ and $\frac{\mathbb{P}_i(\tau > t+s)}{\mathbb{P}_i(\tau > t)} \xrightarrow{t\to\infty} 1$, then, defining*

$$h_i = c_i, \text{ for } i \in S$$

and

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

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

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

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

and

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

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

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

and

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

we have that

$$\frac{s_j(t-v)}{s_i(t)} \xrightarrow{t\to\infty} \frac{h_j}{h_i} \text{ for all } v \geq 0 \text{ and for all } i, j \in \hat{S}.$$

In general 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 $\tilde{P}$.

Proof. The first claim is essentially a restatement of the conditions for convergence in (3.4).

To prove the second statement, first notice that we may write

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

where $(\tilde{H}^n)_{n\geq 1}$ are a sequence of iid random variables with distribution that of the holding time in 0 conditioned on its lying in (0,1), N is a Geometric(e^{-q_0}) r.v. and the R^n's are as in section 2 and all are independent.

Now each R^n is a mixture of $\tau_0^{(i)}$s, so, by Lemma 3.5,

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

Now it follows from Lemma 3.2 that $(\tilde{H}^n + R^n)$ is tail equivalent to R^n and is subexponential and then we deduce, from Lemmas 3.3 and 3.4 that $\mathbb{P}(\tau^{(i)} > t) \sim (a_i + m(e^{q_0} - 1))\mathbb{P}(\tau_0^{(1)} \geq t) = m(e^{q_0} - 1)h_i \mathbb{P}(\tau_0^{(1)} \geq t)$. The last statement follows from the fact that $\tilde{X}$ is non-explosive and it is then easy to check (by considering the chain at the epochs when it leaves and returns to 0) that then h is $\hat{P}$-harmonic if a is $\tilde{P}$-harmonic □

Theorem 3.10. *Suppose that $\tilde{X}$ is a recurrent birth and death process on $\mathbb{Z}^+$ and, for some i, $\tau_0^{(i)}$ is subexponential, then $\mathbb{P}(\tau_0^{(j)} > t) \sim \frac{\beta_j}{\beta_i}\mathbb{P}(\tau_0^{(i)} > t)$, where β is the unique $\tilde{P}$ harmonic function on $\mathbb{N}$ with $\beta_1 = 1$.*

Proof. Notice that, since $\tau_0^{(i)}$ is subexponential, it follows that $\mu_C = 0$ and hence, by Theorem 5.1.1 of [10], 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 $j \le i$,

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

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

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

□

The following is an immediate consequence of Theorems 3.9 and 3.10.

Corollary 3.11. *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 3.9 holds.*

Remark 3.12. *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 3.13. *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 that

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

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] gives 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

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

By adapting the proof of Theorem 3.6, it is easy to see that if τ is subexponential but $\mu^{C_l} > 0$ for some $l \in L$, then $\frac{s_i(t-v)}{s_j(t)} \overset{t\to\infty}{\longrightarrow} 1$ for $i, j \in C_l \cup \{\{0\} \times [0, 1)$ and so, as in Theorem 3.7, weak convergence of the conditioned chains is not possible if each $\mu^{C_l} > 0$. Conversely, if $\min\limits_{l \in L} \mu^{C_l} = \mu^{C_{l^*}}$ 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 1.4 will apply.

References

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

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

[3] R A Doney and L Chaumont: On Lévy processes conditioned to stay positive (2005), to appear in *JLMS*.

[4] W Feller: "An Introduction To Probability Theory and its Applications Vol I", 3rd Edn. (revised printing). Wiley, New York (1970).

[5] W Feller: "An Introduction To Probability Theory and its Application Vol II". 2nd Edition. Wiley New York (1971)

[6] P A Ferrari, H Kesten, S Martinez, and P Picco: Existence of quasi stationary distributions. A renewal dynamical approach. *Ann. Probab.*, 23 (2), 501–521 (1995)

[7] Saul Jacka, Zorana Lazic and Jon Warren: Conditioning an additive functional of a Markov chain to stay non-negative I: survival for a long time. *Adv. Appl. Prob.* 37 (4), 1015–1034, (2005)

[8] Saul Jacka, Zorana Lazic and Jon Warren: Conditioning an additive functional of a Markov chain to stay non-negative II: hitting a high level. *Adv. Appl. Prob.* 37 (4), 1035–1055 (2005)

[9] S D Jacka and G O Roberts: Strong forms of weak convergence. *Stoch. Proc. & Appl.* 67(1), 41–53 (1994).

[10] S D Jacka and G O Roberts: Weak convergence of conditioned processes on a countable state space. *J. Appl. Probab.* 32(4), 902–916 (1995).

[11] Saul Jacka and Jon Warren: Examples of convergence and non-convergence of Markov chains conditioned not to die'. *Elec. J. Prob.* 7 (1) (2002)

[12] H Kesten: A ratio limit theorem for (sub) Markov chains on $\{1, 2, ..\}$ with bounded jumps. *Adv. Appl. Prob.*, 27 (3), 652–691 (1995)

[13] Andreas E. Kyprianou and Zbigniew Palmowski: Quasi-stationary distributions for Lévy processes. *Bernoulli*12 (4), 571–581 (2006).

[14] G O Roberts and S D Jacka: Weak convergence of conditioned birth and death processes *J. Appl. Prob.* 31 (1), 90–100 (1994).

[15] G O Roberts, S D Jacka and P K Pollett: Non-explosivity of limits of conditioned birth and death processes. *J. Appl. Probab.* 34(1), 35– 45 (1997).

[16] E Seneta:"Non-negative Matrices and Markov Chains", Springer-Verlag, New York, Heidelberg, Berlin (1981).

[17] E Seneta and D Vere-Jones: On quasi-stationary distributions in discrete time Markov chains with a denumerable infinity of states. *J. Appl. Prob.* **3**, 403–434 (1966).

[18] K Sigman: Appendix: A primer on heavy-tailed distributions *Queueing Systems*, 33, 261–275 (1999).
[19] D Williams: "Diffusions, Markov processes, and martingales: Vol. I". Wiley, New York (1979).

Department of Statistics, University of Warwick, Coventry CV4 7AL, UK

E-mail address: s.d.jacka@warwick.ac.uk